

Edited by

Raffaele De Caterina

J. Alfredo Martinez

Martin Kohlmeier

Principles of Nutrigenetics and Nutrigenomics

Fundamentals for Individualized Nutrition



PRINCIPLES OF NUTRIGENETICS
AND NUTRIGENOMICS

This page intentionally left blank

PRINCIPLES OF NUTRIGENETICS AND NUTRIGENOMICS

Fundamentals of Individualized Nutrition

Edited by

RAFFAELE DE CATERINA

*Cattedra di Malattie dell'Apparato Cardiovascolare
Università degli Studi di Pisa, Italy*

J. ALFREDO MARTINEZ

*Ciencias de la Alimentación y Fisiología
Universidad de Navarra, Navarra, Spain*

MARTIN KOHLMEIER

*Human Research Core and Nutrigenetics Laboratory, UNC Nutrition Research Institute, and
Schools of Medicine and Public Health Department of Nutrition Kannapolis, NC, United States*



ACADEMIC PRESS

An imprint of Elsevier

Academic Press is an imprint of Elsevier
125 London Wall, London EC2Y 5AS, United Kingdom
525 B Street, Suite 1650, San Diego, CA 92101, United States
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom

Copyright © 2020 Elsevier Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloging-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN: 978-0-12-804572-5

For information on all Academic Press publications visit our website at
<https://www.elsevier.com/books-and-journals>

Publisher: Charlotte Cockle
Acquisition Editor: Megan R. Ball
Editorial Project Manager: Devlin Person
Production Project Manager: Paul Prasad Chandramohan
Cover Designer: Miles Hitchen

Typeset by TNQ Technologies



Contents

Contributors	ix		
Editor Biographies	xv		
Preface	xvii		
Introduction	xix		
I			
The Biological Basis of Heritability and Diversity			
<hr/>			
1. The Nature of Traits, Genes and Variation	3	10. Metabolomics for Biomarker Discovery and to Derive Genetic Links to Disease	75
JOSÉ L. VIZMANOS, NICHOLAS C.P. CROSS, AND FRANCISCO J. NOVO		SUSAN C.J. SUMNER, SUSAN MCRITCHIE, AND WIMAL PATHMASIRI	
2. Molecular Biology of Genetic Variants	11	11. Metagenomics	81
MARTIN KOHLMEIER		ISABEL MORENO-INDIAS AND FRANCISCO J. TINAHONES	
3. Regulation of Gene Expression	17	12. A Broader View on Omics and Systems Biology	89
MARIA LUISA BONET AND ANDREU PALOU		PATRICIA CASAS-AGUSTENCH AND CRISTINA ANDRES-LACUEVA	
4. The Role of Nutrition in DNA Replication, DNA Damage Prevention and DNA Repair	27	13. Study Designs in Genomic Epidemiology	99
MICHAEL FENECH		JOSÉ L. SANTOS AND FERNANDO RIVADENEIRA	
5. Genotyping and Sequencing	33	14. Epistemology of Nutrigenetic Knowledge	107
TERESA EZPONDA, GORKA ALKORTA-ARANBURU, FELIPE PRÓSPER, AND XABIER AGIRRE		MARTIN KOHLMEIER	
6. RNA Analyses	41	II	
FRANCISCO JAVIER RUIZ-OJEDA, JULIO PLAZA-DIAZ, AND ANGEL GIL		Nutrigenetics	
7. Methods for Global Nutrigenomics and Precision Nutrition	49	<hr/>	
ADELA CHIRITA-EMANDI AND MIHAI NICULESCU		15. Nutrigenetics and the Early Life Origins of Health and Disease: Effects of Protein Restriction	113
8. Epigenetic Analyses Tools for Nutrition Research	59	SIMON LANGLEY-EVANS	
MARIA L. MANSEGO, FERMÍN I. MILAGRO, AND J. ALFREDO MARTÍNEZ		16. Maternal Perinatal Nutrition and Offspring Programming	121
9. Proteomic Analyses	69	PAUL CORDERO, MANLIO VINCIGUERRA, AND JUDE A. OBEN	
FERNANDO J. CORRALES AND LETICIA ODRIOZOLA		17. Introduction to Epigenetics	129
		LARA K. HYDE, SIMONETTA FRISO, AND SANG-WOON CHOI	
		18. Genetics of Chrononutrition	141
		MARTA GARAULET, JESUS LOPEZ-MINGUEZ, AND PURIFICACIÓN GÓMEZ-ABELLÁN	
		19. Screening for Inborn Errors of Metabolism	153
		ADELA CHIRITA-EMANDI	

20. Genomics of Eating Behavior and Appetite Regulation BAYAN MESMAR AND NANETTE STEINLE	159	33. Genetic Determinants of Vitamin E Status PATRICK BOREL AND CHARLES DESMARCHELIER	255
21. Genetics of Body Composition: From Severe Obesity to Extreme Leanness JOSÉ L. SANTOS AND VÍCTOR CORTÉS	167	34. Vitamin K JÖRG SAUPE	263
22. Genetic Regulation of Energy Homeostasis: Obesity Implications ITZIAR ABETE, SANTIAGO NAVAS-CARRETERO, MARIA ANGELES ZULET, AND J. ALFREDO MARTINEZ	175	35. Vitamins as Cofactors for Energy Homeostasis and Their Genomic Control, With Special Reference to Biotin, Thiamine, and Pantothenic Acid ANTONIO VELAZQUEZ-ARELLANO AND ALAIN DE J. HERNANDEZ-VAZQUEZ	271
23. Genetic Variations With Influence on the Individualized Response to Weight Loss Diets: FTO as Evidence DANIEL ANTONIO DE LUIS, ROCÍO ALLER, OLATZ IZAOLA, AND J. ALFREDO MARTINEZ	181	36. Impact of Genetic Variation on Vitamin C Status MARTIN KOHLMEIER	279
24. Genetic Variation in the Response to Exercise Training: Impact on Physical Fitness and Performance LOUIS PÉRUSSE	187	37. Riboflavin ANNET M. BOSCH	283
25. Genetic Variations Impacting the Response to Defined Diets TIANGE WANG AND LU QI	197	38. Niacin GEORGE RASTI, NICOLAS G. SIMONET AND ALEJANDRO VAQUERO	287
26. Carbohydrates and Nutrigenetics CLARA IBÁÑEZ AND ANA RAMÍREZ DE MOLINA	203	39. Folate and Vitamins B ₆ and B ₁₂ NATALIA I. KRUPENKO	295
27. Gene-Diet Interactions and Cardiovascular Diseases: Saturated and Monounsaturated Fat JOSE M. ORDOVAS	211	40. The Nutrigenetics of Choline STEVEN H. ZEISEL	303
28. Nutrigenetics: Omega-3 and Omega-6 Fatty Acids WILLIAM S. HARRIS	223	41. Electrolytes OLIVIA DONG AND TIM WILTSHIRE	309
29. Nutrigenetics and Blood Cholesterol Levels in Response to Plant Sterols ITZEL VAZQUEZ-VIDAL AND PETER J.H. JONES	227	42. Iron MARTIN KOHLMEIER	317
30. Genetic Individuality and Alcohol Consumption ROSA M. LAMUELA-RAVENTOS, RAMÓN ESTRUCH, AND RICHARD KIRWAN	231	43. Trace Elements MARTIN KOHLMEIER	321
31. Vitamin A and Other Carotenoids JOHANNES VON LINTIG AND DARWIN BABINO	237	44. Polyphenols and Nutrigenetic/Nutrigenomic Associations With Obesity-Related Metabolic Diseases EUN-YOUNG KWON AND MYUNG-SOOK CHOI	327
32. Vitamin D KARINA FISCHER	245	45. Caffeine GIULIA RENDA AND RAFFAELE DE CATERINA	335
		46. Phytanic Acid Metabolism in Health and Disease RONALD J.A. WANDERS	341
		47. Nutrigenetics, Fatty Acids, and Cognition CAROL L. CHEATHAM	347

48. The Nutrigenetics of Cardiovascular Disease ROSALINDA MADONNA AND RAFFAELE DE CATERINA	355	59. Nutrients and Gene Expression in Type 2 Diabetes DOMINIQUE LANGIN AND NATHALIE VIGUERIE	441
49. Hemostasis and Thrombosis ANTONIO CAPURSO AND CRISTIANO CAPURSO	361	60. Gene Expression in Dyslipidemias CÁTIA LIRA DO AMARAL AND RUI CURI	447
50. Diet—Gene Interactions: Perspectives in Insulin Resistance and Type 2 Diabetes JESSICA C. RALSTON AND HELEN M. ROCHE	371	61. Nutrients and Gene Expression in Inflammation CHARLES N. SERHAN	457
51. Nutrigenetics of Bone Health JOSÉ-LUIS PÉREZ-CASTRILLON, MARTA RUIZ-MAMBRILLA, AND JOSÉ A. RIANCHO	377	62. Nutrients and Gene Expression in Cardiovascular Disease MARIKA MASSARO, EGERIA SCODITTI, NADIA CALABRISO, MARIA ANNUNZIATA CARLUCCIO, PAUL HUGENHOLTZ, AND RAFFAELE DE CATERINA	469
52. Nutrients and Genes in the Liver: The Central Role of Liver X Receptor in Lipogenesis MARICA CARIELLO, ELENA PICCININ, OIHANE GARCIA-IRIGOYEN, AND ANTONIO MOSCHETTA	383	63. Nutrients and Gene Expression in Cancer ANA B. CRUJEIRAS AND DAVID DE LORENZO	483
53. Hyperuricemia and Gout GEETHA CHITTOOR AND VENKATA SAROJA VORUGANTI	389	64. Nutrients and Gene Expression Affecting Bone Metabolism JOSÉ LUIS PÉREZ-CASTRILLÓN AND JOSÉ ANTONIO RIANCHO DEL MORAL	489
54. Aging: Mechanisms, Signaling and Dietary Intervention DEYANG YU, LUIGI FONTANA, AND DUDLEY W. LAMMING	395		
55. Epigenetics in Food Allergies: The Missing Piece of the Puzzle SHIMRIT BAR-EL DADON AND RAM REIFEN	403		
56. Gut Microbiota and Their Influence on Response to Food CLARISSE MAROTZ, STEVEN BROWN, AND AMIR ZARRINPAR	411		

III Nutrigenomics

57. Nutrients and Gene Expression in Development DOLORES BUSSO, NICOLÁS SANTANDER, FRANCISCA SALAS-PÉREZ, AND JOSÉ LUIS SANTOS	423		
58. Nutrients, Obesity and Gene Expression FERMÍN I. MILAGRO, MARÍA J. MORENO-ALIAGA, AND J. ALFREDO MARTÍNEZ	431		
		IV Translational Nutrigenetics and Nutrigenomics	
		65. Precision Nutrition Interventions Based on Personalized Genetic Advice RODRIGO SAN-CRISTOBAL, SANTIAGO NAVAS-CARRETERO, MARTIN KOHLMEIER, AND J. ALFREDO MARTÍNEZ	499
		66. Genetically Informed Nutrition Guidance MARTIN KOHLMEIER	509
		67. Nutrigenomics of Food Pesticides LAURA BORDONI AND ROSITA GABBIANELLI	513
		68. Pharmaconutrigenetics: The Impact of Genetics on Nutrient—Drug Interactions OLIVIA DONG AND TIM WILTSHIRE	519
		69. Setting Genome-Directed Guidelines and Algorithms MARTIN KOHLMEIER	525

70. Direct-to-Consumer Testing LYNNETTE R. FERGUSON AND VIRGINIA R. PARSLow	529	73. Impact of Nutrigenetics and Nutrigenomics on Society ARTEMIS P. SIMOPOULOS	549
71. Privacy Risks and Protective Measures FRANCESCO VISIOLI	539	Index	557
72. Ethical Considerations in Nutrigenetics and Nutrigenomics ULF GÖRMAN, JENNIE AHLGREN, AND KARIN NORDSTRÖM	543		

Contributors

- Itziar Abete** Nutrition Research Center, University of Navarra, Pamplona, Navarra, Spain; Centro de Investigación Biomédica en red: Fisiopatología de la Obesidad y Nutrición (CiberObn), Madrid, Spain
- Xabier Agirre** Area de Hemato-Oncología, IDISNA, CIBERONC, Centro de Investigación Médica Aplicada (CIMA), University of Navarra, Pamplona, Spain
- Jennie Ahlgren** School of Education and Communication, Jönköping University, Jönköping, Sweden
- Gorka Alkorta-Aranburu** Area de Hemato-Oncología, CIMA Lab Diagnostics, IDISNA, Centro de Investigación Médica Aplicada (CIMA), University of Navarra, Pamplona, Spain
- Rocío Aller** Center of Investigation of Endocrinology and Nutrition, Medicine School and Department of Endocrinology and Nutrition, Hospital Clínico Universitario, University of Valladolid, Valladolid, Spain
- Cristina Andres-Lacueva** Biomarkers & Nutrimetabolomics Laboratory, Department of Nutrition, Food Sciences and Gastronomy, XaRTA, INSA, Faculty of Pharmacy and Food Sciences, University of Barcelona, Barcelona, Spain; CIBER de Fragilidad y Envejecimiento Saludable (CIBERFES), Instituto de Salud Carlos III, Barcelona, Spain
- Daniel Antonio de Luis** Center of Investigation of Endocrinology and Nutrition, Medicine School and Department of Endocrinology and Nutrition, Hospital Clínico Universitario, University of Valladolid, Valladolid, Spain
- Darwin Babino** Case Western Reserve University School of Medicine, Department of Pharmacology, Cleveland, OH, United States
- Shimrit Bar-El Dadon** The Center for Nutrigenomics and Functional Foods, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Israel
- Maria Luisa Bonet** Nutrigenomics and Obesity (NUO) Research Group, Laboratory of Molecular Biology, Nutrition and Biotechnology, Universitat de les Illes Balears, Palma de Mallorca, Spain; CIBER de Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Palma de Mallorca, Spain; Institut d'Investigació Sanitària Illes Balears (IdISBa), Palma de Mallorca, Spain
- Laura Bordoni** Unit of Molecular Biology, School of Pharmacy, University of Camerino, Camerino, Italy
- Patrick Borel** C2VN, Aix Marseille Univ, INRA, INSERM, Marseille, France
- Annet M. Bosch** Department of Pediatrics, Division of Metabolic Disorders, Amsterdam University Medical Centers, Amsterdam, the Netherlands
- Steven Brown** Division of Gastroenterology and Hepatology, UC San Diego, La Jolla, CA, United States
- Dolores Busso** Department of Nutrition, Diabetes, and Metabolism, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile
- Nadia Calabriso** National Research Council (CNR), Institute of Clinical Physiology, Lecce, Italy
- Antonio Capurso** Geriatrics and Internal Medicine, University of Bari Medical School, Bari, Italy
- Cristiano Capurso** Internal Medicine, University of Foggia, Foggia, Italy
- Marica Cariello** Department of Interdisciplinary Medicine, "Aldo Moro" University of Bari, Bari, Italy
- Maria Annunziata Carluccio** National Research Council (CNR), Institute of Clinical Physiology, Lecce, Italy
- Patricia Casas-Agustench** Biomarkers & Nutrimetabolomics Laboratory, Department of Nutrition, Food Sciences and Gastronomy, XaRTA, INSA, Faculty of Pharmacy and Food Sciences, University of Barcelona, Barcelona, Spain; CIBER de Fragilidad y Envejecimiento Saludable (CIBERFES), Instituto de Salud Carlos III, Barcelona, Spain
- Carol L. Cheatham** Department of Psychology & Neuroscience, Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, NC, United States
- Adela Chirita-Emandi** Genetics Department, Center of Genomic Medicine, University of Medicine and Pharmacy "Victor Babes", Timisoara, Romania
- Geetha Chittoor** Department of Nutrition and UNC Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, NC, United States
- Myung-Sook Choi** Center for Food and Nutritional Genomics Research, Department of Food and Nutrition, Kyungpook National University, San-Kyuk Dong Puk-Ku, Daegu, Korea
- Sang-Woon Choi** Chaum Life Center, CHA University, Seoul, Korea
- Paul Cordero** Institute for Liver and Digestive Health, University College London, London, United Kingdom
- Fernando J. Corrales** Functional Proteomics Laboratory, Centro Nacional de Biotecnología, CSIC, CIBERhed-ISCIII, PRB2, ProteoRed-ISCIII, Madrid, Spain

- Víctor Cortés** Department of Nutrition, Diabetes, and Metabolism, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile
- Nicholas C.P. Cross** Wessex Regional Genetics Laboratory, Salisbury NHS Foundation Trust, Salisbury, United Kingdom; Faculty of Medicine, University of Southampton, Southampton, United Kingdom
- Ana B. Crujeiras** Epigenomics in Endocrinology and Nutrition Group, Instituto de Investigación Sanitaria (IDIS), Complejo Hospitalario Universitario de Santiago (CHUS/SERGAS) and Santiago de Compostela University (USC), Santiago de Compostela, Spain; CIBER Fisiopatología de la Obesidad y Nutrición (CIBERObn), Madrid, Spain
- Rui Curi** Institute of Physical Activity and Sport Sciences, Cruzeiro do Sul University, Sao Paulo, Brazil
- Raffaele De Caterina** Cattedra di Malattie dell'Apparato Cardiovascolare, Università degli Studi di Pisa, Italy
- David de Lorenzo** Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, CEXS-UPF-PRBB, Barcelona, Catalonia, Spain; Centro de Estudios en Genómica y Nutrición-CESGEN, Parc Científic i Tecnològic Agroalimentari de Lleida-PCi TAL, Lleida, Catalonia, Spain
- Charles Desmarchelier** C2VN, Aix Marseille Univ, INRA, INSERM, Marseille, France
- Olivia Dong** Duke Center for Applied Genomics and Precision Medicine, Duke University School of Medicine
- Ramón Estruch** Faculty of Medicine and Health Sciences, Hospital Clinic, University of Barcelona, CIBEROBN (Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y Nutrición), Barcelona, Spain
- Teresa Ezponda** Area de Hemato-Oncología, IDISNA, CIBERONC, Centro de Investigación Médica Aplicada (CIMA), University of Navarra, Pamplona, Spain
- Michael Fenech** Genome Health Foundation, North Brighton, SA, Australia
- Lynnette R. Ferguson** Discipline of Nutrition and Dietetics, Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand
- Karina Fischer** Clinical Sciences & Epidemiology, R&D Cube Philip Morris International, Neuchatel, Switzerland; Department of Geriatrics and Aging Research & Centre on Aging and Mobility, University Hospital Zurich & University of Zurich, Zurich, Switzerland
- Luigi Fontana** Division of Geriatrics and Nutritional Sciences, Washington University in St. Louis, MO, United States; Department of Clinical and Experimental Sciences, Brescia University Medical School, Brescia, Italy; CEINGE Biotecnologie Avanzate, Napoli, Italy
- Simonetta Friso** University of Verona, School of Medicine, Verona, Italy
- Rosita Gabbianelli** Unit of Molecular Biology, School of Pharmacy, University of Camerino, Camerino, Italy
- Marat Garaulet** Department of Physiology, Faculty of Biology, University of Murcia, IMIB-Arrixaca, Murcia, Spain
- Oihane Garcia-Irigoyen** Department of Interdisciplinary Medicine, "Aldo Moro" University of Bari, Bari, Italy
- Angel Gil** Department of Biochemistry and Molecular Biology II, School of Pharmacy, University of Granada, Granada, Spain; Institute of Nutrition & Food Technology "Jose Mataix", Biomedical Research Center, University of Granada, Armilla, Spain; CIBEROBN (Physiopathology of Obesity and Nutrition CB12/03/30038), Instituto de Salud Carlos III (ISCIII), Madrid, Spain
- Purificación Gómez-Abellán** Department of Physiology, Faculty of Biology, University of Murcia, IMIB-Arrixaca, Murcia, Spain
- Ulf Görman** Centre for Theology and Religious Studies, Ethics division, Lund University, Lund, Sweden
- William S. Harris** Department of Internal Medicine, University of South Dakota; and OmegaQuant, LLC, Sioux Falls, SD, United States
- Alain de J. Hernandez-Vazquez** Genetics of Nutrition Unit, Institute of Biomedical Investigations, Universidad Nacional Autónoma de México, National Institute of Pediatrics, Ciudad de Mexico, Mexico
- Paul Hugenholtz** Erasmus Medical Centre, Rotterdam, The Netherlands
- Lara K. Hyde** Nourishable.tv, Litchfield, NH, United States
- Clara Ibáñez** Cancer Program, IMDEA Food Institute, Madrid, Spain
- Olatz Izaola** Center of Investigation of Endocrinology and Nutrition, Medicine School and Department of Endocrinology and Nutrition, Hospital Clinico Universitario, University of Valladolid, Valladolid, Spain
- Peter J.H. Jones** Nutritional Fundamentals for Health Inc., Vaudreuil-Dorion, QC, Canada
- Richard Kirwan** School of Natural Sciences and Psychology, Faculty of Science, Liverpool John Moores University, United Kingdom; Department of Nutrition, Food Science and Gastronomy, School of Pharmacy and Food Science, INSA-UB, University of Barcelona, Barcelona, Spain
- Martin Kohlmeier** Human Research Core and Nutrigenetics Laboratory, UNC Nutrition Research Institute; and Schools of Medicine and Public Health, Department of Nutrition, Kannapolis, NC, United States
- Natalia I. Krupenko** Department of Nutrition, Nutrition Research Institute, University of North Carolina Chapel Hill, Chapel Hill, NC, United States
- Eun-Young Kwon** Center for Food and Nutritional Genomics Research, Department of Food and Nutrition, Kyungpook National University, San-Kyuk Dong Puk-Ku, Daegu, Korea
- Dudley W. Lamming** Department of Medicine, University of Wisconsin—Madison, Madison, WI, United States; William S. Middleton Memorial Veterans Hospital, Madison, WI, United States; Molecular & Environmental Toxicology Center, University of Wisconsin—Madison, Madison, WI, United States

- Rosa M. Lamuela-Raventos** Department of Nutrition, Food Science and Gastronomy, School of Pharmacy and Food Sciences, INSA-UB, University of Barcelona, CIBEROBN (Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y Nutrición), Barcelona, Spain
- Dominique Langin** Institut National de la Santé et de la Recherche Médicale (Inserm), UMR1048, Institute of Metabolic and Cardiovascular Diseases, Toulouse, France; University of Toulouse, UMR1048, Institute of Metabolic and Cardiovascular Diseases, Paul Sabatier University, Toulouse, France; Franco-Czech Laboratory for Clinical Research on Obesity, Third Faculty of Medicine, Prague and Paul Sabatier University, Toulouse, France; Toulouse University Hospitals, Laboratory of Clinical Biochemistry, Toulouse, France
- Simon Langley-Evans** School of Biosciences, University of Nottingham, Loughborough, United Kingdom
- Cátia Lira do Amaral** Universidade Estadual de Goiás, Campus de Ciências Exatas e Tecnológicas, Anápolis, Goiás, Brazil
- Jesus Lopez-Minguez** Department of Physiology, Faculty of Biology, University of Murcia, IMIB-Arrixaca, Murcia, Spain
- Rosalinda Madonna** Center of Aging Sciences and Translational Medicine, CESI-MeT, "G. D'Annunzio" University, Chieti, Italy
- Maria L. Mansego** Bioinformatic Department, Making Genetics S.L., Noain, Spain
- Clarisse Marotz** Department of Pediatrics, UC San Diego, La Jolla, CA, United States
- J. Alfredo Martinez** Ciencias de la Alimentación y Fisiología, Universidad de Navarra, Navarra, Spain
- Marika Massaro** National Research Council (CNR), Institute of Clinical Physiology, Lecce, Italy
- Susan McRitchie** Nutrition Research Institute, University of North Carolina Chapel Hill, Kannapolis, NC, United States
- Bayan Mesmar** University of Maryland School of Medicine, Baltimore, MD, United States
- Fermín I. Milagro** Department of Nutrition, Food Science, and Physiology, Center for Nutrition Research, University of Navarra, Pamplona, Spain; CIBERObn, Centro de Investigación Biomédica en Red de la Fisiopatología de la Obesidad y Nutrición, Carlos III Health Institute, Madrid, Spain
- María J. Moreno-Aliaga** Department of Nutrition, Food Science, and Physiology, Center for Nutrition Research, University of Navarra, Pamplona, Spain; CIBERObn, Centro de Investigación Biomédica en Red de la Fisiopatología de la Obesidad y Nutrición, Carlos III Health Institute, Madrid, Spain
- Isabel Moreno-Indias** Unidad de Gestión Clínica de Endocrinología y Nutrición, Laboratorio del Instituto de Investigación Biomédica de Málaga (IBIMA), Hospital Universitario de Málaga (Virgen de la Victoria), Málaga, Spain; Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y la Nutrición (CIBEROBN), Madrid, Spain
- Antonio Moschetta** Department of Interdisciplinary Medicine, "Aldo Moro" University of Bari, Bari, Italy
- Santiago Navas-Carretero** Nutrition Research Center, University of Navarra, Pamplona, Navarra, Spain; Centro de Investigación Biomédica en red: Fisiopatología de la Obesidad y Nutrición (CiberObn), Madrid, Spain
- Mihai Niculescu** Genetics Department, Center of Genomic Medicine, University of Medicine and Pharmacy "Victor Babes", Timisoara, Romania; Advanced Nutrigenomics, Hillsborough, NC, United States
- Karin Nordström** Theologische Fakultät, Universität Luzern, Luzern, Switzerland
- Francisco J. Novo** Department of Biochemistry & Genetics, School of Sciences, University of Navarra, Pamplona, Spain
- Jude A. Oben** Institute for Liver and Digestive Health, University College London, London, United Kingdom; Department of Gastroenterology and Hepatology, Guy's and St. Thomas' Hospital, NHS Foundation Trust, London, United Kingdom
- Leticia Odriozola** CIMA, University of Navarra, Pamplona, Spain
- Jose M. Ordovas** JM-USDA-HNRCA at Tufts University, Boston, MA, United States; USA IMDEA-Alimentacion and Centro Nacional Investigaciones Cardiovasculares (CNIC), Madrid, Spain
- Andreu Palou** Nutrigenomics and Obesity (NUO) Research Group, Laboratory of Molecular Biology, Nutrition and Biotechnology, Universitat de les Illes Balears, Palma de Mallorca, Spain; CIBER de Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Palma de Mallorca, Spain; Institut d'Investigació Sanitària Illes Balears (IdISBa), Palma de Mallorca, Spain
- Virginia R. Parslow** Department of Optometry and Vision, Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand
- Wimal Pathmasiri** Nutrition Research Institute, University of North Carolina Chapel Hill, Kannapolis, NC, United States
- José Luis Pérez-Castrillón** Center of Investigation of Endocrinology and Nutrition, Medical School and Department of Internal Medicine, Hospital Universitario Río Hortega, University of Valladolid, Valladolid, Spain
- Louis Pérusse** Department of Kinesiology, Université Laval, Québec, QC, Canada
- Elena Piccinin** Department of Interdisciplinary Medicine, "Aldo Moro" University of Bari, Bari, Italy
- Julio Plaza-Diaz** Department of Biochemistry and Molecular Biology II, School of Pharmacy, University of Granada, Granada, Spain; Institute of Nutrition & Food Technology "Jose Mataix", Biomedical Research Center, University of Granada, Armilla, Spain

- Felipe Prósper** Area de Hemato-Oncología, IDISNA, CIBERONC, Centro de Investigación Médica Aplicada (CIMA), University of Navarra, Pamplona, Spain
- Lu Qi** Department of Epidemiology, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA, United States; Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, United States
- Jessica C. Ralston** Nutrigenomics Research Group, UCD Conway Institute of Biomolecular Research and School of Public Health, Physiotherapy, and Sports Science, University College Dublin, Dublin, Republic of Ireland
- Ana Ramírez de Molina** Cancer Program, IMDEA Food Institute, Madrid, Spain
- George Rasti** Chromatin Biology Laboratory, Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain
- Ram Reifen** The Center for Nutrigenomics and Functional Foods, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Israel
- Giulia Renda** G. d'Annunzio University of Chieti and Center of Excellence on Aging CeSI-Met, Chieti, Italy
- José A. Riancho** Department of Medicine, School of Medicine, Department of Internal Medicine, Hospital Universitario Marques de Valdecilla University of Cantabria, IDIVAL, Santander, Spain
- José Antonio Riancho del Moral** Department of Medicine, School of Medicine, Department of Internal Medicine, Hospital Universitario Marques de Valdecilla, University of Cantabria, Santander, Spain
- Fernando Rivadeneira** Department of Internal Medicine, Erasmus University Medical Centre, Rotterdam, The Netherlands
- Helen M. Roche** Nutrigenomics Research Group, UCD Conway Institute of Biomolecular Research and School of Public Health, Physiotherapy, and Sports Science, University College Dublin, Dublin, Republic of Ireland; Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, Belfast, Northern Ireland, United Kingdom
- Marta Ruiz-Mambrilla** Medical Center of Rehabilitation and Speech Disorders and School of Medicine, University of Valladolid, Valladolid, Spain
- Francisco Javier Ruiz-Ojeda** Department of Biochemistry and Molecular Biology II, School of Pharmacy, University of Granada, Granada, Spain; Institute of Nutrition & Food Technology "Jose Mataix", Biomedical Research Center, University of Granada, Armilla, Spain
- Francisca Salas-Pérez** Department of Nutrition, Diabetes, and Metabolism, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile
- Rodrigo San-Cristobal** Cardiometabolic Nutrition Group, Precision Nutrition and Cardiometabolic Health, IMDEA Food Institute, Madrid, Spain
- Nicolás Santander** Department of Nutrition, Diabetes, and Metabolism, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile
- José L. Santos** Department of Nutrition, Diabetes, and Metabolism, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile
- Jörg Saupe** MVZ Diaverum Erkelenz GmbH, Erkelenz, Germany
- Egeria Scoditti** National Research Council (CNR), Institute of Clinical Physiology, Lecce, Italy
- Charles N. Serhan** Center for Experimental Therapeutics and Reperfusion Injury, Hale Building for Transformative Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States
- Nicolas G. Simonet** Chromatin Biology Laboratory, Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain
- Artemis P. Simopoulos** The Center for Genetics, Nutrition and Health, Washington, DC, United States
- Nanette Steinle** University of Maryland School of Medicine, Baltimore, MD, United States; VA Maryland Healthcare System, Baltimore, MD, United States
- Susan C.J. Sumner** Nutrition Research Institute, University of North Carolina Chapel Hill, Kannapolis, NC, United States
- Francisco J. Tinahones** Unidad de Gestión Clínica de Endocrinología y Nutrición, Laboratorio del Instituto de Investigación Biomédica de Málaga (IBIMA), Hospital Universitario de Málaga (Virgen de la Victoria), Málaga, Spain; Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y la Nutrición (CIBEROBN), Madrid, Spain
- Alejandro Vaquero** Chromatin Biology Laboratory, Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain
- Itzel Vazquez-Vidal** Richardson Centre for Functional Foods and Nutraceuticals, University of Manitoba, Winnipeg, MB, Canada
- Antonio Velazquez-Arellano** Genetics of Nutrition Unit, Institute of Biomedical Investigations, Universidad Nacional Autónoma de México, National Institute of Pediatrics, Ciudad de Mexico, Mexico
- Nathalie Viguerie** Institut National de la Santé et de la Recherche Médicale (Inserm), UMR1048, Institute of Metabolic and Cardiovascular Diseases, Toulouse, France; University of Toulouse, UMR1048, Institute of Metabolic and Cardiovascular Diseases, Paul Sabatier University, Toulouse, France; Franco-Czech Laboratory for Clinical Research on Obesity, Third Faculty of Medicine, Prague and Paul Sabatier University, Toulouse, France
- Manlio Vinciguerra** Institute for Liver and Digestive Health, University College London, London, United Kingdom; Center for Translational Medicine (CTM), International Clinical Research Centre (ICRC), St. Anne's University Hospital, Brno, Czech Republic

- Francesco Visioli** Department of Molecular Medicine, University of Padova, Padova, Italy; IMDEA-Food, CEI UAM+CSIC, Madrid, Spain
- José L. Vizmanos** Department of Biochemistry & Genetics, School of Sciences, University of Navarra, Pamplona, Spain
- Johannes von Lintig** Case Western Reserve University School of Medicine, Department of Pharmacology, Cleveland, OH, United States
- Venkata Saroja Voruganti** Department of Nutrition and UNC Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, NC, United States
- Ronald J.A. Wanders** Laboratory Genetic Metabolic Diseases, Departments of Paediatrics and Clinical Chemistry, Emma Children's Hospital, Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands
- Tiange Wang** Shanghai Institute of Endocrine and Metabolic Diseases, Rui Jin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; Department of Epidemiology, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA, United States
- Tim Wiltshire** UNC Eshelman School of Pharmacy and UNC Center for Pharmacogenomics and Individualized Therapy, Kannapolis, NC, United States
- Deyang Yu** Department of Medicine, University of Wisconsin–Madison, Madison, WI, United States; William S. Middleton Memorial Veterans Hospital, Madison, WI, United States; Molecular & Environmental Toxicology Center, University of Wisconsin–Madison, Madison, WI, United States
- Amir Zarrinpar** Division of Gastroenterology and Hepatology, UC San Diego, La Jolla, CA, United States; VA San Diego Health Sciences, La Jolla, CA, United States
- Steven H. Zeisel** Nutrition Research Institute, School of Public Health and School of Medicine, The University of North Carolina at Chapel Hill, Kannapolis, NC, United States
- Maria Angeles Zulet** Nutrition Research Center, University of Navarra, Pamplona, Navarra, Spain; Centro de Investigación Biomédica en red: Fisiopatología de la Obesidad y Nutrición (CiberObn), Madrid, Spain; Instituto de Investigación Sanitaria de Navarra (IDISNA), Recinto de Complejo Hospitalario de Navarra, Pamplona, Navarra, Spain

This page intentionally left blank

Editor Biographies

Prof. Raffaele De Caterina, MD., PhD

Cattedra di Malattie dell'Apparato Cardiovascolare, Università degli Studi di Pisa, Italy

Raffaele De Caterina is a professor of cardiology at the University of Pisa and director of the University Cardiology Division at Pisa University Hospital, Pisa, Italy. He has been president of the International Society on Nutrigenetics/Nutrigenomics (2009–13) and vice-president of the European Society of Cardiology (2008–10). He is editor-in-chief of the journal *Vascular Pharmacology* and author of over 600 peer-reviewed manuscripts. His major research interests include cardiovascular pharmacology, atrial fibrillation, coronary artery disease, and cardiovascular thrombosis, as well as nutrigenetics and nutrigenomics.

J. Alfredo Martínez, MD., PhD

Ciencias de la Alimentación y Fisiología, Universidad de Navarra, Navarra, Spain

Prof. J. Alfredo Martínez has been involved in several landmark intervention trials and more than 50 national and international projects. He has supervised more than 70 PhD students and published more than 800 peer-reviewed manuscripts and book chapters in the areas of obesity and personalized nutrition, including precision nutritional omics. His works have been cited more than 26,000 times, and his HFactor is over 70. He has been president of the International Society of Nutrigenetics/Nutrigenomics (2015–2017) and is president of the International Union of Nutritional Sciences (IUNS) for the 2017–21 period. Prof. Martínez has enjoyed training or invited stays at the University of Nottingham, Berkeley College, MIT, Harvard University, Oxford University, and King's College London. He has been an advisor for EU projects or a guest speaker or main keynote lecturer in important forums concerning precision nutrition in obesity and cardiometabolic diseases, including IUNS, ISNN, OMS, FAO, IAEA, NIH, SLAN, and FENS.

Martin Kohlmeier

Human Research Core and Nutrigenetics Laboratory, UNC Nutrition Research Institute; and Schools of Medicine and Public Health, Department of Nutrition, Kannapolis, NC, United States

Dr. Martin Kohlmeier, MD, PhD, is a professor of nutrition; director of the Human Research Core and the Nutrigenetics Laboratory at the UNC Nutrition Research Institute in Kannapolis, NC; Director of Nutrition in Medicine for online nutrition education of medical students and health professionals; visiting fellow of Wolfson College at Cambridge University, UK; founding editor-in-chief of *BMJ Nutrition, Prevention and Health*; and president of the International Society of Nutrigenetics/Nutrigenomics. His professional interests include investigating genotype-specific nutrition responses of humans in health and disease, developing decision-support algorithms and tools for precision nutrition, and teaching health care professionals how to use nutrigenetics in practice. His previously authored textbooks in the field include *Nutrigenetics* (Academic Press, 2013) and *Nutrient Metabolism* (2nd ed., Academic Press, 2015).

This page intentionally left blank

Preface

Individualized nutrition based on nutrigenetics and nutrigenomics knowledge is becoming an important area of scientific and health interest globally. This should increase current awareness and applications to benefit personalized well-being, as well as public health. The accessibility of modern technologies, the availability of recognized and new biomarkers based on genomic and other omics data, and a deeper understanding of genotype information are three relevant aspects that have been developed and investigated in relation to personalized nutrition.

Nutrition and dietary intake are essential factors in the interactions between environment and genes for achieving a healthy status. Personalized nutrition should be based on the principle that food or nutrients may be either risk or protection factors for preventing and treating various diseases, depending on the genetic predisposition of the individual (nutrigenetics) and on their ability to regulate gene expression (nutrigenomics). The implementation of personalized nutrition therefore depends on genetic background information (e.g., heritage and epigenetic markers), biological/medical conditions (history of diseases, intolerances or allergies, family considerations, etc.), environmental issues (physical activity and dietary patterns), and cultural variations (e.g., food preferences, religion, and food accessibility) affecting the individual.

Nutritional “omics” approaches (e.g., transcriptomics, proteomics, metabolomics, foodomics, lipidomics, metagenomics, etc.) to explain how foods/nutrients and genes interact are evolving to reveal differences in individual phenotypes, including individualized health or disease conditions. Research in nutrigenetics, nutrigenomics, nutriepigenetics, and systems biology can interactively contribute to personalized nutrition, and need to be considered to provide reliable biomarkers of dietary response or to estimate individual requirements. The occurrence of “single nucleotide polymorphisms” (SNPs) in candidate genes may interact with the energy, micronutrients, and macronutrients provided by the diet. Interindividual differences in disease susceptibility depend not only on the DNA sequence (e.g., SNPs) but also on epigenetic factors affecting gene expression such as DNA methylation, covalent histone modifications, chromatin folding, and the regulatory actions of miRNA. From an epigenetics perspective, the identification of those individuals that at an early age present changes in the methylation profiles of specific genes could help to predict their susceptibility to developing obesity and noncommunicable chronic diseases in later life.

Advances in nutrigenetics and nutrigenomics may allow the monitoring of health maintenance and disease progress. They may also facilitate research and development of newer preventive and therapeutic approaches as well as protocols for tailored dietary treatments based on algorithms. However, some challenges remain concerning nutrient \times gene interactions for the prevention and treatment of nutritional conditions. One of them is that the prevalence of most SNPs differs in importance depending on ethnic background. This means that studies with large and/or mixed populations, and with different racial subgroups, are required. The cost of genetic analyses and personalized advice will have to be reduced for the use of nutritionally related omics in public health and population practices. Implementation of regulations and provision of reliable information about the benefits and limitations of using genetic tests for consumers are still required as well as the clarification of emerging ethical issues.

The International Union of Nutritional Sciences (IUNS), particularly through the IUNS Task Force “Nutrition–gene interactions: knowledge for action and Precision Nutrition”, has been very supportive of the training and dissemination efforts of the International Society of Nutrigenetics/Nutrigenomics (ISNN). This collaboration must be seen in the context of numerous past and future seminars, courses, and sessions held in Geneva (ISNN), Pamplona (ISNN), Granada (20th ICN-IUNS), Gold Coast, Australia (ISNN), Chapel Hill, USA (ISNN), Tel-Aviv (ISNN), Buenos Aires (21st ICN-IUNS), Cambridge, UK (ISNN), and in Tokyo (22nd ICN-IUNS). The efforts of ISNN finally culminated in the completion of this seminal work.

The number of publications related to nutrigenetics and nutrigenomics is continuously growing (more than 1000 peer-reviewed documents with more than 17,500 citations since 2001), paving the way to successful and efficient precision nutrition. This textbook consolidates up-to-date information on nutrigenetics and nutrigenomics with

examples of important applications for nutritionists and other health professionals to become familiar with genetic information and its translation into personalized nutrition. Education at undergraduate and postgraduate levels concerning nutrigenetics and nutrigenomics sciences should be encouraged as prompted by this textbook.

Prof Catherine Geissler
IUNS General Secretary
Professor Emerita of Human Nutrition, King's College London

SUPPLEMENTARY DATA

Please find the supplementary PPT and video files at the companion site: <https://www.elsevier.com/books-and-journals/book-companion/9780128045725>.

Password: k7pogk8t2e

Introduction

Following the seminal discoveries of gene structure and of the genetic code in the 1950s, the past half century has heralded incredible progress in the understanding of human development and gene expression regulation. Emphasis has thus gradually shifted from broad interpretation of common interindividual diversities among humans to make each person an individual, with unique features differing even among monozygotic twins because of the complex interaction of genes and the environment. Today's omics techniques allow the relatively inexpensive sequencing of each individual's genome and the tracing of complex gene variations in diverse populations. These advanced technologies have helped to understand that common variants occur in all regions of the genome. Many of the most common variants are neutral, meaning that they do not alter the function or expression of a gene under most conditions, but we may still learn that important circumstances will uncover underlying differences. A polymorphic variant of a gene can be associated with a difference in expression or with the production of a different form of the protein; such differences may be associated with an increased likelihood of developing a disease, or with a difference in susceptibility to exposures such as foods, nutrients, body composition, or physical activity. Even with exactly the same gene sequence, nutritional exposures may however result in differing gene expression, due to the occurrence of epigenetic changes that do not alter the gene sequence and yet change gene expression. Appreciation of the plasticity of gene expression has been a major theme in genetics and indeed precision medicine in recent years. Our molecular blueprint, including both genetic and epigenetic information, provides the instructions for our responses to nutrition and other exposures. This situation also applies to the way we respond to amounts and composition of dietary intake. What we are responding to are food patterns, individual foods such as milk, nutrients, bioactives, and more, always both in terms of quantity and quality.

Nutrition is a main environmental variable able to affect gene expression. The general theory of the operon, first proposed in a short paper in the Proceedings of the French Academy of Science in 1960, suggested that genes can be negatively controlled by a repressor acting at a single operator located before the first gene. Later, it was discovered that genes could be positively regulated and also regulated at steps that follow transcription initiation. The development of the concept is considered a landmark event in the history of molecular biology and yielded the Nobel Prize in Physiology or Medicine to François Jacob, André Michel Lwoff, and Jacques Monod in 1965. In their first description of the operon system in *Escherichia coli*, Jacob and Monod found that one nutrient, lactose, when added to the model bacterium *E. coli* was able to regulate the expression of a cluster of sequential genes able to drive lactose metabolism. Today several examples of gene expression regulations by nutrients have been described, making nutritional environment the most important way of regulating the expression of our genes.

Because of the increasing awareness of the relevance of genetic and epigenetic variation in conditioning the response to nutritional intakes, and on the parallel development of knowledge about how the nutritional environment may affect gene expression, the new discipline of nutrigenetics/nutrigenomics was born, and took shape in the establishment of a scientific society, the International Society of Nutrigenetics/Nutrigenomics (ISNN) in September 2005. Although differently used in different frameworks, the two terms of *nutrigenetics* and *nutrigenomics* were carefully defined in the early steps of establishing the International Society. Concepts and research on genetic variation and dietary response (i.e., individuals responding differently to the same diet by having different levels of, for example, serum cholesterol and blood pressure because of genetic variation) are encompassed by the term nutrigenetics; and studies on the genome-modifying aspects of diet and the role of nutrients in gene expression are encompassed by the term nutrigenomics. In addition to nutrients, nonnutritive dietary phytochemicals, for example phenolic compounds, are being studied for their effects on various aspects of human metabolism. Nutrigenomics could provide a framework for the development of novel foods that will be genotype dependent for the promotion of health and prevention and management of chronic diseases, thus yielding the fundamentals of personalized nutrition, within the broader context of personalized medicine.

Within this framework, we editors, all former or current presidents of the ISNN, embarked in the first attempt at defining the scope of this newborn discipline with an ambitious comprehensive textbook, now seeing the light after several years of incubation. *Principles of Nutrigenetics and Nutrigenomics-Fundamentals for Individualized Nutrition* aims at covering most of the topics broadly falling under the umbrella of the newborn discipline. Thus, after an introductory section dealing with concepts of the nature of genetic traits, genes and variation, analytical methods and statistical genetics, and molecular biology of genetic variants, a series of chapters devoted to topics dealing with genetic variation and dietary response (nutrigenetics) and a series of chapters properly devoted to nutrients and gene expression (nutrigenomics) follow. The last part of the textbook is devoted to translational aspects in nutrigenetics/nutrigenomics, also taking into account ethical and legal aspects connected with the use of genetic information for nutritional advice.

This textbook must also be seen in the context of fruitful collaborations with the NutriGenomic Organization (NuGO) organization and specially with the International Union of Nutritional Sciences (IUNS) task forces about “Nutrition–gene interactions: knowledge for action and Precision Nutrition” and shared efforts on numerous past and future seminars, courses, and sessions held in Geneva (ISNN), Pamplona (ISNN), Granada (20th ICN-IUNS), Gold Coast, Australia (ISNN), Barcelona (NuGO), Chapel Hill, USA (ISNN), Copenhagen (NuGO), Tel-Aviv (ISNN), Buenos Aires (21st ICN-IUNS), Cambridge, UK (ISNN), and in Tokyo (22nd ICN-IUNS). Furthermore this textbook is a culmination of the publications of the proceedings of different ISNN Congresses as well as three guides and position papers of the ISNN board in the *Journal of Nutrigenetics and Nutrigenomics*.

Despite the difficulties encountered in handling such a broad set of topics for the first time with a comprehensive scope, substantially differing from previous attempts aimed essentially at collecting anecdotal coverage, we, the editors, on behalf of the ISNN, are proud to offer this reference text to scientists, professionals, and practitioners approaching the discipline from the many angles that it may involve. We trust this will be an important step forward in defining the discipline, its most relevant background elements, and—as the subtitle says—offering, as a practical fallout, the fundamentals of personalized nutrition.

Raffaele De Caterina
J. Alfredo Martinez
Martin Kohlmeier

S E C T I O N I

The Biological Basis of
Heritability and Diversity

This page intentionally left blank

The Nature of Traits, Genes and Variation

José L. Vizmanos¹, Nicholas C.P. Cross^{2,3}, Francisco J. Novo¹

¹Department of Biochemistry & Genetics, School of Sciences, University of Navarra, Pamplona, Spain; ²Wessex Regional Genetics Laboratory, Salisbury NHS Foundation Trust, Salisbury, United Kingdom; ³Faculty of Medicine, University of Southampton, Southampton, United Kingdom

Glossary

Alleles The variants of a locus; for autosomal loci, each diploid individual has two alleles.

Autosomes The chromosomes that are not different between sexes. In humans there are 22 pairs of autosomes.

Chromosomes Linear molecules of DNA packaged with proteins.

Codominant Alleles that are fully expressed in the heterozygote.

Continuous, metric or quantitative Phenotypic traits that do not show distinguishable forms but a continuous gradation in the population. These traits are also inherited but they are multifactorial characters controlled by multiple genes (**polygenic** characters), and affected to varying degrees by the environment and epigenetic modifications.

Discrete or discontinuous Traits that are often encoded by only one gene (**monogenic** traits) with two different alleles in the population, resulting in alternative forms of the trait.

Dominant An allele that is fully expressed either in the homozygote or in the heterozygote. In **intermediate dominance** one of the alleles is not completely recessive or the other is not completely dominant.

Epistasis A type of gene interaction. Some traits are affected by more than one gene so that the effect of one gene (epistatic) masks or modifies the effect of another gene (hypostatic).

Expressivity Reflects the degree in which a trait is expressed among individuals with identical genotype.

Gene A segment of the genome that contains the instructions for a functional product. It is considered the molecular unit of heredity.

Genome The heritable material that is transmitted from parents to offspring and confers the biological identity of an individual. It is composed of large molecules of nucleic acids whose sequence contains the instructions to build other functional molecules such as ribonucleic acid (RNA) or proteins.

Genotype The genetic makeup of an organism, the whole set of genes. For a given gene, the genotype is the combination of alleles in a particular individual.

Hemizygous Are males for the genes located in the X chromosome because they only have one copy.

Heritability Of a trait is the contribution of genetic variation to phenotypic variation and it can be defined as the fraction of total phenotypic variance that is due to genetic variance.

Heterozygote An individual with two different alleles (in sequence or function) at a locus.

Homologous chromosomes A pair of chromosomes that carry the same genes but are inherited from a different parent. A diploid

individual ($2n$) harbors a number n (haploid) of pairs of homologous chromosomes.

Homozygote An individual with two identical alleles (in sequence or function).

Linkage Is due to the fact that alleles are located along chromosomes and tend to be inherited in the same combination that is present in parental chromosomes. Linkage is broken by recombination during meiosis that creates novel combinations of alleles in the chromosome.

Locus (plural **loci**) A position (place) in a chromosome; the physical region occupied by a gene, a marker or any genetic element.

Mendelian Traits that display Mendelian inheritance typical of diploid organisms.

Penetrance The proportion of individuals that show the typical phenotype associated with a genotype.

Phenotype The collection of observable characteristics of the organism. It is determined by the genotype but also by environmental factors and epigenetic modifications (both inherited and noninherited).

Quantitative trait locus (QTL) A locus controlling a quantitative trait.

Recessive An allele that is fully expressed only in the homozygote.

Sexual chromosomes The chromosomes that are different in each sex. In humans, females have two X chromosomes, whereas males have one X and one Y chromosome.

Single-nucleotide polymorphisms (SNPs) Polymorphic sites at single nucleotides.

Structural variants (SVs) Polymorphisms that affect relatively large segments of the genome (from Kb to Mb). A type of SV is the

Copy number variations (CNVs) regions that are repeated at different number of times in different individuals in the population.

PHENOTYPIC TRAITS

Gene, Genome, Genotype, and Phenotype

The biological features of living organisms are encoded in genomes. The **genome** is the heritable material that is transmitted from parents to offspring and confers the biological identity of an individual. It is composed of large molecules of nucleic acids (like deoxyribonucleic acids, DNA) made up of four different

monomers (adenine, guanine, cytosine, and thymine). The sequence of these monomers contains the instructions to build other functional molecules such as ribonucleic acids (RNAs) or proteins. The segment of the nucleic acid that contains the instructions for one of these molecules is a **gene**. Therefore, a gene encodes a functional molecule and is considered the molecular unit of heredity.

The **genotype** is the genetic makeup of an organism, the whole set of genes. Genotype is one of the determinants of the **phenotype**, the observable characteristics of the organism known as **traits**. These characteristics range from biochemical and physiological properties to morphology or behavior. The phenotype is also determined by environmental factors together with inherited and noninherited epigenetic modifications.

The genome of eukaryotic organisms is distributed in two main cellular locations: most of the genome is in the nucleus (**nuclear genome**), but some of it is outside (the **extranuclear genome**) into organelles like mitochondria and chloroplasts (in plants). Both genomes have a different structure and mode of inheritance.

The nuclear genome is composed of linear molecules of DNA packaged with proteins into **chromosomes**. Diploid organisms, like humans, have two copies of each chromosome, each of them inherited from one parent. Thus, for each nuclear gene, there are two copies or **alleles**. Both alleles may be identical (in sequence or function) or different, the individual being **homozygous** or **heterozygous**, respectively. Any gene is a segment of a chromosome (it holds a **locus**, plural **loci**), so the chromosome can be considered a linear array of genes. Chromosomes that carry the same genes but are inherited from a different parent are called **homologous**. Therefore, a diploid individual ($2n$) harbors a number n (haploid) of pairs of homologous chromosomes. In humans, n is 23, so we have 46 chromosomes. Any individual (regardless of the sex) has the same 22 pairs of chromosomes, known as **autosomes**. The remaining pair, known as **sexual chromosomes**, is different in each sex: females have two X chromosomes and males have one X and one Y chromosome.

The fact that genes (and variants) are located along chromosomes leads to a situation known as **linkage**. During recombination in meiosis, fragments of homologous chromosomes will be shuffled due to crossing-over, in order to create greater genetic diversity. This will recombine the parental genotypes to yield novel combinations of alleles, but this process depends on the distance between genes: loci that are far apart will be recombined more frequently than genes located very close to each other in the same chromosome.

Linkage analysis has been crucial to create genetic maps and to identify loci responsible for or associated with genetic diseases.

The other eukaryotic genome, the extranuclear genome (**mitochondrial genome** in mammals), is composed of a circular DNA molecule structurally similar to a prokaryotic genome, repeated in many copies inside the organelle so that a cell includes hundreds of copies. As the mitochondria of the zygote come from the ovum, this genome is inherited only from the mother.

In humans, the nuclear genome accounts for 99.9995% of genetic information and mitochondrial genome for the remaining 0.0005%, with only 37 protein-coding genes mainly involved in oxidative phosphorylation.

Types of Traits: Discrete Traits and Continuous Traits

Some phenotypic traits show easily distinguishable forms, like cleft chin (chin with a dimple) in humans. These **discrete** or **discontinuous** traits are often encoded by only one gene (**monogenic** traits) with two different alleles in the population, resulting in the alternative forms of the trait. Sometimes one of the alleles is **recessive** and the other is **dominant**. In order to show the recessive form of the trait, the individual must be homozygous for that allele, whereas the dominant form is expressed when only one or two copies of the dominant allele are present. If the heterozygous individual shows an intermediate phenotype, both alleles could be **codominant** (if both products are expressed) or have an **intermediate dominance** (one of them is not completely recessive or the other is not completely dominant). If there are multiple alleles in the population, there could be variable interactions between them (some of them being dominant/recessive and some of them being codominant or with intermediate dominance).

Conversely, some phenotypic traits do not show distinguishable forms but a continuous gradation in the population (like height and weight). These traits are called **continuous**, metric or **quantitative** traits. Continuous traits are also inherited but do not manifest typical patterns of dominant or recessive inheritance. In fact, they are multifactorial characters: controlled by multiple genes (**polygenic** characters), but also affected to varying degrees by the environment and epigenetic modifications. Many susceptibility traits show this mode of genetic determination, so that a disease state will only be apparent when a threshold value is surpassed.

Mendelian Transmission and Polygenic Inheritance: QTLs and Heritability

Many of the genetic diseases due to mutations in one gene are discrete traits. Discontinuous traits display typical Mendelian inheritance, so sometimes they are called **Mendelian** traits. Mendelian inheritance is typical of diploid organisms. For any given gene, any individual inherits two copies or alleles, one from each parent, which separate randomly (**segregate**) from each other during the production of gametes (ovum or sperm). Gametes are haploid cells and carry only one allele of the gene. During fertilization, both gametes fuse and a new diploid individual arises again with two gene copies, one from each gamete. The mode of inheritance depends on whether a gene is located in an autosome or in a sexual chromosome, because males possess only one allele for the genes located in the chromosome X (they are **hemizygous**).

Thus, when a genetic disorder is caused by a mutated allele, the inheritance of this disorder (or trait) can be **autosomal recessive**, **autosomal dominant**, **X-linked recessive**, **X-linked dominant**, or **Y-linked**. For an autosomal trait, any individual of any sex has two alleles; if the trait is dominant, heterozygous individuals will express the character, but if it is recessive they will not. For X-linked traits, genes are located in the X chromosome so inheritance in females will be similar to autosomal traits, but in males (who only have one chromosome X) the trait will be expressed regardless of whether it is dominant or recessive. Y-linked traits (holandric traits) will be expressed only in males and transmitted from father to son (Fig. 1.1).

Continuous traits involve many genes (polygenic traits), each having a small effect further modified by environmental factors. The result is a normally distributed trait in which each contributing gene shows

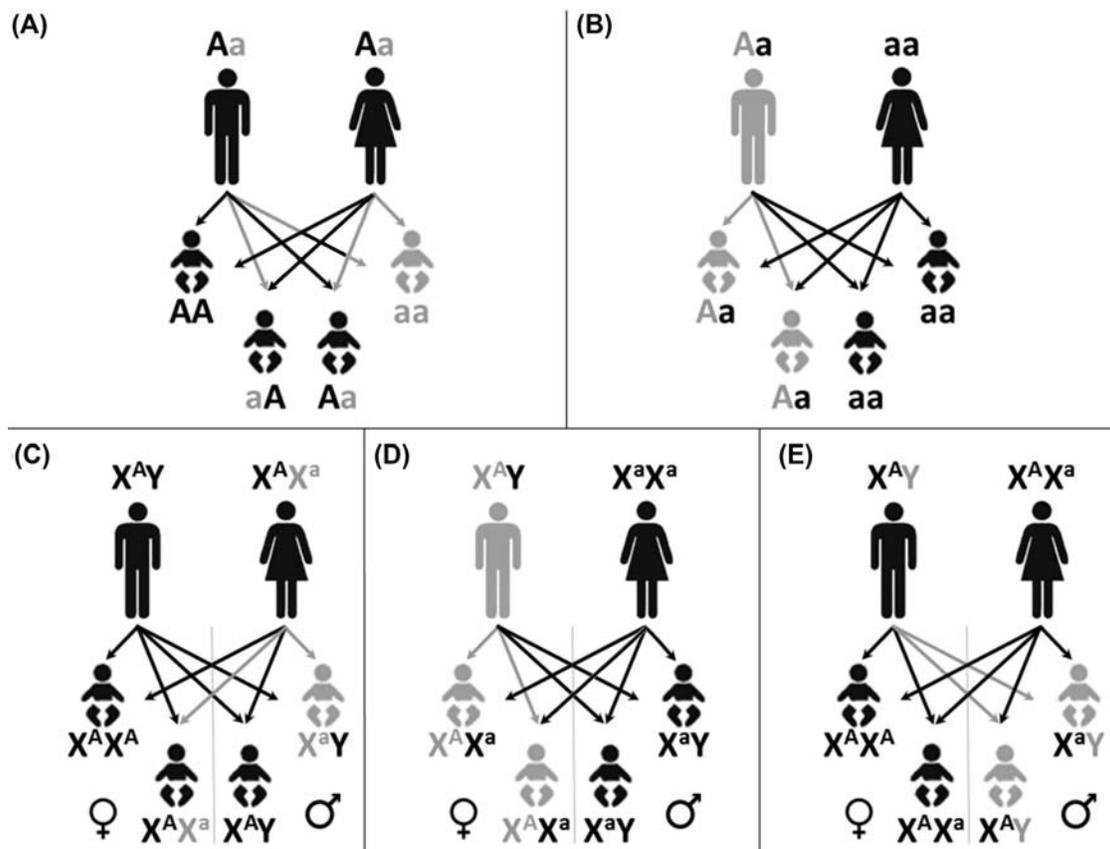


FIGURE 1.1 Types of Mendelian inheritance. (A) **Autosomal recessive inheritance.** The character (e.g., a disease) is caused by a recessive allele (lower case). To show disease manifestations, the individual must be homozygous. Each allele has been transmitted from each of the parents. Heterozygous individuals (carriers) do not show the disease. (B) **Autosomal dominant inheritance.** The disease is caused by a dominant allele (capital letter), individuals show the disease with only one copy of this allele. (C) **X-linked recessive inheritance.** Similar to (A) but the gene is located on the X chromosome. Females have two X chromosomes but males only one, so they will have only one allele (are hemizygous) and the disease could appear with one recessive allele. The X chromosome of males is transmitted to all of his daughters. (D) **X-linked dominant inheritance.** Similar to (B) but, again, the gene is located on the X chromosome. Transmission will be different if the allele comes from the father or from the mother; if the father is affected, then all his daughters will be affected. (E) **Y-linked inheritance.** The gene is located on the Y chromosome, so it is transmitted only in the male lineage.

Mendelian inheritance (Fig. 1.2). A gene involved in a quantitative trait is termed a **quantitative trait locus (QTL)**, which can be identified by several analytical approaches. Consequently, QTLs are frequently associated with a single quantitative trait. QTL identification can explain the genetic architecture of a trait, for example, if it is controlled by many genes of small effect or by a few genes of large effect. The study of these traits relies on statistical analysis to describe the distribution of the trait and to infer and compare population data.

Environmental factors play a key role in quantitative traits. At the beginning of 20th century, Francis Galton used the terms **nature** and **nurture** to distinguish the roles played by biologic inheritance and environmental factors on the development of a trait. As quantitative traits depend on both heredity and environment, it is not always possible to know the exact contribution of each factor, although the relative contribution of each to the variability of the trait in the population can be estimated. Variation of a trait in a population (total variance) is the result of genetic variation (genetic variance) and environmental variation (environmental variance) plus some interaction between both factors. The contribution of genetic variation to phenotypic variation is known as the **heritability** of a trait. This concept can be defined as the fraction of total phenotypic variance that is due to genetic variance. Although measurement of phenotypic variance can be relatively straightforward, this is not always the case with genotypic variance, but some methods have been developed

to do this. It is important to stress that heritability does not measure the degree in which a trait is determined by genes, but the proportion of the phenotypic variation among individuals that is due to genetic variation.

Key Concepts About Gene Interaction

Few traits are determined by just one gene. An organism typically has many thousands of genes in its genome and the development of genetic traits can be the result of complex interactions between alleles of a gene, between genes, and between genes and the environment. Thus, alleles can show different types of interactions: they can be dominant, recessive, codominant, or incompletely dominant. Some traits, e.g., quantitative traits, are the result of polygenic inheritance and gene-environment interactions. Other traits are affected by more than one gene in a different mode of interaction known as **epistasis**. In epistasis, the effect of one gene (epistatic) masks or modifies the effect of another gene (hypostatic). For example, in metabolic reactions, successive biochemical transformations are required so that loss of activity of an initial enzyme in the pathway can mask the activities of enzymes involved in subsequent transformations. Sometimes the presence of some alleles in different genes can induce **novel phenotypes** by novel gene combinations not present in previous generations.

In addition, phenotypic expression is not always a direct reflection of the genotype. This can be due to

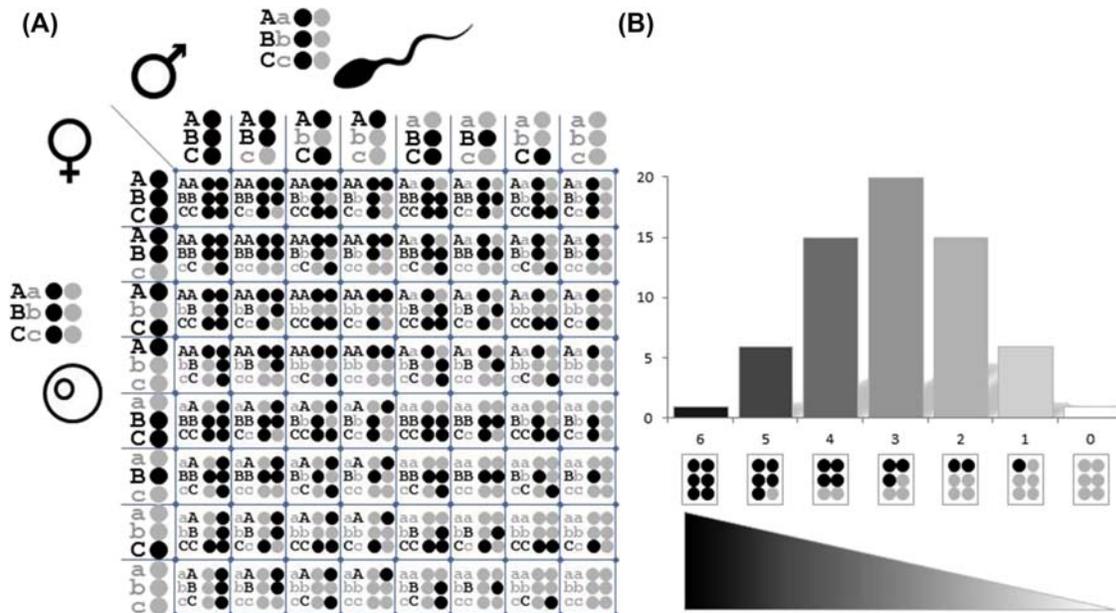


FIGURE 1.2 Mendelian inheritance of quantitative traits. In this case, the trait is determined by three different genes, each with two alleles (black and gray circles). The effects of these genes are additive. In this hypothetical case, crossing between two individuals heterozygous for the three genes will render multiple possibilities (A) that are distributed normally (B) In the population, the distribution of the character will be determined by the frequencies of the different alleles.

environmental factors, epigenetic modifications, or the presence of unknown gene interactions. Whatever the cause, individuals with a particular genotype sometimes do not show the typical phenotype associated with that genotype; in this case this trait displays incomplete **penetrance**. By contrast, **expressivity** reflects the degree in which a trait is expressed among individuals with identical genotype.

Finally, in some cases the expression of a specific phenotype is strictly limited to one sex although the gene that determines the character is located in an autosomal chromosome. Such traits show **sex-limited inheritance**, and they can be explained, for example, by hormonal factors. In traits with **sex-influenced inheritance**, the heterozygous genotype exhibits one phenotype predominantly in one sex and the opposite phenotype in the other. Although the trait is much more prevalent in one sex, it is not limited to that sex.

HUMAN GENOMIC VARIATION

The Landscape of the Human Genome

The completion of the **Human Genome Project**, with the publication of a draft sequence in the year 2000, afforded for the first time in history an overview of the structure and function of our genome. Various research projects over the last decade have expanded our knowledge about many important aspects of genome biology. The **HapMap Project**, followed by the **1000 Genomes Project**, created a comprehensive catalog of human genetic variation. A few years ago, the **ENCODE Project** identified a huge number of functional genomic elements not only in humans but also in some model organisms. More recently, the **Roadmap Epigenomics Project** has led to a much more detailed understanding of specific genomic features that are directly connected to gene function. As a result of all these endeavors, we have reached an unprecedented degree of knowledge about genome structure and function.

One of the first striking findings about the human genome concerned the **number of genes** and the related issue of **what is a gene**. Initial estimates of the number of human genes were focused on **protein-coding genes**, but the unexpected finding of a surprisingly high number of **nonprotein-coding genes** led to a reformulation of the definition of gene, which should now include not only protein-coding genes but also genes coding for RNAs that are not translated into proteins. As of the time of writing (November 2017), the latest Gencode catalog of annotated genes (version 28) includes 19,901 protein-coding genes, 23,348 noncoding RNA genes (long and small), and 14,723 pseudogenes. All these result in more than 200,000 distinct transcripts,

suggesting that human genes undergo a substantial amount of **alternative splicing**.

Within the class of nonprotein-coding genes, **long noncoding RNAs** (lncRNAs) and **micro RNA** (miRNAs) are at present the subject of a considerable amount of attention. These genes are almost as abundant as protein-coding genes, but they are more restricted in their tissue-specific expression and tend to result in fewer alternative transcript isoforms. Although they still constitute a poorly annotated set of genes, noncoding RNAs are now being implicated in many different aspects of genome regulation and cell biology and so they are thought to play important roles in the development of human diseases.

Another burgeoning area of research in genome biology is related to the identification and characterization of **regulatory elements**, which are preferentially located in intronic or intergenic regions of the genome. These elements are short regions to which transcription factors are bound, thus acting as enhancers or silencers of transcription with a crucial role in restricting gene expression to specific tissues or to certain stages of development or cell differentiation. Years of research have identified some genomic features defining these regulatory elements, such as open chromatin (DNAse I hypersensitivity) and specific **histone modifications**. For instance, the ENCODE project found several million regions in the human genome that meet the basic requirements for a regulatory element: DNAse I hypersensitivity and acetylation of lysine 27 of histone 3 (H3K27ac). Other projects such as Fantom5 or the Roadmap Epigenomics Project have subsequently refined this catalog and predicted a function for these regions. Most of them are enhancer/silencer elements, characterized (in addition to DNAse I hypersensitivity and H3K27ac) by the binding of specific transcription factors and by increased mono-methylation of lysine 4 of histone 3 (H3K4me). Also, about 115,000 genomic regions show preferential trimethylation of lysine 4 of histone 3 (H3K4me3), the typical signature for an active promoter. In total, this “**regulome**” comprises about one-fifth of the total sequence of the haploid reference human genome, and it is believed to play a critical role in susceptibility to disease (see below).

Initial analyses of the sequence of the human genome showed that almost 50% of the sequence of the reference haploid genome is made up of various types of **repeats**, including mobile elements. Dispersed repeats, in particular, are extremely abundant due to the fact that they are **mobile elements**. They are organized in several classes and families: short interspersed nuclear elements (SINEs) include nonautonomous retrotransposons such as *Alu* elements, with more than a million copies in a typical genome. Long interspersed nuclear elements (LINEs) are a large family of autonomous retrotransposons,

a few of which remain active and generate one new germline insertion every 200 births in human populations. Endogenous retroviral elements represent the remnants of the integration of retroviral genomes throughout primate history, so they are characterized by the presence of retroviral sequences, such as long terminal repeats (LTRs).

Interindividual Genetic Variation

One of the major achievements of genomic research since the completion of the Human Genome Project has been the creation of a deep catalog of **interindividual genetic variation**. The genome of each person is unique and contains millions of genetic variants when compared to the reference sequence or to the genomes of other individuals. Although many of these variants are supposed to be functionally neutral, a substantial proportion of them will be implicated not only in morphological or physiological features phenotypic features but also in traits such as susceptibility to disease, response to therapeutic interventions, or behavioral traits.

There are basically two types of genomic variation in human populations: polymorphic sites at single nucleotides (**single nucleotide polymorphisms** or SNPs) and polymorphisms that affect relatively large segments of the genome, generally referred to as **structural variants** or SVs. A type of these SVs are the **copy number variations (CNVs)** or variants in the number of repeats of regions from Kb to Mb in the genome. The number of these repeats varies between individuals in a population. After several years of work, the 1000 Genomes Project discovered and characterized **88 million variant sites** in the genomes of humans from different populations throughout the planet. Populations from African ancestry contain the highest amount of genetic variability, in agreement with the out-of-Africa model of human dispersal. This data set, the most comprehensive to date, reveals that an individual typically differs from the reference human genome at 4–5 million sites. The vast majority of these variants are SNPs and short indels (insertion/deletion polymorphisms), although SVs affect more nucleotides due to their larger size. For instance, an integrated map in unrelated individuals from 26 different populations found 68,818 SVs, with a typical genome containing about 2000 of these including large deletions, copy number variants, and insertions of mobile elements. Overall, these structural variants affect around 20 Mb of sequence, compared to 4–5 Mb for SNPs.

A very important feature of any genomic variant is its **population frequency**, because this has implications for its clinical relevance. The 1000 Genomes Project describes variants as “**common**” if their frequency is

>5% or “**rare**” if their frequency is <0.5%. Variants in the middle range (0.5%–5%) are considered as “**low**” frequency. Due to the deep sampling of many genomes from different populations, the majority of variants in the latest catalog (about 64 million) are rare, whereas only 8 million are common. However, the majority of variants found in a typical genome (compared to the reference sequence) are common, with no more than 200,000 rare variants per genome. This bias toward common variants has profound implications in the design and interpretation of **genome-wide association studies (GWASs)**, as will be explained in other chapters of this book.

Functional Significance of Genetic Variation

Not all variants present in a genome will necessarily have a functional impact. Many of them will probably be **neutral**, particularly those in intergenic regions or in nonfunctional segments of genes such as introns. The catalog created by the 1000 Genomes Project found that a typical genome contains about 2000 variants that had been previously associated with a complex trait through a GWAS. Additionally, a genome harbors 25 to 30 variants known to be involved in rare disease.

In the search for variants that could be implicated in human disease, **coding variants** initially attracted most of the attention as they can affect gene function in various ways: change the amino acid sequence of the protein encoded by the gene, introduce premature stop codons and thus lead to truncated proteins, change splicing patterns, etc. A typical genome will include about 150–180 protein truncating variants and 10,000–12,000 sites with a variant that changes the protein sequence. However, most of the variants implicated in the susceptibility to complex disease that were found in GWAS appear to be **noncoding variants**, so the focus is now changing to explore the functional relevance of variants located in noncoding regions of the genome.

The ENCODE Project had already estimated that 54% of the variants identified in GWAS studies were located in noncoding genomic regions that could be annotated with some role in gene regulation. The latest catalog observed around half a million variant sites that overlap known regulatory regions, such as untranslated regions (UTRs), promoters, enhancers, or transcription factor binding sites. Additionally, some variants affect the splicing patterns of the corresponding transcripts, leading to potential changes in protein levels or function. Therefore, noncoding genetic variation could have huge medical implications, and its functional relevance is now being explored in much greater detail.

In terms of susceptibility to complex disease, many genetic variants behave as QTLs. When the trait under

consideration is gene expression, these variants are known as expression QTLs or **eQTLs**. As they can affect gene expression positively or negatively, eQTLs are very attractive candidates to account for the phenotypic effects of noncoding genetic variants. It is important to keep in mind that eQTLs include not only SNPs but other types of variants such as short indels. Around 15% of eQTLs affect the expression of nonprotein-coding genes (lncRNAs), something that will undoubtedly change our view of the mechanisms leading to common disease and have profound implications for personalized genomic medicine.

Other genetic variants at SNPs are associated with DNA methylation levels at CpG sites located in the vicinity. Since they behave as QTLs, these sites are known as methylation QTLs or **meQTLs**. It is believed that these variants are located at binding sites for mediators of CpG methylation, and they could also have a huge impact on individual susceptibility to complex disease.

Suggested Reading

- Albert, F.W., Kruglyak, L., 2015. The role of regulatory variation in complex traits and disease. *Nat Rev Genet* 16, 197–212.
- Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., Abecasis, G.R., 2015. A global reference for human genetic variation. *Nature* 526, 68–74.
- GTEx Consortium, 2015. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 348, 648–660.
- Hill, W.G., 2012. Quantitative genetics in the genomics era. *Curr Genom* 13, 196–206.
- Mackay, T.F., Stone, E.A., Ayroles, J.F., 2009. The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet* 10, 565–577.
- Slatkin, M., 2008. Linkage disequilibrium – understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* 9, 477–485.
- Visscher, P.M., Hill, W.G., Wray, N.R., 2008. Heritability in the genomics era – concepts and misconceptions. *Nat Rev Genet* 9, 255–266.
- Wei, W.H., Hemani, G., Haley, C.S., 2014. Detecting epistasis in human complex traits. *Nat Rev Genet* 15, 722–733.

This page intentionally left blank

Molecular Biology of Genetic Variants

Martin Kohlmeier

Human Research Core and Nutrigenetics Laboratory, UNC Nutrition Research Institute; and Schools of Medicine and Public Health, Department of Nutrition, Kannapolis, NC, United States

Glossary

Allele Alternative form of a genetic variant; originally also of a phenotypic trait.

Epistatic Describing the impact of a gene, often at a distance, on the effect of another gene.

Genotype Set of alleles present in a particular individual; for biallelic variants there are three possible genotypes.

Micro RNA, miRNA Short RNA segment of regulatory significance that is usually the transcription product in a noncoding region.

Monoallelic expression Indicates that only one of alternative alleles is expressed by a cell, contrary to the usual expectation of joint expression.

Nutritope Nutrition environment to which local populations are exposed to and which over time can shape their genome.

Phenotype Set of observable traits that is commonly observed together, such as obesity, elevated blood pressure, and insulin resistance.

Trait Observable characteristic, such as height, blood pressure, or a metabolite concentration.

Transcription initiation site Position where the transcription of DNA into mRNA starts.

chromosome to the other, has to be unwound before it can be replicated for cell division.

While most heritable information resides in the chromosomal DNA of the cell nucleus, each mitochondrion contains its own DNA. The mitochondrial DNA is a loop of double-stranded double helix with 16,569 base pairs encoding only 37 genes. This means that each cell contains as many mitochondrial DNA loops as it has mitochondria.

Numerous DNA sequence patterns determine the functioning of the human genome, coordinating its work in particular cells, tissues, and organs, aligning with hormonal, metabolic, and other biological inputs. The slightly more than 21,000 protein-coding genes, which ultimately produce specific peptides, are the most widely known functional units, typically consisting of both expressed (exons) and nonexpressed (introns) sequences; about the same number of expressed sequences are noncoding (Perteau *et al.*, 2019). They may be extremely large such as FTO with 456,821 bases (originally named fatso because of its large size, and only later discovered to be involved in the regulation of thermogenesis and control of body weight), encoding alpha-ketoglutarate-dependent dioxygenase (EC 1.14.11.-). Or it could be as small as the HIST1H3A gene with 469 bases, which has no introns. This tiny gene encodes one of the histones that constitute nucleosomes for compacting the DNA strand into a chromosome.

Binding of a complex assembly of factors to promoter regions, typically upstream of the transcription start site, control the expression of individual genes. A broad array of further factors, such as hormones, proteins, peptides, lipids, nutrients, and various types of RNAs with signaling function, bind to additional gene-related sequences. An important example are the estrogen-binding sequences that also respond to many plant-

HUMAN GENOME STRUCTURE

Most heritable traits are transmitted from parents to offspring by the more than 3.3 billion individual nucleotide bases strung together in deoxynucleic acid (DNA) polymers in the cell nucleus organized in 23 pairs of chromosomes. Each chromosome contains a single double helix strand with many millions of individual bases. These extremely long double-helix strands with a linear chain of about 50–300 million bases are tightly wound into superhelical structures that have to be unwound partially before the targeted sequence transcribed into ribonucleic acid (RNA) copies as the messenger RNA (mRNA) templates for protein synthesis (Fig. 2.1) or as various types of short RNA with functions in their own right. The entire strand, from one end of a

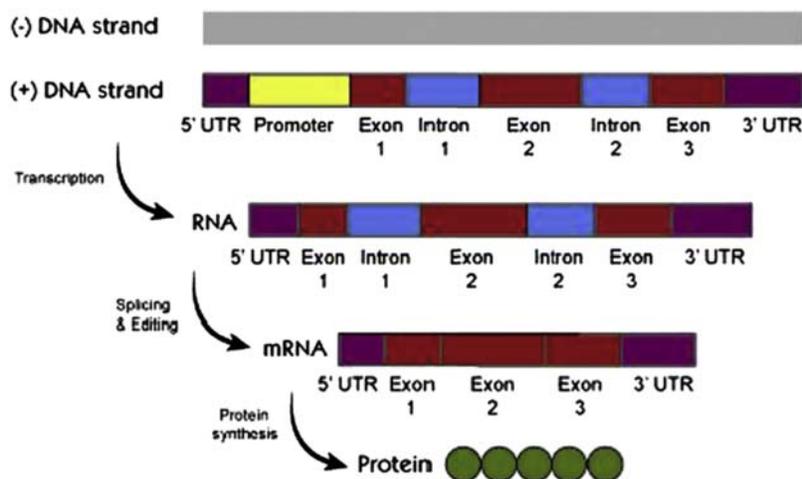


FIGURE 2.1 Genes provide the DNA sequences that are transcribed into RNA, which is then spliced and edited to messenger RNA (mRNA), the template for the synthesis of peptides and proteins. Some short DNA sequences, usually in noncoding regions, are only transcribed to generate microRNAs and other functionally important types of RNA. From Kohlmeier, M., 2013. *NutriGenetics: Applying the Science of Personal Nutrition*. Academic Press, London, UK, pp. 28.

derived compounds such as flavonoids. It is important to understand that in many cases the binding sites can be quite distant from the gene the control. One enhancer region of importance for limiting the expression of lactase to enterocytes in the proximal intestine of infants is located about 14,000 bases upstream of the initiation site, right in the middle of an unrelated neighboring gene.

An all-important consideration relates to the fact that most genes give rise to multiple transcripts and protein versions, depending on where transcription starts, which exons are transcribed, how the RNAs produced from the various transcribed exons are spliced, and how they are edited. For example, the enzyme encoded by APOBEC-1 converts with narrow specificity at position 6666 of mRNA transcribed from APOB the native C into U, which generates a stop codon. This shortens the length of the resulting ApoB protein to 48% of the original length (therefore called ApoB48). This happens only in the proximal intestine of humans and generates the form of apoB that is best suited for chylomicrons. The APOB mRNA from liver for VLDL synthesis is not edited in this way in humans (unlike in rodents). Thus, the APOB-3500 variant (rs144467873) will only have an effect on LDL and other VLDL-derived lipoproteins, but not on chylomicron-derived lipoproteins because the translation of the latter stops before the variant position is reached. The expression patterns and protein products of all kinds of genes with common variants have to be considered very carefully when attempting to link effects to a specific locus (position) or sequence.

In addition to the DNA regions defined by genes and their associated regulatory segments, a growing number

of transcribed sequences are recognized that are not part of the traditional protein-encoding genes. These include, for example, short segments that generate microRNA (miRNA), molecular factors that contribute to the regulation of gene expression.

Methylation of specific cytosine bases in the DNA usually hinders expression of nearby sequences or genes. Such epigenetic modifications (superimposed on the DNA sequence) are very common and constitute another critical factor for the regulation of genome functions as described in later chapters of this book.

FREQUENCY AND RELEVANCE OF GENOMIC VARIATION

There is great genetic diversity in all human populations, and much more so across populations in different regions of the world. It has recently been reported that the genome of a typical individual differs from the reference genome at more than 4 million positions ([1000 Genomes, 2015](#)). At the vast majority of these positions (more than 99.9%) the difference consists of a single base or of a few bases that are missing or added (insertions or deletions). This typical individual also will have large deletions at about 1000 positions, a few hundred insertions of extraneous sequences (Alu, L1, and SVA), about 160 copy number variants, and a handful of inversions and nuclear mitochondrial DNA segments ([Chiang et al., 2017](#)). About 8 million of variants occur in at least 5% of people.

The version someone has at a position with a variant is called an allele. For most single-base differences, two alleles are possible (biallelic variant), less often three

(triallelic variant). Variants extending over several or even many bases can have many different alleles.

The combination of the (usually) two alleles is called the genotype. If the two copies of the variant are the same, the genotype is homozygous; with one copy each of two alleles, the genotype is heterozygous. Depending on the nature of the variant, its effect may be seen only when two copies of the variant allele are present, in which case it is called recessive. This usually happens when expression of just one copy is enough to sustain the function of this gene or DNA segment. However, the sufficiency depends on the specific circumstances, and very often on nutritional exposure. A well-understood example is the effect of variant MTHFR 677 T (rs1801133) on homocysteine concentration in blood. The 677 T allele encodes a 5-methyltetrahydrofolate reductase allozyme with only about half the activity of the version encoded by the 677 C allele, assuming that the sequence is otherwise identical. This lower activity matters only if folate availability in the cell expressing the allozyme is insufficient. It is easy to demonstrate both in vitro and in people that providing plentiful folate diminishes this metabolic consequence and homocysteine can be equalized.

Typically, if a DNA segment is expressed (transcribed into RNA) in a particular cell and type of tissue, all of the alleles present in an individual's genome are expressed. However, when the response to nutrition factors is different between genotypes, only one allele is expressed more often than in variants unresponsive to nutrition (Carpenter et al., 2018). This is called monoallelic expression and is typically due to selectively higher methylation of the nonexpressed allele.

TYPES OF GENOMIC VARIATION

Single Nucleotide Polymorphism

The most common type of heritable DNA variant changes a single base pair anywhere in the genome. If the population frequency of any variant is high (typically when it exceeds a nominal threshold of 1% in a particular population), it is called a polymorphism. Hence, variants characterized by having a different nucleotide in a particular position than is found in most other people is called a single nucleotide polymorphism (SNP). It should be obvious that the acronym SNP (often casually pronounced "snip") does not properly apply to other types of variants described below. SNPs are usually biallelic, which means that there are two different versions (alleles) at the variant position. More rarely, the variant position can be occupied by one of three alternative bases, which would make it a triallelic SNP.

Diverse labels have been used in the past to name variants, often using a combination of gene name, location relative to the transcription initiation site and the altered bases. Use of location estimates and other ambiguities can lead to considerable confusion and misunderstandings, particularly when reading earlier reports. An example is the naming of a variant in the alcohol dehydrogenase 1B gene (ADH1B), typically associated with the rapid development of flush in the face and headache after alcohol consumption. Since it is characterized by the replacement of a guanine with an adenosine base in position 143 when the mRNA is reverse transcribed into complementary DNA (cDNA), it is often named c.143G > A (the c. indicates the position in the cDNA). Alternatively, some use the designation Arg47His or Arg47→His to indicate that the variant results in an allozyme (enzyme with an altered sequence) with an arginine instead of histidine in position 47 of the resulting protein sequence. However, genes may have different transcription initiation sites and more than one gene product. As it happens, the transcription initiation site for the ADH1B gene has been reconsidered and the position of the amino acid substitution is now given as 48, hence the revised name for the same variant is now Arg48His, or 48His for short. This leads to obvious problems when searching older, or even current reports, and some of the published work may be missed. It is much preferable to have a uniform nomenclature that can then be cross-referenced with all the alternative names. It is now widely accepted practice to use unique rs accession numbers where the rs stands for reference SNP cluster ID (Sayers et al., 2011). These rs numbers are listed in the Single Nucleotide Polymorphism Database (dbSNP). The above-mentioned variant ADH1B should therefore be described as rs1229984. Yet another nomenclature uses asterisks and a number following the gene name. This convention is typically used for reporting of clinical pharmacogenetic results (Caudle et al., 2018). Thus, the ADH1B*2 designation in clinical pharmacogenetics reports is equivalent to rs1229984.

When a new mutation changes the ancestral base at a particular position, there are different possible outcomes. If the position is in a protein-coding region, the change may be nonsynonymous, resulting in a stop codon that terminates translation right there. Truncated proteins tend to be removed or at least they have little or no functional activity. Other altered three-base codons may result in encoding a different amino acid. Most random mutations, and that applies particularly to those causing altered amino acid sequences, result in a decrease or even complete loss of function. However, polymorphisms are not mutations and should never be referred to as such, since it takes selective advantage to enrich a mutation enough over many generations to meet the definition of polymorphism (at least 1%

frequency in a population). What makes this even more confusing is that the ancestor with the original mutation, like all of us, has carried numerous variants near the altered position (locus) and the selective advantage for that new variant then tends to drag along those other variants for many generations. Comparing the alleles at a locus of interest in different primate and nonprimate species often provides hints about the ancestral allele (i.e., often the allele that is present in our primate cousins). Such comparisons suggest some trends for nutritionally relevant polymorphisms. First, not all variants that have newly arisen in humans are associated with a loss of function. A case in point is the ADH1B rs1229984 variant discussed above, where the nonancestral x allele encodes for an alcohol dehydrogenase allozyme that is about 60 times more active than the ancestral form. Second, many, if not a majority, of the nutritionally important variants are located in noncoding regions and affect regulatory sequences, such as the promoter region, enhancer regions, and sequences that bind signaling compounds or metabolites. An important nutrigenetic example are variants associated with persistence of lactase intestinal expression that about a third of the global human population carries and enables them to drink large amounts of milk without undue consequences. There are several of these common variants, including rs4988235 (LCT -13910 C > T, in many people with Central or North-European ancestry), rs41380347 (LCT -13915 T > G, particularly in people with Arab ancestry), and rs41525747 (LCT -13907 > G, in people with Ethiopian and other African ancestry). All of them affect an enhancer region that controls the tissue- and age-specific expression of the lactase gene.

Nutritionally important SNPs may also be located in the 3' untranslated region (3' UTR) because this is where the cyclization of mRNA is controlled. Circular mRNA is the form used as template for protein synthesis in ribosomes, going round over and over again as long as the loop remains intact. Specific sequences in the 3' UTR bind miRNAs and other factors that influence the longevity of the circular mRNA in the ribosome and thus the number of protein copies an individual mRNA molecule can produce. For example, the variant DHFR rs7387829 (829 C > T) is located 14 bp downstream of a miRNA-24 binding site and influences dehydrofolate reductase abundance dependent on the genotype.

The change of a single base will not alter the amino acid in the encoded protein sequence if it generates a triplet codon that codes for the same amino acid as the original codon. Such synonymous SNPs are common and usually have no further consequences. However, functional differences are still possible, for instance, due to interfering with the normal splicing of one exon

to the next. The change in DNA sequence may generate a new donor or acceptor site that then results in an abnormal mRNA and ultimately dysfunctional protein product. A modification of the three-dimensional configuration resulting from the altered DNA or mRNA sequence can be enough to cause exon skipping or splicing failure (Livingstone et al., 2017).

Deletions, Insertions, and Repeats

Many people have sequences in their genome that are anywhere from just one base to millions of bases shorter than the corresponding sequence in other people. Similarly, many have sequences that are one or many more bases longer than in other people. The terminology must not be taken to mean that these deletions and insertion occurred recently because they go back many generations, in some cases even millions of years. If the variants are common, we must assume that some selective advantage has enriched them, often due to alignment with the prevailing nutritope, the nutrition environment that the ancestral group was exposed to.

The addition or removal of one or two bases to a coding DNA sequence will lead to a frameshift. Most triplet codons downstream of the altered site will be changed to nonsense and usually after a few of these randomly generated codons a stop codon will occur. Short tandem repeats (STRs, also called microsatellites or simple sequence repeats, SSRs) tend to consist of multiple replications of a motif with just a few base pairs, usually just four to eight and rarely more than 50. Such repeats become functionally relevant mainly when the repeated motifs interact with regulators. For example, a polymorphic region of the gene ALOX5 contains a variable number of the six-base sequence GGGCGG, which is an Sp1-binding pattern. The variant with four repeats of this pattern appears to make carriers more vulnerable to the atherogenic effects of high arachidonic acid consumption than the variant with five repeats (Dwyer et al., 2004).

In some individuals, microdeletions (microscopically undetectable but still thousands of base pairs long) are responsible nutritionally treatable phenotypes, but those tend to be rare cases with grave clinical impact.

COMPLEX PATTERNS OF GENOMIC VARIABILITY

Haplotypes

There are hundreds of common variants in any average-sized gene, some of which may alter functional characteristics whereas many other do not. Determining the functional relevance of all these variants is difficult

enough, but the characteristics often depend on the presence of additional nearby variants. This should not be surprising because the function of proteins is usually strongly influenced by their three-dimensional configuration. Thus, what may seem a great distance between amino acid moieties in a linear configuration may actually be directly adjoining due to folding in the active configuration. It is important to consider that for such combinatorial effects, the variants have to lie on the same strand, what is called in cis position. Only then will they be translated into a protein strand that contains the relevant combination of amino acid moieties.

Certain variants of the MTHFR gene may illustrate the reality of such interactions of distant variants located on the same DNA strand. As can be seen in Fig. 2.2, the allozyme *6 with alanine in position 429 (429 Ala) has about the same activity as the most common *1 form when expressed in a cell-based model system. Likewise, the allozyme *7 with histidine in position 519 (519 His) has at least 100% of the normal activity, if not more. If these two variants are combined in the same strand and then expressed as allozyme *16 (429 Ala/519 His), the resulting activity is greatly diminished. This means that the activity of variant combinations cannot be reliably predicted based on observation with one variant alone. To the extent possible, variant effects should be related to complex haplotypes instead of individual variants. This is possible because in many nutritionally relevant cases only a few extended haplotypes, maybe just four to six, are present in the majority of members of a particular population.

Hemizyosity

Men have only one copy of the X chromosome and any detrimental alteration cannot be compensated by

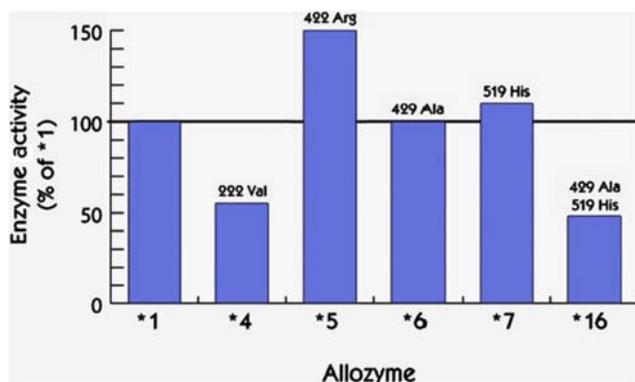


FIGURE 2.2 Both MTHFR allozyme *6 (429 Ala) or *7 (519 His) from recombinant constructs retain normal or even slightly increased activity. The allozyme *16 containing both 429 Ala and 519 His has greatly reduced activity. From Kohlmeier, M., *Nutrigenetics: Applying the Science of Personal Nutrition*. Academic Press, London, UK, 2013, pp. 45.

the second copy as it would for the numbered (autosomal) chromosomes. The worst-case scenario is that a longer sequence is lost and therefore does not fulfill its function. The same situation occurs when a variant diminishes the function of an X-chromosomal gene. This abnormality is called hemi-insufficiency. An important example is the X-chromosomal gene OCT, which encodes the enzyme ornithine transcarbamoylase (EC 2.1.3.3). There are numerous deleterious variants that diminish or abolish the function of the enzyme and thereby disrupt the capacity for safe elimination of ammonia as urea (Kohlmeier, 2013). Before birth the mother's metabolism is sufficient to remove excess ammonia from the circulation of the fetus, which is coupled with hers. After birth, males with low activity OCT variants usually do not survive. The approximately one in 15,000 female carriers of heterozygous defect variants usually grow up healthy without obvious pathology. However, some of those carriers are affected worse than others due to unequal inactivation of the second chromosome copy in their livers. They may have significantly more cells with the defective variant than with the fully functional variant and be even more vulnerable. Unaccustomed high protein intake, for instance at a celebration feast, or a protein catabolic state due to trauma, childbirth, or systemic infection, can lead to acute decompensation and rapidly fatal increase in ammonia concentration.

Copy Number Variants

Gene duplications are common, and in our evolutionary history have generated multiple versions of genes with similar, but very distinct, characteristics. The large number of odorant receptors (ORs) readily illustrates this principle. Humans have close to 900 distinct OR genes. Most, if not all, of them are variations of ancestral versions adapted to detect specific volatile molecules with exquisite sensitivity. In each individual fewer than half are functional, largely due to copy number variants (Veerappa et al., 2013). The number of copies for those affected ORs typically ranges from 0 to 2, which means that a particular OR gene is present in some but not in many others.

Conventionally, copy number variants are defined as multiplications of segments that encompass more than 1000 base pairs, not necessarily a full gene sequence. For example, novel version Hp2, of the HP gene (encoding haptoglobin), has apparently come about through the duplication of exons 3 and 4 of the ancestral Hp1 form combined with some further modifications of the original structure. The protein translated from the now most common variant Hp2 tends to aggregate into polymers with eight and more units of both Hp1 and Hp2.

Individuals homozygous for the Hp1 version only form dimers. The larger polymeric aggregates bind free hemoglobin less effectively than the dimers and are associated with higher vitamin C requirement due to inferior hemoglobin-binding capacity (see more details in Chapter 29, Vitamin C).

A much larger range of copy numbers also occurs. A nutrition-related example is the *AMY1* gene, which encodes salivary amylase (see in [Kohlmeier, 2013](#)). Different individuals may carry variants with anywhere between none and more than 10 copies arrayed in sequence. The variants with more copies appear to have evolved in populations with high starch consumption, while herders traditionally getting little starch tend to have distinctly fewer copies. The health consequences, particularly regarding risk of greater adiposity and type 2 diabetes, are not fully understood.

References

- Carpenter, B.L., Zhou, W., Madaj, Z., DeWitt, A.K., Ross, J.P., Grønbaek, K., Liang, G., Clark, S.J., Molloy, P.L., Jones, P.A., 2018. Mother-child transmission of epigenetic information by tunable polymorphic imprinting. *Proc Natl Acad Sci USA* 115 (51), E11970–E11977. PMID:30509985.
- Caudle, K.E., Keeling, N.J., Klein, T.E., Whirl-Carrillo, M., Pratt, V.M., Hoffman, J.M., 2018. Standardization can accelerate the adoption of pharmacogenomics: current status and the path forward. *Pharmacogenomics* 19 (10), 847–860. PMID:29914287.
- Chiang, C., Scott, A.J., Davis, J.R., Tsang, E.K., Li, X., Kim, Y., Hadzic, T., Damani, F.N., Ganel, L., GTEx Consortium, Montgomery, S.B., Battle, A., Conrad, D.F., Hall, I.M., 2017. The impact of structural variation on human gene expression. *Nat Genet* 49 (5), 692–699. PMID:28369037.
- Dwyer, J.H., Allayee, H., Dwyer, K.M., Fan, J., Wu, H., Mar, R., Lusis, A.J., Mehrabian, M., 2004. Arachidonate 5-lipoxygenase promoter genotype, dietary arachidonic acid, and atherosclerosis. *N Engl J Med* 350 (1), 29–37. PMID:14702425.
- Kohlmeier, M., 2013. *Nutrigenetics: Applying the Science of Personal Nutrition*. Academic Press, London, UK.
- Livingstone, M., Folkman, L., Yang, Y., Zhang, P., Mort, M., Cooper, D.N., Liu, Y., Stantic, B., Zhou, Y., 2017. Investigating DNA-, RNA-, and protein-based features as a means to discriminate pathogenic synonymous variants. *Hum Mutat* 38 (10), 1336–1347. PMID:28649752.
- Perlea, M., Grosu, O.M., Terinte, C., Porocho, V., Velenciuc, N., Lunca, S., 2019. Chess: a new human gene catalog curated from thousands of large-scale RNA sequencing experiments reveals extensive transcriptional noise. *Medicine (Baltimore)* 98 (3), e14111. PMID:30653135.
- Sayers, E.W., Barrett, T., Benson, D.A., Bolton, E., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., DiCuccio, M., Federhen, S., Feolo, M., Fingerman, I.M., Geer, L.Y., Helmberg, W., Kapustin, Y., Landsman, D., Lipman, D.J., Lu, Z., Madden, T.L., Madej, T., Maglott, D.R., Marchler-Bauer, A., Miller, V., Mizrachi, I., Ostell, J., Panchenko, A., Phan, L., Pruitt, K.D., Schuler, G.D., Sequeira, E., Sherry, S.T., Shumway, M., Sirotkin, K., Slotta, D., Souvorov, A., Starchenko, G., Tatusova, T.A., Wagner, L., Wang, Y., Wilbur, W.J., Yaschenko, E., Ye, J., 2011. Database resources of the National Center for Biotechnology information. *Nucleic Acids Res* 39 (Database issue), D38–51. PMID:21097890.
- The 1000 Genomes Consortium, 2015. A global reference for human genetic variation. *Nature* 526, 68–74. PMID:26432245.
- Veerappa, A.M., Vishweswaraiah, S., Lingaiah, K., Murthy, M., Manjogowda, D.S., Nayaka, R., Ramachandra, N.B., 2013. Unraveling the complexity of human olfactory receptor repertoire by copy number analysis across population using high resolution arrays. *PLoS One* 8 (7), e66843. PMID:23843967.

Regulation of Gene Expression

Maria Luisa Bonet^{1,2,3}, Andreu Palou^{1,2,3}

¹Nutrigenomics and Obesity (NUO) Research Group, Laboratory of Molecular Biology, Nutrition and Biotechnology, Universitat de les Illes Balears, Palma de Mallorca, Spain; ²CIBER de Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Palma de Mallorca, Spain; ³Institut d'Investigació Sanitària Illes Balears (IdISBa), Palma de Mallorca, Spain

List of Abbreviations

AMPK AMP-activated protein kinase
 CTD carboxy terminal domain
 eRNA enhancer RNA
 GTF general transcription factor
 HDAC histone deacetylase
 mTOR mechanistic target of rapamycin
 PIC preinitiation complex
 PPP promoter-proximal pausing
 PTM posttranslational modification
 PUFAs polyunsaturated fatty acids
 TF transcription factor
 TSS transcription start site

Glossary

5'-capping Addition of a 7-methylguanosine cap to the 5' end of the nascent pre-mRNA, while elongation is in progress. The 5' cap protects the nascent transcript from degradation from its 5' end and assists in ribosome binding during translation.

Carboxy terminal domain (CTD) Carboxy terminal domain of the RNA polymerase II largest subunit consisting in tandem repeats of a heptad sequence (consensus: Tyr1Ser2Pro3Thr4Ser5Pro6Ser7, 52 repeats in humans); the CTD phosphorylation status changes through the transcription cycle to allow its progression and the cotranscriptional processing of the nascent mRNA, through the recruitment of protein factors involved.

Chromatin A complex of DNA and proteins that forms chromosomes within the nucleus of eukaryotic cells.

Chromatin looping A chromatin loop occurs when stretches of genomic sequence that lie on the same chromosome are in closer physical proximity to each other than to intervening sequences.

Chromatin remodeling Dynamic modification of chromatin architecture to allow or suppress the access of regulatory machineries.

Cohesin A protein complex that holds sister chromatids together after DNA replication until anaphase, when removal of cohesin leads to separation of sister chromatids, and which also functions as a chromatin looping factor.

Core promoter Minimal stretch of contiguous DNA sequence that is sufficient to direct accurate initiation of transcription by the RNA polymerase II machinery; it encompasses the transcription start site(s).

Coregulators Proteins with which sequence-specific DNA-bound transcription factors interact to exert their effect on target gene transcription, but that do not bind themselves to the DNA in a sequence-specific manner. Among the coregulators, coactivators are required for the action of activator transcription factors and corepressors for the action of repressive transcription factors.

Enhancer A short (50–1500 bp) DNA stretch containing multiple activator and/or repressor binding sites, to which a complex of transcription factors binds to affect the transcription from related core promoters. A promoter can be controlled by multiple enhancers, and a given enhancer can control more than one promoter.

Enhancer-derived RNA (eRNA) Noncoding RNA transcripts produced from genomic regions that are presumable enhancer DNA elements; eRNA interact with looping factors (e.g., cohesin complex) and facilitate/stabilize chromosomal looping between the enhancer and promoter(s) of cognate target gene(s).

General (or basal) transcription factors (GTFs) Proteins that are not part of RNA polymerases but are needed for accurate initiation of transcription and are part of the transcription preinitiation complex.

Mediator A multisubunit assembly that appears to be required for regulating expression of most RNA polymerase II transcripts, acting both as a GTF and as a coactivator.

Nucleosome Any of the subunits that repeat in chromatin, consisting of a length of DNA (147 bp) coiled around an octamer of histone proteins consisting of an H3-H4 tetramer flanked by two H2A-H2B dimers.

Poly A tail A stretch of adenine nucleotides that is sequentially added to the mRNA molecule following 3'-specific cleavage, as part of pre-mRNA processing; the poly A tail increases the stability of the mRNA molecule and facilitates translation. 3'-cleavage and polyadenylation of the mRNA transcript is cotranscriptional and is required for termination of transcription by RNA Pol II.

Preinitiation complex (PIC) An entry form of Pol II in a complex with general transcription factors in which the polymerase is bound to the core promoter DNA but has not yet initiated RNA synthesis.

Promoter-proximal pausing (PPP) Pausing of Pol II during early transcript elongation, 20–60 nucleotides downstream of the transcription start site; release of the PPP constitutes a widespread regulatory mechanism in gene transcription in higher eukaryotes.

Proteasome Main nonlysosomal intracellular protease complex, which digests proteins previously marked for degradation and is itself made up of proteins.

Sequence-specific transcription factors Transcription factors that bind to regulatory DNA sequences, such as sequences contained in enhancers, to stimulate or repress transcription of related genes.

Transcription start site Location where transcription starts at the 5'-end of a gene sequence.

mechanisms through which dietary factors impact gene transcription illustrated.

INTRODUCTION

Gene expression is the process through which information encoded in DNA is converted into functional protein or RNA occurring in the cell or the body. Expressed genes include genes for coding messenger RNA (mRNA), which are molecules that are translated into proteins, as well as genes for noncoding but functional RNA molecules such as transfer RNAs, ribosomal RNAs, small nuclear RNAs, small nucleolar RNAs, microRNAs, and others.

Gene expression is a highly regulated process. All cells in the human body arise from a single cell, the fertilized egg or zygote, and hence essentially have (or at least had originally) the same genome, yet the fraction of expressed information varies from cell type to cell type. Some cellular proteins are very abundant, while others are typically present in just a few copies per cell. There are genes that are expressed at a relatively constant level (constitutive genes) and genes whose level of expression can greatly vary in response to internal and/or external cues (inducible genes). Often, genes encoding different products required to respond to a given stimulus are synchronously activated so that the various gene products are produced simultaneously at the appropriate relative levels. Furthermore, mechanisms have evolved during biological evolution to expand the transcriptome and proteome from a limited genome.

Changes in gene expression play essential roles in the regulation of cell fate and differentiation, as well as in the adaptation to stress and metabolic control, and hence are key to execute developmental programs, to correct internal errors, and to maintain metabolic homeostasis in a changing environment. However, changes in gene expression can also be part of, and causally contribute to, pathological responses/processes and disease states.

Expression of protein-coding genes can be regulated at any stage of the biological information flux, from mRNA synthesis to the half-life of the protein product, and it is commonly coordinately regulated at multiple stages. Although important posttranscriptional regulatory mechanisms are known—such as alternative mRNA splicing, alternative mRNA polyadenylation, RNA editing, alternative translation initiation, and gene-silencing through microRNAs (miRNAs)—the emphasis (in this chapter) is on gene expression control at the level of transcription. The transcription process and its overall regulation are reviewed, and main

INTERPLAY BETWEEN CHROMATIN STRUCTURE AND TRANSCRIPTION FACTOR-DNA INTERACTIONS IN THE REGULATION OF GENE EXPRESSION

Transcription of protein-coding genes requires the formation of a *preinitiation complex* (PIC) on the *core promoter*. The latter is defined as the minimal DNA stretch that directs accurate initiation of transcription and encompasses the *transcription start site* (TSS, referred as the +1 position). The PIC includes RNA polymerase II (Pol II) and its *general transcription factors* (GTFs), which are proteins different from RNA Pol II needed for transcription initiation of all protein-coding genes.

The efficacy of PIC assembly can be low in a *chromatin* environment. Eukaryotic DNA is wrapped around nucleosomes and forms higher order chromatin structures that restrict the access of the general transcription machinery to core promoters. Due to this organization, the “ground state” of eukaryotic DNA expression is largely restrictive. To achieve optimal levels of transcription in a nucleosomal context, PIC multiprotein assembly often requires help from *sequence-specific DNA-binding transcription factors* (hereafter, TF). Some of these TFs can bind their cognate sequences in DNA within regular nucleosomes and initiate events leading to *chromatin remodeling* and gene activation; these are called initiator or pioneering TFs. Other TFs are unable to bind regular nucleosomes and depend on initiator TFs or on a preset nucleosomal structure for accessing their target sequences in chromatin. More rarely, gene regulation involves inhibition of PIC assembly by an active, sequence-specific transcriptional repressor. Gene repression can also result from repressive chromatin states or from the action of inhibitory proteins that counteract the activity of transcriptional activators.

TFs function after binding to specific, short regulatory DNA sequence elements (located outside the core promoter) and are required for the activated transcription of specific sets of target genes. DNA regulatory elements for TF binding are generally found within several hundred bases upstream of the TSS to which they are linked, although they can be located many thousands of base pairs away, either upstream or downstream of the TSS, or within the transcribed region of the gene (or of other genes). Often, regulatory DNA elements are grouped together in *enhancers*. An enhancer can control more than one gene, and a gene can be controlled by more than one enhancer (on average, each protein-coding gene in the human genome interacts with four to five enhancers, according to genomic studies). The actual level

of transcription of a given gene depends on the coordinated interactions of the various transacting TFs involved, whose activity and accessibility to the genome often depends on the status of cellular signaling pathways sensitive to external and/or internal stimulus.

How can TFs bound to a distal enhancer modulate PIC assembly at the core promoter? An important mechanism is through *chromatin looping*, which brings the corresponding DNA regulatory elements close together through the looping away of the intervening region, thus allowing productive contact between the bound proteins (e.g., a TF and a GTF). The principles and mechanisms governing chromatin looping have only recently begun to be understood. De novo formation of enhancer-promoter loops may depend on the association of TFs that enable loop formation, by themselves or by recruiting additional protein factors that help establish the loop, such as mediator or *cohesin*; such loops dissolve upon depletion of the associated TF. Chromatin looping may also depend on chromatin flexibility and the action of certain long noncoding RNAs, such as *enhancer-derived RNAs*. Other loops are preestablished and TF binding contributes to stabilize them.

TFs use a variety of nonmutually exclusive mechanisms to regulate the transcription of target genes. Many TFs function by establishing direct protein-protein interactions with GTFs and/or RNA pol II that facilitate (activators) or impede (repressors) PIC assembly, initiation of transcription or transcription reinitiation. Additionally, many TFs possess enzymatic activity that enables local chromatin remodeling at the level of the core promoter or other cis-regulatory DNA elements, making these elements more or less accessible to the corresponding transacting binding factors (GTFs or TFs), or recruit other protein complexes with such activity. TFs may also help chromatin looping or function postinitiation, for instance by facilitating the transition of paused RNA Pol II to productive elongation (see next section).

Most TFs exert their action through the recruitment of transcriptional *coregulators*, which are proteins required for the TF effect on target gene transcription but that do not directly bind themselves to a specific DNA sequence. Coregulators reach the gene promoter through interaction with bound cognate TFs, and contain in their molecule additional domain(s) for interaction with GTFs and/or local chromatin remodeling. Activator TFs work together with coactivators, while repressor TFs work together with corepressors. Regulation of transcription by a given TF may involve the sequential action of different coregulators. This dynamic, as well as the turnover of refreshed TF molecules on its cognate cis DNA element, often involves the action of the nuclear *proteasome*, which is intimately connected to transcription through associated proteolytic and nonproteolytic activities.

TRANSCRIPTION OF PROTEIN-CODING GENES: GENERAL ASPECTS

Pol II Core Promoters

Two main types of core promoters are recognized: focused and dispersed. Focused promoters contain a single predominant TSS, are typical of regulated genes, extend from approximately nucleotides -40 to $+40$ relative to the TSS, and contain combinations of certain canonical nucleotide sequence motifs that have a positional bias with respect to the TSS (though none of them is universal). Dispersed promoters contain multiple weak TSSs that spread over 50 to 100 nucleotides, are typical of constitutively expressed genes, lack canonical core promoter sequence elements, and are often embedded in CpG islands or ATG deserts (DNA stretches rich in CpG dinucleotides or lacking ATG trinucleotides, respectively). Transcription initiation has been most studied from focused, regulated promoters, even though these represent only about 30% of total Pol II promoters in metazoan genomes.

Besides specific DNA sequence motifs/features, active core promoters display specific chromatin signatures, namely a nucleosome-depleted region and the enrichment of specific histone modifications in the *nucleosomes* located upstream and downstream of the TSS, which are strongly positioned (fixed) nucleosomes. All these features facilitate PIC formation. Nucleosome positioning and depletion at promoters is a dynamic process influenced by DNA sequences (sequences inherently stiff such as CpG- and AT-rich regions favor nucleosome depletion and are overrepresented at promoters), histone variants (combinations such as H3.3 plus H2A.Z make nucleosomes prone to be displaced), DNA-binding transcription factors (for which binding sites are often found within the nucleosome-depleted DNA segment), and remodeling by chromatin modifiers.

PIC Assembly and Transcription Initiation

PIC assembly can be viewed as a multistep recruitment process of the GTFs and Pol II that can be facilitated (or, less often, inhibited) by additional regulatory proteins. Collectively, GTFs of Pol II (TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, *Mediator*) function in core promoter recognition, Pol II recruitment, interaction with regulatory factors, DNA unwinding, TSS recognition, and promoter clearance. After PIC assembly is completed, TFIIH catalyzes the ATP-dependent unwinding of approximately 10 base pairs of duplex promoter DNA. Pol II then locates the TSS on the template DNA strand and initiates RNA synthesis. This is followed by promoter clearance, when many of the contacts of Pol II with GTFs are lost, so that the

enzyme can move downstream the DNA to continue RNA transcript elongation. Promoter clearance associates with Pol II CTD phosphorylation (in Ser5) catalyzed by a kinase activity contributed by TFIIF. After promoter clearance, some (though not all) GTFs may remain at the core promoter, constituting a scaffold for transcription reinitiation. This process is thought to happen during activated transcription, which may also involve the formation of an intragenic loop between the core promoter and the 3'-ending (terminator) region of the gene.

Subsequent Steps in Transcription

Pol II often pauses after synthesis of nascent RNA 20–60 nucleotides in length and it remains in a paused state on the DNA—without releasing the nascent transcript—until additional signals promote the transition to productive elongation. Escape of Pol II from this *promoter-proximal pausing* (PPP) is, together with the recruitment of the basal transcription machinery, an important rate-limiting, regulated step in transcription. PPP affects >70% of metazoan genes according to genomic approaches, being particularly common at genes in stimulus-responsive pathways. Regulated escape from PPP facilitates a fast response and synchronous gene activation in neighboring cells in a given tissue. The PPP has been also related to the 5' *capping* of the nascent pre-mRNA and other functions. The PPP is promoted by protein factors such as negative elongation factor (NELF) and DRB sensitivity-inducing factor (DSIF), though mechanisms involved are not well known, and resolved by the protein kinase positive transcription elongation factor b (P-TEFb), through the phosphorylation of NELF, DSIF, and the Pol II CTD in Ser2. P-TEFb is recruited to paused Pol II complexes through interactions with DNA-binding transcription activators, transcriptional coactivators (such as *Mediator*) or chromatin factors (e.g., acetylated histone binding proteins).

Pol II does not dissociate from template DNA until transcription termination, which is closely linked to cotranscriptional 3'-end processing of the nascent RNA transcript. The transcript is cleaved at a sequence-specified location (poly(A) site); this cleavage releases the mRNA, to whose 3'-end the *poly A tail* is immediately added. The remaining RNA transcript is selectively unraveled and degraded, which leads to changes in the synthesizing Pol II that trigger termination of transcription, past the end of the gene. Capping of the 5'-end and splicing of the mRNA also take place cotranscriptionally. mRNA export to the cytoplasm is largely dependent on mRNA processing.

Chromatin Remodeling and Remodelers

Chromatin remodeling is key for efficient PIC assembly and RNA transcript elongation. Many TFs, coregulators, and elongation factors are themselves chromatin remodelers or recruit chromatin remodeler complexes. Two main types of such complexes are recognized: (1) ATP-dependent chromatin remodelers, whose ATPase activity helps to move, eject, or restructure nucleosomes, and (2) enzymes that catalyze covalent posttranslational modifications (PTMs) of histones such as acetylation, methylation, phosphorylation, and others. Histone PTMs affect histone-DNA and nucleosome-nucleosome interactions, and may also create or remove sites for the binding of regulatory proteins (e.g., bromodomain and chromodomain proteins, which bind to acetylated and methylated histones, respectively), hence impacting gene expression.

Nucleosomal histones contain many modifiable amino acid residues, in particular Arg and Lys residues that concentrate in the histone N-terminal tail. Lys acetylation/deacetylation is an important PTM with functional consequences. Acetylation (catalyzed by histone acetyltransferases) favors transcription, whereas deacetylation (catalyzed by histone deacetylases, HDACs) represses transcription. This is explained because acetylation neutralizes the positive charge of the modified Lys residue, weakening the histone tail interactions with the negatively charged sugar-phosphate DNA backbone, and disrupts nucleosome-nucleosome interactions required for higher order chromatin structure, thus increasing the accessibility of the transcription machinery to DNA.

Dynamic, reversible chromatin remodeling is impacted by signaling pathways and the cell's nutritional/metabolic status and contributes to short-term control of gene expression, being inherent to transcription itself. Besides transient remodeling, there are more persistent chromatin epigenetic modifications, such as DNA methylation and certain histone modifications (e.g., the methylation of H3 on Lys 9) that can be passed to the next cell generation during mitosis or meiosis and hence affect transcription and other genome outcomes (e.g., stability) in the long term, intergenerationally and perhaps even transgenerationally. Epigenetic modifications are influenced by age, genetics, and environmental factors including dietary exposures. Epigenetic modifications may affect chromatin structure and capabilities for transacting factor recruitment, and from this gene expression. For instance, DNA methylation of CpG islands promoters usually (though not always) associates with gene silencing through the action of methylated DNA-binding proteins that interfere with efficient PIC assembly, directly or indirectly (see chapters on epigenetics).

SIGNALS AND MECHANISMS IN TRANSCRIPTION REGULATION

The interplay between chromatin structure, with its dynamic organization in nucleosomes and higher order structures, and TF-DNA interactions is key to the regulation of gene expression at the level of transcription (Fig. 3.1). Regulation of transcription is largely achieved by regulating the regulators, i.e., by altering the abundance and/or activity of TFs, coregulators, chromatin remodelers, and other proteins involved. Signals impacting gene expression include nutritional and metabolic signals, hormones, second messengers, and different forms of stress (thermic stress, oxidative stress, organelle stress, hypoxia, DNA lesions, etc.). Some signals function in a rather direct manner, for instance by binding to a TF molecule, while others function indirectly, through effects on metabolic fluxes and/or sensing signaling pathways that in turn impact gene regulatory mechanisms, often by conditioning covalent modifications of TFs or TF-interacting proteins (Fig. 3.2).

TFs are a common target of regulation. Changes brought about by gene stimulatory signals may for instance (1) unmask a DNA binding domain or a transactivation domain in the TF molecule; (2) favor the concentration of the TF in the nucleus (by exposing a nuclear localization signal and/or masking a nuclear export signal in the TF molecule, or by affecting the import and/or export operating mechanisms); (3) increase TF stability, for instance, when the modified TF molecule becomes resistant to degradation by the ubiquitin proteasome system; (4) promote effective TF dimerization (many TF function as dimers and can dimerize with different partners, leading to multiple,

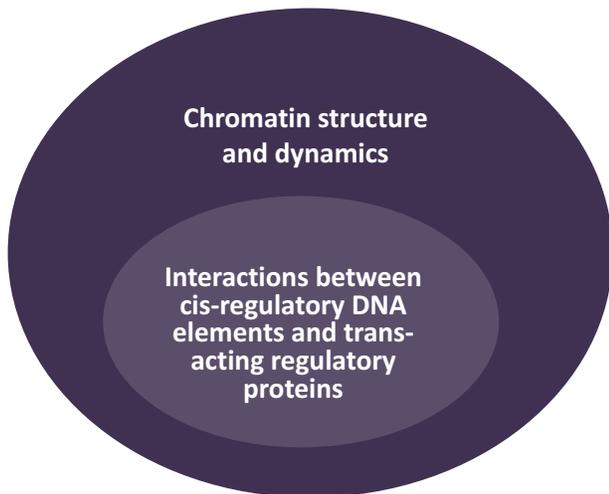


FIGURE 3.1 Regulation of transcription depends on the interplay between chromatin structure and transcription factor-DNA interactions.

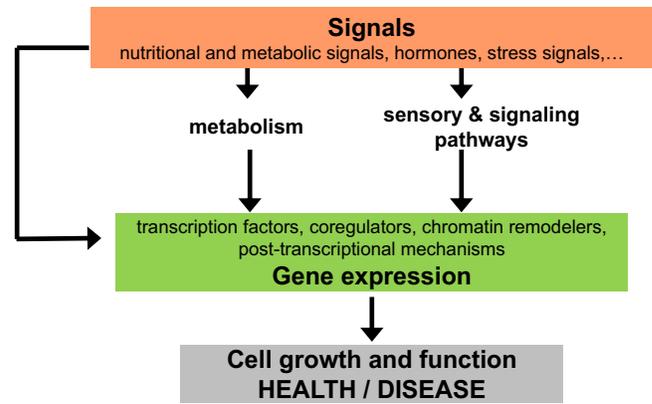


FIGURE 3.2 Signals and general mechanisms in gene regulation.

functionally nonequivalent combinations); (5) promote the proteolytic cleavage of a TF precursor to render the active TF molecule; and/or (6) disrupt protein-protein interactions that were keeping the TF in an inactive state, thus leading to TF activation. These various mechanisms are not mutually exclusive.

Ligand-modulated TFs are best exemplified by the nuclear receptors (NRs), a superfamily of structurally related proteins with 48 members in humans and 49 in mice. NRs bind to defined cis-regulatory DNA sequence elements and affect target gene transcription in a ligand-dependent manner. Some NRs exert also ligand-dependent modulation of other TFs through protein-protein interactions. Among the NRs, class I (steroid hormone receptors) bind classic steroid hormones, require hormone binding for translocation from the cytoplasm into the cell nucleus and DNA binding, and function as homodimers. Whereas class II are receptors for dietary and metabolic lipid signals, are bound to cognate DNA regardless of ligand-binding status but require ligand binding for interchanging corepressors by coactivator proteins, and function as heterodimers with retinoid X receptor (RXR, an NR itself).

Coregulators are equally important as the TFs. Since coregulators are usually promiscuous, meaning that each interacts with multiple TFs, the control of coregulator expression levels and/or activity through regulatory PTMs represents an efficient manner to orchestrate complex gene regulatory programs in cells.

CONTROL OF GENE EXPRESSION BY DIETARY FACTORS

The following general considerations about dietary regulation of gene expression are outlined here (Fig. 3.3):

- Diet-related factors affecting gene regulation include: specific food components (macronutrients,

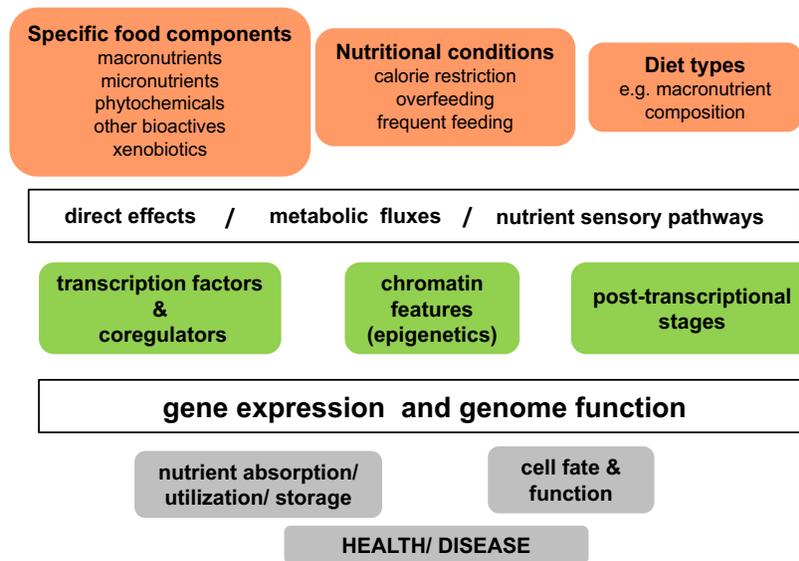


FIGURE 3.3 Overview of gene expression control by dietary factors. Dietary factors including specific food components, nutritional conditions and diet types impact gene expression directly or indirectly and at multiple stages, from transcription to translation and protein activity. At the transcriptional level, they affect the activity of trans acting factors as well as chromatin regulation in the short- and long-term. Changes in gene expression induced by dietary factors can help the cell/organism to utilize or store the dietary nutrients, condition cell fate and function (for instance, by adjusting cell growth and proliferation to nutrient availability), and be linked to health maintenance or the development of diet-related diseases.

micronutrients, phytochemicals and other bioactives, xenobiotics) and their combinations; diet types (e.g., diets of different macronutrient composition); and nutritional conditions (e.g., calorie restriction, overfeeding, frequent feeding).

- Dietary factors impact gene expression directly or indirectly, and at multiple stages, from transcription to translation and protein activity. At transcription level, they affect both TF activity and chromatin regulation in the short and long term.
- Changes in gene expression induced by dietary factors can help the cell/organism to utilize/store dietary nutrients, condition cell fate (e.g., cell growth and proliferation) by adapting it to nutrient availability, and be linked to health maintenance or the development of diet-related diseases.
- The impact of dietary factors on gene expression can be tissue-specific and be dependent on genetic polymorphisms, DNA, and histone modifications (epigenetics), other concurrent factors in the individual (e.g., their health status), or other components of the diet. These sources of variability add additional layers of complexity to the control of gene expression by dietary factors and to research in this area. Gene expression in blood cells may reflect changes going on in different tissues due to genetics or environment, and is thus emerging as a useful tool for advances in nutrigenetics and nutrigenomics.

DIETARY REGULATION OF TRANSCRIPTION FACTORS

Main mechanisms of TF regulation by dietary factors are illustrated next through relevant examples.

Direct Control of TF Activity

Specific food compounds behave as ligands of NR superfamily members or as precursors of endogenous NR ligands. For instance, retinoic acid receptors are activated by the vitamin A derivative retinoic acid, the vitamin D receptor (VDR) is activated by active vitamin D (1,25-dihydroxycholecalciferol), peroxisome-proliferator activated receptors (PPARs) can be activated by certain dietary fatty acids and other dietary compounds, and pregnane X receptor and constitutive androstane receptor are both activated by xenobiotics in foods. Upon ligand binding, these NRs gain affinity for coregulators that provide transcription transactivation activities.

The following examples highlight the significance and therapeutic potential of the NR gene regulatory circuitries controlled by dietary factors. In the enteric tract, NRs help to “absorb the good and neutralize the bad”: in the enterocytes, vitamin D-activated VDR induces genes for proteins involved in the intestinal absorption of dietary calcium, while xenobiotic-activated NRs induce genes for proteins involved in the metabolism and

detoxification of these potentially toxic compounds. As another example, diet is the vehicle of PPAR α activating molecules of potential interest for health, such as certain polyunsaturated fatty acids (PUFA); 9-oxo-10(E),12(E)-octadecadienoic acid (found in tomatoes); pterostilbene (a resveratrol analogue abundant in blueberries); astaxanthin (a carotenoid abundant in seafood); and phytanic acid (a branched-chain fatty acid found in dairy products). PPAR α transcriptional activity suppresses inflammatory pathways and favors mitochondrial substrate oxidation and fatty acid catabolism, and preclinical studies support hypolipidemic and antiobesity effects of PPAR α activating ligands of dietary origin.

Direct Control of TF Abundance

This mechanism is exemplified by the negative control by PUFA of the lipogenic TF sterol regulatory element-binding protein-1 (SREBP-1). SREBP-1 is synthesized in a precursor form that is bound to the endoplasmic reticulum (ER) membrane, and PUFAs inhibit the proteolytic processing of the precursor that renders the active SREBP-1 form, by stabilizing an inhibitory protein in the process, Insig-1. The direct target of PUFA is Ubx8, an ER membrane-bound protein that normally facilitates Insig-1 degradation: Ubx8 is inhibited upon interaction with PUFA. Additionally, PUFA potentially lowers SREBP-1 mRNA levels: the PUFA-mediated inhibition of SREBP-1 maturation disrupts an autostimulatory loop on SREBP-1 gene transcription, and, in addition, PUFAs stimulate SREBP-1 mRNA decay and may function as antagonist ligands of the liver X receptor (LXR), a nuclear receptor that normally transactivates the SREBP-1 gene.

Besides repressing SREBP-1, PUFAs repress the nuclear import of a second important lipogenic TF, carbohydrate response element binding protein (ChREBP). Overall, by promoting SREBP-1 and ChREBP inhibition and PPAR α activation, PUFAs promote a shift in hepatic fatty acid metabolism, from synthesis and storage to oxidation, resulting in a blood lipid-lowering effect.

Indirect Control of TFs Following Changes in Metabolic Fluxes

This mechanism is exemplified by glucose-mediated regulation of hepatic lipogenic capacity. High-carbohydrate low-PUFA diets increase liver lipogenic capacity independently of insulin, a response that favors the storage as fat of excess carbohydrate-ingested energy. When cellular glucose levels are high, increases in the flux through glucose-metabolizing pathways such as glycolysis, the pentose phosphate pathway, and the hexosamine biosynthetic pathway ensue the activation of the lipogenic TFs ChREBP, SREBP-1, and

LXR. These three TFs synergistically transactivate genes for enzymes involved in glucose utilization and lipid synthesis, such as L-pyruvate kinase, fatty acid synthase, and acetyl-CoA carboxylase, among others. LXR transactivates SREBP-1 and ChREBP gene transcription.

Mechanistically, three glucose-derived metabolites, glucose-6-phosphate (the first intermediate in intracellular glucose metabolism), fructose-2,6-bisphosphate (the major regulator of glycolysis), and xylulose 5-phosphate (a metabolite of the pentose phosphate pathway), have been implicated as positive modulators of ChREBP translocation to the nucleus and activation, through allosteric effects and effects on the ChREBP phosphorylation status. The hexosamine biosynthetic pathway, in its turn, produces uridine diphosphate N-acetylglucosamine (UDP-GlcNAc), which is the donor substrate for enzyme-catalyzed regulatory O-GlcNAcylation of many cellular proteins. O-GlcNAcylation of ChREBP stabilizes the ChREBP protein and increases its transcriptional activity toward lipogenic genes. Similarly, O-GlcNAcylation of LXR enhances its transcriptional activity toward the SREBP-1 gene and possibly other lipogenic targets including the ChREBP gene.

Indirect Control of TFs Following Changes in Nutrient/Energy Sensing Signaling Pathways

Under conditions of caloric restriction, cellular levels of AMP and NAD⁺ increase, leading to the activation of two important intracellular energy sensors and master regulators of cell metabolism, AMP-activated protein kinase (AMPK) and sirtuin 1 (SIRT1). AMPK and SIRT1 catalyze, respectively, regulatory phosphorylation and NAD⁺-dependent deacetylation of target proteins, among them TF and coregulators involved in the transcriptional control of energy, lipid and glucose metabolism. AMPK and SIRT1 share many common target protein substrates and can regulate each other. In general, AMPK and SIRT1 activation favors a more efficient use of lipid energy sources and respiratory metabolism. Both sensors enhance mitochondrial biogenesis and mitochondria quality maintenance mechanisms, owing to their action on PPAR gamma coactivator 1 α (PGC-1 α) and on proteins related to the autophagy of damaged mitochondria. PGC-1 α is a transcriptional coregulator that coactivates a constellation of TFs—such as estrogen-related receptor α , nuclear respiratory factors 1 and 2, and PPARs—to induce mitochondrial gene expression. Interestingly, certain dietary chemicals such as resveratrol, other polyphenols, and the NAD⁺ precursor nicotinamide riboside have been shown to activate SIRT1 and/or AMPK and to promote metabolic fitness in experimental animal models, independently of caloric restriction.

Another important nutrient/energy sensing pathway in cells is the mechanistic target of rapamycin (mTOR) pathway. Thus mTOR is a Ser/Thr protein kinase that is activated in response to insulin, growth factors, and amino acids, and inhibited under conditions of cellular glucose deprivation or energetic stress. mTOR drives cell growth by stimulating anabolic processes including protein, lipid, and nucleotide synthesis, and by inhibiting degradative catabolic processes such as autophagy. mTOR activity upregulates cellular protein synthesis and ribosome biogenesis by activating translation factors and by inducing ribosomal RNA and ribosomal protein expression. In addition, mTOR acts on specific TFs and coregulators such as signal transducer and activator of transcription 3, SREBPs, hypoxia-inducible factor 1 α , PGC-1 α , PPAR γ , and PPAR α to favor cell growth, proliferation and survival, lipid synthesis, angiogenesis, mitochondria biogenesis and adipogenesis, and to block hepatic fatty acid oxidation, among other responses.

DIETARY FACTORS AND METABOLISM AS CHROMATIN REGULATORS

Cellular levels of essential cofactors of chromatin remodelers are dependent on the cell's nutritional/metabolic status, which allows the integration of metabolic information into transcriptional control through chromatin changes. Nutrient abundance boosts the nuclear acetyl-CoA pool, thus favoring histone acetylation and gene transcription; interestingly, increased histone acetylation is predominantly observed on cell growth-promoting genes and genes related to nutrient utilization and storage, which in this way become upregulated. The opposite condition, i.e., nutrient deprivation/caloric restriction, also impacts the chromatin landscape, in this case through increases in the cellular levels of NAD⁺ and AMP. The NAD⁺-dependent sirtuins may affect gene expression through direct deacetylation of histones (sirtuins are class III HDACs), and AMPK activation regulates histone acetylation in a complex manner, through a variety of mechanisms. Dietary factors may condition methylation reactions as well, because specific nutrients (methionine, choline, folate, vitamin B12, vitamin B6, zinc, selenium) participate as substrates or necessary cofactors of enzymes in the one-carbon metabolism pathway that produces S-adenosylmethionine (SAM), the universal cellular methyl donor. Dietary manipulations affecting the intake of these nutrients have been shown to affect the methylation status of DNA in cells, with consequences on gene expression (DNA methylation generally associates with transcriptional silencing).

Additionally, there are food compounds able to allosterically modulate chromatin remodelers in cells. For instance, dietary phytochemicals behave as DNA methyltransferases inhibitors (e.g., polyphenols from black raspberry, apples, and green tea) or as type I and II HDACs inhibitors (e.g., curcumin, soy isoflavones, isothiocyanates from cruciferous vegetables, sulfur compounds from garlic, green tea polyphenols). Butyrate and other short-chain fatty acids produced upon fermentation of some types of dietary fiber by the colonic flora also inhibit type I and type II HDACs. Inhibition of DNA methyltransferases and/or HDACs after exposure to such dietary compounds has been shown to result in the reactivation of silenced tumor-suppressing genes in cancer cell lines and animal models of cancer, suggesting that these activities may contribute to anticancer action.

Exposure to certain dietary factors at sensitive periods such as the periconceptual period, uterine life, and early postnatal life may trigger persistent epigenetic changes, particularly of DNA methylation, affecting health outcomes in adulthood. This is concluded from animal studies and sustained by the (so far still limited) evidence available in humans. Dietary factors that have been implicated in long-term metabolic programming include the intake levels of one-carbon metabolism related nutrients (see above), overnutrition, malnutrition, alcohol exposure, and milk leptin.

DIETARY FACTORS AS MODULATORS OF MIRNAS

miRNAs are endogenous, small noncoding RNA molecules (of about 22 nucleotides) that function in post-transcriptional regulation of gene expression by binding to complementary sequences in target mRNAs, this resulting in mRNA cleavage, destabilization of the mRNA through shortening of its poly(A) tail, or a less efficient translation by ribosomes. It has been shown that dietary factors (e.g., macronutrients, minerals, trace elements, vitamins, plant bioactives) alter the production of mammalian microRNAs. Additionally, there is some initial evidence to suggest that, following absorption, secretion, and tissue distribution, microRNAs contained in foods of plant (e.g., rice) and animal (e.g., cow milk, human breast milk) origin may affect gene expression in mammalian tissues.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Regulation of gene expression is essential to life and health and is sensitive to endogenous and dietary and

other environmental factors that exert their action at multiple levels and by multiple mechanisms. There are important and intimate links between diet/nutrition, metabolic status, signaling pathways, and gene regulation. Postgenomic technologies including microarray hybridization and massive next-generation sequencing (RNA-Seq) enable nowadays high-throughput analysis, at genome-wide level, of the transcriptome and its dynamics (presence and quantity of RNAs, RNA expression patterns). Further, changes induced by dietary or other factors in TF binding site occupancy, chromatin features, global patterns of transcript decay, and translational profiling can also be studied nowadays at a genome-wide level. Integration of results from these varied technologies will ultimately allow a comprehensive, system-wide understanding of gene expression control and of diet-gene-health relationships in modern nutrition science.

Suggested Reading

- Danino, Y.M., Even, D., Ideses, D., Juven-Gershon, T., 2015. The core promoter: at the heart of gene expression. *Biochim Biophys Acta* 1849, 1116–1131.
- Gerhauser, C., 2018. Impact of dietary gut microbial metabolites on the epigenome. *Philos Trans R Soc Lond B Biol Sci* 373 (1748).
- Li, X., Egervari, G., Wang, Y., Berger, S.L., Lu, Z., 2018. Regulation of chromatin and gene expression by metabolic enzymes and metabolites. *Nat Rev Mol Cell Biol* 19, 563–578.
- Reynés, B., Priego, T., Cifre, M., Oliver, P., Palou, A., 2018. Peripheral blood cells, a transcriptomic tool in nutrigenomic and obesity Studies: current state of the art. *Compr Rev Food Sci Food Saf* 17, 1006–1020.
- Ross, S.A., Davis, C.D., 2014. The emerging role of microRNAs and nutrition in modulating health and disease. *Annu Rev Nutr* 34, 305–336.
- Schmidt, D.R., Mangelsdorf, D.J., 2008. Nuclear receptors of the enteric tract: guarding the frontier. *Nutr Rev* 66, S88–S97.
- Stover, P.J., James, W.P.T., Krook, A., Garza, C., 2018. Emerging concepts on the role of epigenetics in the relationships between nutrition and health. *J Intern Med* 284, 37–49.
- Weake, V.M., Workman, J.L., 2010. Inducible gene expression: diverse regulatory mechanisms. *Nat Rev Genet* 11, 426–437.

This page intentionally left blank

The Role of Nutrition in DNA Replication, DNA Damage Prevention and DNA Repair

Michael Fenech

Genome Health Foundation, North Brighton, SA, Australia

Glossary

Chromosomal instability A cellular phenotype that exhibits a recurring and abnormally high rate of structural and numerical chromosomal aberrations.

Chromosome aberrations Structural and numerical chromosome aberrations. Structural aberrations include chromosome breaks and chromosome rearrangements. Numerical aberrations include abnormal number of chromosomes mainly arising due to chromosome mal-segregation during mitosis.

Comet assay A molecular cytogenetic assay in which DNA strand breaks are visualized and quantified as the amount of DNA migrating out of nuclei embedded in an agarose gel after electrophoresis.

Dietary reference values A set of recommendations for nutritional intake based on currently available scientific knowledge on the nutritional intake requirements for healthy living across life stages.

DNA damage Damage to DNA in its various forms including DNA strand breaks, DNA base modifications mainly due to chemical adducts on DNA, cross-links between DNA strands and within DNA strands, and DNA-protein cross-links. DNA damage can be observed and studied at the molecular and cytogenetic level.

DNA methylation A chemical modification of the nucleotide cytosine involving the addition of a methyl group (a carbon atom combined with three hydrogen atoms). Methylated cytosine in CpG sites plays an important role in control of gene expression and the structural stability of important noncoding regions of chromosomes such as centromeres and subtelomere sequences.

DNA repair The process by which cells sense DNA damage and activate the DNA repair machinery to rectify abnormalities in DNA sequence or structure. The latter includes proteins that recognize the various types of DNA damage lesions and recruit the appropriate DNA repair enzymes to excise the damaged bases, replace them with normal nucleotides, and ligate broken ends of the DNA.

DNA replication The process by which DNA molecules are replicated within a cell during the DNA synthesis phase of mitosis. This involves DNA polymerases and nucleotides required for the replication processes.

DNA replication stress The inability of a cell to efficiently and correctly replicate DNA due to inadequate or inappropriate supply of nucleotides, defects in DNA polymerases, deficiency in nutrient cofactors required for nucleotide synthesis and DNA polymerase function, the presence of DNA damage in the nucleotide sequence

such as DNA strand breaks, and DNA adducts that can stall the DNA replication process.

Epigenetics Epigenetics is the study of heritable or nonheritable changes in gene expression that occur due to subtle chemical changes such as methylation of the nucleotide cytosine in gene promoter sequences or acetylation of the histone proteins with which DNA is structurally associated. These changes do not involve gene sequence mutations or deletions.

γ H2AX assay Phosphorylated γ H2AX is a protein that accumulates at sites of DNA breaks and its detection by fluorescent antibodies is used to determine the amount of DNA strand breaks within a nucleus.

Micronuclei Whole chromosomes or chromosome fragments that have been excluded from the main nuclei during mitosis and that have been enclosed within a separate nuclear membrane. They have the appearance of small nuclei usually with a diameter that is within the range of 1/3 to 1/16 of that of a normal main nucleus. Micronuclei are scored in postmitotic cells such as exfoliated buccal cells, erythrocytes, and lymphocytes.

Minerals Minerals are inorganic substances required by the body in small amounts for a variety of different essential roles such as cofactors for enzymes or as part of the integral structure of proteins.

Nutriome The nutrient profile of a dietary pattern or nutrient supplement based on its dietary components and their amounts or concentration.

Vitamins Organic compounds that cells and organs in the body require for their essential metabolic functions, but which cells cannot synthesize and can only be obtained from foods or dietary supplements.

INTRODUCTION

Life as we know it depends entirely on the capacity of cells to utilize energy and molecules in the environment for cellular function and reproduction. Multicellular animal organisms, including humans, acquire energy and essential nutrients from foods. Some of these essential nutrients are required for DNA synthesis, maintenance of normal chromosome structure, repair of DNA

damage caused by nutrient deficiency and/or environmental genotoxins, and for the control of gene expression by epigenetic mechanisms. This review provides a brief outline of the role of nutrition in DNA replication, DNA damage prevention, and DNA repair.

ROLE OF NUTRITION IN DNA REPLICATION, DNA DAMAGE PREVENTION, AND DNA REPAIR

At the most basic level nutrition plays an important role by providing essential precursor molecules for the de novo synthesis of purine and pyrimidine nucleotides that determine the genetic code in DNA (Lane and Fan, 2015). Examples of precursor molecules provided by nutrition that are required for nucleotide synthesis are methyl donors such as folate, vitamin B12, and methionine. Folate plays a critical role in the de novo synthesis of purines such as 10-formyl tetrahydrofolate, and in the synthesis of pyrimidines such as 5,10-methylenetetrahydrofolate by donating formyl and methyl moieties, respectively (Fig. 4.1). Furthermore, 5-methyltetrahydrofolate is required for the synthesis of methionine and S-adenosyl methionine, which is required for the conversion of cytosine to 5-methylcytosine, the fifth nucleotide in DNA. 5-methylcytosine plays a critical role in (1) structural chromosome stability, particularly in the pericentromeric region; and (2) in the maintenance of DNA methylation patterns that control normal cellular gene expression and phenotype in response to environmental cues (Fenech, 2012).

Vitamin B12 is vital in DNA metabolism because it determines the availability of folate for nucleotide and methionine synthesis due to its role as an essential

cofactor for the enzyme methionine synthase, which enables the transfer of a methyl group from 5-methyltetrahydrofolate to homocysteine to generate methionine and tetrahydrofolate. The latter, which is the form of folate that can be stored in cells after polyglutamation, is required as a precursor to generate (1) 10-formyl tetrahydrofolate essential for purine nucleotide (ATP, GTP) synthesis; and (2) 5,10-methylenetetrahydrofolate, required for pyrimidine nucleotide synthesis (TTP, CTP). When vitamin B12 and/or folate are deficient, the cell cannot generate enough nucleotides to properly replicate DNA, leading to DNA replication stress, DNA strand breaks, and chromosome aberrations (Fenech, 2012).

Minerals from foods, such as zinc and magnesium, are essential as cofactors for DNA polymerases required for nuclear and mitochondrial DNA synthesis (Sharif et al., 2012; Hartwig, 2001). Furthermore, it was discovered that nuclear iron-sulfur (Fe-S) cluster proteins conduct functions in DNA replication processes involving the enzymes POLD1, PRIM2, and DNA2 (Fuss et al., 2015). Deficiencies in these minerals may also cause DNA replication stress and result in DNA breaks, deletions, and point mutations.

There is overwhelming evidence that a large number of micronutrients (vitamins and minerals) are required as cofactors for enzymes or as part of the structure of proteins (metalloenzymes) involved in DNA repair, prevention of oxidative damage to DNA, as well as maintenance and methylation of DNA. The role of micronutrients in the maintenance of genome stability has been extensively reviewed (Ames, 2006; Ferguson and Philpott, 2008; Fenech, 2010). Examples of micronutrients involved in various genome stability processes are given in Table 4.1, and some of the various possible

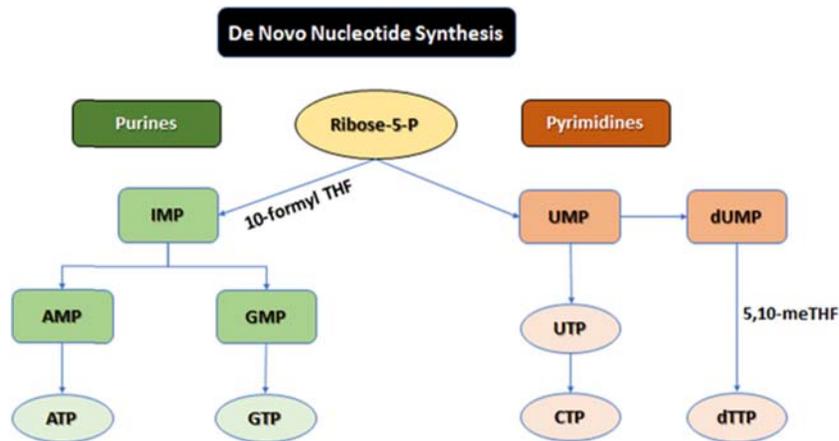


FIGURE 4.1 The role of folate, as 10-formyl tetrahydrofolate (10-formyl THF) and 5,10-methylene tetrahydrofolate (5,10-meTHF) in de novo nucleotide synthesis. 5,10 meTHF, 5,10- methylenetetrahydrofolate; 10-formyl THF, 10-formyl tetrahydrofolate; AMP, adenosine monophosphate; ATP, adenosine triphosphate; CTP, cytosine triphosphate; dTTP, deoxy-thymine triphosphate; dUMP, deoxy-uridine monophosphate; GMP, guanosine monophosphate; GTP, guanosine triphosphate; IMP, inosine monophosphate; ribose-5-P, ribose-5-phosphate; UMP, uridine monophosphate; UTP, uridine triphosphate.

TABLE 4.1 Examples of the Role and the Effect of Deficiency of Specific Micronutrients on Genomic Stability.

Micronutrient/s	Role in Genomic Stability	Consequence of Deficiency
Vitamin C, vitamin E, antioxidant polyphenols (e.g., caffeic acid)	Prevention of DNA and lipid oxidation	Increased baseline level of DNA strand breaks, chromosome breaks, oxidative DNA lesions, and lipid peroxide adducts on DNA.
Folate and vitamins B2, B6, and B12	Maintenance methylation of DNA; synthesis of nucleotides including dTTP from dUMP and efficient recycling of folate.	DNA replication stress. Uracil misincorporation in DNA, increased chromosome breaks, and DNA hypomethylation. Telomere shortening and dysfunction.
Niacin (also known as nicotinic acid and nicotinamide, which are precursors of nicotinamide adenine dinucleotide (NAD) and nicotinamide adenine dinucleotide phosphate (NADP))	Required as substrate for poly[ADP-ribose] polymerases (PARP), which detect single-stranded DNA breaks (SSBs) and recruit DNA repair complexes to the site of SSBs. PARP activity is also required for telomere length maintenance.	Increased level of SSBs in DNA, increased chromosome breaks and rearrangements, and sensitivity to mutagens. Impaired telomere length regulation.
Zinc	Required as a cofactor for Cu/Zn superoxide dismutase, endonuclease IV, function of p53, Fapy glycosylase, and in Zn finger proteins required for genome maintenance such as PARP in DNA damage recognition, hOGG1 for repair of oxidized guanine and PrimPol, a primase-polymerase involved in nuclear and mitochondrial DNA replication.	DNA replication stress. Increased DNA oxidation, DNA breaks, and elevated chromosome damage rate.
Iron	Fe–S cluster proteins conduct important functions in DNA replication (POLD1, PRIM2, DNA2) as well as in DNA repair processes (XPD) or the regulation of telomere length (RTEL1). Also required as component of ribonucleotide reductase and mitochondrial cytochromes.	Increased DNA replication stress. Reduced DNA repair capacity and increased propensity for oxidative damage to mitochondrial DNA. Impaired telomere length control.
Magnesium	Required as cofactor for a variety of DNA polymerases, in nucleotide excision repair, base excision repair, and mismatch repair. Essential for microtubule polymerization and chromosome segregation.	Reduced fidelity of DNA replication. Reduced DNA repair capacity. Chromosome segregation errors.
Manganese	Required as a component of mitochondrial Mn superoxide dismutase.	Increased susceptibility to superoxide damage to mitochondrial DNA and reduced resistance to radiation-induced damage to nuclear DNA.
Calcium	Required as cofactor for regulation of the mitotic process	Mitotic dysfunction.
Selenium	Selenoproteins involved in methionine metabolism and antioxidant metabolism (e.g., selenomethionine, glutathione peroxidase I).	Increase in DNA strand breaks, DNA oxidation, and telomere shortening.

mechanisms by which micronutrient deficiency could cause DNA damage, accelerate senescence and chromosomal instability are illustrated in Fig. 4.2. A key point is that micronutrient deficiency or excess can cause genome damage of the same order of magnitude—if not greater—than the genome damage caused by exposure to significant doses of environmental genotoxins, such as chemical carcinogens, ultraviolet radiation, and ionizing radiation. For example, chromosomal damage in cultured human lymphocytes caused by reducing folate concentration from 120 nmol/L to 12 nmol/L is equivalent to that induced by an acute exposure to 0.2 Gy of low linear energy transfer ionizing radiation (e.g., X-rays), a dose of radiation that is approximately 10 times greater than the annual allowed safety limit of exposure for the general population (Fenech, 2010). The sensitivity of DNA damage to micronutrient deficiency is underscored by the fact that there are at least eight human DNA repair glycosylases dedicated to the removal of the type of DNA base damage (e.g., 8-hydroxydeoxyguanosine, uracil) that is produced when either antioxidant micronutrients (such as zinc, vitamin C, and vitamin E) or folate and vitamin B12 are deficient, respectively (Wood et al., 2001).

Results from a population study suggest that at least nine micronutrients affect genome stability in humans in vivo (Fenech et al., 2005). This cytogenetic epidemiological study on 190 healthy individuals (mean age 47.8 years, 46% males) was designed to determine the association between dietary intake, estimated using a food frequency questionnaire, and genome damage (chromosome breakage and/or loss) in lymphocytes,

measured using the cytokinesis-block micronucleus assay. Multivariate analysis of baseline data showed that (a) the highest tertile of intake of vitamin E, retinol, folate, nicotinic acid (preformed), and calcium is associated with significant reductions in micronucleus (MN) frequency (i.e., -28% , -31% , -33% , -46% , and -49% , respectively (all $P < .005$) relative to the lowest tertile of intake; and (b) the highest tertile of intake of riboflavin, pantothenic acid, and biotin was associated with significant increases in MN frequency (i.e., $+36\%$ ($P = .054$), $+51\%$ ($P = .021$), and $+65\%$ ($P = .001$), respectively, relative to the lowest tertile of intake). Mid-tertile β -carotene intake was associated with an 18% reduction in MN frequency ($P = .038$) relative to the lowest tertile. However, the highest tertile of intake ($>6400 \mu\text{g}/\text{day}$) resulted in an 18% increase in MN frequency relative to the lowest tertile. In interpreting the data from this study, it is important to note that micronutrients usually exhibit metabolic dose-response effects, in which both deficiency and excess can be deleterious. It is also probable that, in a specific mixed diet, depending on intake level of an individual, some of the micronutrients may be outside the intake range that is optimal for prevention of genome instability. The results for β -carotene suggest an optimum for genome stability between 4000 and 6000 $\mu\text{g}/\text{day}$, with a tendency for marked increase in genome damage at higher or lower intake levels. On the other hand, the apparent preventive effects of genome damage associated with vitamin E, retinol, folic acid, preformed nicotinic acid, and calcium were still increasing at the highest tertile of intake, suggesting that an optimum could be achieved at even higher levels

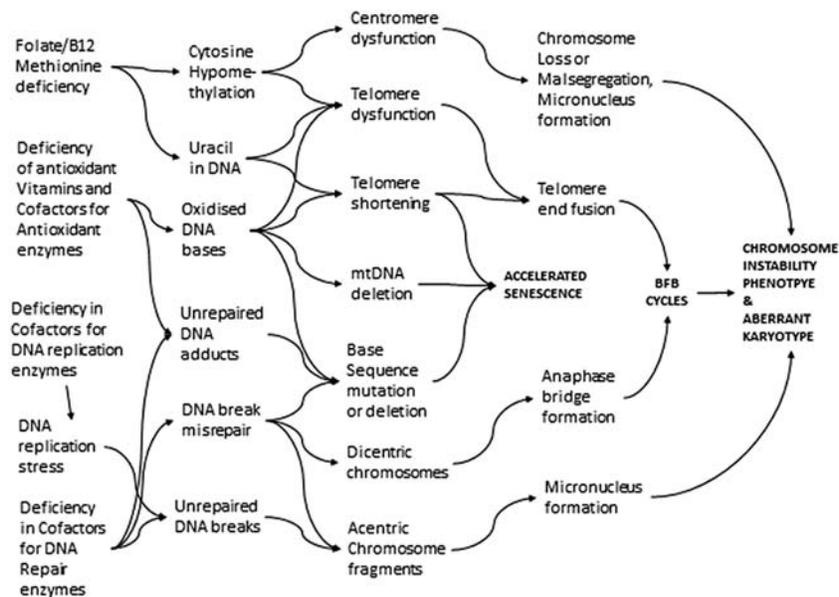


FIGURE 4.2 The various mechanisms by which micronutrient deficiency leads to DNA damage and chromosomal instability. BFB, chromosome breakage-fusion-bridge.

of intake or that the maximum beneficial effect is achieved at these levels of intake. For example, the highest tertile of intake for folate was $>256 \mu\text{g}/\text{day}$, consistent with a number of studies showing that developmental defects and cancer, as well as biomarkers for cardiovascular disease risk—such as homocysteine—are minimized at folate intake levels of $400 \mu\text{g}/\text{day}$ or greater (Fenech, 2012; Bailey et al., 2015).

It is also important to consider the combined effects of micronutrients such as calcium or riboflavin with folate, because epidemiological evidence suggests that these dietary factors may interact in modifying the risk of cancer (Fenech, 2010; Lamprecht and Lipkin, 2003). Interactive additive effects were observed, including the protective (-46%) effect of increased calcium intake, and the exacerbating ($+42\%$) effect of higher riboflavin consumption on increased genome damage caused by low folate intake. The results from this study illustrate the strong impact of a wide variety of micronutrients and their interactions on genome integrity, depending on level of intake. The effects of these interactions highlight the need to consider not only individual micronutrients but also micronutrient combinations at varying dosages. The term *nutriome* was indeed introduced to define this important aspect of nutritional requirements that needs much attention (Fenech, 2014). The ultimate goal is to define, for each individual, the nutriome that matches their genome to allow optimal genome stability to be achieved.

The amounts of micronutrients that appear to be protective against genome damage vary greatly between foods, and careful choice is needed to design dietary patterns optimized for genome health maintenance. Because dietary choices vary between individuals, due to taste preferences (which may be genetically determined) or cultural or religious constraints, several options are required, and supplements may be needed to cover gaps in micronutrient requirements. Clearly, the development or identification of nutrient-dense foods and ingredients that are rich in micronutrients required for DNA replication and repair and for prevention of genome-damaging events is essential in making it feasible for individuals to achieve their daily nutrient requirements for genome integrity maintenance without intake of excess calories.

It was recently shown that bioinformatic, data-driven analysis methods, including network-based approaches, can also be applied to food and nutrition to determine the “nutritional landscape” of foods to design healthy diets that meet specific nutritional requirements (Kim et al., 2015). A similar approach might be used to identify the minimum sets of foods required for prevention of DNA damage.

An important development in recent years is the observation that, although DNA damage measured by

both molecular and cytogenetic biomarkers tends to increase with age (Fenech, 2010), appropriate dietary changes or supplementation by specific micronutrient combinations can attenuate the rate of increase, or even reduce, the level of these biomarkers. A limitation of most these intervention studies is that they are usually performed over brief periods of time (3–6 months) and limited to single tissues, usually blood cells and single assays of DNA damage. A more robust approach should include: (a) measurements performed in easily accessible multiple tissues, such as blood lymphocytes and neutrophils, as well as buccal cells, the latter being representative of epithelial cells, which comprise the bulk of the body; and (b) a comprehensive set of complementary biomarkers of genome damage. These should measure both chromosomal instability events—readily performed using micronucleus cytome assays—as well as molecular lesions, such as DNA strand breaks (comet assay or γH2AX assay), DNA hypo- or hypermethylation, telomere length, DNA oxidation, and mitochondrial DNA deletions.

FUTURE DIRECTIONS

Given this burgeoning knowledge on the association of micronutrients with well-validated DNA damage biomarkers (e.g., micronuclei, telomere length), it has become feasible to define dietary reference values based on requirements to prevent the DNA damaging effects of nutritional deficiency. The earliest evidence of DNA damage induced by malnutrition was that reported by Armendarces et al., in 1971, showing a 5.5-fold increase in chromosome aberrations in children with protein-calorie malnutrition compared to well-nourished children (Armendarces et al., 1971). In contrast, a 2.4-fold increase in micronuclei and 6.4-fold increase in DNA strand breaks was observed in children who are overweight and obese (Scarpato et al., 2011). Thus, determining dietary reference values for macronutrients should also be feasible.

In addition, data from both in vitro models and cross-sectional in vivo data indicate the possibility of significant nutrient-genotype interactions that can modify nutritional requirements for DNA damage prevention. Although it appears feasible to define personalized nutritional requirements for prevention of DNA damage using genotype information, the level of evidence available remains inadequate, at this stage, to be actionable based on recently published guidelines (Grimaldi et al., 2017). A research road map is required to determine the type and quality of data necessary to start defining personalized nutritional advice for DNA damage prevention, preferably focused, initially, on well-studied nutrients such as folate and vitamin B12.

CONCLUSIONS

Nutrition plays an important role in providing essential molecules required for DNA synthesis and DNA repair. The extent of DNA damage induced by nutrient deficiency and excess is as high as that induced by known chemical and physical mutagens. For this reason, and given the known consequences of developmental defects and increased risk of degenerative diseases and accelerated aging associated with increased genomic instability, much more attention should be given to defining the dietary reference values for DNA damage prevention. The prospect of personalized nutrition for DNA damage prevention exists, but more data are required to reliably translate this concept into practice.

References

- Ames, B.N., 2006. Low micronutrient intake may accelerate the degenerative diseases of aging through allocation of scarce micronutrients by triage. *Proc Natl Acad Sci U S A* 103 (47), 17589–17594.
- Armendares, S., Salamanca, F., Frenk, S., 1971. Chromosome abnormalities in severe protein calorie malnutrition. *Nature* 232, 271–273.
- Bailey, L.B., Stover, P.J., McNulty, H., Fenech, M.F., Gregory 3rd, J.F., Mills, J.L., Pfeiffer, C.M., Fazili, Z., Zhang, M., Ueland, P.M., Molloy, A.M., Caudill, M.A., Shane, B., Berry, R.J., Bailey, R.L., Hausman, D.B., Raghavan, R., Raiten, D.J., 2015. Biomarkers of nutrition for development-folate review. *J Nutr* 145 (7), 1636S–1680S.
- Fenech, M.F., 2010. Dietary reference values of individual micronutrients and nutriomes for genome damage prevention: current status and a road map to the future. *Am J Clin Nutr* 91 (5), 1438S–1454S.
- Fenech, M., 2012. Folate (vitamin B9) and vitamin B12 and their function in the maintenance of nuclear and mitochondrial genome integrity. *Mutat Res* 733 (1–2), 21–33.
- Fenech, M.F., 2014. Nutriomes and personalised nutrition for DNA damage prevention, telomere integrity maintenance and cancer growth control. *Cancer Treat Res* 159, 427–441.
- Fenech, M., Baghurst, P., Luderer, W., et al., 2005. Low intake of calcium, folate, nicotinic acid, vitamin E, retinol, beta-carotene and high intake of pantothenic acid, biotin and riboflavin are significantly associated with increased genome instability -results from a dietary intake and micronucleus index survey in South Australia. *Carcinogenesis* 26 (5), 991–999.
- Ferguson, L.R., Philpott, M., 2008. Nutrition and mutagenesis. *Annu Rev Nutr* 28, 313–329.
- Fuss, J.O., Tsai, C.L., Ishida, J.P., Tainer, J.A., 2015. Emerging critical roles of Fe-S clusters in DNA replication and repair. *Biochim Biophys Acta* 1853 (6), 1253–1271.
- Grimaldi, K.A., van Ommen, B., Ordovas, J.M., Parnell, L.D., Mathers, J.C., Bendik, I., Brennan, L., Celis-Morales, C., Cirillo, E., Daniel, H., de Kok, B., El-Sohehy, A., Fairweather-Tait, S.J., Fallaize, R., Fenech, M., Ferguson, L.R., Gibney, E.R., Gibney, M., Gjelstad, I.M.F., Kaput, J., Karlsen, A.S., Kolossa, S., Lovegrove, J., Macready, A.L., Marsaux, C.F.M., Alfredo Martinez, J., Milagro, F., Navas-Carretero, S., Roche, H.M., Saris, W.H.M., Traczyk, I., van Kranen, H., Verschuren, L., Virgili, F., Weber, P., Bouwman, J., 2017. Proposed guidelines to evaluate scientific validity and evidence for genotype-based dietary advice. *Genes Nutr* 12, 35. <https://doi.org/10.1186/s12263-017-0584-0>. eCollection 2017.
- Hartwig, A., 2001. Role of magnesium in genomic stability. *Mutat Res* 475 (1–2), 113–121.
- Kim, S., Sung, J., Foo, M., Jin, Y.S., Kim, P.J., 2015. Uncovering the nutritional landscape of food. *PLoS One* 10 (3), e0118697.
- Lamprecht, S.A., Lipkin, M., 2003. Chemoprevention of colon cancer by calcium, vitamin D and folate: molecular mechanisms. *Nat Rev Cancer* 3, 601–614.
- Lane, A.N., Fan, T.W., 2015. Regulation of mammalian nucleotide metabolism and biosynthesis. *Nucleic Acids Res* 43 (4), 2466–2485.
- Scarpato, R., Verola, C., Fabiani, B., Bianchi, V., Saggese, G., Federico, G., 2011. Nuclear damage in peripheral lymphocytes of obese and overweight Italian children as evaluated by the gamma-H2AX focus assay and micronucleus test. *FASEB J* 25 (2), 685–693.
- Sharif, R., Thomas, P., Zalewski, P., Fenech, M., 2012. The role of zinc in genomic stability. *Mutat Res* 733 (1–2), 111–121.
- Wood, R.D., Mitchell, M., Sgouros, J., Lindahl, T., 2001. Human DNA repair genes. *Science* 291 (5507), 1284–1289.

Genotyping and Sequencing

Teresa Ezponda¹, Gorka Alkorta-Aranburu², Felipe Prósper¹, Xabier Agirre¹

¹Area de Hemato-Oncología, IDISNA, CIBERONC, Centro de Investigación Médica Aplicada (CIMA), University of Navarra, Pamplona, Spain; ²Area de Hemato-Oncología, CIMA Lab Diagnostics, IDISNA, Centro de Investigación Médica Aplicada (CIMA), University of Navarra, Pamplona, Spain

Glossary

ATAC Assay for transposase accessible chromatin
 cDNA Complementary DNA
 ChIP Chromatin immunoprecipitation
 CNV Copy number variant
 ddNTPs di-deoxynucleotidetriphosphates
 DNA Deoxyribonucleic acid
 EMR Electronic medical records
 FAIRE Formaldehyde-assisted isolation of regulatory elements
 FPKM Fragments per kilobase per million
 GWAS Genome-wide association study
 LD Linkage disequilibrium
 mRNA Messenger ribonucleic acid
 LOD Limite of detection
 NGS Next generation sequencing
 PCR Polymerase chain reaction
 RNA Ribonucleic acid
 RPKM Reads per kilobase per million
 rRNA Ribosomal ribonucleic acid
 SNP Single nucleotide polymorphism
 TPM Transcripts per million

A SHORT BACKGROUND

Our inherited genome determines how our organism processes nutrients. These effects are studied by nutrigenetics, which focuses on identifying (and understanding) the modifying effect by the genome and the epigenome in micronutrient uptake and metabolism, as well as dietary effects on health. On the other hand, diet affects our genome (for example, determining DNA damage at both the molecular and chromosome levels) and epigenome (DNA methylation and histone modifications), as well as the expression of RNAs, proteins, and metabolites. Nutrigenomics focuses on identifying (and understanding) the molecular-level effect of nutrients and other dietary bioactives with the use of high-throughput technologies, including genomic, epigenomic, transcriptomic, proteomic, and metabolomic methods.

All of these high-throughput approaches can be studied independently or in an integrated manner. In all cases, the design of the experiment, the generation of “omics” data, and the statistical analyses that follow are key steps that need to be carefully planned to draw reliable conclusions about the interactions between genes and nutrients. In this chapter, we will discuss the types of data and analyses that are employed in nutrigenetics and nutrigenomics, mainly focusing on the two most used approaches (genomics and transcriptomics), paying special attention to critical steps that must be considered in the design and analysis of data.

GENETIC AND GENOMIC DATA: FROM GENETICS TO GENOMICS

Genetics refers to the study of genes, genetic variation, and heritability in living organisms, which is usually applied to the study of a limited number of genes or genetic variants. Genomics, on the other hand, is the study of the complete set of genetic material within an organism. The shift from genetic to genomic studies has come with the development of high-throughput data analyses (single nucleotide polymorphisms—[SNPs] and gene expression arrays or next-generation sequencing [NGS]) as well as computational biology capable of handling high-output data (bioinformatics).

Before the genomics era, the candidate gene approach, in which a gene to be studied was selected based on its known putative function, was the only method used to identify genes containing genetic variation that were capable of affecting nutrient uptake or metabolism. Thanks to Sanger sequencing, genomic variants within a set of samples were identified, which were then tested for their association with the phenotype of

BOX 5.1 NGS DATA ANALYSIS PROCESS

In contrast to Sanger sequencing, NGS is a **multiday process** that takes the isolated nucleic acid through a multistep library preparation protocol, followed by a massive parallel sequencing procedure. Once sequencing data are generated, a complex data analysis process needs to be carried out over weeks or even months. The complexity of NGS makes it error-prone, requiring NGS-focused laboratory setups and trained staff (**technicians, bioinformatician/computational biologists, and data analysts**). NGS is sensitive to nucleic acid quantity, but also to its **quality**: non-degraded long nucleic acid molecules improve the quality of sequencing data. Thus, NGS-based laboratories are constantly improving and optimizing nucleic acid isolation protocols in addition to incorporating new laboratory equipment to assess whether isolated nucleic acids samples fulfill defined quantity and quality metrics before processing and sequencing procedures start. The processing of nucleic acids for NGS or massive parallel sequencing is known as “library preparation,” for which a wide variety of protocols exist. Not choosing the most adequate **library preparation protocol** based on the sample type, as well as isolated nucleic acid quantity, and quality metrics will negatively affect the quality of the sequencing data. Although the sequencing steps of NGS protocols are simple, a wide variety of sequencing machines exist, each with their own sequencing speed as well as inherited error types and rates. To avoid result biases, once generated, the large amount of NGS data needs to be cautiously analyzed considering not only the specific error types of the chosen sequencing machine but also the characteristics of the input sample and the library

preparation protocol. Because NGS data can be biased by many factors, quality assurance and control approaches of nucleic acid isolation/extraction, library preparation, sequencing, and data themselves are critical before conclusions are drawn. NGS is a new and heterogeneous technology; thus, best practice guidelines for its use are still being developed and implemented; for example, the bioinformatics scripts and thresholds chosen to filter the data need to be validated to identify (1) higher than expected sequencing error rates (i.e., estimated by including in the sequencing run a control sample provided by the company—PhiX for Illumina instruments); (2) a low proportion of bases with Q scores below Q30, with or without an unexpected drop in per-base sequencing quality (where Q is the estimated probability of the base call being wrong; thus, a quality score of 30 represents an error rate of 1 in 1000, with a corresponding call accuracy of 99.9%); (3) a high proportion of duplicated sequences (i.e., polymerase chain reaction [PCR] duplicates); (4) a high proportion of reads unmapped to the reference genome; and (5) a low sequencing depth and/or low percentage of genome covered at defined sequencing depth. Sequencing depth describes the number of times a nucleotide has been sequenced by NGS. Therefore, although higher values are better, minimum requirements depend on the error rate of the sequencing technology as well as on the NGS data application itself: inherited versus somatic genetic variation identification, SNP versus CNV identification, or genome sequencing versus gene expression studies (see the discussion on RNA sequencing [RNA-seq] in the transcriptomic section).

interest. Sanger sequencing has been the most widely used sequencing method for several decades. This method employs the four standard deoxynucleotide triphosphates (dNTPs) as well as smaller amounts of the four modified di-deoxynucleotide triphosphates (ddNTPs), each labelled with a different fluorescent dye (ddATP, ddCTP, ddGTP, and ddATP). The random incorporation of ddNTPs causes early termination of chain elongation during in vitro DNA replication by DNA polymerase, and the different fluorescent dye helps identify the incorporated nucleotide.

SNPs are the most common type of sequence variation in the human genome. DNA microarrays were developed to genotype known SNPs within one gene or a few genes at first, and quickly moved to the development of genome-wide SNP arrays capable of genotyping over one million known SNPs. This was the dawn of the era of genome-

wide association studies (GWAS), which test whether any of the genome-wide genotyped genetic variation associates with the phenotype of interest. Although SNPs are the most common type of genetic variation being tested in GWAS, copy number variants (CNVs) are also known to influence food response (i.e., the increase in amylase gene copy number is associated with increased enzymatic activity and starch digestion). However, in addition to data quality, expertise, and sequence context, the limit of detection (LOD) of both Sanger and SNP arrays varies depending on the type of genetic variation. In addition, insertions and deletions are not easily identified by Sanger sequencing or even SNP arrays. Furthermore, Sanger or SNP arrays cannot detect low-level genetic variation in mitochondria or even microbiome studies (see subsequent discussion on metagenomics). Therefore, despite its intrinsic complexity (Box 5.1, Fig. 5.1), advances in

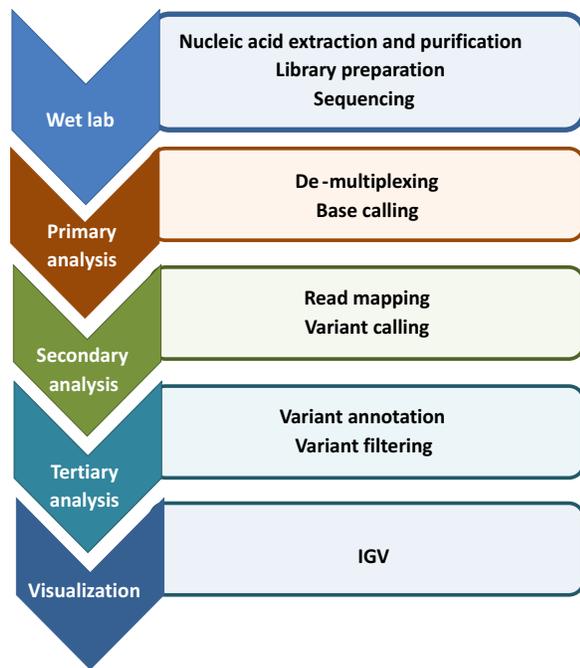


FIGURE 5.1 Stepwise representation of next-generation sequencing analyses, including wet lab experiments (nucleic acid extraction and purification, library preparation, and sequencing) as well as several computational analyses (de-multiplexing, base-calling, read aligned, variant calling, variant annotation, and filtering). Finally, aligned reads/identified variants can be visualized with tools such as the Integrative Genomics Viewer (IGV).

technology with the development of NGS have been instrumental in improving the detection of low-level and more diverse genetic variation (single nucleotide variants, and even complex insertions, deletions, duplications or even inversions) and also in transitioning from genetic to genomic approaches.

Indeed, genomic approaches have allowed genetic variation to be studied without the need to define gene candidates to follow up. For example, studies have begun to apply genome-wide scans to identify previously unknown genetic variants in known and even unknown genes that harbor variations that modify the response to diet. Genome-wide SNP arrays revealed this possibility, but the variation to be tested had to be previously known. With the development of NGS, one can identify both known and novel genetic variation influencing the response to diet within specific genes via targeted NGS, or all genes via whole-genome sequencing (i.e., a single sequencing process that determines the complete DNA sequence of an organism). Thus, with those genome-wide genetic variation datasets, GWAS have emerged as an approach to identifying genes harboring genetic variations that affect the uptake and metabolism of nutrients.

GWAS is a powerful tool for understanding the genetic architecture of the response to diet in nutrigenetics. However, several factors need to be considered to avoid spurious associations between genetic variation and our

phenotype of interest. For example, **multiple test correction** (e.g., the Bonferroni correction) is necessary because of a very high cumulative likelihood of finding false positives when hundreds of thousands to millions of tests are performed within a GWAS. In addition, different studies may have diverse sets of SNPs genotyped; thus, to assess the effect of the same allele across multiple distinct studies (i.e., metaanalysis), **GWAS datasets are imputed**. Linkage disequilibrium (LD) describes the degree to which an allele of one SNP is inherited or correlated with an allele of another SNP but within a population. Thus, the reference panel used for imputation must contain haplotypes drawn from the same study population, and the analysis of imputed genotypes should account for the uncertainty in genotype state generated by the imputation process. In addition, it is true that the SNP influencing a biological system that ultimately leads to the phenotype (functional SNP) can be directly genotyped in the study and may be statistically associated with the trait. However, LD creates another possibility: the influential SNP is not typed directly; instead, a **tag SNP** in high LD with the influential SNP is typed and is statistically associated with the phenotype. Thus, additional studies are required to map the precise location of the influential SNP to avoid spurious results.

Statistically, **quantitative traits** improve the power to detect a genetic effect among others, because traits without well-established quantitative measures have increased error associated with individuals being classified as either case or control. Furthermore, a phenotype that can be **uniformly** studied by clinicians will ease the **validation** of any genetic study during replication. The use of electronic medical records (EMRs) is increasingly widespread. However, EMRs were established for clinical care and administrative purposes—not for research: thus, biases may be introduced to studies. Therefore, **covariate adjustment** becomes necessary to reduce spurious associations owing to sampling artifacts or biases in the study design, although it may reduce statistical power. In a sample with multiple ethnicities, population structure is one of the most important covariates to consider. Allele frequencies differ across human subpopulations, so when phenotype differences also exist among ethnic groups, ethnic-specific SNPs (functional or not) will likely be associated with the trait as a result of population stratification. This is why a replication sample should be drawn from the same population as the GWAS.

SNP-array based genotyping data have been important for GWAS. With the arrival of affordable NGS technology, one million SNPs of SNP arrays have been replaced by the entire genomic sequence of three billion nucleotides, challenging bioinformatics with data storage, quality control, and analysis first, but also offering

the possibility of merging genomic data with those from other high-throughput technologies, such as the transcriptome, proteome, or metabolome.

GENE EXPRESSION VIA TRANSCRIPTOMICS

Transcriptomics is the study of the landscape of a complete set of RNA transcripts expressed by a cell at a specific moment. Transcriptomic analyses compare expression profiles between distinct cell populations and identify gene expression changes in response to different treatments such as dietary factors. These analyses require high-throughput methods such as microarray analysis or RNA-seq, by which RNA profiles are identified using NGS.

Before the genomics era, and similar to SNP array-mediated analysis for GWAS, the first transcriptome studies were performed employing **gene expression microarrays**. Expression microarrays consist of probes of oligonucleotides complementary to known transcripts, in which a sample of fluorescently labeled complementary DNA (cDNA) generated from the RNA of origin is hybridized. This tool has been widely used for transcriptome profiling. Nevertheless, its use has limitations that need to be considered. First, some technical issues are inherent in probe manufacturing, such as slide heterogeneity, spotting variation, printing irregularities, and background noise. Second, special caution must be taken when analyzing expression microarray data, because several probes targeting the same transcript may show distinct results according to the different hybridization reactions of each probe. Finally, expression microarray studies detect known transcripts, but identification of novel transcripts or novel types of RNA genes can be achieved only with the use of NGS. Thus, besides avoiding all of the issues mentioned for expression microarrays, a main advantage of RNA-seq is that both discovery and quantification of transcripts can be combined in a single high-throughput assay. RNA-seq has become widely employed and its use has become standard for transcriptome analysis by the entire life sciences research community, well beyond the genomics community. Tools for RNA-seq have continuously evolved with different technologies available. In this section, we will describe important aspects of technologies that need to be considered when performing these types of studies.

A key point to planning a successful RNA-seq analysis is that the generated data need to have the potential to answer the biological questions of interest: therefore, every RNA-seq study needs to be optimized for both its design and analysis. A good experimental design is crucial for this type of study. Several factors need to be

considered, including the selection of the number of replicates to be studied, the library type, and the sequencing depth itself. Moreover, several potential biases in the sequencing process itself need to be considered in the design of the analysis, as we will discuss next.

Number of replicates—The first consideration when planning an RNA-seq experiment is to determine the number of replicates needed for the study. This number depends on both the biological variability of the system under study, which can be difficult to control, and the technical variability of the RNA-seq protocol, with some steps such as RNA extraction and library preparation accounting for the higher number of biases. Although for many experiments three **biological replicates** are considered the minimum for reliable inference from the results, the statistical power also needs to be considered to be able to detect significant differences among the groups being studied.

Generation of libraries—When executing RNA-seq experiments, one of the first steps is to generate a cDNA library from the RNA of the study before sequencing. Because messenger RNA (mRNA) constitute only about 1%–2% of the total RNA in a cell, the production of RNA-seq libraries requires the prior removal of ribosomal RNA (rRNA), which is by far the most abundant type of RNA, so that the reads from the mRNA are not masked by rRNA. Because of the high content of the latter type of RNA, the method used to remove it is crucial to the experimental design. Two main approaches are currently used in eukaryotes: enriching for mRNA using poly(A) selection or depleting rRNA. The first approach is the most widely employed but it requires a high proportion of mRNA with minimal degradation. Thus, in some experimental situations, such as those starting from tissues, RNA cannot be obtained with enough integrity to produce good poly(A) RNA-seq libraries therefore, such cases require ribosomal depletion. In the case of bacterial samples, in which mRNA is not polyadenylated, the only option is ribosomal depletion.

Another important point of the experimental design is whether the information from transcripts generated from both strands of DNA needs to be conserved. In this sense, the use of approaches that do not retain information expressed in the DNA strand is very common, but their analysis and quantification are more complex in regions where different transcripts overlap. This problem is solved with strand-specific protocols, which have received increasing attention and are the option of choice for studies that seek to identify new transcripts such as novel long non-coding RNAs. Another aspect to consider when generating the library is the size of fragments to be sequenced, which entirely depends on the aim of the study. Although fragments of less than 500 bp are normally generated, when the experiment

aims for isoform expression analysis, longer reads are preferable because they are more easily mapped. Finally, the sequencing experiment can be either single-end or paired-end, depending on whether the library is sequenced from one or two sides of the cloned DNA, respectively. The latter option is usually chosen for experiments with the goal of identifying novel transcripts.

Sequencing depth or library size—Once the library is generated, the sequencing part of the experiment is carried out. The sequencing depth or library size, refers to the total number of reads generated in a single experiment, which will depend on the complexity of the targeted transcriptome. Augmenting the sequencing depth results in the detection of more transcripts and more precise quantification. Although such a situation is generally desirable for samples with good quality, a profound depth is not advisable for samples in which the quality is a limitation, because it can result in the detection of transcriptional noise.

Analysis of RNA sequencing data—Once the sequencing stage of the RNA-seq procedure is finished, analysis of the output data starts. This is a long process with multiple steps, including obtaining raw reads, read alignment, data normalization, quantification of gene

expression, and detection of differential gene expression among samples. At each point, quality controls are carried out to ensure the quality of the data. First, **raw reads** are inspected to detect PCR artifacts, sequencing errors, and the potential contamination of the libraries. The analysis focuses on sequence quality, duplicated reads, the presence of adaptors included in the library generation, and overrepresented sequences. Values for each of these parameters may vary among organisms, but they should be the same among samples in the same experiment. Once raw reads pass these quality controls, they are **mapped** to either a reference annotated transcriptome or a reference genome, when the latter is available (Fig. 5.2). An important parameter to consider is the percentage of mapped reads, which indicates the sequencing accuracy and the presence of contaminating DNA. Some reads, especially when aligned to a transcriptome, may map to more than one position (“multi-reads”), but they should not be discarded. Normally, reads mapped against the genome are higher than those against the transcriptome, and they usually represent 70%–90% of raw reads. The uniformity of read coverage should be also considered when mapping reads, because it indicates the quality of the starting RNA.

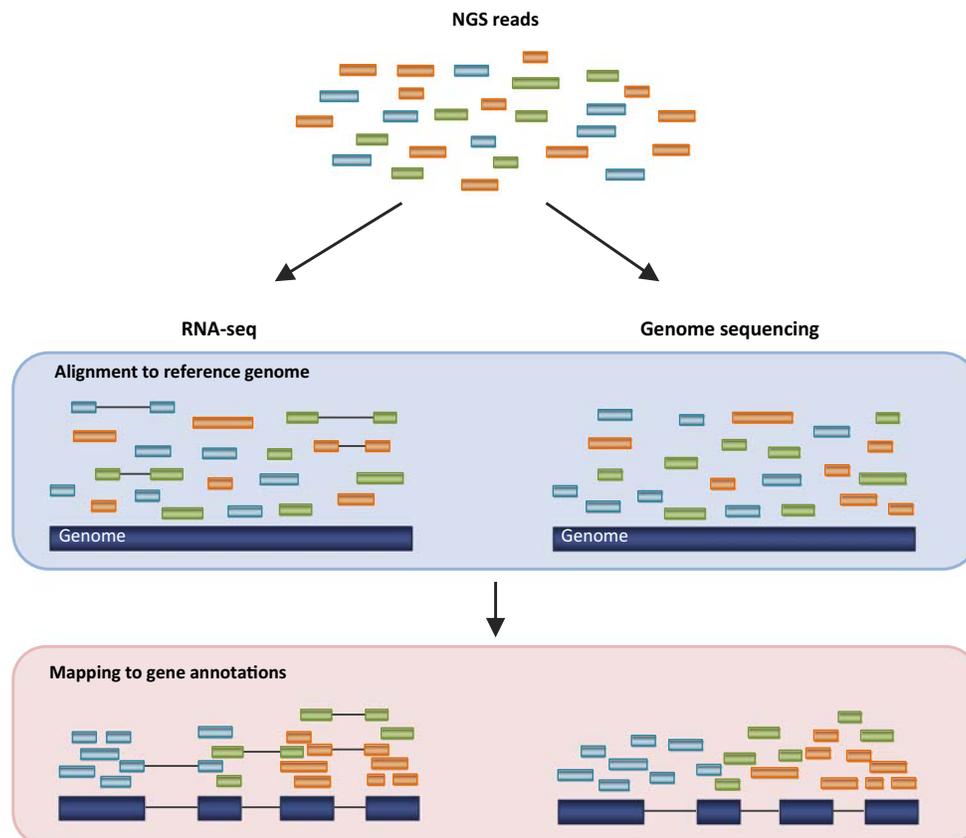


FIGURE 5.2 Representation of alignment and mapping of reads from next-generation sequencing (NGS) experiments. As displayed, reads from RNA sequencing (RNA-seq) experiments are generated from single exons or they can span multiple exons, whereas reads from genomic sequencing analyses map to both exonic and intronic regions.

Once the reads are mapped, the most common procedure to analyze RNA-seq data is to perform **transcript quantification**. To compare expression levels among samples, raw read counts need to be normalized by transcript length, library size, and sequencing biases. Many algorithms have been developed with this aim; they render normalized values such as reads per kilobase per million reads (RPKM), fragments per kilobase per million reads (FPKM), and transcripts per million (TPM), among others. Generated values should be checked for their guanine-cytosine content and gene length biases. At this point, the expected composition of the samples can be checked (i.e., for the absence of rRNA), which indicates the quality of the RNA purification step. At this point, reproducibility among technical replicates should also be checked; this should be generally high (Spearman $R^2 > 0.9$). In the case of biological replicates, reproducibility is usually much lower, although whether samples from the same biological group cluster together in a principal component analysis can be checked. Once transcript quantification is performed, **differential expression analyses** among samples can be carried out. The use of RPKM, FPKM, or TPM is usually encouraged despite read counts, as it normalizes for sequencing depth, which is the most important issue when comparing gene expression among samples. Many packages have been developed to assess differential gene expression, but the outcome of the analysis will greatly depend on the method used. Thus, it is generally recommended to note which method is the most appropriate for each experimental situation and to consider using more than one package.

Integration of results— An important aspect of RNA-seq is the possibility of integrating output data with other types of genome-wide studies, which allows for exploration of the connection between the regulation of gene expression and other aspects of molecular biology, such as chromatin regulation. Thus, integration of data moves toward an understanding of the system as a whole. Integrating **RNA and DNA sequencing** data can be used for several purposes, such as RNA editing analyses, SNP discovery, CNV analysis, and expression quantitative trait loci mapping. Moreover, this integrative analysis has been shown to improve mutation detection in low-purity tumors. The combination of RNA-seq data and **DNA methylation** data consists of analyzing the correlation between the expression of individual genes and CpG DNA methylation profiles, for which general linear models and logistic regression models have been used. In general, an inverse correlation is expected, in which the higher methylation is associated with low gene expression and vice versa. Nevertheless, significant correlations observed accounted for a small number of genes, which suggests that such analysis may be too simplistic to explain gene expression.

Another approach to integrating gene expression and DNA methylation data is to use a network-interaction-based analysis, a method that identifies sets of genes that have simultaneous differential DNA methylation and distinct gene expression. The integration of RNA-seq and **chromatin immunoprecipitation and sequencing (ChIP-seq)** of transcription factor data can provide information about the activating or repressive action of the transcription factor under study. When ChIP-seq experiments involve histone marks, integration of the data can elucidate epigenomic gene regulatory mechanisms. Finally, integration of RNA-seq data and information regarding **chromatin accessibility**, such as DNase-seq, formaldehyde-assisted isolation of regulatory elements-seq (FAIRE-seq), or assay for transposase accessible chromatin-seq (ATAC-seq), provides information about the chromatin architecture of the region of interest in relation to the expression of overlapping genes.

METAGENOMICS

Improved understanding about our genome and transcriptome is bringing us closer to personalized nutrition and medicine. However, the impact of our diet is also affected by microbial communities in our gastrointestinal tract, because they respond to and modify the food we ingest. For example, the ability of our microbiome to digest substrates that our own enzymes cannot process is associated with the development of human diseases such as obesity. Until the development of NGS-based technology, only genomes of microbial species capable of growing in culture were sequenced. The term “metagenomics” refers to the sequencing of the collective genomes of any culture-free microbial community sample. Similar to genomics and transcriptomics, in this study, the choice of the sequencing platform is important. The choice and validation of tools to perform optimal assembly, binning (transforming a continuous characteristic into a finite number of intervals [the bins], as well as annotation and comparative analysis are also important to avoid result biases. NGS technologies providing longer sequencing reads are preferred. Assembly tools specifically developed to work with genome mixtures are critical, followed by binning methods employing supervised machine-learning algorithms and annotation methods that take advantage of constantly updated public genomic repositories. Sequences that have no similarity with known repository sequences can be novel genes that have no sequence or function similarity to known genes; but the mismatch can also result from sequencing or gene prediction errors. Moreover, DNA isolation methods, as well as primer pairs used for amplification,

sequencing platforms, and the analysis software, can bias data in a way that comparing outputs derived from different workflows make it difficult to obtain trustworthy results for their correct interpretation. For this reason, steps are being taken to develop more standardized methodologies and consensus parameters and threshold values that allow minimizing biases.

CONCLUSION

NGS is a complex and heterogeneous technology that enables faster and easier collection of big datasets, but its use implies the presence of inherited biases that need to be understood and considered before conclusions are drawn or studies are compared. However, advances in NGS-based host and microbial genome and transcriptome studies provide a rapid and —more comprehensive understanding of how nutrition affects humans, highlighting the need to collect “omic” data in addition to nutritional, clinical, or environmental factors. First, GWAS established the association between genetic variation and the physiological response to dietary compounds. NGS-mediated genomic and transcriptomic analysis greatly widened the amount of information available for studying the interaction between nutrients and genes. Then, another layer of complexity was added, with early studies associating genetic variation within the microbial community. Thus, nutrigenetic and nutrigenomic studies may require a systems biology approach to interpret the dynamic microbial community

that surrounds us, increasing the need for properly validated sets of bioinformatics tools needed not only for the acquisition, management, storage, and retrieval of such high-throughput datasets, but also for the extremely important steps of quality control and analysis.

Suggested Reading

- Bush, W.S., Moore, J.H., 2012. Chapter 11: Genome-wide association studies. *PLoS Comput Biol* 8 (12).
- Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szczesniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X., Mortazavi, A., 2016. A survey of best practices for RNA-seq data analysis. *Genome Biol* 17, 13.
- Goodwin, S., McPherson, J.D., McCombie, W.R., 2016. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* 17 (6), 333–351.
- Hardwick, S.A., Deveson, I.W., Mercer, T.R., 2017. Reference standards for next-generation sequencing. *Nat Rev Genet* 18 (8), 473–484.
- Johnson, R.C., Nelson, G.W., Troyer, J.L., Lautenberger, J.A., Kessing, B.D., Winkler, C.A., O'Brien, S.J., 2010. Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics* 11, 724.
- Quince, C., Walker, A.W., Simpson, J.T., Loman, N.J., Segata, N., 2017. Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol* 35 (9), 833–844.
- Stephens, Z.D., Lee, S.Y., Faghri, F., Campbell, R.H., Zhai, C., Efron, M.J., Iyer, R., Schatz, M.C., Sinha, S., Robinson, G.E., 2015. Big data: astronomical or genetical? *PLoS Biol* 13 (7), e1002195.
- Tuohy, K.M., Gougoulas, C., Shen, Q., Walton, G., Fava, F., Ramnani, P., 2009. Studying the human gut microbiota in the trans-omics era — focus on metagenomics and metabonomics. *Curr Pharmaceut Des* 15 (13), 1415–1427.
- Zentner, G.E.1, Henikoff, S.2, 2014 Dec. High-resolution digital profiling of the epigenome. *Nat Rev Genet* 15 (12), 814–827.

This page intentionally left blank

6

RNA Analyses

Francisco Javier Ruiz-Ojeda^{1,2}, Julio Plaza-Diaz^{1,2}, Angel Gil^{1,2,3}

¹Department of Biochemistry and Molecular Biology II, School of Pharmacy, University of Granada, Granada, Spain; ²Institute of Nutrition & Food Technology “Jose Mataix”, Biomedical Research Center, University of Granada, Armilla, Spain; ³CIBEROBN (Physiopathology of Obesity and Nutrition CB12/03/30038), Instituto de Salud Carlos III (ISCIII), Madrid, Spain

INTRODUCTION

Ribonucleic acid (RNA) was discovered in 1869 by J. F. Miescher. This molecule is a polynucleotide consisting essentially of chains with a repeating backbone of phosphate and ribose units to which nitrogenous bases are attached. RNA is unique among biological macromolecules in that it can encode genetic information, serves as an abundant structural component of cells, and also possesses catalytic activity [Krebs et al. \(2014\)](#).

RNA has been divided in two major classes ([Table 6.1](#)), the protein-coding RNAs and messenger RNAs (mRNAs), and two forms of abundant nonprotein coding RNAs delineated on the basis of size: the large ribosomal RNAs (rRNA) and long noncoding RNAs (lncRNA), and the small noncoding RNAs transfer RNAs (tRNA), the small nuclear RNAs (snRNA), and the micro and silencing RNAs (miRNA and siRNA). Circular RNAs (circRNAs) are a type of small non-coding RNAs produced from precursor mRNA (pre-mRNA) back-splicing of exons in eukaryote genes.

In all prokaryotic and eukaryotic organisms there are different RNA molecules, predominantly mRNA, rRNA, and tRNA. These types of RNA are directly involved in protein synthesis whereas the other RNAs participate in either mRNA splicing (snRNAs) or modulation of gene expression by altering mRNA function (mi/SiRNAs) and/or expression (lncRNAs) ([Table 6.1](#)).

In all species, the transcription begins with the binding of the RNA polymerase complex (or holoenzyme) to a special DNA sequence at the beginning of the gene known as the promoter. Activation of the RNA polymerase complex enables transcription initiation, and this is followed by elongation of the transcript. In turn, transcript elongation leads to clearing of the promoter,

and the transcription process can begin yet again. Transcription can thus be regulated at two levels: the promoter level (*cis* regulation) and the polymerase level (*trans* regulation). These elements differ among bacteria and eukaryotes.

In bacteria, all transcription is performed using a single type of RNA polymerase. This polymerase contains four catalytic subunits and a single regulatory subunit. Eukaryotic cells are more complex than bacteria in many ways, including in terms of transcription. Specifically, in eukaryotes, transcription is achieved by three different types of RNA polymerase (RNA polymerase I-III). These polymerases differ in the number and type of subunits they contain, as well as the class of RNAs they transcribe; that is, RNA polymerase I transcribes rRNAs, RNA polymerase II transcribes RNAs that will become mRNAs and also small regulatory RNAs, and polymerase III transcribes small RNAs such as tRNAs.

The process of RNA synthesis involves the general steps of initiation, elongation, and termination with 5–3' polarity; use of ribonucleotides (U replaces T as the complementary base for A in RNA); large, multi-component initiation complexes; and adherence to Watson–Crick base-pairing rules [Rodwell et al. \(2015\)](#).

EXTRACTION, ISOLATION, QUANTIFICATION, AND PURIFICATION OF RNA

RNA Extraction and Isolation

RNA extraction is one of the most fundamental methods used in molecular biology and gene expression evaluation. This process is recognized to be influenced by the quantity and quality of starting RNA. Purity

TABLE 6.1 Types, Characteristics, and Major Functions of RNA.

RNA	Abundance	Size (nt)	Stability	Function
CODING RNAS				
mRNA	2%–5% of total	1000–1500	Unstable to very stable	Carries the genetic information
NONCODING RNAS (NCRNAS)				
Short ncRNAs				
miRNAs	< 1% of total	19–31	Stable to very stable	Regulation of proliferation, differentiation, and apoptosis
siRNAs	< 1% of total			Posttranscriptional gene silencing
snRNAs	< 1% of total			mRNA splicing
tRNAs	15% of total			Decode a mRNA sequence into a protein
circRNAs	5%–10% of total	> 50		Protein production and transporting miRNAs inside the cell/ Regulation of mRNA splicing Chen (2016)
Long ncRNAs				
rRNA	80%	> 200	Very stable	Essential for protein synthesis

circRNA, circular RNA; *mRNA*, messenger RNAs; *miRNA*, microRNAs; *nt*, nucleotides; *rRNA*, ribosomal RNAs; *siRNA*, silencing RNA; *tRNA*, transfer RNAs.

and integrity of RNA are critical elements for the overall success of RNA-based analyses. RNA can be isolated from any biological material such as living or conserved tissues, cells, virus particles, or other samples for analytical or preparative purposes. High-quality intact RNA is a starting point in molecular biology, especially in real-time quantitative reverse-transcription polymerase-chain-reaction (qRT-PCR), microarrays, ribonuclease-protection-assay, in situ hybridization, Northern blot analysis, RNA mapping, in vitro translation, cDNA library construction, and any kind of array application; in all these procedures the integrity of the used total RNA should be checked appropriately.

Commonly, it is well known that RNA is an unstable molecule and sensitive to degradation by postmortem processes and inadequate sample handling or storage. Special care and precautions are required for RNA isolation due to the ubiquitous presence of RNases, which are enzymes present in blood, all tissues, as well as most bacteria and fungi in the environment. Therefore, successful RNA purification requires four important steps: effective disruption of cells or tissue; denaturation of nucleoprotein complexes; inactivation of nucleases (RNase); and staying away from contamination.

The quality and quantity of purified RNA is variable, and after extraction during long storage it is rather unstable. In particular, long mRNA fragments up to 10 kb are very sensitive to degradation. This can happen through cleavage of RNases introduced by handling with RNA samples. The most obvious problem concerns the degradation of the RNA and is best addressed by insisting that every RNA preparation should be rigorously assessed for quality and quantity. The extraction and purification procedure of total RNA must fulfill the following criteria:

- > free of protein (checked by absorbance 260 nm/280 nm ratio greater than 1.8)
- > free of genomic DNA
- > should be undegraded (28S:18S ratio should be roughly between 1.8 and 2.0, with low amount of short fragments)
- > free of enzymatic inhibitors for reverse transcription (RT) and PCR reaction, which is strongly dependent on the purification and cleanup methods
- > free of any substances that complex essential reaction cofactors, like Mg^{2+} or Mn^{2+}
- > free of nucleases for extended storage.

Reproducibility can be affected by a substantial quantity of predicament. First, all materials, including plastic and glassware, must be sterile. The source of RNA, sampling techniques, such as biopsy material, single cell sampling, and laser microdissection; RNA isolation techniques often differ significantly between processing laboratories. The RNA quality can be different between two extraction methods; the isolated total cellular RNA with the liquid extraction, e.g., Trizol (Roche Diagnostic, Germany) or TriFast (Peqlab, Germany), has different RNA quality, whereas only the type of homogenization is changed. The best RNA yield is obtained from tissue that has been diced into small fragments with a scalpel prior to being frozen by submerging them into liquid nitrogen. As well, the samples must be homogenized. Additionally, human or animal tissue sampling techniques and the time dependency until the tissue is stored safely in RNase inhibitors or RNA-later (Ambion, USA) need to be considered.

Extraction Methods

Phenol:Chloroform Method

Phenol:chloroform method is based on the sample lysis in cationic detergent guanidinium thiocyanate (GTC), followed by organic extractions and alcohol precipitation. Moreover, GTC is effective at inactivating endogenous ribonucleases. For better removal of DNA from the aqueous phase, addition of low-pH phenol is recommended. Chloroform is a purely organic solvent, unable to mix with the cell lysate. A biphasic emulsion forms when phenol and chloroform are added. The solution is properly interspersed, and then the centrifugation enables effective separation of the "upper phase" and "lower phase" containing the cell lysate and chloroform, respectively. At the interface, a white ring of precipitated proteins appears. Proteins, lipids, carbohydrates, and cell debris are removed through extraction of the aqueous phase with the organic mixture of phenol and chloroform. The "upper phase," which contains nucleic acids, is transferred into a new sterile test tube, where it undergoes a reaction with isopropanol resulting in precipitation of RNA. After further centrifugations, a white pellet representing total RNA is recovered. This characteristic white coloring is due to the presence of residual salts precipitating together with RNA. Such salts are removed by washing the pellet using ethanol (concentration varies from 75% to 90%, according to the manual instructions). Pure RNA is then dissolved in RNase-free water or in an appropriate buffer. To inhibit the RNase activity, a mixture of phenol:chloroform:isoamyl alcohol (25:24:1) is used [Tan and Yiap \(2009\)](#); [Vomelova et al. \(2009\)](#).

Adsorption Methods

Adsorption methods are based on the ability of RNA to create a linkage to specific surfaces in the presence of GTC as detergent. The sample is lysed in GTC, then exposed to the particular adsorbing material, and vigorously shaken, to facilitate binding. The contaminants are then eased off by centrifugation. After, the beads with adsorbed RNA are washed again. RNA is released using a specific GTC salt-free buffer. Following spinning, beads devoid of RNA are sedimented while the RNA remains in the supernatants. RNA binding beads will disperse more easily when the temperature of the mixture is warmer. Beads are fully resuspended before consecutive steps; otherwise, formation of aggregates will negatively influence RNA recovery. Beads have to be shaken thoroughly and their overdrying must be avoided. Afterwards, the aspiration of binding beads when removing supernatants should be prevented. This method is quick and easy, and kits are commercially available.

Isopycnic Gradient Method

This method is based on the dissolution of the sample of experimental material using GTC as detergent. The following equilibrium gradient centrifugation is run on CsCl or cesium trifluoroacetate, which are both soluble in ethanol. The RNA captured in the appropriate density fraction is then precipitated by ethanol. Both salts (chloride, trifluoroacetate) are strong chaotropic agents that effectively suppress RNases.

Special Methods

Eukaryotic mRNA features the polyadenylate "d (A)" tail (sequence at 3'), which is effectively used for mRNA trapping. Since synthetic oligo "d (T)" chains can be covalently bound to special matrices or beads, mRNA can be easily captured and selectively fastened to a particular surface. The resulting mRNA can be obtained by chromatographic methods, or by magnetic field, depending on the sort of adsorbing strategy. Commercial kits are available from most suppliers, differing slightly in specific materials used for mRNA adsorption; the oligo d (T) sequence can be bound to cellulose matrix, magnetic beads, polystyrene-latex particles, or glass filter. Residual salt contamination may inhibit mRNA purification substantially. Thus, in case of mRNA separation from the total cellular RNA, salt remnants have to be removed thoroughly by washing the RNA pellet with 70% ethanol.

Cell-Free mRNA

The cell-free nucleic acids, also called circulating nucleic acids, are present in the plasma, serum, or in

other biological fluids such as urine, bronchial lavage, and amniotic fluid. As a promising diagnostic target of the future, the importance of cell-free mRNA analysis is dramatically rising since it might represent an extremely potent biological marker of various tumors, viral infections and in prenatal diagnostics. For all purposes, extremely sensitive methods, capable of detecting low numbers of circulating nucleic acid copies, are needed.

Small RNA Molecules

RNA purification relies on organic extraction followed by alcohol precipitation. However, this procedure removes the majority of small RNAs, such as miRNA, from the whole RNA. To overcome this challenge, appropriate kits have been developed. For example, miRACLE miRNA Isolation Kit from Stratagene (La Jolla, CA) employs organic extraction followed by small RNAs binding on a silica fiber matrix. MirVana miRNA Isolation Kit (Ambion) provides a unique tool permitting purification of native RNAs, including small RNAs such as miRNA and siRNA, in parallel with the native protein. The kit employs organic extraction followed by immobilization of RNA on glass-fiber filters to purify either total RNA, or RNA enriched for small species. It is important to perceive that specific miRNAs have a characteristic tissue distribution, and some tissues are absolutely devoid of detectable levels of them at all.

RNA Quantity and Quality Assessment

RNA integrity assessment can be done by various methods: the classical gel optical density (OD) measurement, modern OD measurement via Nano-Drop, old-fashioned denaturing agarose gel-electrophoresis, or with highly innovative lab-on-chip technologies like Bioanalyzer 2100 (Agilent Technologies, USA) and Experion (Bio-Rad Laboratories, USA). Quantity and quality assessment using a UV/VIS spectrophotometer should be performed at multiple wave lengths at 240 nm (background absorption and possible contaminations), 260 nm (specific for nucleic acids), 280 nm (specific for proteins), and 320 nm (background absorption and possible contaminations). Based on the OD 260 the quantity, the quality is based on the ratio of OD 260/280, OD 260/240, or OD 260/320. An OD 260/280 ratio greater than 1.8 is usually considered an acceptable indicator of good RNA quality. Genomic DNA can compromise the OD 260 measurement and lead to overestimation of the actual RNA concentration. Moreover, the used buffer and high salt concentrations will interfere with the result of the optical measurement and therefore the calculated RNA concentrations might

be over or underestimated. The OD 260/280 method has been questioned, with a value of 1.8 corresponding to only 40% RNA, with the remainder accounted for by protein (Fleige and Pfaffl, 2006).

Modern spectrophotometric methods, like the Nano-Drop (ND-3300, NanoDrop Technologies, USA) in combination with RNA RiboGreen dye (Molecular Probes, Invitrogen, USA) can be used as a UV/VIS spectrophotometer for ultrasensitive quantification of RNA. The major advantage of the system is the very low sample consumption of 1–2 μ L, which is especially important when using precious materials like human biopsy or laser-dissected samples. Further, the ND-3300 measures a spectrum of the sample covering 400–750 nm, providing more information about the RNA integrity and other chemical contamination or the extracted RNA.

Another check involves gel electrophoresis with RNA either stained with SYBR Green dye (Molecular Probes) or the less sensitive ethidium bromide. However, the assessment of RNA integrity by inspection of the 18S and 28S ribosomal RNA bands using denaturing gel electrophoresis is a cumbersome, low-throughput method and requires significant amounts of precious RNA. Using the RiboGreen (Molecular Probes) reagent, the detection as little as 1 ng RNA/mL is possible, and can be measured in a reproducible manner.

The highly innovative lab-on-chip technologies like microfluidic capillary electrophoresis have become widely used, particularly in the gene expression profiling platforms. This was designed to perform RNA quality and quantity, beyond the RiboGreen reagent that only measures polymeric nucleic acids.

The Agilent 2100 Bioanalyzer (Agilent Technologies, USA) and the Experion (Bio-Rad Laboratories, USA) provide a framework for the standardization of RNA quality control. RNA samples are electrophoretically separated on a microfabricated chip and subsequently detected via laser-induced fluorescence detection. This technology requires only a very small amount of RNA sample down to 200 pg total RNA. The use of an RNA ladder as a mass and size standard during electrophoresis allows the estimation of the RNA band sizes. Integrity may be assessed by visualization of the 18S and 28S ribosomal RNA bands. An elevated threshold baseline and a decreased 28S:18S ratio, both are indicative of degradation. Degradation of the RNA sample produces a shift in the RNA size distribution toward smaller fragments and a decrease in fluorescence signal as dye intercalation sites are destroyed. The 28S/18S ratios are automatically generated by software applications both in Experion and 2100 Bioanalyzer. The RNA measurement using the lab-on-chip technology appears stable and relatively uninfluenced by contamination. RNA from tissue samples is typically classified according to the observation that the 28S rRNA peak

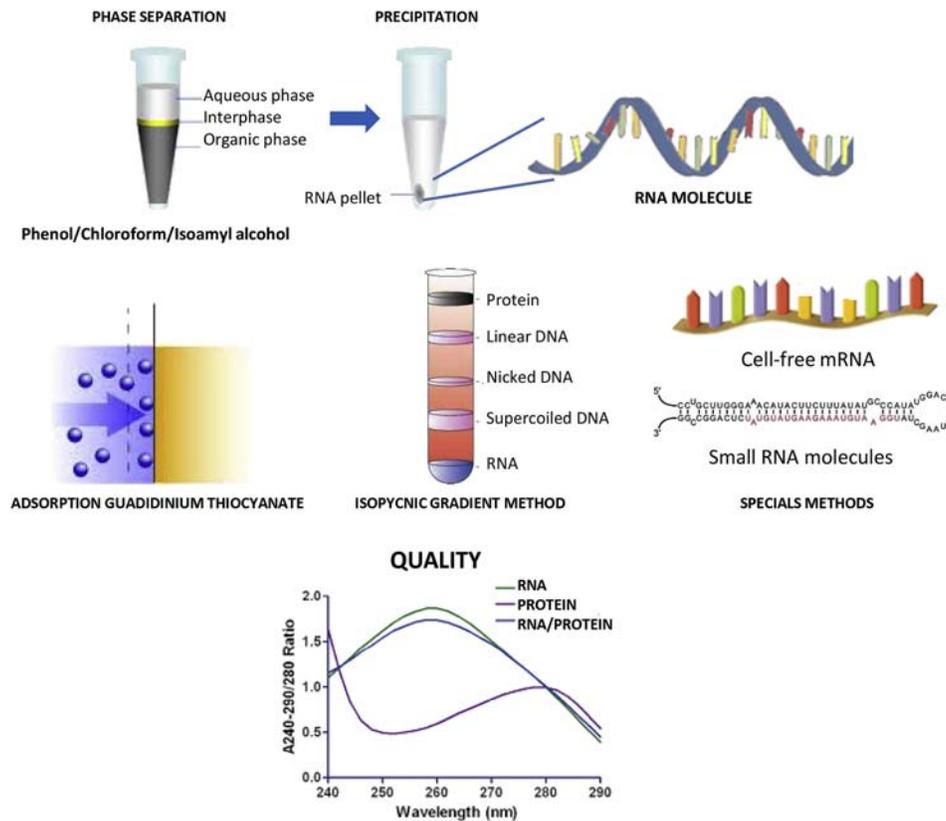


FIGURE 6.1 Isolation methods.

area should be approximately twice the quantity of that of the 18S in total RNA samples for the mRNA quality to be acceptable. In general, a 2.0 ribosomal ratio is regarded as perfect. However, this value is hardly obtained. The 28S/18S ratio may reflect unspecific damage to the RNA, including sample mishandling, postmortem degradation, massive apoptosis or necrosis, but it can reflect specific regulatory processes or external factors within the living cells. Therefore, the 28S/18S ratio has to be assessed for every single experiment and this is regarded as inadequate for the assessment of the quality O'Brien et al. (2012) (Fig. 6.1).

REVERSE TRANSCRIPTION AND REAL-TIME QUANTITATIVE PCR

Integrity of RNA and its Effects on Real-Time qRT-PCR

The determination of RNA quality is an important first step in obtaining meaningful data of gene expression. Thus, qRT-PCR shows sensitivity, good reproducibility, and a wide quantification range. For successful qRT-PCR and microarray experiments it is important to use intact RNA and cDNA synthesis and labeling

efficiency. Many factors as hemoglobin, fat, glycogen, cell constituents, Ca^{2+} , high genomic DNA concentration, and DNA-binding proteins are important contaminants that may inhibit RT as well as the PCR. Exogenous contaminants such as glove powder and phenolic compounds from the extraction process or the plastic can also have an inhibitory effect.

Reverse Transcription

The RT is an essential step. RT converts RNA into cDNA and the detailed description of the protocol and reagents should be provided. This documentation must include the amount of RNA reverse transcribed, priming strategy, enzyme type, volume, temperature, and duration of the reverse transcription step. The reverse transcription step must be carried out in duplicate or triplicate and that the total RNA concentration should be the same in every sample is recommended.

Real-Time Reverse Transcription Polymerase Chain Reaction (qRT-PCR)

Reverse transcription polymerase chain reaction (RT-PCR) is commonly used in molecular biology to detect RNA expression. While RT-PCR is used to qualitatively detect gene expression through creation of complementary DNA (cDNA) transcripts from RNA, quantitative—

real time PCR (qPCR) is used to quantitatively measure the amplification of DNA using fluorescent probes. qPCR can be utilized for quantification of RNA, in both relative and absolute terms.

The structure of the nucleic acid target has a substantial impact on the efficiency of reverse transcription and the PCR. Therefore, the positions of primers, probes, and PCR amplicons must take the folding of RNA templates into consideration.

The Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines were proposed by [Bustin et al. \(2009\)](#); this is a set of guidelines that describe the minimum information necessary for evaluating qPCR experiments. Included is a checklist to accompany the initial submission of a manuscript to the publisher. Regarding RNA quality, the MIQE guidelines describe that the quantification of RNA in the extracted samples is important because it is advisable that approximately the same amounts of RNA be used when comparing different samples. It is also important to test for and report the extent of genomic-DNA contamination and to record the threshold cutoff criteria for the amounts of such contamination that are tolerable. It is essential to show whether the RNA sample has been treated with DNase (including the type of DNase used and the reaction conditions) and to report the results from a comparison of crossing points (Cqs) obtained with positive and no-RT controls for each nucleic acid target, and also to document the quality assessment of RNA templates. RT-qPCR steps are required to be performed as a continuous, single-tube experiment. Key information to report includes data on RNA quantity, integrity, and the absence of RT or PCR inhibitors. The A_{260}/A_{280} ratio must be measured in a buffer at neutral pH. Instead, one should provide gel electrophoresis evidence at the least or, better yet, results from a microfluidics-based rRNA analysis or a reference gene/target gene integrity assay. Inhibition of RT activity or PCR should be checked by dilution of the sample (preferably) or use of a universal inhibition assay such as SPUD (a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations). If the RNA sample is found to be partially degraded, it is crucial that this information be noted because the assay's sensitivity for detecting a low-level transcript may be reduced and relative differences in the degradation of transcripts may produce incorrect target ratios.

The normalization by an internal reference gene reduces or even diminishes tissue-derived effects on qRT-PCR. Reliability of any relative RT-PCR experiment can be improved by including an invariant endogenous control in the assay to correct for sample-to-sample variations in qRT-PCR efficiency and errors in sample quantification. So-called relative quantification determines the changes in steady-state mRNA levels of a gene across multiple samples and gives a result relative to

the levels of an internal control RNA. It is based on the expression levels of a target gene versus an internal reference gene, often nonregulated housekeeping genes are prominent candidates.

An intact RNA does not guarantee good results because an RNA sample may contain inhibitors that can reduce reaction efficiency. These factors include length of the amplicon, secondary structure, and primer quality. The shapes of amplification curves differ in the steepness of any fluorescence increase and in the absolute fluorescence levels at plateau depending on background fluorescence levels. Thus, PCR efficiency has a major impact on the fluorescence history and the accuracy of the calculated expression result and is critically influenced by PCR reaction components.

Additionally, PCR efficiency during real-time qRT-PCR is influenced by the RNA quality. RNA samples of optimal quality can serve as a template for all product lengths, whereas for degraded RNA primer pairs for shorter amplicon are more suitable. To be on the safe side with primer pairs it would be helpful to prove the RNA quality before starting the run.

Digital Polymerase Chain Reaction (dPCR)

Digital polymerase chain reaction (digital PCR, DigitalPCR, dPCR, or dePCR) is a biotechnology refinement of conventional PCR methods that can be used to directly quantify and clonally amplify nucleic acids, including DNA, cDNA, or RNA. dPCR allows the precise quantification of nucleic acids, facilitating the measurement of small percentage differences and quantification of rare variants. dPCR may also be more reproducible and less susceptible to inhibition than qPCR. The Minimum Information for Publication of Quantitative Digital PCR Experiments (MIQE dPCR) guidelines was published by [Huggett et al. \(2013\)](#) to standardize the experimental protocols. Regarding RNA integrity, reverse transcription dPCR (RT-dPCR) is more complex than simply measuring DNA by dPCR because it requires the additional RT step to convert RNA to cDNA. Consequently, the assumption that DNA measurement by dPCR can be precise, reproducible, and absolute cannot be readily extrapolated to the measurement of RNA, and RNA measurement by RT-qPCR is notoriously variable, depending on experimental design, including RT protocol, RNA secondary structure, and choice of reagents, with a varying degree of sensitivity dependent on the numbers of RNA molecules being converted to cDNA. Thus, RT-dPCR may perform well when measuring the relative amounts of the same RNA from different samples. However, assays targeting different parts of the same molecule may well give different results; it is not recommended that cross-assay comparisons are made unless they are calibrated to control for differences in the generation of cDNA.

Microarrays

Over last 20 years, high-throughput methods have been developed that enable the monitoring of thousands of genes or millions of sequences simultaneously, such as microarrays. Microarrays consist of a series of thousands of microscopic spots of oligonucleotides each containing picomoles of a specific nucleic acid sequence of a part of a gene or other nucleic acid element serving as probes to hybridize with cDNA or cRNA from the analyzed sample. Total RNA is processed to produce fluorescently labeled probes. Fluorescent signal is detected by a scanner, and intensity correlates directly with the abundance of mRNA present in the sample. In addition to mRNA expression analysis, different microarray platforms have been developed for numerous applications. Nucleic acid hybridization is then performed between the labeled molecules and complementary probes immobilized on a solid surface, allowing for the determination of what genes are expressed and their relative or absolute abundance. There are standard microarrays where the probes are attached to a solid surface by a covalent bond, or microscopic beads. Arrays now exist to assay genomic sequence, microRNA abundance, DNA–protein interactions (ChIP on chip), single nucleotide polymorphisms, alternative splicing (exon arrays), and DNA methylation. Many of these applications have more recently been approached using next-generation sequencing.

The Kierzek and Turner groups (2007) developed a new approach to study RNAs with oligonucleotide microarrays. In this approach, the probes are five or six nucleotides long and based on the natural 2'-O-methylated backbone because single-stranded regions within natural RNAs are short and 2'-O-methylRNA/RNA duplexes are more stable than RNA/RNA or DNA/RNA duplexes. Moreover, these probes are modified to make hybridization to unstructured RNA independent of base sequence, i.e., probes are isoenergetic [Kierzek et al. \(2015\)](#). Isoenergetic oligonucleotide microarray mapping is based on interactions of oligonucleotide probes with target RNAs. Thus, hybridizations can be performed over a wide range of buffers, cations, and temperature, time, and pH conditions. The same concerns chemical mapping since to react with target RNA the chemical reagents require conditions that could influence the structure of RNA. For example, some reagents used for chemical mapping require the presence of ethanol or dimethylsulfoxide in the reaction mixture. Furthermore, this microarray mapping is simple and fast. The 850 penta/hexanucleotide probes that form hybridization duplexes thermodynamically stable enough for mapping can be multiply printed on a large number of single microscope slides that can then be used to map any RNA. In addition, some limitations have been described of isoenergetic microarrays as pentanucleotide probes can

have multiple complementary binding sites, particularly in large RNAs. Moreover, U binds to A and G with roughly equal stability. Finally, microarray mapping is particularly useful for RNAs containing up to 300–500 nucleotides. Unfortunately, isoenergetic microarrays are not yet commercially available ([Fig. 6.2](#)).

Sequencing

RNA-Sequencing (RNA-Seq) is a powerful sequencing-based method that enables researchers to break through the inefficiency and expense of legacy technologies, such as qRT-PCR and microarrays. Because the method does not require predesigned probes, the data sets are unbiased, allowing for hypothesis-free experimental design. This type of next-generation sequencing analysis is a powerful tool for transcript and variant discovery studies that are not possible using traditional microarray-based methods. RNA-Seq quantifies individual sequence reads aligned to a reference sequence, which produces discreet (digital) read counts. The digital nature of this process and the ability to control coverage levels supports an extremely broad dynamic range, with absolute rather than relative expression values. Assuming 10–50 million mapped reads, the dynamic range of RNA-Seq spans five orders of magnitude ($>10^5$) and is typically several orders of magnitude higher than most array technologies (10^3). As a result, RNA-Seq has been shown to detect a higher percentage of differentially expressed genes compared to expression arrays, especially genes with low abundance. RNA-Seq provides sensitive, accurate measurement of gene expression at the transcript level; generates both qualitative and quantitative data; captures splice junctions, fusions, coding, and multiple forms of noncoding RNA, such as siRNA, miRNA, snoRNA, and tRNA; covers an extremely broad dynamic range; delivers superior performance with degraded RNA, such as FFPE tissue samples; maintains and tracks strand-specific information in the data; delivers a more powerful method for discovery applications; and finally, scales for large studies and high sample numbers [Wang et al. \(2009\)](#); [Ozsolak and Milos \(2011\)](#); [Wilhelm et al. \(2009\)](#).

In summary, the RNA molecule has been extensively investigated in recent years and different molecular biology techniques have focused on improving RNA extraction methods. Moreover, purity and integrity are essential elements to challenge RNA-based analysis. Additionally, adsorption, isopycnic gradient and chromatographic methods have been developed as new RNA extraction methods. Finally, RNA analysis using microarrays has grown extremely in recent years. RNA-sequencing (RNA-Seq) is a powerful sequencing-based method that enables researchers to break through the inefficiency and expense of legacy technologies, such as qRT-PCR and microarrays.

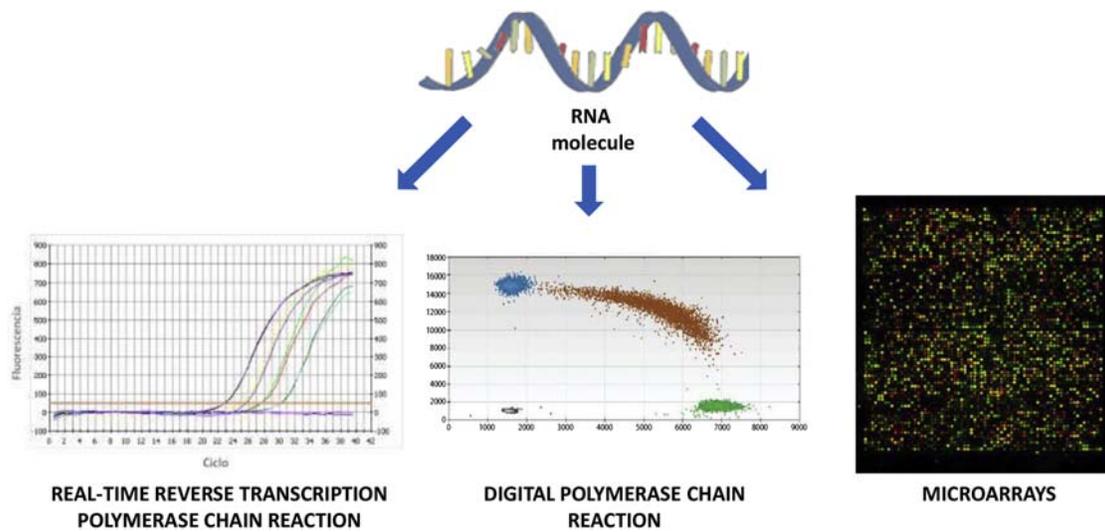


FIGURE 6.2 RNA analyses.

References

- Bustin, S.A., Benes, V., Garson, J.A., Hellems, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55, 611–622.
- Chen, L.L., 2016. The biogenesis and emerging roles of circular RNAs. *Nat Rev Mol Cell Biol* 17, 205–211.
- Fleige, S., Pfaffl, M.W., 2006. RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Asp Med* 27, 126–139.
- Huggett, J.F., Foy, C.A., Benes, V., Emslie, K., Garson, J.A., Haynes, R., Hellems, J., Kubista, M., Mueller, R.D., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., Bustin, S.A., 2013. The digital MIQE guidelines: minimum information for publication of quantitative digital PCR experiments. *Clin Chem* 59, 892–902.
- Kierzek, R., Douglas, H., Turner, D.H., Kierzek, E., 2015. Microarrays for identifying binding sites and probing structure of RNAs. *Nucleic Acids Res* 43, 1–12.
- Krebs, J.E., Goldstein, E.S., Kilpatrick, S.T., 2014. *Lewin's Genes XI*, eleventh ed. Jones & Bartlett Learning, United States of America.
- O'Brien, M.A., Costin, B.N., Miles, M.F., 2012. Using genome-wide expression profiling to define gene networks relevant to the study of complex traits: from RNA integrity to network topology. *Int Rev Neurobiol* 104, 91–133.
- Ozsolak, F., Milos, P.M., 2011. RNA-sequencing: advances, challenges and opportunities. *Nat Rev Genet* 12, 87–98.
- Rodwell, V.W., Bender, D.A., Botham, K.M., Kennely, P.J., Weil, P.A., 2015. *Harper's Illustrated Biochemistry*, thirtieth ed. McGraw Hill, New York.
- Tan, S.C., Yiap, B.C., 2009. DNA, RNA, and protein extraction: the past and the present. *J Biomed Biotechnol* 10, 574398.
- Vomelova, I., Vaničková, Z., Šedo, A., 2009. Methods of RNA purification. All ways (should) lead to Rome. (RNA/mRNA/purification/automation/gene expression). *Folia Biol* 55, 243–251.
- Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10, 57–63.
- Wilhelm, B.T., Landry, J.R., 2009. RNA-seq—quantitative measurement of expression through massively parallel RNA-sequencing. *Methods* 48, 249–257.

Methods for Global Nutrigenomics and Precision Nutrition

Adela Chirita-Emandi¹, Mihai Niculescu^{1,2}

¹Genetics Department, Center of Genomic Medicine, University of Medicine and Pharmacy “Victor Babes”, Timisoara, Romania; ²Advanced Nutrigenomics, Hillsborough, NC, United States

Glossary

Epigenomics investigates mechanisms that alter gene expression, other than changes in the DNA sequence, mainly focusing on DNA methylation, DNA hydroxymethylation, and histone modifications by gene-specific analysis, global methylation/modification content, or genome-wide scans.

Genomics the study of genetic variations (e.g., next-generation sequencing) or changes in chromosomal structure and architecture as a result of alterations in the site–site interaction matrix within a chromosome (e.g., Hi-C, a further development of the chromosome conformation capture method).

Metabolomics the study of the dynamic multiparametric metabolic response (small-molecule metabolite profiles) of living systems to pathophysiological stimuli or genetic modification, or the attempt to identify unknown metabolites, biomarkers, or targets for treatment.

Metagenomics the study of the collective genome of microorganisms from an environmental samples to provide information about the microbial diversity and ecology of a specific environment. Microbial diversity can interact with and alter gene expression, protein synthesis, and posttranslational modifications, and the metabolome.

Proteomics a large-scale study of proteins, particularly their structures and functions, and the interaction between specific proteins with other structures.

Transcriptomics the study of the set of all RNA molecules, including messenger RNA, ribosomal RNA, transfer RNA, and other noncoding RNA produced in one cell or a population of cells.

which allows both researchers and nutrition practitioners to understand individual nutrition requirements better and prescribe personalized treatments (Kohlmeier and Kohlmeier, 2013). Multiple genetic variants were found to influence the metabolism of nutrients; in addition, they were associated with different health outcomes in people who may have similar nutrient intakes but differ in their genotypes. To benefit effectively from the complexity and diversity of nutrigenomic methods and data output (ranging from biomolecular assays to statistical methods), one needs to choose appropriate methods for a given experiment design, but also to ensure that data were interpreted using adequate statistical and bioinformatic tools and methods. Nutrigenomics uses many methods developed within both nutrition and genomic sciences, including high-throughput “omics” technologies such as genomics, transcriptomics, proteomics, metagenomics, metabolomics, and epigenomics (Fig. 7.1).

This chapter introduces the methods most commonly used in nutrigenomic studies and nutrigenomic assessment, spanning from laboratory techniques in the omics fields to statistical aspects to be considered when interpreting data generated by such technologies and platforms.

INTRODUCTION

As a result of the massive development of laboratory and bioinformatic techniques, the generation of genetic-related information has increased tremendously, allowing for an opportunity to explore interindividual genetic differences and their relationship with nutrition-related metabolic variations (Camp and Trujillo, 2014). One consequence has been the development of the concept of precision nutrition,

METHODS SUMMARY

Table 7.1 summarizes several methods that can be used in nutrigenomic research:

- Genomic methods detect genetic variations (e.g., next-generation sequencing [NGS]) or changes in chromosomal structure and architecture as a result of alterations in the site–site interaction matrix within a

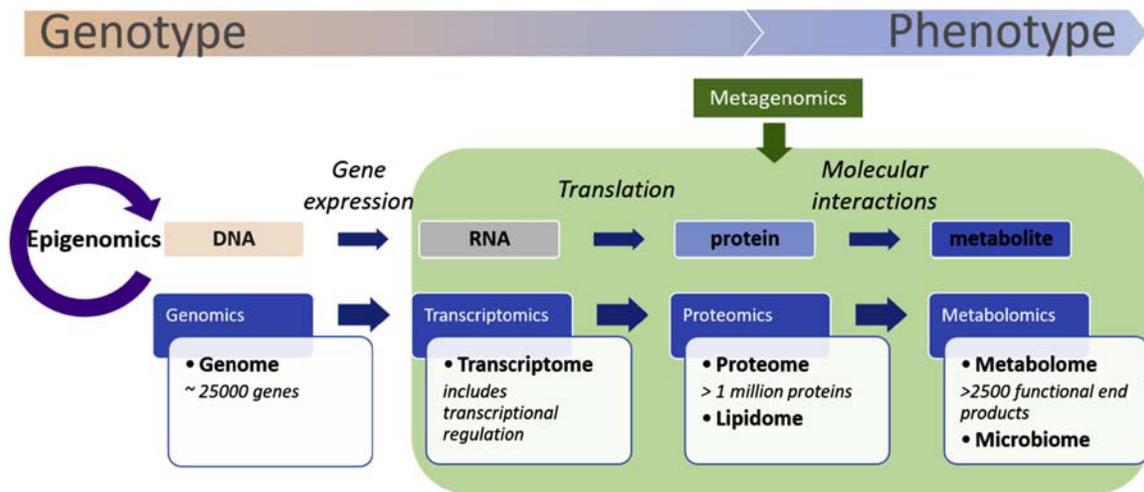


FIGURE 7.1 From genotype to phenotype. Interacting “omics.” A depiction of the relations among genetic information (explored using genomic methods), gene expression (transcriptome), protein expression (proteome), and metabolic consequences (metabolome). In addition, epigenetic regulation is involved in gene expression and influences RNA levels and gene silencing.

chromosome (e.g., Hi-C, a further development of chromosome conformation capture method) (van Berkum et al., 2010).

- Transcriptomic methods assess (either quantitatively or qualitatively) the amount of specific mRNA (gene expression) or the expression of other noncoding RNA species (microRNA [miRNA], other short noncoding RNA, long noncoding RNA, etc.). Two relevant examples are gene expression microarrays and real-time reverse-transcriptase polymerase chain reaction (RT-PCR).
- Proteomic platforms measure the amount of specific proteins or assess the protein structure (e.g., protein sequencing) or the interaction between specific proteins and other structures (e.g., DNA–protein interactions at binding sites).
- Metabolomic approaches investigate the level of known metabolites or attempt to identify unknown metabolites, biomarkers, or targets for treatment.
- Metagenomics studies the collective genome of microorganisms from environmental samples, to provide information about the microbial diversity and ecology of a specific environment. Microbial diversity (including one’s own microbiome) can interact with and alter gene expression, protein synthesis, and posttranslational modifications and the metabolome (including changes in lipid metabolism [lipidomics]).
- Epigenomic approaches investigate mechanisms that alter gene expression, other than changes in the DNA sequence. These methods mainly focus on DNA

methylation, DNA hydroxymethylation, and histone modifications by gene-specific analysis, global methylation/modification content, or genome-wide scans.

Another field of research and assessment refers to the interaction between classes of molecules, such as DNA–protein interactions (e.g., chromatin immunoprecipitation [ChIP]), protein–protein interactions, and so on, which can help in understanding molecular mechanisms involved in the metabolic pathways of nutrients as well as the signaling role of some nutrients.

The last step required for data analysis is a series of assessments that interpret the significance of data generated:

- Statistical methods, enabling identification of the significance of hypothesized changes or of changes for variables that were not considered ab initio;
- The integration of changes within a functional framework at a biological level. Because the ultimate goal of research is to find meaningful and actionable elements, it is essential to be able to interpret the biological data appropriately. Bioinformatics has an essential role; several categories of tools can be used:
 - Tools that help in identifying the placement of an identified change (e.g., alignment tools used in NGS, enabling the correct placement within the genome of sequenced data); and
 - Tools that help in understanding the biologic roles of identified changes (e.g., pathway analysis and gene ontology analysis).

TABLE 7.1 Methods in Nutrigenomics.

Outcome	Methods	Advantages (+) and Limitations (–) of Methods
GENOMICS		
Single nuclear polymorphism (SNP)	Polymerase chain reaction (PCR) (many types)	(+) simple, cheap and widely used method in which minute amounts of DNA are amplified and then quantified. (–) specific primers are needed prior to sequence; errors in primer design produce misleading results; there is a limited sample number, and a single specific region of interest per experiment.
	SNP array	(+) provide information on an extended number of SNPs. (–) expensive and laborious
	Next generation sequencing (NGS)	See below
Gene sequences Exome sequences Genome-wide sequences	NGS platforms Hotspot/shotgun sequencing platforms	(+) high throughput methods to sequence small or large DNA segments (targeted DNA analysis, or whole exome sequencing) or entire genome (whole genome sequencing). (–) high price for equipment and consumables, large amount of data for analysis; issues related to storage and safety of data, issues related to comparison with a “reference genome”.
Chromosomal structure	Karyotyping Fluorescent in situ hybridization (FISH)	(+) qualitative methods that provide information on chromosomal structure at different resolutions (chromosomes in karyotype; specific regions in FISH). (–) karyotype can show only large abnormalities in chromosome structure, while FISH provides information solely related to the region of interest.
Copy number variation and loss of heterozygosity	Microarrays platforms	(+) comparative technique. Information on many genes/all the genes. (–) high price for equipment and consumables, large amount of data for analysis and safe storage, comparison with a “reference genome”.
Chromosomal 3-dimensional (3D) structure	Hi-C platforms	(+) Detection of chromosomal interactions between functional elements, such as promoters and enhancers, using chromosome conformation capture; shows presence of chromosome territories and the preferential association of small gene-rich chromosomes (–) Expensive method with relatively few applications
Protein–genome interactions	Chromatin immunoprecipitation-on chip microarrays	(+) Advanced technology allowing genome-wide analysis of protein of modification of interest distribution (–) Limited resolution compared with sequencing; there may be low quality/affinity of available antibodies for a protein of interest. Relatively costly experiments
TRANSCRIPTOMICS		
Messenger RNA (mRNA), microRNA (miRNA), and other RNA species	RNA sequencing	(+) Deep sequencing technologies provide transcriptome profiling (high-throughput) by sequencing the complementary DNA; provide precise gene expression (–) Expensive; isoforms of the transcriptome are generally indistinguishable from each other (unless the primer pool is designed specifically for this purpose); large RNA molecules need fragmentation before sequencing
	Microarray platforms	(+) Relatively straightforward and available in most laboratories (–) There may be limited range of detection owing to saturation of signals, lack of normalization methods; comparing across experiments is difficult; mRNA abundance is not always a reliable indicator of protein activity
PROTEOMICS		
Protein separation Quantitative expression profiling	2D gel electrophoresis; Fluorescence 2D difference gel electrophoresis	(+) Semiquantitative, can assess posttranslational modifications (–) Poor separation of acidic, basic, hydrophobic, and low-abundant proteins; fluorescent gel electrophoresis requires special equipment for visualization and fluorophores are expensive
Quantitative proteomics Differential expression pattern	Isotope-coded affinity tag; Isobaric tags for relative and absolute quantitation	(+) Sensitive and reproducible, detects peptides with low expression levels, quantitation is straightforward (–) Some proteins are not detected

Continued

TABLE 7.1 Methods in Nutrigenomics.—cont'd

Outcome	Methods	Advantages (+) and Limitations (–) of Methods
Protein-protein interactions	Multidimensional protein identification technology	(+) High separation potential for large protein complex identification (–) Not quantitative; difficulties to identify isoforms
Quantitate specific proteins	Protein array	(+) High-throughput and specific (–) Limited protein levels are more difficult to detect (sensitivity); antibodies against different proteins have different affinity; difficult to compare levels between different proteins
Protein identification and characterization	Mass spectrometry	(+) High sensitivity (for most proteins) and specificity; high-throughput, qualitative and quantitative, shows posttranslational modifications (–) No single method to identify all proteins; not sensitive enough to identify minor or weak spots; matrix-assisted laser desorption/ionization and electrospray ionization do not favor identification of hydrophobic peptides and basic peptides; dedicated equipment (high costs)
METABOLOMICS		
Metabolites	Gas/liquid chromatography (GC/LC) coupled to mass spectrometer (MS) Capillary electrophoresis coupled to MS Nuclear magnetic resonance spectroscopy	(+) Semiquantitative high-throughput methods that provide targeted or untargeted analysis of metabolome. Can be designed as quantitative with appropriate standards and internal controls (–) Samples require fractionation before analysis; reference values are scarce for general population and even more for certain groups such as male/female/pediatric classifications; the metabolome is highly dynamic, time-dependent, and metabolites are sensitive to many environmental changes; metabolites are chemically diverse; data interpretation requires skillful bioinformatics; integration of metabolomic profiles with molecular data from other high-throughput technologies is crucial and difficult; need dedicated and costly equipment, dedicated operator, sometimes expensive standards
METAGENOMICS		
Microbial ecology	Amplicon sequencing of the 16S or 18S ribosomal RNA gene or internal transcribed spacer (ITS)	(+) Fast and cost-effective identification of a wide variety of bacteria and eukaryotes (–) Does not capture gene content other than the targeted gene. Amplification bias
Functional genome analyses	Random shotgun sequencing of DNA or RNA	(+) No amplification bias, detects bacteria, archaea, viruses, and eukaryotes. Enables de novo assembly of genomes (–) Many reads may be from host. Requires reference genomes for classification
Transcriptional profiling of what is active	Sequencing of mRNA NGS	(+) Identifies active genes and pathways (–) mRNA is unstable. Multiple purification and amplification steps can lead to more noise
EPIGENOMICS		
Gene-specific DNA methylation	Bisulfite-pyrosequencing Methylation specific PCR Methylation specific real-time PCR	(+) Are cheap and straightforward techniques. When designed appropriately, can provide single-site resolution (pyrosequencing and several PCR versions) (–) Qualitative/semiquantitative methods. Low throughput
Global DNA methylation	High-performance liquid chromatography (HPLC), GC/MS Immunostaining with anti-5mC antibodies Digestion with methylation-sensitive enzymes	(+) Versatile and relatively low costs (except for HPLC-GC/MS) (–) HPLC-GC/MS platforms: Need dedicated and costly equipment, dedicated operator, sometimes expensive standards
Genome-wide scans for DNA methylation	Microarrays NGS platforms	(+) Accurate, high throughput (–) Expensive and laborious; data interpretation requires skillful bioinformatics

TABLE 7.1 Methods in Nutrigenomics.—cont'd

Outcome	Methods	Advantages (+) and Limitations (–) of Methods
Histone modifications	Western blot, immunohistochemistry, immunocytochemistry, enzyme-linked immunosorbent assay Chromatin immunoprecipitation, ChIP-on-chip	(+) Reliable, semiquantitative methods, fast output (–) Sometimes there are issues related to specificity and affinity of antigen binding sites; depending on assay, sometimes expensive experiments
miRNA	Microarray profiling Quantitative PCR assays	(+) High-throughput, high-resolution NGS of small RNAs can quickly and accurately discover new miRs and confirm the presence of known miRs (–) Requires 0.2–2 µg RNA. Potential cross-hybridization of related miRs. Cannot measure low uptakes (+) reliable, high specificity, used for confirmation. (–) Primer for cDNA is based on complementarity short sequences at the 3' end. More prone to external contamination due to amplification

GENOMICS

Genomics employs several methods to identify genetic variations, from single nuclear polymorphisms (e.g., TaqMan SNP assays) to more comprehensive methods of sequencing an entire gene (or a DNA target of defined length), a panel of genes, the whole exome, or the whole genome (e.g., microarrays, microchips, NGS) (Park and Kim, 2016).

Small Variations in DNA Sequence

Most nutrigenetic studies focused on identifying associations between genetic variations and an individual's nutritional needs in different physiological or pathological states. Small and localized variations in DNA sequence, such as SNPs or short insertions/deletions may be the most well-studied categories of genetic variations associated with nutrition. SNPs are evaluated using PCR, DNA arrays, or sequencing.

PCR is a simple, yet powerful and versatile biochemical method for amplifying specific segments of DNA. PCR uses a DNA polymerase to guide the synthesis of DNA using deoxynucleotide substrates on a single-stranded DNA template. The polymerase adds nucleotides to the 3' end of a custom-designed oligonucleotide when it is annealed to a longer template DNA. Consequently, if a synthetic oligonucleotide is annealed to a single-stranded template that contains a region complementary to the oligonucleotide, DNA polymerase will use the oligonucleotide as a primer and elongate its 3' end to generate an extended region of double-stranded DNA.

In the past, the nutrigenetic approaches were less complex, focusing on one SNP at a time, a single dietary component, and single disease risk factors. Currently,

the challenges are to identify functional genetic variations, define the interplay among several genes, and understand gene–nutrient interactions. Therefore, several other techniques were developed to meet these purposes.

Microarrays are used in genomic, transcriptomic, and proteomic research and provide increased throughput. Microarrays consist of DNA or protein fragments (probes) placed as small spots onto a slide (miniaturized chemical reaction areas). The studies typically involve looking for changes in gene structure, gene expression, or protein expression patterns using selected cells or tissues under different experimental conditions. Microarrays provide a platform for evaluating changes in thousands of genes or proteins simultaneously.

Technological advances have also provided high-throughput, reliable, and quick array-based genotyping platforms. SNP array platforms were successfully developed and used to genotype several thousand genes in one experiment.

DNA sequencing is the fastest developing area for assessing genetic variations. Its versatility enables an investigation of variations in DNA sequences, changes in gene expression, or the presence of copy number variations (CNVs). The short time required to run the method, the capability of high customization, and the low costs per sequenced nucleotide make the newest sequencing methods (NGS) the most efficient and preferred method in both research and diagnostic laboratories.

Haplotype Analysis

The possibility of sequencing the entire genome enabled the identification of associations between polymorphisms that are contiguous on the same

chromosome and relatively close (haplotypes). Multiple genetic variations can present in specific combinations, with a higher probability when such variations are closer on a chromosome, owing to the crossing-over phenomenon that occurs during meiosis. Although it is a recent area of research, the study of haplotypes to define nutritional requirements might be more promising than using individual polymorphisms (see earlier discussion) because it allows for an analysis of several pooled polymorphisms and the collecting of individuals into fewer groups for defining nutrient requirements.

Copy Number Variations

Variations in the number of alleles (CNVs) is a recently developed field of study in nutrigenomics; therefore, there are fewer relevant studies. However, identifying variations in the number of alleles had resulted in a better understanding of the role of the nutritional environment in the evolution of human populations owing to the multiplication of segments of DNA, because these populations may benefit from a selective advantage in a particular environment. CNVs can be assessed using DNA microarrays, comparative genomic hybridization techniques, PCR, or DNA sequencing. Some CNVs have no apparent influence on phenotype, whereas others have been definitively linked with health outcomes. Evidence also indicated that interactions with additional genetic or environmental factors may influence whether CNVs have a detectable phenotypic effect.

Gene–Gene Interactions

Because the metabolic fate for each nutrient is controlled by multiple genes, gene–gene interactions should be considered when defining nutritional needs. Similar to haplotype analysis, gene–gene interactions analyze a health outcome in relation to a combination of genotypes not necessarily grouped in haplotypes. Such combined genetic variations can encompass genes that are not in close proximity with each other and are the subject of multivariate statistical analyses. The advantage of this approach is that it better defines the roles of multiple genetic variations in the context of their interacting roles upon a given nutrient. Therefore, nutritional requirements can be even further refined; however, they require complex statistical methods for a valid interpretation.

Gene–Protein Interactions

ChIP-Chip is a genome-wide location analysis that combines ChIP and DNA microarray analysis to

identify, protein–DNA interactions that occur in living cells. The process consists of copurification of DNA fragments associated with a specific protein (by chemical cross-linking), and includes cell lysis, DNA fragmentation, and immunoaffinity-based purification of the desired protein. The enriched DNA population is labeled, combined with a differentially labeled reference sample, and applied on DNA microarrays to detect enriched signals. Alternatively, the enriched DNA fragments can be identified and quantified by PCR or even DNA sequencing. Furthermore, the analysis uses computational and bioinformatic approaches to normalize the enriched and reference channels, to connect signals to the portions of the genome that are represented on the DNA microarrays and provide confidence metrics and generate maps of protein–genome occupancy.

Limitations of Genomic Assessment in Nutrition

Factors that limit the accuracy of determining individual nutritional needs related to genetic variations consist of the limitations of current techniques for sequencing and genotyping. They also depend on the proper application of statistical and bioinformatical methods. Furthermore, nutrigenetic data are obtained through epidemiological studies in which participants are most often located in defined regions and ecosystems and sometimes belong to defined ethnic structures and certain age groups and have different socioeconomic and lifestyle factors. It is therefore difficult to generalize the data obtained to other populations. There are inherent limits to sequencing and genotyping techniques that can generate misleading results for some genotypes. Therefore, it is important to recognize the likelihood of obtaining incorrect results. Validation is required using different methods. The statistical methods used should be an appropriate multivariate analysis, to assess the false discovery rate and the specificity and sensitivity of the employed method.

TRANSCRIPTOMICS

Transcriptomics reports data obtained using RNA pools within a biological system (e.g., body fluids, cells, or tissues). Genome-wide monitoring of gene expression enables the assessment of transcription of thousands of genes in cells and tissues in relation to their exposure to different nutrients or their metabolites. Cellular functions are regulated by gene expression (synthesis of messenger RNA [mRNA]). miRNA and other noncoding RNA species are recently discovered classes of

RNAs that do not code for proteins; instead, they participate in various processes that regulate gene expression, mRNA stability, and epigenetic regulation. For instance, cells use miRNA to regulate the amount of protein synthesized by a gene through translational inhibition and mRNA stability.

For **gene expression microarrays**, the sample used to assess the amount of mRNA (or other RNA species) can be complementary DNA (cDNA) or an oligonucleotide. The sample is amplified by PCR; then, a fluorescent dye is added and it is hybridized on a microarray slide. These microarray glass slides are often called chips. The amount of the two dyes, represented by the intensity of each fluorophore (one for the sample and another for a reference) represents the gene expression in different samples (this is a comparative method). A laser is used to scan the slide, which detects the amount of fluorescent signal for each gene, and the image is then analyzed.

Gene expression can also be assessed using **NGS** and **real-time RT-PCR**.

PROTEOMICS

Proteomics uses mass spectrometry methods for the investigation and identification of analytes. Ions are created from neutral proteins or peptides, which are then separated according to their mass-to-charge ratio and detected based on their mass spectrum using known reference molecules.

There are several analytical techniques in mass spectrometry, each of which has different advantages and limitations in terms of instrument sensitivity, resolution, mass accuracy, dynamic range, and throughput. A number of techniques are frequently used to analyze the proteome or metabolome. In proteomics, electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) have improved protein characterization, whereas issues have arisen regarding the reproducibility, accuracy, mass range, and dynamic range of another technique, surface-enhanced laser desorption/ionization. Differential image gel electrophoresis has improved quantitative analysis. This technique uses fluorescent tags in a gel-based platform, and strategies of isotopically coded affinity tag labeling are used along with tandem mass spectrometry.

Perhaps the most important issues in a proteomic experiment are protein concentration, sample purification, and protein digestion, in addition to affinity capture and sample fractionation (using gel-based or chromatography techniques) needed to reduce the complexity of the specific fluid to be analyzed. For example, urine provides different analytical challenges compared with plasma, including the need for preparation techniques such as ultrafiltration and precipitation,

which are necessary to remove salts and concentrate urinary proteins. Dissimilarities can arise for each fluid, which adds to interindividual variation. Therefore, the setup stage and appropriate choice of detection platform are essential for the success of a research project aiming at identifying proteomic alterations.

METABOLOMICS

Metabolomics offers a framework for analyzing specific metabolites in a given biological sample. The range of molecular mass for different metabolites can be tens to thousands of daltons, whereas their concentrations vary tremendously between and within tissues. The metabolome is highly dynamic and time-dependent, and the storage of biological samples for metabolic assessment is sensitive to many environmental factors. Metabolites are extensively exchanged with the environment (e.g., food intake, excretion, inhalation, secondary metabolites such as medications, flavorings, and recreational drugs, which can be further processed by the gut microbiome or organs). Furthermore, metabolites are diverse in chemical structure (polarity, charge, dissociation constant, solubility, volatility, stability, and reactivity). Therefore, no single method can capture and analyze the entire metabolome at once. Many extraction methods have been developed to identify and quantify specific groups of metabolites. The most commonly used methods to analyze the metabolome are gas chromatography (GC), liquid chromatography (LC) or capillary electrophoresis, coupled with a mass spectrometer (MS), as well as nuclear magnetic resonance spectroscopy (Garcia-Canas et al., 2010).

Mass spectrometry is an analytical method that acquires data in the form of the mass-to-charge ratio (m/z) and relative intensity of the measured compounds. The chromatographic separation technique is based on the interaction of different metabolites in the sample with the adsorbent materials inside the chromatographic column. The retention time is used together with the m/z MS values to generate the two axes of the LC-MS and GC-MS spectral data (Alonso et al., 2015).

Samples used in metabolomics require fractionation (usually chromatography or electrophoresis) before analysis. Fractionation techniques use the different chemical and physical properties of different molecules and enable the separation of proteins, peptides, and metabolites in liquid or gas phase.

Nuclear magnetic resonance spectroscopy is a fast, reproducible spectroscopic technique based on energy absorption and re-emission of the atom nuclei. The results (spectral data) enable the quantification of concentrations of metabolites and the identification of their

chemical structure. This technique sometimes has low sensitivity and can reliably detect and quantify metabolites present in only relatively high concentrations. Spectroscopy also has limited sensitivity and a lower ability to identify metabolites in complex samples. Not all omic techniques can be interpreted with the same reliability and each analytical technique offers different advantages and limitations (Horgan and Kenny, 2011).

EPIGENOMICS

Epigenetics refers to mechanisms that persistently alter gene expression without actual changes to the DNA sequence. Known epigenetic mechanisms are DNA methylation, chromatin modification, posttranslational modifications of histones, and insertion of transposons. In addition, the epigenetic control of gene expression can be regulated by miRNA and other post-transcriptional mechanisms. These mechanisms have been associated with numerous epigenetic processes, resulting in alterations of the cellular phenotype, such as in carcinogenesis, cell differentiation, inflammation, cell apoptosis, and angiogenesis. The effects of nutrients on gene expression depend on the quantity, bioavailability, exposure time, and chemical structure, and their metabolic fate. Some molecules in foods are also natural epigenetic modifiers or provide methyl donors required for the methylation of DNA and histones.

The investigation of **posttranslational modifications of histones** is usually performed for a limited number of targets or across the whole genome. Antibodies against specific proteins or protein modifications are available for use in western blot, immunohistochemistry, immunocytochemistry, and enzyme-linked immunosorbent assay techniques. Histone posttranslational modifications (including methylation, phosphorylation, acetylation, ubiquitylation, and sumoylation) that can affect gene expression by changing chromatin structure or recruiting histone modifiers can be also quantified with specific assays using colorimetric or fluorometric platforms. To identify the genomic localization of histone modifications, the most commonly used method is ChIP. ChIP uses antibodies to isolate a protein or a modified protein of interest, together with any bound DNA (see also **ChIP-Chip** discussed earlier). This can then be used to identify where the protein alterations (histones in this case) are located within the genome and to determine the relative abundance of proteins for each location. Therefore, the assessment of histone modifications can cover gene-specific analysis or the entire genome using genome-wide scans.

DNA methylation techniques are classified as (1) gene-specific analysis (qualitative or quantitative), (2) an assessment of global methylation content (shows the total methylation in a DNA sample regardless of its distribution within the genome), and (3) high coverage methods such as genome-wide scans (arrays). Gene-specific analysis by bisulfite-pyrosequencing provides an overview of the methylation pattern; it can also identify the methylation percentage of individual single-stranded linear sequences. Moreover, it provides flexibility in positioning the sequencing primer to analyze any single-stranded linear sequence of interest. Another method for analyzing DNA methylation is methylation-specific PCR, which is inexpensive and easy to perform but is qualitative/semiquantitative and does not provide single-site resolution. Real-time PCR is easily accessible, but it has low precision and does not provide single-site resolution. MALDI-TOF is a quantitative method; it provides single-site resolution over extended sequences, but it has dedicated equipment (implying high costs) (Tollefsbol, 2011). Methylation-sensitive high-resolution melting is a method used to discriminate between DNA strands that are differentially methylated. This is initially achieved employing real-time PCR using bisulfite-treated DNA. The differences in methylation status can be obtained at the end of the reaction, based on differences in melting point temperatures among different samples. However, the PCR products can be subject to further sequencing to refine and better characterize changes in DNA methylation at specific loci.

An analysis of global methylation is obtained by directly measuring the methyl group content with high-performance LC, GC/MS, immunostaining with anti-5mC, or digestion with methylation-sensitive enzymes.

Genome-wide scans use either microarrays or NGS platforms. Examples of microarrays platforms are Illumina Infinium for DNA methylation using bisulfite-treated DNA and Nimblegen arrays for DNA methylation or histone modifications (antibody [Ab]-based). NGS employs various platforms using bisulfite (DNA methylation) or Ab-based (DNA methylation or histone modifications) treatment.

Other methods can be used to detect hydroxymethylated DNA (Ab-based).

Nutrimiromics studies the influence of the diet on the modification of gene expression caused by epigenetic processes related to miRNAs, in their relation to risk for the development of chronic diseases. miRNAs are a class of noncoding endogenous RNA molecules, involved in posttranscriptional gene silencing by inducing mRNA degradation or translational repression

by binding to a target mRNA. Methods commonly used are microarray profiling and sensitive and ligation-independent quantitative PCR assays (Quintanilha et al., 2017).

BIOINFORMATICS

Nutrigenomic techniques are generally high-throughput and analyze large numbers of genes, gene expression profiles, or proteins and molecules, in a single procedure or a combination of procedures. Bioinformatics is a key requirement for studying vast amounts of data generated from extremely complex biological samples in large numbers and with high sensitivity and specificity. Next-generation analytical tools require improved robustness, flexibility, and cost efficiency. The process of bioinformatic analysis includes data processing and molecule identification, statistical data analysis, pathway analysis, and data modeling in a system-wide context. Several categories of tools can be used:

- tools that help in identifying the placement of a detected change (e.g., alignment tools used in NGS, enabling the correct placement of sequenced data within the genome); and
- tools that help in understanding the biologic roles of identified changes (e.g., pathway analysis and gene ontology analysis).

There have been efforts to create computational tools to integrate omic data. These need to address three different aspects:

- i. the need to identify the variables measured correctly. For instance, sequencing data requires proper identification and localization throughout the genome of the sequenced DNA strands. Another example is correct identification of the molecular structure from mass spectrometry data;
- ii. to identify the network of changes by delineating connections that exist among cellular components; to decompose it to understand the overall network structure; and
- iii. to develop cellular or system models that simulate and predict particular phenotypes (at the cellular level or in an entire living organism).

There are significant challenges to analyzing modern postgenomic datasets:

- Many technological platforms, both hardware and software, are available for several omic data types, but some of these are prone to introducing technical artifacts.

- Standardized data representations are not always adopted, which complicates cross-experiment comparisons.
- Data quality, context, and variations among laboratories represent other important challenges (Mayer, 2011).

STUDY DESIGN IN NUTRIGENOMICS

There are different methods for answering diverse hypothesized questions. One needs to balance the number of samples versus coverage and account for price, and, in the case of epigenetic assessment, gene expression, proteomics, and metabolomics, acknowledge that such measurements are tissue-specific and might represent other tissues or organs. In addition, the adequate use of bioinformatic methods is critical to integrate data among different platforms.

Some critical questions related to study design are:

- Which tissue is the most appropriate for the research question?
- Are there existing cohorts or studies that have already collected samples? What types of samples are collected, and what are the rules for their use? What information can be obtained?
- How stable are the epigenetic marks within samples over time?
- Has data analysis accounted for cell type effects and confounders?
- How can too many false-positive results (as well as false-negative ones) be avoided?
- Which design responds best to the statistical need to identify true significance thresholds when dealing with multiple variables (e.g., false discovery rates, or other statistical methods controlling for multivariate analysis)?

CONCLUSION

Nutrigenomic assessments need a systems biology approach that uses a number of the high-throughput omics platforms to identify nutrient-sensitive metabolic processes. Each analytic platform (genomics, proteomics, transcriptomics, metabolomics, or epigenomics) provides essential and complementary information that needs to be integrated appropriately to produce a better understanding of the mechanisms underlying diet and health.

Acknowledgments

The work was funded, in part, by POSCCE Project ID: 1854, Cod SMIS: 48749, Contract 677/09.04.2015, and by POC Project Nutrigen, SMIS: 104852, Contract 91/09.09.2016, ID P_37-684.

References

- Alonso, A., Marsal, S., Julià, A., 2015. Analytical methods in untargeted metabolomics: state of the art in 2015. *Front Bioeng Biotechnol* 3.
- Camp, K.M., Trujillo, E., 2014. Position of the academy of nutrition and dietetics: nutritional genomics. *J Acad Nutr Diet* 114, 299–312.
- Garcia-Canas, V., Simo, C., Leon, C., Cifuentes, A., 2010. Advances in Nutrigenomics research: novel and future analytical approaches to investigate the biological activity of natural compounds and food functions. *J Pharm Biomed Anal* 51, 290–304.
- Horgan, R.P., Kenny, L.C., 2011. 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. *Obstet Gynaecol* 13, 189–195.
- Kohlmeier, M., Kohlmeier, G.Z., 2013. *Nutrigenetics: Applying the Science of Personal Nutrition*, first ed. Elsevier/Academic Press, Amsterdam, Boston, p. 1. online resource (xi, 384 pp.) ill.
- Mayer, B., 2011. *Bioinformatics for Omics Data*. Humana Press in Springer Science, pp. 379–547.
- Park, S.T., Kim, J., 2016. Trends in next-generation sequencing and a new era for whole genome sequencing. *Int Neurolog J* 20, S76–S83.
- Quintanilha, B.J., Reis, B.Z., Duarte, G.B.S., Cozzolino, S.M.F., Rogero, M.M., 2017. Nutrimiromics: role of microRNAs and nutrition in modulating inflammation and chronic diseases. *Nutrients* 9.
- Tollefsbol, T., 2011. *Epigenetics Protocols*, second ed. Humana Press.
- van Berkum, N.L., Lieberman-Aiden, E., Williams, L., Imakaev, M., Gnirke, A., Mirny, L.A., Dekker, J., Lander, E.S., 2010. Hi-C: a method to study the three-dimensional architecture of genomes. *J Vis Exp* 39 e1869:1-7.

Epigenetic Analyses Tools for Nutrition Research

Maria L. Mansego¹, Fermín I. Milagro^{2,3}, J. Alfredo Martínez²

¹Bioinformatic Department, Making Genetics S.L., Noain, Spain; ²Department of Nutrition, Food Science, and Physiology, Center for Nutrition Research, University of Navarra, Pamplona, Spain; ³CIBERObn, Physiopathology of Obesity, Carlos III Institute, Madrid, Spain

Glossary Terms and Keywords

5-carboxylcytosine (5-caC) The final oxidized derivative of 5-methylcytosine.

5-formylcytosine (5-fC) The oxidized derivatives of 5-methylcytosine demethylation is a rare base found in mammalian DNA, thought to be involved in active DNA demethylation.

5-hydroxymethylcytosine (5-hmC) A DNA pyrimidine nitrogen base. It is formed from the DNA base cytosine by adding a methyl group and then a hydroxy group.

5-methylcytosine (5-mC) A methylated form of the DNA base cytosine that may be involved in regulating gene transcription.

Alu element A short stretch of DNA originally characterized by the action of the *Arthrobacter luteus* restriction endonuclease.

ChIP Abbreviation for chromatin immunoprecipitation.

Chromatin A complex of DNA and proteins within the nucleus of mammalian cells.

Chromosome The organized, condensed, and packaged structure of DNA and histone proteins located in the nucleus.

CpG Island Location within a DNA sequence (greater than 200 base pairs in length) that contains a high level of CpG sites (>50%), sometimes located consecutively.

CpG site Location within a DNA sequence in which a cytosine and guanine appear consecutively.

DNA methylation Addition of a methyl (CH₃) group from S-adenosylmethionine to a cytosine nucleotide or lysine or arginine residue.

DNMT1 (DNA methyltransferase) Isoforms methylate hemimethylated CpGs, maintaining methylation patterns during DNA replication.

Epigenetics Study of heritable changes in phenotype that do not involve changes in the underlying DNA sequence.

Heterochromatin Tightly packed form of chromatin not actively being transcribed or that is less accessible for transcription.

Hypermethylation Increase in normal methylation levels.

Hypomethylation Decrease in normal methylation levels.

LINE-1 Transposable element in the DNA of some organisms that belongs to the group of long interspersed nuclear elements.

MBD Abbreviation for methyl-CpG binding domain protein.

MeDIP Abbreviation for methylated DNA immunoprecipitation.

MicroRNA Small noncoding RNA (about 22 nucleotides long) that functions in RNA silencing and posttranscriptional regulation of gene expression.

MS-HRM Abbreviation for methylation-sensitive high-resolution melt analysis.

MS-MLPA Abbreviation for methylation-specific multiplex ligation-dependent probe amplification.

N6-Methyladenosine (6-mA) Methylated form of the DNA base adenine.

NGS Abbreviation for next-generation sequencing.

Noncoding RNA RNA molecule that is not translated into a protein.

Nucleosome The DNA-histone complex, which consists of 146 base pairs of double-stranded DNA wrapped around eight histone proteins.

Pyrosequencing Method of DNA sequencing (determining the order of nucleotides in DNA) based on the “sequencing by synthesis” principle.

RRBS Abbreviation for reduced representation bisulfite sequencing.

TET (Ten-eleven translocation) enzymes Enzymes involved in methyl group oxidation with the production of 5-hydroxymethylation as an intermediate.

WGBS Abbreviation for whole-genome bisulfite sequencing.

Nutrition contributes to modify the gene expression profile by regulating transcriptional and posttranscriptional mechanisms. Moreover, nutrients may induce temporary or permanent changes in the epigenetic mechanisms that modulate the expression of genes involved in metabolic processes and other biological networks, which could be factors leading to the development of pathologies such as obesity, type 2 diabetes, cancer, and neurological and cardiovascular diseases.

The term “epigenetics” refers to heritable states of gene expression resulting from a set of reversible modifications in chromatin structure without alterations in the DNA sequence. These changes may be induced in response to environmental or endogenous factors.

They include several components of epigenetic marks, such as nucleic acid modification, histone modifications, or noncoding RNAs (Fig. 8.1). Regarding nucleic acid changes, DNA methylation, mainly of cytosine nucleotide at the carbon 55 position (5-mC), is a common epigenetic mechanism involved in regulation of gene expression of many eukaryotes and is often found in the sequence context CpG (cytosine followed by a guanine). DNA methylation located in the promoter region of coding genes is typically associated with gene overexpression and underexpression, whereas genic methylation usually correlates with transcriptional activity and chromatin accessibility. Nonetheless, most variably methylated regions are located more often in gene bodies and intergenic regions rather than in promoters and upstream regulatory areas; the underlying mechanisms are still unclear.

The development of new technologies such as next-generation sequencing (NGS) has enabled the detection of deep epigenetic modifications belonging to the methylation cycle, such as 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC), and 5-carboxycytosine (5-caC) in genomic DNA or N(6)-methyladenosine (6-mA) in RNA (Fig. 8.2). Moreover, sensitive techniques have been developed to analyze different conditions and types of samples (biopsies, low amounts of input sample DNA, formalin-fixed paraffin-embedded samples (FFPE), whole peripheral blood, urine, saliva, plasma, or serum) (Mansego et al., 2013).

DNA methylation analysis technologies can be suitable from a global determination to a single-nucleotide

resolution and pass through an intermediate stage as methylation levels at specific loci in DNA. These techniques include methylation-specific or dependent restriction enzyme polymerase chain reaction (PCR) and genome-wide DNA methylation profile through microarray analysis or NGS-based methods.

METHODS FOR NUCLEOTIDE METHYLATION ANALYSIS

There are several techniques for studying DNA or RNA methylation. In this chapter, we describe a series of methodologies to quantify mainly 5-mC levels, as well as some other nucleotide modifications such as 5-hmC. All of these approaches are based on one of three main pretreatment techniques: chemical modification, affinity enrichment, and enzyme restriction digestion (Table 8.1).

A scheme of the various techniques that can be used to quantify methylation according to the pretreatment the DNA receives, the number of modified nucleotides, and the resolution of the detection (global, specific region, or genome-wide methylation) is shown in Table 8.2.

During bisulfite conversion, unmethylated cytosines in single-stranded DNA are deaminated to produce uracil while leaving methylated cytosine intact. Sodium bisulfite treatment is the most conventional approach and is considered the standard for quantifying DNA methylation levels because it could offer

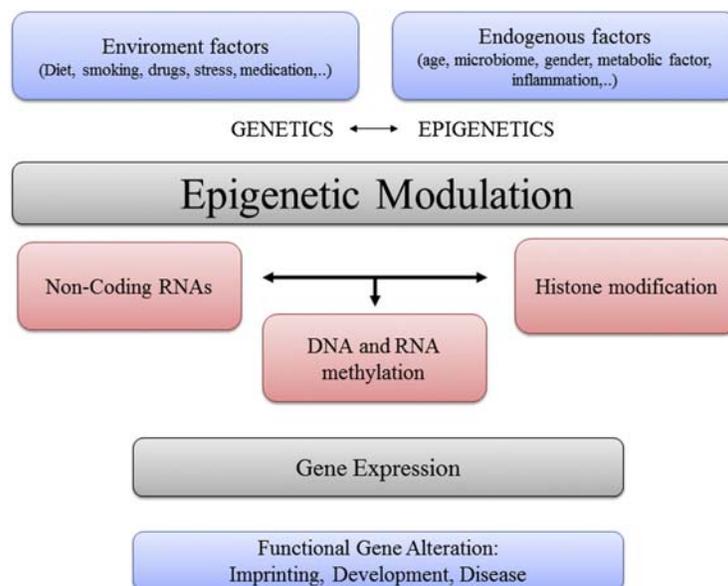


FIGURE 8.1 The epigenetic mechanisms. Many environmental factors such as diet, smoking, and stress, and some internal factors such as inflammation or age may affect gene regulation, which leads to epigenetic modification in the genome. Epigenetic modifications involve DNA methylation, histone modification, and gene regulation by noncoding RNAs. These mechanisms modulate gene expression and affect various gene functions.

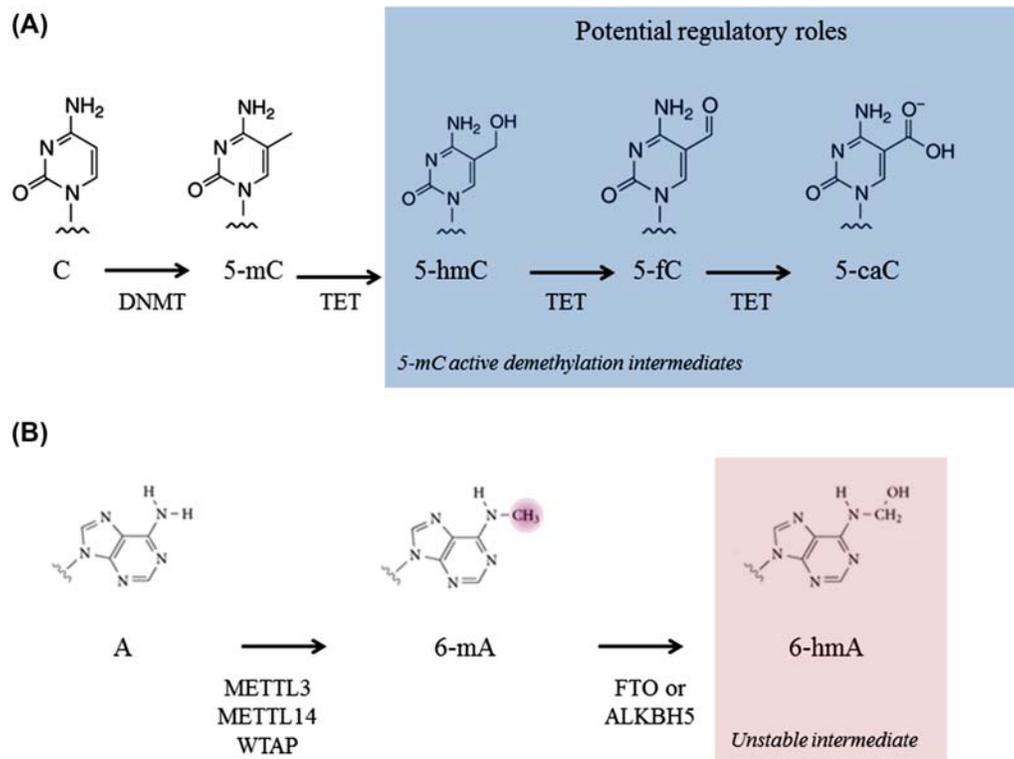


FIGURE 8.2 The methylation cycle of (A) DNA and (B) RNA. 5-*caC*, 5-carboxycytosine; 5-*fC*, 5-formylcytosine; 5-*hmC*, 5-hydroxymethylcytosine; 5-*mC*, 5-methylcytosine; 6-*hmA*, N6-hydroxymethyladenosine; 6-*mA*, N6-methyladenosine; A, adenosine; *ALKBH5*, alkB homolog 5, RNA demethylase; C, cytosine; *DNMT*, DNA methyltransferase; *FTO*, fat mass and obesity associated; *METTL3*, N6-adenosine-methyltransferase; *METTL14*, methyltransferase-like 14; *TET*, ten-eleven translocation; *WTAP*, Wilms tumor 1-associated protein.

single-CpG resolution. However, this treatment has some disadvantages: (1) it typically calls for inputting a large amount of sample DNA, which can degrade after chemical treatment; (2) it can be limited by incomplete conversion of all unmethylated cytosines to uracils; and (3) sodium bisulfite conversion is unable to discriminate between 5-*mC* and 5-*hmC*. Moreover, affinity enrichment treatment is the simplest way to enrich methylated DNA, which was first proved with the methyl-binding protein MECP2. Enrichment-based methods are based on the immunoprecipitation of DNA or RNA with specific antibodies for methylated cytosine to map to 6-*mA* methylated RNA. An important point regarding affinity-based techniques is that they measure the density of methylation in a specific region and the resolution depends on the fragment size of the enriched methylated DNA (usually 300 base pairs [bp]). Therefore, affinity enrichment treatments return a qualitative analysis of methylation. Finally, as these techniques are based on restriction enzyme digestion, they are confined to recognition elements and can only interrogate a subset of methylation sites.

Different combinations of pretreatment methods and subsequent molecular biology methodologies, such as microarrays and NGS, generate a plethora of feasible

techniques for mapping DNA methylation on a genome-wide scale.

GLOBAL QUANTIFICATION OF DNA METHYLATION

Several techniques can be applied to detect the total 5-*mC* content in the genome, including some that also quantify other RNA and DNA modifications. The nucleic acid can be digested into single nucleotides and total genomic 5-*mC* can be quantitated by high-performance liquid chromatography or liquid chromatography/mass spectroscopy, but these techniques are laborious. To bypass these problems, PCR-based methods have been developed to estimate global level of DNA methylation. The detection of repetitive DNA elements such as Alu elements and long interspersed nuclear elements (i.e., LINE-1) is used to assess the DNA methylation level. Interestingly, the average of the relative amounts of C in the four CpG sites of LINE-1 using methylation-sensitive high-resolution melt analysis (MS-HRM) technology is positively associated with a healthier lifestyle (Marques-Rocha et al., 2016). In addition, there is widespread interest in finding correlations between the

TABLE 8.1 Characteristics of Pretreatments for Nucleic Acid Modification Analysis.

Treatment	Description
Chemical modification	<p>Bisulfite conversion of unmethylated cytosine to uracil. Methylated cytosines are protected from this conversion, enabling a determination of DNA methylation. Also, combined treatment with potassium perchlorate discriminates between cytosine nucleotide at the carbon 5 position (5-mC) and 5-hydroxymethylcytosine (5-hmC) by selective oxidation of the latter.</p> <p><i>Pros</i> Reference standard. Single nucleotide resolution. Quantitative analysis.</p> <p><i>Cons</i> Incomplete conversion. Degradation of DNA during bisulfite treatment. Primer design is more difficult.</p>
Affinity enrichment	<p>Enrichment of (hydroxy)methylated DNA using specific antibodies (hydroxymethylated MeDIP or MeDIP) or proteins (MBD) that specifically bind methylated CpG sites in fragmented genomic DNA.</p> <ul style="list-style-type: none"> - MeDIP is based on the affinity purification of methylated DNA using an antibody directed against 5-(hydroxy) mC. - MBD is based on the very high affinity of an h6-GST-MBD fusion protein for methylated DNA. <p><i>Pros</i> Fast, cheap, works easily on a large scale. Differentiation between 5-mC and 5-hmC. Maps 6-mA methylated RNA.</p> <p><i>Cons</i> Resolution depends on the fragment size (300 base pairs). Qualitative analysis. Antibodies not specific to target will precipitate nonspecific DNA or RNA modification</p>
Restriction enzyme-based digestion	<p>Treatment with (hydroxy)methylation-sensitive or dependent restriction enzymes for DNA methylation analysis at specific sites.</p> <p><i>Pros</i> Easy to use</p> <p><i>Cons</i> Determination of methylation status is limited by the enzyme recognition site.</p>

5-hmC, 5-hydroxymethylcytosine; 5-mC, 5-methylcytosine; 6-mA, N6-methyladenosine; MBD, methyl-binding domain; MeDIP, methylated DNA immunoprecipitation.

genomic 5-mC levels and diet, lifestyle, and clinical disorders. In nutritional studies, it would be important to quantify global patterns of DNA methylation to evaluate the effect of nutrient intake, and a specific phenotype such as a different dietary pattern (i.e., diet characterized

by a high intake of vegetables and fruits) may protect against global DNA hypomethylation.

In parallel, simple methods have been developed for the rapid detection and quantitation of global 5-mC, 5-hmC, and 6-mA, including competitive enzyme

TABLE 8.2 Nucleic Acid Modification Analysis Methods Depending on Pretreatment and Resolution.

Treatment	Global	Locus-specific	Genome-wide Array
Chemical modification	Alu/LINE-1-PCR, pyrosequencing, Alu/LINE-1-HRM	MSP/PCR, MethyLight, SMART-MSP, MS-HRM, pyrosequencing, MassARRAY	Infinium BeadChip 450K or MethylationEPIC, Golden Gate
Affinity enrichment	Enzyme-linked immunosorbent assay for 5-mC, 5-hmC and 6-mA	MeDIP-PCR, MIRA	MeDIP-Chip, MIRA-Chip
Restriction enzyme-based digestion	HPLC, LC-MS, LUMA	HpaII-PCR, MS-MLPA, MS-FLAG	Differential methylation hybridization, methylated CpG island amplification and microarray, HELP, CHARM, Mmass

5-hmC, 5-hydroxymethylcytosine; 5-mC, 5-methylcytosine; 6-mA, N6-methyladenosine; CHARM, comprehensive high throughput arrays for relative methylation; HELP, HpaII tiny fragment enrichment by ligation-mediated PCR; HPLC, high-performance liquid chromatography; LC-MS, liquid chromatography-mass spectroscopy; LUMA, luminometric methylation assay; MethylCap, methyl-DNA binding domain capture technique; MIRA, methylated CpG island recovery assay; Mmass, microarray-based methylation assessment of single samples; MSCC, methylation-sensitive cut counting; MS-HRM, methylation-specific high-resolution melting analysis; SMART-MSP, sensitive melting analysis after real-time methylation-specific PCR; -seq, followed by sequencing; WGSBS, whole-genome shotgun bisulfite sequencing.

immunoassays (enzyme-linked immunosorbent assays) for serum, plasma, urine, isolated cells, or tissues.

LOCUS-SPECIFIC QUANTIFICATION OF METHYLATION

To provide a deep resolution of DNA methylation levels and more information on the genomic positions at which nucleotide methylation is altered, other approaches are needed to associate such changes with functional outcomes. Site-specific DNA methylation levels normally tend to be lower than global methylation levels. About 80% of all CpGs have been found to be methylated in human genomes, and some methylated CpGs are located in CpG islands. Approximately 60% of all genes have a CpG island in their promoter region, and these promoter-associated CpG islands may be susceptible to methylation to modulate their gene expression (the promoter regions contain about 70% of CpG islands in human genes).

Furthermore, most studies focused on quantifying the methylation status of these CpG islands. Most of these gene-specific methylation analysis methods are mainly PCR-based. Specific primers are designed to amplify both methylated and unmethylated regions with equal efficiency, because the methylation level is analyzed by different approaches (Fig. 8.3).

Methylation-Specific Polymerase Chain Reaction

This technique involves initial modification of DNA by sodium bisulfite treatment and subsequent amplification with specific primers for methylated versus unmethylated DNA. Methylation-specific PCR requires only small quantities of DNA and is sensitive to 0.1% methylated region of a given CpG island locus.

One important improvement of this method is the quantitative real-time PCR assay that is able to determine DNA methylation levels using oligonucleotide primers that anneal differentially to bisulfite-converted DNA according to the methylation status. Moreover, the fluorescence-based technique called MethyLight uses three primers or probes (forward and reverse primers and an interpositioned probe) allowing for a high degree of specificity, sensitivity, and flexibility in methylation detection.

Methylation-Sensitive High-Resolution Melt Analysis

A further technique to determine DNA methylation levels is MS-HRM, a quantitative real-time PCR-based technique initially designed to distinguish genetic polymorphisms. The PCR products are analyzed directly by dissociation curve analysis and the resulting release of an intercalating fluorescent dye during melting. The

	Methylation-Specific PCR (MSP)	Methylation-Sensitive HRM	Pyrosequencing	MassARRAY (MALDI-TOF)
Method	PCR based on specific primers for methylated versus unmethylated DNA	PCR followed by dissociation curve analysis (analysis of temperature ramping)	DNA sequencing based on the "sequencing by synthesis" principle	PCR amplification followed by an enzymatic base specific cleavage.
Primers	Two types of primers, for methylated and unmethylated sequences	Primers do not include CpGs	Primers do not include CpGs	Primers do not include CpGs
Advantages	It is particularly useful to interrogate CpG islands with high methylation density	Simplicity and high reproducibility	"Gold standard" quantitative method	Sequence up to 400 bp. It allows for interrogation of numerous CpG sites in a cost-efficient manner
Limitations	For astringency, the primer set has to contain as many CpG sites as possible	It assesses methylation in the amplified region as a whole rather than at specific CpG sites	Expensive	Expensive

FIGURE 8.3 Main procedures to analyze locus-specific DNA methylation after sodium bisulfite treatment. HRM, high-resolution melt; MALDI-TOF, matrix-assisted laser desorption ionization–time of flight; MassARRAY, mass spectrometry methylation assay; PCR, polymerase chain reaction.

status of methylation, as represented by the C-to-T content in the PCR product, determines the rapidity of melting and consequent release of the dye.

MS-HRM allows direct quantitation in a single-well assay. It is particularly useful for assessing methylation in short fragments of DNA (a range of 150 bp) derived from FFPE biopsies, which are often moderately to highly degraded and have small amounts of tissue. MS-HRM also provides accurate distinction between mixtures of methylated and unmethylated DNAs, allowing discrimination between DNA from 0% to 100% methylated as long as the bisulfite conversion is completely homogeneous. However, heterogeneous methylation across the CpG sites could influence the melting profile of the PCR product and thus the estimates of methylation obtained by this technique.

Pyrosequencing

Pyrosequencing is the standard quantitative method used for bisulfite sequencing. Similar to the prior techniques, pyrosequencing requires the use of bisulfite-converted DNA and has the same limitations derived from this treatment. Although MS-HRM is simpler and relatively less expensive than pyrosequencing to quantify the methylation of specific CpG sites, this method provides information about the methylation status of single CpG sites. Using PCR, the DNA is amplified and tagged by a primer that is biotinylated. Pyrosequencing is based on the release of pyrophosphate when nucleotides incorporate into the sequencing primer only if it is complementary to the template DNA sequence. This method is widely used in disease research and in nutritional studies. As an example, a report described the consequences of maternal undernutrition during gestation and lactation on DNA methylation and the expression of leptin, a gene involved in regulating body weight and food intake (Jousse et al., 2011). Differences in leptin promoter methylation between mice fed a low-protein diet and control mice were associated with lower body weight or adiposity and higher food intake.

Mass Spectrometry Methylation Assays

The mass spectrometry methylation assays (MassARRAY) is a highly sensitive detection method based on the difference in fragment size that has been cleaved according to the methylation status of multiple CpG sites. This strategy has the ability to sequence reads up to 400 bp, which is considerably longer than other methods, but it also requires the use of bisulfite-converted DNA and primers designed in regions without CpG nucleotides

or polymorphisms. The specific DNA sequence is amplified by PCR with tagged primers in vitro transcribed into RNA, and base-specifically cleaved by an endoribonuclease. Mass spectra of cleavage products are obtained by MassARRAY analyses based on matrix-assisted laser desorption ionization–time of flight mass spectrometry. Fragments differ in mass depending on the sequence changes introduced by the initial bisulfite treatment. This method is usually used for validations of whole epigenome screening experiments (e.g., arrays or NGS technology) and candidate gene promoter analyses.

Methylation-Specific Multiplex Ligation-Dependent Probe Amplification

Methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) is a semiquantitative method for DNA methylation analysis. It is a variant of the MLPA technique in which copy number detection is combined with the use of a methylation-sensitive restriction enzyme. The MS-MLPA protocol is similar to the standard multiplex ligation-dependent probe amplification method, except that each MS-MLPA reaction generates two samples: one undigested sample for copy number detection and one digested sample for methylation detection. Besides, MS-MLPA probes for target sequence contain the restriction site of the methylation-sensitive endonuclease HhaI. This technique has been widely used to detect epigenetic alterations in imprinting diseases such as Prader–Willi, Angelman, Beckwith–Wiedemann, and Russell–Silver syndromes. Furthermore, MS-MLPA is an excellent choice for quantitative analysis of methylation in archival FFPE samples. Both the quality and quantity of DNA may limit subsequent analyses because it requires neither a large amount of input nor bisulfite conversion, and it has the ability to quantify multiple CpGs per sample at the same time.

Methylated DNA Affinity Enrichment Assay

In affinity-based approaches such as methylated DNA immunoprecipitation (MeDIP) and the methylated CpG island recovery assay, fragmented DNA is immunoprecipitated using a monoclonal antibody for 5-mC, 5-hmC, or 6-mA or by methyl DNA-binding proteins (MBD). The enriched DNA can be quantified by PCR-based methods to assess the methylation status of specific regions. Alternatively, MeDIP can be combined with large-scale analyses like microarrays, as described subsequently.

GENOME-WIDE DNA METHYLATION ANALYSIS

Rapidly improving methodologies are increasing the capacity to assess either locus-specific (epigenetics) or genome-wide (epigenomics) DNA methylation.

Different principles used to analyze locus-specific nucleic acid methylation patterns are combined with microarrays or NGS. However, microarrays are gradually being replaced by sequencing-based approaches, which hold the promise of providing biological insights and new avenues for translational research and clinical applications.

Microarray-Based Methylation Analyses

Three main classes of microarray-based methods have been developed to quantify 5-mC or 5-hmC distribution in genomes: (1) those based on sodium bisulfite modification and, in the case of 5-hmC, followed by a treatment with potassium perrhenate, which enables discrimination between 5-mC and 5-hmC; (2) those using methylation-sensitive restriction enzymes; and (3) those enriching for highly methylated regions using a specific antibody for 5-mC or 5-hmC, or by methyl-binding proteins.

In bisulfite-based epigenotyping, microarrays are designed using pairs of oligonucleotide hybridization probes targeting the CpG sites of interest. One probe is complementary to the unaltered methylated sequence; the other is complementary to the unmethylated sequence. The Illumina Methylation Assay (Infinium or Golden Gate technologies) is a method that applies bisulfite sequencing technology at a microarray level to generate genome-wide methylation data, because it is able simultaneously to interrogate over than 850,000 CpG sites per sample at single-nucleotide resolution with the last platform (Infinium MethylationEPIC Kit).

An increasing number of scientific contributions are applying these technologies in nutrition research. For instance, Milagro et al. used a combination of a high-throughput microarray-based assay with a second, more specific technique (MassARRAY), including more individuals than in the first technique, which allowed the researchers to find new biomarkers and reliably validate them (Milagro et al., 2011). With this dual approach, the authors reported that a caloric restriction intervention promoted DNA methylation modifications in overweight and obese men. Moreover, a baseline DNA methylation pattern might be used as an early indicator of response to the metabolic effects of a weight loss program.

One of the enzyme-based methods coupled with array-based analysis is methylated CpG island

amplification and microarray, which uses the differential methylation sensitivities and cutting behaviors of two endonucleases (SmaI and XmaI) with a final array hybridization step. An alternative method is differential methylation hybridization (DMH), which combines digestion of a pool of DNA with a methylation-sensitive restriction enzyme and a mock digestion of another pool. This approach produces parallel DNA pools that are amplified and labeled with different fluorescent dyes for two-color array hybridization. The relative signal intensities can be used to extract DNA methylation information at the corresponding loci on the array. An adaptation of DMH, known as MethylScope, uses the endonuclease McrBC to cut sheared DNA randomly. The combination with the tiling array (a subtype of microarray that functions by hybridizing labeled DNA or RNA target molecules to probes fixed onto a solid surface) design and data processing is referred to as comprehensive high-throughput arrays for relative methylation. Another method, known as HpaII tiny fragment enrichment by ligation-mediated PCR (HELP), uses ligation-mediated PCR for the amplification of HpaII or MspI genomic restriction fragments followed by array hybridization. Finally, another affinity enrichment-based method implicates the coupling of MeDIP with DNA microarrays to obtain relative methylation levels at the loci represented on the array.

However, methylation changes identified by microarray experiments will continue to require validation by one or more of the locus-specific methods. Moreover, the relative finesse of DNA methylation changes induced by nutritional exposures will motivate the continued development of analytical methods that should be rapid, inexpensive, highly sensitive, and quantitatively accurate. On the contrary, although methylation microarrays are a powerful tool for epigenetic studies, they have an important limitation because they analyze only a small part of the CpG sites of the genome.

Next-generation Sequencing—Based Methylation Analyses

In an effort to reach deeper coverage (where coverage is number of reads representing a given nucleotide in the sequence) and single-nucleic resolution of methylation analysis, NGS has revolutionized genomic research (Zhang and Jeltsch, 2010). This technology has been adapted to epigenomic research including DNA/RNA methylation profiling at high resolution and low cost. Several sequencing approaches are available, and the choice of platform and methodology depends on the scientific application (Table 8.3). Moreover, the cost and capacity for extensive bioinformatics analysis must be considered when a researcher uses this application.

TABLE 8.3 Next-Generation Sequencing Approaches for Nucleic Acid Modification Analysis.

Nucleic acid modification	Resolution	Enrichment	Whole genome	Arrays
5-mC	Single CpG	RRBS, mRRBS, LCM-RRBS, or scRRBS	WGBS, T-WGBS, or PBAT	Infinium BeadChip –450K or MethylationEPIC
	Peak	MRE-seq, MeDIP-seq, MBD-seq, or MethylCap-seq	–	CHARM or MeKL-ChIP
5-mC oxidation derivatives	Single CpG	RRHP, reduce representation sequencing with TAB-seq, oxBS-seq, CAB-seq, fCAB-seq, or redBS-seq	TAB-seq, oxBS-seq, CAB-seq, fCAB-seq, redBS-seq	OxBS-450K or OxBS-MethylationEPIC
	Peak	DIP-seq, anti-CMS, hMe-Seal, fC-Seal, GLIB, JBP1, EpiMark, or Aba-seq	–	–
6-mA	Peak	MeRIP-seq	–	–

5-hmC, 5-hydroxymethylcytosine; *5-mC*, 5-methylcytosine; *6-mA*, N6-methyladenosine; *Aba-seq*, DNA-modification-dependent restriction endonuclease; *anti-CMS*, anti-cytosine-5-methylenesulfonate; *CAB-seq*, chemical modification-assisted bisulfite sequencing; *CHARM*, comprehensive high-throughput arrays for relative methylation; *DIP-seq*, DNA immunoprecipitation and shotgun sequencing; *fCAB-seq*, 5fC chemical modification-assisted bisulfite sequencing; *fC-Seal*, 5-formylcytosine selective chemical labeling (fC-Seal) approach for affinity purification and genome-wide profiling of 5fC; *GLIB*, glucosylation, periodate oxidation and biotinylation; *JBP1*, J-binding protein; *LCM-RRBS*, laser-capture microdissection-reduced representation bisulfite sequencing; *LHC-BS* (pre- and postconversion), liquid hybridization capture-based bisulfite sequencing; *MBD-seq*, methyl-CpG-binding domain protein sequencing; *MeDIP-seq*, methylation DNA immunoprecipitation sequencing; *MeKL-ChIP*, methylated DNA, kinase pretreated ligation-mediated PCR amplification-chromatin immunoprecipitation; *MeRIP-seq*, methylation RNA immunoprecipitation sequencing; *MethylCap-seq*, methylation DNA capture sequencing; *MRE-seq*, methylation restriction enzyme sequencing; *mRRBS*, multiplexed reduced representation bisulfite sequencing; *OxBS-MethylationEPIC*, oxidative bisulfite Infinium MethylationEPIC; *oxBS-seq*, oxidative bisulfite sequencing; *PBAT*, postbisulfite adaptor tagging; *redBS-seq*, reduced bisulfite sequencing; *RRBS*, reduced representation bisulfite sequencing; *RRHP*, reduced representation 5-hmC profiling; *scRRBS*, single-cell reduced representation bisulfite sequencing; *TAB-seq*, TET-assisted bisulfite sequencing; *T-WGBS*, transposase-based library construction; *WGBS*, whole-genome bisulfite sequencing.

Sodium bisulfite conversion followed by massively parallel sequencing (bisulfite-seq) has become an increasingly used method for investigating epigenetic profiles. Bisulfite-seq is well-suited to the investigation of epigenetic profiles from clinical tissue samples and can be applied to small quantities of DNA including FFPE samples.

Whole-genome bisulfite-seq provides coverage at a single-bp resolution and is the most comprehensive technique, covering more than 90% of cytosines in the human genome. This method requires the most extensive bioinformatics tools and is the most expensive because more sequences are needed to cover the entire genome. An alternative to whole-genome bisulfite-seq is reduced representation bisulfite sequencing (RRBS), which combines restriction enzymes and Bisulfite-seq in order to enrich for CpG islands and gene regulatory sequences. Specifically, this technology covers approximately 12% of all CpG sites and 84% of all CpG islands in the human genome.

As with microarray analysis, approaches have been described focused on affinity-based sequencing, such as MeDIP-seq or MethylCap (MBD capture technique),

or enzyme-based sequencing methods such as methyl-seq and HELP-seq.

Moreover, a chemistry called oxidative bisulfite sequencing (oxBS-seq) has enabled the quantification of 5-hmC in whole-genome studies. Although other techniques are able to quantify this modification, such as MeDIP (array or sequencing) with specific antibodies for 5-hmC, 5-fC, or 5-caC, oxBS-seq development facilitates the detection of 5-hmC at single-base resolution. This method is based on the selective chemical oxidation of 5-hmC to 5-fC, which enables bisulfite conversion of 5-fC to uracil. In this sense, high levels of 5-hmC were found in CpG islands associated with transcriptional regulators and in long interspersed nuclear elements, which suggests that these regions might undergo epigenetic reprogramming. A new genome-wide approach named methylation-assisted bisulfite sequencing has been presented. It enables single-base resolution mapping of 5-fC and 5-caC and measures their abundance. DNA methyltransferase methylates unmodified C to generate 5-mC, which can be successively oxidized by ten-eleven translocation to generate 5-hmC/5-fC/5-caC (Neri et al., 2015).

However, despite these advances, and probably because of the high cost, the study of methylation changes using these approaches is only slowly finding its way into metabolic and nutritional studies. Despite this, further studies in this area will begin to emerge in coming years.

References

- Jousse, C., Parry, L., Lambert-Langlais, S., Maurin, A.C., Averous, J., Bruhat, A., Carraro, V., Tost, J., Letteron, P., Chen, P., Jockers, R., Launay, J.M., Mallet, J., Fafournoux, P., 2011. Perinatal undernutrition affects the methylation and expression of the leptin gene in adults: implication for the understanding of metabolic syndrome. *FASEB J* 25, 3271–3278.
- Mansego, M.L., Milagro, F.I., Campión, J., Martínez, J.A., 2013. Techniques of DNA methylation analysis with nutritional applications. *J Nutrigenetics Nutrigenomics* 6, 83–96.
- Marques-Rocha, J.L., Milagro, F.I., Mansego, M.L., Mourão, D.M., Martínez, J.A., Bressan, J., 2016. LINE-1 methylation is positively associated with healthier lifestyle but inversely related to body fat mass in healthy young individuals. *Epigenetics* 19, 1–12.
- Milagro, F.I., Campión, J., Cordero, P., Goyenechea, E., Gómez-Uriz, A.M., Abete, I., Zulet, M.A., Martínez, J.A., 2011. A dual epigenomic approach for the search of obesity biomarkers: DNA methylation in relation to diet-induced weight loss. *FASEB J* 25, 1378–1389.
- Neri, F., Incarnato, D., Krepelova, A., Rapelli, S., Anselmi, F., Parlato, C., Medana, C., Dal Bello, F., Oliviero, S., 2015. Single-base resolution analysis of 5-formyl and 5-carboxyl cytosine reveals promoter DNA methylation dynamics. *Cell Rep* 10, 674–683.
- Zhang, Y., Jeltsch, A., 2010. The application of next generation sequencing in DNA methylation analysis. *Genes* 1, 85–101.

This page intentionally left blank

Proteomic Analyses

Fernando J. Corrales¹, Leticia Odriozola²

¹Functional Proteomics Laboratory, Centro Nacional de Biotecnología, CSIC, CIBERhed-ISCIH, PRB2, ProteoRed-ISCIH, Madrid, Spain; ²CIMA, University of Navarra, Pamplona, Spain

Glossary Terms

Biomarker Biomolecule indicating a particular biological condition or process.

Electrophoresis Separation of biomolecules by their different behaviors in an electric field.

Liquid chromatography Separation of biomolecules by their different degrees of interaction with a given sorbent.

Mass spectrometry Analytical technique that arranges ions based on their mass to charge ratio.

Proteoform Different protein species encoded by a given gene.

Proteome Complete set of protein species of a cell or tissue.

Proteomics Science that studies the proteome.

the scientific community to develop proteomics in an attempt to characterize the proteinaceous component or proteome of a biological system. Proteomics has experienced a huge expansion and has attracted the interest of many scientists because proteins are the drivers of cellular pathways in addition to regulating them. Besides, the human proteome configures the universe for the discovery of novel biomarkers and nutritional/therapeutic targets that are expected for population stratification and to promote the development of individualized interventions tailored for the specific needs of particular individuals.

The term “proteome” was initially coined by Wilkins in 1995, although the overall concept of proteomics was introduced in modern biology in 1982 by Anderson, who proposed a human protein index based on two-dimensional electrophoresis (2DE). Accordingly, proteomics can be defined as a discipline focused on the study of the proteome using technologies that allow for large-scale analysis. The proteome can be defined as the set of proteins expressed by a genome, but this concept requires additional considerations if the proteome must be regarded in its whole dimension. The proteome is highly dynamic; it must be defined within genetic and environmental constraints that will determine the repertoire of protein species. The plasticity of the proteome is essential to finely tune cellular functions by control mechanisms in addition to those that are merely transcriptional, enabling fast and efficient reactions to external stimuli orchestrated in adaptive responses to ensure cell survival. It would be easy to define the proteome as a simple translation of the 20,000 human genes, but the different mechanisms of posttranscriptional regulation makes it necessary to revise the rule of one gene, one protein. RNA transcript splicing, protein processing, and posttranslational modifications (PTMs) (more

INTRODUCTION

The massive amount of information generated from the many genome-wide sequencing projects has greatly improved our interpretation of human biology in health and disease. However, the sequence of the 3,120,000,000 base pairs integrating the human genome cannot in itself explain the biological complexity of the human body. The roughly 20,000 human protein coding genes and the yet uncounted nonprotein coding structures are common to more than 200 differentiated cell types that specialize in performing specific functions. The genome must be considered the first information layer of a sophisticated network that configures the structural and functional diversity and is not enough to provide a full understanding of phenotypes, the dynamic and regulatory mechanisms orchestrating adaptive responses and pathogenic processes. Thus, it appears evident that in addition to specific gene expression programs and splicing events, a definition of protein profiles as well as their regulation is essential to translate genes into biological functions. Realizing that proteins are the functional cellular effectors, and the availability of unprecedented technological resources have moved

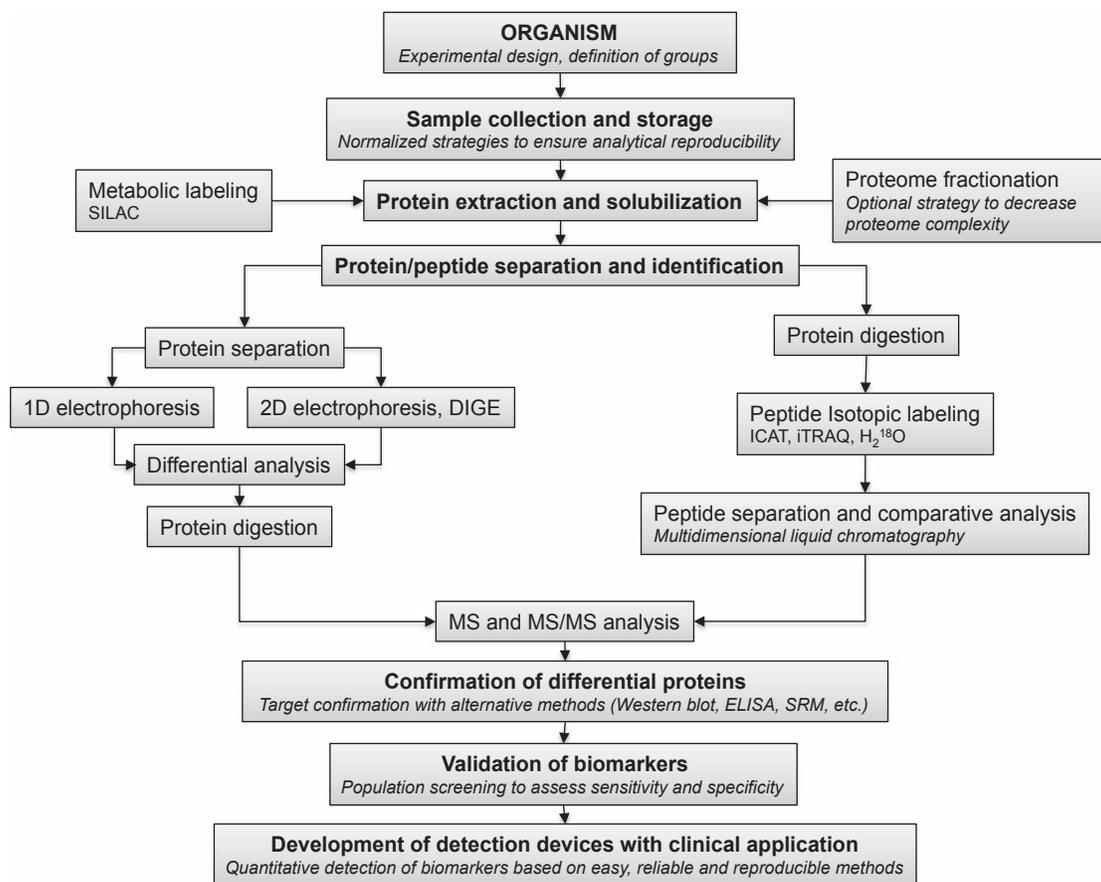


FIGURE 9.1 Schematic representation of the different steps of a proteomics pipeline. *1D*, one-dimensional; *DIGE*, difference gel electrophoresis; *ELISA*, enzyme-linked immunosorbent assay; *ICAT*, isotope-coded affinity tag; *iTRAQ*, isobaric multiplexing tagging system; *MS*, mass spectrometry; *SILAC*, stable isotopic labelling with amino acids in cell culture; *SRM*, selected reaction monitoring.

than 300 different types of PTMs have been described) generate the necessary diversity of gene products to define phenotypes and increase the complexity of the proteome landscape up to at least hundreds of thousands of proteoforms.

The huge dimension of the proteome, the physicochemical heterogeneity of proteins, and the wide dynamic range across protein species in biological matrices (up to 10 logs) make it necessary to elaborate sophisticated pipelines integrating different analytical strategies to study the proteome or specific subproteomes in depth. A typical proteomics experiment can be structured in four different steps: sample collection, handling, and storage; protein extract preparation; protein/peptide separation; and protein/peptide identification, characterization, and quantitation. The sample type as well as the aims of the study will determine the optimal combination of analytical methods (Fig. 9.1). One critical issue for a successful proteomic experiment is the design of sample collection and storage procedures that ensure the maintenance of the proteome integrity and minimize sample heterogeneity. In this regard, interaction with biobanks (BBMRI-ERIC

(<http://www.bbMRI-eric.eu>) is of utmost importance to obtain access to reliable sample collections generated under standardized protocols. Second, proteins must be solubilized, minimizing contamination with other biomaterials (lipids, nucleic acids, etc.). Although many standardized and reproducible procedures are available, no protocol provides an efficient extraction of all proteins of a given biological sample. High concentrations of chaotropic agents, nonionic detergents, and reducing agents are common components of extraction buffers because they allow efficient protein (even hydrophobic protein) solubilization preventing artifactual oxidation events throughout manipulation. However, the method must be adapted to requirements imposed by the sample type (cell, tissue type, storage conditions, etc.), subsequent analytical steps (avoiding interfering substances), and the aims of the study, which may require protein/peptide labeling, subcellular fractionation, enrichment of specific protein/peptide families, and so on, as will be discussed in the following sections. Especially relevant in biomedical research is comparative proteomics, because it enables the identification of protein mediators of biological responses or pathogenic

processes. Different combinations of protein/peptide separation strategies and mass spectrometry (MS) approaches provide reliable and accurate workflows for unsupervised extensive proteome analyses generally used in discovery studies as well as for the validation of the newly generated hypotheses.

BOTTOM-UP PROTEOMICS

Gel-based Strategies

A differential proteomics analysis can be conceived of as a gel-based or a gel-free approach. In the first case, proteins are separated by 2DE based on orthogonal methods resolved by two physical protein properties: the isoelectric point in the first dimension (isoelectrofocusing) and the molecular weight under denaturing conditions (sodium dodecyl sulfate–polyacrylamide gel electrophoresis [SDS-PAGE]) in the second (Rabilloud, 2011). Upon Coomassie, silver, or fluorescent staining, gels are commonly scanned, and the digital replicates are subjected to image analysis for spot detection and normalization followed by statistical analysis to determine the differential spots across the experimental conditions under comparison. A substantial improvement of this method was introduced by two-dimensional fluorescence difference gel electrophoresis, in which proteins from the samples under study are differentially labeled using specific fluorescence dyes (Cy2, Cy3, or Cy5), mixed in equal amounts in a pairwise mode and resolved by 2DE. Scanning gels with laser beams at different wavelengths, as required for optimal excitation of each Cy dye, generate individual images of the corresponding protein spot map from each sample. Then, protein levels can be compared based on the relative intensity of the spots within a single gel, which reduces variability and improves the accuracy of the method. Differential spots are then excised and incubated with proteases (trypsin is the most commonly used because the tryptic digestion is robust and generates peptides with a positive charge in the C terminal at acidic pH), although other proteases or even combinations of two proteolytic enzymes have also been implemented to circumvent the limitations of trypsin digestion in the case of certain proteins. The resulting peptide mixture is analyzed by mass spectrometric methods to enable protein identification. A simple approach to identifying proteins is peptide mass fingerprinting (PMF), in which tryptic digests are analyzed by matrix-assisted laser desorption/ionization (MALDI) time of flight (TOF) MS and the identifications are deduced from the specific set of peptide m/z ratios upon comparison with the theoretical PMFs of genome-wide protein sequence databanks (Karas and Hillenkamp, 1988). The use of

MALDI TOF equipment greatly increased the reliability of identification. Alternatively, peptides can be separated by liquid chromatography (LC) (nanoLC) and analyzed by tandem MS (MS/MS), which provides information relative to the amino acid sequence (Fenn et al., 1989). The chromatographic column is connected online with the MS/MS instrument through a nano-electrospray ionization (ESI) source that allows ionization of the eluted material. Once ionized, a particular peptide or ion is isolated and fragmented by collision with an inert gas, generating a fragment spectrum that is compared with the theoretically predicted fragments for all peptide sequences in a database to identify the peptide sequence and then the precursor protein.

2DE is a relatively simple method that enables the visualization and analysis of 2000–3000 protein spots, the detection of posttranslationally modified species modifying protein isoelectric point (pI) values, and an estimation of differentially expressed proteins. However, some constraints are also associated with 2DE: low abundant (below 1000 copies), hydrophobic, extreme pI or Mr proteins are barely detected, preventing a complete description of the proteome.

Shotgun Analysis

Shotgun proteomics emerges to circumvent some of the major constraints of 2DE; it proposes a gel-free alternative based on the tandem mass spectrometric analysis (MS/MS) of all peptides resulting from the proteolytic digestion of whole-protein mixtures, upon multidimensional LC separations (Zhang, 2013; Gillet et al., 2016). Hence, protein identification and quantitation are deduced from MS and MS/MS data. For differential analyses, two major workflows are available that rely on protein/peptide stable isotope labeling or a label-free strategy. In the first case, proteins can be metabolically labeled in cell culture (stable isotopic labeling with amino acids in cell culture) or peptide mixtures from different samples can be chemically (isotope-coded affinity tag [ICAT]); isobaric multiplexing tagging system [TRAQ], or enzymatically (digestion in $H_2^{18}O$) labeled using light and heavy isotopes, and then combined equally for subsequent simultaneous nanoLC ESI-MS/MS analysis. To increase the resolution of the experiment, different prefractionation strategies are routinely implemented. Protein extracts can be resolved in SDS-PAGE gels and in gel protein digestion of a variable number of gel line slices, providing fractionated peptide pools that are subsequently analyzed by nanoLC (conventionally reversed phase C18 chromatography) coupled online with the mass spectrometer through an ESI source. Alternatively, the whole labeled peptide population resulting from the combination of all protein extracts can first be resolved by cation exchange or C18

at high pH LC to analyze the obtained fractions by LC-ESI-MS/MS, as indicated earlier. MS and MS/MS data provide information allowing peptide assignment to an amino acidic sequence based on information inferred from the fragments generated by collision-induced dissociation (CID), most frequently, electron capture dissociation (ECD), or electron transfer dissociation (ETD), among other fragmentation strategies and also allow for the identification of PTM sites. Search engines are software that provide the means for sequence assignment from spectral information and protein identification with statistical significance (Cottrell, 2011). Relative quantification is also achieved in parallel from the intensity ratio of differentially labeled precursor ions (ICAT, for instance) or reporter ions generated upon fragmentation (iTRAQ). Label-free quantitation involves the independent analysis of at least three replicates of each sample, usually with no previous fractionation, and the abundance ratio is estimated based on the precursor signal intensity or on spectral counting (spectral counting counts the number of spectra identified for a given peptide in different biological samples and integrates the results for all measured peptides of the protein[s] that are quantified).

TOP-DOWN PROTEOMICS

Bottom-up proteomics methods have greatly contributed to extend our knowledge of the human proteome, but an analysis of fragments derived from protein enzymatic or chemical digestion has disadvantages. Peptides might not be specific to a protein or protein form, leading to inconclusive identification; large protein regions may not be mapped, losing relevant information such as PTMs and sequence variants; the relations among variations or modifications within the same polypeptide chain will be lost. In this scenario, top-down MS proposes an analysis of intact proteins to provide an overview of all intact proteoforms, PTMs localization, and sequence variants (Catherman, 2014). The large complexity of most proteomic samples imposes offline separation step(s) (either electrophoretic or chromatographic) before LC-MS/MS analysis. Top-down studies require high-performance mass spectrometers because high resolution and mass accuracy (in some cases, at the millidalton level) are critical to separate and assign precursor spectra containing multiple intact proteoforms or MS2 spectra with hundreds of fragment ions. Moreover, high sensitivity is also needed because proteins have broad isotopic distributions, scattering the signal of a given protein across many peaks. The use of ECD and ETD provides fragmentation patterns complementary to CID and increases the capacity to dissect molecular complexity, facilitating the quantification of proteoforms, localization of

PTMs and variations, assignment of positional isomers, and assessment of PTM interdependence. Although there are still several challenges for top-down proteomics, including protein solubilization, separation, and the detection of large intact proteins, the improvement in separation tools, the development of new mass spectrometers with unprecedented capacity, and the computational tools already available to identify and characterize proteoforms reliably allow the study of complex proteomes and expand its use in biomedical research (Toby et al., 2016).

TARGETED PROTEOMICS

Nontargeted MS-based proteomics has been embraced by researchers studying human biology and disease. These wide-screening procedures offer a panoramic picture of the relative expression and modification of large numbers of proteins and are increasingly used to generate de novo functional hypothesis and as target identification methods in discovery studies. However, they are not optimal for absolute quantitative analysis of predefined targets or for testing the selected targets across large numbers of samples. Complementary procedures such as sequential window acquisition of all theoretical mass spectra (SWATH MS) and selected reaction monitoring (SRM) have emerged as methods of choice for MS-based targeted proteomics.

Selected Reaction Monitoring

Any protein can theoretically be detected and quantitated in any biological matrix selectively recording precursor to fragment transitions of its proteotypic peptides, which are specific and allow for the unequivocal identification of the target protein. The sample is prepared according to general procedures as mentioned earlier, and peptides are prefractionated when required or directly analyzed by nanoLC-ESI-MS/MS. Triple quadrupole instruments are typically used; they are capable of filtering selected ions in the first and third quadrupoles, with the second serving as the collision cell. A set of proteins to be monitored in the MS instrument is selected in a number inversely proportional to the sensitivity and accuracy required in the analysis. Then, specific precursor ions (peptides) and transitions are selected to configure the SRM method, which allows specific identification and accurate quantitation (Anderson, 2006). The use of preexisting MS information from SRM data and multiple shotgun analyses available in databanks including SRM atlas (Kusebauch et al., 2016; <http://www.srmatlas.org>), GPMDB (<http://www.thegpm.org>), and PRIDE (<http://www.ebi.ac.uk/pride/>) significantly reduces the time required to

develop the assay. Because specific m/z values for the parent and fragment ions are specified in the SRM method, modified peptides would not be detected unless specifically targeted.

SRM emerges as a reference method for protein quantification because it has a high dynamic range, is absolute rather than relative, and is reproducible and reliable even across different laboratories when stable isotope labeled standards are spiked in the sample. For these reasons, it is increasingly used to characterize model systems, clinical samples, and toxicology studies, among others (Kearney et al., 2018). The possibility of multiplexing and redefining SRM assays for optimal performance in sample screening programs has made this strategy an attractive alternative to antibody-based methods for targeted protein analysis.

SWATH Mass Spectrometry

A main limitation of SRM is that the number of targeted measurements in a single run is restricted to about 100. SWATH MS has emerged as an alternative method to mitigate this restriction: data independent acquisition MS technology based on targeted data extraction instead of targeted MS acquisition, as in SRM (Liu, 2013). In SWATH MS, MS2 data are recorded by cycling through 32 consecutive 25-Da precursor isolation windows and all fragment ions are monitored at high resolution (10 parts per million), ensuring the specificity of peptide identification. The result is a digital map of MS2 ions derived from the fragmentation of all precursor ions present in the sample in a given m/z window at a certain time. Fragment ion chromatograms for each peptide of interest can be then extracted from the digital maps to perform the targeted analysis. SWATH MS measurements rely on the use of MS reference maps of spectral libraries collecting a large set of validated peptide MS2 spectra (PeptideAtlas, Human Proteinpedia, GPM Proteomics Database, PRIDE, etc.). Combining SWATH MS data with reference maps supports high sensitivity detection and proteome coverage, gives accurate and reproducible quantification, and provides permanent maps that can be reinterrogated to validate emerging biomarkers or hypotheses. The possibility of monitoring all proteins of a sample at a constant sensitivity and reproducibility in large cohorts and the capacity to obtain a personalized proteomic profile from SWATH maps including different tissues for iterative biomarker studies will significantly extend the application of this technology in clinical research.

PROTEIN REGULATION: POSTTRANSLATIONAL MODIFICATION ANALYSIS BASED ON ENRICHMENT PROCEDURES

PTMs enable the rapid and efficient modulation of biological functions and promote efficient adaptive responses to environmental conditions that ensure cell survival. PTMs result from the covalent bonding of different chemical moieties to specific amino acid side chains. Massive PTM screening is especially challenging because the relative stability of some PTMs and the substoichiometric proportion of the modified species hinder their detection and quantification. Protocols to preserve specific PTMs and enrichment procedures to increase the relative concentrations of modified proteins are needed to increase the depth of the study substantially.

Protein Phosphorylation

Reversible protein phosphorylation is a complex network of signaling and regulatory events that affects virtually every cellular process. An understanding of phosphorylation networks in the physiological context remains limited largely owing to inefficient isolation of serine/threonine-phosphorylated peptides. Application of enrichment approaches at the level of phosphorylated proteins and/or peptides using a combination of ion exchange and immobilized metal affinity or TiO₂ chromatography or specific phosphotyrosine or antibodies has allowed the identification of thousands of unique phosphorylation sites, which configure a detailed map to guide the comprehensive exploration of human biological functions that are regulated by phosphorylation (<http://kinase.com/human/kinome/>).

Protein Acetylation

Protein lysine acetylation is a finely tuned process controlled by acetyltransferases and deacetylases that, like other PTMs, regulate a variety of physiological processes including enzyme activity, protein–protein interactions, gene expression, and subcellular localization; it has been associated with numerous disease states including cancer, aging, and cardiovascular, metabolic, and alcoholic liver disease. Characterization of the acetyl-proteome has been greatly extended by the combination of enrichment of lysine (K)-acetylated proteins/peptides based on specific antibodies and high sensitivity and accuracy MS instruments.

Protein Glycosylation

Protein glycosylation is the most versatile and common protein modification that has important roles in various biological processes and disease progression, including cancer. The identification of glycopeptides in a complex protein digest is difficult and requires enrichment and isolation before MS analysis. In this regard, various strategies are used, such as the periodate/hydrazide protocol and hydrophilic and lectin-mediated affinity chromatography. Lectin affinity chromatography has been extensively used to isolate glycoproteins and glycans from various biological sources such as plasma, bile, urine, and cells. These enrichment approaches appear to be complementary because little overlap is observed when they are applied to the same sample, and they greatly enhance the MS signals of glycopeptides, facilitating the sequence and glycoside assignments in typical shotgun analyses. An additional difficulty of analyzing N-glycopeptides by MS is their size, which is often far from the preferred mass range of shotgun proteomics standards. Digestion with proteases in addition to trypsin partially circumvents these limitations resulting from primary structure constraints, as well as from the effect of the modifying hydrocarbon moiety that might eventually block the access of trypsin to an adjacent K or R residue. Finally, a targeted method to determine N-glycosylation occupancy was developed based on the conversion of asparagine (N) to aspartate (D) that occurs upon peptide-N-glycanase F digestion (this amino acid change introduces a mass difference of 0.984 Da). Combinations of these experimental approaches have been delineated to configure optimized workflows aimed at studying liver glycoproteome to define novel mechanisms of liver disease progression, mainly hepatocellular carcinoma, and to identify potential biomarkers.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Proteomics has experienced incredible improvements, providing tools to analyze complex proteomes with unprecedented sensitivity and accuracy. Hundreds upon thousands of proteins can be analyzed in a single experiment, resulting in a precious source of data to understand human biology and disease. Identification of pivotal proteins driving cellular pathways and pathogenic processes sheds light on underlying molecular

mechanisms and provides biochemical changes that enable the development of predictive biomarkers that can trigger earlier nutritional and clinical interventions. Moreover, validation of candidate biomarkers require targeted methods to circumvent the constraints inherent to the vast complexity of the proteome. Tailored methods to quantitate specific proteins facilitates the analysis of large cohorts of clinical samples and might probe the translational relevance of the differential features identified. Selective procedures to enrich specific sub-proteomes together with targeted hypothesis-driven analysis are increasingly being used by the biomedical community and are starting to bridge the gap between analytical and clinical proteomics.

References

- Anderson, L., Hunter, C.L., 2006. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol Cell Proteom* 5, 573–588.
- Catherman, A.D., Skinner, O.S., Kelleher, N.L., 2014. Top down proteomics: facts and perspectives. *Biochem Biophys Res Commun* 445, 683–693.
- Cottrell, J.S., 2011. Protein identification using MS/MS data. *J Proteomics* 74, 1842–1851.
- Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., Whitehouse, C.M., 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246 (4926), 64–71.
- Gillet, L.C., Leitner, A., Aebersold, R., 2016. Mass spectrometry applied to bottom-up proteomics: entering the high-throughput era for hypothesis testing. *Annu Rev Anal Chem* 9, 449–472.
- Karas, M., Hillenkamp, F., 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60 (20), 2299–2301.
- Kearney, P., Boniface, J.J., Price, N.D., Hood, L., 2018. The building blocks of successful translation of proteomics to the clinic. *Curr Opin Biotechnol* 51, 123–129.
- Kusebauch, U., Campbell, D.S., Deutsch, E.W., Chu, C.S., Spicer, D.A., Brusniak, M.Y., Slagel, J., Sun, Z., Stevens, J., Grimes, B., Shteynberg, D., Hoopmann, M.R., Blattmann, P., Ratushny, A.V., Rinner, O., Picotti, P., Carapito, C., Huang, C.Y., Kapousouz, M., Lam, H., Tran, T., Demir, E., Aitchison, J.D., Sander, C., Hood, L., Aebersold, R., Moritz, R.L., 2016. Human SRMAtlas: a resource of targeted assays to quantify the complete human proteome. *Cell* 166, 766–778.
- Liu, Y., Hüttenhain, R., Collins, B., Aebersold, R., 2013. Mass spectrometric protein maps for biomarker and clinical research. *Expert Rev Mol Diagn* 13, 811–825.
- Rabilloud, T., Lelong, C., 2011. Two-dimensional gel electrophoresis in proteomics: a tutorial. *J Proteomics* 74, 1829–1841.
- Toby, T.K., Fornelli, L., Kelleher, N., 2016. Progress in top-down proteomics and the analysis of proteoforms. *Annu Rev Anal Chem* 9, 499–519.
- Zhang, Y., Fonslow, B.R., Shan, B., Baek, M.C., Yates 3rd, J.R., 2013. Protein analysis by shotgun/bottom-up proteomics. *Chem Rev* 113, 2343–2394.

Metabolomics for Biomarker Discovery and to Derive Genetic Links to Disease

Susan C.J. Sumner, Susan McRitchie, Wimal Pathmasiri

Nutrition Research Institute, University of North Carolina Chapel Hill, Kannapolis, NC, United States

Glossary

Biomarker A biomarker is a measurable indicator of a condition or state.

Endogenous metabolites Endogenous metabolites are produced by the biochemical pathways of the host system.

Endometabolome The complete set of endogenous metabolites.

Exogenous metabolites Exogenous metabolites are derived from exposure of the host system to external sources of drugs, medications, environmentally relevant chemicals, or foods.

Exometabolome The complete set of exogenous metabolites.

Metabolomics Metabolomics involves the study of the biochemical fingerprint that cellular processes leave behind.

Metabotype The unique set of metabolites that exist in an individual's (or system's) biological fluids, tissues, or cells at any point in time.

Targeted analysis Technologies used to detect and/or quantitate specific analytes, typically to test hypotheses regarding specific pathways or classes of compounds.

Untargeted metabolomics The use of analytical methods to capture tens of thousands of signals, which are subsequently prioritized for assignment based on the clinical or biological significance of the signals.

molecular weight metabolites and have been used in metabolomics investigations of a variety of disease, therapeutic, and intervention research.

Through an analysis of samples from model systems and human subject investigations, metabolomics signatures have been correlated with genetic makeup (e.g., gender, age, ethnicity, polymorphisms), environmental exposures (e.g., medications, illicit drugs, environmental pollutants, contaminants in food and water), lifestyle factors (e.g., tobacco use, exercise, stress, diet), gut microbes, physical and clinical phenotypes (e.g., height, weight, blood pressure, heart rate), mental health status (e.g., depression, cognitive scores), dietary intake and nutrient status, and many disease states (Stewart et al., 2015; Sumner et al., 2018). The ability to metabotype biospecimens from humans or model systems provides a means to reveal patterns of signals, determine specific metabolites, reveal pathway perturbations associated with health and disease, reveal associated genetics determinants, and inform treatment or intervention.

INTRODUCTION TO METABOLOMICS

Metabolomics involves the study of the low-molecular complement of cells, tissues, and biological fluids. An individual's metabotype is the signature, or biochemical fingerprint, of low-molecular weight metabolites that are present at any time in tissues or biological fluids. This signature is composed of endogenous metabolites (those that map to biochemical pathways) that are present as a readout of the genome, as well as exogenous metabolites that are derived from external sources of exposures. Relatively noninvasive biological fluids (e.g., urine, plasma, serum, umbilical cord blood, and plasma), excreta (e.g., urine, stool, sweat, breath), invasive fluids (cerebrospinal fluid), and organ tissue (e.g., muscle, liver, kidney, brain) are rich in low-

STUDY DESIGNS

Metabolomics studies are designed to reveal biomarkers associated with specific study phenotypes, such as states of health or disease, diet, exercise, or responsivity to treatment or intervention. These study designs address the need to develop more sensitive and early markers for disease detection and staging; monitor treatment for efficacy, adverse response, or relapse; or monitor intervention. Once metabolites are revealed that associate with the study phenotypes, pathway mapping can be used to gain mechanistic insights to identify targets for pharmacological development or nutritional

intervention, or to point to genetic polymorphisms. Studies are designed to ensure that the link between the metabolomics profiles and the phenotypic anchors (e.g., obese versus lean, case versus control, exposed versus nonexposed) are not biased by factors (e.g., gender, race/ethnicity, age, tobacco use, medication use, stress, diet) that are known to influence the metabolome, and that the study is sufficiently powered in light of the heterogeneity of the study population.

TARGETED AND UNTARGETED ANALYSIS

The terms “targeted metabolomics” and “untargeted metabolomics” are used to describe a variety of analytical data capture approaches that were previously reviewed (Stewart et al., 2015; Sumner et al., 2018). A targeted analysis of analytes is conducted to address specific hypotheses, normally focused on a particular pathway. Examples include determining whether concentrations of acylcarnitines are higher in obese subjects compared with healthy weight ones (Muoio and Newgard, 2006; Fiamoncini et al., 2018) or the role of choline and its closely associated one-carbon metabolites in subjects with or without fatty liver (Corbin and Zeisel, 2012).

Untargeted metabolomics is not focused on specific hypotheses; rather, it captures signals for a wide range of small-molecular analytes (typically less than 1500 Da) that cover a variety of biological pathways. In untargeted analysis, the endogenous complement of the metabolomics signatures is composed of thousands of signals derived from amino acids, aromatic amino acids, biogenic amines, sugars, polysaccharides, phospholipids, sphingolipids, carnitines, catecholamines, purines, pyrimidines, hormones, steroids, and so on (endogenous metabolites that comprise the endometabolome). Untargeted platforms also afford the simultaneous capture of signals for metabolites of medications, illicit drugs, cosmetics and personal products, environmentally relevant chemicals, additives in foods, and food components (exogenous metabolites that comprise the exometabolome). Gas chromatograph and liquid chromatography coupled with time of flight or orbitrap mass spectrometry systems, and nuclear magnetic resonance spectroscopy are predominantly used for untargeted metabolomics analysis (Stewart et al., 2015; Sumner et al., 2018).

REVEALING BIOMARKERS AND MECHANISTIC INSIGHTS

Standard statistical analysis (e.g., t test, Wilcoxon test) are used to determine significance or fold change for peaks (or features) between and among study

phenotypes to point to potential biomarkers. Supervised and unsupervised multivariate statistics are used to identify patterns of signals (or suites of markers) and their rank order of importance in defining study phenotypes. These signals are annotated using public databases or identified through signal matching to in-house physical standards libraries.

One of the biggest challenges in untargeted metabolomics is related to the large number of signals (tens of thousands) that are detected in a single biospecimen, whereas only hundreds are identified in any one study. By mapping annotated and identified metabolites to biochemical pathways, perturbations in metabolism that associate with the study phenotypes can be suggested and integrated through pathway mapping and literature curation to derive genetic links. Stable isotopes are used to confirm specific metabolic pathways, metabolic flux, and underlying mechanisms (Fan and Lane, 2011). Metabolome-wide association studies (Gieger et al., 2008) and metabolite quantitative trait loci mapping (Hedjazi et al., 2015) are used in large-scale epidemiology studies.

EXAMPLE STUDY 1: BIOMARKERS OF PLACENTAL ABRUPTION AND THE LINK TO GENETIC POLYMORPHISMS

Placental abruption (PA), an ischemic placental disorder that results from premature separation of the placenta before delivery, is reported to occur in 1% of all pregnancies. PA has severe health consequences for the mother (e.g., maternal hemorrhagic shock) and neonate, including preterm birth and death. There is no universally accepted diagnosis for PA. The earliest but nonspecific symptoms include vaginal bleeding or abdominal pain; they appear for some women in the third trimester of pregnancy. This study was conducted to reveal biomarkers in the second trimester of pregnancy that could serve as predictors for third-trimester PA.

For this study, serum samples collected from women at approximately 16 weeks of gestation were selected from a cohort of women who were delivered at the Swedish Medical Center, Washington. At 16 weeks of gestation, no women presented with signals of pregnancy complication. Half of the samples selected from the biorepository were from women who had PA (sonographic diagnosis) in the third trimester and half were from women who had a normal vaginal delivery. Nine metabolites significantly differentiated the case–control groups, including acylcarnitines, amino acids, biogenic amines, and glycerophospholipids (Gelayer et al., 2016).

Logistic regression was used to model the probability of PA in the third trimester based on serum biomarkers in the second trimester. The area under the receiver

operator curve (AUC) was used to evaluate model performance. The AUC for the model with only metabolites in the second trimester (before any pregnancy complication was apparent) was 0.68; that for the model using only pregnancy vaginal bleeding was 0.63. The AUC for the model with both vaginal bleeding and metabolites was 0.76.

The probability of PA increased with an increase in acylcarnitines and a decrease in phosphatidylcholine. These metabolites arise from pathways that are related through diacylglycerol (DAG). DAG is transformed to 2-arachidononylglycerol (2-AG), and 2-AG is converted to prostaglandin glycerol esters via prostaglandin-endoperoxide synthase-2 (PTGS-2). PTGS-2 is a major target of nonsteroidal antiinflammatory drugs, including aspirin, which is a current treatment (at low doses) for pregnant women who are at high risk for adverse pregnancy outcome. Vaswani et al. demonstrated that PTGS-2 is differentially expressed in the aging placenta and suggested a role for PTGS-2 in preterm labor (Vaswani et al., 2015). Our findings also suggest a role for PTGS-2 in the pathogenesis of abruption.

There is a wealth of literature that describes the role of choline in women's health (Council, 2015), and whole-genome expression microarray analysis has associated higher maternal choline levels with lower concentrations of fms-like tyrosine kinase-1, a kinase known to disable proteins that cause blood vessel growth (Jiang et al., 2013). Furthermore, phosphatidylethanolamine-*N*-methyltransferase (PEMT) rs12325817 has been identified as one of three gene variants known to increase dependence on dietary choline (Silver et al., 2015). PEMT converts phosphatidylethanolamine to phosphatidylcholine. Taken together, these results point to the need to measure PEMT as a susceptibility factor for PA and to conduct a clinical trial using choline or phosphatidylcholine for intervention against PA.

EXAMPLE STUDY 2: BIOMARKERS OF RESPONSE TO WEIGHT-LOSS INTERVENTION FOR ADOLESCENT OBESITY, TYPE 2 DIABETES, AND GENETIC LINKS

Some individuals who enter intervention trials for obesity have clinically significant weight loss (a positive response to treatment) whereas others do not. Noninvasive biomarkers are needed to predict, before intervention, individuals who would have a positive response, so that the biomarkers could be used for stratification. This would ensure that people who were clearly not suited for the trial would not be enrolled. In addition, noninvasive biomarkers are needed to monitor the response to intervention during the trial

and provide insights into underlying mechanisms of weight loss. A study (Pathmasiri et al., 2012) used metabolomics to compare the metabolotypes of first morning void urine obtained from obese adolescents (body mass index [BMI] > 95th percentile) enrolled in a non-drug related healthy lifestyle program. Urine and psychosocial and health-behavior inventories were collected before intervention initiated (baseline) and during the intervention program. Subjects with a decrease in BMI (<0.5 BMI units) between baseline and the end of the 3-week program were categorized as nonresponders. Multivariate statistics of urinary metabolomics data demonstrated a unique pattern for intervention in the responders versus nonresponders. Psychosocial variables of improvement in self-esteem and depression did not correlate with weight loss. Individuals who did not have a positive response to intervention had higher levels of branch chain amino acids (BCAA) compared with those who had a positive response. This study in obese adolescents is consistent with those in the adult population showing that higher levels of BCAAs are strongly associated with increased risk for type 2 diabetes (Wang et al., 2011).

Metabolomics investigations of BCAAs (e.g., leucine, isoleucine, valine, 2-aminoisobutyric acid) led to an investigation of the genetic predisposition of impaired metabolism in type 2 diabetes. In a genome wide-study (Lotta et al., 2016) of over 16,000 individuals, BCAA levels were significantly associated with a genome signal near (21 kb upstream) the PPM1K gene, which encodes protein phosphatase, Mg²⁺/Mn²⁺ 1K (3.1.3.16), an activator of mitochondrial branched-chain α -keto acid dehydrogenase (EC 1.2.4.4), which in turn is a rate-limiting enzyme of BCAA catabolism. Type 2 diabetes case (>47,000)–control (>260,000) analysis significantly ($P < 1E-5$) associated 1 standard deviation in amino acid concentrations with an increased risk for developing type 2 diabetes (odds ratio = 1.44–1.84 for isoleucine, leucine, and valine).

EXAMPLE STUDY 3: METABOLOMICS TO INFORM NUTRITIONAL INTERVENTION FOR DYSFUNCTIONS GENE VARIANTS ARE NOT WELL-UNDERSTOOD

Although autism has a strong genetic component in some individuals, known genetic variants account for only a small portion of diagnosed patients. Metabolomics has been used to reveal perturbations in levels of metabolites for children diagnosed with autism (cases) compared with controls (Wang et al., 2016), with the goals of revealing biomarkers for disease diagnosis and identifying nutritional interventions. Metabolomics was conducted using blood collected from a Chinese

Han population (aged 3–6 years) in two phases: discovery (136 subjects) and then validation (200 subjects).

In this study, 11 metabolites were validated as being associated with autism, and multiple logistic regression modeling showed both docosahexaenoic acid (DHA) and sphingosine 1-phosphate (S1P) were most consistently associated with autism. If validated in additional larger samples, DHA and S1P could be used for clinical diagnosis of autism or clinical evaluation of treatment in patients with autism. Lower levels of DHA and polyunsaturated fatty acids in autism were previously reported, and clinical trials showed that supplementation with both DHA and 9,10-epoxyoctadecenoic acid (also validated in this study) were efficacious in reducing behavioral patterns in children associated with autism. Prior investigations also showed a role for disruptions in sphingomyelin metabolism in development of the white matter of the brain. Metabolites associated with autism in both the discovery and validation phases include uric acid, phospholipids, and long chain fatty acids. Validation of these of these metabolites in additional cohorts is needed to inform the use of these compounds in diagnosis or treatment. Omega-3 long chain polyunsaturated fatty acid supplementation (*n*-3 LCPUFA) is used to manage autism spectrum disorder (ASD). Analysis of case–control and randomized clinical trials of supplementation and ASD (Mazhery et al., 2017) showed that populations with ASD had lower *n*-3 LCPUFA status, which is associated with improvement of some ASD symptoms.

FUTURE PERSPECTIVES

Technological advances in mass spectrometry have enabled the detection of tens of thousands of signals in complex biological systems that comprise the endometabolome and exometabolome. Computational tools, databases, and big data analytics are being used to derive the complete annotation and identification of these signals. This will enable the identification of new molecules, classes of compounds, and metabolic pathways not previously considered in studies of human health. Public databases are being established for metabolomics, genomics, microbiome, and proteomics that are applicable to a wide range of disease and therapeutic areas. Metabolomics is being integrated with these other –omics data streams to take a systems biology approach to understanding health and wellness. The future of precision medicine and precision nutrition will involve fusing integrated –omics with study designs specific to revealing nutritional targets for prevention and intervention at the time when the metabolotype deviates from states of wellness.

References

- Corbin, K.D., Zeisel, S.H., 2012. Choline metabolism provides novel insights into nonalcoholic fatty liver disease and its progression. *Curr Opin Gastroenterol* 28 (2), 159–165.
- Council, C.I., 2015. Choline Information Council; Research Shows Choline Plays Critical Role in Pregnancy and Women's Health. *Women's Health Weekly*. Serial Search Publications, Atlanta, p. 211.
- Fan, T.W., Lane, A.N., 2011. NMR-based stable isotope resolved metabolomics in systems biochemistry. *J Biomol NMR* 49 (3–4), 267–280.
- Fiamoncini, J., Rundle, M., Gibbons, H., Thomas, E.L., Geillinger-Kastle, K., Bunzel, D., Trezzi, J.P., Kiselova-Kaneva, Y., Wopereis, S., Wahrheit, J., Kulling, S.E., Hiller, K., Sonntag, D., Ivanova, D., van Ommen, B., Frost, G., Brennan, L., Bell, J., Daniel, H., 2018. Plasma metabolome analysis identifies distinct human metabolotypes in the postprandial state with different susceptibility to weight loss-mediated metabolic improvements. *FASEB J* 32 (10), 5447–5458.
- Gelaye, B., Sumner, S.J., McRitchie, S., Carlson, J.E., Ananth, C.V., Enquobahrie, D.A., Qiu, C., Sorensen, T.K., Williams, M.A., 2016. Maternal early pregnancy serum metabolomics profile and abnormal vaginal bleeding as predictors of placental abruption: a prospective study. *PLoS One* 11 (6), e0156755.
- Gieger, C., Geistlinger, L., Altmaier, E., Hrabé de Angelis, M., Kronenberg, F., Meitinger, T., Mewes, H.-W., Wichmann, H.E., Weinberger, K.M., Adamski, J., Illig, T., Suhre, K., 2008. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. *PLoS Genet* 4 (11), e1000282.
- Hedjazi, L., Gauguier, D., Zalloua, P.A., Nicholson, J.K., Dumas, M.E., Cazier, J.B., 2015. mQTL.NMR: an integrated suite for genetic mapping of quantitative variations of (1)H NMR-based metabolic profiles. *Anal Chem* 87 (8), 4377–4384.
- Jiang, X., Bar, H.Y., Yan, J., Jones, S., Brannon, P.M., West, A.A., Perry, C.A., Ganti, A., Pressman, E., Devapatla, S., Vermeylen, F., Wells, M.T., Caudill, M.A., 2013. A higher maternal choline intake among third-trimester pregnant women lowers placental and circulating concentrations of the antiangiogenic factor fms-like tyrosine kinase-1 (sFLT1). *FASEB J* 27 (3), 1245–1253.
- Lotta, L.A., Scott, R.A., Sharp, S.J., Burgess, S., Luan, J., Tillin, T., Schmidt, A.F., Imamura, F., Stewart, I.D., Perry, J.R., Marney, L., Koulman, A., Karoly, E.D., Forouhi, N.G., Sjogren, R.J., Naslund, E., Zierath, J.R., Krook, A., Savage, D.B., Griffin, J.L., Chaturvedi, N., Hingorani, A.D., Khaw, K.T., Barroso, I., McCarthy, M.L., O'Rahilly, S., Wareham, N.J., Langenberg, C., 2016. Genetic predisposition to an impaired metabolism of the branched-chain amino acids and risk of type 2 diabetes: a Mendelian randomisation analysis. *PLoS Med* 13 (11), e1002179.
- Mazahery, H., Stonehouse, W., Delshad, M., Kruger, M.C., Conlon, C.A., Beck, K.L., vonHurst, P.R., 2017. Relationship between Long Chain n-3 Polyunsaturated Fatty Acids and Autism Spectrum Disorder: Systematic Review and Meta-Analysis of Case-Control and Randomised Controlled Trials. *Nutrients* 9 (2).
- Muoio, D.M., Newgard, C.B., 2006. Obesity-related derangements in metabolic regulation. *Annu Rev Biochem* 75, 367–401.
- Pathmasiri, W., Pratt, K.J., Collier, D.N., Lutes, L.D., McRitchie, S., Sumner, S.C.J., 2012. Integrating metabolomic signatures and psychosocial parameters in responsiveness to an immersion treatment model for adolescent obesity. *Metabolomics* 8, 1037–1051.
- Silver, M.J., Corbin, K.D., Hellenthal, G., da Costa, K.A., Dominguez-Salas, P., Moore, S.E., Owen, J., Prentice, A.M., Hennig, B.J., Zeisel, S.H., 2015. Evidence for negative selection of gene variants that increase dependence on dietary choline in a Gambian cohort. *FASEB J* 29 (8), 3426–3435.
- Stewart, D.A., Dhungana, S., Clark, R.F., Pathmasiri, W.W., McRitchie, S.L., Sumner, S.J., 2015. Omics technologies used in

- systems biology. In: Fry, R. (Ed.), *Systems Biology in Toxicology and Environmental Health*. Academic Press, Waltham, MA, USA, pp. 57–84.
- Sumner, S.C.J., Pathmasiri, W., Carlson, J.E., McRitchie, S.L., Fennell, T.R., 2018. Metabolomics. In: Smart, R., Hodgeson, E. (Eds.), *Molecular and Biochemical Toxicology*. John Wiley & Sons Inc., Hoboken, NJ, USA.
- Vaswani, K., Chan, H.W., Peiris, H.N., Nitert, M.D., Bradley, R.J., Armitage, J.A., Rice, G.E., Mitchell, M.D., 2015. Gestation related gene expression of the endocannabinoid pathway in rat placenta. *Mediat Inflamm* 2015, 850471.
- Wang, H., Liang, S., Wang, M., Gao, J., Sun, C., Wang, J., Xia, W., Wu, S., Sumner, S.J., Zhang, F., Sun, C., Wu, L., 2016. Potential serum biomarkers from a metabolomics study of autism. *J Psychiatry Neurosci* 41 (1), 27–37.
- Wang, T.J., Larson, M.G., Vasan, R.S., Cheng, S., Rhee, E.P., McCabe, E., Lewis, G.D., Fox, C.S., Jacques, P.F., Fernandez, C., O'Donnell, C.J., Carr, S.A., Mootha, V.K., Florez, J.C., Souza, A., Melander, O., Clish, C.B., Gerszten, R.E., 2011. Metabolite profiles and the risk of developing diabetes. *Nat Med* 17 (4), 448–453.

This page intentionally left blank

Metagenomics

Isabel Moreno-Indias^{1,2}, Francisco J. Tinahones^{1,2}

¹Unidad de Gestión Clínica de Endocrinología y Nutrición, Laboratorio del Instituto de Investigación Biomédica de Málaga (IBIMA), Hospital Universitario de Málaga (Virgen de la Victoria), Málaga, Spain; ²Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y la Nutrición (CIBEROBN), Madrid, Spain

BACKGROUND

Each human intestine harbors trillions of bacteria as well as other microorganisms such as bacteriophages, viruses, fungi, and archaea, which constitute a complex and dynamic ecosystem living in symbiosis with humans throughout our lifetime. This ecosystem is referred to as gut microbiota, although the term “gut microbiome” is also used without distinction. Bacteria represent almost 99% of this gut microbiota, so most studies focus on these microorganisms.

The gut microbiota is related to the promotion of health and the initiation and maintenance of different gastrointestinal and non-gastrointestinal diseases. In this manner, human-associated microbial communities are implicated in a wide range of diseases. Changes in the composition of human gut microbiota are associated with the development of chronic diseases including type 2 diabetes, obesity, and colorectal cancer, among others (Young, 2017). Gut microbiota encompassing trillions of bacteria is considered a microbial organ with a key role within the human organism. Members of the microbiota communicate among themselves and exchange signals with the host's systems, participating in the body homeostasis. Gut microbiota takes part in multiple functions within the body: in the metabolism with roles in synthesizing vitamins, digesting compounds, fermenting nondigestible substances, producing energy, protecting functions in stimulating the immune system, antimicrobial effects, creating a physical barrier against pathogenic bacteria, and so on. Along this line, study of the gut microbiome community has led to important research and renewed scientific efforts to understand the whole functionality of the human body.

The importance of the gut microbiota was hypothesized a long time ago, but advances in the technology

used to identify and analyze the population of microorganisms were necessary. Traditionally, bacteria were studied by culture-based methods, but the difficulty of culturing most anaerobically living commensal gut microorganisms led to the development of culture-independent methods. This rapid development and, most important, the substantial cost reduction of next-generation sequencing (NGS) or high-throughput sequencing have revolutionized the field of microbial ecology. Thus, this technology has led to the establishment of the field of metagenomics. Metagenomics refers to the direct genetic analysis of the genomes contained within a sample: that is, the functional and sequence-based analysis of the collective genomes of a sample. However, many researchers have extended this general term within microbiota studies to the study of certain genes of interest, such as the 16S ribosomal DNA gene. The powerful combination of genome sequencing and bioinformatics-driven analysis of sequence data has transformed our knowledge about how bacteria function, evolve, and interact with each other, with their hosts, and with the environment, providing new avenues of inquiries and advances for translational impact.

HISTORY OF METHODS USED

Advances in technology have primarily driven the current revolution in microbiology, particularly in the study of the gut microbiota. The development of culture-independent methods has provided the scientific community with the chance to deal with new hypothesis and paradigms.

Once the culture methods were not a consideration, molecular methods burst onto the scene, helped by the design of universal 16S ribosomal RNA gene

polymerase chain reaction (PCR) primers for the analysis of microbes. Thus, the use of quantitative PCR and the development of arrays with these primers kicked off microbiota studies. When the semiquantitative technique of denaturing gradient gel electrophoresis was shown to be an effective method for comparing microbiota compositions, it became the reference standard for the rapid initial screening of bacterial communities of initial unknown composition before NGS methods were completely developed.

Thus, high-throughput sequencing or NGS reached bacteriology in the first decade of the 2000s, leading to multiple platforms for scientists (Kozłowska et al., 2019), accompanied by new bioinformatics approaches. Briefly, the first of the technologies that revolutionized genomics and metagenomics was pyrosequencing, led by the 454-sequencing platform from Roche. The principle of this technology was simple: a one-by-one nucleotide addition cycle, in which the pyrophosphate released from the DNA polymerization reaction is transformed in a luminous signal detected by the machine and translated into nucleotide sequences. However, despite the revolutionary nature of this technology for metagenomics, it is now obsolete. An analogous technology was the ion torrent platform. The Ion Torrent Personal Genome Machine, and the Ion S5 (from Thermo Fisher Scientific) are able to detect changes in hydrogen potential generated each time a proton is released after a nucleotide is added in the sequencing reaction. However, read length reduction is another approximation in some platforms to reduce sequencing costs. This is the case of Illumina technology (Illumina, Inc.), which has become one of the most popular technologies owing to its low cost and high yield. The basis of Illumina chemistry is reversible-termination sequencing by synthesis with fluorescently labeled nucleotides. DNA fragments are attached and distributed in a flow cell, in which the sequencing reaction occurs by adding a labeled nucleotide. When a labeled nucleotide is incorporated into the samples, a laser excites its fluorescent molecule and the machine registers the signal.

Although these technologies are most commonly used for metagenome projects, the development of sequencing has continued to solve the known biases of these strategies and to offer a better trade-off among yield, cost, and read length. This is the case of other platforms such as PacBio (developed by Pacific Biosciences). PacBio is based on single-molecule real-time (SMRT) sequencing and is classified as third-generation sequencing. SMRT sequencing is built on two important innovations: zero-mode waveguides, which allow light to illuminate only the bottom of a well in which a DNA polymerase–template complex is immobilized, and phospholinked nucleotides, which enable an examination of the immobilized complex as the DNA

polymerase produces a completely natural DNA strand. PacBio sequencing offers much longer read lengths and faster runs than the other methods, but it is hindered by a lower throughput, higher error rate, and higher cost per base. Finally, the latest developed technology is Nanopore sequencing (Oxford Nanopore Technologies), with the portable device MinION, which provides long reads and fits in a pocket. Nanopore technology is based on nanoscale holes. DNA sequencing is based on the passage of a strand of DNA through the nanopore. The current is changed as the bases pass through the pore in different combinations.

Data obtained from second- or third-generation sequencing technologies have certain computational requirements for their analysis. The bigger the dataset generated, the higher computational resources and more complex bioinformatics analyses are necessary. Moreover, the strengths and weaknesses of second- and third-generation sequencing are complementary, which motivates researchers to integrate both techniques: the use of high-throughput and high-accuracy short read data from the second sequencing generation to correct errors in the long reads of the third sequencing generation, to reduce the required amount of the costly long-read sequence data and salvage the relatively long but more error-prone subreads.

METAGENOMIC APPROACHES

Experimental design is the most important part in a scientific study; it should fit each project's objectives to answer the biological question hypothesized. Different approaches permit contrasting results and it is important to keep in mind to choose the best approach. At this point, two metagenomics approaches for studying of microbiota will be introduced according to the target of sequencing.

Targeted Methods

The sequencing of specific microbial amplicons, mostly the bacterial 16S ribosomal RNA gene, has been extensively used for gut microbiota studies. Indeed, it is the technique that has most contributed to the establishment of the science of gut microbiota. Although this amplicon-based sequencing considers only one microbial gene, it is frequently grouped under the umbrella of metagenomics as a way to perform taxonomic, phylogenetic, or functional profiling.

Amplicon sequences can be directly matched to reference taxa or, more frequently, first grouped into clusters sharing a fixed level of sequence identity (often 97% or 99%) and known as operational taxonomic units

(OTUs). These OTUs are often more specific than genera, but in most cases they are less specific than species. Methods are able to control errors that amplicons can be clustered according to single-nucleotide differences over the sequenced gene region without constructing OTUs, with better resolution of the posterior taxonomy assignment. In this manner, sequences are clustered into amplicon sequence variants (ASVs). In either case, individual reads, OTUs, or ASVs are then assigned to specific taxa based on the sequence homology to a reference genomic sequence, a process known as binning. Thus, the main advantage of this approach is that the 16S ribosomal RNA (rRNA) gene provides the availability of several large databases of reference sequences and taxonomies, such as greengenes, SILVA, and the Ribosomal Database (RBD) Project. However, the species-level resolution might be unfeasible using this information alone, achieving the family or genus levels as the finest taxonomic levels in most cases.

Targeted approaches such as this were first restricted to taxonomic and phylogenetic profiling. However, the development of bioinformatics software has added functional profiling to this approach. The most well-known method is Phylogenetic Investigation of Communities by Reconstruction of Unobserved States, a bioinformatics software package designed to predict the metagenome functional content from targeted amplicons such as the 16S rRNA gene surveys (although it may also be used for full genomes). Moreover, the use of strategies to refine results could improve the disadvantages of this technique. Besides the mentioned ASVs, strategies such as oligotyping have been proposed to improve OTU resolution, using a sequence entropy-based approach to identify maximally informative sites within the 16S rRNA gene.

Untargeted Methods

Despite advances in amplicon sequencing, whole-genome shotgun (WGS) sequencing is the preferred method for studying microbial communities. In a different approach to amplicon sequencing, shotgun-sequencing methods are applied to millions of random genomic fragments sampled from a microbial community. Assembly does not reconstruct entire genomes, so it is necessary to bin genome fragments. Thus, assembly is the process of combining sequence reads into contiguous fragments of DNA called contigs, based on sequence similarity between reads. Then, the presence of paired reads in two different contigs allows those contigs to be linked into a larger noncontiguous DNA sequence called a scaffold. The goal of WGS is to represent each genomic sequence in one scaffold, although this is not always possible.

WGS is able to profile the bacterial community even to the strain level, owing to its ability to identify variation throughout microbial genomes. The genomic content of a community describes its functional potential, because of the knowledge of genes contained in each individual. Advantages of WGS sequencing are the ability to provide information on genome assembly, species-level taxonomy abundance, gene predication, and metabolic pathway reconstruction. However, each stage of the analysis is complicated by incomplete coverage, the high volume of data, the short length of reads, and intrinsic errors caused by parallelism sequencing. Thus, the data obtained will be able to decipher presence in the community (taxonomic profiling), as the 16S approach does, although with better accuracy, and the profile of the genes present in that population (functional profiling). This approach is ideal for novel discovery. Many new genes, phylotypes, regulators, and/or pathways are discovered using shotgun metagenome sequencing.

The combination of the lower cost and the decreased resolution of amplicon sequencing with the higher cost and increased resolution of WMS sequencing enables richer experimental designs. Along this line, two-stage study designs are a good option: beginning by surveying a large number of individuals using amplicon sequencing and then following up with a subset of samples (i.e., representative individuals of the group) using WMS sequencing. Another option could be time-course studies that combine amplicon sequencing, to survey a large number of time points, with WMS sequencing, to analyze a subset of time points (such as the first and last) in greater detail.

WORKFLOW IN METAGENOMIC STUDIES

For amplicon sequencing and whole metagenome sequencing, the workflow is similar (Song et al., 2018). Although these suggestions may be inferred for the rest of metagenome studies, we will focus on gut microbiome studies (Fig. 11.1):

1. Processing of biological samples. As in any other study, the experimental design is the critical point of the results. Selection of the least invasive method for the patient in collecting the samples will ensure the subject's cooperation during the whole study, especially if it will be a longitudinal study. Fortunately, in gut microbiota studies, feces are a good representative sample for colonic microbiota.
 - a. Sample collection: There are plenty of commercial products for collecting fecal samples. However, the critical step for gut

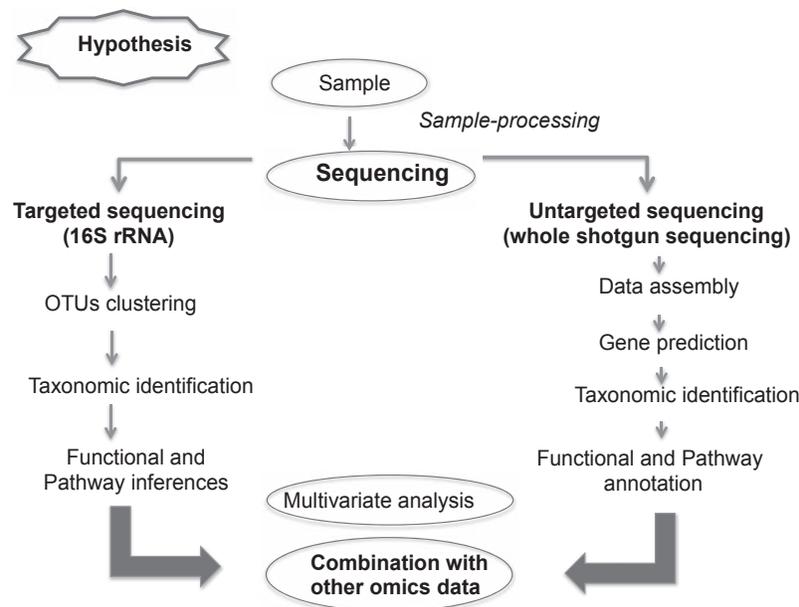


FIGURE 11.1 Workflow for microbiome analysis. *OTUs*, operational taxonomic units; *rRNA*, ribosomal RNA.

microbiota analysis is the correct storage of samples. Thus, freezing the sample at -80°C within 2 hours after sample collection is recommended. However, if this cannot be done, there are some collection commercial kits on the market with stabilizers that permit storage even at room temperature.

- b. DNA extraction: Variations in DNA extraction methods can have dramatic impacts on the results of metagenomic studies, especially in high-diversity communities such as those in feces. Although there is much controversy regarding the best way to extract DNA, most extraction methods use the same basic steps: cellular lysis, removal of non-DNA macromolecules, and collection of DNA. The researcher should choose the method that best fits with the outcome.
2. Sequencing: After the successful extraction of DNA, the proper selection of the sequencing approach, targeted or nontargeted, and the type of sequencer will depend on the goal of the researcher. Surprisingly, commonly used sequencers have different error rates and patterns, but their effects on taxonomic and functional composition are minimal. Balancing sequencing depth with the number of samples per sequencing run depends on the biological question and the complexity of the community. The sampling effort generally depends on variations in the microbial communities to be compared. For communities that share great similarity, deeper sequencing is needed to distinguish treatment effects on microbial communities.
3. Bioinformatics data analysis tools: The large amount of data produced in the high-throughput analysis requires sophisticated analysis tools. Thus, direct manipulation of the data is no longer feasible. Standardization and metadata access are imperative to reduce the magnitude of different experimental protocols and data analysis methods. However, the magnitude of these effects is often small relative to biological variations.
 - a. For targeted amplicon sequences such as 16S rRNA, multiple friendly tools are available for nonexpert bioinformatics researchers. Of those, Quantitative Insights Into Microbial Ecology (QIIME) and its new release, QIIME2, as well as Mothur are among the most used because of the good support they offer, with multiple tutorials and good and quick feedback from the developers. However, in large laboratories, the existence of bioinformaticians who implement and automatize their own pipelines is necessary.
 - b. For whole-metagenome studies, the pipeline is complicated. Early attempts at metagenomic data assemblies used tools initially implemented for single-genome data assemblies. However, assembly tools have significantly evolved since then and multiple tools have been specifically designed to assemble samples containing multiple genomes. Gene prediction, taxonomic identification, functional and pathway annotation, and finally data analysis will be assessed using different strategies (see Roumpeka et al. (2017) for a resume of available software; every month, new programs are being developed).

One of the greatest challenges in metagenomics is identifying significantly different features between communities. The large number of available information requires good knowledge about statistical tests. Moreover, it may be difficult to understand the biological importance of results. Thus, multidisciplinary teams make it possible to deal with the biological hypothesis in a deeper and more reliable way.

OTHER MICROORGANISMS

Eukaryotes have important roles in almost all ecological niches on the earth; presumably, this also happens within the human body. However, the study of these organisms by NGS techniques has been challenging because they are not well-described in sequence databases. This lack of reference eukaryotic genomes is due in part to the difficulty of their genome assembly and annotation. Moreover, the low abundance of eukaryotes compared with bacteria, as well as avoiding amplification of the DNA of the host, are other challenges. However, targeted approaches have used the 18S rRNA gene in a model similar to that of the 16S rRNA gene, and some of the software used for 16S rRNA gene analysis, such as QIIME, included tools to analyze these amplicons.

Viruses are also present in the human microbiome. A typical healthy human carries an abundance of viral particles at around 10^{12} , consisting mainly of bacteriophages and, to a minor extent, eukaryotic viruses. Similar to bacteria, the virome of adults seems to be stable over time. However, the lack of a gene that is universally present in all viruses means that amplicon-based studies cannot be used to characterize the virome. Notwithstanding, bacteriophages have critical roles in the microbiome ecosystem, and more efforts should be made to study them.

CURRENT GUT MICROBIOTA STATUS

The amount of variability in the microbiota within human subjects is immense; big projects such as the Human Microbiome Project and MetaHIT as well as smaller ones have done much to define this variability. However, causality has been reached in only a few cases. Germ-free mice, mice without microorganisms living in them, represent a model system to study the effect of gut microbes on host physiology. Experimental manipulations of murine models in gut microbiota research allow functional and mechanistic research into host–microbe interactions, which helps to assess causality in disease-associated alterations in gut microbiota

composition. Thus, this mouse model is increasingly being used to study the role and function of gut microbiota and its association with diseases.

Humanized gnotobiotic mice, which result from the inoculation of a human gut microbiota sample into germ-free mice, provide a powerful tool for gut microbiota studies because these models can recapitulate a large part of the human gut microbiota phylogenetic composition (Lundberg et al., 2016). This approach has been widely employed in many studies because it allows perturbations in a human-like system and is considered to be the reference standard for confirming associations and trying to prove causality in gut microbiota research. However, the host–microbe relationships in these humanized models do not necessarily reflect the real ones seen in humans, because the gut microbiota is transplanted into a host with which it has not coevolved.

Fecal microbiota transplants (FMTs) are the standard for revealing causality in humans. FMT is defined as the administration of a solution of fecal matter from a donor into the intestinal tract of a recipient to change the recipient's microbial composition directly and confer a health benefit (Camarota et al., 2017). A classic example is the case of persistent *Clostridium difficile* infections, for which this treatment reaches an average of 90% cure rate. However, much attention is being paid to other conditions, such as inflammatory bowel disease and even metabolic diseases.

The process is simple and usually involves selecting a donor without a family history of autoimmune, metabolic, and malignant diseases and screening for potential pathogens. The feces are then prepared by mixing with water, normal saline, or any other buffer, followed by a filtration step to remove particulate matter. The mixture can be administered through the upper digestive tract (nasogastric tube, nasojejunal tube, or esophagogastroduodenoscopy) or the lower digestive tract (colonoscopy or retention enema). Moreover, noninvasive approaches have been tested, such as using frozen or lyophilized pills.

Although it is still considered an experimental treatment and its effectiveness must be confirmed in every disease, its low cost, low risk, and high effectiveness positions FMT as a good therapy because of its repercussions in the gut microbiota.

PERSPECTIVES

As demonstrated earlier, DNA sequences are plenty of information. However, although the genomic content of a community describes what the community is capable of doing, it does not provide a direct measure of its community functional activity in a particular

time; most important, it is unable to discern between living, functionally active and dead cells. To understand fully the determinants of function, additional multiomic data types such as transcriptomics, proteomics, and metabolomics are needed. Each -omics technology has its strengths and weaknesses and must be selected based on the biological questions and objectives of the study.

Metatranscriptomics informs us of genes that are expressed by the community as a whole. Thus, metatranscriptomics sheds light on the active functional profile of a microbial community; in other words, the metatranscriptome provides a picture of the gene expression in a given sample at a given moment and under specific conditions by capturing the total messenger RNA (mRNA). After the rRNA is removed, microbial mRNA is converted into cyclic DNA (cDNA) and sequenced by standard methods. With appropriate barcoding of DNA and cDNA samples, metagenomic and metatranscriptomic sequencing could be carried out together. Combining metagenomics with metatranscriptomics can also reveal changes in functional activity in response to interventions, such as those with diets or drugs.

Metaproteomics is the study of all proteins expressed by organisms within an ecosystem at a specific time. Measuring protein abundance provides a more direct measure of the functional activity of a cell or community. Protein abundance can be determined in a high-throughput manner using next-generation proteomics (metaproteomics in the microbiome context). Proteomic methods rely on mass spectrometry-based shotgun quantification of peptide mass and abundance. Metaproteome datasets reveal information about microbial community structure, dynamics, and functional activities that are important for a better understanding of various community aspects, such as microbial recruiting, how participating organisms cooperate and compete for nutrient resources, and how these organisms distribute metabolic activities across the community.

Metabolomics is another rapidly expanding field of gut microbiota research that evaluates metabolites and other small molecules associated with the interrelationship of host-bacterial metabolism that has implications for health and disease. Metabolomics relies on chromatography techniques (such as high-performance liquid chromatography) to separate compounds, which are then identified and quantified using mass spectrometry. Short chain fatty acids provide a good example of the importance of metabolites in microbiota-host interactions.

Indeed, the use of multiple -omic data types better captures the functional activity of a community. The

addition and combination of multiple forms of -omic data (transcriptomics, proteomics, and/or metabolomics) with metagenomics in an integrated framework will improve the mechanistic models of the microbial community structure and function. Future metagenomics studies should go in this direction.

CONCLUSIONS AND REMARKS

Interest in the human microbiome has increased considerably. A significant driver has been the confirmation of some of important roles of the commensal microorganisms that comprise the human microbiota, which are not simply passengers in the host. However, the real driver has been available methodologies, mainly NGS methods, which have made metagenomic analysis more affordable for small to medium laboratories. However, compared with targeted 16S-based metagenomic sequencing, WGS generates exponentially more sequences that necessitate large storage requirements and produce large numbers of unknown species that demand more computational resources. In addition, metagenomic sequencing faces a crucial limitation: it is unable to distinguish between living-functional and dead bacteria. Longitudinal analysis could be advantageous for studying microbial community perturbations. Moreover, because microbiome science is growing rapidly, introduction of the study of new members such as viruses, bacteriophages, and fungi is convenient. On the other hand, an additional integrated framework of multiomic data could be (and should be) the future of microbiome analysis, the best approach to describe a microbial community fully, such as data on the levels of community RNA (transcriptomics), proteins (proteomics), and metabolites (metabolomics). Rapid and continuous advances of molecular high-throughput technologies and high-performance computational tools ensure that researchers will be able to model and predict the behaviors of microbial communities in the near future (Nayfach and Pollard, 2016).

Nevertheless, massive amounts of novel data on host-microbe interactions are coming to light. This complexity of metagenomic analyses is a fundamental limitation for explaining the microbiome roles within the host. Despite this, numerous health outcomes have been explained, which have been used for clinical diagnosis and treatment. Although there is enthusiasm about the opportunities that gut microbiota presents for new diagnostics, prognostics, and therapeutics, more studies need to be carried out.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at <https://doi.org/10.1016/B978-0-12-804572-5.00011-2>

Acknowledgments

This work was supported by a grant from the Fondo de Investigación Sanitaria of Instituto de Salud Carlos III and co-founded by Fondo Europeo de Desarrollo Regional–FEDER (PI15/01114). Isabel Moreno-Indias acknowledges support from the “Miguel Servet Type I” Program (CP16/00163).

References

- Cammarota, G., Ianiro, G., Tilg, H., Rajilić-Stojanović, M., Kump, P., Satokari, R., Sokol, H., Arkkila, P., Pintus, C., Hart, A., Segal, J., Aloï, M., Masucci, L., Molinaro, A., Scaldaferrì, F., Gasbarrini, G., Lopez-Sanroman, A., Link, A., de Groot, P., de Vos, W.M., Högenauer, C., Malfertheiner, P., Mattila, E., Milosavljević, T., Nieuwdorp, M., Sanguinetti, M., Simren, M., Gasbarrini, A., 2017. European consensus conference on faecal microbiota transplantation in clinical practice. *Gut* 66 (4), 569–580. <https://doi.org/10.1136/gutjnl-2016-313017>.
- Kozińska, A., Seweryn, P., Sitkiewicz, I., 2019. A crash course in sequencing for a microbiologist. *J Appl Genet* 60 (1), 103–111. <https://doi.org/10.1007/s13353-019-00482-2>.
- Lundberg, R., Toft, M.F., August, B., Hansen, A.K., Hansen, C.H.F., 2016. Antibiotic-treated versus germ-free rodents for microbiota transplantation studies. *Gut Microb* 7, 68–74.
- Nayfach, S., Pollard, K.S., 2016. Toward accurate and quantitative comparative metagenomics. *Cell* 166, 1103–1116.
- Roumpeka, D.D., Wallace, R.J., Escalettes, F., Fotheringham, I., Watson, M., 2017. A review of bioinformatics tools for bio-prospecting from metagenomic sequence data. *Front Genet* 8, 23.
- Song, E.J., Lee, E.S., Nam, Y.D., 2018. Progress of analytical tools and techniques for human gut microbiome research. *J Microbiol* 56, 693–705.
- Young, V.B., 2017. The role of the microbiome in human health and disease: an introduction for clinicians. *BMJ* 356, j831.

This page intentionally left blank

A Broader View on Omics and Systems Biology

Patricia Casas-Agustench^{1,2}, Cristina Andres-Lacueva^{1,2}

¹Biomarkers & Nutrimentalomics Laboratory, Department of Nutrition, Food Sciences and Gastronomy, XaRTA, INSA, Faculty of Pharmacy and Food Sciences, University of Barcelona, Barcelona, Spain; ²CIBER de Fragilidad y Envejecimiento Saludable (CIBERFES), Instituto de Salud Carlos III, Barcelona, Spain

Glossary

Big Data A large and complex data set and information that makes it very difficult to process using conventional database management tools.

Foodomics A discipline that involves the study of food and nutrition domains through the application and integration of advanced omics tools.

Holistic nutrition A revolutionary approach that integrates different nutriomics levels.

Network biology A developing area of research that recognizes that biological processes are controlled by a complex system-level network of molecular interactions.

Networks Graphs that do not necessarily show state changes.

Nutriogenetics The effect of genetic differences between individuals in their response to a specific dietary pattern, a functional food, or a supplement for a specific health outcome.

Nutrigenomics Study of the effects of foods and food constituents on gene expression.

Nutrimetabolomics The measurement of all metabolites to access the complete metabolic response of an organism to a nutritional stimulus.

Nutrimicrobiomics The understanding of the interactive effects of nutrients and the microbiome.

Nutriproteomics The large-scale analysis of the structure and function of proteins as well as of protein-protein interactions in a cell to identify the molecular targets of diet components.

Nutritranscriptomics The genome-wide study of mRNA expression levels in one cell or in a population of biological cells for a given set of nutritional conditions.

Pathways Graphs that show overall changes in state.

Precision nutrition An approach that integrates different kinds of big data to understand the complex relationships between food and human biology.

Systems biology The integrated approach for studying biological systems, at the level of cells, organs, or organisms, by measuring and integrating genomics, transcriptomics, proteomics, and metabolomics together with bioinformatics.

Systems thinking A set of synergistic analytical skills used to improve the capability of identifying and understanding systems, predicting their behaviors, and devising modifications to them in order to understand the complexity of health and disease processes.

INTRODUCTION

Novel nutritional research explores the link between diet and health aiming at healthy aging promotion and disease prevention through the innovative precision nutrition field. The precision nutrition approach integrates different kinds of big data to understand the complex relationships between food and human biology. In this scenario, precision nutrition opens the door to develop efficient nutritional interventions for optimal health status and future health outcomes by improving population health and reducing risk factors for lifestyle-related noncommunicable diseases, such as cardiovascular diseases, type 2 diabetes, cancer, obesity, or metabolic syndrome.

The foundation of knowledge created during the postgenomic era has had a huge impact in the application of omics sciences (genomics, epigenomics, transcriptomics, proteomics, metabolomics) to study the interaction between diet, lifestyle, and health. The integration of these omics with the nutrition science in the current era of big data offers unique opportunities to unravel the complex mechanisms underlying the health-modifying capacities of nutritional molecules with the revolutionary systems biology approach. How this field can be optimized to our understanding of lifestyle-related noncommunicable diseases and its nutritional management is a great challenge.

THE OMICS TOOLS: DESCRIPTION, APPLICATION, AND INTEGRATION

The omics technologies generate large amounts of data from gene sequencing and expression to protein

and metabolite patterns (Table 12.1). In 2009, the term *foodomics* was for the first time defined as a new discipline that involves the study of food and nutrition domains through the application and integration of advanced omics tools. Accordingly, foodomics involves the application and integration of advanced omics technologies to nutrition research, including nutrigenetics, epigenomics, nutrigenomics (nutritranscriptomics, nutriproteomics, and nutrimetabolomics) as well as microbiomics (nutrimicrobiomics), as foodomics has also been described as a part of the host-microbiota-exposome interplay. The integration of these powerful methods employed for the analysis of the genome and epigenome (DNA), the transcriptome (RNA), the proteome (protein), the metabolome (metabolites), and the microbiome (microbiota), together with biostatistics, chemometrics, and bioinformatics tools, allows the evaluation of complex biological systems, a more extensive analysis of the individual response to a nutritional intervention, and a global understanding of how food components influence health status (Fig. 12.1). The convergence of all these omics approaches is necessary to build a robust precision nutrition (Fig. 12.2). Therefore, main nutriomics research focuses on dissecting the molecular mechanisms that mediate the physiological effects of nutrients, foods, or dietary patterns, which provides a more reliable picture of real food intake, including the characterization of how foods are processed, absorbed, and metabolized in the organism. Understanding the mechanisms underlying these processes will define the cellular pathways involved and may lead

to the discovery of novel nutritional biomarkers. These metabolic signatures or metabolotypes will be used to identify the responders to certain dietary interventions for use in personalized nutrition. These biomarkers will facilitate the development of more efficient nutritional strategies and applications to achieve the goal of precision nutrition.

INTERINDIVIDUAL VARIABILITY AND OMICS TOOLS IN THE DEVELOPMENT OF NOVEL NUTRITIONAL BIOMARKERS

In the development of novel nutritional biomarkers or dietary exposure markers there are some biological and analytical considerations that must be taken into account. One important biological aspect is the interindividual variability, which is defined by the phenotypic differences between individuals that contribute to health status. The effect of dietary changes on phenotypes (e.g., blood pressure, plasma lipid or insulin measures, body weight, waist circumference, etc.) differs significantly between individuals. The precision nutrition approaches include the study of different factors such as genetics, gender, environmental features (e.g., eating habits, eating behavior, or physical activity), microbiota, and the metabolome in which the challenge lies in characterizing the phenotypic differences between individuals that contribute to health status. It is well recognized that interactions of genetic and epigenetic signatures with environmental factors, such as dietary intake or physical activity, play an important role in determining individual phenotypes (Casas-Agustench et al. 2014; de Gonzalo-Calvo et al. 2015). The characterization of genetic and epigenetic biomarkers may be useful as biomarkers of disease prevention, nutritional status or function as potential molecular targets that are modulated by dietary interventions.

In the investigation of new biological markers related to diet, metabolomics has also emerged as a powerful tool for the discovery of new nutritional biomarkers, both intake and effect. Metabolomics is the science that studies the metabolome (Fig. 12.3), which is the set of metabolites (defined as those intermediate molecules and metabolic products with a molecular weight of less than 1500 DA) present in a biological system (cell, tissue, or fluid). Two different fractions of the human metabolome are influenced by diet: (1) endogenous metabolism, which includes all the metabolites produced by the organism, and (2) the food metabolism, which includes all the external metabolites derived from dietary exposure (Scalbert et al. 2014). Nutrimetabolomics is the science that studies how the diet affects the whole metabolome. Food intake induces changes

TABLE 12.1 Definition of Omics Technologies and Molecules of Interest.

"Omics"	Definition	Molecules of interest
Genomics	Study of genes and their functions	DNA
Transcriptomics	Study of the complete set of RNA transcripts (mRNA) that are produced by the genome	RNA
Proteomics	Large-scale study of the set of proteins produced in an organism	Proteins
Metabolomics	Large-scale study of metabolites within cells, organisms, tissues, or biofluids	Metabolites (sugars, amino acids, lipids, hormones)
Microbiomics	Large-scale study of the microbes in the gut and body	DNA and metabolites

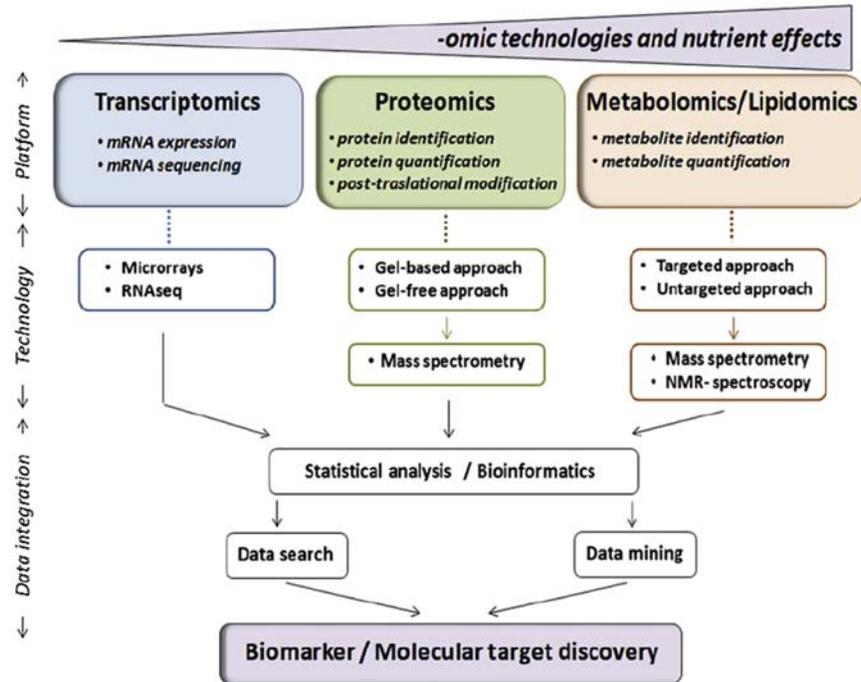


FIGURE 12.1 Integration of omics technologies. Source: Badimon L., Vilahur G. and Padro T., *Systems biology approaches to understand the effects of nutrition and promote health. Br J Clin Pharmacol.* 83 (1), 2017, 38–45.

in the metabolism of the organism, which can be evaluated by analyzing endogenous and exogenous metabolites in body fluids (e.g., blood or urine). These metabolites can be used as objective and precise biomarkers of food consumption and/or the effects of a dietary intervention. Food metabolome includes the set of

all metabolites derived from food intake, its absorption, and its biotransformation in the tissues or organs of the organism and the microbiota. In this sense, the application of the metabolomic approach has become a new strategy in obtaining new biomarkers related to the evaluation of the nutritional state of an individual, food

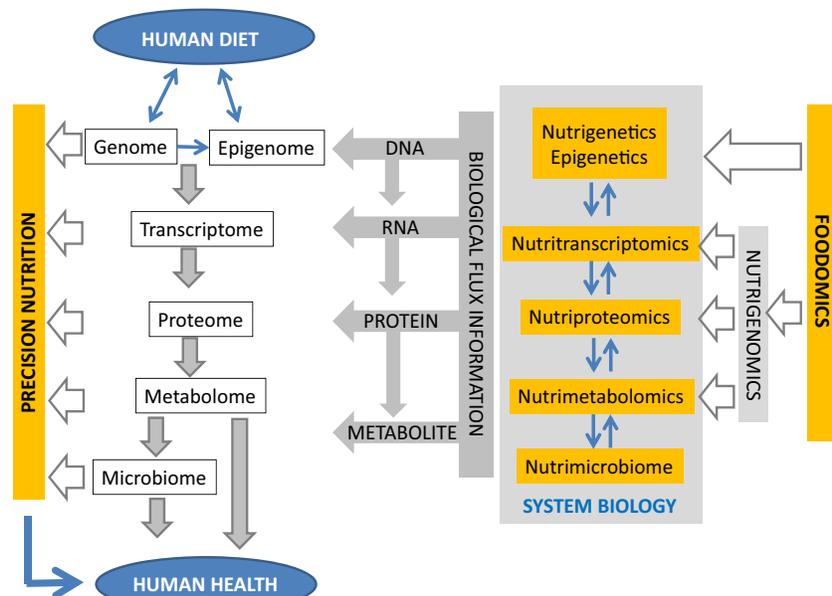


FIGURE 12.2 Role of foodomics in human health through the development of precision nutrition. Adapted from Arola-Arnal A., del Blas J.M., Caimari A., Crescenti A., Puiggròs F., Suárez M., Arola L.I., 2013. How does foodomics impact optimal nutrition?. In: Cifuentes, A. (Ed.). *Foodomics Advanced Mass Spectrometry in Modern Food Science and Nutrition*. Wiley, pp. 309.

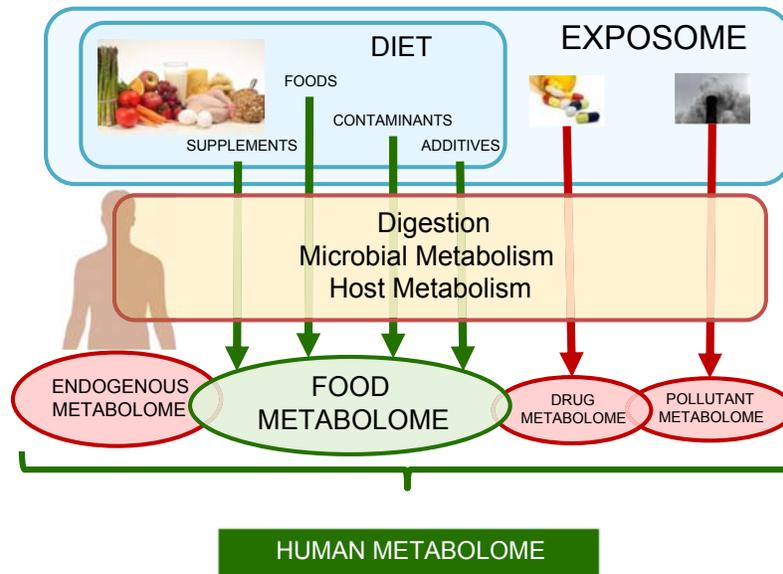


FIGURE 12.3 The human metabolome. Adapted from Scalbert, A., Brennan, L., Manach, C., Andres-Lacueva, C., Dragsted LO, Draper, J., Rappaport, S.M., van der Hooft, J.J., Wishart, D.S., 2014. The food metabolome: a window over dietary exposure. *Am J Clin Nutr* 99 (6), 1286–1308.

consumption, the biological consequences produced after a nutritional intervention, or the study of the metabolic mechanisms in response to the diet according to a certain metabolic phenotype. The food metabolome presents a high complexity and variability, since it is estimated that the foods we consume contain >25,000 different compounds, most of which will suffer various metabolic processes in the body. This particular characteristic turns the food metabolome into a very important source of information about the diet of individuals, whose characterization would allow monitoring of eating habits in an objective and precise manner, as well as study of how foods influence the risk of developing diseases (Scalbert et al. 2014).

DEFINING AN ACCURATE MEASUREMENT OF FOOD INTAKE INCLUDING OMICS TOOLS

An accurate measurement of food intake is the cornerstone of understanding the links between diet and health status. Dietary assessment has been traditionally performed using conventional methodologies of food surveys such as food frequency questionnaires, 24-h dietary recalls, or food records. Nevertheless, the accuracy of the dietary intake and nutritional status is frequently challenged due to the subjective nature of these dietary instruments. This limitation can be improved by the application of metabolomics to characterize dietary exposure. Furthermore, there are some factors not present in the traditional dietary assessment instruments that could misrepresent biomarker

measures of dietary intake. These factors include genetic variability (e.g., biological variation in nutrient absorption and metabolism, epigenetic variation, or gene-gene interactions), lifestyle/physiological factors (e.g., smoking, alcohol consumption, physical exercise, or influence of microbiota), dietary factors (e.g., nutrient bioavailability or nutrient-nutrient interactions), biological samples, and analytical methodology (Scalbert et al. 2014). Further research is necessary to address this issue and identify the best emerging dietary biomarkers. Therefore, as biomarkers cannot replace conventional dietary assessment methods, the use of conventional dietary instruments together with dietary biomarkers is considered the best strategy for tackling the complexity of dietary exposure fingerprinting (Fig. 12.4).

Individual biomarkers of dietary intake have been used to assess exposure to specific foods or food groups. However, this strategy has important limitations and only in some cases has it been successful. The more relevant limitations can be grouped into a wide distribution of food components (low specificity of biomarkers), high interindividual variation, and the microbiota metabolism, among others. Therefore, the use of a combination of food-derived metabolites, namely a multimetabolite biomarker panel (MBP), would be more likely to capture dietary exposure and improve the accuracy and precision of dietary assessment. The rationale behind the use of MBPs is that a wider range of metabolites would improve the measurement of dietary intake, capturing a broader perspective of the diet, and thereby giving a more complete coverage of dietary exposure. It opens a new framework in the research area of nutritional biomarkers. However, while almost all

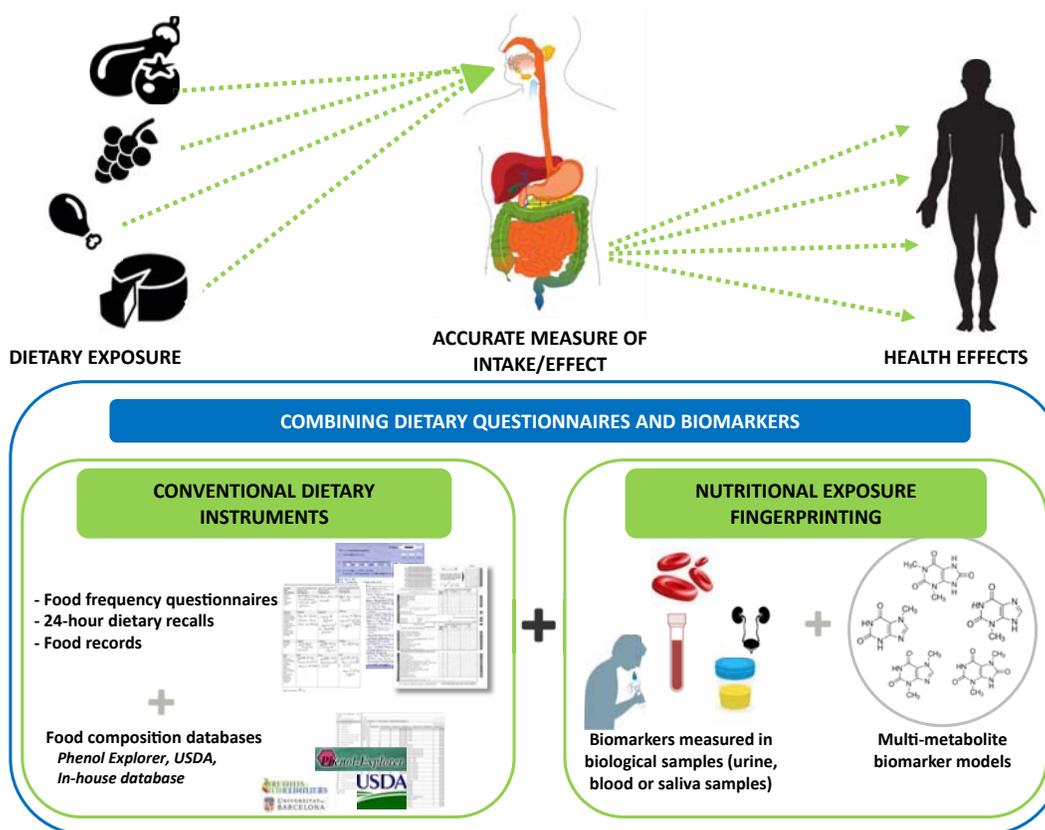


FIGURE 12.4 Schematic diagram for combining dietary questionnaires with biomarkers. Source: Garcia-Aloy, M., Rabassa, M., Casas-Agustench, P., Hidalgo-Liberona, N., Llorach, R., Andres-Lacueva, C., 2017. Novel strategies for improving dietary exposure assessment: multiple-data fusion is a more accurate measure than the traditional single-biomarker approach. *Trends Food Sci Technol* 69, pp. 220–229.

studies investigating dietary biomarkers have focused on single-candidate biomarkers, MBPs have remained practically unexplored. Therefore, future studies will face up to the complexity of evaluating the use of MBPs in combination with conventional dietary questionnaires to assess dietary exposure fingerprinting. These advances will enable more detailed information to be obtained about the associations between diet and health (Garcia-Aloy et al. 2017).

The assessment of dietary patterns to elucidate the beneficial effects in complex disease processes is increasing. The Mediterranean diet is considered a dietary pattern with beneficial effects on human health. As previously addressed, metabolomics is an interesting tool for assessing the nutritional status of an individual, as well as for studying the biological consequences or metabolic mechanisms following a nutritional intervention. The effect on urine metabolome of a long-term (1- and 3-year) dietary intervention with a Mediterranean diet (MD) plus either extra virgin olive oil or nuts as supplements compared to a low-fat diet intervention (LFD) was studied in a subcohort from the PREDIMED nondiabetic subjects using a metabolomic approach (high-throughput screening ^1H NMR spectroscopy) (Vazquez-Fresno et al., 2015). As shown in Fig. 12.5, the

one- and three-year intervention follow-ups showed a marked effect on urinary metabolomic phenotype or metabolome in the volunteers. MD groups had a different metabolic fingerprinting compared to baseline and the control (LFD) groups. The most prominent hallmarks of these changes concerning MD groups were related to the metabolism of carbohydrates (3-hydroxybutyrate, citrate, and cisaconitate), creatine, creatinine, amino acids (proline, N-acetylglutamine, glycine, branched-chain amino acids, and derived metabolites), lipids (oleic and suberic acids), and microbial cometabolites (phenylacetylglutamine and p-cresol). Otherwise, hippurate, trimethylamine-N-oxide, histidine and derivatives (methylhistidines, carnosine, and anserine), and xanthosine were predominant after LFD. The application of NMR-based metabolomics enabled the classification of individuals regarding their dietary pattern and highlights the potential of this approach for evaluating changes in the urinary metabolome at different time points of follow-up in response to specific dietary interventions.

Mass spectrometry was used to identify biomarkers of bread intake in a population in free living conditions. Bread is one of the most consumed foods in most dietetic patterns. This has caused its impact on human health to

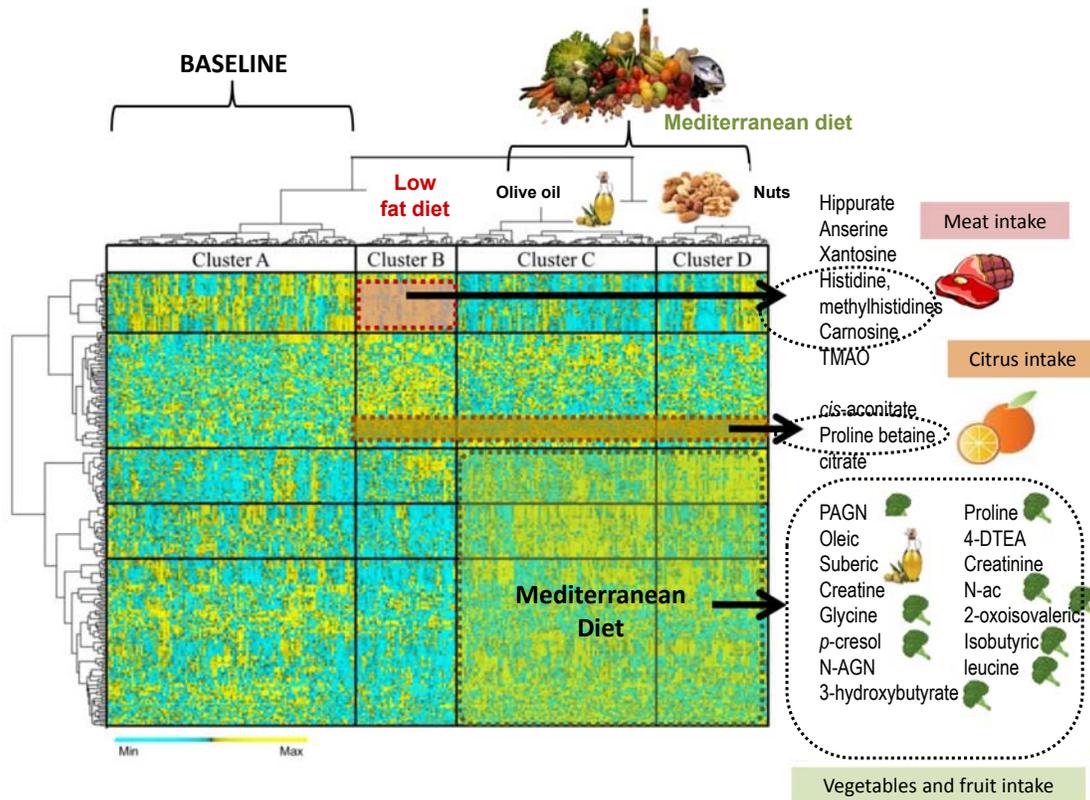


FIGURE 12.5 Metabolic pattern analysis after a Mediterranean diet intervention in a nondiabetic population. Adapted from Vázquez-Fresno, R., Llorach, R., Urpi-Sarda, M., Lupianez-Barbero, A., Estruch, R., Corella, D., Fitó, M., Arós, F., Ruiz-Canela, M., Salas-Salvadó, J., Andres-Lacueva, C., 2015. Metabolomic pattern analysis after mediterranean diet intervention in a nondiabetic population: a 1- and 3-year follow-up in the PREDIMED study. *J Proteome Res* 14 (1), 531–5400. The figure shows a two-way hierarchical clustering analysis (processed with PermutMatrix). Heat map representation of the clustered data matrix in which each colored cell represents the intensity of appropriate NMR signals, according to the color scale at the bottom of the figure. Rows: NMR signals (253, VIP > 1.5). Columns: urine samples baseline and Low-fat diet (1y + 3y); Mediterranean diet + Extra virgin olive oil (1y + 3y), and Mediterranean diet + Nuts (1y + 3y).

be of special interest to researchers. A nontargeted HPLC-q-ToF-MS approach, together with multivariate analysis (OSC-PLS-DA), analyzed the urine of 155 stratified individuals in three groups according to their usual bread consumption: nonbread users ($n = 56$), consumers of white bread ($n = 48$), and consumers of whole wheat bread ($n = 51$) (Fig. 12.6). The most differential metabolites included compounds derived from plant phytochemicals, such as benzoxazinoid metabolites and alkylresorcinols and compounds produced by the microbiota (enterolactones, hydroxybenzoic acid, and dihydroferopolic acid metabolites). The pyrroline, 3-indol carboxylic acid glucuronide, riboflavin, 2,8-dihydroxyquinoline glucuronide, and N- α -acetylcitrulline were also tentatively identified. The results of this study showed that a daily bread list has a significant impact on urinary metabolome, although it is evaluated in free living conditions. It was also shown that the predictive capacity of a combination of various biomarkers of dietary exposure is better than that of a single biomarker.

SYSTEMS BIOLOGY IN NUTRITION RESEARCH

Both previous studies explored the association between diet and health by using the metabolomics approach with NMR and mass spectrometry technologies, respectively. Exploring the molecular-level associations between diet and health using a multiomics analysis is complex. Emergence of systems biology has the potential to increase our understanding of how nutrition influences health status. Systems biology integrates the various components of living organism (genes, proteins, metabolites, and phenotypes) in order to obtain a holistic and quantitative understanding of how biological systems work and respond to environmental stimuli (e.g., food intake). Systems biology uses mathematical modeling and biological information to integrate all omics information levels in cells, organs, or organisms, to comprehensively characterize the physiology of living organisms. The revolutionary vision of systems thinking is required to understand the

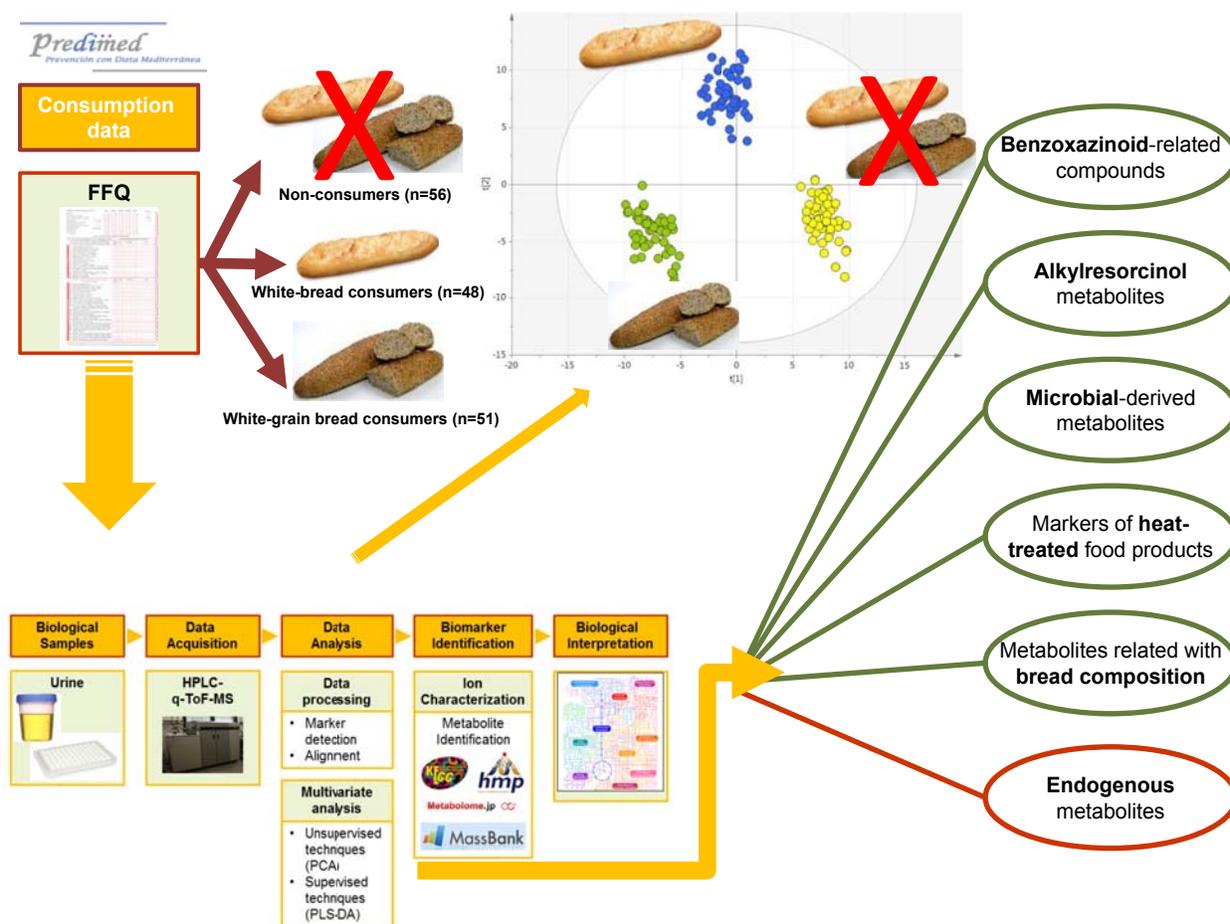


FIGURE 12.6 Nutrimetabolomics fingerprinting to identify biomarkers of bread exposure. Adapted from Garcia-Aloy, M., Llorach, R., Urpí-sarda, M., Tulipani, S., Salas-salvadó, J., Martínez-González, M.A., Corella, D., Fitó, M., Estruch, R., Serra-Majem, L., Andres-Lacueva, C., 2015. Nutrimetabolomics fingerprinting to identify biomarkers of bread exposure in a free-living population from the PREDIMED study cohort. *Metabolomics* 11, 155

complexity of all types of dynamic processes in the organism as a whole. In parallel, network biology recognizes that biological processes are controlled by a complex system-level network of molecular interactions. The network biology analysis is of extreme interest to develop solutions for reducing risk factors for lifestyle-related noncommunicable diseases. For example, cardiovascular diseases are complex, involving numerous biological entities from genes and small molecules to organ function. Data integration and network analysis has the potential to uncover novel players in cardiovascular disease and understanding its progression and prevention (Parnell et al., 2014).

It is well accepted that systems and network biology has the potential to increase our understanding of how nutrition influences metabolic pathways and homeostasis, how this regulation is disturbed in a diet-related disease, and to what extent individual genotypes contribute to such diseases. The interface between biological systems and nutritional as well as other environmental factors represents the next level of complexity

involving time-dependent interactions between nutrients, host metabolism, and gut microbiota that still have to be understood (van Ommen et al., 2008). Therefore, nutritional systems biology will provide a better understanding of the molecular mechanisms by which nutrients exert their effects, as well as the identification of new biomarkers (transcripts, proteins, or metabolites) of food consumption and health promotion (Fig. 12.7).

FUTURE PERSPECTIVES

The near future will bring more emphasis on three key aspects aiming to identify novel biomarkers foundational for precision nutrition: (1) using the high dimensionality of “omics” platforms, (2) the use of mathematical models for an integrative analysis of big data, and (3) the visualization of systems biology data. Given the human complexity and the elusive concept of health, there is still much research to be done in this field. However, the integration of these omics with the

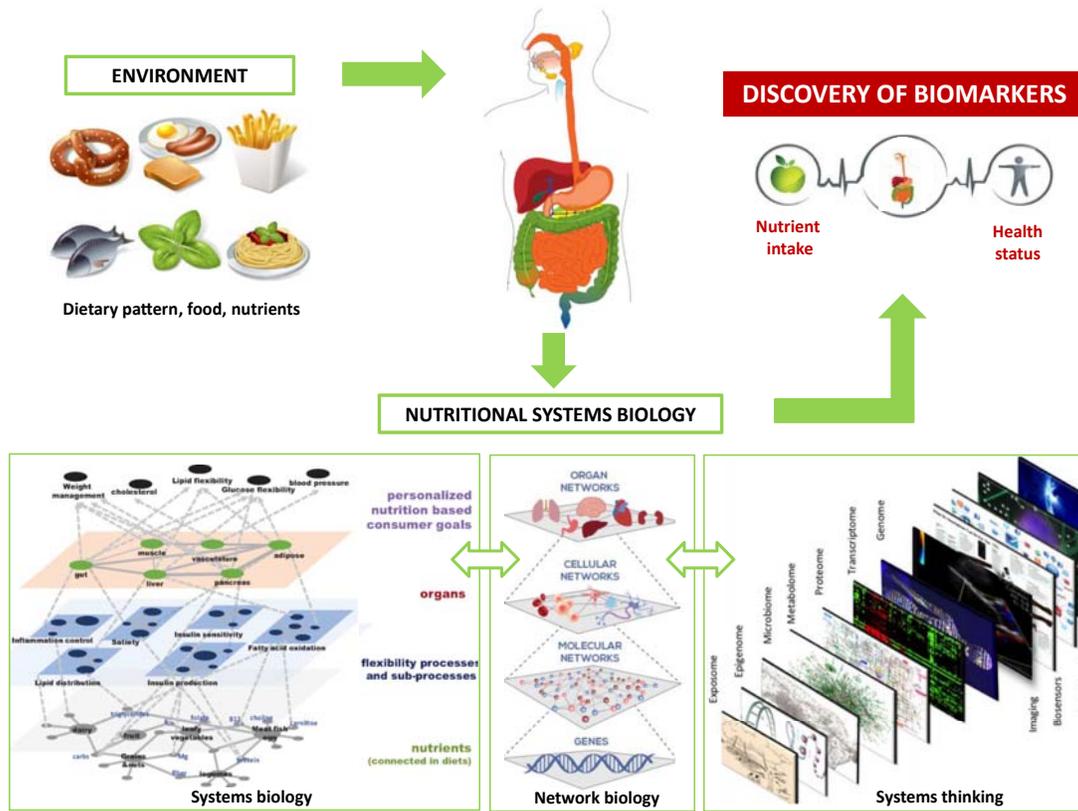


FIGURE 12.7 Discovery of biomarkers by nutritional systems biology. Adapted from van Ommen B., van den Broek T., de Hoogh I., van Erk M., van Someren E., Rouhani-Rankouhi T., Anthony J.C., Hogenelst K., Pasman W., Boersma A. and Wopereis S., *Systems biology of personalized nutrition*, *Nutr Rev* 75 (8), 2017, 579–599. Topol, E.J., 2014. *Individualized medicine from pre-womb to tomb*. *Cell* 157 (1), 241–253.

nutrition science in the current era of big data opens a very innovative area to understand complex biological systems and their role on human health.

References

- Arola-Arnal, A., del Blas, J.M., Caimari, A., Crescenti, A., Puiggròs, F., Suárez, M., Arola, LL, 2013. How does foodomics impact optimal nutrition? In: Cifuentes, A. (Ed.), *Foodomics Advanced Mass Spectrometry in Modern Food Science and Nutrition*. Wiley, p. 309.
- Badimon, L., Vilahur, G., Padro, T., 2017. Systems biology approaches to understand the effects of nutrition and promote health. *Br J Clin Pharmacol* 83 (1), 38–45.
- Casas-Agustench, P., Arnett, D.K., Smith, C.E., Lai, C.Q., Parnell, L.D., Borecki, I.B., Frazier-Wood, A.C., Allison, M., Chen, Y.D., Taylor, K.D., Rich, S.S., Rotter, J.I., Lee, Y.C., Ordovás, J.M., 2014. Saturated fat intake modulates the association between an obesity genetic risk score and body mass index in two US populations. *J Acad Nutr Diet* 114 (12), 1954–1966.
- de Gonzalo-Calvo, D., Dávalos, A., Montero, A., García-González, Á., Tyshkovska, I., González-Medina, A., Soares, S.M., Martínez-Cambor, P., Casas-Agustench, P., Rabadán, M., Díaz-Martínez, Á.E., Úbeda, N., Iglesias-Gutiérrez, E., 2015. Circulating inflammatory miRNA signature in response to different doses of aerobic exercise. *J Appl Physiol* (1985) 119 (2), 124–134.
- García-Aloy, M., Rabassa, M., Casas-Agustench, P., Hidalgo-Liberona, N., Llorach, R., Andres-Lacueva, C., 2017. Novel strategies for improving dietary exposure assessment: multiple-data fusion is a more accurate measure than the traditional single-biomarker approach. *Trends Food Sci Technol* 69, 220–229.
- García-Aloy, M., Llorach, R., Urpi-sarda, M., Tulipani, S., Salas-salvadó, J., Martínez-González, M.A., Corella, D., Fitó, M., Estruch, R., Serra-Majem, L., Andres-Lacueva, C., 2015. Nutrimetabolomics fingerprinting to identify biomarkers of bread exposure in a free-living population from the PREDIMED study cohort. *Metabolomics* 11, 155.
- Parnell, L.D., Casas-Agustench, P., Iyer, L.K., Ordovas, J.M., 2014. How gene networks can uncover novel CVD players. *Curr Cardiovasc Risk Rep* 8, 372.
- Scalbert, A., Brennan, L., Manach, C., Andres-Lacueva, C., Dragsted, L.O., Draper, J., Rappaport, S.M., van der Hooft, J.J., Wishart, D.S., 2014. The food metabolome: a window over dietary exposure. *Am J Clin Nutr* 99 (6), 1286–1308.
- Topol, E.J., 2014. *Individualized medicine from pre-womb to tomb*. *Cell* 157 (1), 241–253.
- van Ommen, B., Keijer, J., Kleemann, R., Elliott, R., Drevon, C.A., McArdle, H., Gibney, M., Müller, M., 2008. The challenges for molecular nutrition research 2: quantification of the nutritional phenotype. *Genes Nutr* 3 (2), 51–59.
- van Ommen, B., van den Broek, T., de Hoogh, I., van Erk, M., van Someren, E., Rouhani-Rankouhi, T., Anthony, J.C., Hogenelst, K.,

- Pasman, W., Boorsma, A., Wopereis, S., 2017. Systems biology of personalized nutrition. *Nutr Rev* 75 (8), 579–599.
- Vázquez-Fresno, R., Llorach, R., Urpi-Sarda, M., Lupianez-Barbero, A., Estruch, R., Corella, D., Fitó, M., Arós, F., Ruiz-Canela, M., Salas-Salvadó, J., Andres-Lacueva, C., 2015. Metabolomic pattern analysis after mediterranean diet intervention in a nondiabetic population: a 1- and 3-year follow-up in the PREDIMED study. *J Proteome Res* 14 (1), 531–540.

This page intentionally left blank

Study Designs in Genomic Epidemiology

José L. Santos¹, Fernando Rivadeneira²

¹Department of Nutrition, Diabetes, and Metabolism, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile; ²Department of Internal Medicine, Erasmus University Medical Centre, Rotterdam, The Netherlands

Glossary

Family aggregation study Type of study design used to evaluate the propensity of disease to cluster in families.

Classical twin study Type of study design used to disentangle the effects of genetics and the environment in family clustering. Based on a comparison of intraclass correlation coefficients of traits (for example, body mass index) in monozygotic versus dizygotic twins.

Association study General term often used to refer to the comparison of allele/genotype frequencies across groups of unrelated individuals (for example, affected versus unaffected cases from case-control or cross-sectional studies).

Genome-wide association study Association study in which simultaneous measurements of millions of simple genetic variants are used to evaluate gene-trait or gene-disease associations derived from linkage disequilibrium.

Linkage disequilibrium Coinheritance of alleles at two (or more) loci in a given haplotype that occurs more often in the population than what would be expected by chance.

Cohort studies Type of association study in which the incidence rate of new cases across defined groups (for example, genotypes) are compared with estimate absolute and relative risks for disease.

Case-only study Strategy to assess gene-environment interactions by testing association of gene variants with an environmental factor only in cases, under the assumption of gene-environment independence in the source population.

Case-parent triad design Type of family-based study design used to assess gene-disease association by selecting affected patients and their parents. Evaluates whether the apparent probability of transmission differs from the expected value of 0.5.

Mendelian randomization study Type of study design in which genetic variants are used as proxies for exposures of interest (many times, of nutritional nature), with the final purpose of establishing the causal relation between a given exposure influenced by gene variants and a health-related outcome.

Linkage analysis Approach that uses genetic markers in families to map genes causing diseases. Most effectively used in multigenerational pedigrees for finding monogenic causes of diseases.

assays and their application in the study of multiple biomedical traits. In addition to the importance of advances in DNA technology, it is clear that the choice of an appropriate study design represents a crucial step in deciphering genetic causes involved in health-related traits and diseases. Classical research on the observation of familial resemblance among relatives, such as twin and adoption studies, have been useful in assessing the genetic contribution to many diseases. In addition, classical epidemiologic studies based on large cohorts of unrelated subjects have been used in genome-wide association studies (GWAS) of complex multifactorial diseases. On the other hand, linkage studies in multigenerational families using genome-wide coverage of genetic markers were successful in the initial discoveries of many genomic regions and genes related to monogenic diseases. Moreover, the discovery of genetic alterations in monogenic diseases, either single mutations or structural variants, has been greatly accelerated with the advent of next-generation sequencing and high-throughput microarray genotyping technologies. In this chapter, we will briefly describe the main types of study designs used to discover novel genes and gene variants or to assess the impact of known genetic variants in human diseases.

FAMILIAL AGGREGATION OF HEALTH-RELATED MULTIFACTORIAL TRAITS AND DISEASES

Common multifactorial diseases result from the interactions of multiple genes with environmental factors. Contrary to the well-defined family aggregation patterns of monogenic diseases (for example, autosomal recessive, dominant, sex-linked), the familial recurrence risk of multifactorial diseases can be estimated as the frequency

INTRODUCTION

Genomic epidemiology is an evolving field that has had a rapid expansion in the scientific literature, partly owing to the increased use of high-throughput genetic

of disease in relatives (of type R, e.g., siblings) of affected individuals. In *family aggregation studies*, the relative recurrence risk ratio (also called Lambda), is defined as the ratio of the familial recurrence risk to the risk for disease in the general population. Values for sibling recurrence risk ratio (Lambda-sibs) in multifactorial diseases are relatively high in type 1 diabetes (~15) but much lower for most other complex diseases. In this context, it has been proposed that family history can be used as an adequate tool to identify persons at increased risk for common chronic diseases, given that this variable summarizes both genetic susceptibility and shared environment that cluster in families. As an example, the Avon Longitudinal Study of Parents and Children cohort identified that obesity in both parents provides an odds ratio > 10 in relation to obesity occurring in the child.

Twin and adoption studies are the classical study design strategies used to disentangle the effects of genetics and the environment in family clustering of diseases and traits. In adoption studies, familial correlation estimation involving adoptees with either biological or adopted relatives are compared to infer the relative importance of genetic and environmental factors in determining a given trait. Twin studies are based on the comparison of intraclass correlation coefficients of traits (for example, body mass index [BMI]) in monozygotic (MZ) versus dizygotic (DZ) twins. Any excess of similarity in MZ versus DZ twins is considered as an evidence of the existence of a genetic component in the studied trait. From MZ and DZ comparisons as well as from other family clustering studies it is possible to calculate several measures of heritability that represent the contribution of the genetic component in explaining a given phenotypic trait. Analysis of heritability of complex human diseases based on a large sample of individuals from the United Kingdom Biobank revealed that heritability using simple family-based statistical models were inflated compared to previous estimates, concluding that both genetics and familial environment make substantial contributions to familial aggregation of traits and diseases (Muñoz et al., 2016). On the other hand, *studies of phenotypically discordant twins* provide the unique opportunity to compare subjects with the same genetic background in relation to a wide range of metabolic, epigenetic, and behavioral nutrition-related traits. Additionally, the *prospective cotwin-control design* identifies MZ twin pairs who differ in specific risk factors, and to follow them for the occurrence of health-related outcomes.

Fig. 13.1 shows different study designs commonly used to assess trait resemblance and genetic effects in families without directly measuring genotypes in DNA. Table 13.1 shows rough estimations of type 1 diabetes risk for relatives of index cases affected with the disease. This table also illustrates how genetic and

environmental factors are involved in a multifactorial disease for which a strong genetic factor (class II human leukocyte antigen [HLA] gene variation) is well-established.

ASSOCIATION STUDIES ON MULTIFACTORIAL DISEASES USING THE CANDIDATE-GENE APPROACH

Classical epidemiologic approaches to assess gene-disease associations, such as *cohort*, *classical case-control*, or *cross-sectional studies*, have been widely used in the field of genetic epidemiology. Cohort studies refer to the comparison of the incidence rate of new cases across defined groups, whereas case-control studies are characterized by the comparison of the frequency of risk factors between affected cases and appropriate sets of controls. Hybrid designs have been also widely used in epidemiologic studies, such as the *nested case-control study* (involving the selection of controls for each case from cohort participants) or the *case-cohort study* (involving comparison of cases with a random sample from the cohort). In the field of genetic epidemiology, “association study” is a common term that refers to any research aiming to find the co-occurrence of genetic markers with diseases and health-related traits in unrelated participants (Cardon and Bell, 2001). In the last part of the past century, the candidate-gene approach was the dominant strategy for evaluating the contribution of genetic variants to disease. Genetic markers used in such studies ranged from ABO blood markers to HLA variants, and in general, single nucleotide polymorphisms (SNPs). Apart from several consistent and strong association signals replicated in many studies (for example, Class II HLA genes in type 1 diabetes, or *APOE* variants in Alzheimer disease), many of the gene-disease associations reported in the literature during those years showed problems related to a lack of replication, lack of population stratification control, poor case-control matching, inadequate subgroup analysis, low statistical power, and the presence of publication bias. In the current century, technology advances have allowed the use of genomic arrays of genetic markers able to interrogate millions of single nucleotide variants in a single assay through GWAS.

GENOME-WIDE ASSOCIATION STUDIES

Tremendous progress in the genetic mapping of complex traits was achieved through the implementation of GWAS; it changed the strategy of hypothesis-driven

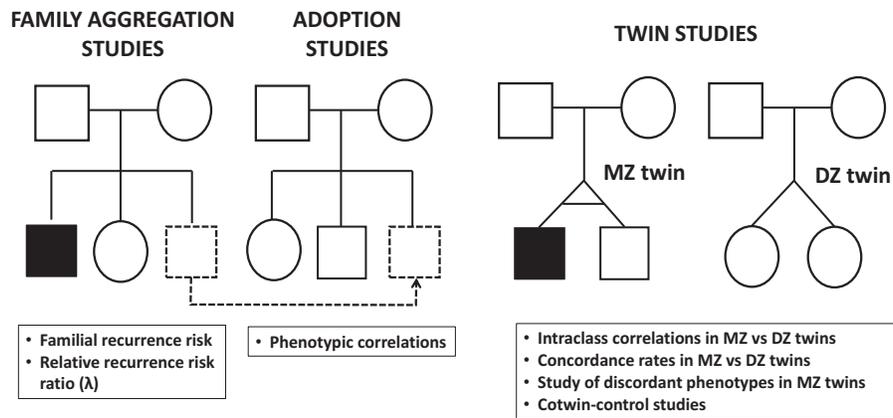


FIGURE 13.1 Study designs commonly used to assess trait resemblance and genetic effects in families without direct DNA measurement. *DZ*, dizygotic; *MZ*, monozygotic.

TABLE 13.1 Approximate Risk Estimates for Developing Type 1 Diabetes in Relatives of Index Patients Diagnosed with Type 1 Diabetes.

Family relation with the index case	Disease risk before adulthood
No family relation (general population) ^a	~0.4%
Sib	~6%
Monozygotic twin	~45%
Dizygotic twin	~10%
HLA identical sib	~15%
HLA haploidentical sib	~5%
HLA discordant sib	~1%

HLA, human leukocyte antigen.

^aThis estimate shows important variations among populations.

studies (based on candidate genes) to hypothesis-free studies (based on the simultaneous evaluation of millions of SNPs) (Visscher et al., 2012). The GWAS approach was facilitated by knowledge derived from international projects on the linkage disequilibrium (LD) structure of the genome across populations. Briefly, LD refers to the coinheritance of alleles at two (or more) loci in a given haplotype that occurs more often than what would be expected by chance. The HapMap project was initially established to describe the haplotypic structure of the human genome across four populations (Phases I and II) and then extended to eight diverse populations (Phase III). Subsequently, the 1000 Genomes Project (<http://www.1000genomes.org/>) was established as the main resource for accessing population genomic data and LD structure in human populations. In association studies, the introduction of susceptibility alleles in an ancestral haplotype creates a particular pattern of LD in cases versus controls that enables

genetic variants associated with diseases to be captured through the detection of such specific LD patterns across the genome. Although the disease-causing or susceptible SNP might not be directly genotyped, LD between the disease-causing SNP and genotyped SNPs results in an association signal at a particular genetic locus. Nevertheless, further fine mapping becomes necessary to determine the causal genetic variants. At a population level, LD structure will decay as a succession of meiotic events that occurs across generations. Then, recent populations arising from few founders (i.e., population isolates) will have longer stretches of LD in the genome compared with LD found in outbred populations, easing the discovery of disease-causing genes.

In GWAS, associations arising in any locus need to be replicated in other independent studies and ultimately meta-analyzed. As such, after replication in tens to hundreds of thousands of individuals, it is likely that genome-wide significant signals associated at a stringent P value $< 5 \times 10^{-8}$ will contain associated variants that represent true underlying biological mechanisms influencing a given trait or disease. In addition, a considerable fraction of other (real) loci have not been identified at genome-wide significance owing to a lack of power (false negatives). From this perspective, meta-analyses of GWAS are expected to keep increasing their sample size as the best means to identify genetic variants with small effects. However, this does not imply that other approaches to improving study design (i.e., better phenotype definition, avoidance of epidemiological or technological biases) can be set aside.

Common genetic variants of minor allele frequency (MAF) greater than 10% with weak (yet true) effects have been the main focus of GWAS (<http://www.gwascentral.org>). Such variants were hypothesized to contribute to the explanation of complex health-related traits and the underlying susceptibility (risk) for having multifactorial diseases. It became evident early that, for

most complex traits and common diseases, the underlying genetic architecture encompasses hundreds (if not thousands) of risk variants of mild effect and with different allele frequencies. According to this perspective, it is necessary to perform well-powered studies incorporating independent populations (for replication), together with a focus on the well-defined selection of SNPs, genes, and gene regions. Moreover, a robust control for multiple hypotheses testing is mandatory to identify genuine genetic effects. Few examples can be mentioned regarding common variants exerting large effects (measured as odds ratios) on complex diseases, such as class II HLA genes in type 1 diabetes, complement factor H in age-related macular degeneration, *APOE* in Alzheimer disease, or *APOA5* in hypertriglyceridemia, and it is unlikely that many such others remain to be identified. Moreover, odds ratios found in GWAS are relatively mild to moderate for most common nutrition-related diseases, such as the *TCF7L2* gene for type 2 diabetes or the *FTO* gene for obesity. In addition, there are several examples of genetic variants with a strong influence on nutrition-related traits such as variants near the lactase (*MCM6-LCT*) gene for lactose intolerance, the taste receptor *TAS2R38* gene in the detection of phenylthiocarbamide (bitter taste), or the copy number variation of the salivary amylase gene *AMY1A* in relation to salivary amylase content or activity.

SNPs achieving genome-wide significance in GWAS generally explain only a small part of the phenotypic variance of many complex traits such as glycemia, plasma lipids, or BMI. In many multifactorial diseases, such variants do not significantly improve the prediction of common diseases compared with existing clinical tools such as, for example, a family history of disease. In this context, the concept of missing heritability refers to the gap between low percentages of the phenotypic variance explained by common SNPs that achieved genome-wide significance compared with the high heritability estimated from twin-, family-, or population-based studies. However, studies have shown that the complete set of SNPs (not only those that achieved genome-wide significance) may explain up to 56% of variance for human height and 27% for BMI (Yang et al., 2015).

Low-frequency genetic variants in the range of 0.5–5% MAF represent an active objective of research in current genome epidemiology studies. The focus in this type of variant takes advantage of the existence of ethnically diverse sequenced references (1000 Genomes Project and UK10K) and the use of specific statistical tests to detect the effects of variants of MAF less than 5%. As an example of the importance of these

variants, low-frequency loss-of-function mutations in *PCSK9* (population frequency 2–4%) were associated with a healthy cardiometabolic profile. Such an observation has led to the development of *PCSK9* inhibitors as a class of low-density lipoprotein (LDL) cholesterol-lowering drugs. Finally, rare or very low-frequency variants of very small effect may also exist, but they are unlikely to be identified by strategies derived from GWAS.

PATHWAY-BASED ASSOCIATION STUDIES

As mentioned, GWAS emerged as a hypothesis-free strategy that has been useful in detecting gene variants with genome-wide significance related to nutrition-related traits and diseases. However, most individual GWAS are probably underpowered to find SNPs with weak (but real) effects, detecting only those with the strongest effect size. Consequently, analyses of most GWAS focusing on individual SNPs often leave undetected gene variants with mild effects. In addition, performing repetitive association tests on millions of single SNPs is inappropriate if we wish to dissect the complex architecture of multifactorial diseases, which are characterized by the joint influence of multiple genes acting in intricate metabolic pathways within a hierarchical network (see, for example, the Kyoto Encyclopedia of Genes and Genomes pathways at <http://www.genome.jp/kegg>). Then, it is expected that a pathway-based association analysis may represent an alternative way to assess the joint effect of multiple variants of mild effects within a specific set of genes belonging to the same metabolic pathway (Wang et al., 2010). This strategy is based on the examination of an initial collection of predefined gene sets for pathways based on prior biological knowledge, which is the significance of each pathway summarized based on gene–disease associations. Difficulties in applying this strategy reside in the choice of the most adequate representation of each gene or pathway by multiple SNPs, the handling of LD between SNPs, the correct adjustment of varied sizes of genes, as well as the adequate use of additional information coming from expression or gene methylation databases. Pathway-based approaches have been successfully used to detect gene variants of the interleukin (IL)12/IL23 pathway associated with Crohn disease, the Wnt signaling pathway in type 2 diabetes, or sterol transport and metabolism genes in plasma high-density lipoprotein cholesterol levels.

FAMILY-BASED ASSOCIATION STUDIES

Family-based association studies have been used as an alternative to classical case–control studies to circumvent the distorting effect of population stratification by ethnicity, which is an important confounding effect in genetic epidemiology studies (Cardon and Palmer, 2003). When affected cases are sampled from case–parent trios, associations between genetic markers and the disease cause the apparent probability of transmission to differ from the expected value of 0.5. In the *case–parent triad design*, the transmission/disequilibrium test (TDT) is generally used to assess the transmission of alleles from heterozygous parents to affected patients. The most important advantage of this approach is that TDT is a valid test of association even if population stratification is present. Disadvantages of this design are related to the difficulty and increased cost of enrolling such trios, and the loss of information derived from ascertaining homozygous parents. This type of study is most adequate when differences are expected in both the prevalence of the disease of ethnic subgroups within the population and the frequency of alleles across different ethnic groups (Fig. 13.2). An example of the adequate use of this type of study design is an assessment of the association between HLA gene variation and type 1 diabetes in ethnically diverse populations.

ASSESSMENT OF GENE–ENVIRONMENT INTERACTIONS AND MEDIATION EFFECTS

An important topic in genetic epidemiology studies is the assessment of gene–gene and gene–environment

interactions (Thomas, 2010). Interaction in the epidemiologic context refers to departures from additive or multiplicative models of disease risk. Interaction (also called effect modification) is frequently presented as the influence of a third variable (effect modifier) on the differential magnitude of the association between genetic markers and health-related outcomes. An example of this approach is given by studies in which the intake of sugar-sweetened beverages interacts with genetics (risk score for obesity derived from GWAS) in determining weight gain in children. Gene–environment interactions can be also tested in *randomized trials*, in which a health intervention is randomly assigned to study groups. Here, the incorporation of genetic markers provides the opportunity to test gene–environment interactions in this context.

It is known that the sample size to detect gene–environment interactions is much greater than the one needed to find true gene–disease associations, which results in the requirement for large and costly studies. An efficient alternative approach to assessing gene–environment interactions is exemplified by the *case-only design*. In this type of study, testing gene–environment associations in patients indicates gene–environment interactions under the assumption of gene–environment independence in the source population.

A global approach to the studying the effects of genetic background and environment without measuring genotypes is provided by *migration studies*. In this type of study, incidence rates of a given disease are compared in an ethnically homogeneous population living in a confined location versus the same population after migrating to a different environment (usually a different country or region), in which where the incidence rate is

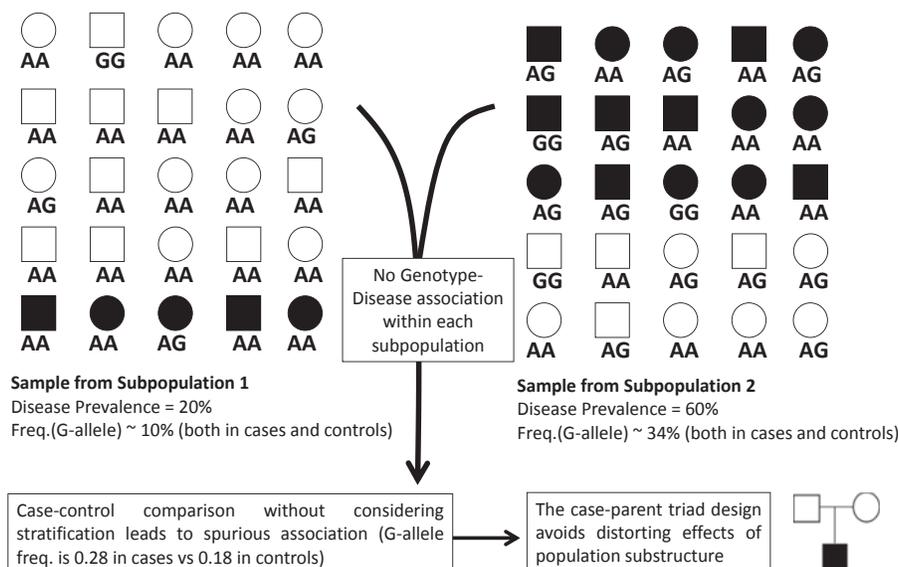


FIGURE 13.2 Patient–parent triad design to avoid distorting effects of population stratification by ethnicity. *Freq.*, frequency.

known to be sharply different. A classic example is the migration of people from the Italian island of Sardinia (with a high incidence of type 1 diabetes) to continental Italy (with medium to low incidence rates of the disease).

A different approach from evaluating gene–environment interactions is the assessment of mediation effects among variables in *mediation studies*. A mediator is defined as a variable that is simultaneously associated with the genetic factor and the outcome but is considered an intermediate step in the causal chain between the independent and dependent variables. Thus, a mediator variable is conceptually different from an effect modifier and from a confounder. As an example of this approach, epidemiologic studies have been used to assess the mediation effect of eating behavior traits on the relation between genetic scores derived from GWAS and BMI in childhood.

In general, multipurpose cohort studies frequently consider multiple measurements of many traits, phenotypes, environmental exposures, and metabolic

variables apart from genetic markers, which provides the opportunity to assess the relation among variables at various complex levels. Fig. 13.3 shows several possible interpretations that may arise in observational cohort studies involving multiple measurements.

MENDELIAN RANDOMIZATION STUDIES

In observational epidemiology, causality inference regarding the assessment of associations between nutrition-related variables and the risk for a disease is threatened by unmeasured or residual confounding and reverse causation effects. In this context, an approach to studying the contribution of metabolic or nutritional risk factors to health-related outcomes is represented by *Mendelian randomization studies*. In this type of study, genetic variants are treated as instrumental variables used as proxies for exposures of interest (many times, they are nutritional), with the final purpose of determining the causal relation between the given

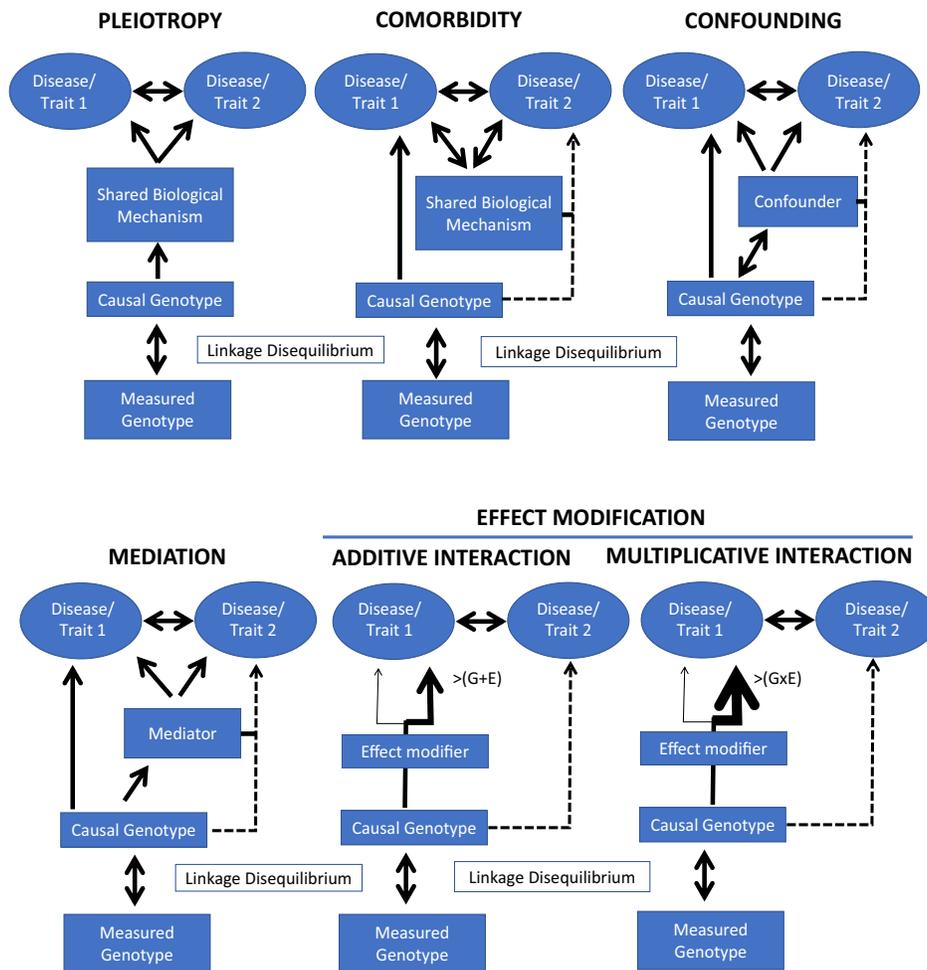


FIGURE 13.3 Relations among study variables in genomic epidemiology observational studies. G, Genetics; E, Environment.

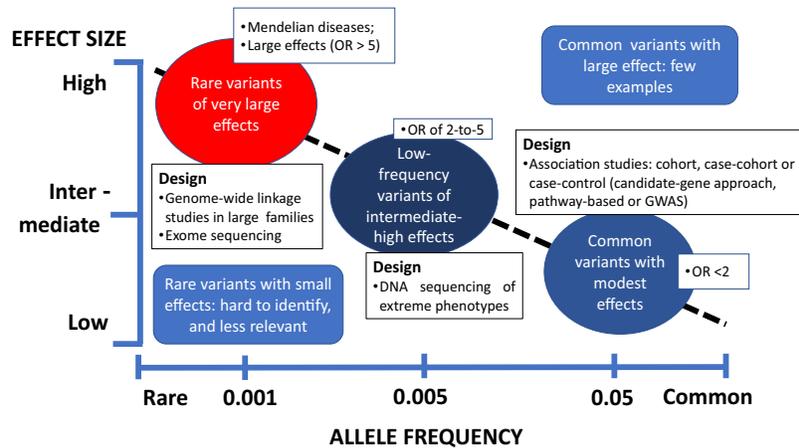


FIGURE 13.4 Different study designs used in genetic epidemiology of rare and common diseases. GWAS, genome-wide association study; OR, odds ratio.

exposure influenced by SNPs and a health-related outcome. An example of this approach is an evaluation of the associations of LDL cholesterol-lowering *PCSK9* variants with type 2 diabetes to assess the effects of *PCSK9* inhibitors on the risk of diabetes. An example in the field of nutrition is an assessment of the association between cruciferous vegetables and lung cancer after stratification by *GSTM1* and *GSTT1* genes, which are implicated in the elimination of isothiocyanates, the likely chemopreventive compound present in such vegetables.

LINKAGE STUDIES AND NEXT-GENERATION SEQUENCING METHODS IN EXTREME PHENOTYPES AND RARE DISEASES

Rare variants affecting severe genetic diseases such as cystic fibrosis, neurofibromatosis, and Huntington disease were successfully mapped during the last part of the 20th century using *genome-wide linkage studies*. In this type of study, mapping the causative gene for monogenic diseases depended on finding families with known Mendelian inheritance of the disease, preferably applied to multigenerational pedigrees with multiple affected patients. *Segregation studies* were also useful in determining whether the disease occurrence in families fit any particular inheritance pattern (dominant, recessive, etc.). In so-called parametric linkage studies applied in monogenic disorders, genotypes of polymorphic genetic markers across the genome (microsatellite markers in the past, or SNPs) are traced in family members to provide evidence of the location of the disease-causing gene, after the specification of a set of parameters (penetrance or allele frequencies). A straightforward type of linkage study is represented

by the *homozygosity mapping* approach, in which segments of shared homozygosity are searched in patients from consanguineous families affected with autosomal recessive diseases. Despite great achievements in the application of linkage studies in many monogenic disorders, this strategy had little success in identifying underlying mutations in genetic diseases with low reproductive fitness, diseases caused by de novo mutations, late-onset diseases, or diseases with early lethality. In addition, nonparametric linkage studies involving families with affected sib-pairs were evaluated in multifactorial disorders with limited success, given the difficulty of detecting small genetic effects among the multiplicity of factors involved in disease etiology.

Next-generation sequencing techniques offer an unprecedented ability to analyze the whole human genome or targeted sequencing projects. Thus, studies focusing on evaluating the exome (the part of the genome that encodes for proteins) have been successfully used to identify genetic causation in unsolved monogenic diseases. The field of human genetics has also expanded its technological armamentarium to evaluate the genome-wide landscape of insertions and deletions through arrays of comparative genome hybridization, enabling the diagnosis of diseases caused by gains and losses of genomic regions. In this context, the combination of exome sequencing, linkage analysis, and metabolomics has been shown to be a powerful weapon in precision medicine to find new genetic causes of rare genetic diseases. For example, a syndrome characterized by epilepsy and neurological impairment was finally found to be a genetic metabolic disease caused by mutations in the branched chain ketoacid dehydrogenase kinase gene that was proposed to be ameliorated with dietary branched-chain amino acids (Novarino et al., 2012). In this family,

identification of disease-causing mutations was carried out through a multistep process initiated with exome sequencing, homozygosity mapping, the study of serum metabolome, and in silico and in vivo tests. This research represents a paradigmatic example on how genomics and metabolomics can be applied to find genetic causes of diseases that may eventually lead to personalized therapies. Both genome-wide linkage analysis and genome/exome sequence analysis are hypothesis-free strategies, which enable the identification of mutations in known genes or novel mutations in genes that were not previously involved in the etiology of the disease.

RECAPITULATION

The choice of an appropriate study design is a critical first step in any genomic epidemiology study; it becomes even more important considering the massive information derived from genomics and other “-omics” disciplines such as epigenomics, proteomics and metabolomics. [Figs. 13.1 and 13.4](#) show a graphical summary of different study designs commonly used in genomic epidemiology studies.

References

- Cardon, L.R., Bell, J.I., 2001. Association study designs for complex diseases. *Nat Rev Genet* 2 (2), 91–99.
- Cardon, L.R., Palmer, L.J., 2003. Population stratification and spurious allelic association. *Lancet* 361 (9357), 598–604.
- Muñoz, M., Pong-Wong, R., Canela-Xandri, O., Rawlik, K., Haley, C.S., Tenesa, A., 2016. Evaluating the contribution of genetics and familial shared environment to common disease using the UK Biobank. *Nat Genet* 48 (9), 980–983.
- Novarino, G., El-Fishawy, P., Kayserili, H., Meguid, N.A., Scott, E.M., Schroth, J., et al., 2012. Mutations in BCKD-kinase lead to a potentially treatable form of autism with epilepsy. *Science* 338 (6105), 394–397.
- Thomas, D., 2010. Gene–environment-wide association studies: emerging approaches. *Nat Rev Genet* 11 (4), 259–272.
- Visscher, P.M., Brown, M.A., McCarthy, M.I., Yang, J., 2012. Five years of GWAS discovery. *Am J Hum Genet* 90 (1), 7–24.
- Wang, K., Li, M., Hakonarson, H., 2010. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet* 11 (12), 843–854.
- Yang, J., Bakshi, A., Zhu, Z., Hemani, G., Vinkhuyzen, A.A.E., Lee, S.H., Robinson, M.R., Perry, J.R., Nolte, I.M., van Vliet-Ostaptchouk, J.V., Snieider, H., LifeLines Cohort Study, Esko, T., Milani, L., Mägi, R., Metspalu, A., Hamsten, A., Magnusson, P.K., Pedersen, N.L., Ingelsson, E., Soranzo, N., Keller, M.C., Wray, N., 2015. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet* 47 (10), 1114–1120.

Epistemology of Nutrigenetic Knowledge

Martin Kohlmeier

Human Research Core and Nutrigenetics Laboratory, UNC Nutrition Research Institute; and Schools of Medicine and Public Health, Department of Nutrition, Kannapolis, NC, United States

Glossary

Biomarker Measurable biological indicator of a state or outcome of interest, usually a molecule that can be analyzed in a sample.

Nutritope Nutrition environment to which local groups are exposed to and which can shape their genome over time.

Replication Implementation of a follow-up study or experiment based on the same design with the expectation of finding a similar result.

INTRODUCTION

The rapid growth of research and practice makes it imperative to consider how we know what we think we know about nutrigenetics. We want to understand the foundations of our knowledge about genetically inherited dispositions affecting the impact of nutrition on health.

Without going too much into depth about our ability to know anything at all, this discussion will focus on some practical aspects of nutrigenetic knowledge. In the end, we need to be prepared to answer the almost child-like question, “How do you know?” The reason we should ask ourselves and our peers that question is that rationally rooted nutrition knowledge has a better chance of holding up over time. We often find that the most worthwhile nutrition questions unavoidably involve genetic factors, although often in the negative.

Distinguishing and enumerating the individual factors involved in seemingly simple nutrition–health relations is a never-ending task. This goes for all involved variables, such as specific foods and food ingredients, metabolic and regulatory events, and functional circumstances at all levels, from molecular changes to physical conditions and population health.

EPISTEMOLOGY AND THE INITIAL BUILDING BLOCKS OF KNOWLEDGE

The term “epistemology” (“making sense of knowledge”) was used by the Scottish philosopher Ferrier in his early exploration of the “laws of our knowing and of our thinking” (Ferrier, 1854). This branch of philosophy wants to find the nature of our knowledge, where it comes from, and whether we are justified in trusting it. We need to ask ourselves the same questions about nutrigenomics.

When we read about typical nutrition guidelines, we rarely spend much time thinking about the myriad underlying assumptions we all bring to the table. It would be impossible to consider all of the different ways in which one individual differs from the next. We therefore ignore most of those differences and lump many others into broad categories. Current nutrition recommendations (Dietary Reference Intakes), for example, use just 22 groups defined by age, gender, pregnancy, and lactation status. Individuals in each group are given the same intake targets for the obvious reason that we cannot agree even about what the next level of subdivisions should be and then allocate sufficient resources to collect the necessary data.

The first requirement in building knowledge is a precise statement of each constituent item. In practical terms, this means that we use terms with clearly defined meaning understood by stakeholders. Although this may appear needlessly reductionist, it promises a sustainable path toward steadily growing strength of knowledge about human biology and how to achieve desirable health outcomes.

A common problem is that we find it hard to recognize our biases. After all, this is why we hold on

to them. Thus, ethnicity is too often assumed in practical terms as a deviation from a background norm without thinking much about it. For example, the ability of adults to digest lactose has been perceived as normal for the longest time. The absence of this digestive function in adults has been perceived as a loss of function and the resulting lactose intolerance as a health condition or disease. But who had this understanding as a matter of course and could not conceive of any other view? It was physicians and scientists in Central and Northern Europe and in the United States. Researchers in Asia would not have come to that conclusion because they knew otherwise from practical experience. Now we understand that adult lactose intolerance is the normal state present in a majority of the world's populations and that specific mutations conferred the ability to tolerate lactose only a few hundred generations ago. It is now clear that people with lactose intolerance are the ones who are normal, and the milk aficionados are the mutants. But perhaps it is better not to assign such loaded terms as "normal" or "deviation" and just be neutrally descriptive.

Another nutrigenetic example in which perspective matters relates to our definition of disease. Most will agree that the disease definition appropriately applies to celiac disease. But should we still consider it a disease even long after symptoms have been reversed by careful avoidance of gluten prompted by early diagnosis? Is it a disease because a sensitized individual is actually sick or because a health risk can be triggered at any time by exposure to gluten? If it is the latter, any nutrigenetic disposition that makes a nutrition exposure detrimental to the carrier's health should possibly also be called disease.

For all practical purposes, there are far too many variables to consider for any hypothesis, individual study, or intervention guideline. Therefore, the critical question becomes which we can ignore or group with others and what justifies our decisions.

Most of the time, variable selection just happens without us even realizing that a choice has been made. An important benefit of exploring the epistemology of specific knowledge areas is that unjustified assumptions may become apparent and thus open to investigation. This is particularly true for most nutritional effects. We have to ask ourselves when genetic variation must be considered. Table 14.1 lists a few such commonly used variables that we use regularly, probably without giving much thought to the specific meaning we intend.

COLLECTING DATA

These are golden days for genetic studies. It has become so much easier and less expensive to obtain

TABLE 14.1 We Commonly Refer to Variables Without Defining Exactly What We Mean, Often Because We Do Not Know That Different Meanings Exist.

PERSONAL VARIABLES

Gender (could be self-reported or anatomically confirmed or based on genetic markers, etc.)

Ethnicity (could be race, self-reported ancestry, country of origin, genetic markers, etc.)

Genotype (could be a single-nucleotide polymorphism (SNP), an SNP as one in a pair, an SNP within a long-range haplotype, etc.)

EXPOSURE VARIABLES

Food (could be defined by name, origin, varietal, quality grade, season, etc.)

Ingredient (could be a natural or complex group with a shared name, natural plus added, etc.)

OUTCOME VARIABLES

Risk (could be acute events, morbidity, composite outcomes, all-cause mortality, etc.)

Disease (could be ongoing ill health, a state of compensated vulnerability, etc.)

reliable genetic information. The quality of common genotyping platforms is commonly very high. Nonetheless, performance has to be carefully examined, particularly when using large-scale analyses with millions of genome positions interrogated, possibly even the full genome. The limiting factors will usually be nongenetic variables. Quantitative assessment of dietary intakes for many study participants is still a tough barrier involving high cost, high day-to-day variability, poor accuracy, and fluctuations over long periods. Although major efforts have been made to make assessment instruments comparable across different populations, cultural differences continue to cause major practical problems for nutrigenetic research.

What too often gets lost when focusing on particular hypotheses are major differences in exposures outside the narrow research focus. For example, major items in the food supply in the United States and other countries are fortified, whereas this is not the case in Europe and elsewhere. This means that, for instance, the investigation of genotype-specific effects of folate on cardiovascular outcomes should not be expected to yield consistent results. Blood pressure data in an Irish population with different methylenetetrahydrofolate reductase (MTHFR) rs1801133 genotypes (MTHFR 677 C > T) were found to be consistently different (Wilson et al., 2012). However, this is not simply related to a difference in folate status, but to riboflavin intake. This would not be seen in an American population because with fortification, average riboflavin intake levels are much higher there than in European cohorts. A controlled feeding trial could demonstrate exactly that, with a sizable effect

of low-dose riboflavin (1.6 mg/day, about as much as the amount coming from fortification in Americans) on systolic and diastolic blood pressure (Wilson et al., 2013). This is one of the most obvious examples because we know the accurate amount of added vitamin coming from fortification. Another time, a discrepancy could arise from a different cultivar of a food item with the same name or dietary patterns between regions.

This is not to say that analytical challenges are less important. It is just that measurement of standard biomarkers is getting much more attention and presents more finite problems. Analytical methodologies continue to present difficult questions, such as when deciding whether to determine folate concentration with a microbial assay or by mass spectrometry. These are obviously not trivial issues because each method captures different metabolites best and the resulting trade-offs are not easily reconciled.

An easier question relates to analytical errors and what is tolerable. This is particularly important when wondering why a genotype-specific effect or interaction was observed in some studies and not in others.

TURNING DATA INTO KNOWLEDGE

Some irreproducible reports are probably the result of coincidental findings that happen to reach statistical significance, coupled with publication bias. Another pitfall is overinterpretation of creative ‘hypothesis-generating’ experiments, which are designed to uncover new avenues of inquiry rather than to provide definitive proof for any single question. Collins and Tabak (2014).

Now it gets even harder because data do not speak for themselves. Probably the most important requirement is diligent replication of studies, because, as Collins and Tabak pointed out, coincidental findings are the bane of nutrigenetic studies. Correcting the statistical significance for testing multiple contrasts is an insufficient solution, because with so many different investigators looking for statistically significant associations, some will get lucky and submit for publication an interesting association with apparently legitimate numbers (Goodman et al., 2016).

Proper replication should help, but before considering what this means, it must be recognized that few original reports of nutrigenetic interactions are replicated in a timely manner (Hirschhorn et al., 2002). Replication studies need an adequate size and comparable study design. The possibility of publication bias is a serious concern. Many investigations that do not find a previously reported result will often never be published because the earlier finding has already become dominant. The investigators may be too discouraged to complete all of the work needed to submit a manuscript for

publication. If they do, they may find that reviewers are more critical than they would be with a confirmatory report, or editors are less interested and are worried about an imperfect replication. Goodman and colleagues highlighted the difficulty of even using a consistent conceptual framework to achieve comparability. They pleaded for reproducing research as an “imperfect surrogate for scientific truth” rather than an end in itself (Goodman et al., 2016). The MTHFR story mentioned earlier should be enough reason to pay close attention to studies that fail to find a comparable result. Although this is to be expected when the original finding was strictly coincidental, real interactions are easily missed because one or more critical variables were different without the authors knowing about it.

Avoiding over interpretations is closely related to having too much confidence in the findings absent their replication. Another even more important issue is the unsupported use of extrapolations to another gender or other age groups, regions, ethnicities, or health status. A lot can and often will go wrong when assuming that people with different characteristics will show the same genotype-specific response found in a particular study. The integration of diverse types of study data, including from cross-sectional surveys, prospective cohorts, and targeted intervention trials, requires a fine balance between forced bundling and thoughtful alignment of the collected information (Ioannidis et al., 2008). The experience with MTHFR, folate and riboflavin intake, and blood pressure outcomes shows what can be done with less challenging hypotheses. It may also serve as a reminder that it makes a lot of sense to focus on proximal effects and interactions first, such as blood pressure, which tends to respond fairly quickly. The question about a much more complex disease group such as stroke could then build on that prior knowledge base.

Another weakness, though often unavoidable, is the expectation that findings from cross-sectional population studies can be used to predict response to an intervention. The understandable argument is that in most cases it is impractical to demonstrate a genotype-specific response to a nutrition intervention for which it will take many years or even decades to yield results, such as with the prevention of diabetes, cardiovascular disease, or cancer. There are no easy answers, but only with full awareness of the involved issues can risk–benefit assessments of predictive uses of genetics data even be discussed, much less effectively used in practice.

FURTHER CONSIDERATIONS

Knowledge should always be considered tentative, even after multiple confirmatory observations. There is

always the possibility that important variables have not been considered and that underlying mechanisms have been misunderstood. Many times, the science has been considered final, only to give way to better understanding. A key criterion for accepting new knowledge items is improved ability to predict effects and improve outcomes. Consistent agreement of observations with the expected behavior increases confidence but does not establish certainty.

References

- Collins, F.S., Tabak, L.A., 2014. NIH plans to enhance reproducibility. *Nature* 505 (7485), 612–613.
- Ferrier, J.F., 1854. *Institutes of Metaphysics. The Theory of Knowing and Being*. W. Blackwood and Sons, Edinburgh. §55, p. 44ff.
- Goodman, S.N., Fanelli, D., Ioannidis, J.P., 2016. What does research reproducibility mean? *Sci Transl Med* 8 (341), 341ps12.
- Hirschhorn, J.N., Lohmueller, K., Byrne, E., Hirschhorn, K., 2002. A comprehensive review of genetic association studies. *Genet Med* 4 (2), 45–61.
- Ioannidis, J.P., Boffetta, P., Little, J., O'Brien, T.R., Uitterlinden, A.G., Vineis, P., Balding, D.J., Chokkalingam, A., Dolan, S.M., Flanders, W.D., Higgins, J.P., McCarthy, M.I., McDermott, D.H., Page, G.P., Rebbeck, T.R., Seminara, D., Khoury, M.J., 2008. Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol* 37 (1), 120–132.
- Wilson, C.P., Ward, M., McNulty, H., Strain, J.J., Trouton, T.G., Horigan, G., Purvis, J., Scott, J.M., 2012. Riboflavin offers a targeted strategy for managing hypertension in patients with the MTHFR 677TT genotype: a 4-y follow-up. *Am J Clin Nutr* 95 (3), 766–772.
- Wilson, C.P., McNulty, H., Ward, M., Strain, J.J., Trouton, T.G., Hoeft, B.A., Weber, P., Roos, F.F., Horigan, G., McAnena, L., Scott, J.M., 2013. Blood pressure in treated hypertensive individuals with the MTHFR 677TT genotype is responsive to intervention with riboflavin: findings of a targeted randomized trial. *Hypertension* 61 (6), 1302–1308.

S E C T I O N I I

Nutrigenetics

This page intentionally left blank

Nutrigenetics and the Early Life Origins of Health and Disease: Effects of Protein Restriction

Simon Langley-Evans

School of Biosciences, University of Nottingham, Loughborough, United Kingdom

Abbreviations

Apbb1	Amyloid beta precursor protein binding family B member 1
FXR	Farnesoid X receptor
LP	Low protein
LXR	Liver X receptor
Mdm2	Murine double minute 2
PCR	polymerase chain reaction
Pmp22	Peripheral myelin protein 22
RNASeq	RNA sequencing
RXR	Retinoid X receptor
Taf10	TATA-box binding protein associated factor 10
Tsg101	Tumor susceptibility gene 101
UTR	Untranslated region
Wee1	G2 checkpoint kinase

EARLY LIFE PROGRAMMING OF DISEASE

An ever-growing body of epidemiological evidence indicates that risk of noncommunicable diseases of adult life is, in part, determined by the environment encountered in early life (Calder et al., 2018). Follow-up studies of historic cohorts show that cardiovascular disease and type-2 diabetes are more prevalent in individuals who were of lower weight at birth and who showed rapid catch-up growth in childhood. Exposure of the embryo, fetus, or infant to a diet that does not meet requirements (either maternal diet or infant feed) elicits adaptive responses that, as they occur during periods of organogenesis and organ maturation, are irreversible. This process by which early life nutrition permanently alters physiology and metabolic competence is referred to as programming (Langley-Evans, 2015). It has been suggested that the responses to an adverse maternal environment during development become maladaptive if the individual is born into a different environment.

For example, disease risk will be established in early life if famine during development gives way to an obesogenic environment later in life (Lillycrop and Burdge, 2015).

Programming by maternal diet has been demonstrated in a number of experimental animal studies, with physiological and metabolic consequences widely reported in offspring in response to maternal micronutrient or macronutrient restriction, and dietary excess (Vithayathil et al., 2018; Hay et al., 2016). Experiments in which pregnant rodents are subject to moderate dietary protein restriction show that cardiovascular and metabolic regulation, renal function, and longevity are all compromised in the associated offspring (Calder et al., 2018). For example, offspring of rats fed a low-protein (LP) diet in pregnancy have high blood pressure from the time of weaning and develop profound hepatic steatosis with aging. The latter is attributed to programming of the lipogenic pathway. As they age they exhibit an insulin resistant state and excessive adiposity (Erhuma et al., 2007).

MECHANISMS OF PROGRAMMING

Offspring of animals subject to under- and overnutrition during pregnancy show evidence that organ structure is altered by the experience. Rats exposed to LP in fetal life have kidneys that contain fewer nephrons, and reduced numbers of islets of Langerhans in the pancreas (Calder et al., 2018; Swali et al., 2011). Similar observations are made in humans where nephron number is associated with weight at birth and renal function is impaired in adults exposed to undernutrition in utero. As organ development is largely complete around the

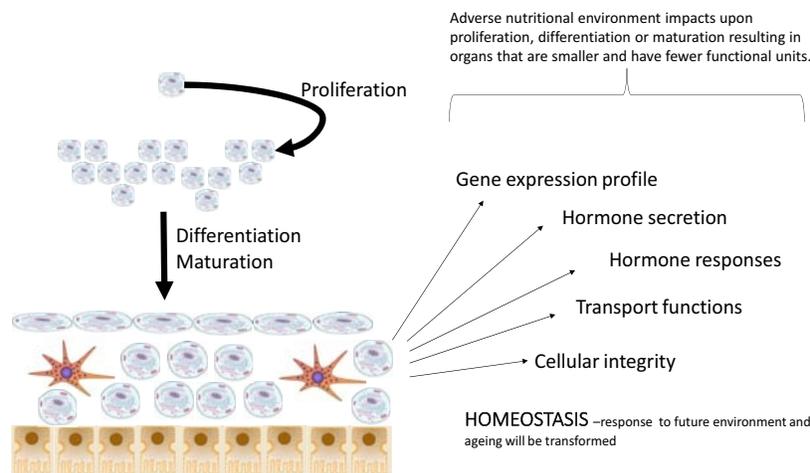


FIGURE 15.1 The concept of tissue remodeling. Organs and tissues develop according to a genetically determined pattern during embryonic and fetal life. Programming of long-term tissue function will occur due to environmental impacts on the genetic pattern, with disruption of proliferation and differentiation resulting in tissues with fewer functional units.

time of birth, any deficits in functional units (e.g., nephrons in the kidney, islets in the pancreas) cannot be recovered, permanently altering structure. Whilst during earlier stages of life the capacity to fulfill organ function will be present, with aging, the organs can no longer meet demands, leading to renal failure, cardiovascular disease, and metabolic incompetence (Calder et al., 2018).

As shown in Fig. 15.1, all tissues and organs develop from progenitor cells, which must first proliferate into progenitor lines. Waves of proliferation and apoptosis build up complex structures, and differentiation of progenitors into specialized cell types gives the tissues their characteristic morphology and functional capacity. Undernutrition at any of these stages can impact the numbers of cells or type of cells in a tissue. With these irreversible changes, all tissue functions and ultimately homeostatic capacity will be altered and effectively set the organism up for either a healthy aging path or ill-health in later life.

Whilst tissue remodeling provides a direct link to early life exposure to undernutrition, the mechanism through which the developing tissues are stimulated to remodel is equally important but is less well understood. Ultimately this situation must involve changes in gene expression that upset the normal developmental pattern of events, but the initiating factors are not well characterized. Studies generally have focused on the downstream effects of adverse nutrition, for example, the expression of the lipogenic pathway in adult offspring exposed to LP (Erhuma et al., 2007), and miss the fact that the true drivers of programming will only be seen whilst the nutritional insult is ongoing. Disturbance of gene expression may only be transient, lasting for a matter of hours or days, and yet can still elicit developmental delay and permanently remodel structures.

Several hypotheses have been advanced to explain how poor maternal nutrition can have a permanent impact upon organ development and function. There is good evidence, for example, that some of the effects of maternal undernutrition arise as a result of downregulation of a key placental barrier to maternal glucocorticoid transfer to the fetal compartment. The programming effects of LP in rats are dependent upon maternal glucocorticoid synthesis and can be replicated by treating protein-replete pregnant rats with carbenoxolone, which promotes glucocorticoid transfer across the placenta (Bogdarina et al., 2010). Similarly, there is evidence that is consistent with the hypothesis that maternal undernutrition modulates gene expression by altering DNA methylation marks on specific gene promoters and other aspects of the epigenetic regulatory machinery (Bogdarina et al., 2010; Altobelli et al., 2013). All of these factors will contribute to tissue remodeling during development.

PROTEIN AND THE PLACENTAL TRANSCRIPTOME

As the interface between the mother and fetus, the placenta plays a key role in developmental programming. This may result from nutrition-driven changes in maternofetal endocrine cross-talk (as discussed earlier), through alterations in the capacity to supply the fetus with substrates and changes in the expression of signaling molecules. Protein restriction alters the expression of placental amino acid transporters, increases fetal glucocorticoid exposure, and remodels placental vasculature.

The impact of maternal protein restriction on the placental transcriptome at day 13 gestation was investigated using RNASeq as a means of capturing the impact of the dietary insult upon the tissue (Daniel et al., 2016). As shown in Fig. 15.2 the effect was extremely modest with only 91 out of 29,516 identified loci showing differential expression in the LP group. However, bioinformatic analysis showed that there was strong enrichment of a number of cellular processes. The main pathways and processes in placenta that were associated with these differentially expressed genes were atherosclerosis signaling, LXR/FXR activation of RXR, and other processes strongly linked to atherosclerosis including inflammation. Follow-up PCR studies of genes from these pathways showed major upregulation (18–32-fold) with protein restriction. The strongest effects were noted for cubilin, retinol binding protein-4, microsomal triglyceride transfer protein, and the apolipoproteins A2 and C2. Interestingly, the same genes were shown in a separate study to be sensitive to maternal fat intake in rats (upregulated gene expression as with LP) and maternal body mass index (upregulated gene expression with obesity) in human placenta.

Collectively these features are all suggestive of maternal protein restriction resulting in increased transport of cholesterol from mother to fetus. The uptake of cholesterol by the embryo and fetus is critical for normal development, and defects of endogenous cholesterol synthesis are known to be lethal. Cholesterol will also play an important role in placental function as it is the

precursor for all steroid hormone synthesis. Disturbances of placental transport or endogenous fetal cholesterol synthesis can have a number of effects on growth, cell proliferation, metabolism, and the organization of tissues, all of which are pertinent in terms of the remodeling of tissues that is associated with maternal undernutrition. A number of studies have also shown that exposure of rodent and human fetuses to maternal hypercholesterolemia can establish atherosclerotic lesions in the fetal vasculature.

PROTEIN AND THE EMBRYONIC AND FETAL TRANSCRIPTOME

The impact of the maternal diet upon the placenta is one route through which the pregnancy environment can determine how the fetal tissues develop. It is clear though that there are also direct effects upon the developing embryo and fetus, independent of the placenta, that translate into remodeling of tissues and hence long-term physiological and metabolic functions. There has been considerable interest in how quite diverse nutritional insults elicit very similar programmed phenotypes (McMullen et al., 2012). For example (Fig. 15.3), maternal protein restriction, maternal iron deficiency, and maternal obesity in rodents all result in offspring with high blood pressure, renal impairment, and insulin resistance. This suggests that a relatively narrow range of mechanisms may be responsible for

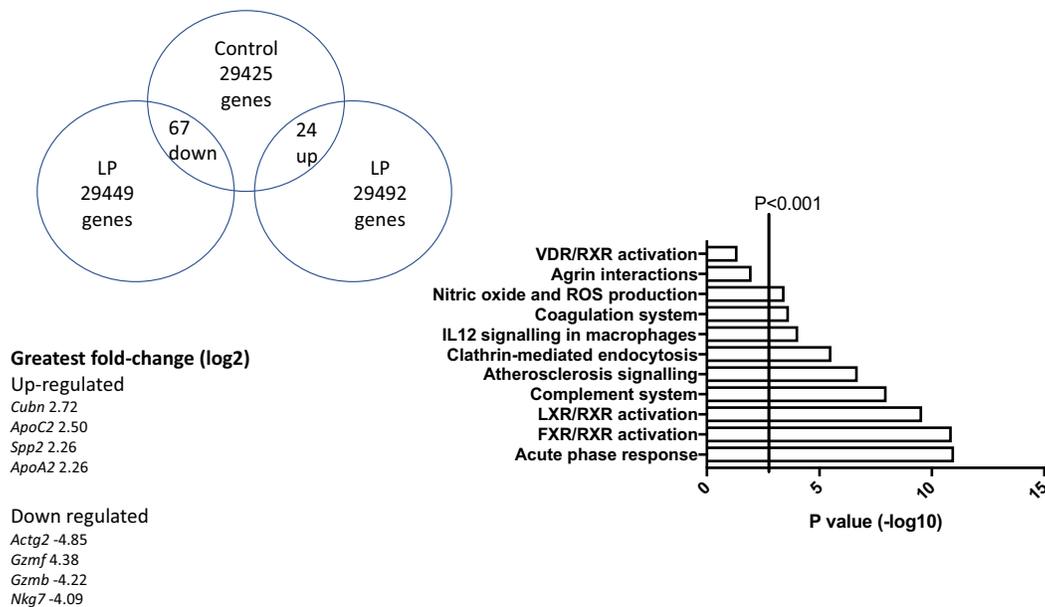


FIGURE 15.2 The impact of protein restriction on the rat placental transcriptome. RNASeq showed that maternal protein restriction altered expression of 91 genes in rat placenta and that pathways associated with cholesterol metabolism and transport were highly enriched in the differentially expressed set of genes. Data drawn from Daniel, Z., Swali, A., Emes, R., Langley-Evans, S.C., 2016. The effect of maternal undernutrition on the rat placental transcriptome: protein restriction up-regulates cholesterol transport. *Genes Nutr* 11, 27.

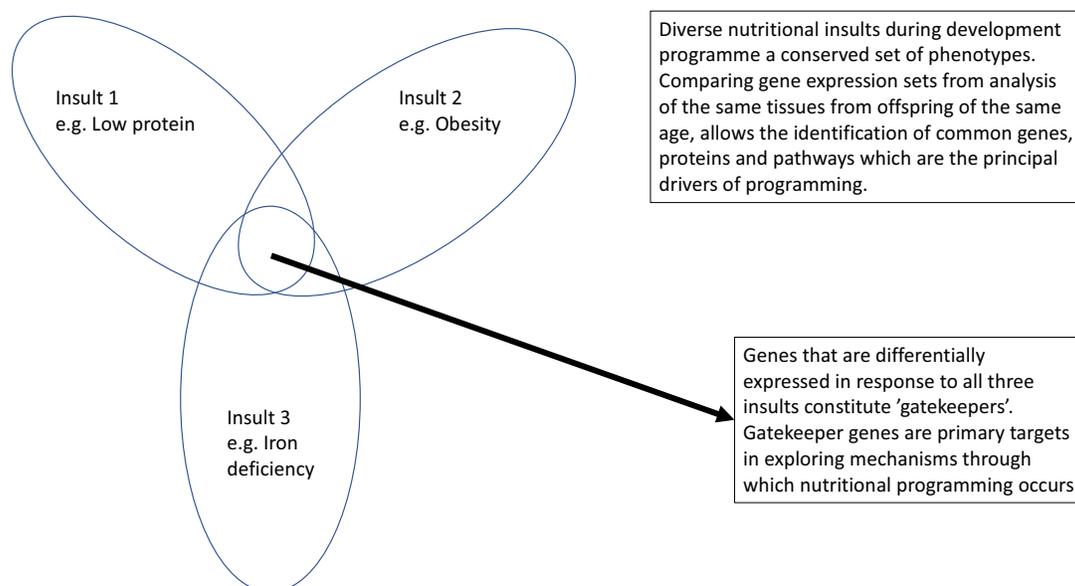


FIGURE 15.3 The gatekeeper approach to explore mechanisms of early life programming. Exposure of rat fetuses to diverse nutritional insults in development results in a common phenotype comprising high blood pressure, renal impairment, and metabolic disturbance. Using a transcriptomics approach to identify genes that are differentially expressed in response to all three insults, it is possible to identify sets of gatekeeper genes, proteins, and pathways that play an integral role in the nutritional programming effect. *Adapted from McMullen, S., Langley-Evans, S.C., Gambling, L., Lang, C., Swali, A., McArdle, H.J., 2012. A common cause for a common phenotype: the gatekeeper hypothesis in fetal programming. Med Hypotheses 78, 88–94.*

programming, and by comparing effects of diverse dietary insults upon the fetal transcriptome it should be possible to identify “gatekeeper” genes, proteins, pathways, and processes that are responsible for tissue remodeling and its sequelae.

The search for potential gatekeepers was reported by Swali et al. (2011) who collected day 13 gestation rat embryos that had been subjected to either maternal LP feeding or maternal iron deficiency. Adult rats exposed to both protocols exhibited the common phenotype of reduced nephron number. Microarray analysis of gene expression in the whole embryo (minus head) found that in the embryos over 2200 genes were differentially regulated with LP and iron deficiency. The study identified that 67 genes were differentially expressed (40 upregulated and 27 downregulated) in both maternal dietary conditions, and these were designated as putative gatekeeper genes. Pathway analysis showed strong enrichment of pathways associated with cytoskeletal remodeling and regulation of the cell cycle. Both processes are of course highly involved in the proliferation and differentiation of cells during embryonic and fetal growth and so can be readily associated with tissue remodeling.

The studies of gatekeeper genes in embryos are well-matched by the findings of Altobelli et al. (2013), whose microarray analysis focused upon the hepatic transcriptome in neonatal rats exposed to LP in utero. With LP exposure, 131 genes were differentially expressed. These

related to processes associated with DNA maintenance, repair, and cell cycle mechanisms. Interestingly, the differential expression associated with LP was abrogated by supplementation of the LP diet with folic acid.

When considered together, these studies suggest that maternal protein restriction limits the capacity of embryonic and fetal cells to divide, potentially due to inhibition of DNA repair mechanisms, resulting in a greater proportion of cells with damaged DNA or DNA that has not been faithfully replicated. This outcome would result in a higher proportion of cells arresting at the G1/S and G2/M checkpoints during cell division. Interestingly, the study of Swali et al. (2011) noted that genes associated with these checkpoints (notably *Taf10*, *Apbb1*, *Mdm2*, *Pmp22*, *Wee1*, and *Tsg101*) were downregulated with maternal protein restriction. As these genes essentially control the phased waves of cell division and apoptosis that normally occur during development, protein restriction (possibly operating through negative effects on DNA integrity or the formation of cytoskeletal elements that make up the mitotic spindle) may be determining cell number and hence size, structure, and functions of organs.

PROTEIN AND THE DNA METHYLOME

The view that fetal programming of adult disease is a consequence of resetting of epigenetic marks is highly

avored in the field, and an array of evidence showing that DNA methylation or histone modifications at specific gene loci are responsive to maternal dietary insults supports this hypothesis (Lillycrop and Burdge, 2015; Calder et al., 2018; Langley-Evans, 2015). For example, Bogdarina et al. (2010) showed that methylation of the angiotensin II type 1b receptor in adrenal, and subsequently the expression of the gene, was sensitive to maternal LP feeding in a glucocorticoid-sensitive manner. The epigenetic hypothesis is attractive as epigenetic marks can be passed on to subsequent generations through both male and female gametes. A number of studies have shown that the programming effects of maternal undernutrition, including LP, are persistent across two or even three generations. Moreover, research in which pregnant animals are fed diets that are lacking in methyl donors required to maintain epigenetic marks typically finds that offspring are programmed for adverse cardiovascular and metabolic phenotypes (Calder et al., 2018).

Although many studies have focused upon nutrition-induced changes to DNA methylation and expression of specific genes, there has been very little experimental work to evaluate the impact of nutrition in pregnancy upon the global methylome. Altobelli et al. (2013) used an MBD sequencing approach to evaluate the methylation of the whole genome and determine differential methylation in neonatal liver DNA from rats exposed to maternal LP during fetal development. As shown in Fig. 15.4, this approach identified 555 differentially methylated loci that could be mapped to known coding regions (mapped to nearest gene). Genomic annotation indicated that the differentially methylated regions were largely distributed within genes (40%) and between genes (58%) with the remainder overlapping exons or 5'/3' UTRs.

The finding that LP in fetal development could have widespread effects on the methylome was of major interest. These effects were explained by the observation that LP had a programming effect upon the expression of DNA methyltransferases that are responsible for both the establishment and the maintenance of methylation. However, when the differentially methylated regions were mapped against expression as measured by microarray, there were very few loci that were found to be both differentially methylated and differentially expressed (Fig. 15.4). This result tends to function against the hypothesis that epigenetic changes play a direct role in the programming of later physiological and metabolic function through tissue remodeling. Remodeling is a direct response to the prevailing environment and could well be driven by availability of substrates (deficiency or excess) and the hormonal environment. However, the epigenetic changes may have a more subtle long-term impact. Increased methylation will suppress gene expression, but to researchers considering this in tissues, the impact of the methylation will only be apparent if the tissue *needs* to express that gene. In other words, the differential methylation that is associated with LP or other forms of maternal nutritional insult in pregnancy may only have a functional impact when the adult environment (e.g., dietary intake, infection, or inflammation) poses a challenge requiring a cellular response.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Maternal diet is a powerful determinant of lifelong health and well-being. Protein restriction in pregnancy has been shown to determine blood pressure, insulin

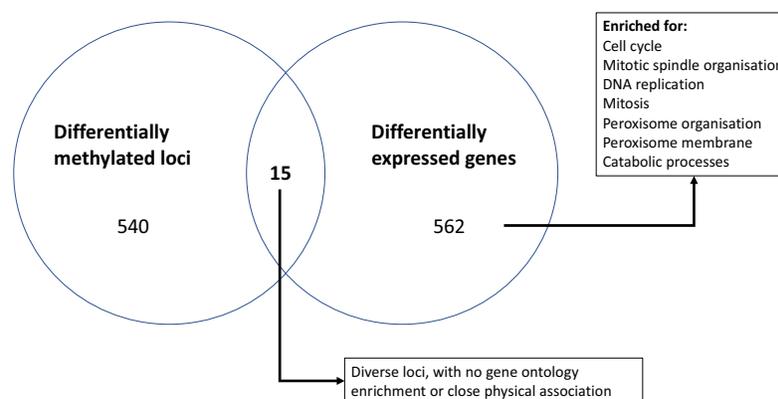


FIGURE 15.4 Mismatch between DNA methylation and gene expression. Analysis of DNA methylation across the whole genome of neonatal rat livers following exposure to maternal protein restriction in utero showed 555 differentially methylated loci. In the same tissue, protein restriction was associated with differential expression of 577 genes. There was little overlap between the differentially methylated and differentially expressed sets. Data redrawn from Altobelli, G., Bogdarina, I.G., Stupka, E., Clark, A.J.L., Langley-Evans, S., 2013. Genome-wide methylation and gene expression changes in newborn rats following maternal protein restriction and reversal by folic acid. *PLoS One* 8.

responses, lipogenesis, and longevity in rodents. These effects can be ameliorated by supplementing the low-protein diet with methyl donors, suggesting that epigenetic regulation may be mechanistically important; or by blockade of maternal glucocorticoid transport across the placenta, suggesting that diet-endocrine interactions contribute to the mechanism. Using high-throughput techniques to determine the effects of protein restriction on gene expression in different tissues, fundamental processes such as the regulation of the cell cycle have been linked to morphological observations that suggest the pregnancy diet can irreversibly shape the development of fetal organs and tissues and hence long-term function.

One of the current challenges is to translate understanding of the mechanistic links between early diet and later disease into new approaches for disease prevention or treatment. The identification of genes, proteins, and pathways that consistently respond to maternal diet provides important clues to where future translational studies should focus. The recent observation that the placenta shows major gene expression changes in a relatively small number of pathways when associated with both maternal under- and overnutrition may be a step toward developing biomarkers that indicate the programming events have occurred during fetal development. This could be the basis for future personalized medicine approaches that draw together the disease risk factors associated with genetic inheritance, early life programming, and adult lifestyle to select preventive and therapeutic management of health.

Moving this area forward poses some significant research challenges. A nutrigenetics approach is relatively straightforward, and, as the previous discussion has shown, it is relatively straightforward to identify gene expression changes in placenta or offspring tissue in response to maternal diet. Follow-up studies using gene editing and knockout mouse strains can then assess the functional significance of the expression of specific genes or pathways as early determinants of long-term health. Nutriepigenetic studies are considerably more challenging, as functional significance of specific variation in DNA or histone modifications is currently impossible to demonstrate *in vivo*. Both approaches are also made more challenging by the nature of the programming phenomenon. Maternal diet elicits a response in the developing embryo or fetus, and that response results in permanent changes to physiology and metabolism, presumably through remodeling of tissue. The health effects of those changes only manifest with aging, so in humans there will be a gap of 6 decades between programming event and measurable impact. The precise timing of the fetal response is not known in any of the model systems studied so far and could well be

extremely transient, potentially even a delay or bringing forward in the expression of developmentally critical genes of just a few hours. For example, studies of mouse fetuses collected from mothers fed a low-protein diet showed delay in limb development and branching morphogenesis at embryonic day 12, followed by rapid catch-up on day 13 (Langley-Evans et al., 2011). Furthermore almost all current and previously published studies evaluate the impact of maternal nutrition at whole organ or tissue level and do not assess programming effects in specific cell types.

The big promise of this field of research has been to deliver evidence-based improvements in maternal nutrition and to develop new approaches to disease prevention. Researchers in the programming field have understandably focused on the easy-to-measure observations that long-term changes in health and disease stem from early nutritional insults and have collected extensive evidence of the, rather secondary, mechanisms that generate pathology. In order to translate current knowledge into impactful change, experiments focused on embryos, perhaps grown *in vitro*, should be prioritized. This will give the mechanistic understanding of nutritional programming that is a prerequisite for developing novel therapeutic and preventive strategies.

References

- Altobelli, G., Bogdarina, I.G., Stupka, E., Clark, A.J.L., Langley-Evans, S., 2013. Genome-wide methylation and gene expression changes in newborn rats following maternal protein restriction and reversal by folic acid. *PLoS One* 8.
- Bogdarina, I., Haase, A., Langley-Evans, S., Clark, A.J.L., 2010. Glucocorticoid effects on the programming of AT1b angiotensin receptor gene methylation and expression in the rat. *PLoS One* 5.
- Calder, P.C., Carding, S.R., Christopher, G., Kuh, D., Langley-Evans, S.C., McNulty, H., 2018. A holistic approach to healthy ageing: how can people live longer, healthier lives? *J Hum Nutr Diet* 31, 439–450.
- Daniel, Z., Swali, A., Emes, R., Langley-Evans, S.C., 2016. The effect of maternal undernutrition on the rat placental transcriptome: protein restriction up-regulates cholesterol transport. *Genes Nutr* 11, 27.
- Erhuma, A., Salter, A.M., Sculley, D.V., Langley-Evans, S.C., Bennett, A.J., 2007. Prenatal exposure to a low-protein diet programs disordered regulation of lipid metabolism in the aging rat. *Am J Physiol Endocrinol Metab* 292, E1702–1714.
- Hay, S.M., McArdle, H.J., Hayes, H.E., Stevens, V.J., Rees, W.D., 2016. The effect of iron deficiency on the temporal changes in the expression of genes associated with fat metabolism in the pregnant rat. *Phys Rep* 4 pii: e12908.
- Langley-Evans, S.C., Daniel, Z.C., Wells, C.A., Ryan, K.J., Plant, R., Welham, S.J., 2011. Protein restriction in the pregnant mouse modifies fetal growth and pulmonary development: role of fetal exposure to [beta]-hydroxybutyrate. *Exp Physiol* 96, 203–215.
- Langley-Evans, S.C., 2015. Nutrition in early life and the programming of adult disease: a review. *J Hum Nutr Diet* 28 (Suppl. 1), 1–14.
- Lillicrop, K.A., Burdge, G.C., 2015. Maternal diet as a modifier of offspring epigenetics. *J Dev Orig Health Dis* 6, 88–95.

- McMullen, S., Langley-Evans, S.C., Gambling, L., Lang, C., Swali, A., McArdle, H.J., 2012. A common cause for a common phenotype: the gatekeeper hypothesis in fetal programming. *Med Hypotheses* 78, 88–94.
- Swali, A., McMullen, S., Hayes, H., Gambling, L., McArdle, H.J., Langley-Evans, S.C., 2011. Cell cycle regulation and cytoskeletal remodelling are critical processes in the nutritional programming of embryonic development. *PLoS One* 6.
- Vithayathil, M.A., Gugusheff, J.R., Ong, Z.Y., Langley-Evans, S.C., Gibson, R.A., Muhlhausler, B.S., 2018. Exposure to maternal cafeteria diets during the suckling period has greater effects on fat deposition and Sterol Regulatory Element Binding Protein-1c (SREBP-1c) gene expression in rodent offspring compared to exposure before birth. *Nutr Metab* 15, 17.

This page intentionally left blank

Maternal Perinatal Nutrition and Offspring Programming

Paul Cordero¹, Manlio Vinciguerra^{1,2}, Jude A. Oben^{1,3}

¹Institute for Liver and Digestive Health, University College London, London, United Kingdom; ²Center for Translational Medicine (CTM), International Clinical Research Centre (ICRC), St. Anne's University Hospital, Brno, Czech Republic;

³Department of Gastroenterology and Hepatology, Guy's and St. Thomas' Hospital, NHS Foundation Trust, London, United Kingdom

Glossary

Epigenetics Study of the molecular machinery around the genetic code, which, without altering the nucleotide sequence, regulates the transcriptomic profile.

Maternal programming Influence of maternal physiological, metabolic, and environmental profile during periconceptual, in utero, and lactation periods on offspring predisposition to respond to different stimuli during the life course.

Nutrieigenetics Study of the mechanisms by which different epigenetic profiles affect nutrient metabolism.

Nutrieigenomics Study of the mechanisms by which nutrients regulate the epigenetic profiles.

Paternal programming Influence of paternal physiological and metabolic status through the sperm on offspring predisposition to respond to different stimuli during the life course.

Perinatal programming Offspring predisposition to respond to different stimuli during different periods of life depending on maternal environmental insights during periconceptual, in utero, and lactation periods.

lactation, the physiological and metabolic adaptations lead to changes of normal macro- and micronutrient requirements (Green and Hester, 2016; Gluckman and Hanson, 2006; Ramakrishnan et al., 2014). During pregnancy there are circulatory system modifications whose main objective is to provide adequate placental circulation and fetal nutrient supply. Moreover, there is increased blood volume, higher iron requirements, basal metabolism and body weight, together with decreased hematocrit due to hemodilution and hormonal alterations (cortisol, thyroid hormones T3 and T4, adrenocorticotropic hormone or prolactin). During lactation, besides endocrine changes such as increased prolactin and oxytocin hormones in plasma, there is also increased basal metabolism and calcium lost, among other physiometabolic modifications. These maternal transformations require nutritional adjustments.

During pregnancy, the interactions between maternal nutrition and genetic background determine fetus nutrient supply, which also interacts with fetal genetic information. Therefore, there is a double nutrient-gene interaction between the mother and the fetus (Fig. 16.1). Furthermore, the term *perinatal* encompasses different time frames and, thus, different physiological requirement in each one of them. First, during a preconception period, maternal nutrition may affect the oocyte and its fecundation, with possible consequences on the future development of the offspring. Thus, novel research has highlighted that dietary polyphenols appear to inhibit oocyte apoptosis and follicle atresia during ovarian development (Santangelo et al., 2014). Other investigators have shown that the maturation of gametes and early embryonic development can be

INTRODUCTION

The suitable interactions between nutrient intake, in both quality and quantity, and genetic background are crucial for maintaining a healthy metabolic and physiological condition across different lifetime periods. However, nutritional requirements, parallel to physiological condition, are adapted to the different biological necessities from each lifetime stage and environmental situation. It is understood, for example, that the requirements of an adolescent during fast-growth stages are not the same as those of adults or elderly as well as during special situations or diseases. Thus, during perinatal periods, including preconception, pregnancy, and

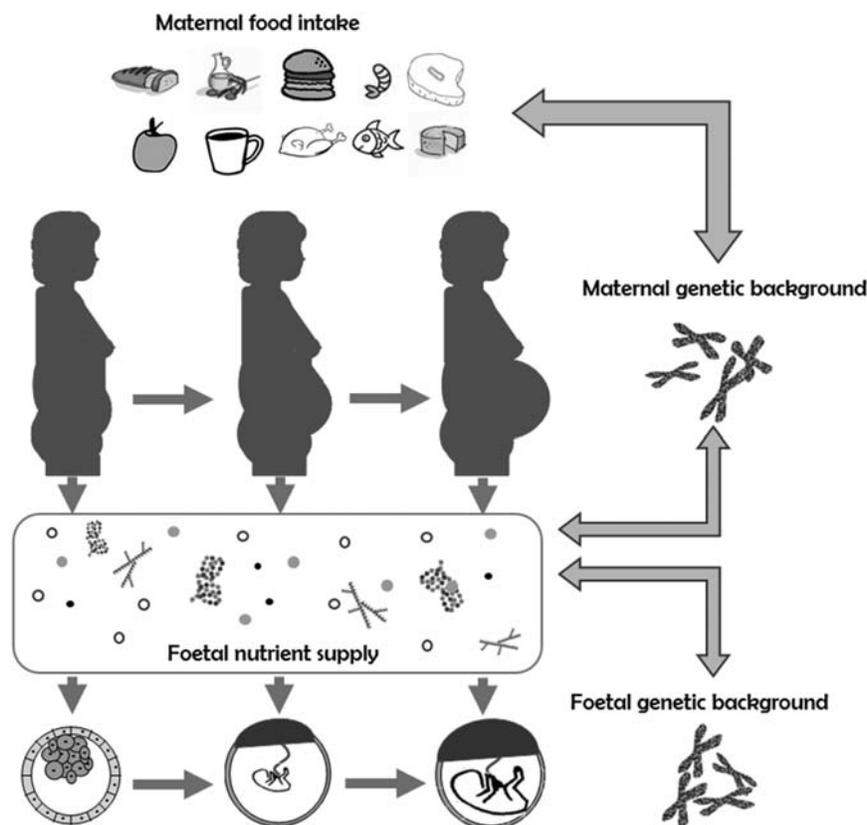


FIGURE 16.1 Nutritional interactions during pregnancy.

compromised in malnutrition situations leading to early pregnancy failure and postnatal increased risks of offspring cardiovascular, metabolic, immune, and neurological morbidities (Fleming et al., 2018). Obesity has been also related with impaired oocyte meiosis, failure of spindle assembly, and increased oxidative stress and abnormal mitochondrial distribution (Jia et al., 2018). After fertilization, fetal development depends on the maternal dietary intake, metabolism, and physiologic features. During this phase of in utero environment, the availability of triglycerides, amino acids, cholesterol, glucose, vitamins, and other metabolites in the maternal plasma leads the development of the fetus. For example, maternal fatness and/or obesogenic nutrition exposes the offspring to a diseased environment (hyperinsulinemia, hyperleptinemia, hypercholesterolemia, etc.) associated with obesity comorbidities, programming offspring physiology and increasing the susceptibility to develop metabolic diseases during adulthood (Green and Hester, 2016). Finally, during early stages after birth, mainly during the lactation period, there is also an effect of nutrient supply in the offspring development and risk of suffering later metabolic diseases during adulthood. In this sense, there is a substantial volume of experimental and observational

data demonstrating the importance of lactation for the development of metabolic diseases such as cardiovascular disease, insulin resistance, or nonalcoholic fatty liver disease (NAFLD), among others (Campion et al., 2009; Jousse et al., 2011; Cordero et al., 2017).

The importance of maternal nutrition in offspring development was initially stated by David Barker and collaborators as the theory of developmental origins of health and disease (DOHaD) (Gluckman and Hanson, 2006). They demonstrated that maternal nutrition during pregnancy affects offspring metabolic singularities and the risk of suffering diseases during different time frames. Changes induced by maternal inputs may follow an adaptive response of the fetus acting through the process of developmental programming, which predispose to long-term changes in its metabolism and physiology to be adapted to the future environment. These offspring adaptations to maternal nutrition are associated with specific transcriptomic profiles across time resulting in the alteration of metabolic pathways and homeostatic regulatory processes (Green and Hester, 2016; Gluckman and Hanson, 2006). First epidemiological studies demonstrating DOHaD theory were focused on historical famines from the 20th century such as the Dutch famine during World War II, the

famine of the siege of Leningrad (September 1941–January 1944), or the Chinese famine of 1959–61. These studies found that individuals whose mothers were perinatally exposed to very low calorie intake exhibited an increased risk of developing cardiovascular diseases, insulin resistance, NAFLD, and other obesity-related comorbidities in adulthood, as well as cognitive problems or some types of cancer. Moreover, the specific perinatal stage in which the famine was suffered may affect differently, but always increasing the prevalence of disorders in the offspring. As a possible explanation, during famine, the fetus should adapt its physiology and metabolism to effective nutrient conservation and storage. However, if later environment presents a normal or augmented quantity of nutrients, the initially safe adaptation will lead to a diseased pathophysiological situation associated with an excess of energy supply and accumulation. Following these initial epidemiological studies that focused on perinatal nutritional deficiencies, there are other nutritionally related outcomes that are attempting to unravel next-generation diseases. Indeed, the current century nutritional profile is driven by socioeconomic changes, migrations from rural to urban areas, or Western pro-obesogenic dietary intake, among other factors.

Besides the direct interactions between nutrients and genetic background, the translation machinery also seems to be regulated by epigenetic mechanisms, which determine the cell- and time-specific manner of the transcriptomic profiles and may also be nutritionally driven (Green and Hester, 2016; Cordero et al., 2015).

MECHANISMS OF PERINATAL PROGRAMMING

Nutrigenetics and nutrigenomics study how the information contained in the genome affect the response to nutrients and how nutrients can regulate gene expression, respectively. These dual interactions are maintained during the life course and, therefore, during pregnancy and other perinatal periods. Furthermore, as previously commented, during these stages, there is a combinatory effect of maternal nutrition and genetic background affecting offspring nutrient availability and genome. However, beyond the genome nucleotide sequence, maternal and offspring genetic material is conditioned by their epigenetic profiles. Epigenetics involves a molecular machinery around the gene sequence, which regulated its transcription without altering the nucleotide sequence (Campion et al., 2009; Cordero et al., 2015). The main epigenetic modifications are DNA methylation, histone modifications, microRNAs, and long noncoding RNAs (lncRNAs). The

most studied, understood, and potentially therapeutically targeted is the DNA methylation, which is based in the addition of a methyl group to a cytosine having a guanine as next nucleotide (CpG site). This covalent reaction is mediated by the DNA methyltransferases (Dnmts). These CpG sites are abundant in the promoter region of the genes, where transcription machinery joins the DNA sequence in order to regulate genome reading. According to the histone marks, covalent reactions (e.g., acetylation, phosphorylation, methylation, ubiquitination) in their amino acidic tails regulate the chromatin conformation and accessibility of the transcriptional machinery to the DNA sequence. MicroRNAs and lncRNAs are short chains of nucleotides and nonprotein coding transcripts longer than 200 nucleotides, respectively, which are able to modulate posttranscriptional regulation of gene expression. The combination of these reactions will determine the cellular physiology and metabolism through affecting gene expression patterns in a cell-specific manner, whilst preserving the nucleotide sequences (Green and Hester, 2016; Campion et al., 2009; Cordero et al., 2015).

Early life nutritional exposure leads to permanent changes in gene expression profiles through the epigenetic regulation of genetic code, leading to an altered risk of suffering metabolic diseases. Thus, increased cardiovascular risk in adult offspring has been associated with changes in expression and methylation patterns of the genes 11 β -hydroxysteroid dehydrogenase type 2 (HSD2), glucocorticoid receptor, and insulin-like growth factor 2 due to maternal unbalanced diets. Others have described that maternal obesogenic feeding-associated glucose homeostasis and insulin resistance modulates hypothalamic proopiomelanocortin (POMC) methylation in offspring, which was also associated with an increase in POMC transcription levels. During offspring development in pregnancy, the accelerated cellular proliferation together with high rates of cellular and tissue differentiation, increase the exposure of the DNA sequence to environmental insights, promoting a high plasticity period in which epigenetic profile is settled. A very interesting example of the effect of nutrition in epigenetic regulation of the physiological status come from studies in honeybees; although all larvae are genetically identical, those fed during very early stages with the royal jelly exhibit epigenetic alterations developing into fertile queen bees. Therefore, different cell types will be epigenetically programmed to express their specific transcriptomic profiles associated with their function in the new organism. Interestingly, when affecting germ line cells, these alterations can be transmitted across successive generations, extending transgenerationally the implications of maternal nutritional programming. Hence, a maternal high-fat diet perinatal

intervention metabolic syndrome features in offspring across five generations, which was parallel to altered histone marks and expression patterns of leptin and adiponectin in adipose tissue (Green and Hester, 2016; Campion et al., 2009; Cordero et al., 2015).

This line of sight allows to state that there is an epigenetic regulation of genes that participate in nutrient metabolism and that the nutrient intake promotes changes in epigenetic profile, stated as nutriepigenetic and nutriepigenomic mechanisms (Fig. 16.2). Some studies have described how differences in DNA methylation patterns are associated with diverse dietary responses, which have been applied as epigenetic biomarkers in weight-loss dietary-based treatments in obese patients. For example, DNA methylation levels of the proinflammatory cytokine tumor necrosis factor alpha in subcutaneous adipose tissue predicted a hypocaloric diet-induced weight loss (Cordero et al., 2011). Moreover, perinatal undernutrition affected the methylation and expression of the leptin gene and the predisposition of developing the metabolic syndrome throughout the lifespan of adult mice, emphasizing that maternal protein undernutrition affects the balance between food intake and energy expenditure in adults (Jousse et al., 2011). Methylation profiled of other obesity-related genes in blood and in adipose tissue has been also described as potential biomarkers for personalized nutrition treatments based on epigenetic criteria. Furthermore, due to the reversibility of epigenetic marks, current research is focused on their target by nutritional compounds for different diseases.

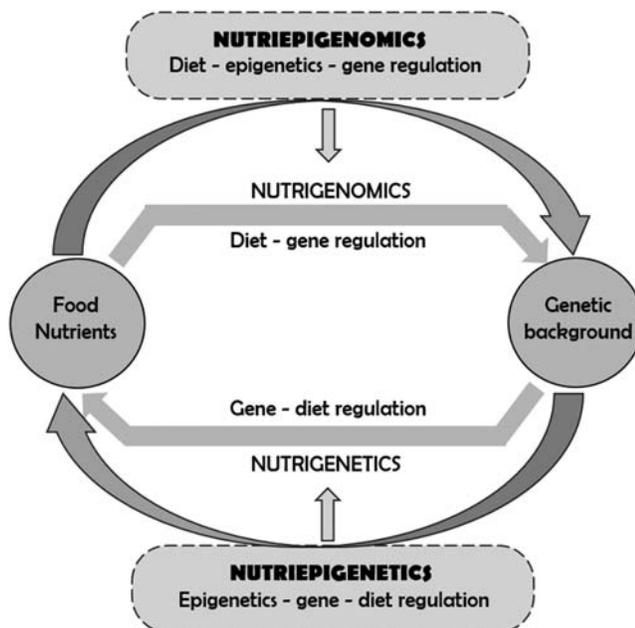


FIGURE 16.2 Nutriepigenetic and nutriepigenomic regulation.

MACRONUTRIENT AND MICRONUTRIENT EXPOSURE

Due to the current Western-style dietary patterns, there is a new consumption profile of dietary macronutrient composition characterized by higher intake of lipids and simple sugars, parallel to increased overall calorie consumption. Observational data regarding the mode that this environment affects offspring development remain limited. As an obesity pandemic and the associated current dietary patterns have been gaining importance only in the last decades, there is not enough follow-up data for the offspring during their adult life. A study in obese woman with diet-induced weight loss prior to pregnancy reprogrammed hepatic DNA methylation patterns turning to healthy expression of genes related to lipid metabolism in their offspring. This research highlights how nutritionally driven changes in maternal physiology potentially influence the metabolism of the offspring by affecting their epigenetic profiles. Others have reported an association between body mass and methylation levels of the energy and cholesterol metabolism homeostasis peroxisome proliferator-activated receptor-gamma coactivator 1- α gene in cord blood (Gemma et al., 2009).

Initial observational human studies investigating undernutrition during fetal development demonstrated that the reduced supply of nutrients in early live environment leads to changes in DNA methylation patterns. There has been described hypomethylation of the insulin growth factor 2 (IGF-2) gene sequence in offspring from mothers suffering famine during pregnancy. Methylation patterns of IGF-2 have been also reported to be altered during periconceptual famine as well as in diets with lower concentrations of folic acid during early pregnancy. A genome-wide DNA methylation analysis in whole blood from adult offspring after periconceptual exposure to famine reported differentially methylated regions located in insulin receptor and carnitine aminotransferase genes, which are associated with glucose homeostasis regulation and lipid metabolism. Other researchers have highlighted the importance of the early gestation compared with mid or late gestation as the most critical time period for offspring epigenetic programming on genes related with lipid and amino acids synthesis and transport in whole blood cells after prenatal exposure to famine (Green and Hester, 2016; Cordero et al., 2015; Moraru et al., 2018). However, after these initial epidemiological-based approaches focused on famines and, parallel to the currently growing increasing obesity rates, the main interest has moved through maternal perinatal obesity and offspring developmental programming. This research is crucial since

during the last 3 decades obesity rates in women at pregnancy age have grown in developed countries from around 10% to up to 35%.

Furthermore, in animal models, the offspring of dams fed obesogenic diets during pregnancy exhibit a higher risk of suffering metabolic diseases, including cardiovascular disease, insulin resistance, and fatty liver, among others. Studies in baboons have reported that obesogenic diet prior to and after conception induces significant alterations in cardiac microRNAs levels that may be associated with cardiovascular disease (Maloyan et al., 2013). Other studies in nonhuman primates demonstrated that maternal obesogenic feeding led to fetal hepatic apoptosis and that elevated maternal insulin levels influenced offspring NAFLD development, being irreversible despite that the descendants were weaning into a healthy diet (Grant et al., 2011).

Epigenetic modifications can also be driven by micronutrient nutritional status. The dietary intake of molecules involved in the homocysteine-methionine cycle or one-carbon metabolism such as vitamin B₁₂, folic acid, methionine, choline, or betaine potentially regulates variations in DNA methylation patterns during pregnancy and lactation and, subsequently, the risk of suffer diseases such as obesity, cardiovascular disease, and NAFLD (Fig. 16.3). Although mechanistically unclear, there is abundant evidence associating the deficit of folic acid intake during periconception and early pregnancy with offspring that have neural tube defects. These malformations are also associated with polymorphisms from genes involved in one-carbon metabolism and with group B vitamins, choline, betaine, and fatty acid maternal nutritional intake (Green and Hester, 2016). Furthermore, maternal supplementation with a one-carbon metabolism-related micronutrient cocktail

in rodent models has been proved to prevent maternal obesity-induced hyper-homocysteinemia and hepatic triglyceride accumulation in the offspring. This protection is parallel to changes in DNA methyltransferases activity and hepatic total DNA methylation. In humans, a randomized controlled trial of periconceptional micronutrient supplementation including folate, zinc, and vitamins showed a reduction of insulin growth factor receptor 2 in girls and maternally expressed 3 in boys in methylation levels from cord blood samples (Cooper et al., 2012).

CURRENT CHALLENGES AND FUTURE APPLICATIONS

Most of the current knowledge and new data in the field of maternal programming and perinatal nutrition are based on findings from animal studies. Rodent experimental models provide direct evidence joining maternal nutritional exposures and developmental programming through epigenetic-mediated changes in gene expression profiles, and subsequent phenotypes. Besides the ethical considerations of human nutritional interventions during pregnancy, rodent models shorten the experimental time, allowing also to study the effect of nutrition during several generations and the responses to different nutritional and environmental stimuli during offspring adult life. Furthermore, the economic cost, genetic and physiologic similarities with humans, and the better control of external insults and confounding factors make it necessary to perform experimental animal models in this field. With these research tools, the amount of data regarding the effect of different dietary macro- and micronutrient composition during

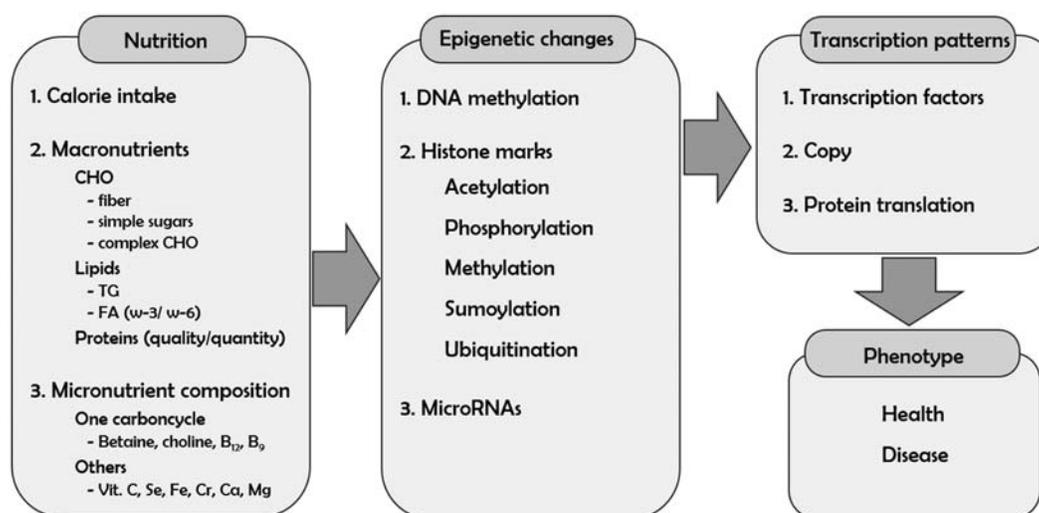


FIGURE 16.3 Nutritional insults affect phenotypical characteristics by epigenetic modifications (C, carbon; Ca, calcium; CHO, carbohydrates; Cr, chromium; FA, fatty acids; Fe, iron; Mg, magnesium; Se, selenium; TG, triglycerides; Vit, vitamin).

pregnancy on developmental programming and adult life is exponentially growing. However, these data should be contrasted with human observational studies before nutritional interventions.

It is important to emphasize that parental programming includes both maternal and paternal influences. Hence, besides maternal nutritional influence during pregnancy, paternal programming influence on offspring metabolism seems to be driven by the paternal epigenetic profile carried by the sperm. This epigenetic information can be nutritionally and physiologically affected. Thus, paternal obesogenic diets and associated phenotype prior to conception have been linked with DNA methylation changes in the descendants in leukocytes extracted from cord blood. Furthermore, animal studies have demonstrated that the percentage of paternal protein intake in the diet is associated with changes in the hepatic expression of fat and cholesterol-related metabolism genes. This research field, focused on paternal nutrition and offspring developmental programming, is currently gaining interest due to the lack of previous knowledge studies and its potential influence on offspring health (Green and Hester, 2016; Hur et al., 2017; Preston et al., 2018).

Epigenetic regulation can be nutritionally driven to get therapeutic interventions. Given the recent interest in unraveling the human epigenome, the important advances in high-throughput techniques, and the current awareness about maternal dietary influences on offspring adult programming, we are now closer to the application of nutritional or nutri-pharmaceutical interventions during pregnancy in order to decrease the risk of suffering metabolic diseases in offspring adulthood (Ramakrishnan et al., 2014). Although general nutrition guidelines during perinatal periods are mostly based on epidemiological evidence, there is still a lack of knowledge regarding the molecular mechanisms implicated in developmental programming. It seems that epigenetic mechanisms due to the plasticity of the epigenome during perinatal stages can be the main influencers in offspring transcriptomic programming. Thus, understanding the different nutriepigenetic profiles affecting nutrient metabolism as well as how to nutriepigenomically regulate them will allow to effectively apply a personalized nutrition treatment to decrease the risk of suffering diseases in the descendants.

In conclusion, although our knowledge of nutri-genetics and nutrigenomics has rapidly progressed in different lifetime frames, there is a lack of data on human studies during perinatal periods due to the difficulty of carrying out safe interventional studies on pregnant women. Taking into account that there is a double effect and interaction with the maternal and fetal

background and that the molecular mechanisms involved seem to go beyond the DNA sequences (through epigenetics) there is still needed more epidemiological and basic science data in order to apply the knowledge in this area to the common population maternal perinatal nutrition. Nevertheless, it is now also clear from meta-analyses that prenatal calcium supplementation, omega-3 fatty acids, and multiple micronutrient supplements reduce the risk of preterm birth, and it is expected that solid nutritional indications will come in the near future from ongoing research.

References

- Campion, J., Milagro, F.I., Martinez, J.A., 2009. Individuality and epigenetics in obesity. *Obes Rev* 10, 383–392.
- Cooper, W.N., Khulan, B., Owens, S., Elks, C.E., Seidel, V., Prentice, A.M., Beltaki, G., Ong, K.K., Affara, N.A., Constancia, M., Dunger, D.B., 2012. DNA methylation profiling at imprinted loci after periconceptional micronutrient supplementation in humans: results of a pilot randomized controlled trial. *FASEB J* 26, 1782–1790.
- Cordero, P., Campion, J., Milagro, F.I., Goyenechea, E., Steemburgo, T., Javierre, B.M., Martinez, J.A., 2011. Leptin and TNF-alpha promoter methylation levels measured by MSP could predict the response to a low-calorie diet. *J Physiol Biochem* 67, 463–470.
- Cordero, P., Li, J., Oben, J.A., 2015. Epigenetics of obesity: beyond the genome sequence. *Curr Opin Clin Nutr Metab Care* 18, 361–366.
- Cordero, P., Li, J., Nguyen, V., Pombo, J., Maicas, N., Novelli, M., Taylor, P.D., Samuelsson, A.M., Vinciguerra, M., Oben, J.A., 2017. Developmental programming of obesity and liver metabolism by maternal perinatal nutrition involves the melanocortin system. *Nutrients* 9, E1041.
- Fleming, T.P., Watkins, A.J., Velazquez, M.A., Mathers, J.C., Prentice, A.M., Stephenson, J., Barker, M., Saffery, R., Yajnik, C.S., Eckert, J.J., Hanson, M.A., Forrester, T., Gluckman, P.D., Godfrey, K.M., 2018. Origins of lifetime health around the time of conception: causes and consequences. *Lancet* 391, 1842–1852.
- Gemma, C., Sookoian, S., Alvarinas, J., Garcia, S.I., Quintana, L., Kanevsky, D., Gonzalez, C.D., Pirola, C.J., 2009. Maternal pregestational BMI is associated with methylation of the PPARC1A promoter in newborns. *Obesity* 17, 1032–1039.
- Gluckman, P., Hanson, M., 2006. *Developmental Origins of Health and Disease*. Cambridge University Press, Cambridge.
- Grant, W.F., Gillingham, M.B., Batra, A.K., Fewkes, N.M., Comstock, S.M., Takahashi, D., Braun, T.P., Grove, K.L., Friedman, J.E., Marks, D.L., 2011. Maternal high fat diet is associated with decreased plasma n-3 fatty acids and fetal hepatic apoptosis in nonhuman primates. *PLoS One* 6, e17261.
- Green, L.R., Hester, R.L., 2016. *Parental Obesity: Intergenerational Programming and Consequences*. Springer-Verlag, New York.
- Hur, S.S., Croypley, J.E., Suter, C.M., 2017. Paternal epigenetic programming: evolving metabolic disease risk. *J Mol Endocrinol* 58, R159–R168.
- Jia, Z., Feng, Z., Wang, L., Li, H., Wang, H., Xu, D., Zhao, X., Feng, D., Feng, X., 2018. Resveratrol reverses the adverse effects of a diet-induced obese murine model on oocyte quality and zona pellucida softening. *Food Funct* 9, 2623–2633.
- Jousse, C., Parry, L., Lambert-Langlais, S., Maurin, A.C., Averous, J., Bruhat, A., Carraro, V., Tost, J., Letteron, P., Chen, P., Jockers, R., Launay, J.M., Mallet, J., Fafournoux, P., 2011. Perinatal

- undernutrition affects the methylation and expression of the leptin gene in adults: implication for the understanding of metabolic syndrome. *FASEB J* 25, 3271–3278.
- Maloyan, A., Muralimanoharan, S., Huffman, S., Cox, L.A., Nathanielsz, P.W., Myatt, L., Nijland, M.J., 2013. Identification and comparative analyses of myocardial miRNAs involved in the fetal response to maternal obesity. *Physiol Genom* 45, 889–900.
- Moraru, A., de Almeida, M.M., Degryse, J.M., 2018. PALTEM: what parameters should be collected in disaster settings to assess the long-term outcomes of famine? *Int J Environ Res Public Health* 15, E857.
- Preston, J.D., Reynolds, L.J., Pearson, K.J., 2018. Developmental origins of health span and life span: a mini-review. *Gerontology* 64, 237–245.
- Ramakrishnan, U., Imhoff-Kunsch, B., Martorell, R., 2014. Maternal nutrition interventions to improve maternal, newborn, and child health outcomes. *Nestle Nutr Inst Workshop Ser* 78, 71–80.
- Santangelo, C., Vari, R., Scaccocchio, B., Filesi, C., Masella, R., 2014. Management of reproduction and pregnancy complications in maternal obesity: which role for dietary polyphenols? *Biofactors* 40, 79–102.

This page intentionally left blank

Introduction to Epigenetics

Lara K. Hyde¹, Simonetta Friso², Sang-Woon Choi³

¹Nourishable.tv, Litchfield, NH, United States; ²University of Verona, School of Medicine, Verona, Italy; ³Chaum Life Center, CHA University, Seoul, Korea

Glossary

Chromatin The packaging of DNA wrapped around histones within the nucleus.

CpG site A cytosine flanked by a guanine to yield 5'-cytosine-phosphate-guanine-3'.

CpG island Regions that are 300–3000 bp long with an increased proportion of CpG sites compared with the rest of the genome. Frequently, CpG islands are located upstream of gene promoters.

DNA methylation The addition of a methyl group to the 5' position of a cytosine.

DNA hydroxymethylation A hydroxymethyl group located at the 5' position on a cytosine. Considered a mark of active demethylation catalyzed by the enzyme ten-eleven-transferase.

Epigenetics The heritable and modifiable phenomenon of chemical marks that are added to the DNA and its supporting structure to modify gene expression without changes to the underlying DNA sequence.

Euchromatin Loosely packaged chromatin that is accessible to transcriptional regulators for gene transcription.

Gene activation Gene transcription.

Gene repression Inhibition of gene transcription.

Heterochromatin Tightly packaged and condensed chromatin that is inaccessible to transcription regulators and associated with gene repression and repetitive regions of the genome.

Histone A highly basic, globular protein composed of two each of H2A, H2B, H3, and H4 subunits. Serves as the protein architecture that DNA wraps around to form chromatin.

Histone modifications Posttranslational modifications added to the tails that extend from the globular body of the histone protein. These modifications include acetylation, methylation, phosphorylation, and ubiquitinylation.

microRNA An 18- to 22–base pair long nonencoding nucleotide that functions to repress expression by complementarily binding to microRNAs to cause degradation or inhibition of protein translation.

Nucleosome DNA wrapped around a histone protein.

Nucleosome remodeling Altering the position of histones along the DNA by moving or removing histones. Performed by large protein adenosine triphosphate–dependent complexes such as switch/sucrose nonfermentable.

Promoter A region upstream of the transcription start site of a gene that serves as a docking site for transcriptional machinery.

INTRODUCTION

In Apr., 2003, the International Human Genome Sequencing Consortium announced the completion of the Human Genome Project. All 3,000,000,000 base pairs (bps) of nucleotides had been sequenced, yielding about 30,000 genes in the human genome. At the time, the predominant opinion was that knowing the sequence of the human genome would reveal the origin of all diseases, thereby unveiling information required to understand human health fully. However, substantial phenotypic disparity and disease susceptibility remained that was not explained by DNA sequence alone. Continued research in molecular biology began to focus on the field of epigenetics as a mechanism of interaction between the environment and genome to gain further insight into disease. Epigenetics is an additional layer of genomic modulation that drives the transition from genotype to phenotype by dynamically regulating where and when genes are expressed. At the time of publication, a search for the term “epigenetics” on PubMed yielded over 20,000 articles, highlighting interest in this topic within the field of health sciences.

The term “epigenetics” means on top of or above genetics. Epigenesis was coined by the Greek philosopher Aristotle to describe his theory of the differentiation of an organism through development, in opposition to the prevailing theories of preformation. The notion of epigenetics was further developed by the inaugural systems biologist Conrad Hal Waddington in the 1930s during his studies on the embryological development of *Drosophila melanogaster*. Waddington created the metaphor of marbles rolling down grooved hills to describe an epigenetic landscape; each marble represented a cell that developed toward a particular fate based on the external influences that drove it down particular grooves, thus determining its final identity (Fig. 17.1).

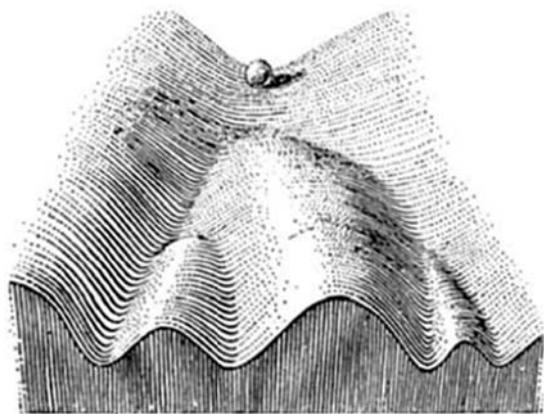


FIGURE 17.1 Waddington's epigenetic landscape. The marble at the top of the hill represents a pluripotent cell with an undifferentiated phenotype. As the marble rolls down the hill, it will follow different valleys to fall to a final resting spot, representing the differentiated cell. Once the marble has followed a particular grooved path, it cannot cross into another neighboring valley. From Smith, C.L., Bolten, A., Nguyen, G., 2010, *Genomic and epigenomic instability, fragile sites, schizophrenia and autism. Curr Genom* 11(6), 447–469.

These philosophical notions of epigenetics were supported with the emergence of molecular biology. Epigenetics is currently defined as a heritable and reversible phenomenon that alters gene expression without changes to the underlying bp sequence. Every somatic cell in an organism has an identical genome, although each cell type is only expressing a subset of the genes in the genome at one time. This specific pattern of gene expression determines the cell identity and dictates how the cell responds and functions within the tissue microenvironment. Epigenetics is the mechanism that determines which genes will be expressed and which will be repressed. Every cell type in an organism has a unique epigenetic pattern. As embryonic stem cells undergo differentiation from a pluripotent state, the epigenetic pattern will be continuously modified to yield distinct germ layers and ultimately differentiated cell types comprising tissues and organs.

Physically, epigenetics entails the addition of chemical tags directly on top of the DNA bps or onto the supporting protein structure without altering the underlying nucleotide sequence. The 3,000,000,000 bps that comprise the DNA double helix form over 1.8 m of a delicate string-like material that must be compacted to fit inside the approximately 2- to 10- μ m nucleus (Rothbart and Strahl, 2014). To fit, DNA is wrapped around globular bead-like proteins called histones to create a nucleosome beads-on-a-string structure; these beads-on-a-string are then further compacted into chromatin fibers and condensed into chromosomes. Active gene transcription requires local relaxation of the highly folded chromatin structure in the region of the gene to enable access for transcriptional machinery. Epigenetic

marks determine the local compaction of chromatin structure and can serve as docking sites for transcriptional regulating machinery to coordinate gene expression. Enzymes that add epigenetic marks onto the DNA and supporting histone structure are called epigenetic writers, whereas enzymes that remove these marks are called epigenetic erasers. Epigenetic readers are proteins that dock onto the epigenetic marks and can serve to alter chromatin compaction or recruit additional transcriptional regulators to modify gene expression. Together, the epigenetic writers, erasers and readers, and multiple levels of epigenetic marks orchestrate the cellular transcriptomic profile.

As a somatic cell undergoes the process of mitosis, the epigenetic patterns of the parental cell are copied onto the nascent DNA of the daughter cells. Accumulating evidence also supports that epigenetic patterns can be heritable from parent to offspring during development, although the exact mechanism remains to be elucidated. Epigenetic patterns are modifiable by the environment, which underlines how cells and organisms respond to changing environmental influences that range from nutrition to stress and toxin exposure. Ongoing research is investigating how various environmental inputs alter the function and targeting of the epigenetic writers, readers, and erasers.

Epigenetic regulation is the mechanism that defines the interface between nature and nurture. The remainder of this chapter will first review the state of the science regarding different levels of epigenetic regulation, including DNA methylation, histone modifications, chromosome remodeling, and small interfering RNAs. It then surveys how these levels of epigenetics drive cell physiology and inform the development of health and disease within the scope of systems biology (Fig. 17.2).

DNA METHYLATION

The best-studied epigenetic mechanism is DNA methylation, which entails the addition of a methyl group to the 5' position on the pyrimidine ring of the nucleotide cytosine (Fig. 17.2). Most commonly, this methylation occurs on a cytosine flanked by a guanine to yield 5'-cytosine-phosphate-guanine-3' (CpG) methylation. Methylation can also occur on cytosines neighboring other nucleotides for non-CpG methylation. The extent of DNA methylation varies between organisms and tissues. In humans, 5% of cytosines are methylated at the global level, with hypermethylation observed in areas of heterochromatin and hypomethylation observed in euchromatin. Across the genome, CpG sites localized to repetitive DNA, such as Alu repeats, are typically methylated, whereas high densities of

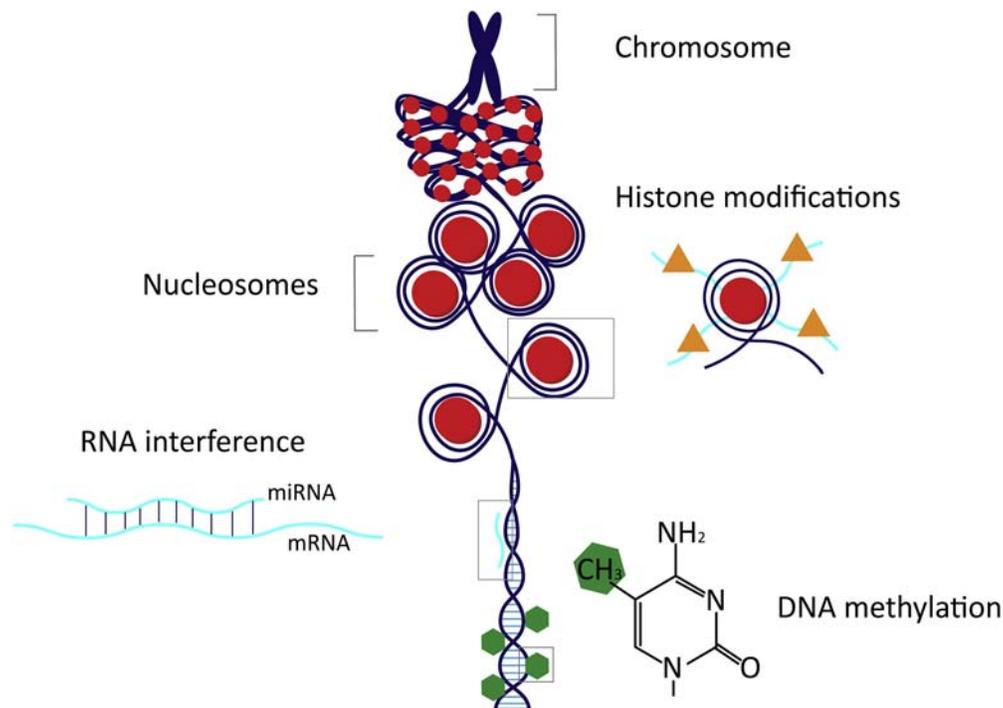


FIGURE 17.2 Epigenetic marks and DNA packaging. Multiple levels of epigenetics regulate the compaction of DNA in the nucleus. DNA serves as the blueprint for transcription of genes into messenger RNA (mRNA), which then serves as the template for protein translation. The DNA double helix is wrapped around histone proteins to create nucleosomes; these nucleosomes are further compacted into chromatin fibers and condensed into a chromosome. DNA methylation is the addition of a methyl group to a cytosine at a CpG site. RNA interference is the transcription of a non–gene coding microRNA (miRNA) that complementarily binds to an mRNA to inhibit protein translation. Histone modifications are the posttranslational addition of a chemical modification to a specific amino acid on a histone tail to alter how tightly the DNA is wound around the histone or to serve as a docking site for transcriptional regulatory machinery.

CpG sites localized more closely to genes are unmethylated.

The effect of DNA methylation on gene expression depends on the genomic context, specifically the localization of methyl-cytosines in relation to the gene and the density of CpGs. Although most DNA methylation occurs at CpG sites, the genome has a lower proportion of CpG sites than expected by chance. According to evolution, this results from the mutability of a methylated cytosine, because spontaneous deamination can yield a thymine and cause a mutation. Whereas most of the genome is CpG poor, there are segments of 300–3000 bp with a 65% observed-to-expected ratio of CpG sites that are called CpG islands (CGIs) (Fatemi et al., 2005). CGIs are localized to the promoter region of 70% of human genes, including nearly all house-keeping genes, tissue-specific genes, and developmentally regulated genes (Deaton and Bird, 2011). Approximately 50% of CGIs in the human genome are associated with transcription start sites of genes, whereas the remaining 50% are considered orphan CGIs that are intragenic or intergenic (Deaton and Bird, 2011). Nearly all promoter-associated CGIs are unmethylated in healthy somatic cells; only about 3%

of promoter-associated CGIs are methylated (Deaton and Bird, 2011). Special cases of promoter-associated DNA methylation in a subset of developmental and tissue specific-genes become methylated during embryonic development and subsequently silence gene expression (Deaton and Bird, 2011).

Tissue differences in DNA methylation are also localized outside promoter-associated CGIs, specifically within 2 kb of CGIs that are denoted CGI shores (Irizarry et al., 2009). There is an inverse relation between CGI shore methylation and gene expression; hypermethylation of the CGI shore is associated with decreased gene expression and hypomethylation is associated with increased gene expression (Irizarry et al., 2009). This supports the notion that DNA methylation differences have a role in differentiating gene expression profiles between tissues.

By comparison, approximately 20%–34% of intragenic CGIs are methylated (Deaton and Bird, 2011). The underlying function of intragenic DNA methylation is not well-understood, although it is hypothesized that this methylation may regulate alternative splicing, the rate of transcriptional elongation, or expression of non-coding RNAs.

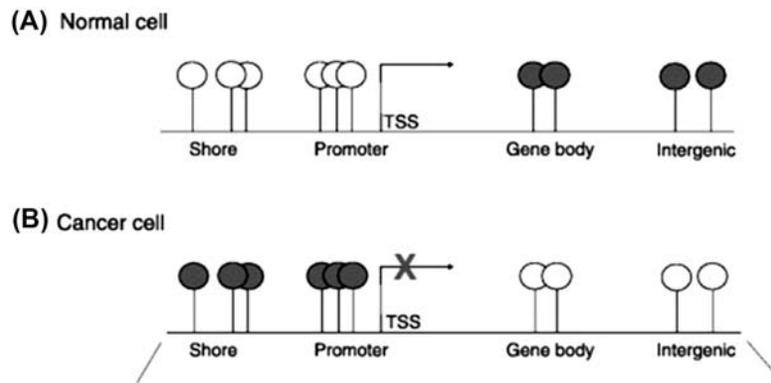


FIGURE 17.3 DNA methylation in healthy and cancer cells. (A) Representative gene in a healthy cell with open circles indicating unmethylated CpGs in the CpG island shore, promoter and transcription start site (TSS) with associated gene transcription and filled circles indicating methylated CpG sites localized in the gene body and intergenic regions. (B) Representative cancer cell with hypermethylation of the CpG island shore and promoter with inhibition of gene transcription and hypomethylation of the gene body and Intergenic regions. *From Witte, T., Plass, C., Gerhauser, C., 2014. Pan-cancer patterns of DNA methylation. Genome Med 6(8), 66–84.*

Investigation into the cancer state has demonstrated an increased frequency of promoter-associated CGI methylation and hypomethylation of CpGs in repetitive DNA elements. Many types of sporadic cancers have been associated with aberrant hypermethylation of CGIs localized to the promoters of tumor suppressor genes. This promoter hypermethylation is associated with suppression of gene expression, thereby enabling cellular hyperproliferation. Compared with healthy somatic cells that show depleted methylation of promoters, cancerous cells show an equivalent level of CGI methylation across promoter-associated and orphan CGIs (Deaton and Bird, 2011). Both hypomethylation and hypermethylation have been observed in CGI shores that were associated with increased expression of oncogenes and reduced expression of tumor suppressors, respectively, in colon cancer tissue compared with patient-matched healthy colonic mucosa (Irizarry et al., 2009) (Fig. 17.3).

Gene-silencing effects of DNA methylation are driven by the inhibition of transcription factor binding to

methylated CpGs and by the recruitment of proteins containing a methyl-binding domain (MBD). Proteins containing an MBD are one type of epigenetic reader. The first protein identified that contained an MBD was methyl-CpG binding protein 2 (MeCP2) which specifically recognizes paired methylated CpGs on both strands of DNA. MeCP2 is a component of the large co-repressor complex NCoR/SMRT, which also contains writers of repressive histone modifications, evidencing cross-talk between multiple levels of epigenetic regulation.

In mammals, the addition or writing of methyl groups to DNA is catalyzed by two families of DNA methyltransferase (DNMT) enzymes; DNMT1 and DNMT3, DNMT1 is involved in maintenance DNA methylation and is responsible for copying the methylation pattern from the parental strand to the nascent daughter strand during DNA replication (Fig. 17.4). The DNMT3 family, which can be further subdivided into DNMT3a and DNMT3b, is responsible for de novo DNA methylation. These enzymes are particularly active during early embryonic development to catalyze imprinting, X-

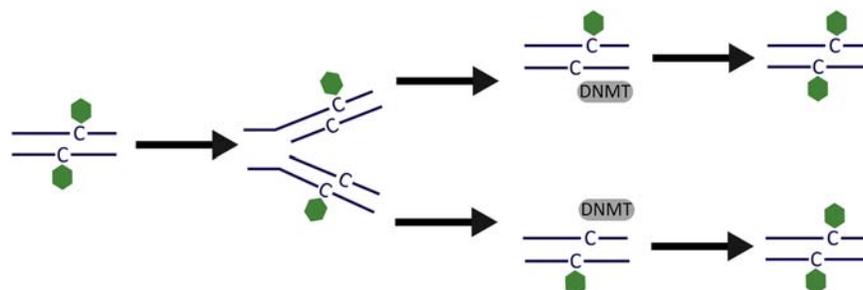


FIGURE 17.4 Heritability of DNA methylation across cell division. Before mitosis, DNA is replicated in a semiconservative nature in which each parental strand serves as the template for a new daughter strand. DNA methyl-transferase (DNMT)1 copies the DNA methylation pattern from the parental strand onto the daughter strand during replication.

chromosome inactivation and tissue-specific methylation pattern reactions, though DNMT3s continue to be expressed and active throughout the lifespan. Taken together, the DNMT enzymes demonstrate how DNA methylation is both heritable and modifiable.

All DNMTs use S-adenosylmethionine (SAM) as a source of methyl groups to yield a methylated cytosine and S-adenosylhomocysteine (SAH). SAM is the universal methyl donor for all biological methylation reactions, including DNA methylation. The availability and regeneration of SAM is regulated through a series of 1-carbon metabolism reactions that are dependent on several B vitamins as co-enzymes. The amino acid methionine is used to synthesize SAM. Following the donation of a methyl group, SAM is converted into SAH and then to homocysteine, an amino acid. Homocysteine can be converted into cysteine through vitamin B6 dependent reactions or can be recycled into methionine by receiving a methyl group from betaine or from 5-methyltetrahydrofolate in a vitamin B12 dependent reaction. In particular, the ratio of SAM to SAH is critical in regulating the rate of biological methylation reactions as SAH inhibits DNMTs.

DNA HYDROXYMETHYLATION

Evidence has indicated the presence of a hydroxymethyl group on the 5' position of cytosine in CpG

dinucleotides; this mark is considered to be a transient step in the active demethylation of DNA. The ten-eleven transferase (TET) enzyme catalyzes an oxidation reaction that converts 5-methylcytosine into 5-hydroxymethylcytosine before subsequent steps that ultimately yield an unmethylated cytosine (Fig. 17.5). Regions of hydroxymethylated DNA are considered to be sites of active demethylation and have been associated with an increase in gene expression (Tammen et al., 2015). The targeting of TET enzymes to specific regions of DNA remains to be elucidated.

HISTONE MODIFICATIONS

Histone modifications are posttranslational modifications (PTM) that are added covalently to the histone protein structure of the nucleosome. These PTMs can act to modify chromatin compaction to alter DNA accessibility to transcriptional regulators and serve as docking sites for epigenetic readers. Histones are octamers of two of each H3, H4, H2A and H2B subunits that bind together to form a highly basic, globular protein. The amino and carboxy terminal tails extend out from the globular portions of each subunit; most PTMs are added onto these histone tails (Rothbart and Strahl, 2014) (Fig. 17.6).

Each histone protein is wrapped by 147 bp of DNA to form a nucleosome. The spacing between each nucleosome varies between 30 and 100 bp and can be modified

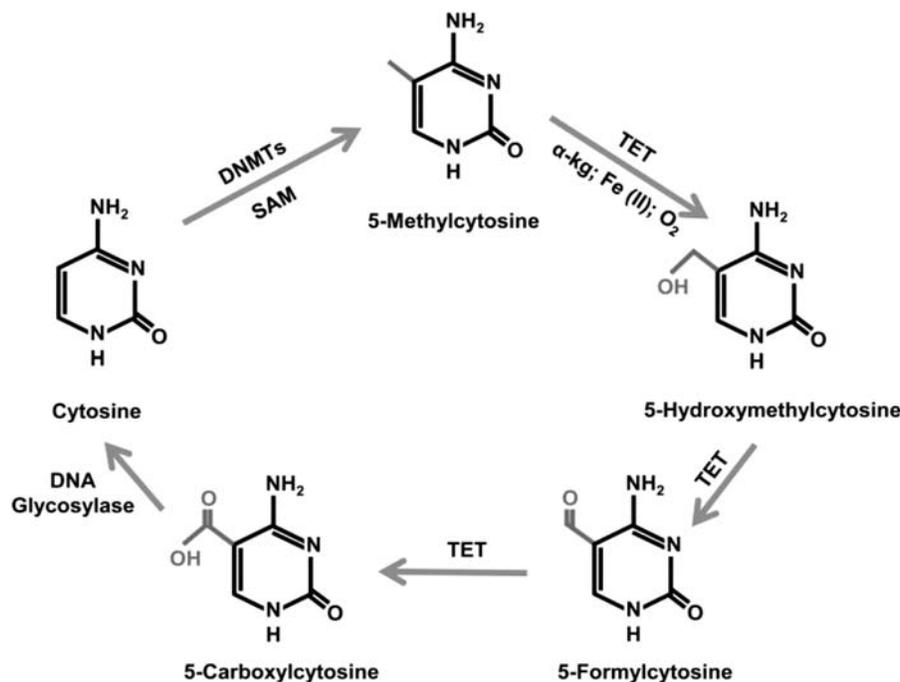


FIGURE 17.5 Proposed mechanism of active demethylation by ten-eleven transferase (TET). The enzyme TET acts on a methylated cytosine to catalyze a series of reactions to yield transitions to hydroxymethylcytosine, formylmethylcytosine, and carboxylcytosine. The enzyme DNA glycosylase catalyzes the final step to yield an unmethylated cytosine. DNMT, DNA methyl-transferase; SAM, S-adenosylmethionine. From Tammen, S.A., Friso, S., Choi, S.W., 2013. Epigenetics: the link between nature and nurture. *Mol Asp Med* 34(4), 753–764.

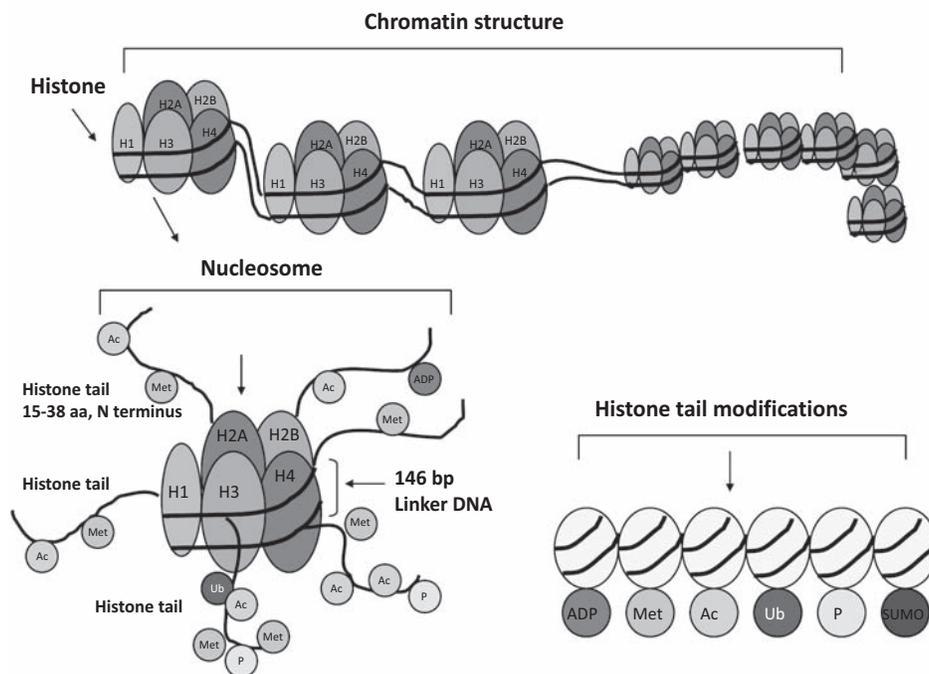


FIGURE 17.6 Map of histone marks. Map of posttranslational histone modifications at specific amino acids within the tails of the histone octamer subunits. From Pandian, G.N., Sugiyama, H., 2013. *Strategies to modulate heritable epigenetic defects in cellular machinery: lessons from nature. Pharmaceuticals* 6(1), 1–24.

by nucleosome repositioning complexes (discussed later). The interaction between nucleosome positioning, histone modifications and DNA methylation determines local chromatin packaging. Highly compacted chromatin is designated as heterochromatin, yielding DNA that is inaccessible to transcription factors and thus associated with gene repression. Gene transcription activation requires the euchromatin state in which DNA is loosely wrapped enabling accessibility to transcriptional regulators (Rothbart and Strahl, 2014).

The effect of histone PTMs on gene expression depends on which chemical modification has been added to specifically which amino acid on a particular histone subunit. The best studied chemical modifications are acetylation, phosphorylation and methylation; though ongoing research is continuing to characterize the effect of other modifications such as SUMOylation, butyrylation, propionylation, crotonylation and addition of O-linked β -N-acetylglucosamine sugar moieties (Rothbart and Strahl, 2014). Gene expression effects of the best-studied histone PTMs are summarized in Table 17.1 and discussed briefly next.

Generally, the addition of acetyl groups to histones is associated with a loosening of the chromatin structure and transcriptional activation (Rothbart and Strahl, 2014). Acetylation of lysine 9 or 27 on histone 3 (H3K9 and H3K27) and lysine 16 of histone 4 (H4K16) are all associated with open, accessible euchromatin and active

gene expression. Specifically, H3K27ac and H3K9ac are implicated in transcriptional initiation at active gene promoters (Rothbart and Strahl, 2014). Histone acetylation is catalyzed by a family of histone acetyl transferase (HAT) enzymes and removed by histone deacetylase (HDAC) enzymes. The acetylation status of histones is dynamically regulated by a balance of HAT and HDAC activity to transition between active gene expression and gene repression. The gene activation effects of histone acetylation are thought to be twofold (Rothbart and Strahl, 2014). First, histone proteins are positively charged and bind tightly to the negatively charged DNA. The addition of a negative acetyl group effectively neutralizes the charge on the histone tail and reduces the attraction to the DNA, weakening the histone–DNA interaction to enable accessibility. Second, many transcription factors and activators of gene expression contain bromodomains that are readers of the histone acetylation mark (Fig. 17.7).

Histone methylation can be activating or repressive, depending on the number of methyl groups added to which residue on the histone tails (Rothbart and Strahl, 2014). For example, methylation at lysine 4 on histone 3 (H3K4me) localizes to enhancers and CGIs and is associated with active or poised promoters. Monomethylation of lysine 9 on histone 3 (H3K9me1) is associated with accessible chromatin and gene activation whereas trimethylation (H3K9me3) at this same site is

TABLE 17.1 Epigenetic Marks and Their Effect on Gene Expression.

Epigenetic mark	Description	Effect on gene expression	Notes
DNA methylation	Addition of methyl group to 5' position on pyrimidine ring of cytosine	Decrease	Effect depends on genomic context. Promoter-associated CpG islands are rarely methylated whereas methylation at CpG Island shores is more variable and inversely associated with gene expression.
DNA hydroxymethylation	Placement of hydroxymethyl group on 5' position of pyrimidine ring of cytosine	Increase	Considered to be a transient mark in the active demethylation process.
H3K27ac	Acetylation of lysine 27 on histone 3	Increase	Associated with transcriptional initiation in euchromatin.
H3K9ac	Acetylation of lysine 9 on histone 3	Increase	Associated with transcriptional initiation in euchromatin.
H4K16ac	Acetylation of lysine 16 on histone 4	Increase	Localized to active genes and enhancer regions.
H3K27me3	Trimethylation of lysine 27 on histone 3	Decrease	Associated with gene repression of inactive developmental loci, placed by polycomb repressive complex.
H3K36me3	Trimethylation of lysine 36 on histone 3	Increase	Colocalizes with RNA polymerase II during elongation.
H3K4me1/me2/me3	Mono/di/trimethylation of lysine 4 on histone 3	Increase	During active transcription.
H3K9me3	Trimethylation of lysine 9 on histone 3	Decrease	Associated with repressive heterochromatin.
H3K9me1	Monomethylation of lysine 9 on histone 3	Increase	Associated with accessible chromatin during active transcription.
H4K20me1	Monomethylation of lysine 20 on histone 4	Increase	Associated with accessible chromatin during active transcription, catalyzed by PR-Set7.
H2AK119Ub1	Monoubiquitination of lysine 119 on histone 2A	Decrease	Catalyzed by Bmi1 subunit of polycomb repressive complex 1.

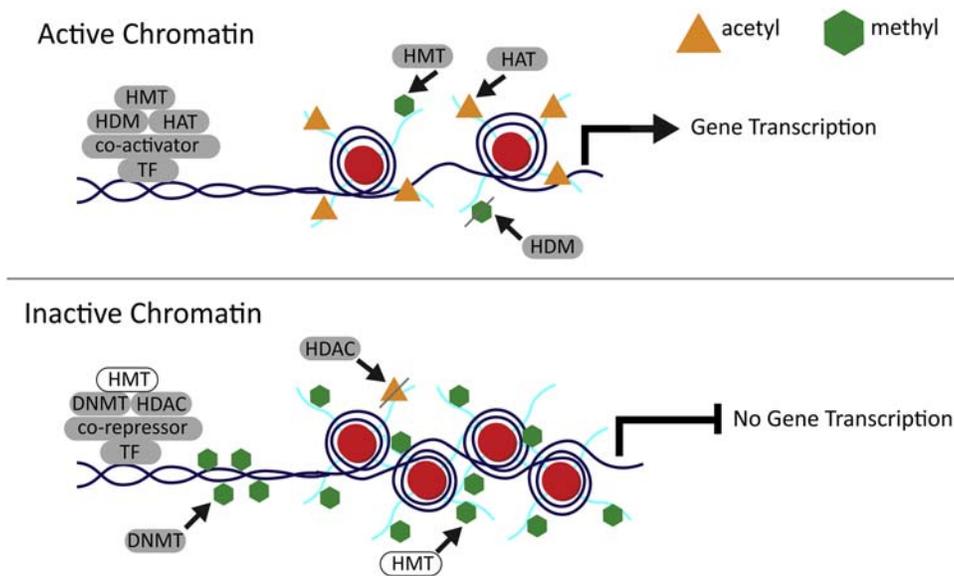


FIGURE 17.7 Dynamic balance of histone acetyl transferases (HATs) and histone deacetylase (HDACs) on gene transcription. Addition of acetyl groups to histones by HAT is associated with an accessible chromatin structure, transcription factors (TF) poised with coactivators including activating epigenetic writers such as HATs, specific histone methyl transferases (HMTs), and specific histone demethylases (HDMs) yielding active gene expression. The removal of acetyl groups from histones by HDAC is associated with a condensed chromatin structure, transcription factors poised with corepressors including HDACs, DNA methyl-transferases (DNMTs), and specific HMTs yielding repressed gene expression.

associated with a repressive heterochromatin state. Trimethylation at lysine 27 of histone 3 (H3K27me3) is associated with gene repression and is localized to inactive developmental loci. The EZH2 subunit of a large protein complex that regulates gene repression called the polycomb repressive complex (PRC) is responsible for the writing of the H3K27me3 mark (Wang et al., 2016).

Evidence indicates that it may be oversimplistic to interpret histone PTMs individually. Rather, it is the collaboration among many histone modifications, DNA methylation, and overall three-dimensional chromatin folding that ultimately informs the impact on gene transcription (Rothbart and Strahl, 2014). International consortiums such as the Encyclopedia of DNA Elements (ENCODE) are enabling genome-wide investigation into the cross-talk and colocalization of multiple levels of epigenetic modification and transcriptional profiles to gain a more complete picture of chromatin architecture.

CHROMATIN REMODELING

Chromatin remodeling entails large multisubunit protein complexes that alter the local chromatin architecture to regulate accessibility to transcriptional machinery. Subunits of these protein complexes can simultaneously catalyze multiple levels of epigenetic reactions and/or reposition nucleosomes to transition between euchromatin and heterochromatin states. Two well-studied

chromatin remodeling complexes include the gene-activating switch/sucrose nonfermentable (SWI/SNF) complex and the transcriptionally silencing polycomb group (PcG) proteins.

SWI/SNF functions in adenosine triphosphate (ATP)-dependent nucleosome repositioning. Originally discovered in *Saccharomyces cerevisiae*, its name was based on its functions in altering yeast mating type and metabolism (Kadoch and Crabtree, 2015). SWI/SNF contains a core ATPase/remodeling domain that harnesses the energy of ATP to destabilize the DNA–histone interaction to reposition the nucleosome effectively (Kadoch and Crabtree, 2015). Binding of transcription factors is inhibited at DNA sequences that are wrapped around histones; therefore, SWI/SNF nucleosome repositioning enables accessibility of DNA to transcriptional regulators. Additional subunits of the SWI/SNF complex include bromodomains that are specific to acetylated histones and subunits with specificity to transcriptional regulators and coregulators (Kadoch and Crabtree, 2015). Biochemical models indicate that transcriptional regulators bind sequence-specifically to a region of DNA and then subsequently recruit SWI/SNF to alter the local chromatin architecture to promote transcription.

In contrast, the PcG proteins are a large protein complex containing enzymatic subunits that catalyze the writing of repressive histone marks to silence gene expression. Originally discovered in *D. melanogaster* for their role in Hox gene silencing in anteroposterior morphogenesis, PcG proteins are implicated in epigenetic silencing during development (Wang et al., 2016). In

mammals, there are two well-characterized members of the PcG family including PRC1 and PRC2, respectively.

PRC1 complexes catalyze the addition of the transcriptionally silencing monoubiquitination mark on lysine 119 of H2A (H2AK119Ub1) by the Bmi1 subunit (Wang et al., 2016). PRC2 complexes include the EZH2 subunit, which is responsible for the trimethylation of H3K27 (Wang et al., 2016). PcG proteins are critical in working with DNA methylation for X inactivation and to maintain somatic cell identity through gene repression.

RNA INTERFERENCE

The continually developing field of epigenetics has expanded with the discovery of RNA interference in 1998 by Andrew Fire and Craig C. Mello while working with *Caenorhabditis elegans*. MicroRNAs (miRNA) are endogenously expressed short RNAs that do not encode for proteins. Initially transcribed as single-stranded pre-miRNAs, they fold over on themselves to form a stem loop structure that is then subsequently processed by RNA-induced silencing complexes to yield 18–22 nucleotide-long, single-stranded mature miRNA transcripts (Laganà et al., 2014). miRNAs silence gene expression by sequence-specifically binding to messenger RNAs (mRNAs) to promote degradation or inhibit translation (Laganà et al., 2014).

RNA interference can also be induced by exogenous small interfering RNAs (siRNAs). siRNAs can be introduced by viruses or inserted synthetically in the laboratory. Currently, siRNAs are used experimentally to modify gene expression and are investigated for possible

targeted therapeutic approaches in disease (Laganà et al., 2014).

EPIGENETICS IN HEALTH AND DISEASE

Multiple levels of epigenetic marks have critical roles in determining the cellular transcriptomic profile required to maintain homeostasis through the lifespan. Epigenetic research is currently focused on characterizing the unique epigenetic profiles across all tissue types in health and how alterations in these profiles are implicated in disease. International collaborations such as ENCODE and the Human Epigenome Project are creating multilevel databases of epigenetic profiles ranging from DNA methylation to histone modifications and chromatin folding in tissues, and cell types serve as a valuable resource for comparison.

The highly dynamic nature of epigenetic remodeling during development is well-established, as evidenced by several developmental diseases associated with impairment in the epigenetic readers, writers, and erasers. For example, Prader–Willi syndrome is associated with impaired genetic imprinting of genes located on chromosome 15 and Rett syndrome is caused by a mutation in the epigenetic reader MeCP2. Epigenetic profiles change over the course of the life span and are considered to have both a physiological and pathophysiological role. Aging is associated with genomic DNA hypomethylation and gene-specific hypermethylation, although it is currently unknown whether these changes are programmatic in the physiological aging process or whether they are stochastic (Fig. 17.8).

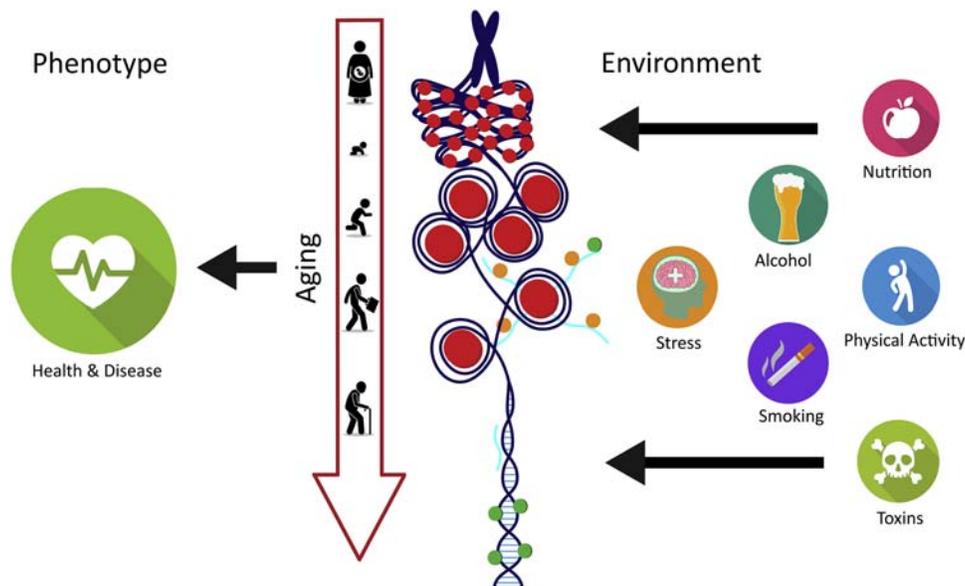


FIGURE 17.8 Environmental influences on epigenetics. Factors that can alter epigenetic patterns through development and aging and in response to environmental influences. Ongoing research is characterizing how these epigenetic changes can alter phenotype.

Much research has focused on investigating how the environment can modify epigenetic profiles and in turn, how these epigenetic profiles can alter susceptibility to disease. Most notably, early development of cancer is associated with epigenetic modifications yielding the expression of oncogenes, silencing of tumor suppressor genes, and a loss of cellular identity. Investigations are ongoing into the etiological role of epigenetic changes in cardiovascular disease, diabetes, and obesity as well as in chronic neurological and psychiatric disorders including autism.

One compelling example of both the heritable and modifiable nature of epigenetics is the influence of parental environmental exposures on offspring epigenetic patterns. Seminal studies were performed using the Agouti mouse model, a strain characterized by a yellow coat, obese phenotype, and increased risk for diabetes and cancer with DNA hypomethylation of the agouti gene.

When the maternal Agouti mouse diet was supplemented with the carbon-metabolism nutrients vitamins B6, B12, and folate, offspring were born with brown coats, a lean phenotype, reduced disease risk, and increased DNA methylation of the agouti gene locus. Historical studies in humans investigated individuals who were in utero during the Dutch Hunger Winter, when Holland experienced a severe famine during the 1944–45 Nazi occupation in World War II, compared with their non-famine exposed siblings. As adults, prenatally famine-exposed individuals demonstrated an increased risk for cardiovascular disease and type 2 diabetes with decreased DNA methylation at the imprinted insulin-like growth factor 2 locus compared with their non-famine exposed siblings. In both of these examples, parental environmental exposures in utero drove changes in epigenetic patterns and associated phenotypic alterations that were persistent through the lifetime of offspring. Numerous studies have characterized the timing and penetrance of different parental environmental exposures on offspring epigenetic patterns to inform best practices for health policy.

In conjunction with epigenomic characterization of pathologies is research into epigenetic targeting drugs. Currently, Clinical [Trials.gov](https://www.clinicaltrials.gov) indicates that HDAC inhibitors are being investigated as a therapeutic option in many types of cancer including leukemia, breast cancer, myeloma, renal cancer, hepatocellular carcinoma, and melanoma. The DNA methyltransferase inhibitor azacytidine has also been used to treat myelodysplastic syndromes and myeloid leukemias. Ongoing research is

attempting to elucidate how to target these drugs specifically to cancerous cells to avoid altering epigenetics in healthy tissues.

CONCLUSIONS

Epigenetics serves as the interface between nature and nurture by responding to changing environmental conditions to bring phenotypic plasticity to the static genotype. Multiple levels of epigenetic regulation, including DNA methylation and hydroxymethylation, posttranslational histone modifications, nucleosome repositioning, and RNA interference, serve to modify gene expression with no changes to the underlying DNA sequence. Epigenetic mechanisms regulate where and when genes are expressed to yield distinct tissue-specific transcriptomic profiles despite identical genomes across all cells.

Early research into epigenetics primarily focused on one level of epigenetic regulation, frequently at only one locus at a time. In the advent of large-scale sequencing tools, investigations have transitioned to characterizing the genome-wide profile of multiple levels of epigenetic patterns across different cell types. These bioinformatic approaches enable exploration of cross-talk between levels of epigenetic regulation to gain a comprehensive view of genome architecture and control of gene expression. Future work will build on this foundation to elucidate how epigenetic patterns change in disease, how therapeutics can target epigenetic modification, and how environmental influences may have a role in preventative health.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at <https://doi.org/10.1016/B978-0-12-804572-5.00017-3> and a companion video at WEBSITE.

References

- Deaton, A.M., Bird, A., 2011. CpG islands and the regulation of transcription. *Genes Dev* 25, 1010–1022.
- Fatemi, M., et al., 2005. Footprinting of mammalian promoters: use of a CpG DNA methyltransferase revealing nucleosome positions at a single molecule level. *Nucleic Acids Res* 33 (20), e.176.

- Irizarry, R.A., et al., 2009. Genome-wide methylation analysis of colon cancer reveals similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 41 (2), 178–186.
- Kadoch, C., Crabtree, G.R., 2015. Mammalian SWI/SNF chromatin remodelling complexes and cancer: mechanistic insights gained from human genomics. *Sci Adv* 1 (5), e1500447.
- Laganà, A., Shasha, D., Croce, C.M., 2014. Synthetic RNAs for gene regulation: design principles and computational tools. *Front Bioeng Biotechnol* 2, 65.
- Pandian, G.N., Sugiyama, H., 2013. Strategies to modulate heritable epigenetic defects in cellular machinery: lessons from nature. *Pharmaceuticals* 6 (1), 1–24.
- Rothbart, S.B., Strahl, B.D., 2014. Interpreting the language of histone and DNA modification. *Biochim Biophys Acta* 1839 (8), 627–643.
- Smith, C.L., Bolten, A., Nguyen, G., 2010. Genomic and epigenomic instability, fragile sites, schizophrenia and autism. *Curr Genom* 11 (6), 447–469.
- Tammen, S.A., et al., 2015. Hepatic DNA hydroxymethylation is site-specifically altered by chronic alcohol consumption and aging. *Eur J Nutr* 1–10.
- Wang, W., et al., 2016. Polycomb Group (PcG) proteins and human cancers: multifaceted functions and therapeutic implications. *Med Res Rev* 35 (6), 1220–1267.
- Witte, T., Plass, C., Gerhauser, C., 2014. Pan-cancer patterns of DNA methylation. *Genome Med* 6 (8), 66–84.

This page intentionally left blank

Genetics of Chrononutrition

Marta Garaulet, Jesus Lopez-Minguez, Purificación Gómez-Abellán

Department of Physiology, Faculty of Biology, University of Murcia, IMIB-Arrixaca, Murcia, Spain

List of Abbreviations

BMAL1 Brain- and muscle aryl hydrocarbon receptor nuclear translocator-like protein-1
PER2 Period 2
CLOCK Circadian Locomotor Output Cycles Kaput
REVERB (or Nr1d2) Nuclear receptor subfamily 1, group D, member 2
CRY Cryptochrome
MTNR1B Melatonin receptor 1B.

Glossary

- Acrophase** The time when the peak of a rhythm occurs. Originally, acrophase referred to the phase angle of the peak of a cosine wave fitted to the raw data of a rhythm.
- Adiponectin** Protein secreted from adipose tissue that modulates a number of metabolic processes such as glucose regulation and fatty acid oxidation.
- Association** When one or more genotypes within a population occur with a phenotypic trait. Genetic association can be between phenotypes, between a phenotype and a genetic polymorphism, or between two genetic polymorphisms.
- BMAL-1** Brain- and muscle aryl hydrocarbon receptor nuclear translocator-like protein-1, a transcription factor that represents a positive element of circadian autoregulatory feedback loop.
- Chronotype** Behavioral manifestation of underlying circadian rhythms of myriad physical processes. A person's chronotype is the propensity for the individual to sleep at a particular time during a 24-hour period. 'Eveningness or evening type' (delayed sleep period) and 'orningness or morning type' (advanced sleep period) are the two extremes; most individuals have some flexibility in the timing of their sleep period.
- Circadian function index** A numerical index that determines the circadian robustness and is based on three circadian parameters: regularity, fragmentation, and amplitude of the rhythm. Fragmentation values were inverted and normalized between 0 and 1, where 0 is a noise signal, and 1 a perfect sinusoid. Finally, circadian function index (CFI) is calculated as the average of these three parameters performed. Consequently, CFI oscillates between 0 (absence of circadian rhythmicity) and 1 (a robust circadian rhythm).
- Circadian rhythmicity** Any biological process that displays an endogenous, entrainable oscillation of about 24 hours. These 24-hour rhythms are driven by a circadian clock and have been widely observed in plants, animals, fungi, and cyanobacteria. The term "circadian" comes from the Latin "circa," meaning "around" (or "approximately") and "diem," meaning "day."
- Circadian system** The circadian system consists of a set of structures involved in the generation of 24-hour rhythms or circadian rhythms in behavioral, physiological, and biochemical variables, as well as in the external and internal synchronization of these variables to environmental cues and to each other, respectively.
- Chronodisruption** Serious disruption of the internal temporal order of the biochemical, physiological, and behavioral circadian rhythms.
- CLOCK** Circadian Locomotor Output Cycles Kaput is a transcription factor that represents one of the positive elements with *BMAL1*, of the circadian autoregulatory feedback loop.
- Cortisol:** A steroid hormone in the glucocorticoid class of hormones produced in humans by the fasciculate zone of the adrenal cortex within the adrenal gland. It is released in response to stress and low blood glucose concentrations.
- CRY** Cryptochrome is a transcription factor that represents a negative element of the circadian autoregulatory feedback loop.
- Epigenetics** Field focused on the study of genetic expression regulators that switch genes on or switch off, altering a person's phenotype.
- HOMA-IR** The homeostatic model assessment (HOMA) is a method used to quantify insulin resistance (IR) and β -cell function. It was first described under the name HOMA by Matthews et al. in 1985. The approximating equation for insulin resistance in the early model used a fasting plasma sample and was derived using the insulin-glucose product, divided by a constant.
- Homozygotes** A cell is said to be homozygous for a particular gene when identical alleles of the gene are present on both homologous chromosomes. An individual who is homozygous dominant for a particular trait carries two copies of the allele that codes for the dominant trait, whereas an individual who is homozygous recessive for a particular trait carries two copies of the allele that codes for the recessive trait.
- Interaction** An action that occurs when two or more objects have an effect upon one another. In our case, this happens when single nucleotide polymorphisms and behaviors interact for a physiological trait.
- Leptin** The "satiety hormone" is made by adipose cells that regulate energy balance by inhibiting hunger. Leptin is opposed by the actions of the hormone ghrelin, the "hunger hormone". Both hormones act on receptors in the arcuate nucleus of the hypothalamus to regulate appetite to achieve energy homeostasis.
- Melatonin** A hormone that anticipates the daily onset of darkness. It is involved in the entrainment (synchronization) of the circadian rhythms of physiological functions including sleep timing, blood pressure regulation, seasonal reproduction, and many others.

Many of melatonin's biological effects in animals are produced through activation of melatonin receptors; others result from its role as a pervasive and powerful antioxidant, with a particular role in protecting nuclear and mitochondrial DNA.

Metabolic syndrome A clustering of at least three of five of the following medical conditions: abdominal (central) obesity, elevated blood pressure, elevated fasting plasma glucose, high serum triglycerides, and low high-density lipoprotein levels. Metabolic syndrome is associated with the risk for developing cardiovascular disease and diabetes.

Monozygotic twin Twins that originate from a single zygote split and share 100% of genes.

MTNR1B Melatonin receptor 1B, a protein that in humans is encoded by the *MTNR1B* gene. Several studies have identified *MTNR1B* receptor mutations that are associated with increased average blood sugar level and around a 20% elevated risk for developing type 2 diabetes. *MTNR1B* messenger RNA is expressed in human islets; immunocytochemistry confirms that it is primarily localized in β cells in islets.

Nutrigenetics A branch of nutritional genomics that aims to identify genetic susceptibility to diseases and genetic variation in the effects of nutrient intake on the genome. Nutrigenetics is not to be confused with nutrigenomics, which focuses on the role specific foods have in activating genes that affect susceptibility to certain illnesses such as Alzheimer disease and cancer. Nutrigenetics is still in its relative infancy compared with other branches of medical science. It aims to offer people personalized disease prevention advice based on genetic makeup.

PER 2 Period 2 is a transcription factor that represents a negative element of the circadian autoregulatory feedback loop.

Peripheral clocks Clocks located in the peripheral tissue such as liver, kidney, heart, or adipose tissue that are regulated or synchronized by the central clock in the suprachiasmatic nucleus.

PPARG Peroxisome proliferator-activated receptor γ , which regulates fatty acid storage and glucose metabolism. *PPARG* has been implicated in the pathology of numerous diseases including obesity, diabetes, atherosclerosis, and cancer.

Rev-erb β (or Nr1d2) Nuclear receptor subfamily 1 group D, member 2, a circadian regulated gene that messenger RNA displays rhythmic expression *in vivo* and in serum-synchronized cell cultures. However, it is currently unknown to what extent *Rev-erb β* contributes to oscillations of the core circadian clock. However, it has been shown to suppress hepatic gluconeogenic gene expression and glucose output through the related *Rev-erb α* receptor, which mediates gene repression.

SIRT1 Sirtuin 1 or nicotinamide adenine dinucleotide-dependent deacetylase sirtuin-1, is a protein that in humans is encoded by the *SIRT1* gene. Studies suggest that human sirtuins may function as intracellular regulatory proteins with mono-adenosine diphosphate-ribosyl transferase activity. *SIRT1* is downregulated in cells that have high insulin resistance; inducing its expression increases insulin sensitivity, which suggests that the molecule is associated with improved insulin sensitivity.

SNP Single nucleotide polymorphism, a genetic polymorphism that occurs in the same population of two or more alleles at one locus, each with appreciable frequency, in which the minimum frequency is typically taken as 1%.

Suprachiasmatic nucleus A tiny region of the brain in the hypothalamus situated directly above the optic chiasm. It is responsible for controlling circadian rhythms. The neuronal and hormonal activities it generates regulate many different body functions in a 24-hour cycle.

Zeitgebers Any external or environmental cue that entrains or synchronizes an organism's biological rhythms to the earth's 24-hour light-dark cycle. The most important Zeitgebers are light,

temperature, social interactions, pharmacological manipulation, exercise, and eating/drinking patterns.

INTRODUCTION

Our physiology changes during the day. It is well-known that several hormones that are related to obesity, such as cortisol, leptin, and adiponectin, among others, display circadian rhythmicity. Alteration of this normal pattern is called chronodisruption (CD) and is related to several disturbances in many systems and organs of our body.

CD can be defined as a serious disruption of the internal temporal order of the biochemical, physiological, and behavioral circadian rhythms. In our modern society, CD may be produced by external situations that are relatively common, such as jet lag, shift work, night light pollution, and overnight recreational activities (social jet lag). Other factors are internal and may produce CD by altering the core machinery of the molecular circadian clock. Brain- and muscle aryl hydrocarbon receptor nuclear translocator-like protein-1 (*BMAL1*), Period 2, and Circadian Locomotor Output Cycles Kaput (*CLOCK*), among others clock proteins, have a specific role in our physiology as well as in the circadian molecular clock. Indeed, several studies performed in mutant animals demonstrated that mutations in clock genes are related to obesity, aging, and other metabolic alterations implicated in several chronic diseases. For example, studies performed in experimental animals showed that animals that have mutations in *Clock* are more obese and display metabolic disturbances.

In humans, mutations are rare. Nevertheless, genetic variations in one single nucleotide (single nucleotide polymorphism [SNP]) are rarer common and may determine differences in vulnerability to disease. Regarding the circadian machinery, it is well-known that several genetic variants are related to obesity, obesogenic behaviors and also the effectiveness of treatments in weight loss programs.

Clock genes are not only associated with obesity, they may interact with several obesogenic behaviors for obesity or weight control parameters. One clear example is the interaction between *CLOCK 3111T/C* and emotional eating behaviors (Lopez-Guimera et al., 2014). Our results demonstrate that during a dietary treatment to lose weight, subjects who were very emotional eaters and carriers of the risk allele C in *CLOCK 3111T/C* lost less weight compared with (1) C carriers who were nonemotional eaters and (2) TT carriers, independent of their emotional status. Many other examples of interactions between clock genes and several behaviors will be explained subsequently.

Thus, through nutrigenetics we know that our behaviors may interact with our genes and thus decrease the deleterious effect of one specific risk variant. In other words, we can change our obesity predisposition: although we cannot change our genetic code, we can definitely change our behaviors.

From epigenetics, the message is even more positive: by changing our behaviors, we can even change our genome. For example, we demonstrated that DNA methylation levels at different CpG sites of *CLOCK* are higher in obese than in nonobese women, and these methylation levels were associated with several obesogenic behaviors such as snacking frequently, eating when bored, and eating from large packages. Therefore, we presume from these data that through some small but stable changes in eating behaviors, we can change DNA superstructure and consequentially gene expression, thus changing our destiny.

CHRONODISRUPTION AND FAILURES IN THE CENTRAL CLOCK

It was previously described that CD can be result from alterations in circadian system at different levels. Impairments can be present in *inputs* or *outputs* to the circadian central clock located in the suprachiasmatic nucleus, but also in the *central clock*. With regard to the *inputs*, failures appear to be result from different circumstances, such as (1) an absence of changes between day and night (with light–dark cycles, timing of food intake [eating or fasting], or programmed exercise [activity or rest] acting as synchronizing agents); (2) different periods or unusual phasing of synchronizing agents, such as light at night, nocturnal feeding, or physical activity; and (3) shifts in the time provided by zeitgebers (i.e., jet lag, shift work). *Outputs* of the central pacemaker may also be altered: suppression of melatonin at night or loss of rhythm of glucocorticoids such as cortisol. Other problems in the *central clock* can result, such as from desynchronization between the central pacemaker and peripheral oscillators, or alterations in so-called clock genes.

Failures in the Central Clock: Mutations in Experimental Animals

Many factors may affect the circadian system at different levels. In relation to clock genes, numerous studies performed in experimental animals with mutations in clock genes have proved that there is a relation between these mutations and further failures in circadian system and illness. In this sense, animals with mutations in clock genes have a higher risk of developing certain diseases, such as cardiovascular disease, cancer, and obesity.

One of the first studies showing the effect of genetic mutations on chronic illnesses was conducted in 2005 by Turek et al. This study proved that homozygous *Clock* mutant mice were hyperphagic and obese, displayed adipocyte hypertrophy, and developed the metabolic syndrome (MetS). However, there is controversy regarding such findings: results have not proved to be consistent across different experiments. For example, Kennaway et al. showed that *Clock* (Delta19) mutation in mice did not cause obesity; by contrast, the authors found reduced plasma free fatty acids and increased insulin sensitivity, together with increased plasma adiponectin, a protective antiinflammatory cytokine.

Further studies showed that mutations in different clock genes are related to a higher risk for developing certain diseases, such as premature aging, as demonstrated by Kondratov et al. In that study, *Bmal1* ($^{-/-}$) mice had a reduced life span (mean life span of knockout animals was 37 weeks versus 120 weeks for wild-type animals) and displayed various symptoms of premature aging, including sarcopenia, changes in the percentage of lymphocytes, and impaired vision, among others.

Genetic Variations of Clock Genes in Humans

Clock Genes Associate With Obesity in Humans

SNPs at some specific position of the genome are present in more than 1% of the general population. These genetic variations underlie differences in our vulnerability to several diseases such as obesity.

In this sense, work performed by Sookoian et al. in 2008 was the first to demonstrate that several variants at *CLOCK* were associated with obesity, especially with abdominal obesity. In addition, in that year, Scott et al. confirmed these results by showing that *CLOCK* could have a relevant role in the development of MetS, type 2 diabetes, and cardiovascular disease.

Our research group replicated the results of Sookoian and Scott and demonstrated that several genetic variants in clock genes were related to obesity and related diseases such as MetS. For example, *CLOCK* SNPs (rs3749474, rs4580704, and rs1801260 [3111T > C]) were associated with body mass index [BMI], energy intake, and diverse variables related to obesity (Garaulet et al., 2009). In fact, our results on the whole showed that carriers of the minor alleles ate more, ate more fat, and were more obese. On the other hand, minor allele carriers (A) of *CLOCK* rs4580704 had a lower risk for developing diabetes (31% lower) and hypertension (46% lower) than did noncarriers (Garaulet et al., 2009). In addition, one genetic variant in *CLOCK*, rs1801260 (3111T > C), was associated with increased obesity and less weight loss (Garaulet et al., 2010a). Moreover, the proportion of

subjects who slept less than 6 h per day was higher among subjects with the C minor variant of this polymorphism than among noncarriers (59% versus 41%). Some of these associations are functionally explained by the presence of a polymorphism involving a change in the structure of the messenger ribonucleic acid (RNAm), in turn leading to a change in gene expression, as is the case of the *CLOCK* rs3749474 polymorphism (Garaulet et al., 2010b). Most of these novel results were replicated in two different populations (Mediterranean and North American) with different genetic backgrounds and dietary habits, which supports the robustness of the findings.

Clock genes interact with several behaviors for obesity in humans

Clock genes may also *interact* with behaviors for obesity. This is the case of behaviors directly related to emotional eating. For example, seeking refuge in food (especially high-calorie foods) is a common strategy to reduce anxiety, sadness, and negative emotions that occur when following a long-term diet or owing to difficult circumstances of daily life. Our results demonstrated that participants of our weight loss program who were carriers of the risk allele C at *CLOCK* 3111T > C and displayed emotional eating behaviors more frequently had more difficulties in losing weight during treatment. Interestingly, C carriers who did not display emotional behaviors despite being potentially at risk owing to their genetic background had weight loss similar to that of TT carriers (protective allele). These results are encouraging because they show that by changing how we eat, we can reduce or even eliminate the deleterious effect of a genetic variant.

Of a total population of 1500 subjects, we found that 60% were emotional eaters, 40% were minor C allele carriers, and 30% were both C carriers and emotional eaters. Considering these results, it may be useful to develop behavioral and cognitive programs aimed at reducing the frequency of emotional eating episodes, particularly for the 30% who carry the risk variant C. These results are encouraging and should be used in clinical practice.

CLOCK GENES ARE RELATED TO WEIGHT LOSS

Genetics and Weight Loss

A main problem in weight loss treatments is dramatic interindividual variability in response to treatment. It is

believed that elucidation of the genetic component will predict weight loss effectiveness. Some studies carried out in monozygotic twins analyzed genetic factors for the effectiveness of weight loss programs, whereas other studies emphasized familial aggregation in their ability to lose weight and the role that parental obesity has in this matter.

Most older genetic studies focused on candidate obesity-related genes (NUGENOB) and found little association between genetics and weight loss, except for some cases (peroxisome proliferator-activated receptor γ). Notably, no association with weight loss was observed for 13 obesity risk SNPs from genome-wide associated studies, which suggests that the genetic architecture of obesity and weight loss effectiveness may differ or that gene–environment effects need to be considered.

Surprisingly, clock genes have a relation to weight loss. In this sense, our most relevant results show an association between *CLOCK* 3111T > C (rs1801260) and weight loss effectiveness (Garaulet et al., 2010a) (Fig. 18.1). This study demonstrated that C carriers were more resistant to weight lost than TT homozygotes when individuals were subjected to an energy-restricted diet (Garaulet et al., 2010a). These data suggest that clock SNPs may predict weight loss in response to a low-energy diet.

Main Barriers to Lose Weight (Environmental Factors)

It is classically known that weight loss effectiveness is highly related to *environmental factors* (Corbalan et al., 2009). In a study performed in 1400 obese subjects from a Mediterranean area in southeast Spain who attended a weight-reduction program, subjects who lost less weight during treatment had more obstacles and a higher barriers to lose weight score (double) compared with those who achieved their goal. One of the main barriers was a lack of motivation. In fact, patients who displayed a higher predisposition to lose motivation (83% more) did not achieve their weight goal. Other barriers such as stress-related diet, thinking in black-or-white terms, and eating while bored were also related to failure to lose weight (Corbalan et al., 2009).

Genetics Associated With Obesogenic Behaviors to Modulate Total Weight Loss

We have also discovered how certain polymorphisms are associated with different environmental factors to modulate total weight loss (Lopez-Guimera et al., 2014). This is the case of *CLOCK* 3111T > C (rs1801260) SNP. Our group showed that minor allele C carriers of *CLOCK* 3111T > C were more of the evening type, had less healthy dietary habits, and tended to sleep less than did T carriers; they also lost less weight during treatment than TT carriers (Lopez-Guimera et al., 2014).

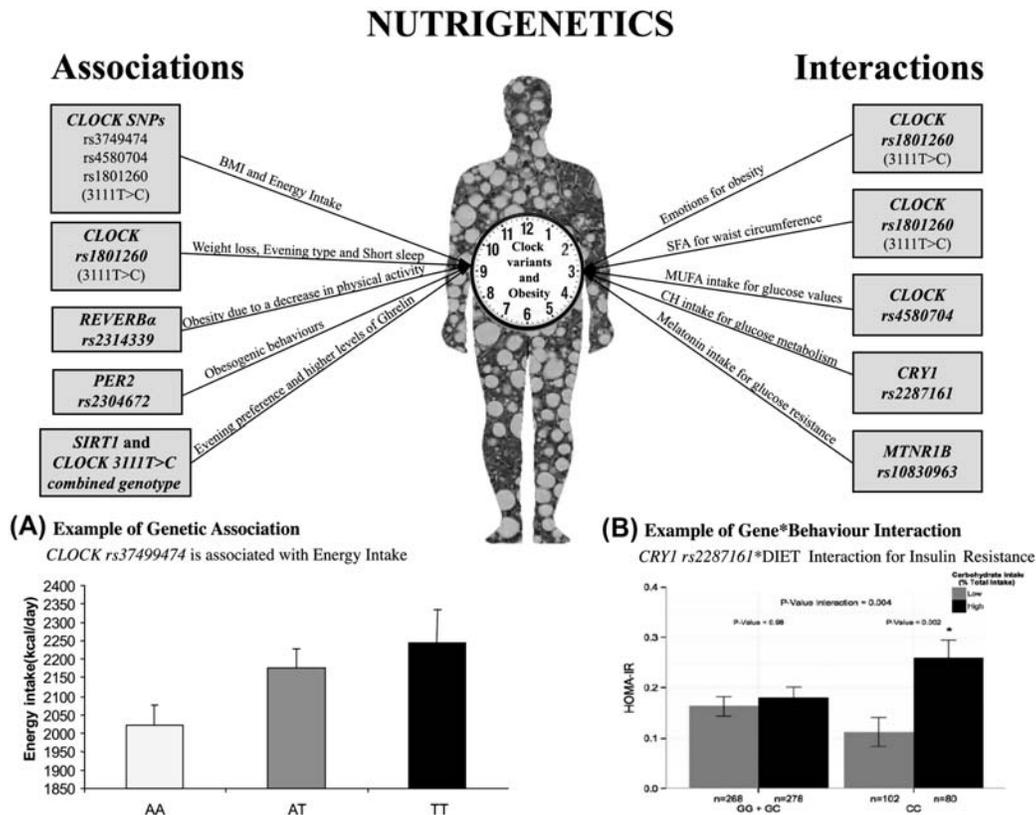


FIGURE 18.1 How several impairments in the central clock are related directly with obesity. In humans, mutations are rare; these impairments are caused by human genetic variants (SNPs). Thus, these human genetic variants may be associated directly with obesity or different behaviors. Moreover, these genetic variants may interact with different behaviors to influence obesity. We can consider several solutions: (A) changes in behaviors directly related to a decrease in obesity and an increase in weight loss; or (B) through genetics, by changes in our behaviors that may interact with SNPs to produce a decrease in obesity and weight loss. Two typical examples are shown of how behaviors may influence genetic variants. Panel A shows an example of genetic association between the genetic variant of *CLOCK* rs37499474 and energy intake, where TT carriers exhibit a greater energy intake than AA carriers. Panel B represents an example of gene–diet interaction for insulin resistance. It is remarkable that when subjects are CC carriers of *CRY1* rs2287161 and also eat high amounts of carbohydrates, their insulin resistance (HOMA) is higher. However, among G carriers (GG + GC), values of HOMA are independent of carbohydrate intake and remain constant. *CH*, cholesterol; *CLOCK*, Circadian Locomotor Output Cycles Kaput; *MUFA*, monounsaturated fatty acids; *PER2*, Period 2; *SFA*, saturated fatty acids.

Moreover, the combination of *SIRT1* and *CLOCK* 3111T > C was associated with evening preference, with less adherence to Mediterranean food intake patterns and higher plasma levels of ghrelin (the hunger hormone) and, finally, with less effective weight loss treatment.

One of the greatest challenges in weight loss treatments is detecting good predictors of success. An analysis of genes, dietary intake, and behavioral factors could be used as a tool to increase the success of such treatments.

Genetic Variants in Clock are Associated With Changes in Circadian Rhythmicity (CD), and These Changes may Predict a Less Effective Treatment.

To achieve a better understanding of the relation between circadian rhythms and obesity, we evaluated

changes in circadian rhythmicity with *CLOCK* 3111T > C SNP in overweight women by continuously recording body peripheral temperature (wrist temperature [WT]), actimetry, and position over 1 week. These variables have been described as good markers of CD (Bandin et al., 2013). Our results showed that risk carriers (C) had significant circadian abnormalities, such as (1) a lower amplitude and (2) greater fragmentation of the rhythm; (3) a less stable circadian pattern; and (4) significantly decreased circadian function, as was shown after assessing the circadian function index). C subjects were also less active, started their activities later in the morning, and were sleepier during the day, showing a delayed acrophase that characterizes ‘evening-type subjects (Bandin et al., 2013).

In a subsequent study recording the same markers of CD (WT, actimetry, and position), we demonstrated that

alterations in circadian rhythms (CD) could be good predictors of less effective weight-loss treatments. Our results showed that subjects who were low responders in treatment displayed a more flattened pattern of WT, characterized by lower amplitude, higher intradaily variability, and deteriorated circadian functionality index compared with individuals who had a greater response to treatment. This work supports the hypothesis that analyzing circadian rhythms of subjects at the beginning of treatment could be useful to predict future weight loss.

NUTRIGENETICS SAYS THAT OUR GENOME MAY INTERACT WITH OUR BEHAVIORS

For many years, researchers have searched for potential interactions between genetic and environmental behaviors, such as nutrition, for various physiological and pathophysiological aspects. In fact, it has been demonstrated that different SNPs in our circadian machinery interact with dietary intake and several behaviors for obesity and for other metabolic risk (MetS)-related variables.

In this sense, in 2009 we revealed several interactions between *CLOCK* variants and dietary intakes for different MetS traits. We showed a significant interaction between *CLOCK* SNP rs4,580,704 and monounsaturated fatty acids (MUFA) intake for glucose plasma values and homeostatic model assessment (HOMA): “the protective effect of the minor allele on insulin sensitivity was only present when MUFA intake was >13.2% of energy” (Garaulet et al., 2009). We also found different effects across *CLOCK* 3111T > C genotypes for saturated fatty acid intake (% of energy): “The deleterious effect of C risk variant on waist circumference was only found with high saturated fatty acid intakes (>11.8%)” (Garaulet et al., 2009).

More recently, one study carried out in 2014 by our group showed a novel *CRY* × diet interaction for insulin resistance in a Mediterranean population that was replicated in a North American population (Dashti et al., 2014). Findings from the metaanalyses indicated that “an increase in carbohydrate intake was associated with an increase in HOMA-IR, in fasting insulin, and with a decrease in quantitative insulin sensitivity check index (QUICKI), but only among individuals homozygous for the minor C allele in *CRY1* polymorphism, rs2287161” (Dashti et al., 2014). *CRY*s (negative elements of clock) are implicated in regulating glucose metabolism. Studies in experimental animals demonstrated that gluconeogenesis can be modulated by circadian changes in the hepatic expression of *Cry*. Therefore, “factors which affect the expression of *CRY* could impair hepatic regulation of glucose homeostasis and thus increase risk for diabetes in humans” (Dashti et al., 2014).

These results can help us design more effective dietary programs. For CC carriers at *CRY1* rs2287161, our advice would focus on helping them reduce carbohydrate intake, so HOMA risk could be even lower than for G carriers (Fig. 18.1).

Another example, but this time for gene–drug interactions, is related to the melatonin receptor 1B (*MTNR1B*) rs10830963 variant. *MTNR1B* rs10830963 G risk variant worsens the effect of exogenous melatonin on glucose tolerance; thus, G carriers should avoid having food together with exogenous melatonin administration (Garaulet et al., 2015). These results could affect subjects who take melatonin (i.e., for sleep problems), but also G carriers who usually have food (at dinner) together with high levels of endogenous melatonin, as happens in natural late eaters or in shift workers. Further studies should answer specific questions about the timing of food intake, genetic risk, and insulin resistance.

Table 18.1 provides advice to prevent genetic and behavioral associations and interactions regarding obesity that may be crucial for genetically informed personalized nutrition in obesity treatment.

All of these examples and many others show how our behaviors may interact with our genes and might decrease the deleterious effect of one specific risk variant. Consequently, regarding the question of whether we are predestined, nutrigenetics answers that no, we are not. Although we cannot change our genome, we can change our behaviors to improve our health.

EPIGENOMICS SAYS THAT WE CAN ALSO CHANGE OUR GENOME WITH OUR BEHAVIORS

It is evident that genes are modulated by a set of regulators that switch gene expression on or off, altering a person’s phenotype. The field focused on the study of these genetic expression regulators is called epigenetics.

The main idea underlying this field is that epigenetics does not change our DNA. However, it decides how much or whether some genes are expressed in different cells of our body. These changes can occur as a result of three processes: (1) DNA methylation, (2) chromatin modifications, and (3) posttranscriptional gene silencing mediated by RNA. The best characterized epigenetic mechanism is methylation, which consists of adding one methyl group to a cytosine of DNA when it stands next to guanine nucleotide (dinucleotide CpG, [–C–phosphate–G–]) or cytosine and guanine separated by one only phosphate.

TABLE 18.1 Advice to Prevent Genetic and Behavioral Associations and Interactions for Obesity that may be Crucial for Genetically Informed Personalized Nutrition in Obesity Treatment.

Single nucleotide polymorphisms (SNP)	Evidence	Recommendation	References
1. Evidence in nutrigenetics			
1.A. Associations			
CLOCK SNPs rs3749474 rs4580704 rs1801260 (3111T > C)	Associated with body mass index, energy intake, and different variables related to obesity (minor allele carriers ate more, ate more fat, and were more obese)	In minor alleles carriers: - Decrease energy intake - Decrease total fat intake	Garaulet et al. (2009)
CLOCK rs1801260 (3111T > C)	Associated with weight loss (C allele carriers were more resistant to weight loss) and were more evening type, slept less, and had less adherence to Mediterranean diet	In C allele carriers: - Sleep at least 8 h/day - Get up earlier in morning - Go to bed earlier - Follow Mediterranean diet patterns	Lopez-Guimera et al. (2014)
RevErba rs2314339	Associated with obesity owing to decrease in physical activity; physical activity was performed in afternoon	In A allele carriers: - Increase physical activity - Perform physical activity in morning	
Period 2 rs2304672	Associated with several obesogenic behaviors such as attrition of weight loss treatment, snacking, stress while dieting, eating while bored, and skipping breakfast (allele minor carriers)	In G allele carriers: - Develop a stronger follow-up plan during dietary therapy - Avoid snacking - Avoid being around food when bored - Try having always breakfast	
SIRT1 and CLOCK 3111T > C combined genotype	Associated with evening preference, with fewer adherences to Mediterranean food intake patterns and with higher plasma levels of ghrelin in subjects carrying minor alleles of both SNPs	In minor alleles carriers of both SNPs: - Get up earlier in morning - Follow Mediterranean diet patterns - Try to have 5 meals/day to prevent hunger	
1.B. Interactions			
CLOCK rs1801260 (3111T > C)	Interact with: 1. Emotions for obesity (among C allele carriers, emotional eaters lost significantly less weight than non-emotional eaters)	Among C allele carriers emotional eaters: - Develop stronger follow-up plan during dietary therapy - Try to have SFA energy intake lower than 11.8%	(Lopez-Guimera et al., 2014; Garaulet et al., 2009)

Continued

TABLE 18.1 Advice to Prevent Genetic and Behavioral Associations and Interactions for Obesity that may be Crucial for Genetically Informed Personalized Nutrition in Obesity Treatment.—cont'd

Single nucleotide polymorphisms (SNP)	Evidence	Recommendation	References
<i>CLOCK</i> rs4580704	2. SFA for waist circumference (among C allele carriers when the SFA energy intake was more than 11.8% had higher waist circumference) Interact with MUFA intake (% of energy) for glucose plasma values and homeostatic model assessment (HOMA) (MUFA intake was higher than 13.2%, had minor allele carriers, had lower glucose plasma concentrations and HOMA than noncarriers)	In CC carriers: - Increase intake of olive oil	Garaulet et al. (2009)
<i>CRY1</i> rs2287161	Interact with carbohydrate intake for glucose metabolism (among CC carriers when they ate high values of carbohydrates, their insulin resistance [HOMA] was higher than G carriers [GG + GC])	Among CC carriers: - Try to eat less carbohydrates from total energy intake	Dashti et al. (2014)
<i>MTNR1B</i> rs10830963	Interact with melatonin intake for glucose resistance	In GG carriers: - Avoid consuming food when melatonin levels are elevated	Garaulet et al. (2015)
2. Evidences in epigenetic			
CpG sites of <i>CLOCK</i>	Eating behaviors, such as high frequency of snacking, eating quickly, eating when bored, or eating from large packages, were positively associated with methylation levels of <i>CLOCK</i> CpG 1	Try to avoid these eating behaviors: snacking, eating slowly, eating when bored, and or eating from large packages	
CpG sites of <i>BMAL1</i>	Evening-type people had more epigenetics modifications (high methylations in CpG 5–9) owing to weight loss intervention than did morning types	Get up earlier in morning	

CLOCK, Circadian Locomotor Output Cycles Kaput; *MUFA*, monounsaturated fatty acids; *SFA*, saturated fatty acids.

TABLE 18.2 Effect of Different Behaviors in Methylation Levels of *CLOCK* CpG1, Showing Importance of how we eat.

Evidences in epigenetic		
<i>CLOCK</i> CpG 1	Methylation levels	P value
Snacking frequency	12×	.026
Eat fast	9×	.08
Eat when bored	3×	.008
Eat from large packages	19×	.004

From Fermín I. Milagro et al. *Chronobiology International*. 2012.

In this sense, it has been shown that circadian clock gene expressions can be regulated by epigenetic mechanisms. Moreover, methylation levels in clock genes have been associated to obesity and metabolic disturbances.

In 2012, our group, in collaboration with Drs. Milagro and Martinez from the University of Navarra, published a study demonstrating significant associations between the methylation levels of several CpGs located in *CLOCK* with MetS, weight loss, and obesity. This study, performed on 60 women, demonstrated that the degree of methylation in the CpG sites of *CLOCK*, such as CpG 1, 5, 6, and 8, increased with obesity. More important, how we eat was related to methylation levels at *CLOCK*. We observed that patients who tended to snack frequently had 12 times higher methylation levels in *CLOCK* CpG1, whereas those who tended to eat when bored or to eat from large packages had 9 to 19 times higher methylation levels. These increases in methylation levels of blood mononuclear cells suggest a suppression of *CLOCK* expression, which has also been related to obesity. Table 18.2 represents the effect of different behaviors in methylation levels of *CLOCK* CpG1, showing the importance of how we eat.

More recently, a second research demonstrated that a weight loss nutritional intervention modified the methylation pattern of *BMAL1*, *CLOCK*, and *NR1D1* in whole blood. More important, these changes in methylation levels of *BMAL1* were associated with a reduction in metabolic risk parameters, i.e., serum lipids. Interestingly, evening types had more epigenetic modifications caused by the weight loss intervention than did morning types. Moreover, baseline methylation of *BMAL1* was positively correlated with energy intake and carbohydrate intake, which suggests that interventions designed to decrease total energy intake and carbohydrates may decrease vulnerability to CD and obesity, especially in evening-type subjects (Table 18.1).

WHAT CAN WE DO? WE CAN CHANGE WHAT, HOW, AND WHEN WE EAT

In this review, we have shown so far that we can interact with our genome and even change it and improve our health through changes in how we eat, what we eat, and other daily behaviors. One novel aspect to consider in dietary interventions may be when we eat: the timing of food intake. If we consider that eating is an external synchronizer of our peripheral clock and that an unusual eating time may disrupt our circadian system, when we eat may have a significant role in obesity treatment. Along this line, a novel observational study performed in 2013 revealed that eating late may influence the success of weight-loss therapy, leading to a decrease in its effectiveness. This study was performed on 411 overweight and obese subjects who underwent dietary weight-loss treatment: 199 subjects were early eaters (had their main meal of the day, lunch, before 3:00 p.m.) and 212 were late eaters (had lunch after 3:00 p.m.). Late eaters lost significantly less weight than did early eaters, although they were of similar age, appetite hormones values, energy intake, sleep duration, or macronutrient distribution. It is remarkable that late eaters were more evening types (Garaulet et al., 2009, 2010b). The study for Spanish subjects concluded that if we could change behavior toward eating lunch earlier, perhaps we could lose more weight during a dietary treatment.

Another observational study conducted in collaboration with Dr. Izquierdo from the University of Barcelona showed that the timing of food intake may also influence the effectiveness of bariatric surgery in severely obese subjects. Many subjects were good responders to

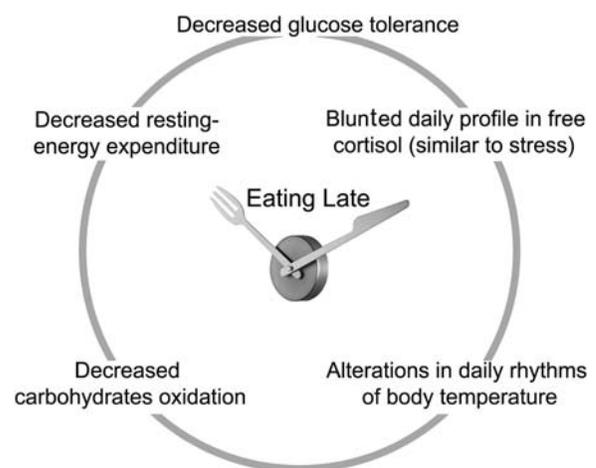


FIGURE 18.2 How eating late affects the metabolic characteristics of lean women. After 1 week of eating late, their metabolic pattern was similar to that of obese women.

TABLE 18.3 Specific Recommendations for Preventing Obesity and Metabolic Syndrome by Improving Circadian System Health, Based on Available Scientific Evidence Proved in (A) Epidemiological, Clinical, and Experimental Animals (A); and Based on (B) Interventional Studies in Humans or in Animal Models and (C) Preliminary Studies or a Low Number of Studies.

Recommendation	Evidence	Type of Evidence
<i>Sleep</i>		
Sleep during night and be active during day	Shift workers face potential health problems. Overall, those who work at night or have rotating shifts seem to have higher risk for insulin resistance, metabolic syndrome, and heart disease	A
Sleep enough hours per day (7–8 h)	Insufficient sleep increases risk for insulin resistance, type 2 diabetes, and obesity, which suggests that sleep restriction may impair peripheral metabolic pathways	A
Go to bed early and get up early	Late chronotypes were previously related to obesity and less healthy and more irregular lifestyles	A
Take short (20-min) naps in the middle of the day	Napping for 90 mins too late in the afternoon can affect nighttime sleep, reducing sleep efficiency and resulting in earlier wake times	B
<i>Light</i>		http://www.ncbi.nlm.nih.gov/pubmed/11560181
Avoid intense light exposure during nighttime and sleep in total darkness when possible	Light is an external synchronizer of circadian rhythms	A
<i>Meal timing</i>		A
Eat the main meal of the day before 3 p.m.	Eating late is associated with several metabolic alterations	B
Eat dinner at least 2.5 h before going to bed	The coincidence of melatonin and food intake (glucose) is associated with impaired glucose tolerance	C
Avoid eating at night	Increasing scores on the night eating questionnaire (NEQ) correlated positively with increasing body mass index and blood pressure in women. NEQ scores correlated positively with waist circumference and triglyceride levels in men	A
<i>Timing for physical activity</i>		
Avoid performing physical activity for 2–3 h before usual sleep time	Evening physical activity alters wrist temperature circadian rhythmicity	C

treatment (68%). They lost 80% of the initial excess weight during the first year after surgery and maintaining the weight loss during the 6 years of follow-up. However, about 11% of the population was defined as primarily weight-loss poor responders. Those subjects lost only about 40% of their initial excess weight during the first year after surgery. From a clinical viewpoint, it is important to detect this group of poor responders before surgery to decide on further treatment. The results of the study demonstrated that the percentage of late eaters was significantly higher among primarily poor weight-loss responders (~70%) than in good weight-loss responders (~37%). Moreover, poor weight-loss responders had lunch later than did good weight-loss responders. Surprisingly, obesity-related variables, biochemical parameters, presurgical total energy expenditure, sleep duration, chronotype, calorie intake, and macronutrient distribution were similar among groups.

To discover why the timing of food intake could influence weight loss, we carried out a randomized study in which 32 lean young women completed two protocols: one including assessments of resting energy expenditure (indirect calorimetry) and glucose tolerance and the other including circadian-related measurements, such as profiles in salivary cortisol and WT (Fig. 18.2). Participants received standardized meals during both meal intervention weeks and were studied at two times: early eating (lunch at 13:00) and late eating (lunch 16:30). Results of the study showed that after 1 week of eating late, these lean women experienced metabolic alterations that usually characterize obese women, such as decreased glucose tolerance, decreased resting energy expenditure, and decreased carbohydrate oxidation. Moreover, late eaters had a flattened pattern of daily cortisol and alterations in daily rhythms of body temperature similar to those that characterize obese women, as previously described.

Table 18.3 gives specific recommendations for preventing obesity and MetS by improving circadian system health, based on available scientific evidence proved in (A) epidemiological, clinical, and experimental animals; based on (B) interventional studies in humans or in animal models and in (C) preliminary studies or a low number of studies. These recommendations can be useful for the general population and in the clinical practice to prevent or treat obesity alterations.

SUMMARY AND CONCLUSIONS

Fig. 18.1 shows the current revision. Several impairments in the central clock are directly related to obesity. In humans, mutations are rare and are caused by human genetic variants. These variants may be associated

directly with obesity or through different behaviors. Moreover, these genetic variants may interact with different behaviors to influence obesity and weight loss. We can consider several solutions (1) changing our behaviors: these changes are directly related to a decrease in obesity and an increase in weight loss; or (2) through genetics, by changes in our behaviors that may interact with SNPs to produce a decrease in obesity and weight loss.

We propose modifying what, how, and when we eat as an effective tool to decrease genetic risk and consequently to diminish CD and decrease obesity. The research addressed in this revision presents a novel and promising area in obesity prevention and treatment.

Financial disclosure

This study was supported by grants from the Spanish Government of Economy and Competitiveness (SAF2014-52480-R) and the European Regional Development Fund (ERDF) to Marta Garaulet.

Conflict of interest: None.

References

- Bandin, C., Martinez-Nicolas, A., Ordovas, J.M., Ros Lucas, J.A., Castell, P., Silvente, T., et al., 2013. Differences in circadian rhythmicity in CLOCK 3111T/C genetic variants in moderate obese women as assessed by thermometry, actimetry and body position. *Int J Obes* 37 (8), 1044–1050. Epub 2012/11/28.
- Corbalan, M.D., Morales, E.M., Canteras, M., Espallardo, A., Hernandez, T., Garaulet, M., 2009. Effectiveness of cognitive-behavioral therapy based on the Mediterranean diet for the treatment of obesity. *Nutrition* 25 (7–8), 861–869. Epub 2009/06/23.
- Dashti, H.S., Smith, C.E., Lee, Y.C., Parnell, L.D., Lai, C.Q., Arnett, D.K., et al., 2014. CRY1 circadian gene variant interacts with carbohydrate intake for insulin resistance in two independent populations: Mediterranean and North American. *Chronobiol Int* 31 (5), 660–667. Epub 2014/02/20.
- Garaulet, M., Lee, Y.C., Shen, J., Parnell, L.D., Arnett, D.K., Tsai, M.Y., et al., 2009. CLOCK genetic variation and metabolic syndrome risk: modulation by monounsaturated fatty acids. *Am J Clin Nutr* 90 (6), 1466–1475. Epub 2009/10/23.
- Garaulet, M., Corbalan, M.D., Madrid, J.A., Morales, E., Baraza, J.C., Lee, Y.C., et al., 2010. CLOCK gene is implicated in weight reduction in obese patients participating in a dietary programme based on the Mediterranean diet. *Int J Obes* 34 (3), 516–523. Epub 2010/01/13.
- Garaulet, M., Lee, Y.C., Shen, J., Parnell, L.D., Arnett, D.K., Tsai, M.Y., et al., 2010. Genetic variants in human CLOCK associate with total energy intake and cytokine sleep factors in overweight subjects (GOLDN population). *Eur J Hum Genet* 18 (3), 364–369. Epub 2009/11/06.
- Garaulet, M., Gomez-Abellan, P., Rubio-Sastre, P., Madrid, J.A., Saxena, R., Scheer, F.A., 2015. Common type 2 diabetes risk variant in MTNR1B worsens the deleterious effect of melatonin on glucose tolerance in humans. *Metab Clin Exp* 64 (12), 1650–1657. Epub 2015/10/07.
- Lopez-Guimera, G., Dashti, H.S., Smith, C.E., Sanchez-Carracedo, D., Ordovas, J.M., Garaulet, M., 2014. CLOCK 3111 T/C SNP interacts with emotional eating behavior for weight-loss in a Mediterranean population. *PLoS One* 9 (6), e99152. Epub 2014/06/07.

This page intentionally left blank

Screening for Inborn Errors of Metabolism

Adela Chirita-Emandi

Genetics Department, Center for Genomic Medicine, University of Medicine and Pharmacy “Victor Babes”, Timisoara, Romania

Glossary

Aminoacidopathies Diseases that result from abnormalities in the breakdown of amino acids in the cytosol.

Disorders of fatty acid oxidation and ketogenesis Caused by defects in mitochondrial fatty acid oxidation required for delivery of energy during fasting, via complete oxidation or through production of ketones in the liver that then serve as an alternative energy source for the brain.

Inborn errors of metabolism Defects in enzymes of the metabolism of proteins, carbohydrates, and fatty acids or energy metabolism of mitochondria.

Lysosomal storage disorders Disorders caused by deficient enzymes or hydrolases required for intracellular breakdown of large lipid and mucopolysaccharide molecules, causing substrate to accumulate and enlarge and/or result in functional impairment of the organ system.

Mitochondrial disorder Disorders of energy metabolism that include genetic defects of the pyruvate dehydrogenase complex, the Krebs cycle, and the respiratory electron transport chain, encompassing pathways of substrate breakdown and production of adenosine triphosphate.

Organic acidurias Deficiency of enzymes in mitochondrial metabolism of coactivated carboxylic acids, most of which are derived from amino acid breakdown. The term “organic acidurias” is preferred to “organic acidemias” because it is most often detected by analysis of the urine.

Peroxisomal disorders Group of genetically heterogeneous metabolic diseases that share dysfunction of peroxisomes, which are cellular organelles that participate in important peroxisome-specific metabolic pathways, such as β -oxidation of very long-chain fatty acids and detoxification of hydrogen peroxide. Peroxisomes are also involved in the production of cholesterol, bile acids, and plasmalogens, which contribute to a major part of the phospholipid content of brain white matter.

Urea cycle disorders Disorders of ammonia detoxification, diseases caused by defects in breakdown of protein that produce large amounts of nitrogen (ammonia) that are highly neurotoxic (which normally would be converted to urea and excreted in the urine).

INTRODUCTION

Inborn errors of metabolism (IEM) are genetic disorders in which specific enzyme defects disturb the metabolism of exogenous or endogenous proteins, carbohydrates, or fat. IEM disorders may manifest at any stage of life. Whereas onset in the newborn period or infancy is well-recognized, onset in adulthood is much less known. Furthermore, owing to early diagnosis and proper management, many adults received a diagnosis of IEM in childhood. Almost all countries around the globe include IEM in state newborn screening programs, but the number of diseases varies greatly between countries and among different states within the United States (Therrell et al., 2015). In principle, the result of reduced or absent enzyme activity with all IEM leads to toxic metabolites that build up in the blood, brain, and other organs, whereas other compounds may become deficient, leading to adverse health outcomes. Therefore, nutritional interventions that can avoid or overcome metabolic consequences are warranted for some IEM. Early diagnosis and treatment, preferably at birth, can often diminish or surmount the adverse effects of some IEM, resulting in almost normal health. For many IEM, nutritional intervention is the primary therapy and is required lifelong (Therrell et al., 2014). Therefore, screening programs represent an excellent opportunity to establish the diagnosis quickly. Diseases to be screened in such a program had to meet the following criteria:

- the disorder occurs with significant frequency;
- it has acceptable costs and a reliable method of testing at 24–48 h after birth, when it would not usually be clinically detected; and

- there are demonstrated benefits of early detection, timely intervention, and efficacious treatment.

Several guidelines have become available over the years since newborn screening started, that are available from the organization Newborn Screening Translational Research Network (NBSTRN) (Lloyd-Puryear et al., 2018). NBSTRN provides tools and resources for researchers to conduct research relevant to newborn screening programs for rare diseases; for analytical and clinical validation of screening tests; and for the collection, analysis, sharing, and reporting of longitudinal laboratory and clinical data on newborn screened individuals.

Usually, newborns are screened for amino acid disorders, urea cycle disorders, organic acid disorders, and fatty acid oxidation defects in many states in the United States, whereas in European countries, the number of disorders in screening varies from 2 to 30 (Vittozzi et al., 2010; Camp et al., 2012). A few states in the United States are planning to include lysosomal storage diseases and peroxisomal disorders among their newborn screening panels. Table 19.1 provides examples of the most common IEM included in screening programs.

Newborn screening programs are based on tandem-mass spectrometry and other newborn technologies. After implementation, these methods offer the possibility of extending the panels without significant investment and cost. Their availability and a significant decrease in the cost of next-generation sequencing have raised the question of replacing or combining metabolic newborn screening methods with whole-genome or whole-exome sequencing. Translating next-generation sequencing technologies to the clinic has become a goal of clinical genomics. The use of sequencing in newborn screening can expand the number of genetic conditions detected before symptoms show (in utero or shortly after birth). However, this approach for newborn screening raises challenging problems ranging from the possibility of false-positive results to different ethical concerns and the issue of extending the role of parental decision-making (Milko et al., 2018).

NUTRITIONAL INTERVENTIONS FOR INBORN ERRORS OF METABOLISM

Existing nutritional interventions for IEM include dietary modifications to exclude specific nutrients that cannot be metabolized, along with medical foods and dietary supplements. A medical food is defined by the

Food and Drug Administration as “a food which is formulated to be consumed or administered enterally under the supervision of a physician and which is intended for the specific dietary management of a disease or condition for which distinctive nutritional requirements, based on recognized scientific principles, are established by medical evaluation.”

Medical foods for IEM are grouped into two categories. One group of medical foods is represented by foods that contain enough nutrients to meet nutritional requirements; it excludes nutrient(s) that cannot be metabolized for a specific IEM. An example, is phenylketonuria, in which phenylalanine is excluded and sometimes tyrosine is added. There are a variety of presentations for medical foods, such as drinks made by reconstituting powders, ready to drink products, customized modular products, and bars. The other group of medical foods is represented by products modified to be low in protein but still providing the required energy in the diet (for example, specially modified flour, cereals, biscuits, meat and cheese substitutes, pasta, and rice) (Camp et al., 2012).

Alongside dietary restriction and medical foods, dietary supplements provide for unmet nutritional needs (e.g., essential amino acids, vitamins, or minerals), enhance enzyme activity (e.g., vitamins), or assist in eliminating toxic metabolites (e.g., carnitine). Dietary supplements can be extracts or concentrates, or they may be tablets, capsules, softgels, gelpcaps, liquids, and powders, or other forms such as bars. Dietary supplements rarely undergo the Food and Drug Administration approval process required for drugs.

The therapeutic effects of nutritional interventions in IEM have the potential to improve cognitive and behavioral disturbance(s), prevent acute metabolic decompensation, and prevent or slow clinical deterioration (van Karnebeek and Stockler, 2012). The effect depends on the onset of the dietary intervention, compliance with treatment, the type of IEM, and the severity of the disease (partly depending on the genotype) (Table 19.2).

A significant number of IEM are amendable to dietary intervention. However, daily adherence is important to obtain metabolic control. Because of the lack of information and challenges for families to comply with specific dietetic regimens, several apps and online tools have been developed. One example is the Metabolic Diet App Suite (<http://www.metabolicdietapp.org>), created as a free online tool for 15 different IEM. General information is provided regarding the nutrient content of foods, as well as calculators and trackers of dietary intake and targets for each IEM (Ho et al., 2016).

TABLE 19.1 Examples of Disorders Most Commonly Included in Newborn Screening.

Group of IEM	Disease
Amino acid metabolism disorders	<ul style="list-style-type: none"> • Phenylketonuria (clinically significant hyperphenylalaninemia variants) • Maple syrup urine disease • Homocystinuria • Tyrosinemia • Citrullinemia
Carbohydrate disorders	<ul style="list-style-type: none"> • Galactosemia • Glycogen storage disease type I, II (Pompe disease), III, IV (Andersen disease), V, VI
Fatty acid oxidation defects	<ul style="list-style-type: none"> • Medium-chain acyl-CoA dehydrogenase deficiency • Long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency • Trifunctional protein deficiency type 1 (LCHAD deficiency) • Trifunctional protein deficiency type 2 (mitochondrial trifunctional protein deficiency) • Carnitine uptake defect (primary carnitine deficiency) • Very long-chain acyl-CoA dehydrogenase deficiency
Organic acid disorders (organic aciduria or organic acidemia)	<ul style="list-style-type: none"> • Isovaleric acidemia • 3-Methylcrotonyl-CoA carboxylase deficiency • Glutaric acidemia type I • Glutaric acidemia type II • 3-Hydroxy-3-methylglutaryl-coenzyme A lyase deficiency • Multiple carboxylase deficiency (biotinidase deficiency, holocarboxylase synthetase deficiency) • Methylmalonic acidemia • Propionic acidemia • β-Ketothiolase deficiency
Mitochondrial disorders	<ul style="list-style-type: none"> • Leber hereditary optic neuropathy • Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes • Mitochondrial neurogastrointestinal encephalopathy disease • Myoclonic epilepsy with ragged-red fibers • Neuropathy, ataxia, and retinitis pigmentosa • Pyruvate carboxylase deficiency
Peroxisomal disorders	<ul style="list-style-type: none"> • Zellweger syndrome spectrum • Adrenoleukodystrophy (α-ALD)
Urea cycle disorders	<ul style="list-style-type: none"> • Citrullinemia • Argininosuccinic aciduria • Carbamoyl phosphate synthetase I deficiency

TABLE 19.2 List of Inborn Errors of Metabolism with Dietary Interventions.

Disease Name	Therapeutic Modalities
Arginine: glycine amidinotransferase deficiency	Creatine supplements
Argininemia, argininosuccinic aciduria, citrullinemia type I and II, carbamoyl phosphate synthetase deficiency, N-acetylglutamate synthetase deficiency	Dietary protein restriction, arginine supplement, sodium benzoate, phenylbutyrate (liver transplantation)
β -Ketothiolase deficiency	Avoid fasting, sick day management, protein restriction
Biotin-responsive basal ganglia disease biotinidase deficiency, holocarboxylase synthetase deficiency	Biotin supplement
Cerebral folate receptor- α deficiency	Folinic acid
Cerebrotendinous xanthomatosis	Chenodeoxycholic acid, 3-hydroxy-3-methyl-glutaryl reductase inhibitor
Coenzyme Q10 deficiency	Coenzyme Q10 supplements
Cobalamin A and B deficiency	Hydroxycobalamin, protein restriction
Cobalamin C and F deficiency, congenital intrinsic factor deficiency, Imerslund–Gräsbeck syndrome	Hydroxycobalamin
Cobalamin D, E, and G deficiency	Hydroxy-/cyanocobalamin
Creatine transporter defect	Creatine, glycine, arginine supplements
Dihydropteridine reductase deficiency	BH ₄ , diet, amine replacement, folinic acid
Ethylmalonic encephalopathy	N-acetylcysteine, oral metronidazole
Guanidino-acetate-N-methyltransferase deficiency	Arginine restriction, creatine, and ornithine supplements
GLUT1 deficiency syndrome	Ketogenic diet
Glutaric acidemia I	Lysine restriction, carnitine supplements
Glutaric acidemia II	Carnitine, riboflavin, β -hydroxybutyrate supplements, sick day management
GTP cyclohydrolase deficiency	BH ₄ , amine replacement
Hyperornithinemia, hyperammonemia, homocitrullinemia syndrome)	Dietary protein restriction, ornithine supplement, sodium benzoate, phenylacetate
3-Hydroxy-3-methylglutaryl-CoA lyase deficiency	Protein restriction, avoid fasting, sick day management,
Homocystinuria	Methionine restriction, \pm pyridoxine, \pm betaine
Isovaleric acidemia	Dietary protein restriction, carnitine supplements, avoid fasting, sick day management
Nonketotic hyperglycinemia	Glycine restriction; \pm sodium benzoate, N-methyl-D-aspartic acid receptor antagonists, other neuromodulating agents
Maple syrup urine disease (variant)	Dietary restriction branched amino acids, avoid fasting, (liver transplantation)
Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes	Arginine supplements
Menkes disease occipital horn syndrome	Copper histidine
3-Methylcrotonyl glycinuria	Dietary protein restriction; carnitine, glycine, biotin supplements; avoid fasting; sick day management
3-Methylglutaconic aciduria type I	Carnitine supplements, avoid fasting, sick day management
Methylmalonic acidemia	Dietary protein restriction, carnitine supplements, avoid fasting, sick day management
2-Methyl-3-hydroxybutyryl-CoA dehydrogenase deficiency	Avoid fasting, sick day management, isoleucine restricted diet

TABLE 19.2 List of Inborn Errors of Metabolism with Dietary Interventions.—cont'd

Disease Name	Therapeutic Modalities
Mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase deficiency	Avoid fasting, sick day management, ±dietary precursor restriction
Molybdenum cofactor deficiency type A	Precursor Z/cyclic pyranopterm monophosphate
Methylenetetrahydrofolate reductase deficiency	Betaine supplements, ± folate, carnitine, methionine supplements
Ornithine transcarbamoylase deficiency	Dietary protein restriction, citrulline supplements, sodium benzoate/phenylbutyrate (liver transplantation)
Pterin-4 α -carbinolamine dehydratase deficiency	BH4
Pyruvate dehydrogenase complex deficiency	Ketogenic diet and thiamine
Phenylketonuria	Dietary phenylalanine restriction ± amino acid supplements (BH4 supplement)
Phosphoglycerate dehydrogenase deficiency, phosphoserine aminotransferase deficiency, phosphoserine phosphatase deficiency	L-serine and ± glycine supplements
Propionic acidemia	Dietary protein restriction, carnitine supplements, avoid fasting, sick day management
6-Pyruvoyltetrahydropterin synthase deficiency	BH4, diet, amine replacement
Pyridoxine-dependent epilepsy	Pyridoxine
Pyrimidine 5-nucleotidase superactivity	Uridine supplements
Succinyl-CoA 3-oxoacid CoA transferase deficiency	Avoid fasting, protein restriction, sick day management
Sjögren–Larsson syndrome	Diet: low fat, medium chain and essential fatty acid supplements and Zileuton
Smith–Lemli–Opitz syndrome	Cholesterol and simvastatin
Sepiapterin reductase deficiency	Amine replacement
Succinic semialdehyde dehydrogenase deficiency	Vigabatrin
Thiamine-responsive encephalopathy	Thiamin supplement
Tyrosine hydroxylase deficiency	L-Dopa substitution
Tyrosinemia type II	Dietary phenylalanine and tyrosine restriction
Wilson disease	Zinc and tetrathiomolybdate

Adapted from van Karnebeek, C.D.M., Stockler, S., 2012. Treatable inborn errors of metabolism causing intellectual disability: a systematic literature review. *Mol Genet Metab* 105(3), 368–381.

CONCLUSION

Without treatment, infants and children with IEM experience serious adverse health outcomes including intellectual disability, abnormal development, nutrient deficiencies, and sequelae that require complex medical care. Newborn screening methods are successful in establishing the diagnosis at birth and appropriate nutritional interventions are available for many IEM, dramatically improving the prognosis.

References

Camp, K.M., Lloyd-Puryear, M.A., Huntington, K.L., 2012. Nutritional treatment for inborn errors of metabolism: indications,

- regulations, and availability of medical foods and dietary supplements using phenylketonuria as an example. *Mol Genet Metab* 107 (1–2), 3–9.
- Ho, G., Ueda, K., Houben, R.F.A., Joa, J., Giezen, A., Cheng, B., van Karnebeek, C.D., 2016. Metabolic Diet App Suite for inborn errors of amino acid metabolism. *Mol Genet Metab* 117 (3), 322–327.
- Lloyd-Puryear, M., Brower, A., Berry, S.A., Brosco, J.P., Bowdish, B., Watson, M.S., 2018. Foundation of the newborn screening translational research network and its tools for research. *Genet Med* 1.
- Milko, L.V., Rini, C., Lewis, M.A., Butterfield, R.M., Lin, F.-C., Paquin, R.S., Powell, B.C., Roche, M.I., Souris, K.J., Bailey Jr., D.B., Berg, J.S., Powell, C.M., 2018. Evaluating parents' decisions about next-generation sequencing for their child in the NC NEXUS (North Carolina Newborn Exome Sequencing for Universal Screening) study: a randomized controlled trial protocol. *Trials* 19 (1), 344.
- Therrell, B.L., Lloyd-Puryear, M.A., Camp, K.M., Mann, M.Y., 2014. Inborn errors of metabolism identified via newborn screening:

- ten-year incidence data and costs of nutritional interventions for research agenda planning. *Mol Genet Metab* 113 (1–2), 14–26.
- Therrell, B.L., Padilla, C.D., Loeber, J.G., Kneisser, I., Saadallah, A., Borrajo, G.J.C., Adams, J., 2015. Current status of newborn screening worldwide: 2015. *Semin Perinatol* 39 (3), 171–187.
- van Karnebeek, C.D.M., Stockler, S., 2012. Treatable inborn errors of metabolism causing intellectual disability: a systematic literature review. *Mol Genet Metab* 105 (3), 368–381.
- Vittozzi, L., Hoffmann, G.F., Cornel, M., Loeber, G., 2010. Evaluation of population newborn screening practices for rare disorders in member states of the European Union. *Orphanet J Rare Dis* 5 (Suppl. 1), P26.

Genomics of Eating Behavior and Appetite Regulation

Bayan Mesmar¹, Nanette Steinle^{1,2}

¹University of Maryland School of Medicine, Baltimore, MD, United States; ²VA Maryland Healthcare System, Baltimore, MD, United States

List of Abbreviations

ADAH Age associated cognitive loss and alcohol dependence
AgRP Agouti-related protein
BDNF Brain-derived neurotrophic factor
BMI Body mass index
CART Cocaine and amphetamine-related transcript
CCK Cholecystokinin
CNS Central nervous system
D1R Dopamine receptor D1
DRD2 Dopamine D2 receptor gene
EnaC Epithelia sodium channel
FTO Fat mass and obesity associated
GABA γ -Aminobutyric acid
GIP Gastric inhibitory peptide
GLP1 Glucagon-like peptide 1
GRM8 Glutamate receptor metabotropic 8
LOC Loss of control
mGluR₁ Glutamate receptor, metabotropic 1
mGluR₄ Glutamate receptor, metabotropic 4
NMB Neuromedin beta
NPY Neuropeptide Y
POMC Pro-opiomelanocortin
PPAR- γ Peroxisome proliferator activated receptor γ
PW Prader–Willi
PYY Peptide tyrosin-tyrosin
SNP Single nucleotide polymorphism
T1R2 Taste receptor type 1 member 2
T1R3 Taste receptor type 1 member 3
T2D Type 2 diabetes
TRPV1 Transient receptor potential cation subfamily V member 1
WAGR Wilms tumor, aniridia, genitourinary anomalies, and mental retardation
WC Waist circumference

INTRODUCTION

Eating behavior affects food selection, and nutrition has an impact on health and risk for chronic disease. Nutrition affects a large variety of chronic diseases,

including obesity, type 2 diabetes, cardiovascular disease, and cancer. Globally, over the past 30 years, the prevalence of obesity and its related comorbidities has increased significantly. Obesity, which is a complex disorder involving genetic, physiological, social, financial, and psychological factors, has become a worldwide epidemic. In the recent past, knowledge has emerged regarding the genetic influences of eating behavior.

Here, we highlight current scientific evidence pointing to links between genetic variation and eating behaviors and focus on eating behavior traits, taste, and hormonal pathways known to affect feeding behavior and appetite. A brief discussion of the human gut microbiome concludes this overview of genetic influences of eating behavior.

BACKGROUND

There are three well-recognized patterns of eating behavior: restraint, disinhibition, and hunger. These factors have been associated with the development and maintenance of obesity. Restraint is the cognitive avoidance of certain foods to maintain body weight and body image. Disinhibition is the lack of restraint over food; it includes a range of eating behaviors such as binge eating, loss of control, emotional eating, and eating in the absence of hunger. Hunger refers to the subjective perception of the need for food; the greater this perception is, the more likely it is that the individual will consume food.

Several family studies investigated the contribution of genetics to these eating behavior traits. A twin study by Neale et al., in which the authors examined whether

genetic factors have an important role in developing certain eating behaviors, found evidence demonstrating that the correlation of disinhibition scores is twice as strong in monozygotic female twins than in dizygotic female twins (Neale et al., 2003). In a study by Steinle et al. investigators explored the genetic link of the three eating behaviors in an Amish population. They first established a statistically significant link between obesity and eating behavior scores as assessed by the Three-Factor-Eating questionnaire. Heritability estimates found these eating behavior traits to be highly heritable. After performing linkage analysis, the authors reported evidence of linkage of restraint scores to chromosomal regions D6S276 and D3S1304; the latter was also associated with percent body fat (Steinle et al., 2002). Functional candidate genes residing in the chromosome 3 region include *peroxisome proliferator activated receptor γ* (PPAR- γ). PPAR- γ is expressed in many tissues, but the highest concentrations are found in adipose tissue and the large intestine. PPAR- γ is an important mediator of adipocyte differentiation and energy storage. PPAR- γ agonists stimulate eating by decreasing leptin concentrations. Furthermore, mutations in PPAR- γ are associated with obesity. Functional candidate genes residing in the chromosome 6 region of linkage include the GLP-1 receptor. GLP-1 receptors are found in the hypothalamus and are the target for hypothalamic neurons containing glucose transporter 2 and glucokinase. GLP-1 secretion after a meal stimulates insulin release, lowers blood glucose, and reduces food intake. In the Quebec Family Study, genetic variants in *glutamate decarboxylase 2* (GAD2) associated with disinhibited eating behavior and susceptibility to hunger (Choquette et al., 2009). GAD2 encodes the glutamic acid decarboxylase enzyme, which may have a role in γ -aminobutyric acid (GABA) regulation. GABA is a neurotransmitter involved in regulating food intake. Also in the Quebec Family cohort, *neuromedin β* , encoding a satiety peptide, was linked to disinhibition as well as susceptibility to hunger and weight gain over time (Bouchard et al., 2004). Increased restraint scores were observed in a Sorb cohort living in Germany who carry a variant in *glutamate receptor, metabotropic 8*, a gene encoding a G protein-coupled glutamate receptor associated with smoking behavior and liability toward alcoholism, implying a role in addiction vulnerability (Gast et al., 2013). More recently, in a Caucasian European cohort living in South Africa, variants in *fat mass and obesity-associated* (FTO), a gene highly predictive of complex obesity, were associated with poorer eating behaviors (higher hunger and disinhibition scores) as well as increased consumption of high fat and simple carbohydrates (Harbron et al., 2014).

THE GENETICS OF TASTE

Taste has an important role in food preference and a direct influence on eating behavior. There is great variation among individuals in the ways in which they perceive taste, and taste is at least in part genetically determined. Five tastes are described in humans: sweet, bitter, sour, salt, and umami. The receptor for sweet taste is encoded by *TAS1R2* and *TAS1R3* genes, located on human chromosome 1. These genes are responsible for producing a heterodimer of two protein subunits: T1R2 (taste receptor type 1, member 2), and T1R3 (taste receptor type 1, member 3). The T1R2/T1R3 heterodimer has been shown to detect natural sugars, such as glucose and fructose, as well as a variety of noncaloric artificial and natural sweeteners, some D-amino acids, and sweet-tasting proteins. The *TAS1R2* gene demonstrates high levels of genetic variation. Studies showed that genetic variation in *TAS1R2* and *TAS1R3* may directly influence consumption of sugar in overweight individuals (Eny et al., 2010). Eny et al. studied two common variants in *TAS1R2*, Ser9Cys and Ile191Val, and the association of these variants with sugar consumption in lean and obese individuals. Individuals homozygous for the Ile allele (Ile/Ile) consume significantly more sugars than do Val carriers. The Ile allele frequency was common among obese subjects. There were no differences in nutrient consumption among lean or obese Ser9Cys individuals (Eny et al., 2010).

Bitterness is considered a key determinant for the acceptance or rejection of food. Bitterness perception is signaled through transmembrane G protein-coupled receptors encoded by type 2 bitter taste receptor genes (*TAS2R*). There are approximately 25–30 *TAS2R* genes, which are located on chromosomes 5, 7, and 12. These receptors are expressed in multiple organs, including the brain, and throughout the gastrointestinal (GI) tract, including the pancreas. Not only are these “bitter” genes responsible for taste detection and perception, they serve to mediate protective responses to the ingestion of toxic chemicals, such as vomiting.

Among the *TAS2R* gene family, *TAS2R38* is the most highly studied. Ligands for the *TAS2R38* receptor include compounds such as phenylthiocarbamide and 6-N-propylthiouracil. Bitter-tasting compounds found in the brassica family of vegetables including Brussels sprouts, broccoli, and cabbage are also activators of this receptor. Studies examined the association between bitter-taste gene mutations and cancer and obesity. In a study conducted in a genetically isolated village in Italy, investigators explored the relationship between *TAS2R38*, body mass index (BMI), and waist circumference (WC). Results showed that females who had a specific haplotype of *TAS2R38* associated phenotypically

with low taste perception to 6-N-propylthiouacil had higher BMI and WC (Tepper et al., 2008). A group in Germany investigated the relation between common variants of this gene and colorectal cancer and found that in an analysis of the diplotypes and the phenotypes of the gene, individuals with haplotypes predicting enhanced taste sensitivity to bitter (specifically subjects harboring the Pro-Ala-Val haplotype) had a lower risk for colorectal cancer compared with the nontaster group (Carrai et al., 2011). However, Basson et al. reported a positive correlation between bitter taste sensitivity and colon polyp number, which suggests a possible relation between vegetable intake and polyp risk, i.e., that persons who are more sensitive to bitter taste eat fewer vegetables and have a higher risk for colon polyps (Basson et al., 2005). Dotson et al. showed that Amish women with the nontaster genotype had higher disinhibition eating behavior scores and higher disinhibition scores were associated with higher BMI (Dotson et al., 2010).

Studies that attempted to identify genetic determinants for salt tasting were not as robust as those related to sweet and bitter tasting. Potential salt tasting receptors thought to mediate salt perception are the sodium specific receptor and the amiloride-sensitive epithelial sodium channel (EnaC). This hypothesis was confirmed by a knockout model of an essential EnaC subunit in taste buds. The presence of the transient receptor potential cation subfamily V member 1 channel, which is believed to be involved in nonspecific salt taste perception, was also proposed as being involved in salt taste (Chaudhari and Roper, 2010; Desimone and Lyall, 2006). Studies in mice showed that genetic variations of receptor subunits can influence the difference in salt taste acuity, but studies in human have been limited.

Amino acids glutamate and aspartate and certain ribonucleic acids are perceived by taste buds as savory or umami. There is evidence that a heterodimeric G protein-coupled receptor consisting of two proteins (TAS1R1 and TAS1R3) functions as an umami receptor in rodents and humans. Splice variants of metabotropic glutamate receptors mGluR₁ (glutamate receptor, metabotropic 1, *Grm1*) and mGluR₄ (glutamate receptor, metabotropic 4, *Grm4*) are also proposed as taste receptors for glutamate. There is a substantial difference in both humans and rodents in umami taste perception, but little is known about the genetic variation of these receptors regarding the degree of umami taste sensitivity. Investigators characterized the general sensitivity of glutamate by correlating it with variations in the umami taste gene (TAS1R3). This can essentially explain why food containing amino acids such as glutamate are perceived as more tasteful or more palatable by some individuals compared with others, at least partly explaining variations in food consumption. In a study from Japan, the authors reported that an umami tasting

disorder in some elderly individuals is associated with loss of appetite, weight loss, and poor health. They were able to assess the umami taste subjectively by creating a filter paper disk method using monosodium glutamate (Sasano et al., 2015).

THE GENETICS OF MEAL SIZE AND FREQUENCY

Although meal size and frequency are largely influenced by the socioeconomic and psychological environment, underlying genetic factors have been linked to these eating patterns. De Korm et al. conducted a study to investigate variations in genes encoding cholecystokinin (CCK) and leptin and test whether those variations are linked to specific eating patterns (De Korm et al., 2007). Leptin is a hormone encoded by the *OB* gene located on chromosome 7; it is produced mainly in adipose tissues as well as the stomach, mammary epithelium, and heart. Leptin acts on the brain through a feedback mechanism to inhibit food intake and regulate energy homeostasis. CCK is a peptide with various molecular forms that is widely present through the GI tract and the central nervous system (CNS) secreted in response to fat and protein intake. CCK increases the sensation of fullness and suppresses the sensation of hunger. De Korm et al. demonstrated that obese carriers of common allelic variants in both leptin and leptin receptor genes had increased risk for extreme snacking behavior, whereas variants in the CCK gene in obese subjects were related to larger meal sizes (De Korm et al., 2007).

Ghrelin is a peptide mainly secreted by the fundus of the stomach. This peptide has a role in stimulating hunger and controlling energy homeostasis. It is considered an orexigenic hormone, levels of which increase preprandially and fall postprandially. Ghrelin directly stimulates food intake and growth hormone secretion in humans and rodents. It is normally downregulated in response to energy excess. The fall in ghrelin levels in response to meals may be blunted in some obese individuals. In human studies, mutations in *Ghrl*, the gene encoding ghrelin and preproghrelin, were identified in obese subjects. Ukkola et al. identified a mutation at amino acid position 51 (Arg51Gln) in the preproghrelin sequence present in 6.3% of obese subjects compared with none in control nonobese subjects (Ukkola et al., 2001). Furthermore, some studies showed that hyperghrelinemia might explain the hyperphagia observed in certain syndromic obesity disorders such as Prader-Willi (PW) syndrome, and that elevated ghrelin levels in those patients is consistent with increased hunger and their tendency to overeat (Cummings et al., 2002; Purtell et al., 2011).

Peptide tyrosin-tyrosin (PYY) is an anorexigenic hormone that acts in opposition to ghrelin. It is secreted by the L cells of the distal gut, suppresses energy intake, and has a role in satiety signaling. Its effect is mediated through the hypothalamus and the brain stem. Other hormonal peptides including amylin and gastric inhibitory peptide, which may have indirect effects on appetite through insulin stimulation, also have roles in regulating appetite and food intake. To our knowledge, there are no robust studies identifying genetic variants in genes encoding these peptides and their subsequent effect on human energy metabolism, or meal sizes and portions.

THE GENETICS OF APPETITE TRAITS

The burden of obesity in general, especially childhood obesity, is being increasingly recognized. It is well-established that environmental factors have a leading role in the development of obesity among individuals, but why certain individuals are more susceptible to obesity than others even when living in a common shared environment remains unanswered. Appetite and energy homeostasis are regulated largely by the hypothalamus and the brain stem (the arcuate nucleus). Variations in genes related to leptin-melanocortin signaling pathways have been associated with changes in appetite patterns and obesity-related phenotypes. Within the arcuate nucleus, leptin acts on two different classes of neurons, one of which expresses anorexigenic peptides such as proopiomelanocortin, as well as cocaine- and amphetamine-related transcript, which reduces food intake; the other coexpresses the orexigenic peptides neuropeptide Y and agouti-related protein, which increase food intake.

Another substance involved in appetite regulatory pathways in the CNS is brain-derived neurotrophic factor (BDNF). BDNF is important for brain development and promotes the survival and maintenance of nerve cells. It is mostly active at the nerve synapsis and regulates synaptic plasticity. Scientific investigation is focusing more on BDNF because it may have an important role in energy homeostasis by acting primarily on the ventromedial hypothalamus. Genetic variation in BDNF has been linked to Wilms tumor, aniridia, genitourinary anomalies, and mental retardation (WAGR) syndrome, as well as to psychiatric conditions such as bipolar disorder and anxiety. Polymorphisms of BDNF have been linked to cases of obesity and hyperphagia in children; a single nucleotide polymorphism (SNP) that is mostly reported is rs6265, which encodes for the substitution, in a valine residue at position 66 of pre-proBNF, with a methionine residue (Val66Met). In a study of 33 children with genetically confirmed WAGR syndrome, 58% had BDNF gene haploinsufficiency,

most of whom were heterozygotes with complete deletions of BDNF. Patients with deletions of BDNF had significantly higher BMI z scores than did children who had no deletions of BDNF. The authors of that article reported that affected children showed signs of hyperphagia (Han et al., 2008).

Several studies provide evidence that the obesity-associated gene *FTO* may be a primary regulator of food intake and appetite. SNPs in *FTO* have been shown in many studies to be linked to obesity. It is believed that this gene may influence human body mass by affecting appetite regulating neural circuits and energy expenditure. Although, the exact mechanism of action of the protein encoded by this gene remains unclear, in addition to obesity, phenotypes associated with *FTO* mutations include age-associated cognitive loss and alcohol dependence syndrome. A study aimed at establishing a link between *FTO* and sensitivity to satiety in children concluded that a particular SNP, rs9939609, is associated with altered food intake behaviors. Children who were homozygous for the high-risk A-allele consumed more food (Wardle et al., 2009). A more recent United Kingdom study involving adolescents enrolled in the Avon Longitudinal Study of Parents and Children investigated associations of BMI-related genes to binge eating. When examining the association of the frequency of binge eating with different genetic variants in nearly 600 adolescents, the authors reported evidence for 32 BMI-related SNPs, and particularly for the rs1558902a SNP in *FTO*, of associations with binge eating, further strengthening evidence that *FTO* influences food intake and appetite (Micali et al., 2015).

Some studies were conducted to determine whether genetic variations in hypothalamic signaling pathways were related to early obesity in children. Wang et al. published a study in which variants in 8 obesity susceptibility genes (*BDNF*, *LEPR*, *FTO*, *PCSK1*, *POMC*, *TUB*, *LEP*, and *MC4R*) were genotyped in 128 children. A genetic predisposition score was assigned and their association with BMI was assessed. The authors found that as the number of risk alleles increased, the risk for higher BMI also increased (Wang et al., 2013). A study in Norway examined the relation of genetic risk score with BMI in children aged 4–8 years, as well as their appetite traits. Appetite traits were assessed via parent-completed questionnaires, exploring the child's eating behaviors; parents answered questions about whether their child enjoyed food, how slow the child ate meals, whether the child looked forward to mealtimes, and whether the child exhibited emotional eating. The study investigators found that children with higher genetic risk scores had higher BMIs and a higher percentage of body fat compared with children with lower risk scores. Furthermore, they concluded that children with higher

genetic risk scores demonstrated lower levels of slowness of eating, indicating higher eating rates; however genetic associations with other appetite modifiers such as food responsiveness, emotional overeating, enjoyment of food, and satiety responsiveness were not statistically significant (Steinsbek et al., 2016).

The mesolimbic pathway, which is dopaminergic, is involved in pleasure-seeking behaviors and controls responses to natural rewards, such as food, certain drugs, and social interactions.

There have been extensive studies looking into the role of the brain dopamine system in influencing food-seeking behaviors. Researchers in a United Kingdom study of 26 overweight or obese humans showed that blocking D3 receptors influenced behavioral responses toward food cues differently compared with to a placebo group (Mogg et al., 2012). Studies have linked drug addiction to polymorphisms in the dopamine D2 receptor gene (*DRD2*), particularly the TaqIA1 and TaqIB1 alleles (Smith et al., 1992). Carpenter et al. demonstrated the possibility of an association with food addiction and craving with these alleles. They noted that mean BMI was significantly higher in individuals who carried more copies of the TaqIA1 allele (Carpenter et al., 2013). Having more copies of the TaqI A1 allele translates into having a lower density of dopamine D2 receptors (*DRD2*) and led to the concept of an insensitive reward system, which contributes to overconsumption of highly palatable food, and which then stimulates the reward system in a more robust manner to compensate for the lack of *DD2R* stimulation, in turn leading to obesity.

Eny et al. investigated another *DRD2* polymorphism, C957T, and demonstrated an association with habitual sugar consumption. C957T is located in exon 7 of the *DRD2* gene and was reported to be linked to the TaqI A1 allele in some populations. In this study, C957C homozygote men consumed more sugar compared with T957T homozygote men. In addition, women who carried the homozygote variant (CC) consumed more sugar than did women who carried the heterozygote variant (CT). These results provide evidence that variations in *DRD2* may contribute to individual differences in sugar consumption (Eny et al., 2009). With respect to impaired dopamine pathway signaling, Zhang et al. reported a mutation in the *Slc35d3* gene linked to reduced membrane dopamine receptor D1 (*D1R*) and impaired dopamine signaling in striatal neurons, leading to metabolic syndrome and obesity in mice. Interestingly, the investigators found that the development of obesity in mutant mice was more related to reduced energy expenditure and physical inactivity than to hyperphagia (Zhang et al., 2014).

Well-described obesity and hyperphagia syndromes include Bardet–Biedl, PW, and Alström syndromes. Genes encoding ciliary proteins have been implicated in the Bardet–Biedl syndrome; more than 18 distinct genetic loci were identified as being involved. Genetic defects in PW and Alström syndromes have been localized to chromosome 15, whereas some cases of PW-like syndromes have been attributed to defects on chromosomes 1, 5, and 6.

Classical eating disorders include anorexia nervosa, bulimia nervosa, and binge eating. These disorders tend to aggregate in families. Twin studies indicated that the genetic influence accounts for nearly 40%–60% of variability noted in these traits. Although genes are clearly associated with the susceptibility to develop eating disorders, gene–environment interaction has a critical role in the phenotypes observed.

Clearly, genetic factors influence eating behavior and appetite traits. More studies are needed to determine how genetic susceptibility affects lifestyle behaviors related to variations in body weight, eating behavior, food choice, and behavior change. Although genetic predisposition to eating behaviors and obesity susceptibility has been identified, the question remains as to how plastic these traits might be in settings of intentional behavior change. Some insight into this question might be gained from data from the Diabetes Prevention Program, a randomized prospective trial designed to determine whether increased activity and prudent diet are effective in bringing about weight reduction and reduced progression to type 2 diabetes (T2D). This landmark study demonstrated a significant reduction in the risk for T2D among those assigned to the intensive lifestyle intervention compared with those receiving usual care or treatment with metformin (The Diabetes Prevention Program Research Group, 2002). In a posthoc analysis, several T2D susceptibility genes were genotyped in the study cohort. Even among subjects with the highest T2D genetic risk, individuals assigned to the intensive lifestyle arm significantly reduced their risk for progressing to T2D. Data from the Diabetes Prevention Study suggest that among persons with genetic susceptibility to T2D, engaging in activity and prudent nutrition resulted in weight loss and reduced diabetes despite elevated genetic risk. Additional studies are needed to determine whether eating behaviors are modifiable in individuals genetically at risk. Might these same principles be observed among persons with at-risk eating behavior trait alleles? Would individuals with genetic eating behavior risk factors be able to modify eating behavior and body weight intentionally when behavior strategies are applied to bring about intentional behavior change?

NEW GENOMIC APPROACHES ON THE HORIZON: THE MICROBIOME RELATION TO FOOD INTAKE AND OBESITY

There has been increased interest in the relation between the human gut microbiota and diseases, including obesity, diabetes, and the metabolic syndrome. The human gut is the host to nearly 100 trillion microorganisms, encompassing about 1000 species of bacteria and over a million genes (see also the dedicated chapters in this textbook). The gut microbiome–obesity association first came to attention when OB mice (obese mice with a nonfunctioning leptin receptor gene) were found to have a different gut microbial composition compared with mice without the leptin mutation.

It has been suggested that intestinal microflora can modulate energy harvesting by producing short-chain fatty acids from indigestible carbohydrates, which include acetate, butyrate, and propionate; these can stimulate gut hormonal mechanisms such as GLP-1 and PYY, which delay gut motility and promote satiety. It is evident that the gut microbiome differs between obese and lean individuals. This theory was reinforced by findings from studies of fecal microbiome transplants from obese to lean mice, which demonstrated increased adiposity in lean mice that received fecal transplants from obese mice despite neutral energy intake. Human studies characterizing fecal microbial communities in adult obese and lean subjects, including those involving monozygotic and dizygotic twins, found that in general, the gut microbiome composition is largely shared among family members. Although each individual has a unique microbial composition, variations are comparable between twins. In most studies, obese subjects have reduced microbial diversity. Reduced gut microbial diversity and its association with host metabolic health are a well-accepted concept. Most studies suggest that obesity is accompanied by an increased abundance of Firmicutes levels and a reduced abundance of Bacteroidetes. However, this finding is not uniformly observed across all studies. Furthermore, studies in mice suggest that deliberate manipulation of the gut microbiome through prebiotics influences gut hormones, feeding behavior, carbohydrate metabolism, and body composition.

Novel approaches to obesity therefore include those that deliberately attempt to manipulate the microbiome therapeutically to prevent, mitigate, or cure obesity and related metabolic disorders, including glucose intolerance and nonalcoholic fatty liver disease. The gut microbiome thus represents a novel target to develop new strategies to affect human eating behavior and health. Both prebiotics and probiotics have the potential to modify the composition and function of the human gut microbiome significantly and affect human health.

Ongoing studies are needed to identify better how the microbiome can be manipulated and which individuals will achieve the most robust and enduring response.

Despite scientific progress in nutrigenetics and nutrigenomics and the degree of our current understanding of genetic influences in eating behaviors obesity and diabetes, the ultimate challenge lies in developing methods that can be practically applied to improve eating behaviors and prevent the continued rise in obesity and diabetes prevalence worldwide.

References

- Basson, M.D., Bartoshuk, L.M., Dichello, S.Z., Panzini, L., Weiffenback, J.M., Duffy, V.B., 2005. Association between 6-n-propylthiouracil (PROP) bitterness and colonic neoplasms. *Dig Dis Sci* 50 (3), 483–489.
- Bouchard, L., Drapeau, V., Provencher, V., Lemieux, S., Chagnon, Y., Rice, T., Rao, D.C., Vohl, M.C., Tremblay, A., Bouchard, C., Pérusse, L., 2004. Neuromedin beta: a strong candidate gene linking eating behaviors and susceptibility to obesity. *Am J Clin Nutr* 80 (6), 1478–1486.
- Carpenter, C.L., Wong, A.M., Li, Z., Noble, E.P., Heber, D., 2013. Association of dopamine D2 receptor and leptin receptor genes with clinically severe obesity. *Obesity* 21 (9), E467–E473.
- Carrai, M., Steinke, V., Vodicka, P., Pardini, B., Rahner, N., Holinski-Feder, E., et al., 2011. Association between TAS2R38 gene polymorphisms and colorectal cancer risk: a case-control study in two independent populations of Caucasian origin. *PLoS One* 6 (6), E20464.
- Chaudhari, N., Roper, S.D., 2010. The cell Biology of taste. *J Cell Biol* 190 (3), 285–269.
- Choquette, A.C., Lemieux, S., Tremblay, A., Drapeau, V., Bouchard, C., Vohl, M.C., Perusse, L., 2009. GAD2 gene sequence variations are associated with eating behaviors and weight gain in women from the Quebec family study. *Physiol Behav* 98 (4), 505–510.
- Cummings, D.E., Clement, K., Purnell, J.Q., Vaisse, C., Foster, K.E., Frayo, S., et al., 2002. Elevated plasma ghrelin levels in prader-willi Syndrome. *Nat Med* 8, 643–644.
- De Korm, M., Vander Schouw, Y.T., Hendriks, J., Ophoff, R.A., Van Gils, C.H., Stolk, R.P., et al., 2007. Common genetic variants in CCK, Leptin and Leptin Receptor genes are associated with specific human eating patterns. *Diabetes* 56, 276–280.
- Desimone, J.A., Lyall, V., 2006. Taste receptors in the gastrointestinal tract III. Salty and sour taste: sensing of sodium and protons by the tongue. *Am J Physiol Gastrointest Liver Physiol* 291 (6), 61005–61010.
- Dotson, C.D., Shawn, H.L., Mitchell, B.D., Munger, S.D., Steinle, N.I., 2010. Variation in the gene TAS2R38 is associated with the eating behavior disinhibition in Old Order Amish women. *Appetite* 54 (1), 93–99.
- Eny, K.M., Corey, P.N., El-Sohemy, A., 2009. Dopamine D2 receptor genotype (C957T) and habitual consumption of sugars in a free-living population of men and women. *J Nutrigenetics Nutrigenomics* 2, 235–242.
- Eny, K., Wolever, T.M., Corey, P., El-Sohemy, 2010. Genetic variations in TAS1R2 (Ile 191 Val) is associated with consumption of sugars in overweight and obese individuals in 2 distinct populations. *Am J Clin Nutr* 92, 1501–1510.
- Gast, M.T., Tönjes, A., Keller, M., Horstmann, A., Steinle, N., Scholz, M., Müller, I., Villringer, A., Stumvoll, M., Kovacs, P., Böttcher, Y., 2013.

- The role of rs2237781 within GRM8 in eating behavior. *Brain Behav* 3 (5), 495–502. <https://doi.org/10.1002/brb3.151>. Epub 2013 Jun 23.
- Han, J., Liu, Q., Hones, M., Levinn, R., Menzie, C.M., Jefferson-George, K.S., et al., 2008. Brain-derived neurotrophic factor and obesity in WAGR syndrome. *N Engl J Med* 359, 918–927.
- Harbron, J., van der Merwe, L., Zaahl, M.G., Kotze, M.J., Senekal, M., 2014. Fat mass and obesity-associated (FTO) gene polymorphisms are associated with physical activity, food intake, eating behaviors, psychological health, and modeled change in body mass index in overweight/obese Caucasian adults. *Nutrients* 6 (8), 3130–3152. <https://doi.org/10.3390/nu6083130>.
- Micali, N., Field, A.E., Treasure, J.L., Evans, D.M., 2015. Are obesity risk genes associated with binge eating in adolescence? *Obesity* 23 (8), 1729–1736.
- Mogg, K., Bradley, B., O'Neill, B., Bani, M., Merlo-Pich, E., Koch, A., et al., 2012. Effects of dopamine D3 receptor antagonism on approach responses to food cues in overweight and obese individuals. *Behav Pharmacol* 23 (5–6), 603–608.
- Neale, B.M., Mazzeo, S.E., Bulik, C.M., 2003. A twin study of dietary restraint disinhibition and hunger. *Twin Res* 6 (6), 471–478.
- Purtell, L., Sze, L., Loughnan, G., Smith, E., Herzog, H., Sainsbury, A., et al., 2011. In adults with Prader-Willi syndrome, elevated ghrelin levels are more consistent with hyperphagia than high PYY and GLP-1 levels. *Neuropeptides* 45 (4), 301–307.
- Sasano, T., Satoh-Kuriwada, S., Shoji, N., 2015. The importance of umami taste in oral and overall health. *Flavour* 410.
- Smith, S., O'hara, B.F., Persico, A.M., Grelick, D.A., Newlin, D.B., Vlahov, D., et al., 1992. Genetic vulnerability to drug abuse. The D2 dopamine receptor Taq IB1 restriction fragment length polymorphism appears more frequently in polysubstance abusers. *Arch Gen Psychiatr* 49 (9), 723–727.
- Steinle, N., Hsueh, W., Snitker, S., Pollin, T., Sakul, T., St jean, P., et al., 2002. Eating behavior in the old order Amish: heritability analysis and genome-wide linkage analysis. *Am J Clin Nutr* 75, 1098–1106.
- Steinsbek, S., Belsky, D., Guzey, I.C., Wardle, J., Wichstrom, L., 2016. Polygenic risk, appetite traits and weight gain in middle childhood: a longitudinal study. *JAMA Pediatrics* 170 (2), e154472.
- Tepper, B.J., Koelliker, Y., Zhao, L., Ullrich, N., Lanzara, C., D'Adamo, P., et al., 2008. Variation in the bitter taste receptor gene TAS2R38 and adiposity in a genetically isolated population in southern Italy. *Obesity* 16 (10), 2289–2295.
- The Diabetes Prevention Program Research Group, 2002. Reduction in incidence of type II diabetes with life style intervention or metformin group. *N Engl J Med* 346 (6), 393–403.
- Ukkola, A., Ravussin, E., Jacobson, P., Snyder, E., Chagonon, M., Sjostrom, L., Bouchard, C., 2001. Mutations in the preproghrelin/ghrelin Gen associated with obesity in humans. *J Clin Endocrinol* 86 (8), 3996–3999.
- Wang, Y., Wang, A., Donovan, S.M., Terran- Garcia, M., 2013. STRONG kinds research team. *Hum Hered* 75 (2–4), 152–159.
- Wardle, J., Llewellyn, C., Sanderson, S., Plomin, R., 2009. The FTO gene and measured food intake in children. *Int J Obes* 33 (1), 45–5.
- Zhang, Z., Hao, C., Li, C., Zang, D., Zhao, J., Li, X., et al., 2014. Mutation of SLC35D3 causes metabolic syndrome by impairing dopamine signaling in striatal D1 neurons. *PLoS Genet* 10 (2), e1004124.

This page intentionally left blank

Genetics of Body Composition: From Severe Obesity to Extreme Leanness

José L. Santos, Víctor Cortés

Department of Nutrition, Diabetes, and Metabolism, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile

Glossary

Body mass index Defined as weight (in kilograms) divided by height squared (in meters). It is a measure of excess weight standardized by height and is highly correlated with adiposity. It is used to define obesity using the cutoff of 30 kg/m².

Congenital generalized lipodystrophy (CGL) Rare disease showing the most extreme form of adipose tissue deficiency in humans, associated with severe ectopic fat accumulation in liver and insulin resistance, among other clinical characteristics. The most studied type of CGL is caused by mutation in the AGPAT2 gene.

Genetic risk scores Risk scores composed by integrating the weighted effects of single nucleotide variants that define subjects with higher genetic susceptibility for a given disease

Genome-wide association study Hypothesis-free study in which simultaneous measurements of millions of simple genetic variants are used to evaluate gene–trait or gene–disease associations.

Heritability in narrow sense Percentage of the total phenotypic variation that is attributed to additive genetic effects.

Missing heritability The gap between relatively low percentages of the complex trait variation explained by common single nucleotide variants calculated from association studies, compared with the high heritability estimated from twin, family, and population studies.

Monogenic obesity A group of rare types of obesity caused by genetic mutations in a single gene that exerts a profound effect on adiposity. Most known causes of monogenic obesity are related to mutations in genes of the leptin–melanocortin pathway, such as *LEP*, *LEPR*, *MC4R*, and *POMC*, among others.

Multifactorial obesity Common disease characterized by increased adiposity resulting from the complex interaction among genetic, metabolic, physiological and environmental, familial, and societal factors in a broad sense (including diverse influences on dietary and physical activity patterns).

Single nucleotide variant (SNV) The most common type of genetic variation derived from the substitution of a single nucleotide, defining three genotypes. Can be either rare or common in the population.

Syndromic obesity Group of diseases in which obesity is accompanied by other additional clinical features, frequently including cognitive disability, dysmorphias, and congenital endocrinopathies affecting multiple organ systems.

INTRODUCTION

Obesity is a high-frequency disease with multifactorial etiology implying multiple complex interactions between genetics and the environment. The increased prevalence of overweight and obesity is the result of changes that have occurred in modern societies (increased urbanization; rapidly increase in available, cheap, high-caloric foods; and sedentary lifestyles) that operate with a special impact on genetically susceptible subjects. In this chapter, we use body mass index (BMI) (in kg/m²) as a surrogate variable for whole-body adiposity, in which obesity is defined as a BMI of 30 kg/m² or greater.

The study of human common genetic variations in relation to BMI, waist-to-hip ratio (WHR), or percent body fat has been carried out mainly through genome-wide association studies (GWAS). These studies revealed more than 500 genetic loci associated with small individual effects at the population level (Loos, 2018). In addition, monogenic obesity is caused by rare genetic mutations exerting a profound effect on body fat accumulation. Most cases with genetic obesity of known origin are caused by the disruption of genes belonging to the leptin–melanocortin pathway (Huvenne et al., 2016). In contrast to obesity, the study of human leanness represents a complementary approach to ascertaining genes related to body weight regulation. The genetic assessment of human thinness is composed of constitutional leanness, psychiatric disorders such as anorexia nervosa (AN), and extreme leanness phenotypes such as congenital or acquired lipodystrophy syndromes. In some cases, obesity and leanness may behave as mirror phenotypes caused by different genetic alterations in the same gene or group

of genes, which trigger contrasted effects. The study of obesity and leanness as opposite phenotypes represents a fruitful approach to assessing causal gene variations related to obesity and to understanding the physiological roles of body fat accumulation.

COMMON GENE VARIANTS IN RELATION TO ADIPOSITY TRAITS IN THE POPULATION

Different heritability estimates for BMI have been calculated from twin-based studies (60%–75%), family-based studies (40%–45%), and population-based studies (20%–40%). Likewise, corresponding heritability estimates for WHR are 30%–60%, 20%–50%, and around 10%, respectively (Loos, 2018). Fig 21.1 shows simulated data on the approximate correlations found in studies of monozygotic (MZ) and dizygotic (DZ) twins in relation to BMI. In this idealised example, the excess in the intraclass correlation coefficient in MZ compared to DZ twins suggests the importance of the genetic component in BMI. Through many GWAS, common variants of multiple genes have been significantly associated with multifactorial metabolic traits and diseases that show phenotype variation at the population level, such as serum lipids, glycemic-related traits, or

BMI. Common single nucleotide variants (SNVs) in more than 500 loci have been significantly associated with adiposity traits, mainly BMI (341 loci) and BMI-adjusted WHR (129 loci), with an additional number of loci identified for percent body fat, lean mass, weight loss, and visceral adiposity (Loos, 2018). Bioinformatic analyses revealed that SNVs significantly related to BMI are preferentially expressed in the central nervous system (CNS), presumably affecting appetite and energy intake phenotypes, whereas genes involved in body fat distribution are preferentially involved in adipocyte biology. Among common genetic variation, affecting BMI, the strongest significant signal is represented by intronic variants in the *FTO* gene (rs9939609, rs1421085, and others). Specifically, common variants in this region show approximately 1 kg of increased body weight per risk-allele, assuming a 1.7-m-tall person (Loos, 2018). Families with loss-of-function mutations of *FTO* gene highlighted the importance of this gene in human biology, presenting a syndrome characterized by severe polymalformations and growth retardation, but without an apparent effect on body fat mass. This finding is concordant with the hypothesis stating that common genetic variation of *FTO* protein is not directly associated with obesity but may act as a regulatory region that controls the expression of other genes in its vicinity related to energy homeostasis

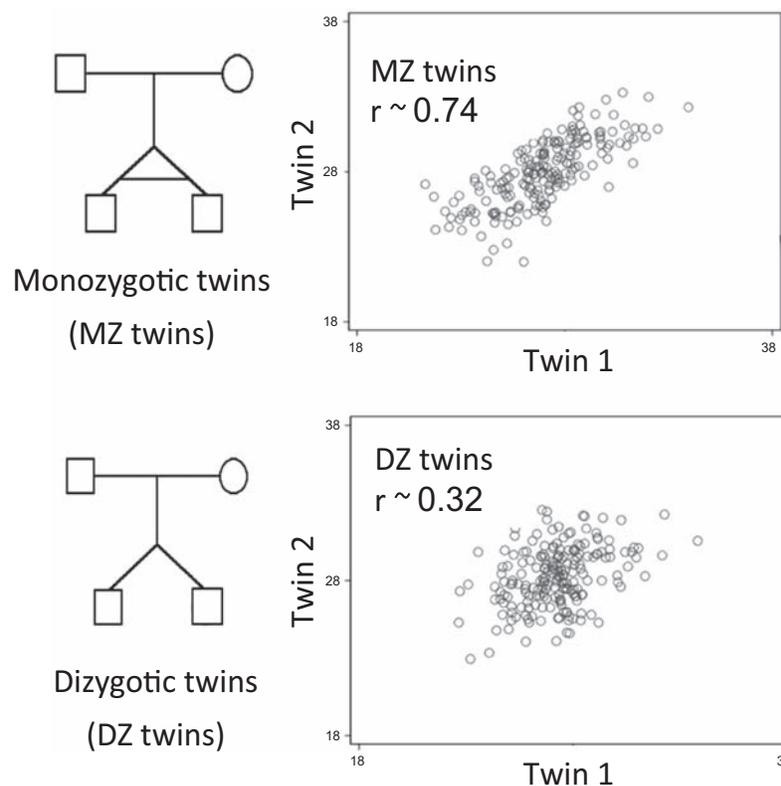


FIGURE 21.1 Simulated data representing BMI correlations in monozygotic and dizygotic twins.

(*IRX3*, *IRX5*, *RPGRIPL*, and *IRXB*). In addition, common variation near *MC4R* gene was reported as a reproducible association signal for BMI in many GWAS. This result was a remarkable finding considering that rare mutations in *MC4R* gene are a known cause of severe obesity and hyperphagia, as we will see later in this chapter. The common polymorphism rs17782313, among others in the same region mapped 188 kb downstream from *MC4R*, have been significantly associated with increased BMI and altered eating behavior in both children and adults. Using measurements obtained from hundreds of thousands of individuals in large epidemiologic studies, variants with genome-wide significance for BMI were also found in or near *TMEM18*, *FAIM2*, *GNPDA2*, *SEC16B*, *TFAP2B*, brain-derived neurotrophic factor (*BDNF*), *NEGR1*, *BCDIN3D*, adenylate cyclase type-3 (*ADCY3*), *ATP2A1*, and *ETV5* genes, among others. Most genes from GWAS were derived from European and Asian population (Akiyama et al., 2017), with low contributions of populations from other parts of the world (Loos, 2018).

Despite great efforts displayed in conducting and analyzing GWAS, the magnitude of genetic effects measured as odds ratios for obesity or per-allele increase in BMI is low for significant common SNVs. In addition, the percentage of variance of BMI explained by common SNVs represents only a small fraction of total phenotypic variation, estimated as less than 4% for SNVs associated with BMI and 1.4%–2.7% for BMI-adjusted WHR (Loos, 2018). In this context, the low degree of involvement of common SNVs in explaining BMI has led to the concept of missing heritability, defined as the gap between the relatively low percentages of phenotypic variation explained by common SNVs calculated from GWAS compared with the high heritability estimated from twin, family, and population studies (ranging from 30% to 70%,

depending on the study). Research showed that the complete set of SNVs (not only those that achieved genome-wide significance) may explain up to 27%–29% for BMI (Akiyama et al., 2017). Taken together, these figures indicate that the missing heritability gap for BMI is small compared with the magnitude initially proposed. An alternative explanation for the missing heritability is the potential involvement of rare SNVs (minor allele frequency [MAF] <1%) and low-frequency variants (MAF of 1%–5%) on BMI. A study based on more than 700,000 participants found low-frequency variants with a large effect in BMI (*ZBTB7B*, *ACHE*, *RAPGEF3*, *RAB21*, *ZFH3*, *ENTPD6*, *ZFR2*, *ZNF169*, kinase suppressor of Ras2 [*KSR2*], and *GIPR*) also able to detect a rare loss-of-function *MC4R* variant (p.Tyr35Ter) previously linked to monogenic obesity (Turcot et al., 2018). Fig. 21.2 shows the effect on weight (in kilograms) per minor allele (assuming a 1.7-m-tall person) for several rare and common variants detected in GWAS.

It has been proposed that it is necessary to refine phenotypes involved in future GWAS with a focus on in-depth body composition assessment (Müller et al., 2018). In contrast, other authors indicated that larger sample sizes focused on simple measures such as BMI are a more useful strategy to find new genes related to adiposity traits (Speakman et al., 2018). In most GWAS studies, common BMI-associated SNVs are typically combined to create genetic risk scores (GRSs) that define subjects with higher genetic risks for increased body weight. Although GRSs show a strong significant association with BMI, such scores have been shown to be insufficient for predicting future obesity compared with models based on variables such as family history (Loos, 2018). Regardless of the low capacity for predicting obesity conveyed by significant SNVs from GWAS, the important contribution of GWAS in gene discovery

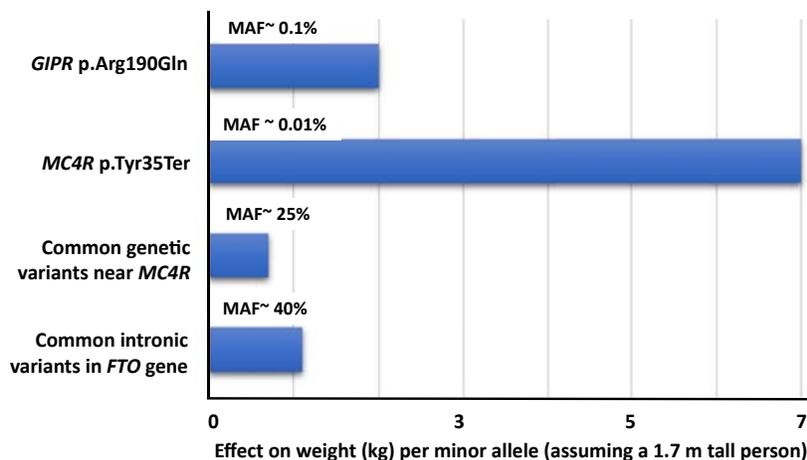


FIGURE 21.2 Effect of selected common and low-frequency genetic variants on body weight. *FTO*, fat mass and obesity-associated; *GIPR*, gastric inhibitory polypeptide receptor; MAF, minor allele frequency; *MC4R*, melanocortin-4 receptor. Figure based on data of Loos, R.J., 2018. *The genetics of adiposity*. *Curr Opin Genet Dev* 50, 86–95.

provides novel knowledge about disease pathophysiology and potentially leads to new biological targets for prevention and treatment.

MONOGENIC AND SYNDROMIC OBESITY

Most cases of rare types of monogenic obesities of known origin are related to the disruption of genes belonging to the leptin–melanocortin pathway (Huv-[enne et al., 2016](#)). The hypothalamic leptin–melanocortin system has a crucial role in eating behavior and energy homeostasis. In brief, this system is composed of (1) the leptin, leptin-receptor isoforms, and downstream signaling system; (2) the proopiomelanocortin (*POMC*), enzymes that process *POMC* (*PCSK1*, among others), and *POMC*-derived peptides (notably, the α -melanocyte stimulating hormones α - and β -melanocyte-stimulating hormone (α -MSH and β -MSH), and the adrenocorticotrophic hormone [ACTH]); (3) five melanocortin G protein–coupled receptors (*MC1R–MC5R*); (4) peptides that antagonize the effect of *POMC*-derived melanocortins on melanocortin receptors such as agouti-

related protein (AGRP); and (5) proteins in distant neurons with actions modulated by melanocortin receptors in downstream signaling. Many studies in humans and mice have delineated the important roles of the leptin-melanocortin system in the control of food intake and whole-body energy homeostasis.

[Table 21.1](#) shows a restricted list of five monogenic obesity types directly caused by mutations in genes of the leptin-melanocortin system: leptin (*LEP*), leptin receptor (*LEPR*), *MC4R* (melanocortin-4 receptor), *POMC* (proopiomelanocortin), and proprotein convertase subtilisin/kexin type 1, encoding proprotein convertase 1/3, also called PC1/3 (*PCSK1*). Other genes with mutations implicated in monogenic obesity (some of them are also part of downstream effectors of the leptin-melanocortin system) are Single-minded 1, *BDNF*, neurotrophic tyrosine kinase receptor type 2, the melanocortin receptor accessory protein 2, SH2B adaptor protein-1 (*SH2B1*), *KSR2*, and *ADCY3*, among others. All of them show a variable degree of syndromic clinical manifestations apart from obesity.

Leptin is a hormone mainly produced by white adipocytes that circulate in plasma with levels that are

TABLE 21.1 Clinical findings in monogenic obesity caused by mutations in selected genes of the leptin-melanocortin system.

	Genetic cause	Clinical findings
Leptin deficiency	Encoding leptin OMIM *164160 Autosomal recessive	<ul style="list-style-type: none"> – Severe early onset obesity – Extreme hyperphagia – Hyperinsulinemia – Hypothalamic hypothyroidism – Hypogonadotropic hypogonadism – Alteration of immune function
Leptin receptor deficiency	Leptin receptor OMIM *601007 Autosomal recessive	<ul style="list-style-type: none"> – Severe early-onset obesity – Hyperphagia – Delayed puberty – Reduced insulin-like growth factor-1 levels – Growth abnormalities – Alteration of immune function
Proopiomelanocortin deficiency	Proopiomelanocortin OMIM *176830 Autosomal recessive	<ul style="list-style-type: none"> – Severe early-onset obesity – Hyperphagia – Adrenal insufficiency – Hypopigmentation with pale skin and red hair
Prohormone convertase 1/3 deficiency	Proprotein convertase subtilisin/kexin type 1, encoding proprotein convertase 1/3 OMIM *162150 Autosomal recessive	<ul style="list-style-type: none"> – Severe early-onset obesity – Adrenal, gonadotropic, somatotrophic, and thyrotrophic insufficiency – Reduced plasma insulin and elevated proinsulin – Postprandial hypoglycemic malaises – Severe malabsorptive neonatal diarrhea – Central diabetes insipidus
Melanocortin-4 receptor deficiency	Melanocortin-4 receptor OMIM *155541 Autosomal dominant, recessive	<ul style="list-style-type: none"> – Severe early-onset obesity – Hyperphagia – High plasma insulin levels – Low systolic and diastolic blood pressure – Increased bone mineral density and lean mass – Accelerated linear growth in early childhood

proportional to body fat, showing a circadian rhythm and with higher plasma concentrations in women than in men. Leptin represents an adequate signal of body fat energy stores because it can cross the blood–brain barrier and interacts with leptin receptors that are widely expressed in the CNS. In the hypothalamic arcuate nucleus, leptin acts as a signal of starvation in such a way that a severe reduction of body fat (as occurs in generalized lipodystrophy) or the genetic deficiency of the leptin gene itself (as occurs in monogenic obesity caused by congenital leptin deficiency) triggers a starvation response characterized by hyperphagia. In 1999, researchers described how subcutaneous leptin administration to obese patients affected with mutations in the leptin gene had a dramatic reduction in energy intake and body weight. After this great success of personalized medicine, there were enormous expectations regarding the use of leptin as an antiobesity agent. However, it became rapidly apparent that patients with common multifactorial obesity scarcely responded to the exogenous administration of leptin in terms of body weight reduction and decreased energy intake. In addition, obese patients with genetic deficiency in the leptin receptor were described. As expected, they showed no response to exogenous leptin.

The most frequent cause among rare severe monogenic obesity types in humans corresponds to mutations in the *MC4R* gene, which may be responsible for up to 5% of severe childhood obesity and approximately to 1% of adult cases of obesity. The *MC4R* protein is a seven-transmembrane, G protein–linked receptor involved in central regulation of energy homeostasis. *MC4R* is widely expressed in the hypothalamus, brain stem, and other brain regions, where it modulates energy intake by a variety of mechanisms including mediating the anorectic response to the adipocyte-derived hormone leptin. *MC4R* is also expressed in dopamine-rich regions of the striatum, and some studies in rodents suggest that melanocortin signaling also modulates food reward. Interestingly, the amount of food consumed ad libitum in tests was related to the degree of dysfunction generated by *MC4R* mutations; an increased amount of food was eaten in patients harboring mutations with a clear effect on receptor functionality (ligand binding capacity and/or intracellular signaling response). Apart from the well-known roles of *MC4R* in regulating food intake, this receptor regulates blood pressure, lipolysis in adipose tissue, and insulin secretion through reciprocal actions on the two divisions of the autonomic nervous system (sympathetic and parasympathetic).

POMC deficiency is another type of monogenic obesity caused by defects in the leptin-melanocortin system. *POMC* is a proprotein that undergoes extensive posttranslational modification to generate a range of smaller peptides such as melanocortins α -MSH and

β -MSH produced in the hypothalamus and skin and ACTH produced in the hypophysis. Affected patients show severe early-onset obesity and hyperphagia owing to the lack of α -MSH that transmits the anorexic effect of leptin through *MC4R*. In addition, these patients have adrenal insufficiency because of the lack of ACTH that promotes steroidogenesis in the adrenal cortex through *MC2R*, leading to low levels of circulating cortisol. Clinical features of these patients include hypoglycemia, hyperbilirubinemia, and hepatic cholestasis. Finally, patients with *POMC* deficiency also have hypopigmentation with pale skin and red hair, caused by the lack of stimulation of *MC1R* by melanocortin peptides in the skin.

The use of *MC4R* agonists has been proposed as a possible therapeutic strategy for treating obesity. However, most reported *MC4R* agonists have the undesirable property of increasing blood pressure, with the notable exception of the *MC4R* activator called RM-493 (setmelanotide). Administration of this agonist to patients with common obesity led to increased energy expenditure and significant reductions in body weight (1 kg per week for 4 weeks). In patients with *POMC* mutations, adrenal insufficiency was successfully treated with hydrocortisone, but no therapy was available for hyperphagia and obesity. However, a study reported a dramatic weight reduction in two patients with mutations in *POMC* after treatment with the *MC4R* agonist setmelanotide, which simultaneously decreased body weight and feelings of hunger (Yeo, 2017).

Syndromic obesities represent a group of diseases in which obesity is accompanied by other clinical features frequently including intellectual disability, dysmorphias, and congenital endocrinopathies affecting multiple organ systems. Frequently mentioned obesity syndromes are Albright hereditary osteodystrophy (caused by mutations in the *GNAS* gene), Alström syndrome (*ALSM1* gene), Bardet–Biedl syndrome (loci *BBS1-19*), and Prader–Willi syndrome (*PWS*). Among them, *PWS* is the most prevalent, with an estimated frequency of 10,000 to 30,000 live newborns. This disease results from the absence of expression of imprinted paternally inherited genes in the 15q11.2-q13 region owing to paternal deletion (65%–75% of *PWS* patients), maternal disomy, or imprinting defects. Clinical manifestations of *PWS* are severe neonatal hypotonia, intellectual disability, behavior difficulties, short stature, growth hormone deficiency, dysmorphic features, and hypogonadism. Induced pluripotent stem cells derived from neurons of *PWS* patients showed decreased levels of PC1/3 (encoded by the *PCSK1* gene), supporting a link between *PWS* and *PCSK1*. The wide range of clinical manifestations and endocrinopathies observed in *PCSK1* (PC1/3) deficiency (Table 21.1) is derived from the participation of this enzyme in the cleavage of

multiple prohormones such as *POMC*, proinsulin, proglucagon, and proghrelin. PWS and PC1/3 deficiency have partially overlapped clinical features in subsets of cases, including linear growth abnormalities and hypogonadotropic hypogonadism.

GENOMIC STUDIES FOCUSED ON LEANNESS PHENOTYPES

Population-based studies of BMI based on the GWAS paradigm assess not only genes related to obesity but also the entire range of body weight from leanness to obesity. Regarding the left part of the BMI distribution, common variation in genes such as *AGRP* or *GPR74* have also been related to healthy constitutional thinness, whereas congenital genetic deficiency in enteropeptidase has been proposed to be a rare disease characterized by a lean phenotype. On the other hand, large genome-wide studies in AN, a psychiatric disorder that is etiologically and phenotypically different from constitutional thinness, has revealed several risk loci

that implicates a possible metabo-psychiatric origin of this disease.

Obesity and leanness are not just opposite phenotypes; they may behave as mirror phenotypes caused by different genetic alterations in the same gene or group of genes. An interesting example of this phenomenon is provided by structural variations of chromosome 16p11.2, in which large deletions in this region containing the *SH2B1* gene (involved in leptin and insulin signaling) are strongly associated with obesity, whereas duplications are strongly associated with leanness. Smith–Magenis syndrome is a rare dominant neurobehavioral disorder of obesity caused by mutations in the *RAI1* gene located on 17p11.2. In contrast, duplication of genomic segments in 17p11.2 causes Potocki–Lupski syndrome, which is characterized by leanness and overexpression of the *RAI1* gene. Another example of mirror phenotypes in BMI is provided by the gain-of-function variants I251L and V103I in the *MC4R* gene, which have been shown to be protective against obesity, in opposition to the loss-of-function mutations related to monogenic obesity.

TABLE 21.2 Classification, causes, and main clinical findings of primary lipodystrophy syndromes.

	Genetic cause	Clinical findings
1. Congenital generalized lipodystrophy		
Berardinelli–Seip syndrome	<i>AGPAT2</i> , <i>BSCL2</i> , <i>PTRF</i> and <i>CAVI</i> OMIM *603100, * 606158, *603198, and *601047	– Generalized lipoatrophy – Hyperlipidemia – Hyperinsulinemia – Diabetes mellitus – Hypoleptinemia
2. Familial partial lipodystrophy		
Type 1 (Köbberling syndrome)	Unknown	Deficiency of subcutaneous fat in limbs and gluteal regions
Type 2 (Dunningan syndrome)	Heterozygous mutations in <i>LMNA</i> OMIM *150330	– Reduced subcutaneous fat in limbs – Lipohypertrophy in trunk, neck, supraclavicular fossae, and face – Diabetes mellitus and hypertriglyceridemia
Types 3, 4, and 5	Heterozygous mutations in <i>PPARG</i> , <i>AKT2</i> , and <i>PLIN</i> OMIM *601487, *164731, and *170290	– Diabetes and dyslipidemia
3. Acquired lipodystrophy		
Lawrence syndrome	Unknown. Associated with panniculitis in 25% of patients	– Generalized lipoatrophy – Insulin resistance – Diabetes mellitus – Dyslipidemia – Hepatic steatosis
Barraquer–Simons syndrome	Unknown. Associated with autoimmune processes	– Partial lipoatrophy in upper half of body – Lipohypertrophy in gluteal regions and thighs – Mild metabolic disorders

EXTREME LEANNESS IN CONGENITAL GENERALIZED LIPODYSTROPHY SYNDROMES

Lipodystrophy is the pathological reduction of adipose tissue mass but also the abnormal distribution of body fat in which areas of lipoatrophy are counterbalanced by areas of excessive localized adiposity. Clinically, lipodystrophies are classified by genetic, etiological, and anatomical considerations (Table 21.2) (Patni and Garg, 2015). The partial lipodystrophy associated with highly active antiretroviral therapy is the most common lipodystrophy syndrome worldwide, affecting 20%–50% of HIV-infected patients. By contrast, congenital generalized lipodystrophy (CGL) (Berardinelli–Seip syndrome), familial partial lipodystrophy, and acquired generalized lipodystrophy are infrequent conditions, with an estimated CGL prevalence of 1:10,000,000 in the general population. Although the genetic bases of some of these diseases are well-established (Table 21.2), molecular and cellular defects that lead to lipoatrophy (sometimes accompanied by lipohypertrophy of other body regions) remain largely unknown.

Generalized lipodystrophies are the most extreme forms of leanness owing to adipose tissue deficiency. This disease is associated with severe insulin resistance, diabetes, hyperlipidemia, fatty liver, polycystic ovary syndrome, and female infertility. From a biological perspective, the study of generalized lipodystrophy has been instrumental in discovering new mechanisms implicated in adipose tissue development, growth, and maintenance, as well as the molecular links between adipose tissue dysfunction and systemic insulin resistance. CGL is caused by mutations in four known genes (1-acylglycerol-3-phosphate-o-acyltransferase-2 [*AGPAT2*], *BSCL2*, *CAV1*, and *PTRF*) and follows an autosomal recessive inheritance pattern. The most studied form of CGL is caused by mutations in the *AGPAT2* gene, which encodes *AGPAT*, an enzyme that converts lysophosphatidic acid to phosphatidic acid in the biosynthesis pathway of triglycerides and glycerophospholipids. Although 10 other *AGPAT* isoforms have been identified by sequence homology, only mutations in *AGPAT2* are implicated in human CGL. Similar to patients with CGL, *AGPAT2*-deficient (*Agpat2*^{-/-}) mice completely lack white and brown adipose tissue and develop severe insulin resistance, diabetes, hypertriglyceridemia, fatty liver, and female infertility.

RECAPITULATION

We have described common genetic variations influencing BMI and adiposity traits in the population, as well as rare mutations causing either monogenic obesity syndromes or extreme leanness. GWAS has revealed more than 500 genetic loci associated with individual small effects on BMI at the population level, covering the whole range of body weight from leanness to obesity. In monogenic obesity, rare mutations in genes of the leptin-melanocortin system (*LEP*, *LEPR*, *MC4R*, and *POMC*, among others) have been identified as causing monogenic massive obesity and hyperphagia. Interestingly, there are selected examples of patients with opposite phenotypes of obesity and leanness related to gain versus losses of DNA segments or gain-of-function versus loss-of-function mutations in key genes related to whole-body energy homeostasis (for example, structural variations of chromosome 16p11.2). Finally, genetic causes of several CGL syndromes are also well-established (*AGPAT2* and *BSCL2*, among others). The evaluation of obesity and leanness as opposite phenotypes represents an interesting approach to finding causal gene variations related to body composition and assessing the metabolic complications of obesity.

References

- Akiyama, M., Okada, Y., Kanai, M., Takahashi, A., Momozawa, Y., Ikeda, M., et al., 2017. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nat Genet* 49 (10), 1458–1467.
- Huvenne, H., Dubern, B., Clément, K., Poitou, C., 2016. Rare genetic forms of obesity: clinical approach and current treatments in 2016. *Obes Facts* 9 (3), 158–173.
- Loos, R.J., 2018. The genetics of adiposity. *Curr Opin Genet Dev* 50, 86–95.
- Müller, M.J., Geisler, C., Blundell, J., Dulloo, A., Schutz, Y., Krawczak, M., et al., 2018. The case of GWAS of obesity: does body weight control play by the rules? *Int J Obes* 42 (8), 1395–1405.
- Patni, N., Garg, A., 2015. Congenital generalized lipodystrophies—new insights into metabolic dysfunction. *Nat Rev Endocrinol* 11 (9), 522–534.
- Speakman, J.R., Loos, R.J.F., O'Rahilly, S., Hirschhorn, J.N., Allison, D.B., 2018. GWAS for BMI: a treasure trove of fundamental insights into the genetic basis of obesity. *Int J Obes (Lond)* 42 (8), 1524–1531.
- Turcot, V., Lu, Y., Highland, H.M., Schurmann, C., Justice, A.E., Fine, R.S., et al., 2018. Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. *Nat Genet* 50 (1), 26–41.
- Yeo, G.S.H., 2017. Genetics of obesity: can an old dog teach us new tricks? *Diabetologia* 60 (5), 778–783.

This page intentionally left blank

Genetic Regulation of Energy Homeostasis: Obesity Implications

Itziar Abete^{1,2}, Santiago Navas-Carretero^{1,2},
Maria Angeles Zulet^{1,2,3}, J. Alfredo Martinez⁴

¹Nutrition Research Center, University of Navarra, Pamplona, Navarra, Spain; ²Centro de Investigación Biomédica en red: Fisiopatología de la Obesidad y Nutrición (CiberObn), Madrid, Spain; ³Instituto de Investigación Sanitaria de Navarra (IDISNA), Recinto de Complejo Hospitalario de Navarra, Pamplona, Navarra, Spain; ⁴Ciencias de la Alimentación Fisiología, Universidad de Navarra, Navarra, Spain

INTRODUCTION

A sustainable energy supply is essential to maintain body functions and a healthy status. Energy homeostasis involves regulating energy intake and energy expenditure; thus, chronic unbalance of this process is a direct cause of energy accumulation and the onset of obesity. Obesity is a major public health problem that leads to the onset of other metabolic diseases.

ENERGY HOMEOSTASIS

Thermodynamic balance in the body includes energy intake, energy expenditure, and energy storage. Changes in body weight over time depend on an inadequate balance between energy intake and expenditure. Humans receive energy from protein, carbohydrates, fat, and alcohol. On the other hand, energy expenditure involves the resting metabolic rate, the thermogenic effect of food, and physical activity. The resting metabolic rate is the amount of energy necessary to fuel the body at rest, which typically is estimated to be at least 50%–60% of the total energy expenditure. The thermic effect of food is the energy cost of absorbing and metabolizing food that is consumed. It is being associated with the food composition intake, which on a typical mixed diet is normally about 8%–10% of the total energy ingested. Physical activity—associated energy expenditure is the most variable and modifiable factor of energy expenditure and consists of the length (minutes) of physical activity performed multiplied by the energy cost (kilojoules or kilocalories per minute) of that activity, and may represent more than 40% of total energy

expenditure. When energy intake equals energy expenditure, the body is in energy balance and body weight and composition remain in a steady state. When energy intake exceeds energy expenditure, a state of positive energy balance occurs; the consequence is an increase in body mass. When energy expenditure exceeds energy intake, a situation of negative energy balance ensues; the consequence is a loss of body mass and eventually body fat content decreases. However, many unknown features remain concerning fuel homeostasis and energy balance regulation, depending on the genetic makeup.

GENETIC INFLUENCE ON ENERGY HOMEOSTASIS

Body weight and composition, as well as the storage of energy as triglycerides in adipose tissue, are determined by interactions among genetic, environmental, and psychosocial factors. The role of genes in body fat regulation is well-established; however, the rising prevalence of obesity has not been caused by a recent change in individuals' genetic background. The propensity for obesity most likely results from a series of changes in the environment, particularly an increased availability of food that is also high in energy (energy-dense food), combined with a reduced energy requirement for physical performance. Nevertheless, genetic factors have an important role in the etiology of obesity, considering that differences concerning fat deposition are found among individuals with a high caloric intake and low physical activity.

Although the cause of obesity can be monogenic (when it is typically caused by a single gene mutation

with severe obesity as the main symptom) or syndromic (when the obesity is a symptom of a genetic syndrome), the vast majority of cases are polygenic (as the result of the interaction of several gene variants and environment). Indeed, there is huge interindividual heterogeneity. Thus, the specific set of genes entailing obesity in one subject is unlikely to be the same in another randomly selected obese subject. Interestingly, these variants predisposing to obesity have been also found in normal weight and even in lean individuals.

Genome-wide association studies have led to the identification of many common genetic variations in the human genome. Considering their individual effects, these allelic polymorphisms have only a minor impact on body weight. Nevertheless, through the sum of their effects, whether beneficial or disadvantageous, these variations define an individual's predisposition to gain weight and present the framework for polygenic obesity, particularly in the face of environmental changes. Identification of these subtle factors predisposing to obesity may be the first step to understanding apparent body weight variations in a given population.

In this chapter, we summarize a number of mutations and polymorphisms or genetic variants in genes encoding factors regulating food and energy intake and factors implicated in energy expenditure.

GENES ENCODING FACTORS REGULATING ENERGY INTAKE

Most obesity-predisposing genes encode molecular components of the physiological systems regulating energy balance. More specifically, leptin (*LEP*), leptin receptor (*LEPR*), ghrelin (*GHRL*), ghrelin receptor (*GHSR*), and brain-derived neurotrophic factor (*BDNF*) genes have been shown to have variants potentially related to the pathophysiology of obesity and accompanying complications.

In this context, different single nucleotide polymorphisms (SNPs) present in these genes are associated with a higher risk for obesity (Table 22.1). In relation to the *LEP* gene, the SNP *LEP* G2548A, which is fairly prevalent in the population, and the presence of the GG genotype combined with the GG allele of the *LEPR* Q223R allele, have been shown to increase the risk for obesity. Also, in the gene encoding for *LEPR*, another two SNPs, K109R and K656N, have been studied. In this case, the controversial results encountered did not confirm the relationship of a specific genotype and a higher risk for overweight or obesity. The influence of *LEPR* polymorphisms (rs2767485) on serum orexigenic (neuropeptide Y [NPY], melanin-concentrating hormone [MCH], and agouti-related peptide [AgRP]) and anorexigenic neuro-peptides (*LEP* and α -melanocyte-stimulating hormone)

was assessed. C-Allele carriers had statistical higher levels of orexigenic markers such as NPY, AgRP, and MCH compared with TT type. However, these differences did not remain after weight loss treatment.

In relation to *GHRL* and *GHSR* genes, a large number of variants have shown a relation to higher risk for obesity (Table 22.1). The *GHRL* SNP most studied was Leu72Met.

A review article compiled the association of common polymorphisms of these genes with overweight/obesity and the related disturbances related to it. The researchers found 17 articles studying the influence of *LEP/LEPR* SNPs on body mass index (BMI) and on metabolic alterations. Of the 17 articles, nine reported a significant association with obesity and/or its related health outcomes, or at least considered them to be possible risk factors. However, eight failed to evidence such relationships. The *LEPR* SNP Q223R was studied in 12 studies. Four reported an association of SNPs with obesity and/or overweight and one study found an association with the risk for type 2 diabetes mellitus. K109R and K656N were related to obesity in children, but not in adults. However, one study associated these SNPs with obesity only when present simultaneously. In another study, SNPs Q223R, K109R, and K656N were significantly associated with obesity in children.

Seven studies reported the influence of *GHRL/GHSR* SNPs on obesity; the Leu72Met SNP was analyzed in six. Two investigations found an association with obesity whereas four failed to show an association with obesity in only two studies. Arg51Gln was included in only one study, which found no association. Other SNPs were also investigated (including +3056 T > C [rs2075356], -1500 C > G [rs3755777], -1062 G > C [rs26311], -994 C > T [rs26312]) and were reported to be significantly related to obesity (Takezawa et al., 2013).

BDNF is a neurotrophin that has a fundamental role in the regulating food intake and body weight control. In the general population, common variants of *BDNF* that affect *BDNF* gene expression or *BDNF* protein processing have also been associated with modest alterations in energy balance and cognitive functioning. *BDNF*-positive or negative humans exhibit hyperphagic behavior and obesity. Intronic SNP rs12291063 has emerged as the strongest predictor of hypothalamic *BDNF* expression and BMI. In a cohort of 837 healthy children who underwent body composition analyses, subjects with the rs12291063 CC genotype had a significantly greater percentage of body fat and fat mass compared with CT and TT subjects. On the other hand, four large-scale genome-wide association studies using 30,000 to nearly 250,000 subjects found seven SNPs in or near the *BDNF* gene (rs4074134, rs4923461, rs925946, rs10501087, rs6265, rs10767664, and rs2030323) that were associated with increased BMI. In an adult Japanese

TABLE 22.1 Single Nucleotide Polymorphisms in Some Genes Related to Energy Homeostasis.

Gene	Polymorphism	Risk Allele	Association With Obesity
<i>Genes involved in appetite regulation and energy homeostasis</i>			
FTO	rs9939609	A	Higher risk for obesity and type 2 diabetes
	rs17817449	G	Higher risk for obesity
	rs3751812	G	Higher risk for obesity
	rs1421085	C	Higher risk for obesity
	rs9930506	G	Association with obesity traits
LEP	rs7799039	A	In haplotype with <i>LEPR</i> Q223R. Increased risk for obesity
LEPR	rs1137101	A	In haplotype with <i>LEP</i> G2548A. Increased risk for obesity
	rs1137100	G	Controversial results. Risk for obesity in children, not adults
	K656N	C	Controversial results. Risk for obesity in children, not adults
	rs2767485	C	Increased orexigenic neuropeptides levels among obese adolescents
GHRL/GHSR	rs696217	T	Possible association with obesity
	rs34911341	T	No association found with obesity
	rs2075356	C	No association found with obesity
	rs3755777	G	No association found with obesity
	rs26311	G	No association found with obesity
	rs26312	A	No association found with obesity
BDNF	rs12291063	C	Association with higher BMI and adiposity
	rs4074134	T	Association with higher BMI and type 2 diabetes
	rs4923461	G	Association with higher BMI
	rs925946	T	Higher risk for obesity
	rs10501087	C	Association with higher BMI
	rs6265	T	Association with higher BMI
	rs10767664	T	Association with higher BMI
	rs2030323	A	Association with higher BMI
<i>Genes involved in energy metabolism and thermogenesis</i>			
ADRB2	rs1042713	A	Association with metabolic syndrome
	rs1042714	G	Association with metabolic syndrome
	rs1800888	T	Increased risk for obesity and cardiovascular disease
ADRB3	rs4994	G	Increased risk for obesity
UCP1	rs12502572	G	Reduced energy expenditure
UCP2	rs17132534	C	No association with BMI
	rs659366	T	Association with central obesity
	rs660339	A	Possible association with obesity and central obesity
	45-bp insertion/deletion		Controversial results for obesity onset
UCP3	rs1800849	A	No clear association with obesity
MC4R	rs17782313	C	Association with higher BMI
	rs12970134	A	Association with higher waist circumference

BMI, body mass index.

population ($n = 2865$), rs4074134, rs4923461, and rs10501087 were significantly correlated with BMIs greater than 30 kg/m^2 . In a Chinese children population, rs6265 was significantly associated with obesity. Finally, a longitudinal study found that rs925946 was significantly associated with greater childhood weight gain and adult obesity. These associations of *BDNF* SNPs with higher BMI suggest that these gene variants have a significant impact on obesity.

Although several studies show certain relationships, results were often contradictory and inconclusive. Thus, the real association of these SNPs with obesity remains to be clarified and warrants further research.

The fat mass and obesity-associated (*FTO*) gene is also associated with energy balance. Genome-wide association studies have revealed that single-nucleotide polymorphisms (rs9939609, rs17817449, rs3751812, rs1421085, and rs9930506) in intron-1 of the *FTO* gene are associated with an increased risk for obesity. At-risk alleles of these SNPs are associated with greater food intake and increased hunger or lowered satiety but are not associated with altered resting energy expenditure or low physical activity in humans. Animal studies have shown a clear influence of the *FTO* gene on food intake. There is an increase in food intake with a reduction in *FTO* expression and a food intake reduction with an enhanced expression of *FTO*. In research, the *FTO* rs9939609 polymorphism was genotyped in 652 adolescents (53% females) to assess whether energy and macronutrient intake were different across *FTO* rs9939609 genotypes. Results showed no significant differences among *FTO* rs9939609 genotypes. In contrast, the authors found that consumption of low-fat diets (<30% of energy) may attenuate the genetic predisposition toward obesity in risk allele carriers.

The melanocortin 4 receptor (*MC4R*) is a protein that has attracted much attention owing to its involvement in regulating appetite and energy homeostasis. *MC4R* is a key regulator of energy homeostasis, inducing energy expenditure and decreasing food intake. *MC4R* common variants are related to the hypothalamic pathway of energy intake. Polymorphism rs17782313, mapped 188 kb downstream from *MC4R*, has been strongly associated with obesity and higher BMI in adults and children. Experimental genetic studies suggested significant associations with higher appetite and energy intake, and each copy of the rs17782313 C allele in the *MC4R* gene has been associated with a 0.2-kg/m^2 increase in BMI. Two metaanalyses on genetic predictors of obesity confirmed that polymorphisms at the *MC4R* locus are independently associated with BMI in adults and children. The rs17782313 variant was also related to BMI in a Korean population.

Genome-wide association studies have provided evidence that several variants of the *FTO* (rs9939609,

rs7202116, and rs9930506) and *MC4R* (rs12970134 and rs17782313) genes are significantly associated with obesity. However, there are controversial results. A study found that *FTO* rs9939609 and/or *MC4R* rs17782313 did not influence changes in resting energy expenditure or increase the genetic susceptibility to greater body mass regain during the maintenance period.

Vitamin D receptor (VDR) expression in adipocytes has a role in regulating energy metabolism and inducing obesity. The importance of VDR in the etiopathology of obesity is associated with the occurrence of polymorphisms such as *Fok1*, *Bsm1*, *Apa1*, and *Taq1*.

GENES ENCODING FACTORS IMPLICATED IN ENERGY EXPENDITURE

Adaptive thermogenesis in humans is closely related to the active mobilization of lipids from fat tissues and attracts special interest in relation to obesity. Central neural pathways responsible for food intake and energy expenditure regulation are tightly interconnected. The peripheral transmission of central commands to fat stores is mediated by the sympathetic nervous system. Adrenergic receptors (*ADRB2* and *ADRB3*) have an important role in lipolysis and thermogenesis, and the polymorphism of these genes causes differences in energy expenditure.

The β_2 -adrenergic receptor gene (*ADRB2*) encodes a major lipolytic receptor protein in human fat cells. Two common polymorphisms of the *ADRB2* gene (*Arg16Gly* rs1042713 and *Gln27Glu* rs1042714) were explored in several diseases such as hypertension and obesity, but outcomes are still controversial. By altering the amino acid sequence in the extracellular N-terminus of the *ADRB2*, rs1042714 allele mutations are believed to alter *ADRB2* function. A metaanalysis analyzed associations between *Gln27Glu* (rs1042714) or *Arg16Gly* (rs1042713) polymorphisms and obesity susceptibility. Results revealed that the *Gln27Glu* (rs1042714) polymorphism was associated with obesity susceptibility, whereas no associations were found in the case of *Arg16Gly* (rs1042713). According to other investigations, the *Glu27* allele has a tendency to increase BMI, body fat mass, fat cell volume, and the waist to hip ratio, as well as type 2 diabetes, and to suppress lipid oxidation. However, the contribution of *Gly16* to obesity is controversial. The rare functional *ADRB2* rs1800888 (Thr164Ile) polymorphism was also studied and was found to be associated with increased blood pressure, increased frequency of hypertension and increased risk of cardiovascular disease (CVD).

The β_3 -adrenergic receptor gene (*ADRB3*) stimulates lipolysis in adipose cells as does *ADRB2*. A common polymorphism of the *ADRB3* gene is the *Trp64Arg* variant (rs4994). A metaanalysis involving 44,833 individuals

showed a significant association between the *Arg64Trp* variant and BMI; *Arg64*-allele carriers had higher BMI compared with noncarriers. Several authors reported that the *Trp64Arg* polymorphism is associated with increased body weight and insulin resistance. *Trp64Arg* polymorphism was investigated in relation to weight changes at 4-year follow-up. Participants with the *Arg* allele significantly increased in body weight, whereas those without the *Arg* allele did not significantly change weight. Thus, in that study, the *Arg* allele was associated with long-term changes in body weight in obese individuals.

Uncoupling proteins (UCPs) are a class of proteins found in the inner mitochondrial membrane that dissipate the proton gradient, releasing stored energy as heat. These proteins, which may contribute to energy metabolism in mitochondria, may be involved in pathogenesis of obesity. Allelic variants in the UCP genes may influence body weight through variations in resting energy expenditure, substrate oxidation, and exercise efficiency. There are five known types in mammals: *UCP1*, also known as thermogenin, *UCP2*, *UCP3*, *UCP4*, and *UCP5*.

An SNP in the exon 2 promoter region of the *UCP1* gene (−3826 A > G) may influence energy body homeostasis. The mutated allele G was associated with lower energy expenditure, which suggests that the polymorphism has an adverse effect on thermogenesis.

Variants in *UCP2* and *UCP3* genes were reported to be associated with obesity. Some mentioned polymorphisms are −866 G/A, *Ala55Val* and 45-bp insertion/deletion (Ins/Del) in the *UCP2* gene and −55C/T in the *UCP3* gene. Results from observational studies suggested that the *UCP2* −866 G-allele is associated with obesity whereas the A-allele seems to be protective against obesity. In some studies, the *UCP2* Ins/Del polymorphism was related to obesity, but there were also investigations evidencing no relationships with obesity. Some investigations found that T-allele carriers of *Ala55Val* polymorphism had an increased risk for overweight and obesity, but there are also studies that did not support this association.

The impact of these polymorphisms on obesity is still under debate; contradictory results are still being reported. Further studies are required to provide more convincing evidence.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Further research on SNPs regarding energy regulation is needed to improve our understanding of the physiology and pathophysiology of the regulation of energy homeostasis and eating behavior, because it might open new approaches to managing and treating obesity. More than 600

genes or chromosome regions have been associated with energy homeostasis, but they account less than 5% of body weight variability, whereas epidemiological studies suggest that the role of inheritance may be as high as 25%–70%.

Suggested Reading

- Corgosinho, F.C., Almeida, S.S., Tock, L., Pesquero, J.B., Araújo, R.C., Clemente, A.P.G., Dal'Molin Netto, B., da Silveira Campos, R.M., Masquio, D.C.L., de Carvalho Ferreira, J.P., de Lima Sanches, P., de Piano Ganen, A., Rogero, M.M., Oyama, L.M., Tufik, S., de Mello, M.T., Dâmaso, A.R., 2017. LEPR polymorphism may affect energy balance during weight loss among Brazilians obese adolescents. *Neuropeptides* 66, 18–24.
- Fairbrother, U., Kidd, E., Malagamuwa, T., Walley, A., 2018. Genetics of severe obesity. *Curr Diabetes Rep* 18, 85.
- Fanning, E., O'Shea, D., 2018. Genetics and the metabolic syndrome. *Clin Dermatol* 36, 9–13.
- Ghalandari, H., Hosseini-Esfahani, F., Mirmiran, P., 2015. The association of polymorphisms in leptin/leptin receptor genes and ghrelin/ghrelin receptor genes with overweight/obesity and the related metabolic disturbances: a review. *Int J Endocrinol Metab* 13, e19073.
- Ingelsson, E., McCarthy, M.I., 2018. Human genetics of obesity and type 2 diabetes mellitus: past, present, and future. *Circ Genom Precis Med* 11, e002090.
- Jiang, L., Penney, K.L., Giovannucci, E., Kraft, P., Wilson, K.M., 2018. A genome-wide association study of energy intake and expenditure. *PLoS One* 13, e0201555.
- Labayen, I., Ruiz, J.R., Huybrechts, I., Ortega, F.B., Arenaza, L., González-Gross, M., Widhalm, K., Molnar, D., Manios, Y., DeHenauf, S., Meirhaeghe, A., Moreno, L.A., 2016. Dietary fat intake modifies the influence of the FTO rs9939609 polymorphism on adiposity in adolescents: the HELENA cross-sectional study. *Nutr Metabol Cardiovasc Dis* 26, 937–943.
- Locke, A.E., Kahali, B., Berndt, S.I., et al., 2015. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518, 197–206.
- Loos, R.J., 2018. The genetics of adiposity. *Curr Opin Genet Dev* 50, 86–95.
- Mărginean, C.O., Mărginean, C., Melit, L.E., 2018. New insights regarding genetic aspects of childhood obesity: a minireview. *Front Pediatr* 6, 271.
- Mou, Z., Hyde, T.M., Lipska, B.K., Martinowich, K., Wei, P., Ong, C.J., Hunter, L.A., Palaguachi, G.I., Morgun, E., Teng, R., Lai, C., Condarco, T.A., Demidowich, A.P., Krause, A.J., Marshall, L.J., Haack, K., Voruganti, V.S., Cole, S.A., Butte, N.F., Comuzzie, A.G., Nalls, M.A., Zonderman, A.B., Singleton, A.B., Evans, M.K., Martin, B., Maudsley, S., Tsao, J.W., Kleinman, J.E., Yanovski, J.A., Han, J.C., 2015. Human obesity associated with an intronic SNP in the brain-derived neurotrophic factor locus. *Cell Rep* 13, 1073–1080.
- Saeed, S., Arslan, M., Froguel, P., 2018. Genetics of obesity in consanguineous populations: toward precision medicine and the discovery of novel obesity genes. *Obesity (Silver Spring)* 26, 474–484.
- Speakman, J.R., 2015. The 'fat mass and obesity related' (FTO) gene: mechanisms of impact on obesity and energy balance. *Curr Obes Rep* 4, 73–91.
- Takezawa, J., Yamada, K., Miyachi, M., Morita, A., Aiba, N., Sasaki, S., Watanabe, S., 2013. Saku control obesity program (SCOP) study group. Preproghrelin gene polymorphisms in obese Japanese women. Minor homozygotes are light eaters, do not prefer protein or fat, and apparently have a poor appetite. *Appetite* 63, 105–111.
- Turcot, V., Lu, Y., Highland, H.M., Schurmann, C., Justice, A.E., Fine, R.S., Bradfield, J.P., Esko, T., Giri, A., Graff, M., Guo, X.,

- Hendricks, A.E., Karaderi, T., Lempradl, A., CHD Exome+ Consortium, EPIC-CVD consortium, ExomeBP Consortium, Global Lipids Genetic Consortium, GoT2D Genes Consortium, EPIC InterAct Consortium, INTERVAL Study, ReproGen Consortium, T2D-Genes Consortium, MAGIC Investigators, Understanding Society Scientific Group, et al., 2018. Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. *Nat Genet* 50, 26–41.
- Walter, S., Mejía-Guevara, I., Estrada, K., Liu, S.Y., Glymour, M.M., 2016. Association of a genetic risk score with body mass index across different birth cohorts. *J Am Med Assoc* 316, 63–69.
- Waterhouse, E.G., Xu, B., 2013. The skinny on brain-derived neurotrophic factor: evidence from animal models to GWAS. *J Mol Med (Berl)* 91, 1241–1247.
- Yamakita, M., Ando, D., Tang, S., Yamagata, Z., 2010. The Trp64Arg polymorphism of the beta3-adrenergic receptor gene is associated with weight changes in obese Japanese men: a 4-year follow-up study. *J Physiol Anthropol* 29, 133–139.
- Yang, Q., Xiao, T., Guo, J., Su, Z., 2017. Complex relationship between obesity and the fat mass and obesity locus. *Int J Biol Sci* 13, 615–629.

Genetic Variations With Influence on the Individualized Response to Weight Loss Diets: FTO as Evidence

Daniel Antonio de Luis¹, Rocío Aller¹, Olatz Izaola¹, J. Alfredo Martinez²

¹Center of Investigation of Endocrinology and Nutrition, Medicine School and Department of Endocrinology and Nutrition, Hospital Clinico Universitario, University of Valladolid, Valladolid, Spain; ²Ciencias de la Alimentación Fisiología, Universidad de Navarra, Navarra, Spain

Glossary

BMI body mass index

FTO gene fat mass and obesity-associated gene

Nutrigenetics focused on how the interaction of genetic variations and nutrients influences metabolism, health conditions, and the risk for diet-related diseases such as obesity

SNPs are genetic variations that could modulate weight and metabolic response after weight loss diets. Understanding nutrigenetic interaction is a milestone to a personalized nutrition on obesity patients.

SNP single nucleotide polymorphism

INTRODUCTION

The increasing prevalence of obesity is an important public health burden associated with a great number of chronic diseases. Such adverse outcomes involve an increased risk for cardiovascular diseases, type 2 diabetes, hyperlipidemia, metabolic syndrome, and some types of cancer. Globally, body mass index (BMI) is epidemiologically used as an adiposity marker; in more than 1.5 billion overweight adults, it is equal to or more than 25 kg/m², whereas obese adults, with a BMI of 30 kg/m² or greater, were estimated at about 500 million by world health organization (WHO).

Obesity is the consequence of a higher dietary energy intake than energy expenditure. It results in an imbalance of energy and an increase in body weight, in which the genetic background is involved. Indeed, obesity is influenced by many factors such as behavior, environment, hormonal status, metabolic processes, perinatal

conditions, and genetic predisposition. A huge number of strategies can be used to induce a negative energy balance and secondary weight loss, such as lifestyle modifications including reducing energy intake and increasing physical activity as well as pharmacological prescriptions or even surgical treatments. Those approaches may result in significant weight loss in obese subjects; however, the individual response is variable. The occurrence of hypo or hyper responders supports the hypothesis that the outcome to weight loss intervention is related to genetic variation, in which potential predictors of successful weight loss are still not well-understood.

A number of studies reported that the intake of dietary calories is a major contributor to obesity associated with the consumption of high-energy food groups beyond recommended amounts, particularly fat- and carbohydrate-rich items or junk food. On the other hand, some investigations reported that foods with appetite-controlling properties and fiber-rich foods (such as fruits, vegetables, and whole grains) were inversely related to the prevalence of obesity. In addition, dietary patterns that disturb regulated dietary energy intake, such as snacking, skipping a meal, and away-from-home meals, were also associated with obesity.

Some studies of families, adopted children, and twins suggested that genetic factors accounted for 50%–75% of interindividual variations in BMI. For example, in the Swedish twins study, within-pair correlations for BMI were 0.70 for men and 0.66 for women among identical twins reared apart. After the completion of the

Human Genome Project, single nucleotide polymorphisms (SNPs) were identified, which constitute the most frequent DNA sequence variants in the genome. Indeed, SNP identification is an important area of investigation in the field of nutrigenetics. In this context, in order to associate the genetic make-up with nutrients and outcomes. Nutrigenetics focuses on how the interaction between genetic variations and nutrients influences metabolism, health conditions, and the risk for diet-related diseases such as obesity.

SINGLE NUCLEOTIDE POLYMORPHISMS AS A FORM OF GENETIC VARIATION

SNPs, as distinguished from the term “mutation,” exist at a frequency of 1% or higher in the human population. SNPs are single nucleotides (A [adenine], T [thymine], C [cytosine], and G [guanine]), with differences in DNA sequences both within and among populations. For example, two sequenced DNA fragments from different individuals, CCTTCA to CCTACA (forward); GGAAGT to GGATGT as (reverse), have a difference in a single nucleotide. A and T are called alleles. There are three potential genotypes: T/T, T/A, and A/A. The genotypes of individuals with A/A or T/T alleles and those of individuals with T/A are assigned as being homozygous and heterozygous, respectively. Differences in the genotype have certain functional consequences, such as altering the expression and function of a protein, which may be followed by the incidence of common diseases. Differences in SNPs among individuals may occur in some cases because some individuals are more susceptible to obesity or respond differently to dietary interventions for weight loss.

RESPONSE TO WEIGHT LOSS DIETS AND SINGLE NUCLEOTIDE POLYMORPHISM

Genetic makeup and genetic variations (SNPs) determine an individual’s susceptibility to obesity through a series of potential mechanisms governing pathways and regulatory systems at different levels, including energy intake and expenditure. Several studies examined interactions between dietary nutrient intake and genetic variations (SNPs) involved in obesity-related variables. These genetic variations may have a role in different metabolic pathways involved in energy use, such as lipid metabolism, adipogenesis, cell differentiation, energy consumption, lipolysis, leptin and endocannabinoid receptors, appetite control, adipocyte function, and inflammation pathways. On the other hand, association and linkage studies have shown evidence of relationships among candidate obesity genes

and body weight, BMI, body fat, fat distribution, energy expenditure, fuel oxidation, and several other phenotypic features of obesity, including adiposity-related health risks. Thus, more than 100 genes have been reported to participate in energy metabolism regulation, and a recent genome-wide association study showed that about 97 genes are related to body weight and adiposity regulation.

Most research in the literature are observational studies without intervention designs, as there are fewer experimental trials assessing short-term weight-loss being rare long-term maintenance designs. It is difficult to compare and interpret to these interventional studies are because they assessed heterogeneous populations (children, men, women, and overweight, obese, diabetic, or prediabetic subjects), and also because of heterogeneity in the duration and type of treatment (dietary, pharmacological, or bariatric surgery), differences in the sample size, lack of adequate control groups, and other potential biases. Table 23.1 lists potential genes candidates that influence weight loss in weight loss strategies.

In this context, huge progress has been made in identifying polymorphisms in genes related to obesity and the weight loss response after a hypocaloric diet. For instance, there is great consistency in the influence of the fat mass and obesity-associated (*FTO*) gene on weight loss outcomes, whereas for other variants there is less information. The role of an SNP of *FTO* with a personalized vision of the influence of a specific SNP (rs9939609) in the *FTO* gene on weight loss with different types of diets was examined.

POLYMORPHISM RS9939609 OF THE FTO GENE AND DIETS FOR WEIGHT LOSS

After its identification as a candidate gene for obesity, *FTO* was predicted to be a 2-oxoglutarate Fe(II)-dependent demethylase closely related to bacterial DNA demethylase AlkB and mammalian AlkB homologs 1 and 2. In humans, a loss-of-function mutation in *FTO* leads to a complex phenotype of postnatal growth retardation, microcephaly, severe psychomotor delay, functional brain deficits, and characteristic facial dimorphism. Despite the severe phenotype seen in *FTO* deficiency, however, a compelling body of evidence has emerged to support a role for *FTO* in controlling energy homeostasis. *FTO* is expressed ubiquitously in human and animal tissues, which is consistent with multiple organ systems being affected in *FTO* deficiency. The highest expression, however, is seen in the brain, including the hypothalamus, where control of food intake is located. Within the arcuate nucleus of the hypothalamus (ARC), *FTO* is bidirectionally regulated as a function of nutritional status, decreasing after a 24- to 48-h fast and

TABLE 23.1 Genes Candidates With an Evaluated Role in Weight Loss Interventions.

Adipogenic Genes
<i>Peroxisome proliferator-activated receptor PPARγ2</i> (Pro12Ala)
Genes Related to Appetite Control
<i>Cannabinoid receptor gene</i> (CNR1) (G1359A) (rs1049353)
<i>Fatty acid amide hydrolase</i> (FAAH) (385 C-NA)
<i>Leptin</i> LEP (C-2549A) (5'-region)
<i>Leptin receptor</i> (LEPr) (Ser343Ser) (T/C) LEPr(Lys656Asn)
<i>Pro-opiomelanocortin</i> POMC (R236G)
<i>Serotonin receptor HTR2C promoter</i> (C-759T)
<i>Neuromedin β</i> (NMB) (Pro73 Thr)
<i>Melanocortin receptor</i> (MCR)
MC4R (rs1943218, rs17066866, rs17066856, rs9966412, rs17066859, rs8091237, rs7290064)
MCR3R (C17A) (C241A)
Genes Related to Energy Expenditure
<i>Adrenoreceptor β3</i> ADRB3 (Trp64Arg)
<i>Uncoupling proteins</i> UCP1 (A-3826G) UCP2 (G-866A), UCP3 (–55CT)
Genes Related to Insulin Metabolism
<i>Glucagon-like peptide-1</i> (GLP)1-receptor (rs6923761)
<i>Insulin receptor substrates/insulin-like growth factor</i> (IGF)
Insulin 2 near INSIG2 (rs7566605)
IRS1 (Gly972Arg)/IRS-2 (G1057D)
IGF (GAA1013GAA)
<i>T-cell factor 7-like 2 gene</i> (TCF7L2) (rs12255372, rs7903146, rs11196205, rs7895340)
Genes Related to Lipid Metabolism
<i>Apolipoproteins (apoenzyme)</i>
APOA4 360His
APO A5 (T1131C)
<i>Fatty acid binding proteins</i> (FABP) (Ala54 Thr)
<i>Hepatic lipase</i> LIPC (G-250G)
<i>Perilipin</i> (PLIN) (G11482A) and PLIN (A14995T)
<i>Acyl-CoA synthetase 5</i> (ACSL5) (rs2419621) (C/T)
Genes Potentially Related to Obesity
<i>Fat mass and obesity-associated</i> FTO (rs9939609, rs17817449, and rs1421085)
TNF- α (G308A)
<i>Plasma factor VII Arg/Gln353</i>
<i>Angiotensin-converting enzyme</i> (ACE) I/D

Continued

TABLE 23.1 Genes Candidates With an Evaluated Role in Weight Loss Interventions.—cont'd

<i>Catechol-O-methyltransferase</i> CYP19/COMT CYP19 11r/Val108/158Met
<i>Adrenoreceptor β3</i> (ADRB3x) UCP(1or3) (Trp64Arg) x (A-3826G) or (–55C Δ T)
CLOCK rs1801260
Genes and Drug Treatment to Weight Loss
<i>Adrenoreceptor</i> ADRB3 (Trp64Arg)
<i>Serotonin transporter</i> SLC6A4 (LS/SS)
<i>G-protein β3 subunit</i> (GNB3) (C825T)
<i>Phenylethanolamine N-methyltransferase</i> (PNMT) (G148A)
a2A Adrenoreceptor C1291G
Genes and Bariatric Surgery
<i>Interleukin-6</i> (IL6) (G-174C)
<i>Uncoupling proteins</i>
UCP2, UCP2 (G-866A)
UCP1 UCP1 (A-3826G)
<i>G proteins</i> GNAS1 (T393C), GNB3 (C825T) (G814A)
<i>Cannabinoid receptor gene</i> (CNR1) (G1359A) (rs1049353)
<i>Fatty acid amide hydrolase</i> (FAAH) (385 C-NA)
LEPR (Ser343Ser) (T/C) LEPR(Lys656Asn)
GLP1-Receptor (rs6923761)
<i>Tumor necrosis factor</i> (TNF- α) (G308A)
FTO (rs9939609)

increasing after 8–10 weeks of exposure to a high-fat diet, whereas modulating *FTO* levels specifically in the ARC can influence food intake. The weight of evidence supports the notion that *FTO* itself can influence energy homeostasis by having a direct effect on food intake.

The *FTO* gene is highly polymorphic, and several variants (rs3751812, rs17817449, rs17818902, and rs9939609) of the gene are associated with obesity or high BMI. One of these genetic variants (rs9939609), located within the first *FTO* intron, is related to an increased risk for obesity. Most studies are based on cross-sectional data and evaluate the relationship between lifestyle influences and the *FTO* rs9939609 gene variant on body weight changes. Few intervention studies explored the interaction between different type of diets and *FTO* gene variants on adiposity indices or metabolic parameters, although some trials were designed to test the effect of *FTO* rs9939609 on weight loss, with different types of diets in obese subjects (Table 23.2). Thus, a sample of 305 obese nondiabetic outpatients was enrolled in a randomized controlled trial. In the basal visit, patients were

randomly allocated to one of two diets for 3 months. Diet I was low in carbohydrate and provided 1507 kcal/day (38% carbohydrates, 26% proteins, and 36% fats). Diet II was low in fat and provided 1500 kcal/day (53% carbohydrates, 20% proteins, and 27% fats). The quality of dietary fat was similar in both diets (45% monounsaturated fatty acids, 30% saturated fatty acids, and 25% polyunsaturated fatty acids). The results showed no association between *FTO* variant rs9939609 and body weight response. However, an interaction was found with the macronutrient distribution of the hypocaloric diets and metabolic changes resulting from weight loss. This finding suggests that prescribing a hypocaloric diet (low fat or low carbohydrate) is beneficial for both groups of subjects (wild- or mutant-type genotype) of the rs9939609 polymorphism, with the same weight-lowering impact and decreases in fat mass, waist circumference, and systolic blood pressure. However, in another study, A allele carriers seemed to have even better weight loss than TT genotypes following a Mediterranean-style diet. In another design, a trend toward lower weight loss was observed for AA carriers in a young population, mean age 10 years. The most relevant data of such a study were that changes in dietary macronutrient composition resulting from a hypocaloric diet may interact with the polymorphism and confer to A allele carriers some advantages to body weight loss owing to low-fat intake. The improvement in insulin levels and homeostatic model assessment for

insulin resistance (HOMA-IR) was statistically significant in wild-type patients with both diets and only in the mutant-type group with a low-fat diet. Moreover, in the latter group, total cholesterol, low-density lipoprotein (LDL) cholesterol, and C-reactive proteins levels decreased, as well. This beneficial metabolic result in A allele carriers with diet II could be the result of an interaction of the macronutrient distribution of the diet (low fat percentage) with the A variant of this *FTO* SNP, because the decrease in weight was similar in both diets. In another interventional study, the lifestyle modification program was a low-fat hypocaloric diet (1520 energy/day) for 3 months, with a macronutrient distribution of 52% carbohydrates, 25% lipids, and 23% proteins, with 50.7% monounsaturated fats, 38.5% saturated fats, and 11.8% polyunsaturated fats. After the dietary treatment, the TT genotype group experienced a greater decrease in weight and waist circumference than the A allele carrying subjects. The AT + AA genotype group experienced a greater decrease in weight, -3.1 (3.6) kg (decrease in TT genotype, -2.4 [4.1] kg; $P = .005$) and in waist circumference, -3.4 (6.4) cm (decrease in TT genotype, -2.6 (4.8) cm; $P = .008$) than the other group. In the TT genotype group, total cholesterol and LDL cholesterol levels decreased significantly. In the TA + AA genotype group and after dietary treatment, total cholesterol, LDL cholesterol, insulin, and HOMA-IR decreased significantly, as well. After dietary treatment and in the

TABLE 23.2 Effect of rs9939609 on Weight and Other Parameters After Weight Loss Diet.

Author	Sample Size	Duration Average Age	Type of Diets	Effect on Weight	Effect on Metabolic Parameters
Reihner et al. (2009)	208	12 months 10.8 years	Lifestyle intervention	AA carriers lowest degree of weight loss	No reported
Razquin et al. (2010)	776	36 months 67–68 years	Mediterranean diet (MD) two arms (PREDIMED study) and control diet	Better weight loss A allele carriers with both MD	No reported
De Luis et al. (2012)	305	3 months 43.5 years	Two arms: diet I low in carbohydrate 28% versus Diet II low in fat 27%	No differences in weight loss A allele carriers	A allele carriers had better response with diet II of total cholesterol, low-density lipoprotein cholesterol, and C-reactive protein
De Luis et al. (2013)	106	3 months 49.5 years	One arm: low-fat diet 25%	Better weight loss A allele carriers	A allele carriers had better response with of insulin and homeostatic model assessment for insulin resistance (HOMA-IR)
De Luis et al. (2015a)	233	3 months 48.3 years	Two arms: diet M high monounsaturated 67.5% versus Diet HP (high polyunsaturated 22%, 7%)	Better weight loss A allele carriers with both diets	A allele carriers had better response with diet P of insulin and HOMA-IR
De Luis et al. (2015b)	195	9 months 50.4 years	Two arms: diet HP (high protein content 34%) versus Diet S (standard protein content) 20%	Better weight loss A allele carriers with both diets	No differences in A allele carriers

TT genotype group, total cholesterol, LDL cholesterol, HOMA-IR, and insulin level decreases were lower than in the TA + AA genotype group. Differences in weight response in the previous studies could have resulted from the duration and style of dietary intervention in the protocols or differences on background characteristics in the study populations (basal weight, sex distributions, average age, and dietary fat, for example).

To explore this last hypothesis (type of dietary fat), a trial was conducted in which patients were randomly allocated to one of two diets for 3 months. Diet M (high monounsaturated fat hypocaloric diet) consisted of a 1341 kcal diet: 46.6% of carbohydrates, 34.1% of lipids, and 19.2% of proteins. The distribution of fats was 21.7% saturated fats, 67.5% monounsaturated fats, and 10.8% polyunsaturated fats. Diet P (high polyunsaturated fat hypocaloric diet) consisted of a 1452 kcal diet: 45.7% carbohydrates, 34.4% lipids, and 19.9% proteins. The distribution of fats was 21.8% saturated fats, 55.5% monounsaturated fats, and 22.7% polyunsaturated fats (7 g/day of w-6 fatty acids, 2 g/day of w-3 fatty acids, and a ratio of w6/w3 of 3.5). In this experiment, consuming a hypocaloric diet (high monounsaturated [M] or high polyunsaturated [P] hypocaloric diet) was beneficial for both groups of subjects (TT or AT + AA genotype) of the rs9939609 DNA variant. However, A allele carriers had the greatest reduction in weight with an additional improvement in HOMA-IR, insulin, and leptin levels after the P diet.

To validate the potential effect of intervention time, another study was performed over 9 months. A total of 195 obese subjects were randomly allocated to one of two diets. The high-protein/low-carbohydrate hypocaloric diet (HP diet) was 1050 kcal/day and consisted of 33% carbohydrates (86.1 g/day), 33% fats (39.0 g/day), and 34% proteins (88.6 g/day). The distribution of fats was 23.5% saturated fats, 63.8% monounsaturated fats, and 12.6% polyunsaturated fats. The standard protein hypocaloric diet (S diet) was 1.093 kcal/day and consisted of 53% carbohydrates (144.3 g/day), 27% fats (32.6 g) and 20% proteins (55.6 g/day). The distribution of fats was 20.9% saturated fats, 67.4% monounsaturated fats, and 11.6% polyunsaturated fats. With the HP diet, BMI (-1.9 ± 1.2 versus -2.10 ± 1.8 ; $P < .05$), weight (-6.5 ± 2.1 versus -10.1 ± 4.1 kg; $P < .05$), fat mass (-3.9 ± 3.2 versus -6.0 ± 3.4 kg; $P < .05$), and waist circumference (-5.7 ± 5.0 versus -9.9 ± 5.5 cm; $P < .05$) decreased in both genotype groups (TT versus AT + AA). With the S diet, BMI (-0.9 ± 1.1 versus -1.8 ± 1.2 ; $P < .05$), weight (-3.2 ± 3.0 versus -9.1 ± 3.6 kg; $P < .05$), fat mass (-3.0 ± 3.1 versus -5.2 ± 3.1 kg; $P < .05$), and waist circumference (-3.1 ± 4.0 versus -8.1 ± 4.9 cm; $P < .05$) decreased in both genotype groups. There were significant differences between the effects of the HP and S diet (on

weight, BMI, and fat mass) in the different genotype groups. After the HP diet, glucose, insulin levels, HOMA-IR, total cholesterol, triglycerides, and LDL total cholesterol decreased in both genotype groups. However, with the S diet, only total cholesterol and LDL total cholesterol decreased in both genotype groups. The evidence revealed an association between the rs9939609 DNA variant and weight loss after high-protein and standard hypocaloric diets. Weight loss was better in A carriers than noncarriers with both diets. Also, an interaction was observed with the type of hypocaloric diets and metabolic changes as a result of weight loss. Metabolic improvement was better with the high-protein hypocaloric diet.

A metaanalysis with nine of the cited studies contributed some data in this area, minimizing the effect of the minor allele of this polymorphism in the response of weight to the diets, without having evaluated the metabolic variables.

CONCLUSION

Many genetic variations affecting the development of obesity are identified in the literature, but the results are sometimes controversial, even exploring the same SNP. Data are being published on *FTO* variants, with potential roles in therapeutic pathways. Therefore, more investigations are still necessary in various genetic variations on obesity, association between diets and weight loss and genetic variations regarding obesity for effective management tools of personalized nutrition on obesity. Other genetic variations such as deletions/insertions, rearrangements, and so on should be investigated to increase knowledge in this area. Future studies will have to be large to assess the effects of multiple polymorphisms and will have to control for many factors other than diet. Caution is needed before applying these results to clinical practice. The molecular basis of variability in weight change has the potential to improve treatment outcomes and can simultaneously consider information from genomic and other sources in devising individualized treatment plans. Our group is currently evaluating these SNPs in response to bariatric surgery and weight loss drugs.

References

- de Luis, D.A., Aller, R., Izaola, O., 2012. Evaluation of weight loss and adipocytokines levels after two hypocaloric diets with different macronutrient distribution in obese subjects with rs9939609 gene variant. *Diabetes Metab Res Rev* 28, 663–668.
- de Luis, D.A., Aller, R., Conde, R., Izaola, O., Gonzalez Sagrado, M., Castrodeza Sanz, J., 2013. The rs9939609 gene variant in *FTO* modified the metabolic response of weight loss after a 3-month intervention with a hypocaloric diet. *J Investig Med* 61, 22–26.

- de Luis, D., Aller, R., Izaola, O., Primo, D., Urdiales, S., Romero, E., 2015a. Effects of a high-protein/low-carbohydrate diet versus a standard hypocaloric diet on weight and cardiovascular risk factors: role of a genetic variation in the rs9939609 FTO gene variant. *J Nutrigenetics Nutrigenomics* 8, 128–136.
- de Luis, D.A., Aller, R., Izaola, O., Pacheco, D., 2015b. Role of rs9939609 FTO gene variant in weight loss, insulin resistance and metabolic parameters after a high monounsaturated *vs* a high polyunsaturated fat hypocaloric diets. *Nutr Hosp* 32, 175–181.
- Razquin, C., Martinez, J.A., Martinez-Gonzalez, M.A., Bes-Rastrollo, M., Fernández-Crehuet, J., Marti, A., 2010. A 3-year intervention with a Mediterranean diet modified the association between the rs9939609 gene variant in FTO and body weight changes. *Int J Obes* 34, 266–272.
- Reinehr, T., Hinney, A., Hebebrand, T.A.M., 2009. Aggravating effect of INSIG2 and FTO on overweight reduction in a one-year lifestyle intervention. *Arch Dis Child* 94, 965–967.
- Finucane, M.M., Stevens, G.A., Cowan, M.J., Danaei, G., Lin, J.K., Paciorek, C.J., Singh, G.M., Gutierrez, H.R., Lu, Y., Bahalim, A.N., Farzadfar, F., Riley, L.M., Ezzati, M., Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Body Mass Index), 2011. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. *Lancet* 377, 557–567.
- Livingstone, K.M., Celis-Morales, C., Papandonatos, G.D., Erar, B., Florez, J.C., Jablonski, K.A., Razquin, C., Marti, A., Heianza, Y., Huang, T., Sacks, F.M., Svendstrup, M., Sui, X., Church, T.S., Jääskeläinen, T., Lindström, J., Tuomilehto, J., Uusitupa, M., Rankinen, T., Saris, W.H., Hansen, T., Pedersen, O., Astrup, A., Sørensen, T.I., Qi, L., Bray, G.A., Martinez-Gonzalez, M.A., Martinez, J.A., Franks, P.W., McCaffery, J.M., Lara, J., Mathers, J.C., 2016. FTO genotype and weight loss: systematic review and meta-analysis of 9563 individual participant data from eight randomised controlled trials. *BMJ* 354, i4707.
- Martinez, J.A., Navas-Carretero, S., Sari, S.W., Astrup, A., 2014. Personalized weight loss strategies—the role of macronutrient distribution. *Nat Rev Endocrinol* 10, 749–760.
- Merrit, D.C., Jamnik, J., ElSohemy, A., 2018. FTO genotype, dietary protein intake, and body weight in a multiethnic population of young adults: a cross-sectional study. *Genes Nutr* 20 (13), 4.
- Milagro, F.I., Moreno-Aliaga, M.J., Martinez, J.A., 2016. FTO obesity variant and adipocyte browning in humans. *N Engl J Med* 374, 190–191.
- Przeliorz, A., Regulska Iłow, B., 2017. The role of macronutrient intake in reducing the risk of obesity and overweight among carriers of different polymorphisms of FTO gene. A review. *Rocz Panstw Sakl Hig* 68, 5–13.
- Rosen, C.J., Ingelfinger, J.R., 2015. Unraveling the function of FTO variants. *N Engl J Med* 373, 964–965.
- Stunkard, A.J., Foch, T.T., Hrubec, Z., 1986. A twin study of human obesity. *J Am Med Assoc* 256, 51–54.

Suggested Reading

- Boissel, S., Reish, O., Proulx, K., Kawagoe-Takaki, H., Sedgwick, B., Yeo, G.S., 2009. Loss-of-function mutation in the dioxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. *Am J Hum Genet* 85, 106–111.
- Bray, M.S., Loos, R.J., McCaffery, J.M., Ling, C., Franks, P.W., Weinstock, G.M., Snyder, M.P., Vassy, J.L., Agurs-Collins, T., 2016. Conference Working Group. NIH working group report—using genomic information to guide weight management: from universal to precision treatment. *Obesity (Silver Spring)* 24, 14–22.
- Claussnitzer, M., Hui, C.C., Kellis, M., 2016. FTO obesity variant and adipocyte browning in humans. *N Engl J Med* 374, 192–193.
- Deram, S., Villares, S.M., 2009. Genetic variants influencing effectiveness of weight loss strategies. *Arq Bras Endocrinol Metabol* 53, 129–138.

Genetic Variation in the Response to Exercise Training: Impact on Physical Fitness and Performance

Louis Pérusse

Department of Kinesiology, Université Laval, Québec, QC, Canada

INTRODUCTION

There is considerable evidence suggesting that there are individual differences in the response to exercise training (i.e., trainability), even when subjects are exposed to the same standardized exercise regimen, and that genetic factors have a role in explaining these differences (Bouchard, 2012; Bouchard et al., 2015; Bouchard et al., 2011a; Rankinen et al., 2011). The interested reader should refer to reviews cited here for a complete list of references published before 2011. This chapter presents a brief review of the contribution of genetic factors in the trainability of various physical fitness and health-related phenotypes. First, evidence for the presence of interindividual differences in the response to exercise training and for a genetic component in explaining the response to exercise training will be reviewed. Second, genomic markers of trainability identified through candidate gene association studies and genome-wide explorations will be reviewed. Only human exercise intervention studies involving chronic exercise training will be considered in the current review; observational studies evaluating gene–exercise interactions are reviewed elsewhere (Rankinen and Bouchard, 2012).

THE RESPONSE TO EXERCISE TRAINING IS HERITABLE

Individual Differences in Response to Exercise Training

The concept that genetic factors may have a role in the response to regular exercise was first proposed several

decades ago. In a series of exercise training studies performed in groups of sedentary young men and women, it was shown that there were large interindividual differences in the trainability of cardiorespiratory fitness, assessed by maximal oxygen uptake (VO_{2max}) and several physical performance and health-related fitness phenotypes. In these studies, the mean gain in VO_{2max} was approximately 25% of the baseline values, but with a range of no gain to a doubling of the pre-training values. The most important study documenting the extent of interindividual differences in the response to exercise training is undoubtedly the Health, Risk Factors, Exercise Training, and Genetics (HERITAGE) Family Study. In that study, 742 healthy and sedentary black and white subjects from 204 families followed a highly standardized and supervised 20-week endurance-training program on a cycle ergometer. Average increase in VO_{2max} was 384 ± 202 mL O_2 /min, but considerable heterogeneity in the training response (defined as the difference between pre- and posttraining values) was observed, with values ranging from 0 to more than 1000 mL O_2 /min. Similar heterogeneity was observed for the trainability of other traits (Bouchard, 2012; Bouchard et al., 2011a), including submaximal exercise capacity, stroke volume and cardiac output, heart rate (HR) and blood pressure, plasma lipid levels, body fatness, and muscular strength (Sarzynski et al., 2011).

Evidence for a Genetic Component in Response to Exercise Training

Studies performed in pairs of monozygotic (MZ) twins and in the HERITAGE Family Study provided clear evidence that adaptation to exercise training is

characterized by significant familial aggregation and is heritable. Basically, those studies showed that the pattern of response to exercise is more similar among individuals sharing genes (within twin pairs or families) than among those not sharing genes (between twin pairs or families). A series of training exercise studies undertaken in pairs of MZ twins submitted to various training regimens and reviewed elsewhere (Bouchard, 2012; Sarnyanski et al., 2011) showed that individual differences in trainability were not randomly distributed and that the variance in training responses was significantly greater between pairs of MZ twins than within pairs for various performance-related phenotypes. For example, there was about three to eight times more variance between pairs than within pairs for changes in VO_{2max} in response to exercise and intraclass correlation coefficients ranged from 0.44 to 0.77. In the HERITAGE Family Study, training changes in age- and sex-adjusted VO_{2max} showed 2.5 times more variance between families than within families, with a maximal heritability estimate reaching 47%. Adjusting the VO_{2max} response for baseline value did not modify the heritability estimate, which suggests that genetic factors underlying VO_{2max} and its response to exercise appear to be different. The trainability of several other phenotypes was found to be influenced by genetic factors, including submaximal aerobic performance, stroke volume and cardiac output, resting and submaximal exercise HR and blood pressure, body composition and fat distribution, plasma lipid and lipoprotein levels, and skeletal muscle enzyme activities. Maximal heritability for these phenotypes ranged from 25% to 55%. A review and meta-analysis of twin and family studies confirmed these findings regarding the significant contribution of genetic factors in trainability (Zadro et al., 2017). The review included twin and family studies up to Mar. 2016, in which body composition and/or cardiorespiratory fitness were assessed after a physical activity intervention and which provided a heritability estimate or within-MZ twin correlation (r_{MZ}). Maximal heritability estimates for the response to physical activity ranged from 0% to 21% for measures of body composition and from 22% to 57% for cardiorespiratory fitness (Zadro et al., 2017).

GENOMIC MARKERS OF TRAINABILITY

Two main approaches have been used to identify genes and molecular markers of trainability. The first approach, and the one that has been most often used in the field of exercise genomics, is the candidate gene approach. In candidate gene associations studies, a variant (most often a single-nucleotide polymorphism [SNP]) in a candidate gene, selected based on its physiological relevance to the trait of interest, is tested for

association with the trait under study. This can be done by comparing allele or genotype frequencies between groups of subjects (e.g., high- versus low-responders to exercise) or, for continuous traits, by comparing mean trait values across genotype groups or between carriers and noncarriers of a specific allele of a genetic variant. The major weakness of the candidate gene approach is that gene selection is limited by our understanding of the physiology underlying the trait under investigation and is thus far from being optimal for identifying all relevant genes. Improvements in microarray-based high-throughput technologies have made it possible to assay hundreds of thousands of DNA sequence variants in a single assay to quantify the expression level of thousands of transcripts simultaneously or analyze genome-wide DNA methylation and, as a result, to use genome-wide screening approaches to identify genomic markers of trainability (Bouchard, 2012). As a result of these improvements, genome-wide association studies (GWAS) have become the method of choice for identifying genes contributing to complex traits and have been used to identify various genes associated with trainability. Generally, 500,000 to one million variants are used in GWAS and tested for association with the phenotype of interest, yielding higher-resolution mapping than genome-wide linkage analyses, which are typically performed with a few hundred genetic markers (most microsatellites) and lead to the identification of chromosomal regions or loci referred to as quantitative trait loci (QTLs). Global gene expression profiling has also been used to identify genes associated with trainability. In this approach, genome-wide RNA expression profiles in a tissue are compared before and after exercise, to identify transcriptomic predictors of the response to exercise training. DNA methylation is a type of epigenetic modification that has an important role in regulating gene expression. Changes in global DNA methylation levels after exercise training have also been used to identify epigenetic markers of trainability.

Candidate Gene Studies

Several genetic variants have been shown to be associated with the response to exercise training (Bouchard, 2012; Rankinen et al., 2011). The last update of the human gene map for performance and health-related fitness phenotypes provided a detailed list of candidate genes associated with the trainability of various health-related fitness traits (Bray et al., 2009). A total of 49 unique genes from 81 studies were reported to be associated with training-induced changes of hemodynamic, body composition, insulin and glucose metabolism and plasma lipid, lipoprotein, and hemostatic phenotypes. However, most of these associations were not

replicated and cannot be considered valid. Only five genes were associated with the trainability of various cardiorespiratory fitness and performance phenotypes in more than one candidate gene study and are summarized in Table 24.1. These studies are briefly reviewed subsequently.

The *ACE* gene encodes the angiotensin-converting enzyme, a peptidyl dipeptidase enzymatic component of the renin-angiotensin system with an important role in regulating blood pressure. An insertion/deletion (I/D) polymorphism located in intron 16 of the gene and linked with variations in plasma and ACE tissue activity has been found to be associated with training-induced left ventricular (LV) growth, as assessed by echocardiography in two studies (Montgomery et al., 1997; Myerson et al., 2001). In the first study, the *ACE* D allele was associated with greater increases in LV mass and septal and posterior wall thickness after 10 weeks of physical training in British Army recruits (Montgomery et al., 1997). A few years later, in a study performed in another cohort of army recruits, the same group reported that the training-induced increase in LV mass was 2.7 times greater in recruits with the D/D genotype compared with the I/I homozygotes (Myerson et al., 2001). Several studies tested the hypothesis of an association between the *ACE* I/D polymorphism and $\text{VO}_{2\text{max}}$ trainability and found no evidence of an association (Bouchard, 2012). A study investigating the association between the *ACE* I/D polymorphism and changes in physical function among mobility-limited older adults after a 12-month physical activity program found that exercise training improved gait speed and the overall score from a battery of tasks designed to examine lower-extremity function in carriers of the D allele, but not in I/I homozygotes (Buford et al., 2014). A polymorphism in another gene of the renin-angiotensin system, the Met235Thr polymorphism of the angiotensinogen (*AGT*) gene, was associated with blood pressure response to exercise training in two different studies. In the HERITAGE Family Study, *AGT* M235M subjects had the greatest reduction in submaximal exercise diastolic blood pressure (Rankinen et al., 2000), whereas in middle-aged Finnish men undergoing a 6-year exercise intervention trial, M235M homozygotes were those exhibiting the most favorable changes in resting systolic and diastolic blood pressure (Rauramaa et al., 2002).

Other genes have been considered candidates for trainability of cardiorespiratory fitness. One is the apolipoprotein E (*APOE*) gene, which encodes a 229–amino acid protein that has three isoforms (apo E2, E3, and E4). apo E3 is encoded by the combination of rs429358T and rs7412T alleles. The association between *APOE* alleles and the trainability of $\text{VO}_{2\text{max}}$ was investigated in four studies. In the HERITAGE Family study, no evidence of an association was found (Leon et al., 2004), whereas

in three other studies, significant associations were reported, but with discordant results. In one study, carriers of the encoding E2 allele had the lowest training response (Hagberg et al., 1999), whereas in another, E2 allele carriers exhibited the greatest improvements (Thompson et al., 2004). A study of Chinese adult subjects aged 18–40 years found that after 6 months of endurance training, $\text{VO}_{2\text{max}}$ improved significantly, but only in E2/E3 and E3/E4 subjects and not in those with other genotypes (Yu et al., 2014). More recently, a study found that improvements in lipid profile of dyslipidemic obese adolescents after physical exercise and diet for 28 days were associated with the *APOE* polymorphism; improvements were observed in subjects with the E2 and E3 alleles, but not in those with the E2 allele (Gultom et al., 2015).

Another genetic variant associated with the response to exercise training is the Gly482Ser polymorphism of the peroxisome proliferator-activated receptor gamma coactivator 1 α (*PPARGC1A*) gene, which has an important role in regulating mitochondrial biogenesis and enzymes involved in oxidative phosphorylation. In a 9-month lifestyle intervention program involving diet and physical activity, a lower increase in anaerobic threshold and insulin sensitivity was observed in carriers of the Ser482 allele compared with noncarriers in response to the intervention (Stefan et al., 2007). In a study undertaken in 28 male subjects aged 59 ± 7 years submitted to supervised 10-week cycling training, changes in skeletal muscle structure and mitochondrial function (assessed from analyses of biopsies from the vastus lateralis muscle) were compared between Gly482Gly subjects ($n = 13$) and carriers of the Ser482 allele ($n = 15$) (Steinbacher et al., 2015). The proportion of slow-twitch muscle fiber increased significantly in response to training, but only in Gly482Gly subjects, which exhibited a relative increase of 19% ($P < .01$); no changes were observed in carriers of the Ser482 allele. All other muscle characteristics (capillary supply, mitochondrial density, mitochondrial enzyme activities, and intramyocellular lipid content) increased similarly in both groups. The authors concluded that the *PPARGC1A* Gly482Ser variant impairs the exercise-induced fast-to-slow muscle fiber transformation generally associated with endurance training (Steinbacher et al., 2015). A study reported that improvements in lipid profile after a 12-week bench-stepping exercise training at lactate threshold intensity in Japanese subjects aged more than 65 years were significant in Gly482Gly subjects but not in carriers of the Ser482 allele (Tobina et al., 2016). The results of these three studies suggest that the Ser482 variant of *PPARGC1A* gene is associated with less favorable changes in response to exercise training.

Another candidate gene of interest for exercise-related traits is the α -actinin-3 (*ACTN3*) gene. Because

TABLE 24.1 Candidate Genes Associated With Response to Exercise Training in At Least Two Different Studies.

Gene	Variant	Chr.	Position (Base Pair)	Participants	Exercise Protocol	Trait	Reference
ACE	Insertion/deletion	16	61 565 900	140 British Army recruits	10-week strength and endurance training	Left ventricular mass	Montgomery et al. (1997)
				141 British Army recruits	10-week strength and endurance training	Left ventricular mass	Myerson et al. (2001)
				283 mobility-limited elderly subjects	12-month walking, strength, flexibility, and balance training	Gait speed, lower-extremity function	Buford et al. (2014)
AGT	Met235Thr	1	230 845 794	476 subjects from HERITAGE	20-week cycling endurance training	Submaximal exercise diastolic blood pressure	Rankinen et al. (2000)
				140 middle-aged men	6-year physical training	Resting systolic and diastolic blood pressure	Rauramaa et al. (2002)
APOE	E3/E4	19	45 411 941	51 overweight men	9-month endurance training	VO _{2max} , HDL-C and HDL2-C	Hagberg et al. (1999)
				120 men and women	Not available	VO _{2max} , TC/HDL-C ratio, low-density-lipoprotein/HDL ratio	Thompson et al. (2004)
				360 men and women	6-month endurance training	VO _{2max}	Yu et al. (2014)
				60 dyslipidemic obese adolescents	28 days diet and physical exercise	Lipid profile	Gultom et al. (2015)
PPARGC1A	Gly482Ser	4	23 815 662	136 men and women	9-month diet and moderate endurance exercise	Anaerobic threshold and insulin sensitivity	Stefan et al. (2007)
				28 male subjects	10-week cycling endurance training	Proportion of slow-twitch muscle fiber	Steinbacher et al. (2015)
				119 subjects over 65 years	12-week bench-stepping exercise	Total cholesterol (TC) and low-density lipoprotein cholesterol (LDL-C)	Tobina et al. (2016)
ACTN3	R577X	11	66 328 095	602 men and women	12-week elbow flexor/extensor resistance training	Dynamic strength	Clarkson et al. (2005)
				157 elderly men and women	10-week knee extensor strength training	Knee extensor concentric peak power	Delmonico et al. (2007)
				206 police recruits	18-week running endurance training	VO _{2max}	Silva et al. (2015)

HDL, high-density lipoprotein; VO_{2max}, maximal oxygen uptake.

of a C/T transition in exon 16 of the gene that results in the replacement of an arginine residue (R) with a stop codon (X) at amino acid 577, a common variation (R577X) in the *ACTN3* gene abolishes the synthesis of functional *ACTN3* protein. Although several studies provided evidence of an association between the *ACTN3* R577X variant and physical performance in humans (McArthur and North, 2011), only a few studies investigated the role of this variant in response to exercise training. Two studies reported an association between the *ACTN3* R577X variant and changes in muscle strength in response to strength training, but they found opposite results. In one study involving young adult men and women enrolled in a 12-week standardized elbow flexor-extensor resistance training program, women homozygotes for the mutant allele (577X) had greater increases in dynamic (one-repetition maximum) strength compared with the homozygous wild-type (RR) women, whereas no association was observed in men (Clarkson et al., 2005). Another study in elderly men and women involved in a 10-week unilateral knee extensor strength training program found greater improvements of muscle strength in R577R women compared with X577X homozygotes (Delmonico et al., 2007). A third study investigating the association between the *ACTN3* R577X variant and peak VO_{2max} in police recruits who underwent an 18-week running endurance training program reported a greater increase in relative peak VO_{2max} in R577R compared with X577X individuals (12% versus 6%; $P = .02$) (Silva et al., 2015).

Genome-Wide Screening Studies

Two genome-wide linkage analyses (Rankinen et al., 2011) were performed in the HERITAGE Family Study to identify genes for the response to exercise training. QTLs for training-induced changes in submaximal exercise (50 W) stroke volume (SV50) and heart rate (HR50) were found on chromosomes 10p11 and 2q33.3-q34, respectively. Fine mapping of the region of the QTL on 10p11 for changes in SV50 revealed evidence of an association with SNPs located in the kinesin family member 5B (*KIF5B*) gene. Kinesins are a family of microtubule-based motor proteins involved in transporting organelles in eukaryotic cells and the *KIF5B* gene encodes a common isoform of kinesin expressed in several tissues, including cardiac myocytes. The QTL for exercise-induced changes in HR50 was localized in a region containing the cyclic adenosine monophosphate-responsive element binding protein 1 (*CREB1*) gene, an abundantly expressed regulator of gene expression involved in regulating several physiological functions. A genetic variant located in the 5'-region of the *CREB1* (rs2253206) gene was associated

with smaller improvements of HR50 in response to endurance training.

Four GWAS of trainability were performed and led to the identification of several variants associated with trainability (Table 24.2). Three of these GWAS were performed in the HERITAGE Family Study and involved changes in VO_{2max} (Bouchard et al., 2011b), HR50 (Rankinen et al., 2012), and triglycerides (TG) (Sarzynski et al., 2015) in response to a standardized 20-week endurance training program. The first GWAS led to the identification of 39 SNPs associated with a VO_{2max} training response. The strongest evidence of an association was observed with an SNP (rs6552828; $P = 1.31 \times 10^{-6}$) located in the first intron of the acyl-CoA synthetase long-chain family member 1 gene (Bouchard et al., 2011b). When all 39 SNPs were analyzed simultaneously in a regression model, retaining only SNPs explaining at least 1% of the trait variance, 16 SNPs (see Table 24.2) collectively accounted for 45% of variance in the trainability of VO_{2max} (Rankinen and Bouchard, 2012). In another GWAS, a total of 41 SNPs showed significant (P -values $< 9.9 \times 10^{-5}$) associations with the HR50 response. The strongest association was observed with an SNP (rs6432018; $P = 8.08 \times 10^{-7}$) located near the tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, theta polypeptide gene encoding a member of a family of signaling proteins involved in apoptosis, cell proliferation, and metabolism. In a multivariate regression model, nine SNPs (Table 24.2) with 34% of trait variance explained the genetic variability of the response (Rankinen et al., 2012). In both GWAS, a genetic score was constructed using the most informative SNPs from the regression models by summing the number of positive training response (increase in VO_{2max} and decrease in HR50) alleles. The differences in VO_{2max} and HR50 training responses between subjects with the lowest and highest genetic score categories were 383 mL/min and 21 beats/min, respectively (Rankinen and Bouchard, 2012). The third GWAS identified 39 SNPs associated ($P < 1 \times 10^{-4}$) with TG changes in response to exercise (Sarzynski et al., 2015). In a multivariate regression model, the top 10 SNPs explained 32% of variance in the training-induced TG changes, but the four top SNPs (Table 24.2) accounting for 17% of variance were sufficient to account for all of the genetic variance (heritability) of the TG response to exercise. More recently, a GWAS was performed in 79 adult subjects to identify genetic variants associated with training-induced changes VO_{2max} and knee peak torque after a 9-week high-intensity training program (Yoo et al., 2016). The authors identified seven SNPs explaining 26% of variance in percent changes of VO_{2max} in response to training. For the knee peak torque, six

TABLE 24.2 Genetic Variants Associated With Response to Exercise Training Identified From Genome-wide Association Studies.

Study	Trait	Single-nucleotide Polymorphism	Chromosome	Position (Base Pair)	Gene
Bouchard et al. (2011b)	VO _{2max}	rs10499043	6	106 247 137	PRDM1 (287 kb)
		rs1535628	9	105 016 749	GRIN3A (516 kb)
		rs4973706	3	18 921 772	KCNH8 (268 kb)
		rs12115454	9	118 720 050	LINC00474 (70 kb)
		rs6552828	4	185 725 416	ACSL1
		rs11715829	3	146 957 166	ZIC4 (146 kb)
		rs884736	1	7 015 105	CAMTA1
		rs10921078	1	192 059 022	RGS18 (69 kb)
		rs6090314	20	61 857 552	BIRC7 (10 kb)
		rs10500872	11	20 245 723	DBX1 (64 kb)
		rs1956197	14	59 477 414	DAAM1 (178 kb)
		rs824205	15	24 008 071	NDN (75 kb)
		rs7933007	11	118 730 669	CXCR5 (24 kb)
		rs12896790	14	38 273 922	TTC6 (55 kb)
		rs4952535	2	42 131 523	LOC388942 (10 kb)
rs2053896	4	137 154 796	TERF1P3 (120 kb)		
Rankinen et al. (2012)	HR50	rs2979481	8	30 262 786	RBPMS
		rs6432018	2	9 721 896	YWHAHQ (2 kb)
		rs2253206	2	208 381 978	CREB1 (13 kb)
		rs1560488	4	90 225 835	GPRIN3
		rs10248479	7	115 606 355	TFEC
		rs857838	1	158 750 550	OR6N2 (3 kb)
		rs909562	6	16 130 333	MYLIP (1 kb)
		rs4759659	12	130 837 288	PIWIL1
Yoo et al. (2016)	VO _{2max}	rs2057368	14	55 304 009	GCH1 (5 kb)
		rs11051548	12	31 877 289	AMN1
		rs2542729	18	28 148 324	CDH2 (383 kb)
		rs1451462	2	53 123 412	LOC730100 (138 kb)
		rs13060995	3	9 224 103	SRGAP3 (202 kb)
		rs6570913	6	149 266 798	UST
		rs11096663	2	20 620 093	RHOB (27 kb)
	rs12613181	2	163 513 04	KCNH7	
	Knee peak torque	rs10072841	5	150 952 884	FAT2 (4 kb)
		rs6564267	16	75 603 925	GABARAPL2
		rs17044554	2	23 218 255	LOC101929251 (30 kb)
rs1341439		13	104 358 672	SLC10A2 (662 kb)	
rs4522375	15	24 202 687	PWRN2 (207 kb)		

TABLE 24.2 Genetic Variants Associated With Response to Exercise Training Identified From Genome-wide Association Studies.—cont'd

Study	Trait	Single-nucleotide Polymorphism	Chromosome	Position (Base Pair)	Gene
Sarzynski et al. (2015)	TG	rs7154161	14	51 854 268	LINC00640 (22 kb)
		rs222158	21	27 782 161	CYYR1
		rs2722171	12	104 449 487	GLT8D2
		rs1906058	16	6 144 648	RBFOX1
		rs2593324	3	22 119 221	ZNF385D

HR50, heart rate at submaximal (50 W) exercise; TG, triglycerides, VO_{2max} , maximal oxygen uptake.

Positions are from GRCh37.p13 Assembly from Phase 3 1000 Genomes Browser. Gene indicates gene located nearest to the variant. Distance to the gene in kilobases (1000 base pairs) is given in parentheses. If no distance is shown, the variant is within the gene itself.

SNPs were identified and accounted for 27% of variance in the relative increase in muscle strength.

A few studies used global gene expression profiling to identify genes associated with the VO_{2max} (21 genes) and TG (11 genes) responses to endurance exercise (Table 24.3). In the first study, Timmons and colleagues used a combination of skeletal muscle gene expression profiling and DNA markers to identify genes associated with the trainability of VO_{2max} (Timmons et al., 2010). Two independent RNA expression profiles of skeletal muscle samples were generated from subjects who underwent supervised endurance training: one ($n = 41$ gene chips) identified and the second ($n = 17$ gene chips) blindly validated a 29-RNA signature that predicted changes in VO_{2max} . Next, tagging SNPs for the 29 predictor transcripts were identified and genotyped in HERITAGE. A regression analysis combining transcriptome-derived SNPs and SNPs previously associated with the trainability of VO_{2max} in HERITAGE led to the identification of a set of 11 genetic variants explaining 23% of variance in the VO_{2max} training response. In second study, a peripheral blood mononuclear cell (PBMC) expression signature was generated from 13 healthy policemen recruits undergoing an 18-week running endurance training and used to identify molecular classifiers of trainability of VO_{2max} (Dias et al., 2015). Baseline PBMC expression patterns enabled the identification of a 98-transcript signature associated with absolute changes in VO_{2max} . None of these PBMCs transcripts were among those reported by Timmons and colleagues in skeletal muscle (Timmons et al., 2010). A smaller panel of 10 transcripts showing a strong association with the VO_{2max} response was able to classify the 13 subjects according to their potential to improve VO_{2max} (Dias et al., 2015). A third study based on HERITAGE used GWAS (results described earlier) and gene expression profiling from skeletal muscle to identify genetic variants associated with the TG response to exercise training (Sarzynski et al., 2015). A molecular

signature based on the baseline expression of 11 genes predicted 27% of TG changes in response to exercise training. Association analyses of SNPs in or near these 11 predictor genes and TG changes revealed that only SNPs from four genes (SLC13A4, FASTK, MACROD1, and EEF2K) were associated with TG changes. A composite score based on the top four SNPs, each from the genomic (rs222158, rs2722171, rs1906058, and rs2593324) and transcriptomic (rs1043968, rs379336, rs594461, and rs11646610) analyses, incorporated into a regression model including over 50 baseline variables, was the strongest predictor variable ($P = 3.0 \times 10^{-68}$), explaining 14.4% of variance in the TG response to exercise training (Sarzynski et al., 2015).

Epigenetic Markers of Trainability

DNA methylation is an epigenetic mark resulting from the incorporation of a methyl group in cytosine-guanine base pairing (CpG), affecting the expression of specific gene transcripts. As reviewed in detail elsewhere (Voisin et al., 2015), there is evidence that both acute and chronic exercises significantly affect DNA methylation in a highly tissue- and gene-specific manner. Results from intervention studies (Voisin et al., 2015) seem to agree that exercise training has an impact on the methylation of many genes involved in metabolism and fuel use, but in almost all studies, the exercise-induced methylation changes did not correlate with the expected changes in gene expression. A study showed that a 7-week strength training program that significantly increased the lower limb muscle mass of eight untrained male subjects was associated with large alterations in the DNA methylation pattern of the skeletal muscle (17,365 differentially modified CpG sites) (Seaborne et al., 2018). Interestingly, after a period of detraining in which muscle mass returned to baseline levels and a subsequent period of training resulted in a larger increase in muscle mass compared with the

TABLE 24.3 Genes Associated With Response to Exercise Training Identified From Transcriptome-wide Profiling.

Study	Trait	Tissue	Gene	Genomic Location
Timmons et al. (2010)	VO _{2max}	Muscle	SVIL	Chr10: 29,746,274–30,025,864
			SLC22A3	Chr6: 160,769,405–160,873,613
			NRP2	Chr2: 206,547,184–206,662,857
			TTN	Chr2: 179,390,716–179,672,150
			H19	Chr11: 2,016,406–2,019,065
			ID3	Chr1: 23,884,421–23,886,285
			MIPEP	Chr13: 24,3034,328–24,463,587
			CPVL	Chr7: 29,035,247–29,226,213
			DEPDC6	Chr8: 120,885,900–121,063,157
			BTAF1	Chr10: 93,683,526–93,790,755
Dias et al. (2015)	VO _{2max}	Peripheral blood mononuclear cells	DIS3L	Chr15: 66,585,633–66,626,236
			C4orf46	Chr4: 159,587,827–159,593,407
			CCDC50	Chr3: 191,046,874–191,116,459
			CDK14	Chr7: 90,225,681–90,839,906
			FAAH2	ChrX: 57,313,011–57,515,620
			HIST1H2BF	Chr6: 26,199,787–26,200,216
			IGLJ3	Ch22: 23,247,168–23,247,205
			PIP5K1B	Chr9: 71,320,188–71,624,092
			SCARNA9L	ChrX: 20,154,184–20,154,531
			SNORA48	Chr17: 7,478,031–7,478,165
Sarzynski et al. (2015)	TG	Muscle	TCF4	Chr18: 52,889,562–53,303,252
			DNAAF4	Chr15: 55,709,953–55,800,432
			ZNF30	Chr19: 35,417,807–35,436,076
			BTG2	Chr1: 203,274,664–203,278,730
			MACROD1	Chr11: 63,766,006–63,933,602
			UBE2L3	Chr22: 21,903,736–21,978,323
			B3GALT5-AS1	Chr21: 40,969,074–40,984,749
			EEF2K	Chr16: 22,217,592–22,300,066
			NCBP2	Chr3: 196,662,273–196,669,464
			FASTK	Chr7: 150,773,708–150,777,970
			C2orf69	Chr2: 200,775,979–200,792,996
			SLC13A4	Chr7: 135,365,979–135,412,943

TG, triglyceride; VO_{2max}, maximal oxygen uptake.

initial training, methylation changes were more important (27,155 differentially modified CpG sites) with a higher frequency of hypomethylation (69% compared with 53% after the initial training). The largest increase in hypomethylation after retraining suggests that human skeletal muscle may possess an epigenetic memory

of hypertrophy facilitating better response to subsequent training (Seaborne et al., 2018).

MicroRNAs (miRNAs) are another class epigenetic markers that could serve as biomarkers of adaptation to exercise. miRNAs are noncoding RNAs that regulate gene expression via degradation of messenger RNA or

inhibition of translation and that can be released in the circulation (Xu et al., 2015). Although several miRNAs were found to be responsive to acute exercise, their role in trainability remains to be determined. Only one study reported a strong correlation between changes in miR-20A and changes in VO_{2max} ($r = 0.73$; $P = .02$) in 10 male rowers after a 3-month aerobic rowing exercise training, which suggests that circulating miR-20a could be a biomarker of trainability in cardiorespiratory fitness (Baggish et al., 2011).

CONCLUSION

This short review emphasized the importance of individual differences in adapting to regular exercise training and provided evidence that genetic factors have an important role in explaining these differences. Based on data from twin and family studies, the heritability of trainability ranges from about 20% to 50%, depending on the phenotype. Although progress has been made in identifying genes underlying trainability, much remain to be done. Only five genetic variants from candidate gene studies have been associated with trainability with evidence of replication. Results from GWAS, global gene expression profiling, and genome-wide DNA methylation yielded more genomic markers that have the ability to predict trainability. Finally, one miRNA has been associated with trainability of cardiorespiratory fitness. Although the results are promising, more research is needed to unravel the molecular basis of trainability and to be able to use genetic markers to predict performance or discriminate between high- and low-responders to exercise training.

References

- Baggish, A.L., Hale, A., Weiner, R.B., Lewis, G.D., Systrom, D., Wang, F., Wang, T.J., Chan, S.Y., 2011. Dynamic regulation of circulating microRNA during acute exhaustive exercise and sustained aerobic exercise training. *J Physiol* 589, 3983–3994.
- Bouchard, C., 2012. Genomic predictors of trainability. *Exp Physiol* 97, 347–352.
- Bouchard, C., Antunes-Correa, L.M., Ashley, E.A., Franklin, N., Hwang, P.M., Mattsson, C.M., Negrao, C.E., Phillips, S.A., Sarzynski, M.A., Wang, P.Y., Wheeler, M.T., 2015. Personalized preventive medicine: genetics and the response to regular exercise in preventive interventions. *Prog Cardiovasc Dis* 57, 337–346.
- Bouchard, C., Rankinen, T., Timmons, J.A., 2011. Genomics and genetics in the biology of adaptation to exercise. *Comp Physiol* 1, 1603–1648.
- Bouchard, C., Sarzynski, M.A., Rice, T.K., Kraus, W.E., Church, T.S., Sung, Y.J., Rao, D.C., Rankinen, T., 2011. Genomic predictors of the maximal O_2 uptake response to standardized exercise training programs. *J Appl Physiol* 110, 1160–1170.
- Bray, M.S., Hagberg, J.M., Perusse, L., Rankinen, T., Roth, S.M., Wolfarth, B., Bouchard, C., 2009. The human gene map for performance and health-related fitness phenotypes: the 2006–2007 update. *Med Sci Sport Exerc* 41, 35–73.
- Buford, T.W., Hsu, F.C., Brinkley, T.E., Carter, C.S., Church, T.S., Dodson, J.A., Goodpaster, B.H., McDermott, M.M., Nicklas, B.J., Yank, V., Johnson, J.A., Pahor, M., Group, L.R., 2014. Genetic influence on exercise-induced changes in physical function among mobility-limited older adults. *Physiol Genom* 46, 149–158.
- Clarkson, P.M., Devaney, J.M., Gordish-Dressman, H., Thompson, P.D., Hubal, M.J., Urso, M., Price, T.B., Angelopoulos, T.J., Gordon, P.M., Moyna, N.M., Pescatello, L.S., Visich, P.S., Zoeller, R.F., Seip, R.L., Hoffman, E.P., 2005. ACTN3 genotype is associated with increases in muscle strength in response to resistance training in women. *J Appl Physiol* (1985) 99, 154–163.
- Delmonico, M.J., Kostek, M.C., Doldo, N.A., Hand, B.D., Walsh, S., Conway, J.M., Carignan, C.R., Roth, S.M., Hurley, B.F., 2007. Alpha-actinin-3 (ACTN3) R577X polymorphism influences knee extensor peak power response to strength training in older men and women. *J Gerontol A Biol Sci Med Sci* 62, 206–212.
- Dias, R.G., Silva, M.S., Duarte, N.E., Bolani, W., Alves, C.R., Junior, J.R., da Silva, J.L., de Oliveira, P.A., Alves, G.B., de Oliveira, E.M., Rocha, C.S., Marsiglia, J.D., Negrao, C.E., Krieger, E.M., Krieger, J.E., Pereira, A.C., 2015. PBMCS express a transcriptome signature predictor of oxygen uptake responsiveness to endurance exercise training in men. *Physiol Genom* 47, 13–23.
- Gultom, L.C., Sjarif, D.R., Sudoyo, H.A., Mansyur, M., Hadinegoro, S.R., Immanuel, S., Setiawati, M., 2015. The role of apolipoprotein E polymorphism in improving dyslipidemia in obese adolescents following physical exercise and National Cholesterol Education Program Step II intervention. *J Pediatr Endocrinol Metab* 28, 597–603.
- Hagberg, J.M., Ferrell, R.E., Katznel, L.I., Dengel, D.R., Sorkin, J.D., Goldberg, A.P., 1999. Apolipoprotein E genotype and exercise training-induced increases in plasma high-density lipoprotein (HDL)- and HDL2-cholesterol levels in overweight men. *Metabolism* 48, 943–945.
- Leon, A.S., Togashi, K., Rankinen, T., Despres, J.P., Rao, D.C., Skinner, J.S., Wilmore, J.H., Bouchard, C., 2004. Association of apolipoprotein E polymorphism with blood lipids and maximal oxygen uptake in the sedentary state and after exercise training in the HERITAGE family study. *Metabolism* 53, 108–116.
- McArthur, D.G., North, K., 2011. The ACTN3 gene and human performance. In: Bouchard, C., Hoffman, E.P. (Eds.), *Genetics and Molecular Aspects of Sport Performance*. Wiley-Blackwell, Chichester, West Sussex, UK, pp. 204–214.
- Montgomery, H.E., Clarkson, P., Dollery, C.M., Prasad, K., Losi, M.A., Hemingway, H., Statters, D., Jubb, M., Girvain, M., Varnava, A., World, M., Deanfield, J., Talmud, P., McEwan, J.R., McKenna, W.J., Humphries, S., 1997. Association of angiotensin-converting enzyme gene I/D polymorphism with change in left ventricular mass in response to physical training. *Circulation* 96, 741–747.
- Myerson, S.G., Montgomery, H.E., Whittingham, M., Jubb, M., World, M.J., Humphries, S.E., Pennell, D.J., 2001. Left ventricular hypertrophy with exercise and ACE gene insertion/deletion polymorphism: a randomized controlled trial with losartan. *Circulation* 103, 226–230.
- Rankinen, T., Bouchard, C., 2012. Gene-exercise interactions. *Prog Mol Biol Transl Sci* 108, 447–460.
- Rankinen, T., Gagnon, J., Pérusse, L., Chagnon, Y.C., Rice, T., Leon, A.S., Skinner, J.S., Wilmore, J.H., Rao, D.C., Bouchard, C., 2000. AGT M235T and ACE ID polymorphisms and exercise blood pressure in the HERITAGE Family Study. *Am J Physiol Heart Circ Physiol* 279, H368–H374.
- Rankinen, T., Sarzynski, M.A., Bouchard, C., 2011. Genes and response to training. In: Bouchard, C., Hoffman, E.P. (Eds.), *Genetics and Molecular Aspects of Sport Performance*. Wiley-Blackwell, Chichester, West Sussex, UK, pp. 177–184.

- Rankinen, T., Sung, Y.J., Sarzynski, M.A., Rice, T.K., Rao, D.C., Bouchard, C., 2012. Heritability of submaximal exercise heart rate response to exercise training is accounted for by nine SNPs. *J Appl Physiol* (1985) 112, 892–897.
- Rauramaa, R., Kuhanen, R., Lakka, T.A., Vaisanen, S.B., Halonen, P., Alen, M., Rankinen, T., Bouchard, C., 2002. Physical exercise and blood pressure with reference to the angiotensinogen M235T polymorphism. *Physiol Genom* 10, 71–77.
- Sarzynski, M.A., Davidsen, P.K., Sung, Y.J., Hesselink, M.K., Schrauwen, P., Rice, T.K., Rao, D.C., Falciani, F., Bouchard, C., 2015. Genomic and transcriptomic predictors of triglyceride response to regular exercise. *Br J Sports Med* 49, 1524–1531.
- Sarzynski, M.A., Rankinen, T., Bouchard, C., 2011. Twin and family studies of training responses. In: Bouchard, C., Hoffman, E.P. (Eds.), *Genetics and Molecular Aspects of Sport Performance*. Wiley-Blackwell, Chichester, West Sussex, UK, pp. 110–120.
- Seaborne, R.A., Strauss, J., Cocks, M., Shepherd, S., O'Brien, T.D., van Someren, K.A., Bell, P.G., Murgatroyd, C., Morton, J.P., Stewart, C.E., Sharples, A.P., 2018. Human skeletal muscle possesses an epigenetic memory of hypertrophy. *Sci Rep* 8, 1898.
- Silva, M.S., Bolani, W., Alves, C.R., Biagi, D.G., Lemos Jr., J.R., da Silva, J.L., de Oliveira, P.A., Alves, G.B., de Oliveira, E.M., Negrao, C.E., Krieger, J.E., Dias, R.G., Pereira, A.C., 2015. Elimination of influences of the ACTN3 R577X variant on oxygen uptake by endurance training in healthy individuals. *Int J Sports Physiol Perform* 10, 636–641.
- Stefan, N., Thamer, C., Staiger, H., Machicao, F., Machann, J., Schick, F., Venter, C., Niess, A., Laakso, M., Fritsche, A., Haring, H.U., 2007. Genetic variations in PPAR α and PPAR γ 1A determine mitochondrial function and change in aerobic physical fitness and insulin sensitivity during lifestyle intervention. *J Clin Endocrinol Metab* 92, 1827–1833.
- Steinbacher, P., Feichtinger, R.G., Kedenko, L., Kedenko, I., Reinhardt, S., Schonauer, A.L., Leitner, I., Sanger, A.M., Stoiber, W., Kofler, B., Forster, H., Paulweber, B., Ring-Dimitriou, S., 2015. The single nucleotide polymorphism Gly482Ser in the PGC-1 α gene impairs exercise-induced slow-twitch muscle fibre transformation in humans. *PLoS One* 10, e0123881.
- Thompson, P.D., Tsongalis, G.J., Seip, R.L., Bilbie, C., Miles, M., Zoeller, R., Visich, P., Gordon, P., Angelopoulos, T.J., Pescatello, L., Bausserman, L., Moyna, N., 2004. Apolipoprotein E genotype and changes in serum lipids and maximal oxygen uptake with exercise training. *Metabolism* 53, 193–202.
- Timmons, J.A., Knudsen, S., Rankinen, T., Koch, L.G., Sarzynski, M., Jensen, T., Keller, P., Scheele, C., Volvaard, N.B., Nielsen, S., Akerstrom, T., MacDougald, O.A., Jansson, E., Greenhaff, P.L., Tarnopolsky, M.A., van Loon, L.J., Pedersen, B.K., Sundberg, C.J., Wahlestedt, C., Britton, S.L., Bouchard, C., 2010. Using molecular classification to predict gains in maximal aerobic capacity following endurance exercise training in humans. *J Appl Physiol* 108, 1487–1496.
- Tobina, T., Mori, Y., Doi, Y., Nakayama, F., Kiyonaga, A., Tanaka, H., 2016. Peroxisome proliferator-activated receptor gamma co-activator 1 gene Gly482Ser polymorphism is associated with the response of low-density lipoprotein cholesterol concentrations to exercise training in elderly Japanese. *J Physiol Sci*.
- Voisin, S., Eynon, N., Yan, X., Bishop, D.J., 2015. Exercise training and DNA methylation in humans. *Acta Physiol (Oxf)* 213, 39–59.
- Xu, T., Liu, Q., Yao, J., Dai, Y., Wang, H., Xiao, J., 2015. Circulating microRNAs in response to exercise. *Scand J Med Sci Sports* 25, e149–154.
- Yoo, J., Kim, B.H., Kim, S.H., Kim, Y., Yim, S.V., 2016. Genetic polymorphisms to predict gains in maximal O₂ uptake and knee peak torque after a high intensity training program in humans. *Eur J Appl Physiol* 116, 947–957.
- Yu, B., Chen, W., Wang, R., Qi, Q., Li, K., Zhang, W., Wang, H., 2014. Association of apolipoprotein E polymorphism with maximal oxygen uptake after exercise training: a study of Chinese young adult. *Lipids Health Dis* 13, 40.
- Zadro, J.R., Shirley, D., Andrade, T.B., Scurrah, K.J., Bauman, A., Ferreira, P.H., 2017. The beneficial effects of physical activity: is it down to your genes? A systematic review and meta-analysis of twin and family studies. *Sports Med Open* 3, 4.

Genetic Variations Impacting the Response to Defined Diets

Tiange Wang^{1,2}, Lu Qi^{2,3}

¹Shanghai Institute of Endocrine and Metabolic Diseases, Rui Jin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; ²Department of Epidemiology, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA, United States; ³Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, United States

Glossary

Genetic variation The difference in DNA sequences among individuals within a population.

Gene–diet interaction Individuals carrying different *genotypes* respond to dietary factors in different ways.

Genetic predisposition A genetic characteristic that influences the possible phenotypic development of an individual organism within a species or population under the influence of environmental conditions.

Genetic risk score A sum of risk alleles of genetic variants in multiple loci and their associated weights.

INTRODUCTION

Over the past few decades, there has been a rapid increase in obesity and related metabolic diseases such as type 2 diabetes and cardiovascular disease. The global epidemic of obesity is believed to be substantially attributed to a recent profound shift in our living environment, such as unhealthy dietary patterns and a sedentary lifestyle that leads to increased energy intake and reduced energy expenditures, and their interactions with genetic susceptibility (Qi and Cho, 2008).

The etiology of obesity is imbalance between dietary energy intake and energy expenditure. Diet is among the most important modifiable factors affecting energy balance and regulation of body weight, together with physical inactivity and genetic predisposition. Many foods may tip the balance of energy input and output. There is growing evidence that consumption of unhealthy foods such as refined carbohydrates, fried foods, and sugary drinks is related to the increased risk for obesity, whereas intake of fruits, vegetables, whole

grains, fish, and legumes, as well as healthy eating patterns such as the Mediterranean diet, may prevent weight gain or promote weight loss. As mainstream efforts struggling against the rapid rise of obesity and its complications, a variety of diet interventions have been proposed to improve weight loss and weight maintenance. Although various diets generally are not significantly different in their effectiveness in weight loss, considerable interindividual heterogeneity has been noted in participants' response, and accumulating evidence suggests that genetic background may at least partly account for such variability (Qi and Cho, 2008). With revolutionary advances in genomic technologies, a large body of genome-wide association studies (GWAS) have identified numerous genetic variants associated with body weight and obesity risk. Interestingly, many of the identified obesity-related genes are highly expressed or act in the central nervous system, involved in regulating appetite and the balance of energy intake and expenditure, which provides a potential biological basis for interactions between genetic variations and diets. In addition, growing data from longitudinal studies and randomized diet intervention trials have shown that genetic variations may interact with dietary factors on obesity, weight change, and related metabolic traits.

This chapter summarizes advances in studies of the interaction between genetic variations and diet in relation to obesity, weight loss, and maintenance. The chapter particularly addresses the modification effects of genetic variations on diet interventions targeting weight management and discusses challenges in the field as well as implications for findings in precision nutrition.

EVOLUTIONARY BASIS FOR INTERACTIONS BETWEEN GENETIC VARIATIONS AND DIET ON BODY WEIGHT AND OBESITY

Traditional genetic studies have estimated that about 40%–60% of the variance in body weight can be explained by genetic components. In past few years, the application of GWAS has made significant advances in identifying specific genetic variants. Thus far, nearly 100 genomic loci containing variants affecting body mass index (BMI) have been detected, and more loci were connected to other obesity-related measures (Locke et al., 2015). Among the identified loci, *FTO* shows the strongest association. Interestingly, many obesity genes are highly expressed in brain, especially regions involved in regulation of appetite, food intake, and energy balance, which highlights a neuronal influence on body weight regulation and a potential biological basis for interactions between genetic variations and diets.

A principle assumption underlying traditional nutrition research is that disease risk conferred by dietary intake is uniform for each individual. Such an assumption appears to be accurate. Large-scale nutrition surveys have shown that shifts from a principally nutritionally dense diet to a more energy-dense one is among the driving forces responsible for the rapid increase in obesity in past decades, and several lines of evidence suggest that genetic variation may also have a role in shaping the epidemic of obesity. Genetic variants identified so far explain only a small proportion of heritability of body weight and obesity, suggesting so-called missing heritability; and such missing heritability may be at least partly caused by interactions between genetic variations and environmental factors such as diet. The interaction between genetic factors and diet has its origins in our evolutionary history. A presentation of gene–diet interactions should involve the concept of the thrifty genotype hypothesis, which was first proposed by the American geneticist James Neel in relation to an enhanced predisposition to type 2 diabetes; later, it was extended to obesity. This hypothesis proposed that widespread obesity in modern societies is a consequence of the mismatch of the makeup of thrifty genes, which would have been advantageous in promoting efficient fat deposition to survive in periods of famine early in history, but have become disadvantageous in modern societies owing to prepared fat deposition for a famine that never occurs. According to the thrifty genotype hypothesis, compared with those who do not carry the thrifty genes, individuals who do are more susceptible to obesity under the same environmental exposure. Since its publication, the thrifty genotype hypothesis has spawned many additional ideas supported by

many data. However, conflicting evidence exists, and more evidence is needed to validate this hypothesis further.

HABITUAL DIETARY INTAKES AND GENETIC PREDISPOSITION TO OBESITY

In the past few decades, potential interactions between genetic variations and various dietary factors have been extensively investigated (Qi and Cho, 2008). Reproducible evidence is emerging to show the interactions of body weight and obesity. In epidemiology studies, the consumption of sugar-sweetened beverages (SSB) has been consistently related to weight gain and an increased risk for obesity. We assessed the interactions of habitual intake of SSB and obesity genetic susceptibility, evaluated by a genetic risk score (GRS) derived from 32 obesity-associated loci, in relation to BMI and obesity risk (Fig. 25.1). We employed a two-stage design consisting of three prospective cohorts: the Nurses' Health Study (NHS), the Health Professional Follow-Up Study (HPFS) in the discovery stage, and the Women Genome Health Study (WGHS) in the replication stage. We observed directionally consistent interactions between genetic susceptibility and SSB in the NHS and HPFS. In the combined samples of these cohorts, increases in BMI (kg/m^2) per 10 risk alleles were 1.00 for an SSB intake of less than 1 serving/month, 1.03 for 1–4 servings/month, 1.39 for

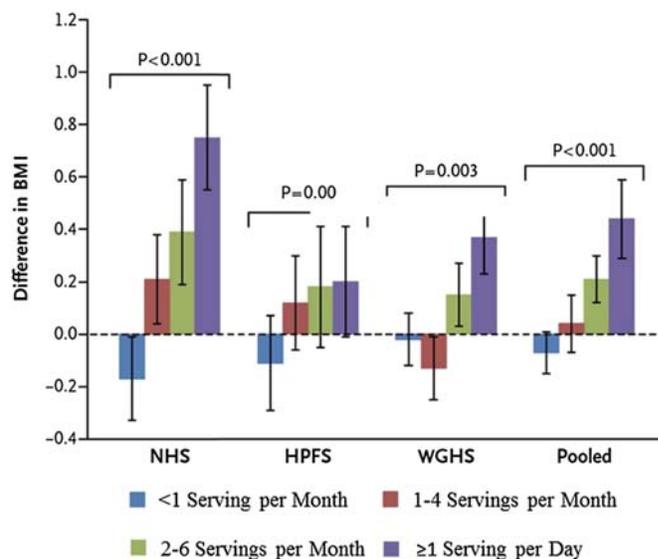


FIGURE 25.1 Gene–diet interaction on body mass index (BMI). Increase in BMI per 10-unit increase in genetic risk score, which was calculated by summing the number of risk alleles of 32 BMI-associated genetic variants, according to intake of sugar-sweetened beverages in the Nurses' Health Study (NHS), Health Professionals Follow-Up Study (HPFS), and Women Genome Health Study (WGHS).

2–6 servings/week, and 1.77 for 1 serving or more/day ($P_{\text{interaction}} < .001$). Findings were successfully replicated in the WGHS ($P_{\text{interaction}} < .001$). We also observed consistent gene–SSB interactions in two cohorts of Swedish adults, the Malmö Diet and Cancer Study and healthy participants from the Gene–Lifestyle Interactions and Complex Traits Involved in Elevated Disease Risk (Brunkwall et al., 2016). Brunkwall et al. found a nominal significant interaction between SSB intake category and the obesity GRS ($P_{\text{interaction}} = .03$). Comparing participants within the top and bottom quartiles of the GRS with each increment in SSB intake was associated with 0.24 and 0.15 higher BMIs, respectively. Higher fried food intake, which increases energy intake, is considered an unhealthy dietary factor that influences the risk for general and central obesity. In another study including the NHS, HPFS, and WGHS (Qi et al., 2014), we provided reproducible evidence to show that fried food consumption interacted with genetic susceptibility (the GRS) in relation to obesity ($P_{\text{interaction}} \leq .001$). Our data indicated that the gene–fried foods interactions were bidirectional. On the one hand, we found that differences in BMI associated with consumption of fried foods were stronger among men and women with a higher GRS compared with those with a lower GRS. On the other hand, the genetic association with adiposity was strengthened with higher consumption of fried foods. In the three combined cohorts, differences in BMI per 10 risk alleles were 1.1, 1.6, and 2.2 for fried food consumption less than once, one to three times, and four or more times a week ($P < .001$ for interaction) and odds ratios (95% confidence intervals) for obesity per 10 risk alleles were 1.61 (1.40–1.87), 2.12 (1.73–2.59), and 2.72 (2.12–3.48) across the three categories of consumption ($P_{\text{interaction}} = .002$). Our study indicated that the *FTO* genotype showed the strongest interaction ($P_{\text{interaction}} < .001$) among all obesity-predisposing variants. In another study, Nettleton et al. performed an analysis on dietary patterns and genetic predisposition to obesity risk. The study included data from 18 cohorts of European ancestry and defined a composite score representing a healthy diet (calculated based on self-reported intakes of whole grains, fish, fruits, vegetables, nuts/seeds, and red/processed meats, sweets, sugar-sweetened beverages, and fried potatoes). The results suggested that associations between genetic predisposition (the GRS) and obesity traits were stronger among individuals with a healthier diet score. The data from these prospective cohorts highlight the potential interactions of habitual intakes of dietary factors with genetic makeup in relation to obesity.

GENOTYPE, DIET INTERVENTIONS, AND WEIGHT LOSS

Testing gene–diet interactions in observational studies may be biased by confounding and reverse causation, and therefore provide twisted information on how genotype modifies dietary effect on body weight. Evidence-based prevention and intervention rely mainly on statistical interpretation of data from randomized clinical trials, in which potential confounding effects are minimized. Emerging data have highlighted the importance of detecting gene–diet interactions in randomized diet intervention trials on weight management.

The Preventing Overweight Using Novel Dietary Strategies (POUNDS Lost) was a clinical trial including 811 overweight or obese adults assigned to one of four weight loss diets varying in macronutrient content for 2 years. At 6 months, participants assigned to each diet had lost an average of 6 kg, but they began to regain weight after 12 months. By 2 years, weight loss remained similar in those who were assigned to various diets. Driven by specific hypotheses, we performed a series of analyses on gene–diet interactions in the POUNDS Lost trial and detected statistical significant interactions between diet interventions and a variety of genetic variations that had a strong association with obesity, diseases (e.g., diabetes, hypertension, dyslipidemia, and cardiovascular disease), and diet preference in relation to weight loss and related changes in metabolic traits. Table 25.1 lists major findings from our analyses. Interestingly, several genetic variants showed consistent gene–diet interactions related to difference metabolic traits. For example, the *FTO* gene is among the main culprits in determining genetic susceptibility to obesity and is highly expressed in the central nervous system. In the POUNDS Lost trial, *FTO* genetic variants were found to modify the impact of dietary protein intake or dietary fat intake on a series of obesity-related traits, such as weight loss, changes in body composition, fat distribution, and improvement in insulin resistance. Another example is the insulin receptor substrate 1 (*IRS1*) gene, which is associated with insulin resistance and hyperinsulinemia. In the POUNDS Lost trial, different studies reported statistically significant interactions between genetic variants in *IRS1* and dietary fat intake in relation to weight loss, improvement in insulin resistance, and management of metabolic syndrome components. Our analysis indicated that a genetic variant in the *FGF21* region, which was identified from GWAS to be related to a preference for macronutrients such as carbohydrates and fat, was associated with distinct weight loss trajectories among overweight or obese participants with different carbohydrate intakes (low versus high) in the POUNDS Lost trial.

TABLE 25.1 Example of Interactions Between Genotypes And Weight Loss Diets on Adiposity And Metabolic Outcomes in The Preventing Overweight Using Novel Dietary Strategies Trial.

Low-fat/high-carbohydrate diet or high-fat/low-carbohydrate diet		High- or low-protein diet	
Genetic variants	Outcomes	Genetic variants	Outcomes
Diabetes genetic risk score	Glycemic traits	Diabetes genetic risk score	Insulin resistance; insulin secretion
<i>IRS1</i> rs1522813, rs2943641	Insulin resistance; metabolic syndrome; body weight;	<i>DHCR7</i> rs12785878	Insulin resistance
<i>FTO</i> rs1558902	Insulin resistance	<i>FTO</i> rs9939609, rs1558902	Body composition and fat distribution; appetite
<i>GIPR</i> rs2287019	Glycemic traits; insulin resistance		
<i>CRY2</i> rs11605924, <i>MTNR1B</i> rs10830963	Energy expenditure		
<i>TCF7L2</i> rs12255372	Body composition		
<i>PCSK7</i> rs236918	Insulin resistance		
<i>APOA5</i> rs964184	Lipid profiles		
<i>LIPC</i> rs2070895	Lipid profiles		
<i>CETP</i> rs3764261	Lipid profiles		
<i>NPY</i> rs16147	Blood pressure		
<i>PPM1K</i> rs1440581	Insulin resistance; body weight		
<i>FGF21</i> rs838147	Body composition		
Adiponectin genetic risk score	Appetite		

Adapted from Heianza, Y., Qi, L. 2017. Gene-diet interaction and precision nutrition in obesity. *Int J Mol Sci* 18 (4). <https://doi.org/10.3390/ijms18040787>.

Several other studies examined gene–diet interactions in randomized intervention trials. For example, transcription factor 7-like 2 (*TCF7L2*) genetic variant rs7903146 interacted with dietary fiber intake in relation to weight loss in a follow-up study of 304 subjects from the Tübingen Lifestyle Intervention Program (TULIP), a trial consisting of an exercise and diet intervention with decreased intake of fat and increased intake of fiber (> 15 g fiber per 1000 kcal). However, the association of *TCF7L2* SNP rs7903146 with successful weight loss during the lifestyle intervention in TULIP was not replicated in another randomized trial, the Diabetes Prevention Program (DPP). The authors speculated that this might be because increased fiber intake was not part of DPP. Replication remains a major challenge in gene–diet interaction analyses in randomized trials. Taking advantage of the similar design of the POUNDS Lost trial and the Dietary Intervention Randomized Controlled Trial, is a 2-year diet intervention trial aimed at testing three different diets (low-fat, Mediterranean, and low-carbohydrate) on weight loss in 322 moderately obese participants, investigators reported reproducible findings on interactions between *CETP* variant rs3764261 and dietary fat in relation to an

improvement in high-density lipoprotein cholesterol and triglycerides (Heianza and Qi, 2017).

These findings from randomized intervention trials suggest that weight loss and related metabolic improvement in response to diet interventions may be contingent on individuals' genomic background. One unique advantage of studies applying diet intervention trials to test gene–diet interactions is that they may provide direct evidence to instruct genotype-targeted diet modifications in future public health practice. However, most existing diet intervention trials are relatively small, which would limit the power for detecting moderate gene–diet interactions.

CONCLUDING REMARKS AND FUTURE DIRECTIONS

The etiology of chronic diseases such as obesity is complex. It involves contributions from both environmental (e.g., diet) and genetic factors. The epidemic of obesity has been ascribed to a collision between genes and the obesogenic environment. Genome-wide profiling efforts have given informative insights into

the biological processes of obesity; however, our understanding of how these genomic factors interact with diet is still unclear. In the past few years, studies on gene–diet interactions have been rapidly increasing. In the face of the modern obesity epidemic, the previously prevailing view that the condition is purely environmentally related has been revised; what is increasingly clear is that the effect of diets on human health such as body weight are likely to depend on our genome makeup. Findings from these studies have broad appeal and will no doubt lead to a great deal of further analysis by the scientific community. However, considering the wealth of variation, the major challenge remains in their meaningful interpretation.

Findings from gene–diet interactions hold great promise for accelerating precision nutrition and personalized interventions in obesity management. Although genomic information is generally not useful for disease prediction, it could be employed along with other risk factors to screen and classify high-risk populations for tailored intervention or treatment. Data from dietary intervention trials suggest that changes in adiposity and metabolic response to low-calorie weight-loss diets could be significantly modified by genetic variants, especially those related to obesity, type 2 diabetes, metabolism, and food preference. Although further external replications and large-scale analyses are necessary to confirm these findings, the positive results obtained thus far tend to support precision dietary interventions and to consider genetic predisposition to diseases, genetic variants determining dietary preference and metabolites, as well as phenotypes and intermediate metabolites. The idea of precision nutrition and dietary intervention is considered because each dietary habit and type of advice is individually tailored to prevent chronic disease on the basis of genetic background, habitual food and beverage consumption, nutrient

intake (especially those contributing to risk for disease), and a person's intermediate metabolite profiles. Research integrating data on genes, nutrition, and dietary interventions combined in an investigation of human health would be an exciting area in the near future. Nevertheless, important gaps in translating the new knowledge into the clinic and preventive practice pose a major challenge. In addition, implementation challenges include difficulties in integrating genetic testing in the current health care system, the setting of appropriate risk thresholds, and ethical, legal, and social issues. Empirical data are lacking on these issues. To accelerate this momentum and move our understanding into this new frontier, solutions need to be applied to integrate phenotypic and genomic data efficiently to interpret them better. More collaborative and multidisciplinary research is needed to integrate the fruits from scientific research into public health and clinical practice.

References

- Brunkwall, L., Chen, Y., Hindy, G., et al., 2016. Sugar-sweetened beverage consumption and genetic predisposition to obesity in 2 Swedish cohorts. *Am J Clin Nutr* 104 (3), 809–815. <https://doi.org/10.3945/ajcn.115.126052>.
- Heianza, Y., Qi, L., 2017. Gene-diet interaction and precision nutrition in obesity. *Int J Mol Sci* 18 (4). <https://doi.org/10.3390/ijms18040787>.
- Locke, A.E., Kahali, B., Berndt, S.I., et al., 2015. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518 (7538), 197–206. <https://doi.org/10.1038/nature14177>.
- Qi, L., Cho, Y.A., 2008. Gene-environment interaction and obesity. *Nutr Rev* 66 (12), 684–694. <https://doi.org/10.1111/j.1753-4887.2008.00128.x>.
- Qi, Q., Chu, A.Y., Kang, J.H., et al., 2014. Fried food consumption, genetic risk, and body mass index: gene-diet interaction analysis in three US cohort studies. *BMJ* 348, g1610.

This page intentionally left blank

Carbohydrates and Nutrigenetics

Clara Ibáñez, Ana Ramírez de Molina

Cancer Program, IMDEA Food Institute, Madrid, Spain

Abbreviations

SNPs Single-nucleotide polymorphisms
 SLC Solute Carrier
 VLDL Very low-density lipoprotein cholesterol
 IDL Intermediate-density lipoprotein cholesterol
 LDL Low-density lipoprotein cholesterol
 HDL High-density lipoprotein cholesterol

Glossary Terms

An overview of the investigation in nutrigenetics and nutrigenomics of carbohydrates is presented as well as a summary of the most relevant applications and achievements in this research area. The Nutritional Genomics studies the interaction between bioactive food components and the genome, which includes Nutrigenetics and Nutrigenomics. The influence of nutrients on genes expression is called Nutrigenomics, while the heterogeneous response of gene variants to nutrients, dietary components and developing nutraceuticals is called Nutrigenetics.

INTRODUCTION

The concept of a “nutritional phenotype” involves the interconnection between diet and the disease/wellness balance with the aim to determine relevant genes, proteins, and metabolic pathways that modulate health status in a certain environment as a starting point to tailor dietary patterns (Zeisel et al., 2005). In the postgenomic era it is already accepted that there exists cross talk between our dietary behavior and response to nutrients with individual genetic and metabolic factors (Lampe et al., 2013) with the important role of microbiota that has gained interest in recent years (Sonnenburg and Bäckhed, 2016).

Among the macronutrients, carbohydrates are the source with higher recommended rates, being 45%–65% total calories according to the U.S. Institute of Medicine (Trumbo et al., 2002). Among the wide group of compounds that include carbohydrates, attending to the number of forming units there are three major categories of carbohydrates, including monosaccharides (fructose,

glucose, galactose), disaccharides (lactose, maltose, sucrose), oligosaccharides (starches), and polysaccharides. In addition, carbohydrates can be differentiated in digestible and nondigestible carbohydrates (fibers, fructans). Some of the nondigestible compounds are prebiotics as they can be fermented by gut microbiota and exert important benefits to the host, for example, by the modulation of the microbiota diversity. To date, all identified and assumed prebiotics are carbohydrate compounds, mainly oligosaccharides (Slavin, 2013).

Different biological processes influence our personal response to carbohydrates when ingested. Accordingly, different genes related to these processes determine how carbohydrates affect our energetic balance and well-being. In this sense, carbohydrates are absorbed in different ways to reach the bloodstream. Oligosaccharides and polysaccharides must be hydrolyzed into their component monosaccharides to be absorbed in the upper small intestine by amylase, α -glucosidases, and disaccharidases (i.e., lactase, maltase, etc.). Then specific transporter proteins play essential roles in facilitating the diffusion (GLUTs proteins) or the active transport (SGLTs proteins) of the molecular units across the apical membrane of enterocytes. A schematic representation of carbohydrate absorption is given in Fig. 26.1. Namely, fructose transport is assisted by fructose transporter 2 and 5 (GLUT2, GLUT5) while glucose and galactose are preferably transported by the sodium/glucose transporter 1 (SGLT1). This process ends with the delivery of carbohydrates in the bloodstream.

In contrast, the concentration and transformation of carbohydrates in the organism is essentially influenced by metabolic pathways performance such as glycolysis, glycogenolysis, and glycogenesis among others. Of all the proteins that are involved in those pathways, some of the most relevant and limiting enzymes have been designed as hexokinase (HK), phosphofructokinase (PFK-1), and pyruvate kinase (PK).

Given the relevance of carbohydrates in our diet, it becomes imperative to deepen the knowledge of genetic

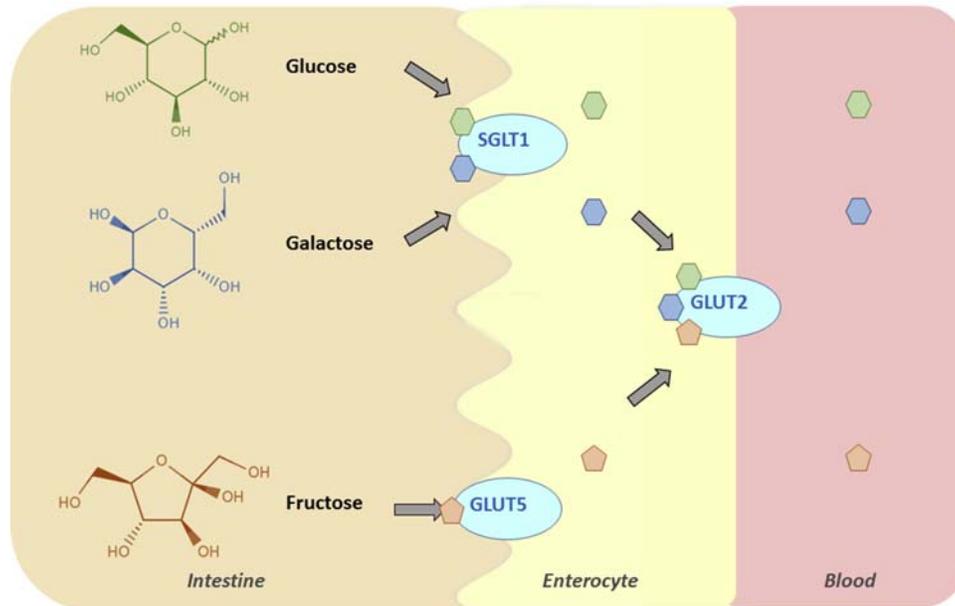


FIGURE 26.1 Scheme of the carbohydrate absorption from the intestine to the bloodstream.

and molecular mechanisms related to their biological effect in our personal health status. This scenario makes nutrigenetics research highly relevant to understanding biological processes involving the role of carbohydrates in the organism and thus potentially affecting individual carbohydrate intake behavior and metabolism in the general population.

This chapter aims to give a broad view of investigations in nutrigenetics and nutrigenomics of carbohydrates.

POLYMORPHISMS LINKED TO CARBOHYDRATE INTAKE AND METABOLISM

Interindividual differences in the response to diet are genotype dependent to a large extent (Lampe et al., 2013). As it will be shown later, different polymorphisms linked to carbohydrate intake or metabolism may affect a person's appetite and increase consumption of carbohydrates, taste perception, as well as the susceptibility to develop obesity.

As well, carbohydrates by acting as prebiotics can also affect absorption of other nutrients, such as minerals. Prebiotics are a group of carbohydrates that have multiple unknown effects and thus are the focus of intense research. Prebiotics resist digestion in the small intestine and reach the colon where they are fermented by the intraluminal microflora; a variety of benefits have been associated with prebiotic consumption. Previous investigations have identified various genetic factors that contribute calcium and magnesium absorption in

pubertal adolescent prebiotic supplementation (Abrams et al., 2005). In this work, it has been reported that daily diet supplemented with prebiotic compounds significantly increased the absorption of minerals with a more acute effect in individuals with the *Fok1* F allele (rs2228570) in the vitamin D receptor gene.

Genetic Variants Linked to Carbohydrates and Obesity

Obesity remains a central public health concern by itself, and because it amplifies risk factors for common chronic illness such as cardiovascular disease (Van Gaal et al., 2006) or many types of cancer (O'Sullivan et al., 2018). For instance, the World Health Organization warns that worldwide obesity has more than doubled since 1980 (World Health Organization, 2016) and each year at least 2.8 million people die as a consequence of being overweight or obese (World Health Organization, 2014). Thus, research into the mechanisms by which obesity can be reduced is nowadays of key importance. In this respect, some approaches that have been explored have targeted molecules related to carbohydrate consumption and metabolism. For instance, the agouti-related protein (AgRP) is a powerful orexigenic molecule. It promotes food consumption when ubiquitously overexpressed or when it is systematically administered (Ilnytska and Argyropoulos, 2008). Interestingly, polymorphisms affecting AgRP have shown to modulate the appetite and to increase carbohydrate consumption at the expense of fat intake (Loos et al., 2005). Namely, lower saturated and monounsaturated fats were consumed by individuals' Ala67Thr

heterozygotes while carbohydrate intake was 2.6% higher when compared to Ala67Ala homozygotes. On the other hand, there are AgRP polymorphisms strongly linked to resistance to fatness and to resistance for developing type 2 diabetes in people with African ancestry (Ilnytska and Argyropoulos, 2008). Although it has been observed that AgRP could modulate metabolism and energy balance via different actions, the detailed mechanism remains unknown.

Another actor related to obesity and carbohydrate intake is the dopamine receptor D2 (*DRD2*) gene, which has been already linked to food addiction. It seems that one of the main reasons for that contribution is related to the function of *DRD2* as a key element in the regulation of the rewarding effects of foods in response to a diet with high content of carbohydrates. In this sense, the C957T single-nucleotide polymorphism (SNP) in the *DRD2* gene has been linked to lower sucrose consumption (Eny et al., 2009). In contrast, higher sugar intake has been linked to a *TUB* gene variant, also associated with higher glycemic load (van Vliet-Ostapchouk et al., 2008). *TUB* gene encodes the tubby protein homolog that has been linked to obesity, eating behavior, and sensorineural degradation processes (Kleyn et al., 1996).

Genetic Variants Linked to Sweetness Perception

Apart from the contribution to obesity from the rewarding effects that high-sugar diet triggers, also highlighted is that the sense of sweetness in energy intake and food acceptance can also contribute to increase body weight. Carbohydrate consumption seems to be also modulated by the sweet taste perceived by each individual. The understanding of interindividual differences on calorie intake due to sweetness in oral and gastrointestinal systems could contribute to the molecular knowledge of obesity and targeted prevention. In this line, Low et al. reviewed the most relevant facts and published results pointing out that those mechanisms could even predispose to develop obesity or overweight (Low et al., 2014).

One potential target in the investigation of sweetness perception is the *TAS1R2/TAS1R3* taste receptor, which mediates the sweet taste of some sugars. Taste receptor type 1 member 2 and 3 (*TAS1R2* and *TAS1R3*) genes encode for heterodimeric G protein-coupled receptors that mediate the human sweet taste perception (Fushan et al., 2009). *TAS1R2* responds to sucrose, fructose, and some artificial sweeteners, and there exists controversy about the glucose ligand. Thus, genetic variant Ile191Val has been connected to an 18% reduction in sugar intake related to taste perception (Eny et al., 2010), while polymorphism Val191Val has been associated with higher

carbohydrate intake and with hypertriglyceridemia (Ramos-Lopez et al., 2016). In addition, a more recent sensory analysis study with 95 volunteers aged 20–29 found that lower sensitivity rates led to higher sugar consumption (Dias et al., 2015). Moreover, eight polymorphisms of *TAS1R2* gene were analyzed, relating them to sucrose taste and sugar intake; the latter was related to body mass index. Other polymorphisms have also been studied in the *TAS1R3* gene (Fushan et al., 2009). Namely, two C/T SNPs located at positions –1572 (rs307355) and –1266 (rs35744813) upstream of the *TAS1R3* coding sequence, more concretely in a non-coding regulatory region of the gene, were found to be strongly correlated with human taste sensitivity to sucrose. Moreover this polymorphism could explain 16% of interindividual variability in sweetness perception in the general population. Other investigations of the same research group established different taste threshold sensitivity for sucrose and sweet taste that could be related to *GNAT3* gene (Fushan et al., 2010).

Genetic Variants Linked to Transporters

Food intake has also been observed to be influenced by the carbohydrate transporter and detector GLUT2 molecule involved in energy production at a cellular level. It has been shown that the contribution to food intake of GLUT2 is via the hypothalamus (Stolarczyk et al., 2010). The brain, especially the hypothalamus, is strongly involved in regulating weight and food intake (Bauer et al., 2009). A genetic polymorphism in GLUT2 (Thr110Ile) has shown to induce 30% increase carbohydrate consumption independently from protein, fat, or alcohol intake or the individual's age (Eny et al., 2008). Although the sugar preferences have not been explained yet, Thr110Ile polymorphism is hypothesized to exert certain glucose sensing deficiency rather than glucose transport insufficiency. A clinical trial involving 54 males and 55 females with type 2 diabetes found that Thr110Ile polymorphism increased average intakes of complex and simple carbohydrates without effects on total energy, protein, fat, alcohol, fiber, and glycemic index (Eny et al., 2007). In addition, Thr110Ile variant has been associated with higher incidence of type 2 diabetes and hypercholesterolemia (Igl et al., 2010; Willer et al., 2007).

Other type of transporters in cells are potassium channels, which include a specific tetramerization domain in the cytoplasmic N-terminal. That domain is contained in the potassium channel tetramerization domain containing 15 (*KCTD15*), a protein encoded by *KCTD15* gene. Interestingly, in a genome-wide association where 1700 females were involved, it was observed that carriers of an identified risk allele in or near

KCTD15 consumed more carbohydrates (per allele effect 2.50 g/day) and less monounsaturated fat than noncarriers (Bauer et al., 2009). In addition, various new genes identified in that study were found to be highly expressed in the brain, and several of them particularly in the hypothalamus, which is concordance with the importance of this organ to food predilections and obesity. Going further, the authors suggest that food preference could be correlated with obesity mechanisms and adiposity measures.

NUTRIGENOMICS OF CARBOHYDRATES

Apart from the nutrigenetic effects that involve various gene polymorphisms in relation to carbohydrate intake or metabolism, it is highly relevant how the genes are affected by carbohydrate consumption, namely nutrigenomics of carbohydrates. Thus, a high carbohydrate-to-fat ratio was found to enhance the histone acetylation on the sucrase-isomaltase gene in the small intestines of mice (Honma et al., 2007).

Fructose is among the most highly consumed carbohydrates worldwide and is therefore an important source of calories. As stated before, one of the transporters of fructose is GLUT5, which is encoded by the *SLC2A5* gene. Its expression has been shown to be modulated by hormones like glucocorticoids and dietary fructose. As this receptor is not modulated by glucose, Suzuki et al. (Suzuki et al., 2011) perfused rat intestines with fructose (n = 10 rats) and glucose (n = 20 rats) in order to elucidate in vivo expression of GLUT5. Acetylation of histone H3 in two regions of the *SLC2A5* gene was found to be higher in rat intestines perfused with fructose than in those perfused with glucose. Microarray hybridization and RT-PCR were used in another study focused on the identification of genes altered by high fructose perfusion and compared to high glucose perfusion in the small intestine of neonatal rats (Cui et al., 2004). As a result more than 20 genes showed increased expression levels related to high fructose perfusion including not only transporter molecules such as GLUT5 and GLUT4 but also important enzymes such as glucose-6-phosphatase, fructose-1,6-bisphosphatase, A2- N-acetyltransferase, phosphoribosylpyrophosphate synthetase, phosphofructokinase, and fructose-2,6-bisphosphatase, among others. In contrast, other genes were downregulated by high fructose perfusion including apolipoprotein C-III, Na⁺-phosphate transporter IIb, or GLUT2. These studies demonstrate that fructose can modify gene expression when they are infused in animal models. A more recent study demonstrated that dietary fructose modified gene expression in liver of rats fed with 63% w/w fructose for 2 weeks (Koo et al., 2008). This intervention increased

expression of the genes encoding fructokinase, aldolase B, phosphofructokinase-1, fructose-1,6-bisphosphatase, and carbohydrate response element binding protein (ChREBP) were observed together with an increase in glycogen and triglycerides in liver. Similarly, fructose was administered to male Wistar rats for 3 weeks and hepatic gene expression was compared to control diet in the same period (Francini et al., 2010). Expression of gene catalase, copper–zinc–superoxide dismutase, and glutathione peroxidase was significantly lower in the fructose-fed group. In line with these data is the observation that catalase and copper–zinc–superoxide dismutase protein expression together with the catalase activity and reduced glutathione content were also lower in rats fed with fructose. As well, the antioxidant system of liver was also evaluated and protein carbonyl content was higher in the fructose group whereas TBARS and membrane susceptibility to lipid peroxidation were not altered after fructose intake (Francini et al., 2010). Going further, an excess of maternal fructose during pregnancy has been related to transcriptome modification and adverse effects in the later life of the offspring (Tain et al., 2016). This is to say that most relevant transcriptome changes have been observed in fructose metabolism, glycolysis, gluconeogenesis, fatty acid metabolism and insulin signaling as summarized in Fig. 26.2. These investigations at prenatal period may be especially useful as early diagnosis and prevention methods in a personalized way.

Glucose is the most abundant monosaccharide in nature and like fructose is a simple sugar highly consumed worldwide that also exerts important effects on gene expression.

Furthermore, it has been observed that high glucose doses downregulated the expression of genes such as sucrase-isomaltase, *SLC2A2*, *SLC2A5*, *SLC5A1* and ketohexokinase in vitro (Le, 2010). Additionally, gene expression has been correlated not only to glucose consumption but to cell growth phases. In this line, levels of expression of hexose-transporter mRNAs (*SGLT1*, *GLUT1-GLUT5*) in relation to the phases of cell growth were evaluated in Caco-2 cell clones differing in glucose consumption rates and expression of sucrase-isomaltase gene (Mahraoui et al., 1994). GLUT1 and GLUT3 mRNAs decreased levels in the stationary phase compared to the exponential phase in correlation with glucose decrease consumption, except for high-glucose-consuming clones, which maintained elevated levels in both phases. In contrast, *SGLT1*, GLUT2, and GLUT5 were only expressed in the low-glucose-consuming clones in correlation with sucrase-isomaltase expression, with augmented levels in the stationary phase compared to the exponential phase. In fact, a glucose-dependent negative regulation of expression of sucrase-isomaltase gene in Caco-2 cell

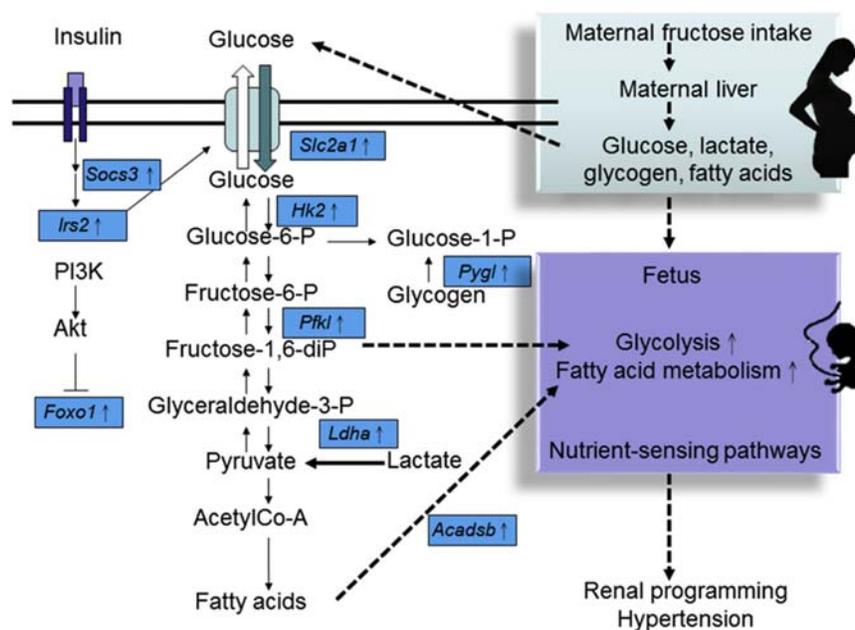


FIGURE 26.2 Schematic representation of changes in the expression of genes regulating glucose metabolism, fatty acid metabolism, and insulin signaling in the kidneys of offspring exposed to maternal high-fat (HF) diet. Solid lines with arrowheads indicate known signaling events and interactions between glucose metabolism, fatty acid metabolism, and insulin signaling. Dashed lines with arrowheads denote proposed mechanisms contributing to maternal HF consumption-induced programmed hypertension. Solid square boxes indicate differentially expressed genes (DEGs) identified by next-generation sequencing (NGS). Reprinted from *Tain et al. (2016)*.

line (Chantret et al., 1994) has been reported. Primarily, it was thought that transcriptome effects due to glucose were correlated to insulin release and not directly influenced by glucose consumption. However, the independent effects of glucose on gene expression have been extensively demonstrated in liver and pancreatic cells (Spears, 2002). Glucose in presence of insulin has demonstrated to overexpress genes encoding pyruvate kinase, glucose transporters, fatty acid synthase, and acetyl-CoA carboxylase while phosphoenolpyruvate carboxykinase and other genes involved in gluconeogenic pathway are downregulated (Girard et al., 1997; Johnston, 1999; Towle et al., 1997; Towle, 2001; Vaulont and Kahn, 1994) leading to augmented substrates derived from glucose to be used into fatty acid synthesis (Rui, 2014). It is important to note that the mechanisms by which genes are modulated are highly dependent on individual tissues and mediators such as hormones. In this case, insulin and glucagon are determinant in the effects of glucose in different tissues. For instance, one of the mechanisms insulin dependent is the increase of glucose that increases insulin levels, and this fact activates genes like *SREBP1* and glucokinase in liver and suppresses other genes such as forkhead transcription factor *FKHR* (Vaulont et al., 2000). Also, glucose also exerts effects directly on target genes. In this case, the genes contain a region in their promoters with specific cis-regulatory elements, called glucose response elements, promoting systemic glucose uptake into cells,

its intracellular metabolism, transformation, and accumulation as lipids (Spears, 2002). The transcription factors that bind the glucose response elements have not been completely identified yet, and regulation through this mechanism is not fully understood. However, one of the possible mechanisms, as described previously, could be mediated by formation of a complex of ChREBP and max-like protein X (Uyeda and Repa, 2006).

In most humans, starch is the most extensively consumed carbohydrate, such as with foods like potatoes and cereals. The molecule is a polymer of glucose molecules joined by glycosidic bonds and is found in most green plants as energy storage. Interestingly, high-starch consumers potentially present more copies of salivary amylase gene *AMY1* than those with low-starch diets (Perry et al., 2007). Resistant starch is a subtype of starch that resists digestion almost as much as dietary fiber. It is largely unaffected by enzymatic hydrolysis to glucose in the proximal small intestine and can be utilized by colonic bacteria. Interestingly, it has been found that pigs fed with resistant starch showed higher concentrations of short-chain fatty acids in cecal and colonic samples and influenced by colonic bacteria composition (Haenen et al., 2013). Those results were accompanied by a significantly higher expression of monocarboxylate transporter 1 (*SLC16A1*) and glucagon (*GCG*) in pig intestine. The effects of dietary resistant starch also have been examined in rats (Han et al., 2004). For that purpose, F344/Du Crj rats were fed

with 5% of pancreatin-resistant fraction obtained from *Phaseolus vulgaris* beans for 4 weeks and compared to control rats fed with cellulose powder for the same time period. As a result, lower amounts of circulating lipids were observed in the group fed with resistant starch compared to controls in terms of VLDL, IDL, LDL, and HDL. On the other hand, hepatic cholesterol 7 α -hydroxylase, LDL receptor, and *SR-B1* mRNA were significantly upregulated in the resistant starch group (Han et al., 2004).

Although this chapter aims to give a broad overview of gene expression modified or regulated by carbohydrates in humans, it is worth mentioning that carbohydrates also modulate gene expression in plants through sugar-sensing systems initiation as extensively investigated and reviewed previously (Koch, 1996; Rolland et al., 2006).

CONCLUSIONS

Investigation of nutrigenetics and nutrigenomic interactions with carbohydrates has become increasingly relevant to deepen understanding of genetic and molecular mechanisms relating to carbohydrates in our diet. Thus, knowledge of biological events involving carbohydrates at the gene level and the role of polymorphisms affecting individual intake behaviors and metabolic processing are key research lines in the so-called nutritional phenotype. These genetic variations among individuals contribute to dietary effects both qualitatively and quantitatively. Some of the differences attributed to these variations influence sweet perception, obesity susceptibility, absorption of minerals, and transport of sugars, among others. This scenario becomes even more complex when the effect of carbohydrate consumption on genes is included in the equation. In this regard, fructose, glucose, and starch are among the best investigated carbohydrates, though mostly in animal models. In summary, future investigations focusing on nutrigenomics and nutrigenetics of carbohydrates are needed to better understand the interactions of carbohydrates and genes in humans. Innovative intervention studies and genetic analysis, ideally assisted with omic technologies, will provide a broader vision of the numerous biological processes dealing with carbohydrates, both from foods and from endogenous metabolism.

Acknowledgments

This work has been supported by Ministerio de Economía y Competitividad del Gobierno de España (MINECO, Plan Nacional I + D + i AGL2016-76736-C3), Gobierno regional de la Comunidad de Madrid (P2013/ABI-2728, ALIBIRD-CM), and EU Structural Funds. C. I.

thanks the MINECO (Ministerio de Economía y Competitividad) for her Juan de la Cierva postdoctoral research contract (FJCI-2014-19601).

Conflicts of Interest

The authors declare no conflict of interest.

References

- Abrams, S.A., Griffin, I.J., Hawthorne, K.M., Liang, L., Gunn, S.K., Darlington, G., Ellis, K.J., 2005. A combination of prebiotic short- and long-chain inulin-type fructans enhances calcium absorption and bone mineralization in young adolescents. *Am J Clin Nutr* 82, 471–476.
- Bauer, F., Elbers, C.C., Adan, R.A., Loos, R.J., Onland-Moret, N.C., Grobbee, D.E., van Vliet-Ostaptchouk, J.V., Wijmenga, C., van der Schouw, Y.T., 2009. Obesity genes identified in genome-wide association studies are associated with adiposity measures and potentially with nutrient-specific food preference. *Am J Clin Nutr* 90, 951–959.
- Chantret, I., Rodolose, A., Barbat, A., Dussaulx, E., Brot-Laroche, E., Zweibaum, A., Rousset, M., 1994. Differential expression of sucrase-isomaltase in clones isolated from early and late passages of the cell line Caco-2: evidence for glucose-dependent negative regulation. *J Cell Sci* 107 (Pt 1), 213–225.
- Cui, X.L., Soteropoulos, P., Tolia, P., Ferraris, R.P., 2004. Fructose-responsive genes in the small intestine of neonatal rats. *Physiol Genom* 18, 206–217.
- Dias, A.G., Eny, K.M., Cockburn, M., Chiu, W., Nielsen, D.E., Duizer, L., El-Soehy, A., 2015. Variation in the TAS1R2 gene, sweet taste perception and intake of sugars. *J Nutrigenetics Nutrigenomics* 8, 81–90.
- Eny, K.M., Fontaine-Bisson, B., Wolever, T.M., El-Soehy, A., 2007. GLUT2 polymorphism modifies the intake of dietary carbohydrates in individuals with Type 2 Diabetes. *FASEB J* 21, A176-A176.
- Eny, K.M., Corey, P.N., El-Soehy, A., 2009. Dopamine D2 receptor genotype (C957T) and habitual consumption of sugars in a free-living population of men and women. *J Nutrigenetics Nutrigenomics* 2, 235–242.
- Eny, K.M., Wolever, T.M., Corey, P.N., El-Soehy, A., 2010. Genetic variation in TAS1R2 (Ile191Val) is associated with consumption of sugars in overweight and obese individuals in 2 distinct populations. *Am J Clin Nutr* 92, 1501–1510.
- Eny, K.M., Wolever, T.M., Fontaine-Bisson, B., El-Soehy, A., 2008. Genetic variant in the glucose transporter type 2 is associated with higher intakes of sugars in two distinct populations. *Physiol Genom* 33, 355–360.
- Francini, F., Castro, M.C., Schinella, G., García, M.E., Maiztegui, B., Raschia, M.A., Gagliardino, J.J., Massa, M.L., 2010. Changes induced by a fructose-rich diet on hepatic metabolism and the antioxidant system. *Life Sci* 86, 965–971.
- Fushan, A.A., Simons, C.T., Slack, J.P., Drayna, D., 2010. Association between common variation in genes encoding sweet taste signaling components and human sucrose perception. *Chem Senses* 35 (7), 579–592.
- Fushan, A.A., Simons, C.T., Slack, J.P., Manichaikul, A., Drayna, D., 2009. Allelic polymorphism within the TAS1R3 promoter is associated with human taste sensitivity to sucrose. *Curr Biol* 19, 1288–1293.
- Girard, J., Ferré, P., Foufelle, F., 1997. Mechanisms by which carbohydrates regulate expression of genes for glycolytic and lipogenic enzymes. *Annu Rev Nutr* 17, 325–352.
- Haenen, D., Zhang, J., Souza da Silva, C., Bosch, G., van der Meer, I.M., van Arkel, J., van den Borne, J.J., Perez Gutierrez, O., Smidt, H., Kemp, B., Muller, M., Hooiveld, G.J., 2013. A diet high in resistant starch modulates microbiota composition, SCFA concentrations, and gene expression in pig intestine. *J Nutr* 143, 274–283.

- Han, K.H., Sekikawa, M., Shimada, K., Sasaki, K., Ohba, K., Fukushima, M., 2004. Resistant starch fraction prepared from kinoto bean affects gene expression of genes associated with cholesterol metabolism in rats. *Exp Biol Med* 229, 787–792.
- Honma, K., Mochizuki, K., Goda, T., 2007. Carbohydrate/fat ratio in the diet alters histone acetylation on the sucrase–isomaltase gene and its expression in mouse small intestine. *Biochem Biophys Res Commun* 357, 1124–1129.
- Igl, W., Johansson, Å., Wilson, J.F., Wild, S.H., Polašek, O., Hayward, C., Vitart, V., Hastie, N., Rudan, P., Gnewuch, C., 2010. Modeling of environmental effects in genome-wide association studies identifies SLC2A2 and HP as novel loci influencing serum cholesterol levels. *PLoS Genet* 6, e1000798.
- Inytska, O., Argyropoulos, G., 2008. The role of the Agouti-related protein in energy balance regulation. *Cell Mol Life Sci* 65, 2721.
- Johnston, M., 1999. Feasting, fasting and fermenting: glucose sensing in yeast and other cells. *Trends Genet* 15, 29–33.
- Kleyn, P.W., Fan, W., Kovats, S.G., Lee, J.J., Pulido, J.C., Wu, Y., Berkemeier, L.R., Misumi, D.J., Holmgren, L., Charlat, O., 1996. Identification and characterization of the mouse obesity gene *tubby*: a member of a novel gene family. *Cell* 85, 281–290.
- Koch, K., 1996. Carbohydrate-modulated gene expression in plants. *Annu Rev Plant Biol* 47, 509–540.
- Koo, H., Wallig, M.A., Chung, B.H., Nara, T.Y., Cho, B.S., Nakamura, M.T., 2008. Dietary fructose induces a wide range of genes with distinct shift in carbohydrate and lipid metabolism in fed and fasted rat liver. *Biochim Biophys Acta* 1782, 341–348.
- Lampe, J.W., Navarro, S.L., Hullar, M.A., Shojaie, A., 2013. Inter-individual differences in response to dietary intervention: integrating omics platforms towards personalised dietary recommendations. *Proc Nutr Soc* 72, 207–218.
- Le, M.T., 2010. Factors Impacting Fructose Bioavailability and its Adverse Metabolic Effects.
- Loos, R.J., Rankinen, T., Rice, T., Rao, D.C., Leon, A.S., Skinner, J.S., Bouchard, C., Argyropoulos, G., 2005. Two ethnic-specific polymorphisms in the human Agouti-related protein gene are associated with macronutrient intake. *Am J Clin Nutr* 82, 1097–1101.
- Low, Y.Q., Lacy, K., Keast, R., 2014. The role of sweet taste in satiation and satiety. *Nutrients* 6, 3431–3450.
- Mahraoui, L., Rodolosse, A., Barbat, A., Dussaulx, E., Zweibaum, A., Rousset, M., Brot-Laroche, E., 1994. Presence and differential expression of SGLT1, GLUT1, GLUT2, GLUT3 and GLUT5 hexose-transporter mRNAs in Caco-2 cell clones in relation to cell growth and glucose consumption. *Biochem J* 298 (Pt 3), 629–633.
- O'Sullivan, J., Lysaght, J., Donohoe, C.L., Reynolds, J.V., 2018. Obesity and gastrointestinal cancer: the interrelationship of adipose and tumour microenvironments. *Nat Rev Gastroenterol Hepatol*. <https://doi.org/10.1038/s41575-018-0069-7>.
- Perry, G.H., Dominy, N.J., Claw, K.G., Lee, A.S., Fiegler, H., Redon, R., Werner, J., Villanea, F.A., Mountain, J.L., Misra, R., 2007. Diet and the evolution of human amylase gene copy number variation. *Nat Genet* 39, 1256–1260.
- Ramos-Lopez, O., Panduro, A., Martinez-Lopez, E., Roman, S., 2016. Sweet taste receptor TAS1R2 polymorphism (Val191Val) is associated with a higher carbohydrate intake and hypertriglyceridemia among the population of west Mexico. *Nutrients* 8, 101.
- Rolland, F., Baena-Gonzalez, E., Sheen, J., 2006. Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annu Rev Plant Biol* 57, 675–709.
- Rui, L., 2014. Energy metabolism in the liver. *Compr Physiol* 4 (1), 177–197.
- Slavin, J., 2013. Fiber and prebiotics: mechanisms and health benefits. *Nutrients* 5, 1417–1435.
- Sonnenburg, J.L., Bäckhed, F., 2016. Diet-microbiota interactions as moderators of human metabolism. *Nature* 535, 56–64.
- Spears, J., 2002. Scientific Advances in Animal Nutrition: Promise for the New Century, pp. 113–126.
- Stolarczyk, E., Guissard, C., Michau, A., Even, P.C., Grosfeld, A., Serradas, P., Lorsignol, A., Penicaud, L., Brot-Laroche, E., Leturque, A., Le Gall, M., 2010. Detection of extracellular glucose by GLUT2 contributes to hypothalamic control of food intake. *Am J Physiol Endocrinol Metab* 298, E1078–E1087.
- Suzuki, T., Douard, V., Mochizuki, K., Goda, T., Ferraris, R.P., 2011. Diet-induced epigenetic regulation in vivo of the intestinal fructose transporter *Glut5* during development of rat small intestine. *Biochem J* 435, 43–53.
- Tain, Y., Chan, J.Y., Hsu, C., 2016. Maternal fructose intake affects transcriptome changes and programmed hypertension in offspring in later life. *Nutrients* 8, 757.
- Towle, H.C., Kaytor, E.N., Shih, H., 1997. Regulation of the expression of lipogenic enzyme genes by carbohydrate. *Annu Rev Nutr* 17, 405–433.
- Towle, H.C., 2001. Glucose and cAMP: adversaries in the regulation of hepatic gene expression. *Proc Natl Acad Sci USA* 98, 13476–13478.
- Trumbo, P., Schlicker, S., Yates, A.A., Poos, M., 2002. Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein and amino acids. *J Am Diet Assoc* 102, 1621–1630.
- Uyeda, K., Repa, J.J., 2006. Carbohydrate response element binding protein, ChREBP, a transcription factor coupling hepatic glucose utilization and lipid synthesis. *Cell Metabolism* 4, 107–110.
- Van Gaal, L.F., Mertens, I.L., Christophe, E., 2006. Mechanisms linking obesity with cardiovascular disease. *Nature* 444, 875–880.
- van Vliet-Ostaptchouk, J.V., Onland-Moret, N.C., Shiri-Sverdlov, R., Van Gorp, P.J., Custers, A., Peeters, P.H., Wijmenga, C., Hofker, M.H., Van Der Schouw, Yvonne, T., 2008. Polymorphisms of the TUB gene are associated with body composition and eating behavior in middle-aged women. *PLoS One* 3, e1405.
- Vaulont, S., Kahn, A., 1994. Transcriptional control of metabolic regulation genes by carbohydrates. *FASEB J* 8, 28–35.
- Vaulont, S., Vasseur-Cognet, M., Kahn, A., 2000. Glucose regulation of gene transcription. *J Biol Chem* 275, 31555–31558.
- Willer, C.J., Bonnycastle, L.L., Conneely, K.N., Duren, W.L., Jackson, A.U., Scott, L.J., Narisu, N., Chines, P.S., Skol, A., Stringham, H.M., Petrie, J., Erdos, M.R., Swift, A.J., Enloe, S.T., Sprau, A.G., Smith, E., Tong, M., Doheny, K.F., Pugh, E.W., Watanabe, R.M., Buchanan, T.A., Valle, T.T., Bergman, R.N., Tuomilehto, J., Mohlke, K.L., Collins, F.S., Boehnke, M., 2007. Screening of 134 single nucleotide polymorphisms (SNPs) previously associated with type 2 diabetes replicates association with 12 SNPs in nine genes. *Diabetes* 56, 256–264.
- World Health Organization, 2016. Obesity and Overweight [Fact Sheet]; Updated June 2016. Trouvé le 13.
- World Health Organization, 2014. Obesity, Situation and Trends. World Health Organization, Geneva (CH).
- Zeisel, S.H., Freake, H.C., Bauman, D.E., Bier, D.M., Burrin, D.G., German, J.B., Klein, S., Marquis, G.S., Milner, J.A., Peltó, G.H., Rasmussen, K.M., 2005. The nutritional phenotype in the age of metabolomics. *J Nutr* 135, 1613–1616.

This page intentionally left blank

Gene-Diet Interactions and Cardiovascular Diseases: Saturated and Monounsaturated Fat

Jose M. Ordovas^{1,2}

¹JM-USDA-HNRCA at Tufts University, Boston, MA, United States; ²USA IMDEA-Alimentacion and Centro Nacional Investigaciones Cardiovasculares (CNIC), Madrid, Spain

Glossary

APOA2 Apolipoprotein A-II
BMI Body mass index
CVD Cardiovascular disease
GxE Gene-environment interactions
GRS Genetic risk score
GOLDN Genetics of Lipid Lowering Drugs and Diet Network
GWAS Genome-wide association studies
HDL High-density lipoprotein
MedDiet Mediterranean diet
MetS Metabolic syndrome
MUFAs Monounsaturated fatty acids
MESA Multi-Ethnic Study of Atherosclerosis
PREDIMED Prevención con Dieta Mediterránea
SFAs Saturated fatty acids
TCF7L2 Transcription factor 7-like 2
T2D Type 2 diabetes

Poor nutrition is a prominent cause of poor health around the world. Therefore, optimizing dietary intake at the individual and population levels is a crucial strategy for health maintenance and disease prevention. Cardiovascular disease (CVD) is a leading cause of death worldwide, and dietary recommendations are essential to CVD prevention and management. However, despite the established connection between diet and heart health, during the past few decades nutrition science and policy have moved in a way that has created uncertainty and confusion about what constitutes optimal and feasible dietary recommendations, especially when it comes to macronutrient intake, such as dietary fats. Moreover, the current nutritional guidelines have not yielded the expected benefits on CVD prevention and bettering of its related risk factors (i.e., dyslipidemia, diabetes, obesity).

The diet-heart hypothesis has its roots in the seminal work of Alexander I. Ignatowski and Nikolai N. Anitschkow over 100 years ago. These investigators

independently reported that rabbits fed either a high-saturated and -cholesterol diet or a cholesterol-rich diet developed fatty deposits in their aortas. The appropriateness of the model, an herbivore animal, cast severe doubts among their contemporary colleagues about the extrapolation of their findings to human atherosclerosis. The concept remained therefore dormant for decades before it was awakened by the epidemiological research, namely the Seven Countries Study, spearheaded and directed by Ancel Keys in the 1950s. During the ensuing decades and until recent times, the dominant thought has been that dietary cholesterol and saturated fat (saturated fatty acids, SFAs) were both positively associated with CVD risk; however, this was not without controversy, even among some of the leading researchers who popularized the diet-heart hypothesis, such as George Mann and E. H. (Pete) Ahrens in the 1970s and 1980s. Nowadays, the cholesterol in the diet has been almost exonerated as a significant driver of circulating blood cholesterol. Concomitantly, the once apparently stable connection between dietary fat and mainly SFAs and CVD appears to be deteriorating as the result of mounting evidence casting doubt about previous beliefs. Thus, whereas the Dietary Guidelines for Americans 2015–20, and several organizations, steadily recommend limiting intake of SFAs to approximately <10% of energy, other organizations, like the Canada's Heart and Stroke Foundation, have opted for eliminating thresholds or limits for dietary SFAs, and instead focus on a balanced diet. Part of the controversy comes from the heterogeneous biological effects of different SFAs (i.e., lauric, palmitic, stearic), the variety of their food sources (plant vs. animal), and the fact that when one component of the diet is removed it will be replaced by another, which will have its own,

and sometimes, unexpected effects. A classic example of this exchange took place a few decades ago with the replacement of high *trans* fatty acids margarine for the traditional butter. It is humbling to recognize that concerns similar to those present in contemporary nutrition-related literature were already being articulated by some of the most knowledgeable nutrition scientists of the late 1940s and early 1950s. Of particular interest is the recognition by some that the association between dietary factors such as cholesterol or SFAs and plasma cholesterol levels and CVD may be present in some, but not in all, individuals.

Monounsaturated fatty acids (MUFAs) have endured through history a more lenient judgment than SFAs. Whereas it is true that MUFAs were found guilty by association during the prosecution against total dietary fat, in general, they were accepted as a health-neutral or even health-positive fat as a component of the Mediterranean diet. Nevertheless, similar to SFAs, their effect on lipid metabolism and other cardiovascular risk factors show significant interindividual variability. These observations highlight the need for more personalized or precision nutrition. Therefore, some of the nutrition research of the last 2 decades has shifted toward the identification of the factors underlying those interindividual differences in response to dietary fats and cholesterol, with greater emphasis placed on the search for genetic factors. Progress in this area has been made possible by the increased understanding of the genetic architecture of complex, common diseases. However, despite the knowledge accumulated, many of the questions remain unanswered, and some of them may be concealed as part of gene-environment (GxE) interactions. In this regard, a few years back, we mined the scientific literature to collect information about such interactions from almost 400 publications for CVD-related traits (e.g., blood lipids, glycemic traits, obesity anthropometrics, vascular measures, inflammation, and metabolic syndrome (MetS)), and generated the CardioGxE (Parnell et al., 2014). The synthesis of this information revealed that the CardioGxE SNPs showed little overlap with variants identified by main-effect genome-wide association studies (GWAS), indicating the importance of environmental interactions with genetic factors on cardiometabolic traits. Moreover, comparison to gene networks responding to plasma cholesterol-lowering or regression of atherosclerotic plaques showed that GxE genes have a more significant role in those responses, mainly through high-energy diets and fat intake, than do GWAS-identified genes for the same traits. In this work, the focus will be on the interaction between genetic variants, SFAs, MUFAs, and cardiovascular risk factors.

The summary of the current knowledge can be seen in Table 27.1. Most references to the research presented in

the table can be found in the freely available file accompanying the Parnell et al. catalog of gene-environment interactions (Parnell et al., 2014) (https://static-content.springer.com/esm/art%3A10.1186%2F1756-0381-7-21/MediaObjects/13040_2014_119_MOESM1_ESM.txt)

The information emerging from these data suggests that the relation between dietary SFAs and MUFAs and CVD-related biomarkers may be mediated by genetic factors. In most cases dietary SFAs enhance the association of genetic variants predisposing to CVD disease. However, based on the current evidence we cannot attribute the effects specifically to SFAs, and the observed effects may be related to the foods or dietary patterns containing the SFAs. The opposite is observed in relation to dietary MUFAs. In this case, higher intake of MUFAs tends to be associated with the quenching of the deleterious main effects associated with certain genetic variants. Again, it is possible that this is not entirely due to the MUFAs themselves but to the foods rich in such fatty acids (i.e., olive oil) or the dietary patterns (i.e., Mediterranean diet).

The most compelling evidence regarding the modulation of the effect of a genotype on cardiometabolic traits and cardiovascular disease by a Mediterranean diet pattern comes from the PREDIMED Study and the Transcription factor 7-like 2 (TCF7L2) gene. Single nucleotide polymorphisms (SNPs) at this locus have been strongly and consistently associated with type 2 diabetes (T2D), but such consistency has not been observed for associations related to plasma lipids and CVD. We explored whether the TCF7L2-rs7903146 (C > T) SNP associations with T2D, glucose, lipids, and CVD incidence were modulated by MedDiet in the PREDIMED Study, a randomized trial (MedDiet vs. low-fat control) including 7018 participants and median follow-up of 4.8 years. Consistent with previous association studies, our results showed that the less common T allele of the TCF7L2-rs7903146 SNP was associated with T2D (odds ratio 1.9 [95% CI 1.6–2.2] for TT compared with CC). Most interesting, MedDiet interacted with rs7903146 on fasting glucose at baseline (prior to the randomization to the diet groups; P interaction = 0.004). Thus, when adherence to the MedDiet was low, TT participants had higher fasting glucose concentrations (132 ± 3.5 mg/dL) than CC + CT (127 ± 3.2 mg/dL) participants ($P = .001$). Conversely, when habitual adherence to a MedDiet pattern was high, there were no differences in fasting glucose concentrations across genotypes ($P = .605$) (Corrêla et al., 2013) (see Table 27.1). Similar effects were also demonstrated for total cholesterol, LDL-cholesterol, and triglycerides (all interactions significant at the level of $P < .05$). When we analyzed the data during the intervention phase of the study, TT subjects consuming the low-fat control diet had a higher stroke incidence (adjusted hazard ratio (HR) 2.9 [95% confidence interval

TABLE 27.1 Summary of research involving gene by diet (saturated and monounsaturated fat) in relation to CVD risk factors and CVD.

Reference	Gene/SNPs	Population	Outcome
Rudkowska et al. (2013)	ABCA1 (rs2230806)	Inuits, men and women (n = 553)	Higher triglyceride (TG) levels were associated with a higher SFA intake in carriers of the C/C genotype.
Schuler et al. (2017)	ACE (rs4343)	Healthy and nonobese twin pairs (n = 46)	This SNP represents a robust nutrigenetic marker for an unfavorable response of blood pressure to high-SFA diets.
Ferguson et al. (2010)	ADIPOQ (rs266729)	LIPGENE-SU.VI.MAX Study of MetS cases and matched controls (n = 1754).	A reduction in plasma SFAs is expected to lower insulin resistance in MetS subjects who are minor allele carriers.
Ferguson et al. (2010)	ADIPOR1 (rs10920533)	LIPGENE-SU.VI.MAX Study of MetS cases and matched controls (n = 1754).	A reduction in plasma SFAs is expected to lower insulin resistance in MetS subjects who are minor allele carriers.
Rudkowska et al. (2013)	AGT (rs699)	Inuits, men and women (n = 553)	Higher total cholesterol (TC) and low-density lipoprotein cholesterol (LDL-C) levels were associated with a higher SFA intake for the T/T genotype.
Rudkowska et al. (2013)	APOA1 (rs5070)	Inuits, men and women (n = 553)	Lower high-density lipoprotein-cholesterol (HDL-C) levels were associated with a higher SFA intake for individuals with the C/C genotype.
Rudkowska et al. (2013)	APOA1 (rs670)	Inuits, men and women (n = 553)	Higher LDL-C levels were associated with a higher SFA intake for the G allele.
Corella et al. (2009)	APOA2 (rs5082)	3462 individuals from 3 populations in the United States: The Framingham Offspring study (1454 whites), the GOLDN study (1078 whites), and BPRHS (930 Hispanics of Caribbean origin)	The influence of the APOA2 -265T > C polymorphism on body-weight -related measures was modulated by SFA in three populations.
Corella et al. (2011)	APOA2 (rs5082)	4602 subjects from two populations: a high-CVD-risk Mediterranean population (n = 907 men and women) and a multiethnic Asian population (n = 2506 Chinese, n = 605 Malays and n = 494 Asian Indians) participating in a Singapore National Health Survey.	The influence of the APOA2 -265T > C polymorphism on body-weight -related measures was modulated by SFA in Mediterranean and Asian populations.
Smith et al. (2013)	APOA2 (rs5082)	GOLDN Study (n = 1116) and in the BPRHS Study (n = 955).	Individuals with the SFA-sensitive APOA2 genotype (CC) who consumed a greater amount of higher-fat dairy foods had greater BMI (compared with those consuming less of the higher-fat dairy foods).

Continued

TABLE 27.1 Summary of research involving gene by diet (saturated and monounsaturated fat) in relation to CVD risk factors and CVD.—cont'd

Reference	Gene/SNPs	Population	Outcome
Campos et al. (1997)	APOA4 (rs5110)	Men (n = 222) and women (n = 236) from rural and urban Costa Rica.	Lifestyles associated with an urban environment, such as increased SFA intake, elicit a more adverse plasma lipoprotein profile (higher LDL-C and lower HDL-C) among Costa Rican carriers of the apoA-IV-2 allele than in apoA-IV-1 homozygotes. Lifestyles associated with an urban environment, such as increased SFA intake, elicit a more adverse plasma lipoprotein profile among Costa Rican carriers of the apoA-IV-2 allele than in apoA-IV-1 homozygotes.
Weggemans et al. (2001)	APOA4 (rs5110)	Mostly normolipidemic subjects (n = 405)	The response of total cholesterol to saturated fat (Table 27.1) was significantly lower in the three subjects with the ApoA 4360-2/2 genotype than in those with the 1/1 or 1/2 genotype. The response of total cholesterol to saturated fat (Table 27.1) was significantly lower in the three subjects with the ApoA 4360-2/2 genotype than in those with the 1/1 or 1/2 genotype. The response of plasma TC to SFA was significantly lower in subjects with the APOA4 360-2/2 genotype than in those with the 1/1 or 1/2 genotype.
Weggemans et al. (2001)	APOB (rs1042031)	Mostly normolipidemic subjects (n = 405)	The response of LDL-C to SFA was significantly lower in subjects with the ApoB EcoRI-1/2 genotype than in those with the 1/1 genotype (most common)
Rudkowska et al. (2013)	APOB (rs693)	Inuits, men and women (n = 553)	T allele had higher TC and LDL-C levels with a higher SFA intake.
Salas et al. (1998)	APOC3 (rs5128)	Male students (n = 115)	The APOC3 gene variant affects insulin response to an OGTT, which could result in reduced sensitivity to insulin, especially when persons consume diets rich in SFA.
Weggemans et al. (2001)	APOC3 (rs5128)	Mostly normolipidemic subjects (n = 405)	The response of HDL-C to SFA was larger in subjects with the ApoC3 SstI-1/1.
Brown, Ordovás and Campos (2003)	APOC3 (T-455C, rs2854116; T-625del, rs11568823)	Randomly selected residents from Costa Rica (n = 336)	Compared to a diet high in SFA fat, a habitually low SFA fat diet is associated with a beneficial lipoprotein profile only among homozygotes of the APOC3 promoter 455T-625T polymorphism.
Rudkowska et al. (2013)	APOC3 (rs5128)	Inuits, men and women (n = 553)	Higher LDL-C levels were associated with a higher SFA intake for the C allele.

TABLE 27.1 Summary of research involving gene by diet (saturated and monounsaturated fat) in relation to CVD risk factors and CVD.—cont'd

Reference	Gene/SNPs	Population	Outcome
Moreno et al. (2004)	APOE (rs429358, rs7412)	Healthy subjects (n = 84)	LDL size was smaller after a CHO diet than after MUFA or SFA diets. After a CHO diet, a significant increase in LDL particle size was noted with respect to the MUFA diet in apoE 4/3 subjects, whereas a significant decrease was observed in the apoE 3/3 individuals.
Yang et al. (2007)	APOE (rs429358, rs7412)	Myocardial infarction (MI) cases (n = 1927) and population-based control (n = 1927)	This study supports the hypothesis that the APOE*2 and APOE*4 variants increase susceptibility to MI in the presence of high SFA .
Rudkowska et al. (2013)	APOE (rs405509)	Inuits, men and women (n = 553)	Lower HDL-C levels were associated with a higher SFA intake for individuals with the T/T genotype.
Fisher et al. (2011)	CAV2 (rs2270188)	Case-control study (n = 192/384) and confirmatory case-cohort study (n = 614/2248) taken from the EPIC-Potsdam study	This SNP interacts with dietary fat and SFA to affect risk of T2DM .
Wallace et al. (2000)	CETP (rs708272)	Free-living individuals (n = 55) completed a double-crossover trial with two dietary regimens, a high-SFA diet and a high-PUFA diet, each phase continuing for 4 weeks.	CETP B1B1 genotype showed an average 0.44 mmol/L (95% CI: 0.22, 0.66) greater change in TC than those with one or more CETP B2 allele when comparing diets high and low in SFA .
Weggemans et al. (2001)	CETP (rs708272)	Mostly normolipidemic subjects (n = 405)	The response of HDL-C to SFA was larger in subjects with the CETP Taq1b-2/2.
Li et al. (2007)	CETP (rs708272)	780 diabetic men (n = 780) who participated in the Health Professionals Follow-up study	An inverse association of the B1 allele with plasma HDL-C existed for those with a high consumption of SFA and MUFA .
Garaulet et al. (2009)	CLOCK (rs1801260)	Participants (n = 1100) in the GOLDN Study	The deleterious effect of gene variants on waist circumference was only found with high SFA intakes (>11.8%) and the protective effect of the minor allele on insulin sensitivity was only present when MUFA intake was >13.2% of energy.
Corella et al. (2016)	CLOCK (rs4580704)	7098 PREDIMED trial participants	Protective effect of the G-allele against hyperglycemia and T2DM incidence especially in subjects consuming the MUFA -rich diet.
Corella et al. (2011)	FTO (rs9939609)	GOLDN[(n = 1069) and Boston Puerto Rican Health study (BPRHS) studies (n = 1094).	Homozygous participants for the FTO-risk allele had a higher mean BMI than the other genotypes only when they had a high- SFA intake.
Phillips et al. (2012)	FTO (rs9939609)	LIPGENE-SU.VI.MAX Study of MetS cases and matched controls (n = 1754).	The SNP was associated with obesity measures, especially in those with the MetS , which was further exacerbated by high dietary SFA intake at baseline and 7.5 y later.

Continued

TABLE 27.1 Summary of research involving gene by diet (saturated and monounsaturated fat) in relation to CVD risk factors and CVD.—cont'd

Reference	Gene/SNPs	Population	Outcome
Zheng et al. (2013)	IRS1 (rs2943641, rs7578326)	GOLDN Study (n = 820) and the BPRHS (n = 844).	rs7578326 G-allele carriers and rs2943641 T-allele carriers and their haplotype G-T carriers had a significantly lower risk of IR and MetS than noncarriers only when the dietary SFA-to-carbohydrate ratio was low (≤ 0.24).
Zheng et al. (2013)	KCNMB3 (rs1183319, rs7645550)	GOLDN Study (n = 820) and the BPRHS (n = 844).	HOMA-IR lower in carriers of the rs1183319 variant only for those with high SFA intake.
Ordovas et al. (2002)	LIPC (rs1800588)	Men (n = 1020) and women (n = 1110) women participating in the Framingham study.	When total SFA intake was high, HDL-C concentrations were lowest among those with the TT genotype.
Zhang (2005)	LIPC (rs1800588)	Diabetic men (n = 780) participants in the Health Professionals Follow-up study	In diabetics, the T allele was associated with higher HDL-C concentrations only in men who were not overweight or who had higher SFA intake.
Nettleton et al. (2007)	LIPC (rs1800588)	11,806 (8897 Whites/2909 African Americans) participating in ARIC	There were significant interactions between LIPC genotype and percent energy from SFA in African Americans. When percent energy intake from SFA was low, TG concentrations were approximately 5 mg/dL lower in CT than in CC or TT genotypes. However, when percent energy intake from SFA was high, TG concentrations were 5–8 mg/dL lower in CC than in CT and TT genotypes.
Lindi et al. (2008)	LIPC (rs2070895)	151 healthy subjects	A MUFA -enriched diet reduced serum LDL-C concentration especially in subjects with the A-250A genotype.
Rudkowska et al. (2013)	LIPC (rs1800588)	Inuits, men and women (n = 553)	Carriers of the T/T genotype of LIPC C-514T had higher TG levels with a higher SFA intake.
Wallace et al. (2000)	LPL (rs320)	Free-living individuals n = 55) completed a double-crossover trial with two dietary regimens, a high-SFA diet and a high-PUFA diet, each phase continuing for 4 weeks.	Individuals with the LPL X447 + allele showed an average 0.45 (95% CI: 0.18, 0.72) mmol/L greater change in TC , than those homozygous for the LPL S447 allele when comparing diets high and low in SFA .
Nettleton et al. (2007)	LPL (rs320)	11,806 (8897 whites/2909 African Americans) participating in ARIC	In whites, but not African Americans, there was an interaction between genotype and dietary SFA and MUFA . Specifically, HDL-C was positively associated with dietary fat intake in CC homozygotes and CG heterozygotes but inversely associated with dietary fat in GG homozygotes.

TABLE 27.1 Summary of research involving gene by diet (saturated and monounsaturated fat) in relation to CVD risk factors and CVD.—cont'd

Reference	Gene/SNPs	Population	Outcome
Smith et al. (2013)	LRP1 (rs1799986)	Population-based sample of Puerto Ricans (n = 920) living in the Boston area	High intake of SFA was associated with higher BMI , waist and hip in minor allele carriers (CT + TT) compared to CC participants.
Huang (2012)	MAT1A (rs3851059)	Population-based sample of Puerto Ricans (n = 994) living in the Boston area	When dietary MUFA intake was low (<11.4% energy/d), plasma Hcy was lower in carriers of the A allele compared to GG subjects.
Ortega-Azorin et al. (2014)	MLXIPL (rs3812316)	7166 participants in the PREvención with Dieta MEDiterránea trial	MedDiet, rich in MUFA , enhances the TG -lowering effect of the MLXIPL -rs3812316 variant and strengthens its protective effect on MI incidence.
Zheng et al. (2013)	PSMD3 (rs4065321, rs709592)	GOLDN study: Men (n = 462) and women (n = 508) and GOLDN study: Men (n = 462) and women (n = 508)	In GOLDN, rs4065321 and rs709592 both significantly interacted with dietary MUFAs and carbohydrate on glucose concentrations.
Garcia-Rios et al. (2012)	PER2 (rs934945, rs2304672)	381 participants with MetS in the European LIPGENE Study	The rs2304672 SNP may influence lipid metabolism by interacting with the plasma total SFA concentration in participants with MetS .
Robitaille et al. (2003)	PPARA (rs1800206)	632 men	In a model including rs1800206, SFA , its interaction, and covariates (smoking habits, and energy and alcohol intake), the interaction explained a significant percent of the variance observed in waist circumference .
Bouchard-Mercier et al. (2011)	PPARA (rs1800206)	674 participants	Among PPARα V162 carriers, subjects with higher SFA intakes had smaller LDL-PPD than those with lower intakes. Among subjects homozygous for the PPARα L162 allele, those with higher SFA intakes had larger LDL-PPD than those with lower SFA intakes.
Robitaille et al. (2003)	PPARG (rs1801282)	A cohort of 720 adults participating in the Québec Family Study (QFS)	SFA intake was significantly correlated with several components of the MetS in P12/P12 homozygotes. None of these expected associations were observed among carriers of the A12 allele.
Lindi et al. (2003)	PPARG (rs1801282)	76 men and 74 women	After the 3-month study period, carriers of the Ala12 allele presented a greater decrease in serum TG concentration in response to n-3 fatty acid supplementation than did subjects with the Pro12Pro genotype when the total dietary fat intake was below 37 E% or the intake of SFA was below 10 E%.
Rosado et al. (2010)	PPARG (rs1801282)	Sixty obese women	In obese women with Pro12Ala/Ala12Ala polymorphisms in the PPARγ2 gene, fat oxidation was negatively correlated with the MUFA (%) intake.

Continued

TABLE 27.1 Summary of research involving gene by diet (saturated and monounsaturated fat) in relation to CVD risk factors and CVD.—cont'd

Reference	Gene/SNPs	Population	Outcome
Garaulet et al. (2011)	PPARG (rs1801282)	1465 subjects enrolled in a behavioral treatment program for obesity based on a Mediterranean diet	Gene-diet interaction between the PPAR γ Pro12Ala SNP and MUFA for BMI and body fat .
Bouchard-Mercier et al. (2011)	PPARG (rs1801282)	674 participants	Gene-diet interactions were found for PPAR γ P12A polymorphism with SFA intake and LDL-PPD .
Rudkowska et al. (2013)	PPARG (rs10865710)	Inuits, men and women (n = 553)	The individuals with the C allele had higher TC and LDLC levels with a higher SFA intake.
Smith et al. (2012)	PLIN1 (rs894160)	GOLDN Study: Men (n = 462) and women (n = 508)	When the ratio of SFA to CHO was high, insulin and HOMA-IR were significantly higher in minor allele carriers but did not differ when the ratio was low.
Zheng et al. (2013)	PSMD3 (rs4065321, rs709592)	GOLDN and BPRHS	Both significantly interacted with dietary MUFAs and carbohydrate on glucose concentrations.
Garaulet et al. (2014)	REV-ERB-ALPHA1 (rs2314339)	2214 subjects, Spanish Mediterranean (n = 1404) and North American (n = 810)	A significant interaction between the REV-ERB-ALPHA1 variant and MUFA intake for obesity was detected in the Mediterranean population.
Perez-Martinez et al. (2005)	SCARB1 (rs4238001)	59 healthy volunteers	Carriers of the G/A genotype have significant increases in insulin sensitivity after a MUFA -rich diet compared with G/G individuals.
Perez-Martinez et al. (2008)	SERPINE1 (rs34857375)	59 healthy volunteers	Subjects carrying the 4G allele (4G/4G and 4G/5G) showed a significant decrease in PAI-1 concentrations after the MUFA diet, compared with the SFA -rich.
Phillips et al. (2009)	STAT3 (rs8069645, rs744166, rs2306580, rs2293152, rs10530050)	LIPGENE-SU.VI.MAX Study of MetS cases and matched controls (n = 1754).	STAT3 SNPs influenced the risk of abdominal obesity , which is modulated by dietary SFA intake.
Phillips et al. (2012)	TCF7L2 (rs7903146)	LIPGENE-SU.VI.MAX Study of MetS cases and matched controls (n = 1754).	The TCF7L2 (rs7903146) SNP influences MetS risk, which is augmented by both gender and dietary SFA intake.
Corella et al. (2013)	TCF7L2 (rs7903146)	7018 participants in the PREvención con DIetaMEDiterránea Study	MedDiet, rich in MUFA , reduces increased fasting glucose and lipids in TT individuals, as well as stroke incidence.
Fernandez-Real et al. (2013)	THRA (rs12939700)	Two population cohorts at baseline (n = 3417 and n = 2265), 6 years later (n = 2139) and (2) in 4734 high cardiovascular risk subjects (HCVR, PREDIMED trial).	Only when SFA intake was high (>24.5 g/d), GG carriers showed a significantly higher BMI than A carriers after controlling for energy intake and physical activity.

TABLE 27.1 Summary of research involving gene by diet (saturated and monounsaturated fat) in relation to CVD risk factors and CVD.—cont'd

Reference	Gene/SNPs	Population	Outcome
Fernandez-Real et al. (2013)	TLR4 (rs5030728)	The Toronto Nutrigenomics and Health Study (n = 676)	SFA intake was inversely associated with HDL-C among GG individuals, whereas a positive relationship was observed for heterozygotes. There was no association between dietary SFAs and HDL-C among AA individuals.
Joffe et al. (2011)	TNF (rs1800629)	Normal-weight (N = 88) and obese (N = 60) white SA women	With increasing SFA intake (%E), serum TC levels decreased for the GG genotype and increased for the GA plus AA genotypes.
Lai et al. (2009)	WDC1 (rs4460661)	BPRHS (n = 1115)	As dietary MUFA intake increased, minor allele carriers of SNP i22835A > G had higher BMIs, whereas major allele carriers had lower BMIs.
Casas-Agustench et al. (2014)	Weighted GRS calculated on the basis of 63 obesity-associated variants	Cross-sectional analyses included 783 white US participants from GOLDN and 2035 from MESA.	Potential dietary recommendations to reduce BMI effectively in populations with high obesity GRS would be to reduce total fat intake mainly by limiting SFAs.

(CI) 1.4–6.2]; $P = .006$) when compared with CC participants consuming such diet. Most interesting, TT participants consuming the MedDiet, had a stroke incidence over the study period that was similar to CC participants on the MedDiet (adjusted HR 0.96 [95% CI 0.49–1.87]; $P = .892$ for TT compared with CC). Whereas the results from this large, long-term intervention study need to be validated in other populations, these initial results strongly support that a MedDiet reduces the expected increase in fasting glucose and lipids associated with the T allele, but, most importantly, it contributes to reducing stroke incidence in this high-risk segment of the population represented by TT subjects.

Whereas in the previous instance we have demonstrated that a “healthy diet” can reduce increased genetic risk, the following example demonstrates how an “unhealthy diet” can augment or bring up the deleterious effect of a specific polymorphism. In this regard, we have reported a consistent interaction between the -265 APOA2 T > C (rs5082) SNP, SFAs, and the body mass index (BMI).

Apolipoprotein A-II (APOA2), a highly abundant serum protein, comprises approximately 20% of high-density lipoprotein (HDL) protein mass. The biological role of APOA2 has remained elusive over more than 4 decades since its discovery, but animal models indicate that APOA2 influences lipoprotein metabolism and dietary intake behaviors. While the role of APOA2 in humans is less clear, our genetic studies indicate an association of this common rs5082 SNP with postprandial

metabolism of large triglyceride-rich lipoproteins. This SNP likely affects a key regulatory cis-element in the APOA2 promoter, altering mRNA expression and APOA2 plasma levels. We have repeatedly shown in global populations of several ancestries a gene-diet interaction where APOA2 CC homozygotes (~16% of the population) show increased risk of obesity when consuming a diet high in SFAs compared to TT subjects (~38% of the population of European descent) (see Table 27.1). Moreover, we have shown that carriers of the CC genotype have, on average, a higher energy intake and preference for protein- and fat-rich foods. These consistently replicated interactions imply that obesity prevention strategies would show greater success when tailored to the genotype, but may also have much wider implications for optimizing diet to improve cardiometabolic health. Despite the consistency of this APOA2 by SFA interaction in populations around the world, evidence is limited to observational studies, and the underlying mechanism remains unknown. In addition, complete investigation of nutritional health effects at the molecular level requires integrating the interplay between the human host and the gut microbiome.

Whereas the APOA2 locus provides a powerful example of the interplay between nutrition, genetics, and BMI for the scientific community, its translation to the clinical practice is limited by the fact that BMI is regulated by scores of genes. In addition, any practical tool to be used in the implementation of precision nutrition has to be consistent with such biological reality.

Thus, for that purpose, it is imperative to combine multiple genetic variants related to obesity into a genetic risk score (GRS). This will improve the identification of individuals at risk of developing obesity and the development of useful tools to implement personalized dietary recommendations to individuals with higher predisposition to obesity. For this purpose, we analyzed the association between an obesity GRS and BMI in the Genetics of Lipid Lowering Drugs and Diet Network (GOLDN) population ($n = 783$), focusing on gene-diet interactions with total fat and SFA intake, and replicated our findings in the Multi-Ethnic Study of Atherosclerosis (MESA) population ($n = 2035$). We calculated a weighted GRS using 63 obesity-associated variants. We used multiple linear regression models adjusted by potential confounders to examine gene-diet interactions between dietary intake of total fat and SFAs and the obesity GRS in determining BMI. We found significant interactions between total fat intake and the obesity GRS using these variables as continuous for BMI (P for interaction for the meta-analysis of both populations = 0.002). Our findings gained statistical strength when we assessed interactions more specifically using SFA intake, the GRS, and BMI (P for interaction <0.001 in the meta-analysis). Therefore, SFA intake interacts with an obesity GRS in modulating BMI, suggesting that successful dietary recommendations to reduce BMI effectively in populations with high obesity GRS would be to reduce total fat intake mainly by restricting SFAs.

Despite the developments in this field and the data that have been accumulating over the years, we lack enough evidence to unequivocally implement nutritional recommendations based on genetic information for the prevention of CVD. In order to achieve this objective we need to support the validity of past and future findings by seeking external validation in multiple populations and by conducting randomized clinical trials in genotype-selected individuals. However, even failure to replicate may be informative and relevant as long as the experimental design is appropriate and we can identify the basis for such differences (i.e., age, gender, ethnicity, diabetes, obesity).

Whereas for the purpose of simplifying the presentation of different nutrients this work focuses only on gene interactions with SFAs and MUFAs in relation to cardio-metabolic traits, it is important to underscore that precision nutrition will need to integrate the information of the genome, the epigenome, the microbiome, and deep phenotyping (i.e., metabolomics) in combination with the exposome. With vast improvements and cost reductions on the required technology and the ability to capture data in real time, this close to utopian undertaking may become a reality in the near future. This should end the long-term search for the “healthy diet” and focus on the identification of which dietary

pattern is “healthy” for the individual in order to achieve and maintains its maximum physical and mental capacity through nutrition.

References

- Bouchard-Mercier, A., Godin, G., Lamarche, B., et al., 2011. Effects of peroxisome proliferator-activated receptors, dietary fat intakes and gene-diet interactions on peak particle diameters of lowdensity lipoproteins. *J Nutrigenet Nutrigenomics* 4 (1), 36–48. <https://doi.org/10.1159/000324531>.
- Brown, S., Ordovás, J.M., Campos, H., 2003. Interaction between the APOC3 gene promoter polymorphisms, saturated fat intake and plasma lipoproteins. *Atherosclerosis* 170 (2), 307–313. PMID: 14612212.
- Campos, H., López-Miranda, J., Rodríguez, C., et al., 1997. Urbanization elicits a more atherogenic lipoprotein profile in carriers of the apolipoprotein A-IV-2 allele than in A-IV-1 homozygotes. *Arterioscler Thromb Vasc Biol* 17 (6), 1074–1081. PMID: 9194757.
- Casas-Agustench, P., Arnett, D.K., Smith, C.E., et al., 2014. Saturated fat intake modulates the association between an obesity genetic risk score and body mass index in two US populations. *J Acad Nutr Diet* 114 (12), 1954–1966. <https://doi.org/10.1016/j.jand.2014.03.014>.
- Corella, D., Peloso, G., Arnett, D.K., et al., 2009. APOA2, dietary fat, and body mass index: replication of a gene-diet interaction in 3 independent populations. *Arch Intern Med* 169 (20), 1897–1906. <https://doi.org/10.1001/archinternmed.2009.343>.
- Corella, D., Tai, E.S., Sorlí, J.V., et al., 2011. Association between the APOA2 promoter polymorphism and body weight in Mediterranean and Asian populations: replication of a gene-saturated fat interaction. *Int J Obes (Lond)*. 35 (5), 666–675. <https://doi.org/10.1038/ijo.2010.187>.
- Corella, D., Carrasco, P., Sorli, J.V., et al., 2013. Mediterranean diet reduces the adverse effect of the TCF7L2-rs7903146 polymorphism on cardiovascular risk factors and stroke incidence: a randomized controlled trial in a high-cardiovascular-risk population. *Diabetes Care* 36 (11), 3803–3811. <https://doi.org/10.2337/dc13-0955>.
- Corella, D., Asensio, E.M., Coltell, O., et al., 2016. CLOCK gene variation is associated with incidence of type-2 diabetes and cardiovascular diseases in type-2 diabetic subjects: dietary modulation in the PREDIMED randomized trial. *Cardiovasc Diabetol* 15, 4. <https://doi.org/10.1186/s12933-015-0327-8>.
- Ferguson, J.F., Phillips, C.M., Tierney, A.C., et al., 2010. Gene-nutrient interactions in the metabolic syndrome: single nucleotide polymorphisms in ADIPOQ and ADIPOR1 interact with plasma saturated fatty acids to modulate insulin resistance. *Am J Clin Nutr* 91 (3), 794–801. <https://doi.org/10.3945/ajcn.2009.28255>.
- Fernández-Real, J.M., Corella, D., Goumidi, L., et al., 2013. Thyroid hormone receptor alpha gene variants increase the risk of developing obesity and show gene-diet interactions. *Int J Obes (Lond)* 37 (11), 1499–1505. <https://doi.org/10.1038/ijo.2013.11>.
- Fisher, E., Schreiber, S., Joost, H.G., et al., 2011. A two-step association study identifies CAV2 rs2270188 single nucleotide polymorphism interaction with fat intake in type 2 diabetes risk. *J Nutr* 141 (2), 177–181. <https://doi.org/10.3945/jn.110.124206>.
- Garaulet, M., Lee, Y.C., Shen, J., et al., 2009. CLOCK genetic variation and metabolic syndrome risk: modulation by monounsaturated fatty acids. *Am J Clin Nutr* 90 (6), 1466–1475. <https://doi.org/10.3945/ajcn.2009.27536>.
- Garaulet, M., Smith, C.E., Gomez-Abellan, P., et al., 2014. REV-ERB-ALPHA circadian gene variant associates with obesity in two independent populations: Mediterranean and North American. *Mol*

- Nutr Food Res 58 (4), 821–829. <https://doi.org/10.1002/mnfr.201300361>.
- García-Ríos, A., Pérez-Martínez, P., Delgado-Lista, J., et al., 2012. A Period 2 genetic variant interacts with plasma SFA to modify plasma lipid concentrations in adults with metabolic syndrome. *J Nutr* 142 (7), 1213–1218. <https://doi.org/10.3945/jn.111.156968>.
- Garault, M., Smith, C.E., Hernández-González, T., Lee, Y.C., Ordovas, J.M., 2011 Dec. PPAR γ Pro12Ala interacts with fat intake for obesity and weight loss in a behavioural treatment based on the Mediterranean diet. *Mol Nutr Food Res* 55 (12), 1771–1779. <https://doi.org/10.1002/mnfr.201100437>. Epub 2011 Nov 21. PubMed PMID: 22102511; PubMed Central PMCID: PMC3951915.
- Huang, T., Tucker, K., Lee, Y., Crott, J., Parnell, L., Shen, J., Smith, C., Ordovas, J., Li, D., Lai, C., 2012 Apr. MAT1A variants modulate the effect of dietary fatty acids on plasma homocysteine concentrations. *Nutr Metab Cardiovasc Dis* 22 (4), 362–368. <https://doi.org/10.1016/j.numecd.2010.07.015>. Epub 2010 Dec 24. PubMed PMID: 21185701; PubMed Central PMCID: PMC3117050.
- Joffe, Y.T., van der Merwe, L., Collins, M., et al., 2011. The -308 G/A polymorphism of the tumour necrosis factor- α gene modifies the association between saturated fat intake and serum total cholesterol levels in white South African women. *Genes Nutr* 6 (4), 353–359. <https://doi.org/10.1007/s12263-011-0213-2>.
- Lai, C.Q., Parnell, L.D., Arnett, D.K., et al., 2009. WDTC1, the ortholog of *Drosophila* adipose gene, associates with human obesity, modulated by MUFA intake. *Obesity (Silver Spring)* 17 (3), 593–600. <https://doi.org/10.1038/oby.2008.561>.
- Li, T.Y., Zhang, C., Asselbergs, F.W., Qi, L., et al., 2007. Interaction between dietary fat intake and the cholesterol ester transfer protein TaqIB polymorphism in relation to HDL-cholesterol concentrations among US diabetic men. *Am J Clin Nutr* 86 (5), 1524–1529. PMID: 17991668.
- Lindi, V., Schwab, U., Louheranta, A., et al., 2003. Impact of the Pro12Ala polymorphism of the PPAR- γ 2 gene on serum triacylglycerol response to n-3 fatty acid supplementation. *Mol Genet Metab* 79 (1), 52–60. PMID: 12765846.
- Lindi, V., Schwab, U., Louheranta, A., Vessby, B., Hermansen, K., Tapsell, L., Riccardi, G., Rivellesse, A.A., Laakso, M., Uusitupa, M., 2008 Feb. KANWU Study Group. The G-250A polymorphism in the hepatic lipase gene promoter is associated with changes in hepatic lipase activity and LDL cholesterol: The KANWU Study. *Nutr Metab Cardiovasc Dis* 18 (2), 88–95. Epub 2007 Feb 26. PubMed PMID: 17327141.
- Moreno, J.A., Pérez-Jiménez, F., Marín, C., et al., 2004. The effect of dietary fat on LDL size is influenced by apolipoprotein E genotype in healthy subjects. *J Nutr* 34 (10), 2517–2522. PMID: 15465740.
- Nettleton, J.A., Steffen, L.M., Ballantyne, C.M., et al., 2007. Associations between HDL-cholesterol and polymorphisms in hepatic lipase and lipoprotein lipase genes are modified by dietary fat intake in African American and White adults. *Atherosclerosis* 194 (2) e131–40. PMID: 17157861.
- Ordovas, J.M., Corella, D., Demissie, S., et al., 2002. Dietary fat intake determines the effect of a common polymorphism in the hepatic lipase gene promoter on high-density lipoprotein metabolism: evidence of a strong dose effect in this gene-nutrient interaction in the Framingham Study. *Circulation* 106 (18), 2315–2321. PMID: 12403660.
- Ortega-Azorin, C., Sorli, J.V., Estruch, R., et al., 2014. Amino acid change in the carbohydrate response element binding protein is associated with lower triglycerides and myocardial infarction incidence depending on level of adherence to the Mediterranean diet in the PREDIMED trial. *Circ Cardiovasc Genet* 7 (1), 49–58. <https://doi.org/10.1161/CIRCGENETICS.113.000301>.
- Parnell, L.D., Blokker, B.A., Dashti, H.S., et al., 2014. CardioGxE, a catalog of gene-environment interactions for cardiometabolic traits. *BioData Min* 7, 21. <https://doi.org/10.1186/1756-0381-7-21>.
- Pérez-Martínez, P., Adarraga-Cansino, M.D., Fernández de la Puebla, R.A., et al., 2008. The -675 4G/5G polymorphism at the Plasminogen Activator Inhibitor 1 (PAI-1) gene modulates plasma Plasminogen Activator Inhibitor 1 concentrations in response to dietary fat consumption. *Br J Nutr* 99 (4), 699–702. PMID: 17903340.
- Pérez-Martínez, P., Pérez-Jiménez, F., Bellido, C., et al., 2005. A polymorphism exon 1 variant at the locus of the scavenger receptor class B type I (SCARB1) gene is associated with differences in insulin sensitivity in healthy people during the consumption of an olive oil-rich diet. *J Clin Endocrinol Metab* 90 (4), 2297–2300. PMID: 15671101.
- Phillips, C.M., Goumidi, L., Bertrais, S., et al., 2009. Dietary saturated fat modulates the association between STAT3 polymorphisms and abdominal obesity in adults. *J Nutr* 139 (11), 2011–2017. <https://doi.org/10.3945/jn.109.110635>.
- Phillips, C.M., Goumidi, L., Bertrais, S., et al., 2012. Dietary saturated fat, gender and genetic variation at the TCF7L2 locus predict the development of metabolic syndrome. *J Nutr Biochem* 23 (3), 239–244. <https://doi.org/10.1016/j.jnutbio.2010.11.020>.
- Phillips, C.M., Kesse-Guyot, E., McManus, R., et al., 2012. High dietary saturated fat intake accentuates obesity risk associated with the fat mass and obesity-associated gene in adults. *J Nutr* 142 (5), 824–831. <https://doi.org/10.3945/jn.111.153460>.
- Robitaille, J., Després, J.P., Pérusse, L., Vohl, M.C., 2003. The PPAR- γ P12A polymorphism modulates the relationship between dietary fat intake and components of the metabolic syndrome: results from the Québec Family Study. *Clin Genet* 63 (2), 109–116. PMID: 12630956.
- Rosado, E.L., Bressan, J., Martínez, J.A., Marques-Lopes, I., 2010. Interactions of the PPAR γ 2 polymorphism with fat intake affecting energy metabolism and nutritional outcomes in obese women. *Ann Nutr Metab* 57 (3-4), 242–250. <https://doi.org/10.1159/000322187>.
- Rudkowska, I., Dewailly, E., Hegele, R.A., et al., 2013. Gene-diet interactions on plasma lipid levels in the Inuit population. *Br J Nutr* 109 (5), 953–961. <https://doi.org/10.1017/S0007114512002231>.
- Salas, J., Jansen, S., López-Miranda, J., et al., 1998. The SstI polymorphism of the apolipoprotein C-III gene determines the insulin response to an oral-glucose-tolerance test after consumption of a diet rich in saturated fats. *Am J Clin Nutr* 68 (2), 396–401. PMID: 9701199.
- Schuler, R., Osterhoff, M.A., Frahnkow, T., et al., 2017. High-saturated-fat diet increases circulating angiotensin-converting enzyme, which is enhanced by the rs4343 polymorphism defining persons at risk of nutrient-dependent increases of blood pressure. *J Am Heart Assoc* 6 (1). <https://doi.org/10.1161/JAHA.116.004465>.
- Smith, C.E., Arnett, D.K., Corella, D., et al., 2012. Perilipin polymorphism interacts with saturated fat and carbohydrates to modulate insulin resistance. *Nutr Metab Cardiovasc Dis* 22 (5), 449–455. <https://doi.org/10.1016/j.numecd.2010.09.003>.
- Smith, C.E., Tucker, K.L., Arnett, D.K., et al., 2013. Apolipoprotein A2 polymorphism interacts with intakes of dairy foods to influence body weight in 2 U.S. populations. *J Nutr* 143 (12), 1865–1871. <https://doi.org/10.3945/jn.113.179051>.
- Wallace, A.J., Mann, J.I., Sutherland, W.H., et al., 2000. Variants in the cholesterol ester transfer protein and lipoprotein lipase genes are predictors of plasma cholesterol response to dietary change. *Atherosclerosis* 152 (2), 327–336. PMID: 10998460.
- Weggemans, R.M., Zock, P.L., Ordovas, J.M., et al., 2001. Genetic polymorphisms and lipid response to dietary changes in humans. *Eur J Clin Invest* 31 (11), 950–957. PMID: 11737237.

- Yang, Y., Ruiz-Narvaez, E., Kraft, P., Campos, H., 2007. Effect of apolipoprotein E genotype and saturated fat intake on plasma lipids and myocardial infarction in the Central Valley of Costa Rica. *Hum Biol* 79 (6), 637–647. PMID: 18494374.
- Zhang, C., Lopez-Ridaura, R., Rimm, E.B., Rifai, N., Hunter, D.J., Hu, F.B., 2005 Jun. Interactions between the -514C->T polymorphism of the hepatic lipase gene and lifestyle factors in relation to HDL concentrations among US diabetic men. *Am J Clin Nutr* 81 (6), 1429–1435. PubMed PMID: 15941898.
- Zheng, J.S., Arnett, D.K., Parnell, L.D., et al., 2013. Modulation by dietary fat and carbohydrate of IRS1 association with type 2 diabetes traits in two populations of different ancestries. *Diabetes Care* 36 (9), 2621–2627. <https://doi.org/10.2337/dc12-2607>.

Nutrigenetics: Omega-3 and Omega-6 Fatty Acids

William S. Harris

Department of Internal Medicine, University of South Dakota; and OmegaQuant, LLC, Sioux Falls, SD, United States

INTRODUCTION

Both the long-chain omega-3 fatty acids (n-3 FAs)—eicosapentaenoic and docosahexaenoic acids, EPA and DHA—and the principal long-chain n-6 FA (arachidonic acid, AA) play important roles in metabolism. As precursors to a wide variety of active biomolecules collectively termed “oxylipins,” higher/lower levels of these FAs impact important and disparate clinical outcomes such as cardiovascular disease (De Caterina, 2011), depression (Sublette et al., 2011), and dry eye disease (Epitropoulos et al., 2016), to name but a few. Levels of the n-3 FAs are determined partly by biosynthesis from the plant-derived precursor alpha-linolenic acid and partly from direct dietary intake of EPA and DHA from seafoods. AA levels, on the other hand, are primarily determined by metabolic conversion from the essential n-6 FA linoleic acid (Fig. 28.1). This is reflected in population coefficients of variation for these three FAs in red blood cell (RBC) membranes, which, based on data from our lab, is 78%, 31%, and 10% for EPA, DHA, and AA, respectively. The role that genetic factors play in determining the levels of these FAs is poorly understood. This chapter will focus on two specific questions: How much do genetic variations influence (1) baseline EPA, DHA, and AA levels in blood, and (2) the blood EPA and DHA response to n-3 supplementation?

EPA, DHA, AND AA LEVELS AND GENETIC VARIANTS

We recently conducted a genome-wide association study (GWAS) in the Framingham Heart Study's Offspring Cohort (Tintle et al., 2015). In this study, we

examined the relations between RBC FA levels and over 2.5M single nucleotide polymorphisms (SNPs) in 2633 subjects. Although there were 191 SNPs associated with one or more FAs, virtually none were significantly associated with EPA or DHA levels. AA levels, on the other hand, were significantly linked with 82 separate SNPs, and the overall variability in AA levels explained by these genetic variables was 13%. This pattern—AA relatively strongly linked with genetic variants and EPA/DHA weakly linked, if at all—is a recurring theme.

RBC FAs were measured in 4457 pregnant women in a study by Koletzko et al. (2011), along with 17 SNPs in the fatty acid desaturase-1 (FADS-1) and -2 regions. Associations were stronger for AA than for EPA and DHA. Nevertheless, they were relatively small across the board as reflected in the variation in FA levels explained by all SNPs: AA explained 1.1% of the variability, DHA, 0.5% and EPA, 0.27%. Another study of pregnant women found similar rank order using plasma instead of RBC FAs (Xie and Innis, 2008). A consortium-based GWAS study reported that FADS1 SNPs explained 2% of EPA and 0% of DHA variance, and variations in the elongase 2 gene (ELOVL2) SNPs explained only 0.4% of EPA and 0.7% of DHA variance (Guan et al., 2014). In 658 participants in the Verona Heart Project, 13 SNPs were genotyped in the FADS1-3 cluster on chromosome 11, and the FA composition of plasma phospholipids and RBC membranes was determined. There was no association with these genetic variants on EPA or DHA levels, but there was a strong relationship with AA levels (Malerba et al., 2008). Atypical findings came from a study in 1504 Chinese subjects where one SNP in FADS1 (rs66698963) was measured along with plasma FA levels. There were three genotypes examined: I/I, I/D, and D/D. Comparing genotype I/I versus D/D, AA levels were 57% higher, EPA levels were 90%

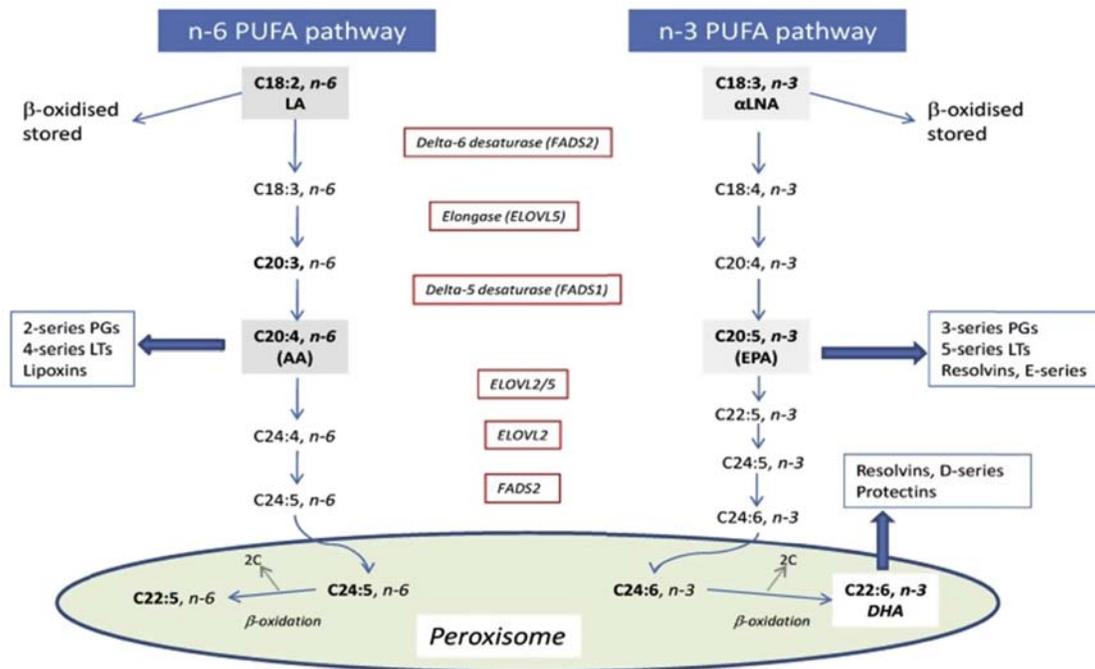


FIGURE 28.1 The biosynthetic pathways for the n-6 and n-3 polyunsaturated fatty acids. From Minihane, A.M., 2016. *Impact of genotype on EPA and DHA status and responsiveness to increased intakes. Nutrients 8, 123 with permission.*

higher, but DHA levels were not different (Li et al., 2018). An examination of the effects of variants of ELOVL2 on baseline PUFA levels in 367 subjects found that all three minor variants were associated with 6% lower DHA levels (Alsaleh et al., 2014) but no effects were seen on EPA levels.

Schaeffer et al. (2006) reported that 7% of the variability in plasma EPA levels was explained by haplotypes in the FADS genes as was about 3% of the DHA variability. For ARA, the variability explained by this haplotype was 29%. In a study of Danish infants at 9 months and 3 years of age, RBC DHA was influenced by FADS genotype at the earlier but not the later time point. However, even at 9 months, the variation in DHA levels explained by variants was only about 3% (Harslof et al., 2013). Overall, there appears to be minimal genetic influence on circulating levels of the major long-chain n-3 PUFAs, where AA levels appear to be more readily affected by gene variants.

GENETIC INFLUENCES ON THE BLOOD N-3 PUFA RESPONSE TO N-3 SUPPLEMENTATION

Not only is there variability in background or baseline n-6, and particularly n-3 PUFA levels, there is also remarkable variability in the *change* in blood n-3 PUFA levels after giving fish oil supplements. The study of Flock et al. (2013) is illustrative (Fig. 28.2). In this study,

115 healthy volunteers were randomized to placebo or one of four doses of EPA + DHA (300, 600, 900, or 1800 mg/day) for 5 months. The average baseline RBC EPA + DHA level (the Omega-3 Index) was about 4.3%. There was a clear dose-response relationship observed, but the variability within each dose group was marked. For example, in the 1800 mg/day group, the change in the Omega-3 Index ranged from 0.5% to 7%—a 14-fold difference in response. Clearly compliance failures could have contributed to this, but the authors reported that compliance was 97% by capsule count. So the observed variability in response was the result of other factors, but the role that genetics plays is unknown.

Very few researchers have examined the question of how genetic factors may influence the blood n-3 PUFA response to an increased EPA and DHA intake. Alsaleh et al. (2014) randomized 367 subjects to three ascending doses of fish oil for 6 months. Variants in the ELOVL2 gene were associated with lower plasma DHA levels at baseline, but a commensurate increase in EPA or docosapentaenoic acid n3 (DPA; as might be expected with lower conversion to the DHA product) were seen. With respect to effects on plasma EPA levels postsupplementation, subjects with the minor alleles got a *smaller* rise in EPA on the 900 mg dose but a *greater* rise in EPA on the 1800 mg dose. The same pattern played out with DHA, except the lower levels were seen with a dose of 450 mg, with no difference between genotypes at 900 mg/d. It is unclear how a mutation in an enzyme that controls the conversion of EPA to DPA (and then

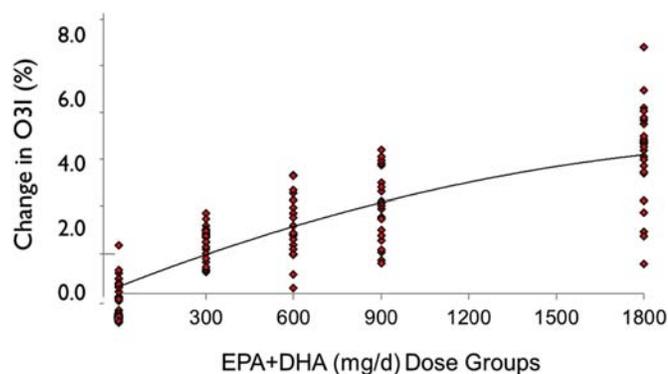


FIGURE 28.2 Variability in the red blood cell EPA + DHA level (the Omega-3 Index, or O3I) in response to supplementation. In this study, 115 healthy volunteers were randomized to placebo or one of four doses of EPA + DHA (300, 600, 900, or 1800 mg/day) for 5 months. The average baseline O3I values were about 4.3%. The changes in percentage points for each individual in each dosage group are shown. Adapted from Flock, M.R., Skulas-Ray, A.C., Harris, W.S., Etherton, T.D., Fleming, J.A., Kris-Etherton, P.M., 2013. Determinants of erythrocyte omega-3 fatty acid content in response to fish oil supplementation: a dose-response randomized controlled trial. *J Am Heart Assoc* 2, e000513.

DHA) could cause a greater rise in plasma n-3 PUFAs at one fish oil dose and a smaller rise with another dose. This finding clearly needs confirmation. In another report from the same study, FADS variants were not associated with the FA response to supplementation (Al-Hilal et al., 2013). There have been no studies examining the effects of genetic factors outside of the FADS and ELOVL on the blood n-3 PUFA responses to fish oil. In a study that examined the effects of EPA and DHA on biosynthetic gene expression, Allaire et al. (2017) found that the former increased ELOVL2 by 24% but did not affect FADS1 or 2, or ELOVL5. DHA supplementation had no effect on any of these genes.

CONCLUSIONS

The study of genetic influences on in vivo PUFA status is in its infancy. We have reasonably good evidence for a relatively strong effect of FADS variants on AA levels (i.e., levels are lower with minor variants), but not of EPA and DHA levels. We have very little evidence regarding the genetic factors that influence the blood EPA and DHA response to fish oil supplementation. Given the marked heterogeneity in the latter that cannot be otherwise explained, coupled with the growing body of evidence for health benefits from higher blood n-3 PUFA levels (Jackson and Harris, 2018), continued research into this question is strongly encouraged.

References

- Al-Hilal, M., Alsaleh, A., Maniou, Z., Lewis, F.J., Hall, W.L., Sanders, T.A., O'Dell, S.D., 2013. Genetic variation at the FADS1-FADS2 gene locus influences delta-5 desaturase activity and LC-PUFA proportions after fish oil supplement. *J Lipid Res* 54, 542–551.
- Allaire, J., Harris, W.S., Vors, C., Charest, A., Marin, J., Jackson, K.H., Tchernof, A., Couture, P., Lamarche, B., 2017. Supplementation with high-dose docosahexaenoic acid increases the Omega-3 Index more than high-dose eicosapentaenoic acid. *Prostaglandins Leukot Essent Fatty Acids* 120, 8–14.
- Alsaleh, A., Maniou, Z., Lewis, F.J., Hall, W.L., Sanders, T.A., O'Dell, S.D., 2014. ELOVL2 gene polymorphisms are associated with increases in plasma eicosapentaenoic and docosahexaenoic acid proportions after fish oil supplement. *Genes Nutr* 9, 362.
- De Caterina, R., 2011. n-3 fatty acids in cardiovascular disease. *N Engl J Med* 364, 2439–2450.
- Epitropoulos, A.T., Donnenfeld, E.D., Shah, Z.A., Holland, E.J., Gross, M., Faulkner, W.J., Matossian, C., Lane, S.S., Toyos, M., Bucci Jr, F.A., Perry, H.D., 2016. Effect of oral Re-esterified omega-3 nutritional supplementation on dry eyes. *Cornea* 35, 1185–1191.
- Flock, M.R., Skulas-Ray, A.C., Harris, W.S., Etherton, T.D., Fleming, J.A., Kris-Etherton, P.M., 2013. Determinants of erythrocyte omega-3 fatty acid content in response to fish oil supplementation: a dose-response randomized controlled trial. *J Am Heart Assoc* 2, e000513.
- Guan, W., Steffen, B.T., Lemaitre, R.N., Wu, J.H., Tanaka, T., Manichaikul, A., Foy, M., Rich, S.S., Wang, L., Nettleton, J.A., Tang, W., Gu, X., Bandinelli, S., King, I.B., McKnight, B., Psaty, B.M., Siscovick, D., Djousse, L., Ida Chen, Y.D., Ferrucci, L., Fornage, M., Mozafarian, D., Tsai, M.Y., Steffen, L.M., 2014. Genome-wide association study of plasma N6 polyunsaturated fatty acids within the cohorts for heart and aging research in genomic epidemiology consortium. *Circulation. Cardiovascular genetics*. 7, 321–331.
- Harslof, L.B., Larsen, L.H., Ritz, C., Hellgren, L.I., Michaelsen, K.F., Vogel, U., Lauritzen, L., 2013. FADS genotype and diet are important determinants of DHA status: a cross-sectional study in Danish infants. *Am J Clin Nutr* 97, 1403–1410.
- Jackson, K.H., Harris, W.S., 2018. Blood fatty acid profiles: new biomarkers for cardiometabolic disease risk. *Curr Atheroscler Rep* 20, 22.
- Koletzko, B., Lattka, E., Zeilinger, S., Illig, T., Steer, C., 2011. Genetic variants of the fatty acid desaturase gene cluster predict amounts of red blood cell docosahexaenoic and other polyunsaturated fatty acids in pregnant women: findings from the Avon Longitudinal Study of Parents and Children. *Am J Clin Nutr* 93, 211–219.
- Li, P., Zhao, J., Kothapalli, K.S.D., Li, X., Li, H., Han, Y., Mi, S., Zhao, W., Li, Q., Zhang, H., Song, Y., Brenna, J.T., Gao, Y., 2018. A regulatory insertion-deletion polymorphism in the FADS gene cluster influences PUFA and lipid profiles among Chinese adults: a population-based study. *Am J Clin Nutr* 107, 867–875.
- Malerba, G., Schaeffer, L., Xumerle, L., Klopp, N., Trabetti, E., Biscuola, M., Cavallari, U., Galavotti, R., Martinelli, N., Guarini, P., Girelli, D., Olivieri, O., Corrocher, R., Heinrich, J., Pignatti, P.F., Illig, T., 2008. SNPs of the FADS gene cluster are associated with polyunsaturated fatty acids in a cohort of patients with cardiovascular disease. *Lipids* 43 (4), 289–299. Epub.2008.Mar.5. 43 (2008).
- Minihane, A.M., 2016. Impact of genotype on EPA and DHA status and responsiveness to increased intakes. *Nutrients* 8, 123.
- Schaeffer, L., Gohlke, H., Muller, M., Heid, I.M., Palmer, L.J., Kompauer, I., Demmelmair, H., Illig, T., Koletzko, B., Heinrich, J., 2006. Common genetic variants of the FADS1 FADS2 gene cluster

- and their reconstructed haplotypes are associated with the fatty acid composition in phospholipids. *Hum Mol Genet* 15, 1745–1756.
- Sublette, M.E., Ellis, S.P., Geant, A.L., Mann, J.J., 2011. Meta-analysis of the effects of eicosapentaenoic acid (EPA) in clinical trials in depression. *J Clin Psychiatry* 72, 1577–1584.
- Tintle, N.L., Pottala, J.V., Lacey, S., Ramachandran, V., Westra, J., Rogers, A., Clark, J., Olthoff, B., Larson, M., Harris, W., Shearer, G.C., 2015. A genome-wide association study of saturated, mono- and polyunsaturated red blood cell fatty acids in the Framingham Heart Offspring Study. *Prostaglandins Leukot Essent Fatty Acids* 94, 65–72.
- Xie, L., Innis, S.M., 2008. Genetic variants of the FADS1 FADS2 gene cluster are associated with altered (n-6) and (n-3) essential fatty acids in plasma and erythrocyte phospholipids in women during pregnancy and in breast milk during lactation. *J Nutr* 138, 2222–2228.

Nutrigenetics and Blood Cholesterol Levels in Response to Plant Sterols

Itzel Vazquez-Vidal¹, Peter J.H. Jones²

¹Richardson Centre for Functional Foods and Nutraceuticals, University of Manitoba, Winnipeg, MB, Canada;

²Nutritional Fundamentals for Health Inc., Vaudreuil-Dorion, QC, Canada

Glossary

APOE The apolipoprotein E (*APOE*) gene encodes a protein called apolipoprotein E (apoE), which is one of the major proteins that the body uses to transport cholesterol and triglycerides around the body.

ABCG5/G8 These proteins are responsible for promoting cholesterol excretion through the reverse cholesterol transport pathway.

CYP7A1 Cytochrome P450 family 7 subfamily A member 1 (*CYP7A1*) gene encodes a rate-limiting enzyme, cholesterol 5 α -hydroxylase, in the bile acid synthesis pathway.

CETP Cholesteryl ester transfer protein (*CETP*) is a protein-coding gene located on chromosome 16 involved in the transfer of cholesteryl ester from high-density lipoprotein (HDL) to other lipoproteins.

NPC1L1 Niemann-Pick C1-like 1 (*NPC1L1*) transporter is a multipass membrane protein that takes up free cholesterol into cells through vesicular endocytosis and is essential for cholesterol and vitamin E absorption.

differences in *APOE*, *CYP7A1*, *ABCG8*, *CETP*, and *NPC1L1* genes. The purpose of this review is to summarize the most recent evidence describing the effect of genetic variation on blood cholesterol concentrations in response to PS consumption (Table 29.1).

The *apolipoprotein E (APOE)* gene encodes an essential protein (apoE) for assembling of triglyceride-rich lipoproteins. Previous studies showed that carriers of the *APOE* ϵ 3 and ϵ 2 alleles had lower levels of TC and LDL-C levels after consuming 1.1 or 2.2 g/day of PS compared to their cholesterol levels when consuming a placebo diet (Sanchez-Muñiz et al., 2009). A recent study showed that in 63 mildly hypercholesterolemic individuals who consumed 2 g/day of PS for 28 days only carriers of the *APOE* ϵ 4 isoform had greater LDL-C lowering compared to carriers of the *APOE* ϵ 3 isoform (-0.31 ± 0.007 vs. -0.13 ± 0.05 mmol/L, respectively) (Mackay et al., 2015a). Therefore, these studies provide evidence that *APOE* polymorphisms play a role in plasma lipid response to PS.

The *cytochrome P450 family 7 subfamily A member 1 (CYP7A1)* gene encodes a rate-limiting enzyme, cholesterol 5 α -hydroxylase, in the bile acid synthesis pathway. In a study including 63 mildly hypercholesterolemic adults using a crossover design, participants were provided 2 g/day of PS for 4 weeks. At the endpoint, the decrease in LDL-C was higher in subjects carrying the G-allele of *CYP7A1*-rs3808607 compared to AA carriers (-0.46 vs. -0.05 mmol/L, respectively) (Mackay et al., 2015a). On the other hand, TT homozygotes showed no benefit in response to PS. Additionally, this study found an additive effect of a dual genotype on the responsiveness of cholesterol to PS. Moreover, carriers of a combination of GG genotype of then *CYP7A1*-rs3808607 and *APOE* ϵ 4 had a larger decrease in LDL-C levels in

INTRODUCTION

Epidemiological data have demonstrated that diets rich in specific lipid components including mono- and polyunsaturated fatty acids (MUFAs and PUFAs), low in saturated fat, and rich in plant sterols/stanols (PSs) reduce the risk of coronary heart disease by 20%–25% (Rosato et al., 2017; Kohler et al., 2017). The similar structure of PS to cholesterol causes a decrease in the synthesis and excretion of bile acids, leading to a reduction in cholesterol absorption and therefore lowering circulating levels (Gylling et al., 2014). Although it is well known that a dosage of 1.0–3.0 g/day of PS decreases LDL-cholesterol levels by approximately 12 mg/dL (0.31 mmol/L), substantial interindividual variability in response to PS consumption exists across individuals (Jones, 2015). This is likely due to genetic interindividual

TABLE 29.1 Summary of studies reporting gene-diet interactions on blood cholesterol concentrations in response to plant sterols (PSs).

Author	Gene	SNP	Study design	Diet	Population (n)	Ethnicity	Gene-nutrient interaction
Granado-Lorencio et al. (2014)	<i>ABCG8</i>	rs6544718	Crossover	750 µg β-cryptoxanthin and 1.5 g/day of PS during 28 days	38 adult F	Spanish	Lower TC levels in women CC carriers compared to CT/TT carriers
Zhao et al. (2008)	<i>ABCG8</i>	rs4148217	Crossover	4 week of fat spread with or without 2 g/day of PS	82 hypercholesterolemic M	Caucasian	Lower TC and LDL-C levels in A-allele carriers with high basal PS levels compared to C-allele carriers
Mackay et al. (2015a)	<i>APOE</i>	rs429258 rs7412	Crossover	2 g/day of PS during 28 days	63 adults	Canadian	Higher reduction in LDL-C levels in <i>APOE</i> ε4 carriers compared to <i>APOE</i> ε3
Mackay et al. (2015b)	<i>CETP</i>	rs5882	Crossover	4 weeks of 2 g/day of PS	71 adults	Canadian	Higher TG reduction in GG carriers compared to AA or AG carriers
Mackay et al. (2015a)	<i>CYP7A1</i>	rs3808607	Crossover	2 g/day of PS during 28 day	63 adults	Canadian	Lower LDL-C levels in G-allele carriers compared to A-allele carriers
Granado-Lorencio et al. (2014)	<i>NPC1L1</i>	rs2072183	Crossover	1.5 g/day during 28 day	19 postmenopausal women	Spanish	Higher TC and LDL-C levels in CC carriers compared to GG carriers
Zhao et al. (2008)	<i>NPC1L1</i>	rs2072183 rs10264715	Crossover	4 week of 2 g/day of PS	82 hypercholesterolemic M	Caucasian	Lower TC and LDL-C levels in G-allele rs2072183 + A-allele rs10264715 compared to carriers of the wild-type allele

Abbreviations: *ABCG8*, ATP-binding cassette subfamily G8; *APOE*, Apolipoprotein E; *CETP*, Cholesteryl ester transfer protein; *CYP7A1*, Cholesterol 7-alpha-hydroxylase; F, Females; LDL-C, Low-density lipoprotein cholesterol; M, Males; *NPC1L1*, Niemann-Pick C1-like 1; PSs, Plant sterols; TC, Total cholesterol.

response to consumption of 2 g/day of PS after 28 days compared to any other genotypes ($\Delta-0.67 \pm 0.22$ mmol/L) (Mackay et al., 2015a). However, carriers of the TT genotype of *CYP7A1*-rs30808607 and carriers of *APOE* ϵ 3 showed an increase in LDL-C levels in response of PS ($\Delta+0.09 \pm 0.08$ mmol/L) (Fig. 29.1) (Data adapted from Mackay et al.: *Am J Clin Nutr* 102(4):951–957, 2015a). These results provide evidence of the high susceptibility of certain individuals to have an adverse response to PS.

Another important key player in the lowering effect of PS on LDL-C concentrations is the influence of genetic variants in the *ATP-binding cassette subfamily G member G5* (*ABCG5*) and *G8* (*ABCG8*) genes. A crossover study including 19 postmenopausal women showed that after consumption of 1.5 g/day of PS in a fruit-based drink (250 mL)/d for 4 weeks, CC carriers of the *ABCG8* SNP rs6544718 had a higher decrease in TC levels compared to decreases seen in CT/TT carriers (-0.10 vs. -0.070 mmol/l) (Granado-Lorencio et al., 2014). Moreover, after consumption of 1.5 g/day of PS plus 750 μ g β -cryptoxanthin (β -Cx), CC carriers showed a greater decrease in TC concentrations compared to that observed after PS alone (-0.10 vs. -0.34 mmol/L). Overall, CC carriers showed a higher decrease in TC levels compared to CT/TT carriers after PS plus 750 μ g β -cryptoxanthin (β -Cx) consumption (-0.34 vs. -0.097 mmol/L, respectively) (Granado-Lorencio et al., 2014). It is possible that only women CC carriers of the *ABCG8* SNP rs6544718 will benefit more from consuming PS plus 750 μ g β -cryptoxanthin to decrease TC levels. Similarly, in a crossover study including 82 hypercholesterolemic men, carriers of the A-allele of *ABCG8* SNP rs4148217 showed 3.9-fold higher reductions in serum TC and LDL-C levels compared to C-

allele carriers in response to 2 g/day of PS for 28 days (Zhao et al., 2008).

Finally, another gene-PS interaction that influences blood cholesterol levels is a common SNP of the Niemann-Pick C1-like 1 (*NPC1L1*) gene. A recent study found that CC homozygous carriers for the *NPC1L1*-rs2072183 SNP showed an increase in TC and LDL-C levels after consuming PS ($+0.38$ and $+0.33$ mmol/L, respectively) (Granado-Lorencio et al., 2014). After consuming PS plus β -Cx, CC carriers displayed a decrease in TC and LDL-C concentrations (-0.24 and -0.27 , respectively). However, these differences did not reach statistical significance due to the small number of participants with this genotype ($n = 9$) (Granado-Lorencio et al., 2014). Another study reported that carriers of two SNPs, rs2072183-G and rs10264715-A (haplotype), showed 1.3 and 2.4-fold in TC and LDL-C concentrations, respectively, compared to carriers of the wild type after consuming 2 g/day of PS for 4 weeks (Zhao et al., 2008). These data showed that some individuals benefit more from PSs than others.

Additionally, consumption of 1.6–2.5 g/day of PS has been previously shown a 6% reduction in triglyceride (TG) concentrations (Naumann et al., 2008). Cholesteryl ester transfer (CETP) protein collects TGs from lipoproteins (VLDL or LDL) and exchanges them for cholesteryl esters from HDL particles. A randomized crossover study showed that carriers of the GG genotypes of *CETP* SNP rs5882 had a more significant decrease in TG levels compared to AA or AG carriers ($\Delta-0.46 \pm 0.13$ vs. $\Delta-0.03 \pm 0.08$ mmol/L) after consumption of 2 g/day for 28 days of PS (Mackay et al., 2015b).

In summary, accumulating evidence suggests that the intake of 2 g/day of PS decreases TC and LDL-C concentrations (Katan et al., 2003). Moreover, these responses appear to be associated with genetic variations in various genes including *ABCG8*, *APOE*, *CETP*, *CYP7A1*, and *NPC1L1*. Also, interactions between polymorphisms in these genes modify the responsiveness of blood cholesterol to PS. However, additional studies are needed to better understand the impact of genetics influencing blood cholesterol concentrations in response to PS to personalized nutrition interventions.

Acknowledgments

The present work has received research funding from the International Life Sciences Institute North America (Grant 320188).

Conflict of Interest

Tltzel Vazquez-Vidal declares that she has no conflict of interest. Peter J.H Jones has received research grants from Nutritional Fundamentals for Health Inc., Mitacs, and the International Life Science Institute. He also owns stock in Nutritional Fundamentals for Health Inc.

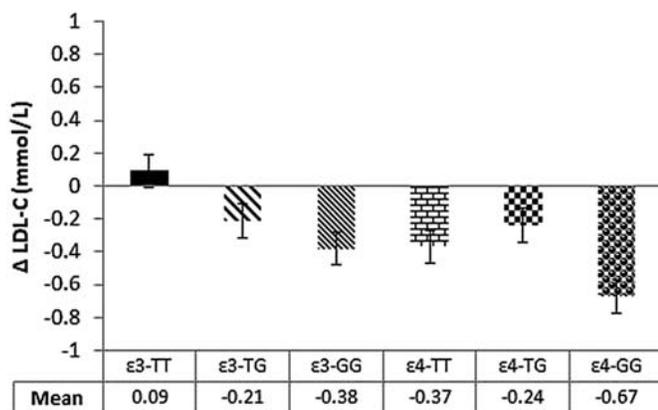


FIGURE 29.1 SNP-SNP interactions between *APOE* and *CYP7A1* genes on changes in LDL-C concentrations in response to consumption of 2 g/day of PS for 28 days. Adapted from Mackay, D.S. et al. 2015. *CYP7A1*-rs3808607 and *APOE* isoform associate with LDL cholesterol lowering after plant sterol consumption in a randomized clinical trial. *Am J Clin Nutr* 102(4), 951–957.

References

- Granado-Lorencio, F., et al., 2014. Beta-Cryptoxanthin modulates the response to phytosterols in post-menopausal women carrying NPC1L1 L272L and ABCG8 A632 V polymorphisms: an exploratory study. *Genes Nutr* 9 (5), 428–435.
- Gylling, H., et al., 2014. Plant sterols and plant stanols in the management of dyslipidaemia and prevention of cardiovascular disease. *Atherosclerosis* 232, 346–360.
- Jones, P.J., 2015. Inter-individual variability in response to plant sterol and stanol consumption. *J AOAC Int* 98 (3), 724–728.
- Katan, M.B., et al., 2003. Efficacy and safety of plant stanols and sterols in the management of blood cholesterol levels. *Mayo Clin Proc* 78 (8), 965–978.
- Kohler, J., et al., 2017. Plant sterol enriched functional food and atherosclerosis. *Br. J Pharmacol* 174, 1281–1289.
- MacKay, D.S., et al., 2015. CYP7A1-rs3808607 and APOE isoform associate with LDL cholesterol lowering after plant sterol consumption in a randomized clinical trial. *Am J Clin Nutr* 102 (4), 951–957.
- Mackay, D.S., et al., 2015. Cholesterol ester transfer protein polymorphism rs5882 is associated with triglyceride-lowering in response to plant sterol consumption. *Appl Physiol Nutr Metab* 40 (8), 846–849.
- Naumann, E., et al., 2008. The baseline serum lipoprotein profile is related to plant stanol induced changes in serum lipoprotein cholesterol and triacylglycerol concentrations. *J Am Coll Nutr* 27 (1), 117–126.
- Rosato, V., et al., 2017. Mediterranean diet and cardiovascular disease: a systematic review and meta-analysis of observational studies. *Eur J Nutr* 2–19.
- Sanchez-Muniz, F.J., et al., 2009. Serum lipid and antioxidant responses in hypercholesterolemic men and women receiving plant sterol esters vary by apolipoprotein E genotype. *J Nutr* 139 (1), 13–19.
- Zhao, H.L., et al., 2008. Genetic variation in ABC G5/G8 and NPC1L1 impact cholesterol response to plant sterols in hypercholesterolemic men. *Lipids* 43 (12), 1155–1164.

Genetic Individuality and Alcohol Consumption

Rosa M. Lamuela-Raventos¹, Ramón Estruch², Richard Kirwan³

¹Department of Nutrition, Food Science and Gastronomy, School of Pharmacy and Food Sciences, INSA-UB, University of Barcelona, CIBEROBN (Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y Nutrición), Barcelona, Spain; ²Faculty of Medicine and Health Sciences, Hospital Clinic, University of Barcelona, CIBEROBN (Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y Nutrición), Barcelona, Spain; ³School of Natural Sciences and Psychology, Faculty of Science, Liverpool John Moores University, United Kingdom; Department of Nutrition, Food Science and Gastronomy, School of Pharmacy and Food Science, INSA-UB, University of Barcelona, Barcelona, Spain

Glossary

ADH Alcohol dehydrogenase
ALDH Aldehyde dehydrogenase
ALI Alcoholic liver injury without cirrhosis
CcO Cytochrome *c* oxidase
MTHFR Methylene tetrahydrofolate reductase
PNPL Patatin-like phospholipase domain-containing protein

When discussing nutrigenetics and nutrigenomics, it is quite normal to immediately consider the interactions of macronutrients, such as carbohydrates, proteins and lipids, and micronutrients, such as vitamins, minerals, and other commonly consumed food components such as polyphenols, etc., with the genome of the individual. However, it may not be as obvious to consider the genetically variable implications of the consumption of a substance, such as alcohol, that is neither a macro- nor a micronutrient but nonetheless may also be consumed in significant amounts.

While the consumption of alcohol is by no means considered essential for a healthy diet, it nevertheless can, at times, form a meaningful portion of calorie intake, especially amongst certain individuals. In addition, considering that alcohol plays a significant role in the social and food cultures of many regions of the world (such as wine in the Mediterranean diet), it is clear that a better understanding of the effects of alcohol on differing genotypes is essential.

Furthermore, the abuse of alcohol can result in a spectrum of progressively serious conditions that are

considered under the umbrella term of alcoholic liver disease (ALD), and this ranges from basic steatosis or alcoholic fatty liver and alcoholic liver injury without cirrhosis (ALI) to alcoholic cirrhosis (AC) and hepatocellular carcinoma (HCC). Thus, the consumption of alcohol in itself can have significant effects on health.

When thinking of the possible health effects of alcohol consumption (specifically *excessive* alcohol consumption), it might be easy to first consider the aforementioned conditions, which have recognizable, physical diagnoses. However, it must be remembered that these conditions often develop from another alcohol-related illness, alcoholism or alcohol use disorder (AUD).

ADH1B AND ALDH2

To better understand how the genotype can affect the development of alcohol dependence, it is important to first understand how the body processes alcohol. The *in vivo* degradation of alcohol (and thus the detoxification from it) is coordinated by two main enzymes: alcohol dehydrogenase (ADH), which oxidizes alcohol to acetaldehyde; and aldehyde dehydrogenase (ALDH), which further oxidizes acetaldehyde to acetate. The intermediate product acetaldehyde, if allowed to accumulate, can cause physical responses such as tachycardia, flushing, hypotension, nausea, and vomiting.

Excessive accumulation of acetaldehyde and the subsequent heightened physical responses may dissuade

sensitive individuals from consuming alcohol due to the expectation of unpleasant side effects. Polymorphisms in the ADH and ALDH genes have been shown to result in the aforementioned accumulation of acetaldehyde.

Various ADH enzymes are responsible for alcohol metabolism in the liver, including ADH1A, ADH1B, ADH1C, etc. ADH1B exists in multiple allelic forms, with ADH1B*2 (rs1229984 A) coding for an enzyme that oxidizes alcohol at a faster rate than ADH1B*1. This increased rate of oxidation of alcohol to acetaldehyde can result in an accumulation of acetaldehyde, resulting in the previously mentioned side effects. Thus, individuals with one or two copies of the ADH1B*2 allele may be more sensitive to the detrimental effects of alcohol and may therefore be less likely to develop alcohol dependence.

Similarly, there are multiple forms of ALDH enzymes. However, in this case the ALDH2*2 allele (rs671 A) codes for an enzyme with low activity compared to the ALDH2*1 allele. Thus, acetaldehyde produced by the action of ADH cannot be converted sufficiently quickly to acetate, and there is therefore an accumulation of acetaldehyde, resulting in the unpleasant side effects of alcohol consumption and an aversion to further use of alcohol.

Indeed, the development of AUD has been shown to differ considerably depending on the genotype of an individual. Luczak et al. (2006) showed this clearly through a meta-analysis of the effects of polymorphisms in both the ADH1B and ALDH2 genes and rates of alcohol dependence in East-Asian populations. They found that, compared to individuals with only ADH1B*1 alleles, those with one or two copies of the ADH1B*2 polymorphism were one-fourth and one-fifth as likely to be

alcohol dependent, respectively. Similarly, they found that, compared to ALDH2*1 homozygotes, those with one ALDH2*2 allele had one-fifth the risk, and those with two ALDH2*2 alleles had one-ninth the risk, of being alcohol dependent. This protective effect was even greater in individuals possessing both of the protective polymorphisms (Figs. 30.1 and 30.2).

It should also be noted that the ADH1B*2 and ALDH2*2 alleles can actually increase an individual's susceptibility to alcohol-related conditions, such as liver disease, pancreatitis, and head and neck cancers, in case that individual were to become alcohol dependent and consume alcohol excessively. This is believed to be due to the role of acetaldehyde in the development of such diseases.

PRE-MIRNA-27A

While the mechanism by which certain polymorphisms can result in increased alcohol intake may be relatively well understood, as in the case of ADH1B*2 and ALDH2*2, as previously mentioned, in other cases, the mechanism remains largely unknown. This is the case of a polymorphism in pre-miRNA-27a, which has also been linked to alcohol consumption.

While ADH1B*2 and ALDH2*2 both code for enzymes involved directly in the degradation of alcohol, pre-miRNA-27a is a micro RNA—defined as a small non-coding RNA of approximately 22 nucleotides—that regulates gene expression. In general, miRNAs are abundant in the brain and have been found to play major roles in many biological processes. In fact, their importance can be inferred by the fact that SNPs are relatively rare in miRNAs, indicating their biological importance.

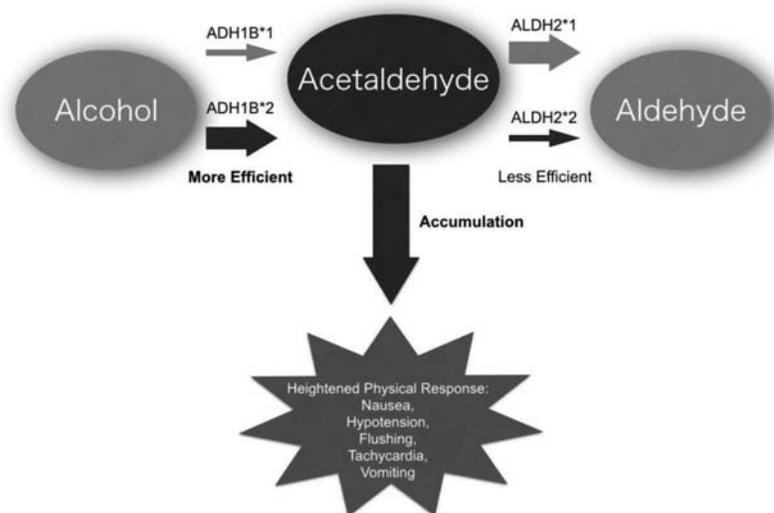


FIGURE 30.1 Alcohol detoxification pathway involving polymorphic alcohol dehydrogenase and aldehyde dehydrogenase. ADH, Alcohol dehydrogenase; ALDH, aldehyde dehydrogenase.

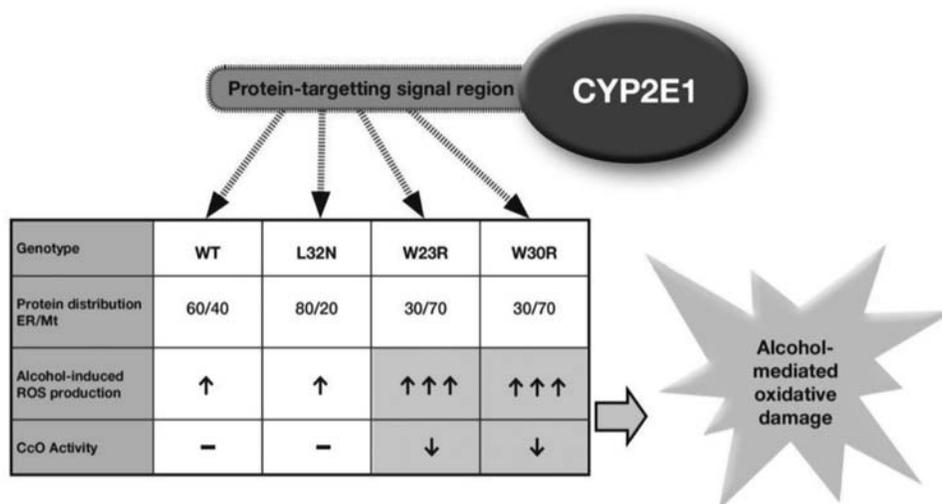


FIGURE 30.2 Protein distribution, reactive oxygen species production and cytochrome *c* oxidase activity of wild type and mutant alleles of the protein-targeting signal region of CYP2E1. *CcO*, cytochrome *c* oxidase; *ER*, Endoplasmic reticulum; *Mt*, Mitochondrial; *ROS*, Reactive oxygen species; *WT*, Wild type.

In a study of an elderly Spanish population, Barragán et al. (2016) investigated the relationship between the pre-miRNA-27a rs895819 A > G polymorphism and alcohol intake. These authors observed that alcohol consumption amongst individuals with the GG variant genotype was considerably higher than amongst AA and AG genotypes, indicating the recessive nature of the GG phenotype (Table 30.1).

Importantly, they also noted that both men and women with the GG genotype were about three times more likely to have a “high” intake of alcohol (considered more than one drink, 13.2 g of alcohol, for women or two drinks for men) compared with the AA genotypes (Table 30.1)

Furthermore, in sensitivity analyses it was found that the increased prevalence of heavier drinkers amongst GG individuals was virtually the same in both obese and nonobese subjects (odds ratio 3.87 and 3.31, respectively).

How does the GG polymorphism bring about this significantly increased alcohol intake? It appears that the G allele variant is not only fully functional but also results in increased levels of miRNA-27a. As miRNAs are common in the brain, it is possible that increased levels of this particular miRNA affect the neurological chemistry involved in hedonistic consumption.

It has been speculated that one possible mechanism of action of the pre-miRNA-27a rs895819 A > G polymorphism is through its interaction with the serpin peptidase inhibitor clade A member I (SerpinA1), a known protein target of miRNA27a. This protein is linked to analgesic tolerance (specifically to morphine), and previous studies have shown that SerpinA1 knockout mice developed less analgesic tolerance than wild types.

TABLE 30.1 Genotypic prevalence of pre-miRNA-27a rs895819 polymorphisms amongst nondrinkers, moderate drinkers, and heavy drinkers and total daily alcohol consumption according to polymorphism.

	Genotype prevalence (%)		
	AA	AG	GG
Nondrinkers (0 g/day)	45.20%	42.00%	39.50%
Moderate drinkers (men <26.4 g/day or women <13.2 g/day)	48.90%	50.40%	43.00%
Heavy drinkers (men >26.4 g/day or women >13.2 g/day)	5.90%	7.60%	17.40%
Total alcohol consumption (g/day)	5.2	5.9	9.1

Adapted from Barragán, R., Coltell, O., Asensio, E.M., Francés, F., Sorlí, J. V., Estruch, R., Salas-Huetos, A., Ordovas, J.M., Corella, D., 2016. MicroRNAs and drinking: association between the pre-miR-27a rs895819 polymorphism and alcohol consumption in a mediterranean population. *Int J Mol Sci* 17, 1–18. <https://doi.org/10.3390/ijms17081338>.

This could potentially reveal a relation between miRNA27a, SerpinA1, and chronic opioid consumption, as well as, potentially, other forms of substance abuse, such as alcohol in this case.

CYP2E1

Mutations in genes can affect the function of their end products through various mechanisms. In addition to knockout mutations, which completely inhibit the production of functioning end product, or mutations that

reduce the activity of end products and are well known to have significant biological effects, this too may occur with mutations in the region of a gene responsible for its localization within a cell or organism. This is the case of a particular set of mutations in the protein-targeting signal region of the CYP2E1 gene, which codes for cytochrome P450 2E1, an enzyme that is involved in the metabolism of alcohol in the liver and intestines. Along with ADH, CYP2E1 converts alcohol to toxic acetaldehyde and then to acetate, alongside the enzyme aldehyde dehydrogenase (ALDH). CYP2E1 is also involved in the detoxification of other low-molecular-weight endogenous and exogenous compounds. This detoxifying action of CYP2E1 is thought to contribute to the production of reactive oxygen species (ROS), which may result in tissue damage through the peroxidation of membrane lipids.

In liver cells, CYP2E1 is found both bound to the endoplasmic reticulum (microsomal location) and within mitochondria. In the presence of ethanol, mitochondrial CYP2E1 was found not only to generate more ROS compared to microsomal CYP2E1 but also to result in direct damage of the mitochondrial cytochrome *c* oxidase (CcO), required to complete the electron transport chain involved in ATP synthesis, thus acting as a further cellular stress.

Bansal et al. (2013) showed that mutations in the protein-targeting signaling region of CYP2E1 affect the targeting efficiency to the endoplasmic reticulum (ER) or of mitochondria. Along with the wild-type (WT) enzyme, they identified three mutations: W23R, W30R, and L32N. WT-transfected cells showed a 60/40% distribution in microsomes and mitochondria, respectively; whereas W23/30R-transfected cells showed a 30/70% distribution, indicating increased mitochondrial localization. Interestingly, L32N-transfected cells showed an 80/20% distribution, showing increased ER localization compared to the WT enzyme. This differential localization was confirmed through analysis of human liver samples, with increased mitochondrial localization in cells containing W23/30R mutations and increased microsomal localization in cells containing the L32N mutation. Here ethanol treatment of mitochondrial fractions resulted in slightly increased production of ROS from WT- and L32N-transfected cells, but mitochondria from W23/30R-transfected cells showed a nearly two-fold higher ROS production. Microsomal fractions, in contrast, showed very low ROS levels, indicating that the mitochondrial compartment of cells is a major contributor to ROS production.

Additionally, cells with W23/30R mutations were shown to have significantly lower (−25%) cytochrome C oxidase activity, which was further reduced (−50%) in the presence of ethanol.

The study authors hypothesized that carriers of the W23/30R mutations, which preferentially localize CYP2E1 to mitochondria, may experience a greater oxidative stress when exposed to alcohol, and may thus be more prone to alcohol-induced liver disease.

PNPLA3

As previously mentioned related to the pre-miRNA-27a polymorphism, while the mechanism of action of certain mutations in their relation to disease is well understood, in other types of mutation this may not be true, despite the existence of a clear association between the mutation and the development of specific disease conditions. This is the case for the patatin-like phospholipase domain-containing protein (PNPL) A3 mutation as it relates to alcoholic liver disease.

Despite the fact that alcohol consumption is considered to be the most important factor in the development of the alcoholic liver disease conditions—steatosis/alcoholic fatty liver, ALI, alcoholic cirrhosis, and hepatocellular carcinoma—less than 20% of heavy drinkers see their alcoholic liver disease progress to alcoholic cirrhosis. This would suggest that other, possibly genetic, host factors are important.

Polymorphisms in the PNPLA3 gene have been associated with hepatic fat content in nonalcoholic fatty liver disease. The review by Salameh et al. (2015), in an excellent overview and meta-analysis investigating the link between the presence of the rs738409 C > G point mutation in the PNPLA3 gene and various conditions of alcoholic liver disease. These authors found that the incidence of ALI, alcoholic cirrhosis, and hepatocellular carcinoma increased with the CG genotype compared to the CC genotype, and the incidence increased even further in the GG genotype (Table 30.2). Interestingly,

TABLE 30.2 Incidence of different forms of alcoholic liver disease in CG and GG mutants of the patatin-like phospholipase domain-containing protein A3 gene, compared to wild-type (CC) individuals.

Condition	Odds ratio (compared with WT)	
	CG vs. CC	GG vs. CC
ALI	1.45	2.22
AC	2.09	3.37
HCC	2.87	12.41

AC, Alcoholic cirrhosis; ALI, Alcoholic liver injury; HCC, Hepatocellular carcinoma; WT, Wild type

Adapted from Salameh, H., Raff, E., Erwin, A., Seth, D., Nischalke, H.D., Falletti, E., Burza, M.A., Leathert, J., Romeo, S., Molinaro, A., Corradini, S.G., Toniutto, P., Ulrich, S., Daly, A., Day, C.P., Kuo, Y.-F., Singal, A.K., 2015. PNPLA3 gene polymorphism is associated with predisposition to and severity of alcoholic liver disease. *Am J Gastroenterol* 110, 846–856. <https://doi.org/10.1038/ajg.2015.137>

there was no increase in the incidence of alcoholic fatty liver, a condition that can be resolved with abstinence from alcohol. ALI, alcoholic cirrhosis, and hepatocellular carcinoma, however, once developed, are not reversible.

Considering the clear relationship between the presence of this polymorphism and the risk of developing severe forms of ALD, screening for such polymorphism amongst patients presenting with liver steatosis might prove useful in clinical or lifestyle interventions aimed at reducing the incidence of these serious conditions.

MTHFR C677T

High levels of circulating homocysteine (Hcy) are associated with an increased risk of both cardiovascular disease (CVD) and certain cancers. Alcohol consumption has been associated with an increase in Hcy, and this may be one of the mechanisms by which excess alcohol consumption may lead to increased CVD and cancer morbidity.

The increase in Hcy observed in cases of alcohol abuse is partially due to an accumulation of biologically active folate (5-methyltetrahydrofolate) in liver cells, along with a reduction in methionine synthase activity, which results in reduced Hcy remethylation. Additionally, methylenetetrahydrofolate reductase (MTHFR) is an enzyme involved in the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, which is required for the remethylation of Hcy to methionine. Reduced activity of MTHFR is thus another mechanism by which Hcy levels can increase.

Nienaber-Rousseau et al. (2013) investigated, in a South African population, the link between a C > T point mutation in base pair 677 (C677T) of the MTHFR gene, which results in reduced enzyme activity, and effects on Hcy concentrations. These authors observed that Hcy increased with both age and alcohol intake, measured by gamma-glutamyltransferase concentrations.

Additionally, the presence of the MTHFR 677 CT genotype was associated with increased Hcy levels, which were even higher in subjects with the TT genotype (although the TT genotype was observed in only 0.8% of the study population). The study concluded that alcohol consumption likely has a significant interaction with the MTHFR genotype on Hcy concentrations. A combination of high alcohol consumption and an unfavorable MTHFR genotype would thus result in increased risk of developing CVD and certain cancers. Supplementation with group B vitamins would potentially not be as effective in MTHFR 677 CT or TT carriers, and an intervention aimed at reducing alcohol consumption may prove more effective.

References

- Bansal, S., Anandatheerthavarada, H.K., Prabu, G.K., Milne, G.L., Martin, M.V., Guengerich, F.P., Avadhani, N.G., 2013. Human cytochrome P450 2E1 mutations that alter mitochondrial targeting efficiency and susceptibility to ethanol-induced toxicity in cellular models. *J Biol Chem* 288, 12627–12644. <https://doi.org/10.1074/jbc.M113.452367>.
- Barragán, R., Coltell, O., Asensio, E.M., Francés, F., Sorlí, J.V., Estruch, R., Salas-Huetos, A., Ordovas, J.M., Corella, D., 2016. MicroRNAs and drinking: association between the pre-miR-27a rs895819 polymorphism and alcohol consumption in a mediterranean population. *Int J Mol Sci* 17, 1–18. <https://doi.org/10.3390/ijms17081338>.
- Luczak, S.E., Glatt, S.J., Wall, T.L., 2006. Meta-analyses of ALDH2 and ADH1B with alcohol dependence in Asians. *Psychol Bull* 132, 607–621. <https://doi.org/10.1037/0033-2909.132.4.607>.
- Nienaber-Rousseau, C., Pisa, P.T., Venter, C.S., Ellis, S.M., Kruger, A., Moss, S.J., Melse-Boonstra, A., Towers, G.W., 2013. Nutritional genetics: the case of alcohol and the MTHFR C677T polymorphism in relation to homocysteine in a black South African population. *J Nutrigenetics Nutrigenomics* 6, 61–72. <https://doi.org/10.1159/000348839>.
- Salameh, H., Raff, E., Erwin, A., Seth, D., Nischalke, H.D., Falleti, E., Burza, M.A., Leathert, J., Romeo, S., Molinaro, A., Corradini, S.G., Toniutto, P., Ulrich, S., Daly, A., Day, C.P., Kuo, Y.-F., Singal, A.K., 2015. PNPLA3 gene polymorphism is associated with predisposition to and severity of alcoholic liver disease. *Am J Gastroenterol* 110, 846–856. <https://doi.org/10.1038/ajg.2015.137>.

This page intentionally left blank

Vitamin A and Other Carotenoids

Johannes von Lintig, Darwin Babino

Case Western Reserve University School of Medicine, Department of Pharmacology, Cleveland, OH, United States

Glossary

Aldehyde Dehydrogenase Family Enzyme members that catalyze the synthesis of retinoic acid from retinal.

ATP-Binding Cassette Transporter 4 Membrane-associated protein within the retina that transports retinal from the disc lumen into the cytosol.

β -Carotene-15,15'-Dioxygenase Catalyzes the oxidative cleavage of provitamin A carotenoids into retinal.

β -Carotene-9,10'-Dioxygenase Catalyzes the oxidative cleavage of carotenoids into apocarotenoids.

Carotenoids Isoprenoid pigments produced by plants, bacteria, and fungi that play critical physiological roles in living organisms.

Carotenoid Cleavage Oxygenases Family of enzymes that catalyze the oxidative cleavage and/or isomerization of carotenoids and/or retinoids.

Cellular Retinal-Binding Protein A soluble binding protein that selectively binds 11-*cis*-retinal.

Cellular Retinoic Acid-Binding Protein Binding protein that binds/carries retinoic acid within the nucleus.

Cellular Retinol-Binding Protein Carrier protein involved in the transport of retinol (vitamin A alcohol).

Cluster of Differentiation 36 A receptor protein for oxidized LDL involved in the cellular uptake of carotenoids.

Cytochrome P450, Family 26 Members of the cytochrome P450 proteins involved in retinoic acid metabolism.

Interphotoreceptor Retinoid-Binding Protein Retinoid carrier protein that facilitates the transport of retinoids between the retinal pigment epithelium and photoreceptor outer segments.

Intestine-Specific Homeobox Transcription factor that regulates the expression of the β -carotene-15, 15'-dioxygenase (BCO1) gene.

Retinal Aldehyde form of vitamin A that when bound to opsins acts as the visual chromophore in animal visual systems.

Retinal Pigment Epithelium-Specific Protein 65 kDa CCO member enzyme responsible for the conversion of all-*trans*-retinyl esters into 11-*cis*-retinol.

Retinoic Acid Acid form of vitamin A that acts as a hormone by binding to retinoic acid receptors and mediating functions of vitamin A required for growth and development.

Retinol Alcohol form of vitamin A that can be converted to other forms of bioactive vitamin A and esterified for storage.

Retinol Binding Protein 4 Delivers retinol from liver stores to the peripheral tissues.

Scavenger Receptor Class B, Member 1 Receptor for high-density lipoprotein involved in cellular carotenoid uptake.

Short-chain dehydrogenases/reductases Enzymes that catalyze the oxidation/reduction of a wide range of substrates, including retinoids and steroids.

Stimulated by Retinoic Acid 6 Multidomain membrane protein that acts as a retinoid pore and facilitates the bidirectional flux of vitamin A between retinol binding protein 4 and the cytoplasm membrane.

Transthyretin Transport protein in the serum and cerebrospinal fluid that forms a complex with serum retinol binding protein.

Carotenoids are isoprenoid pigments produced by plants, bacteria, and fungi. The synthesis of these terpenoids (C40) involves the condensation of eight isoprenoid units (C5), desaturation of double bonds, and introduction of terminal hexyl rings. The enormous diversity of carotenoids (with more than 600 chemically distinct compounds) is achieved by modulation in length of the polyene chromophore, shifts of the conjugated double bonds, and the addition of functional groups.

Some 50 carotenoids exist in the human diet, from which only about 10 are present in significant amounts in human plasma. Human tissues retain considerable amounts of these lipids. Physiological functions attributed to them are associated with their chemical and physical properties as free radical scavengers in lipophilic environments such as membranes and lipoproteins. These pigments serve as blue light filters in the central part of the human retina, the macula lutea, which owes its yellow color to high levels of lutein, zeaxanthin, and *meso*-zeaxanthin.

Carotenoids can be metabolically transformed to apocarotenoids, such as retinoids, which encompass all natural derivatives of vitamin A (all-*trans*-retinol). Vitamin A is the precursor for at least two essential biologically active molecules, 11-*cis*-retinal and all-*trans*-retinoic acid (RA). 11-*cis*-retinal binds to the protein moiety (opsin) of visual pigments that mediate phototransduction in the eyes. The second critical biologically active vitamin A metabolite, RA, is required for a wide range

of biological processes, including reproduction, embryonic development, cell differentiation, immunity, and metabolic control. This hormone-like substance is the ligand of nuclear receptors such as retinoic acid receptors (RARs). RARs form heterodimeric complexes with retinoid X receptors. These ligand-activated transcription factors control the activity of numerous genes in the human genome.

THE METABOLISM OF CAROTENOIDS AND RETINOIDS

The metabolism of carotenoids and their retinoid metabolites can be portrayed in the following framework: To become biologically active, these lipids must be acquired from the diet in the intestine, transported within the body to target tissues, and—in the case of retinoid function—be metabolically converted. Ultimately, these lipids must be eliminated to maintain the physiological state. Research into the molecular biology and biochemistry of this metabolism led to the discovery of novel enzyme classes and signaling proteins, including G protein-coupled receptors and nuclear hormone receptors. Mutations in the corresponding genes have been identified as responsible for pathologic conditions ranging from inherited forms of night blindness to complex microphthalmic syndromes. The importance of this metabolism is emphasized by the finding that genetic polymorphisms in involved genes are associated with diverse chronic disease states. In the following we will describe this metabolism with an emphasis on mutations related to inherited disease. Our descriptions refer to [Table 31.1](#) and [Fig. 31.1](#), which give a systematic summary accompanying this text and depicting major aspects of the complex network of binding proteins, transporters, and metabolic enzymes for this unique class of dietary lipids.

CAROTENOID CONVERSION TO RETINOIDS

The first step in retinoid metabolism is the conversion of the parent C40 carotenoid precursor into a C20 retinaldehyde by symmetric oxidative cleavage at position C15, C15' in the carbon backbone of provitamin A carotenoids such as β -carotene. The reaction is catalyzed by β -carotene-15, 15'-dioxygenase (BCO1). A second structurally related carotenoid cleavage enzyme, the β -carotene-9, 10-dioxygenase (BCO2), is encoded in the human genome and cleaves carotenoids at position C9, C10, and C9', C10'. Both enzymes are members of an ancestral enzyme family of nonheme

iron oxygenases. These proteins adopt a common structural fold (a seven-bladed propeller covered by a half dome) with a central ferrous iron in the active center. The iron is coordinated by four strictly conserved histidine and 2 s shell glutamate residues. The iron is accessible through a nonpolar tunnel, which extracts the lipophilic substrate and exposes the scissile double bond to the active center. The reaction follows a dioxygenase reaction mechanism in which both atmospheric oxygen atoms are incorporated into the apocarotenoid products.

BCO1 is a cytosolic enzyme with limited substrate specificity for provitamin A carotenoids and apocarotenoids (>C20), with at least one unsubstituted β -ionone ring. In contrast, BCO2 displays broad substrate specificity and cleaves most, if not all, carotenoids and apocarotenoids present in blood and tissues. BCO2 resides in the inner membrane of mitochondria. This subcellular compartmentalization of vitamin A production and carotenoid catabolism prevents competition between BCO1 and BCO2 for common provitamin A substrates such as β -carotene. Recent experimental evidence also suggests that a concerted action of BCO2 and BCO1 provides a mechanism to tailor asymmetric carotenoids such as β -cryptoxanthin for vitamin A production. In this pathway, BCO2 removes the noncanonical 3-hydroxy-ionone ring site by cleavage across the C9, C10 double bond. The resulting apocarotenoid is then further processed to retinaldehyde by BCO1. Ongoing research also has implicated BCO2-derived apocarotenoid cleavage products, different than retinoids, as biologically active modulators of physiological process.

Besides their roles in apocarotenoid production, BCO1 and BCO2 play critical roles in carotenoid homeostasis as evidenced by naturally occurring mutations and genetic polymorphisms in the corresponding genes that alter carotenoid levels in blood and tissues of humans and animals. In humans, mutation in the BCO1 gene can cause hypercarotenemia and hypovitaminosis A. Mutations in BCO2 have yet to be reported in humans. However, genetic analyses of animals indicate that BCO2 plays an important role in the homeostatic control of tissue level of nonprovitamin A carotenoids, including the macular pigments of the primate eyes. Intriguingly, BCO2 gene expression is responsive to oxidative stress. This regulation manages the chemistry and biology of carotenoids that act as anti- and prooxidants depending on the subcellular localization and concentration. This finding may provide an explanation for the low carotenoid status of patients affected by chronic disease because they generally are associated with inflammation and oxidative stress.

TABLE 31.1 Genes that Affect Carotenoid and Vitamin A Metabolism, Transport and Distribution and the Resulting Associated Diseases due to their Polymorphisms and Mutations.

Gene	Function	Inherited Diseases
ABCA4	Transports RAL from disc lumen into the cytosol	Stargardt macular dystrophy
CRALBP	Carries 11- <i>cis</i> -RAL as physiological ligand	Fundus albipunctatus, rod-cone dystrophy
CRABP	Carries retinoic acid within the nucleus	Teracarcinoma, embryonal cancer
CRBP1	Transport of retinol to peripheral tissues	Epithelioid cell melanoma, blepharophimosis
CD36	Receptor for oxidized LDL, intestinal uptake of carotenoids and fat soluble vitamins	Fatty acid metabolism, atherosclerosis, Alzheimer disease
DHRS3	Oxidation of at-ROL to at-RAL	Neuroblastoma
IRBP	Retinoid carrier between RPE and POS	Retinitis pigmentosa
ISX	Transcription factor that regulates BCO1 expression	-
LRAT	Catalyzes the esterification of at-ROL into at-REs	Leber congenital amaurosis, retinitis pigmentosa
RPE65	Converts at-REs into 11- <i>cis</i> -ROL	Leber congenital amaurosis, retinitis pigmentosa
RBP4	Carries retinol in the blood	Retinal dystrophy, iris coloboma, microphthalmia
RDH5	Oxidation of 11- <i>cis</i> -ROL to 11- <i>cis</i> -RAL	Fundus albipunctatus, rod-cone dystrophy
RDH8	Reduction of at-RAL to at-ROL	Macular degeneration
RDH10	Oxidation of at-ROL to at-RAL	Congenital malformations, Scimitar syndrome
RDH11	Oxidation of 11- <i>cis</i> -ROL to 11- <i>cis</i> -RAL	Retinal dystrophy, short stature syndrome
RDH12	Reduction of at-RAL to at-ROL	Leber congenital amaurosis, retinitis pigmentosa
SCARB1	Receptor for high-density lipoprotein, carotenoid uptake	-
STRA6	Receptor for retinol-RBP4 complex	Matthew-Wood syndrome
TTR	Carries holo-RBP	Amyloidotic polyneuropathy, cardiomyopathy
BCO1	Vitamin A production	Hypovitaminosis A
BCO2	Carotenoid catabolism	-
CYP26A1	Hydroxylates RA	Keratomalacia, caudal regression syndrome
CYP26B1	Hydroxylates RA	Radiohumeral fusions, skeletal and craniofacial anomalies
CYP26C1	Hydroxylates RA	Focal facial dermal dysplasia
ALDH1A2	Catalyzes the synthesis of RA from RAL	Embryonic lethal in mice
ALDH1A1	Alcohol metabolism	-
ALDH1A3	Catalyzes the synthesis of RA from RAL	Microphthalmia, respiratory distress

CAROTENOID AND RETINOID ABSORPTION, BODY TRANSPORT, AND STORAGE

The hydrophobic nature of carotenoids and retinoids limits their solubility and diffusion in the aqueous environment of the body. Therefore, animals including humans, have evolved membrane transporters and specific binding proteins for the bodily distribution of these compounds. Three classes of retinoid-binding proteins named for their selectivity (i.e., cellular retinol-binding

protein (RBP), cellular retinoic acid-binding protein (CRABP), and *cis*-retinoid-specific cellular retinal-binding proteins (CRALBP)) carry retinoids within cells. Vertebrate CRBPs and CRABPs are members of the lipocalin family of lipid-binding proteins, whereas CRALBP belongs to the Sec14 protein family. Lipocalins possess an eight-stranded, antiparallel, symmetrical β -barrel fold. The ligand binding site is located inside the barrel and determines the specificity for different retinoids.

Retinoid-binding proteins facilitate the retinoids' inter- and intracellular transport and interactions with

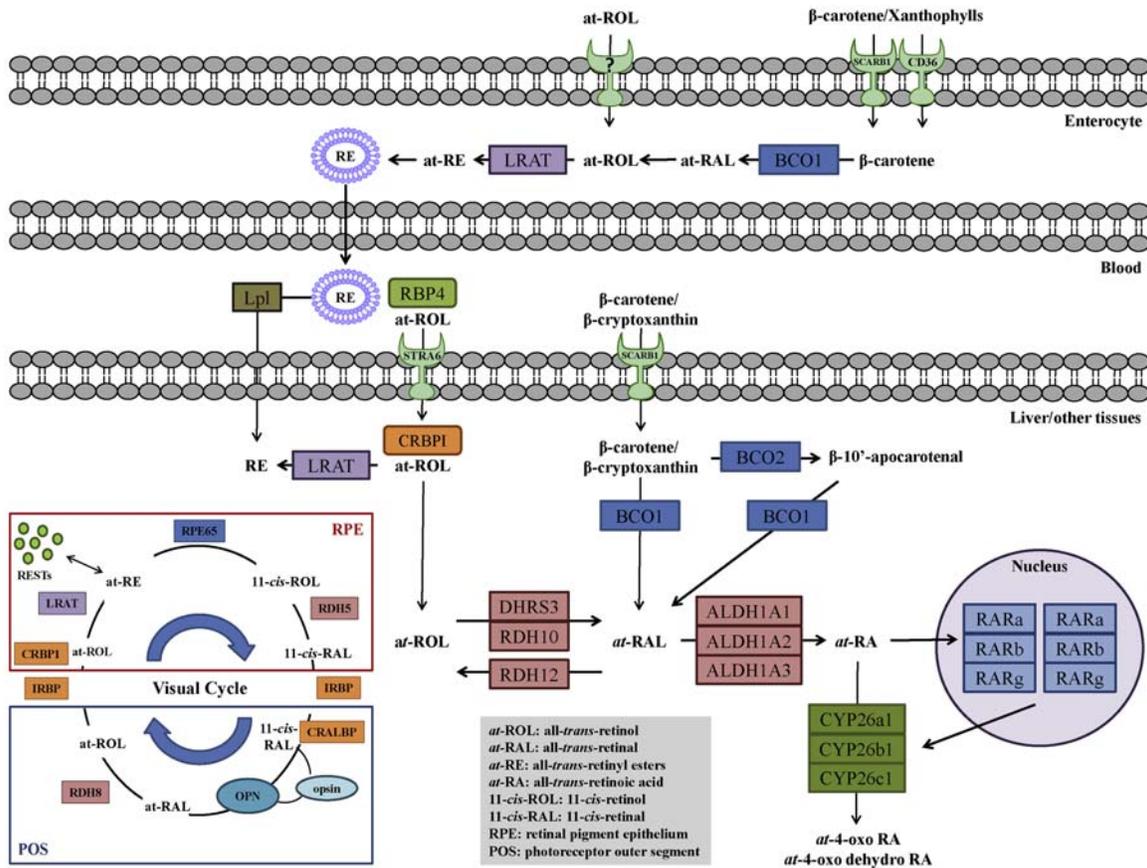


FIGURE 31.1 General scheme of carotenoid and vitamin A metabolism, transport, and distribution.

metabolic enzymes. These proteins also protect retinoids from chemical modification, and vice versa protect their surrounding from chemical reactivity of the bound lipids. Mutations in these genes can impair embryonic development and cause blinding diseases in adults, including retinitis pigmentosa and bothnia dystrophy.

Specific carotenoid-binding proteins of the steroidogenic acute regulatory and the glutathione S transferase protein families have been biochemically described in the primate eyes. These proteins bind carotenoids in a one-to-one molar ratio, and may contribute to the concentration of zeaxanthin, lutein, and *meso*-zeaxanthin in human eyes. The ultimate description of their function requires analyses of the visual consequences caused by mutations in their corresponding genes.

During intestinal absorption and distribution to target cells, carotenoids and retinoids must cross over biological membrane barriers. Though these compounds can participate between their binding proteins and membranes, recent evidence indicates the involvement of specific membrane transporters. The scavenger receptors SCARB1 and CD36 have been implicated in intestinal carotenoid uptake. This is highlighted by a unique transcriptional regulation of *SCARB1* gene expression in enterocytes of the gut by the homeodomain

transcription factor ISX. ISX also controls the abundance of BCO1, which converts absorbed β -carotene to retinaldehyde (see above). In a state of sufficient dietary supplies, ISX suppresses the absorption of carotenoids and other fat-soluble vitamins, and prevents excessive accumulation of these dietary lipids. This negative feedback regulation also adapts vitamin A production to the fluctuating supply of dietary β -carotene, as demonstrated in mouse models. The validity of the concept for regulation of carotenoid absorption is supported by the observation that genetic polymorphisms in involved genes are associated with alterations in blood and tissue levels of carotenoids in humans. Particularly, genetic polymorphisms in 5'-upstream regions of the *BCO1* gene have been associated with this genotype and have been confirmed in repeated genome-wide association studies. The influence of the vitamin A status and genotype on intestinal carotenoid metabolism may also explain the high- and low-converter phenotypes in response to β -carotene supplementation in clinical studies.

In enterocytes, retinoids derived from dietary carotenoids and/or preformed vitamin A are converted to retinyl esters and packaged along with carotenoids and other fat-soluble vitamins into triacylglycerol-rich chylomicrons. Peripheral tissues render retinyl esters

available in a lipoprotein lipase-dependent manner. The large remaining quantity in chylomicron remnants is deposited in the liver for storage in specialized stellate (Ito) cells. Additionally, peripheral tissues such as lung, kidney, and fat store significant amounts of vitamin A in the form of retinyl esters. These retinyl esters are formed by esterification of retinol with medium-chain fatty acids, mainly palmitate. These esters readily aggregate and are sequestered in lipid droplets in specialized cells. Early studies demonstrated two independent enzymatic activities for the formation of retinyl esters, namely a lecithin-dependent acyl transfer facilitated by lecithin retinol acyltransferase (LRAT), which is expressed in most tissues except adipocytes, and an acyl-CoA-dependent activity in a variety of mammalian tissues such as small intestine, liver, adipocytes, skin, testis, and retina. LRAT has been molecularly cloned from several vertebrate species, and the generation of LRAT-deficient mice confirmed its pivotal role in retinoid metabolism and vitamin A homeostasis. LRAT is a 25 kDa protein that localizes to the endoplasmic reticulum and is an integral membrane protein with a single membrane-spanning helix localized at the C-terminus. On the basis of its amino acid sequence and predicted tertiary structure, LRAT is classified as a member of the ancestral NlpC/P60 thiol peptidase protein superfamily. Besides LRAT, the human genome encodes seven genes belonging to this protein family. Their basic structural motif is reminiscent of papain-like proteases, and consists of a four-strand antiparallel β -sheet and three α -helices. The conserved catalytic residues Cys161, His60, and His72 define the active site. LRAT adopts an analogous catalytic strategy as thiol peptidases, whereby the deprotonated Cys161 serves as a nucleophile to attack the carbonyl carbon of an ester bond at the SN1 position of phosphatidylcholine, eventually leading to acyl transfer and transesterification of retinol.

In contrast to LRAT, which utilizes lecithin as an acyl-donor, acyl-CoA:retinol acyltransferase (ARAT) takes advantage of a preactivated acyl-moiety coupled to coenzyme A. ARAT has never been purified or cloned, however, the existence of such vitamin A ester forming enzymes is supported by several lines of evidence. Studies of LRAT-deficient mice indicate that the intestinal absorption of vitamin A decreased to 10% of wild-type animals when challenged with a physiologic dose of dietary retinol. Retinyl ester stores are utilized in times of limited dietary vitamin A supply. Mobilization requires splitting of the ester bond by hydrolases. No particular hydrolase has been assigned to retinyl ester metabolism. Various carboxyl-esterases, hepatic, lipoprotein, pancreatic triglyceride, and hormone-sensitive lipases have been shown to catalyze retinyl ester hydrolysis in the context of different physiological processes in different cell types and tissues.

In the fasting circulation, vitamin A from body stores bound to RBP4 (holo-RBP4) is the major transport mode of retinoids, and its levels are homeostatic in healthy well-nourished individuals, but differ between gender. Once in the blood, holo-RBP4 escapes kidney glomeruli filtrations by forming a complex with the 55 kDa transthyretin (TTR) homotetramer at a 2:1 M ratio.

Vitamin A from holo-RBP4 is acquired by cells via the *stimulated by retinoic acid 6* (STRA6) gene product. This multidomain membrane protein acts as a retinoid pore, and facilitates the bidirectional flux of vitamin A between RBP4 and the cytoplasm membrane. Cellular concentration of retinoids is driven by conversion of retinol into retinyl esters via LRAT. The importance of a specific vitamin A transport system is documented by a rare, inherited condition known as the Matthew Wood syndrome. This autosomal recessive disease is caused by mutations in the *STRA6* gene and characterized by severe bilateral microphthalmia, often in combination with pulmonary dysplasia, cardiac defects, and diaphragmatic hernia, among other anomalies and malformations. The phenotypic manifestation is highly variable even within the same family, and ranges in severity from isolated microphthalmia to complex microphthalmia syndromes. Similarly, mutations in the *RBP4* gene can result in congenital eye malformations, including microphthalmia. Again, large variations in the phenotypic manifestation of RBP4 mutations are reported, even within the same family, indicating that stochastic factors, such as maternal vitamin A supply, modulate the severity of the diseases.

In contrast to retinoids, carotenoids exist in all classes of lipoproteins of the fasting circulation, with β -carotene being abundant in low-density lipoprotein (LDL) particles and zeaxanthin and lutein in high-density lipoprotein (HDL) particles. The cellular uptake of carotenoids from lipoprotein particles may involve LDL receptor-mediated endocytosis, as well as exchanges of carotenoids between lipoproteins and cells mediated by scavenger receptors such as SCARB1 and CD36.

METABOLISM OF RETINOIC ACID

The alcohol and ester form of vitamin A are the predominant retinoids in most tissues. In contrast, the acidic form of the vitamin, retinoic acid (RA), exists only in very low nanomolar concentrations in most tissues and blood. The tissue levels of RA are controlled in a temporal and spatial manner, and the synthesis of small amounts of RA is sufficient to exert profound physiological effects through the activation of RAR complexes. RARs are transcription factors of the nuclear hormone receptor gene family, which, in conjunction with retinoid X receptors, control transcription by binding

to conserved DNA motifs (RA response elements) in promoter regions of about 500 target genes in the human genome. Excessive amounts of RA in tissues are toxic and can cause fatal conditions throughout the life cycle. Therefore, the levels of this hormone-like compound are under strict metabolic control.

For the synthesis of RA, vitamin A is stepwise oxidized. The interconversion of retinol to retinal is catalyzed by cytosolic alcohol dehydrogenases (ADHs) and microsomal retinol dehydrogenases (RDHs). The latter belong to the short-chain dehydrogenase/reductase protein family (SDR). The redox carriers of these reactions are adenine dinucleotide cofactors NAD(H) and NADP(H). The enzymes bind their cofactors by a conserved sequence motif, the Rossmann-fold, which consists of six to seven parallel β -strands flanked by three to four α -helices. ADHs and RDHs use different catalytic mechanisms with a zinc atom and a tyrosine in the active center, respectively. In tests for enzymatic assays, ADHs and RDHs catalyze the bidirectional interconversion of retinol to retinal, dependent on the oxidative state of their redox carriers. Under physiological conditions, the ratio of NAD/NADH is around 700 in cytoplasm, and the ratio of NADP/NADPH is 0.005. Thus, enzymes using NAD as redox carriers catalyze the oxidation and enzymes using NADP to catalyze reduction of retinoids.

Though ADHs and RDHs can use retinol and retinal as substrates, most of these enzymes can metabolize other alcohols including sterols. Much of what we know about the physiological roles of different ADHs and RDHs stems from loss-of-function studies in experimental animals, such as knockout mice. Genetic inactivation of ADH1, 3, and 4 results in viable and fertile offspring on vitamin A sufficient diets. However, ADH-deficient mice are more sensitive than wild-type mice either to deficiency or excess of retinoids. In mouse models, ADH1 protects against vitamin A toxicity, whereas ADH4 promotes survival during vitamin A deficiency. ADH3, which is ubiquitously and redundantly expressed, functions in preventing both scenarios.

Several RDHs (RDH5, 8, 10, 11, 12, 13, and 14) have been identified that participate in ocular retinoid metabolism, and are expressed in the RPE and retina. These enzymes differ with respect to their preferred isomeric substrate and roles in visual chromophore metabolism (see below). For extraocular retinoid metabolism, the critical role of RDH10 has been well established. Loss of function of RDH10 is embryonic lethal. This lethality is due to a lack of embryonic RA synthesis, indicating that RDH10 is essentially required for the oxidation of retinol to retinal in the embryo. The expression of RDH10 in adult tissues, including the eyes, indicates that the enzyme contributes to retinoid homeostasis

throughout the mammalian life cycle. Evidence from experimental mice suggests that RDH1 contributes to retinoid homeostasis in extraocular tissues as well. Though not lethal, mutations in the corresponding gene are associated with increased hepatic retinyl ester stores and altered body fat mass. Additionally, a critical retinaldehyde reductase (DHRS3) has been identified, which contributes to RA metabolism. DHRS3 reduces vitamin A aldehyde to retinol, and the levels of this protein are transcriptionally controlled by RA signaling. Loss-of-function studies in mice indicate that DHRS3 is required to control embryonic RA levels by limiting the availability of retinal for oxidation to RA. Because DHRS3 is expressed in adult tissue, the enzyme might be critical for tissue control of RA levels postdevelopmentally. Esterification of vitamin A also seemingly contributes to the oxidation state of retinoids in tissues. Mice deficient for LRAT display highly increased hepatic levels of CYP26A1, a major RA catabolizing enzyme, indicating that in the absence of esterification retinol is constantly oxidized via retinal to RA, which must be catabolically eliminated via CYP26A1.

The final oxidation step, the conversion of retinal to RA, is catalyzed by specific members of the aldehyde dehydrogenases (ALDHs) family. The cytosolic Class I ALDHs were initially termed ALDH1 in human, Ahd2 in the mouse, and RALDH or RalDHI in rat. Mammalian genomes encode three different family members, which are now termed ALDH1A1, ALDH1A2, and ALDH1A3, respectively. Much of what is known about the physiological roles of embryonic RA synthesis comes from genetic dissection in single and compound knockout mice for these enzymes. These studies clearly identified Aldh1A2 as a major enzymatic system for the production of RA, as indicated by early embryonic death of null mice. Aldh1A3 is expressed in specific compartments later during mouse development. Aldh1A3 null mice display eye defects, and die at birth due to respiratory distress consistent with the enzyme's role at later developmental stages. In humans, Aldh1A3 gene mutations are associated with congenital microphthalmia. Aldh1A1 knockout mice develop normally and are fertile. These mice display several metabolic anomalies during adult life, e.g., they are resistant to diet-induced obesity.

The cytochrome P450-dependent enzymes CYP26A1, CYP26B1, and CYP26C1 catalyze the production of 4-hydroxy and 4-oxo-RA. Initially, hydroxylated RA had been implicated as the active form of RA, since these metabolites can act as teratogens when applied in excessive amounts to vertebrate embryos. However, studies in mouse models indicate that these molecules are not instructive signals during development. This belief is corroborated by the mutually exclusive expression patterns of CYP26A1 and RA-synthesizing enzyme ALDH1A2 in the vertebrate embryo. Furthermore,

developmental malformations of CYP26A1-deficient vertebrate embryos indicate that the absence of this enzyme induces RA toxicity rather than RA deficiency in the embryo. Finally, the responsiveness of CYP26A1 expression to RA indicates that this enzyme contributes to the catabolism of retinoids. This metabolic regulation is found for many CYP enzymes that catabolize dietary lipids and prevent their excessive accumulation. The regulation of the activities of the other RA catabolizing enzymes, CYP26B1 and CYP26C1, is more complex, and the description of their physiological role is a subject of ongoing research.

METABOLISM OF VISUAL CHROMOPHORE

Once absorbed by vertebrate eyes, vitamin A must be converted to the visual chromophore to establish and sustain vision. Individual steps in the retinoid visual cycle have been delineated in biochemical detail, and the function of key enzymes has been confirmed in animal models. Additionally, mutations in genes encoding these proteins are associated with various blinding diseases, have also provided insights into gene functions, and highlight the critical need for homeostatic maintenance of retinoids in vision.

In the disc membranes of photoreceptor outer segments, the light-sensitive proteins, rhodopsin in rods and cone opsins in cones, exist as integral membrane proteins, and the chromophore, 11-*cis*-retinal, is covalently bound via a Schiff base. Varying degrees of light absorption by each holo-opsin induces a *cis*-to-*trans* isomerization of the protein-bound chromophore to initiate phototransduction. Hydrolysis of the Schiff base linkage by bulk water entering from the cytoplasmic side liberates the retinaldehyde photoproduct. Some is released into the disc lumen and must be transferred to the cytosol by ATP-binding cassette transporter 4 (ABCA4). The next step in the visual cycle involves reduction of all-*trans*-retinal to all-*trans*-retinol catalyzed by RDHs—here acting as reductases. Two enzymes, RDH8 in photoreceptor outer segments and RDH12 in photoreceptor inner segments, which employ NADPH as a cofactor, are mainly responsible for catalyzing this reaction in mouse photoreceptors. The redundancy of retinal reductase activity shown in mouse knockout models suggests that photoreceptors contain additional RDHs other than RDH12 and RDH8. This redundancy could be explained by the need for a large enzymatic capacity in order to convert the highly, chemically reactive aldehyde group of the photoproduct to the less chemically reactive alcohol under bright light conditions.

After bright light bleaching of rhodopsin, the photoproduct can exist in millimolar concentrations within cells. The aldehyde group of the photoproduct can form adducts with primary amino groups that exist in many cellular molecules, including lipids, proteins, and ribonucleotides. The natural occurrence of such an aberrant side reaction is documented by the presence of the bisretinoid A2E, formed by a condensation reaction of two molecules of retinaldehyde with the membrane lipid phosphatidylethanolamine. The importance for rapid clearance of the photoproduct is also demonstrated by the consequences of mutations in RDH8 and RDH12 in humans. These mutations have been associated with several ocular diseases, including age-related macular degeneration, early onset autosomal recessive retinal dystrophy, retinitis pigmentosa, and Leber congenital amaurosis. Knockout mouse models with impaired retinal clearance have been established to characterize the underlying pathology. In these models, several mechanisms by which the photoproduct induces photoreceptor cell death have been proposed and may involve oxidative stress, action of Toll-like receptors, and microglia activation.

ROL formed in photoreceptor outer segments (POS) is transported to the retinal pigment epithelium (RPE), where it is esterified. This process is facilitated by two retinoid-binding proteins: interphotoreceptor retinoid-binding protein (IRBP), which binds retinoids in the extracellular space, and cellular retinol binding protein 1 (CRBP1), located within RPE cells. The ester synthase in the RPE, LRAT, comprises an important role in ocular retinoid metabolism, as it is required for the clearance of all-*trans*-retinol from the photoreceptor outer segments and, as aforementioned, the uptake of all-*trans*-retinol bound to RBP4 from the blood. Due to their high hydrophobicity, all-*trans*-retinyl esters constitute a stable storage form of vitamin A within internal membranes and oil droplet-like structures called retinosomes. Additionally, all-*trans*-retinyl ester serves as substrate for the retinal pigment epithelium-specific protein 65 kDa (RPE65), which catalyzes the endothermic transformation of all-*trans*-retinoid to its 11-*cis* conformation. The biochemical consequences of these mutations underlie a severe reduction or complete absence of the 11-*cis*-retinoid product. Disrupting the enzymatic steps of chromophore regeneration in the RPE, especially those involving LRAT and RPE65, or a combination of both, has severe consequences for retinal health. The resulting chromophore deficiency causes slow progressive death of rods that is attributed to continuous activation of visual phototransduction by unliganded opsin. It is estimated that mutations in the RPE65 gene account for 2% of autosomal recessive

retinitis pigmentosa and up to 16% of Leber congenital amaurosis.

The resulting 11-*cis*-retinol is subsequently oxidized in the final catalytic step of the visual cycle to 11-*cis*-retinal by RDH5 and RDH11, with possibly additional 11-*cis*-RDHs participating within the RPE. Similar to RDH8 and RDH12, RDH5 and RDH11 mutations have been associated with ocular diseases including fundus albipunctatus, retinal dystrophy, and juvenile cataracts. Newly synthesized and highly unstable 11-*cis*-retinal is protected by binding to CRALBP, which mediates its transport back to photoreceptor outer segments where the chromophore can once again couple to opsin, thereby completing the cycle.

The retinoid cycle described above most accurately depicts the visual chromophore regeneration between rod photoreceptor cells and the RPE as defined in rod-dominant mouse models. A proposed complementary cone visual cycle has yet to be fully determined, but mutations in many of the molecular players described so far also are responsible for cone dystrophies. The common factor arises from a deficiency of the visual chromophore. A parallel disordered vectorial transport of cone visual pigments lacking bound-chromophore leads to very rapid cone degeneration. The molecular and biochemical description of cone-specific mechanism for chromophore supply is the subject of intense research.

CONCLUSION

Much progress has been made in elucidating the chemistry and metabolism of retinoids and carotenoids, as well as the structures and full extent of physiological functions of the processing proteins. This knowledge provides the framework for furthering our understanding of the genetic polymorphisms and mutations in key genes that limit or ablate the metabolism into active forms, distribution, and transport of these necessary micronutrients. The ever-rising number of associated diseases with these defunct genes highlights the need for a refinement in interventions for individuals and genetic subgroups.

Suggested Reading

- Hall, J.A., Grainger, J.R., Spencer, S.P., Belkaid, Y., 2011. The role of retinoic acid in tolerance and immunity. *Immunity* 35, 13–22.
- Kedishvili, N.Y., 2013. Enzymology of retinoic acid biosynthesis and degradation. *J Lipid Res* 54, 1744–1760.
- Kiser, P.D., Golczak, M., Palczewski, K., 2014. Chemistry of the retinoid (visual) cycle. *Chem Rev* 114, 194–232.
- Rhinn, M., Dolle, P., 2012. Retinoic acid signalling during development. *Development* 139, 843–858.
- Sun, H., 2012. Membrane receptors and transporters involved in the function and transport of vitamin A and its derivatives. *Biochim Biophys Acta* 1821, 99–112.
- von Lintig, J., 2010. Colors with functions: elucidating the biochemical and molecular basis of carotenoid metabolism. *Annu Rev Nutr* 30, 35–56.

Vitamin D

Karina Fischer^{1,2}

¹Clinical Sciences & Epidemiology, R&D Cube Philip Morris International, Neuchatel, Switzerland; ²Department of Geriatrics and Aging Research & Centre on Aging and Mobility, University Hospital Zurich & University of Zurich, Zurich, Switzerland

Glossary Terms and Keywords

Allele Each of two or more alternative forms of a gene or a specific DNA sequence (locus) that arise by mutation and are found at the same place on a chromosome.

Candidate gene (studies) A particular gene selected for a study investigating the association between genetic variation within the gene and a phenotype or disease state of interest based on the known biological, physiological, or functional relevance of the gene to the phenotype or disease.

Carrier (genetic carrier) A person who has inherited one copy of a recessive allele for a genetic trait or mutation but usually does not display that trait or show symptoms of the disease.

CUBN The gene encoding the protein cubilin (also known as intrinsic factor-cobalamin receptor), a receptor protein in the epithelia of the proximal renal tubules that mediates endocytosis and intracellular trafficking of the filtered 25-hydroxyvitamin D (25(OH)D)-vitamin D binding protein (DBP) complex together with the protein megalin.

CYP2R1 The gene encoding vitamin D 25-hydroxylase, a member of the CYP2 family of cytochrome P450 proteins, involved in metabolizing foreign compounds.

CYP24A1 The gene encoding vitamin D 24-hydroxylase, a mitochondrial enzyme responsible for inactivating vitamin D metabolites through the C-24 oxidation pathway.

CYP27B1 The gene encoding 1 α -hydroxylase, an enzyme in the renal proximal tubule that catalyzes the hydroxylation of 25-hydroxyvitamin D (25(OH)D) into 1,25 dihydroxyvitamin D (1,25(OH)₂D).

Cytochrome P450 (CYP-450) mixed-function oxidases Oxidase heme proteins catalyzing a reaction in which each of the two atoms of oxygen in O₂ is used for a different function in the reaction. The microsomal cytochrome P450 mixed function oxidase system is important for metabolizing vitamin D₃ and is present in the liver and skin.

DHCR7 The gene encoding delta-7-sterol reductase, the penultimate enzyme of mammalian sterol biosynthesis that converts 7-dehydrocholesterol to cholesterol.

Enzyme A catalytic protein that accelerates a specific biochemical (metabolic) reaction.

Gene A segment of deoxyribonucleic acid (DNA) encoding a protein or a functionally distinct RNA molecule.

Genetic variation Genetic differences in and among populations fundamentally caused by a permanent change in the chemical structure of chromosomes (mutation) and the exchange of genes during reproduction.

GC (group-specific component) The gene encoding vitamin D-binding protein (also known as human group-specific component (GC) or vitamin D-binding α -globulin), the major vitamin D-binding protein in the plasma.

Genome The sum of the genetic and epigenetic information of an organism.

Genome-wide association study (GWAS) An observational study between a genome-wide set of genetic variants (usually single nucleotide polymorphisms) and a specific trait (e.g., disease) in different individuals to assess whether any of the variants is associated with the trait.

Haplotype (haploid genotype) A group of alleles that are inherited together in a chromosome segment from a single parent.

Linkage disequilibrium A measure that describes how closely the alleles at two loci are statistically correlated.

Locus (pl. loci) The position of a particular DNA segment on a chromosome.

LRP2 The gene encoding megalin (also known as low-density lipoprotein receptor-related protein 2), a receptor-related protein that mediates endocytosis and intracellular trafficking of the filtered 25-hydroxyvitamin D (25(OH)D)-vitamin D binding protein (DBP) complex together with the protein cubilin.

MC1R The gene encoding the melanocortin 11 receptor, a G protein-coupled receptor located on the plasma membrane of specialized cells (known as melanocytes), one of the key proteins involved in regulating mammalian skin and hair color.

Melanin A group of natural pigments produced by oxidation of the amino acid tyrosine and subsequent polymerization in specialized cells known as melanocytes, found in the basal layer of the epidermis.

Metaanalysis Statistical analysis that combines the results of multiple scientific studies.

Mutation Alteration in the order of bases in the nucleotide sequence of the DNA in which only mutations in germ line cells, but not other cells, are transmitted to offspring.

NADSYN1 The gene encoding NAD synthetase that catalyzes the final step in the biosynthesis of nicotinamide adenine dinucleotide (NAD), a coenzyme in metabolic redox reactions including cholesterol and thus also vitamin D₃ synthesis.

Osteomalacia A bone-thinning disorder that occurs exclusively in adults and is characterized by proximal muscle weakness and bone fragility.

Osteoporosis A condition that is characterized by reduced bone mineral density and increased bone fragility.

Phenotype The observable (physical) appearance of an individual with a particular genetic (genotype) and epigenetic information.

- Polymorphism** Commonly occurring difference in the DNA sequence among individuals of the same population with a minimum frequency of typically >1%.
- Rickets** A childhood disease characterized by impeded growth and deformity of the long bones.
- RXR α** The gene encoding the nuclear retinoid X receptor α , a receptor that heterodimerizes with other transcription factors such as the nuclear vitamin D receptor to regulate gene transcription by binding to specific DNA sequences in the promotor region of vitamin D-regulated genes.
- Single nucleotide polymorphism (SNP)** A variation in a single nucleotide of the DNA that occurs at a specific position in the genome with a minimum frequency of typically >1% within a population.
- Ultraviolet B (UV-B) light** Electronic radiation (280–315/320 nm) that has a longer wavelength than x-rays (0.01–10 nm) but shorter than visible light (390–700 nm) and is responsible for the formation of vitamin D in the epidermis of the skin.
- Vitamin D** A group of fat-soluble secosteroids with multiple biological functions including the intestinal absorption of calcium. In humans, the most important compounds are vitamin D₃ (cholecalciferol) and vitamin D₂ (ergocalciferol).
- Vitamin D status** The stores of vitamin D in the body, typically measured by the concentration of 25-hydroxyvitamin D in the blood and the following sufficiency categories: deficiency: <20 ng/mL; insufficiency: 20–29 ng/mL; normal: 30–100 ng/mL
- VDR** The gene encoding the protein nuclear vitamin D receptor, an intracellular hormone receptor that specifically binds 1,25(OH)₂D and mediates its effects.

INTRODUCTION

Vitamin D refers to a group of fat-soluble secosteroid prohormones. It is particularly crucial for maintaining bone health, but it has also other nonskeletal functions (Dastani et al., 2013). In human nutrition, the most important vitamin D metabolite is cholecalciferol (vitamin D₃), whereas ergocalciferol (vitamin D₂) is less common and less biologically effective (Berry and Hypponen, 2011; Bikle, 2014; Kohlmeier, 2015). Vitamin D₃ and D₂ are usually jointly referred to as vitamin D because the metabolism of the two compounds is considered to be similar (Berry and Hypponen, 2011). The major natural source of vitamin D is the synthesis of vitamin D₃ in the skin by the action of sunlight; few foods contain considerable amounts of vitamin D. Because vitamin D from both dermal synthesis and diet is biologically inactive, enzymatic conversion of the inert forms to the biologically active hormonal vitamin D form is needed (Berry and Hypponen, 2011; Bikle, 2014; Kohlmeier, 2015). There has been particular interest in the role of genetic variants in vitamin D metabolism as well as their impact on vitamin D status, various health outcomes, and consequences for dietary and health recommendations (McGrath et al., 2010).

Sources: Vitamin D₃ can be synthesized in the epidermis of the *skin* by sun exposure (UV-B light) from a cholesterol precursor molecule (Bikle, 2014; Kohlmeier, 2015). Skin pigmentation as well as clothing and sunscreen decrease the effective UV-B dose and thus

the dermal production of vitamin D (Bikle, 2014). Major *dietary* sources of vitamin D₃ are fatty ocean fish and eggs together with fortified foods (such as milk, margarine, and breakfast cereals) and, notably, dietary supplements (Berry and Hypponen, 2011; Dastani et al., 2013; Kohlmeier, 2015). Ergosterol, the precursor of vitamin D₂, is found in plants and mushrooms. The synthesis of vitamin D₂ from ergosterol also depends on exposure to UV-B light (Bikle, 2014). Today, consumption of vitamin D with unfortified foods is usually too low to compensate for the lack of sufficient sun exposure. Therefore, fortified foods and vitamin D supplements are important dietary sources (Kohlmeier, 2015). During seasons with low sun exposure (e.g., winter at high latitudes), vitamin D stored in the liver and, to a lesser extent, in extrahepatic tissues can sustain normal vitamin-dependent functions even in the absence of significant dietary vitamin D intake (Kohlmeier, 2015).

Function: Vitamin D has traditionally been linked to calcium and phosphate homeostasis as well as to bone health. More recently, however, vitamin D has also been shown to regulate cell cycle, hormone secretion, and immune function, and thereby also to affect nonskeletal outcomes such as muscle and connective tissues, cancer, cardiometabolic diseases, energy metabolism, and autoimmune diseases including multiple sclerosis (Balvers et al., 2015; Bikle, 2014; Dastani et al., 2013). However, in humans, evidence for most of these nonskeletal actions of vitamin D has been derived from observational studies that do not allow for cause–effect inference (Balvers et al., 2015; Bikle, 2014). Most actions of the biologically active form of vitamin D are mediated by the nuclear vitamin D receptor, a transcription factor that regulates hundreds of genes resulting in widespread actions of vitamin D on various metabolic processes (Bikle, 2014). In addition, vitamin D has a number of nongenomic actions including binding to membrane receptors to stimulate calcium transport across the plasma membrane of cells (Bikle, 2014; Dastani et al., 2013).

Requirement: A young light-skinned person can produce enough daily vitamin D by exposing face and arms about 10–15 min/day to summer *sun* (UV-B light). Dark-skinned and older persons, however, need several times greater exposure (Kohlmeier, 2015). Because of uncertainty about the increased risk of cancer from sunlight, estimated dietary reference intakes of vitamin D usually assume only minimal sun exposure (US Institute of Medicine (IOM); Institute of Medicine, 2011). Accordingly, from *diet*, adults should get at least 600 IU/day (15 μ g/day) (1 μ g = 40 IU) vitamin D and 800 IU/day (20 μ g/day) with advanced age (Institute of Medicine, 2011). Given these recommendations, today's typical North American and European dietary intake of vitamin D tends to be too low to provide an adequate vitamin D status (Kohlmeier, 2015). Moreover, these IOM

recommendations refer to minimum but not optimum vitamin D intakes to improve health (Balvers et al., 2015).

Deficiency and excessive intake: Vitamin D deficiency is a worldwide problem, and several organizations have published guidelines for adequate vitamin D intake and status (Balvers et al., 2015). Severe vitamin D deficiency during childhood causes rickets characterized by bone deformities in the lower limbs. Usually rickets is caused by a combination of having insufficient exposure to sunlight together with genetic vulnerability such as having dark skin pigmentation (Kohlmeier, 2013). At a later age, vitamin D deficiency causes loss of bone mineral (osteoporosis), and in severe cases, loss of bone minerals and connective tissue (osteomalacia) (Kohlmeier, 2015). The safe upper limit of vitamin D intake for adults has been defined as 100 µg/d (4000 IU/d), whereas for children up to 8 years it is 1000–3000 IU (Institute of Medicine, 2011). Prolonged intake of several hundred or 1000 micrograms per day, however, may cause hypercalcemia and soft tissue calcification or even coma and death (Kohlmeier, 2015).

OVERVIEW OF VITAMIN D METABOLISM

Both vitamin D originating from dermal production and vitamin D from dietary sources has to undergo a series of two consecutive enzymatic hydroxylation reactions in the liver and kidney to become biologically active (Fig. 32.1). A third hydroxylation step in the kidney can finally deactivate active vitamin D metabolites (Bikle, 2014; Dastani et al., 2013). These three steps in vitamin D metabolism are all performed by cytochrome P450 mixed-function oxidases (Bikle, 2014). Vitamin D produced in the epidermis binds directly to vitamin D binding protein (DBP) in the dermal blood capillaries to enter the circulation and be transported to the liver (Berry and Hypponen, 2011; Kohlmeier, 2015). In contrast, vitamin D from food sources first has to be absorbed during fat digestion and released into the lymphatic system as part of chylomicrons before it can enter the circulation preferentially bound to DBP and be transported to the liver. The first step in bioactivation of inactive vitamin D is 25-hydroxylation to 25-hydroxyvitamin D (25(OH)D), the major circulating

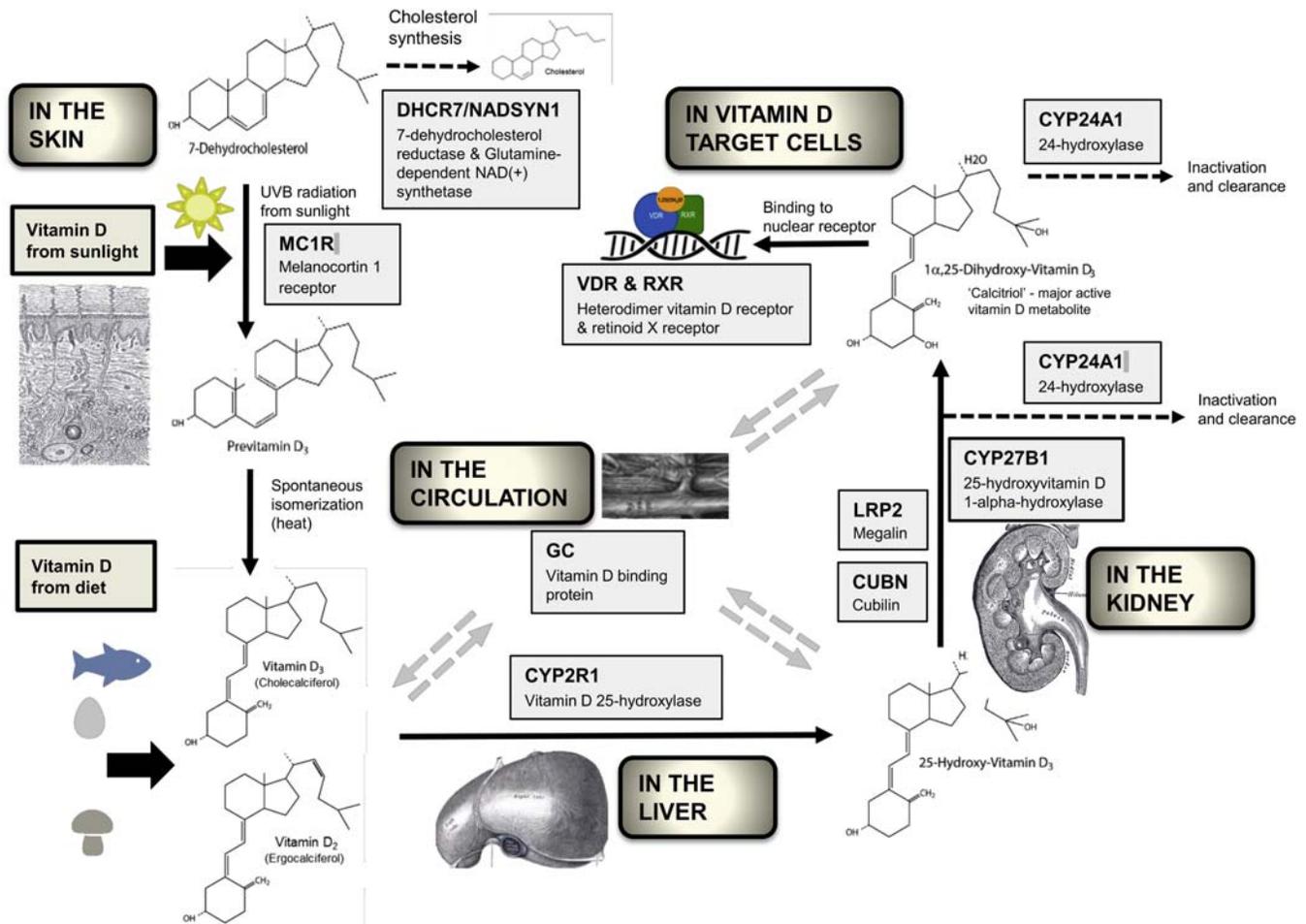


FIGURE 32.1 Genetic variation in vitamin D metabolism.

form of vitamin D in the blood. This hydroxylation takes place primarily in the liver (Bikle, 2014; Dastani et al., 2013). The 25(OH)D produced in the liver can then be stored locally or again be secreted into the circulation as a complex with DBP and transported to the kidneys (Kohlmeier, 2015). The second step in bioactivation of vitamin D is 1 α -hydroxylation of 25(OH)D to 1,25-dihydroxyvitamin D (1,25(OH)₂D), the most potent and biologically active metabolite of vitamin D. This step takes place primarily in the kidney, but also in some other tissues (Berry and Hypponen, 2011; Bikle, 2014; Dastani et al., 2013). Most actions of 1,25(OH)₂D are mediated by the nuclear vitamin D receptor that regulates hundreds of genes involved in vitamin D-related metabolic pathways (Bikle, 2014). Eventually, active vitamin D metabolites can be hydroxylated at the 24 position to be deactivated and further catabolized in nearly all vitamin D target cells (Bikle, 2014). Specific cells, including the keratinocytes in the skin and white blood cells, are equipped with all relevant metabolic enzymes for the activation and inactivation of vitamin D metabolites (Kohlmeier, 2015).

MAJOR PROTEINS AND GENETIC VARIATION IN VITAMIN D METABOLISM RELATED TO VITAMIN D STATUS

In the past decades, many studies reported associations between vitamin D status and various health outcomes. More recently, research has turned to studies on gene-environment interactions investigating whether particular variants of vitamin D-related genes may modify the association between vitamin D supply and vitamin D status or vitamin D-related health outcomes. The first step in understanding the interplay between environmental factors and genetic variation on vitamin D-related outcomes, however, is to understand which particular genetic variants may influence the vitamin D status and whether these specific single nucleotide polymorphisms (SNPs), haplotypes, or other polymorphisms translate to a more or less efficient enzyme, receptor, or binding protein (McGrath et al., 2010).

Several candidate gene studies on polymorphisms involved in vitamin D-related pathways as well as genome-wide association studies (GWAS) (Lasky-Su et al., 2012; O'Brien et al., 2018) and large-scale meta-analyses of GWAS (Ahn et al., 2010; Jiang et al., 2018; Wang et al., 2010) investigating the association of genome-wide SNPs with vitamin D status were reviewed (Berry and Hypponen, 2011; Dastani et al., 2013; McGrath et al., 2010). These studies particularly investigated associations between genetic variation in six to nine of the most recognized vitamin D-related genes (including the DHCR7/NADSYN1 locus, GC, CYP2R1/CYP27A1,

CYP27B1, CYP24A1, or VDR/RXRA) and serum levels of vitamin D. The following section focuses on these genes. It also addresses three other genes (MC1R, CUBN, and LRP2) related to vitamin D metabolism and status. However, most candidate gene studies and GWAS on genetic variation in vitamin D metabolism and vitamin status were conducted in European or Asian populations, and single SNP findings for a specific locus could not be replicated in other ethnic groups/races (e.g., black Americans). Moreover, the distribution of genetic variants such as SNPs or haplotypes in vitamin D metabolism may differ markedly by race (Powe et al., 2013).

In the Skin

In the epidermis of the skin, vitamin D₃ is produced from the cholesterol precursor 7-dehydrocholesterol (7-DHC) through a nonenzymatic two-step process in which one of the rings of 7-DHC is broken by medium-wave UV-B light (280–320 nm) to form previtamin D₃ that is further isomerized to vitamin D₃. Alternatively, 7-DHC can be reduced to cholesterol by the action of the enzyme 7-DHC reductase (encoded by the gene DHCR7) that catalyzes the final step of cholesterol synthesis and thereby reduces the substrate available for vitamin D₃ synthesis and thus the 25(OH)D concentration in blood (Berry and Hypponen, 2011). For this action, DHCR7 uses the reduced form of nicotinamide adenine dinucleotide (phosphate) (NAD(P)), a coenzyme in common metabolic redox reactions. The final step in the biosynthesis of NAD(P) in turn is catalyzed by NAD synthetase 1 (encoded by the gene NADSYN1). Both UV-B intensity and skin pigmentation contribute to the rate of vitamin D₃ formation in the epidermis (Bikle, 2014). Melanin pigments in the skin provide a highly effective protection against the harmful effects of UV light (Kohlmeier, 2013), with the production of melanin increasing owing to sun exposure. However, next to protecting the skin against UV radiation, increased melanin production also blocks UV-B light from reaching 7-DHC and thus limits the production of vitamin D₃ (Bikle, 2014; Kohlmeier, 2013). A key player among more than 100 proteins involved in human skin pigmentation is the protein melanocortin 1 receptor (encoded by the gene MC1R) (Kohlmeier, 2013). This receptor is located on the plasma membrane of specific melanin-producing cells (melanocytes) in the bottom layer of the skin's epidermis (Dessinioti et al., 2011).

DHCR7 (OMIM: 602858) and NADSYN1 (OMIM: 608285) are two neighboring genes encoding the enzymes 7-DHC reductase and NAD synthetase 1, respectively. Because of the high linkage disequilibrium (LD) of SNPs in the region of these two genes, this locus is usually referred to as the DHCR7-NADSYN1 locus (Ahn et al., 2010). Mutations in DHCR7 are known to cause

the inborn Smith–Lemli–Opitz syndrome, in which impaired activity of DHCR7 leads to accumulation of 7-DHC and a deficiency of cholesterol together with congenital abnormalities and intellectual disabilities (Berry and Hypponen, 2011; Kohlmeier, 2013). Similarly, a DHCR7 allele encoding a less active form of DHCR7 can be expected to produce less cholesterol and increase the concentration of 7-DHC, and thus give UV-B light a better opportunity to generate vitamin D₃ and 25(OH)D, whereas an allele that encodes a more active enzyme would produce more cholesterol at the expense of vitamin D₃ (Kohlmeier, 2013). In a metaanalysis of GWAS from five cohorts involving 4501 persons of European ancestry, rs3829251 in NADSYN1 was associated with 25(OH)D blood levels (Ahn et al., 2010). This SNP was in high LD with another SNP (rs1790349) in DHCR7 that was also associated with 25(OH)D concentrations. The association for the DHCR7-NADSYN1 locus with 25(OH)D levels was later confirmed in a replication set for rs11234027 in DHCR7 that was in complete LD with rs3829251 in NADSYN1 (Ahn et al., 2010). Moreover, the association of rs3829251 with 25(OH)D levels was also confirmed in United Kingdom older adults (Jolliffe et al., 2016). Another large-scale metaanalysis of GWAS involving 33,996 individuals of European descent from 15 cohorts identified a further association between rs7944926 in DHCR7 and 25(OH)D blood levels. This SNP was in complete LD with another SNP in DHCR7 (rs12785878) in a replication sample (Wang et al., 2010) and also subsequent GWAS (Lasky-Su et al., 2012; O'Brien et al., 2018) and a large-scale metaanalysis of GWAS (Jiang et al., 2018) in populations of European descent identified rs12785878 in DHCR7 as being associated with 25(OH)D levels. MC1R (OMIM: 155555) is a gene encoding the protein melanocortin 1 receptor, one of the key proteins involved in regulating human skin pigmentation and hair color. Genes controlling skin pigmentation can be regarded as the strongest inherited modifiers of vitamin D synthesis (Dessinioti et al., 2011; Kohlmeier, 2013). In the Caucasian population, MC1R is highly polymorphic; more than 100 alleles have nonsynonymous changes (Dessinioti et al., 2011). For instance, the red hair and pale skin of Dutch people have been associated with the two SNPs rs1805007 (C > T; minor allele T, Arg151Cys) and rs1805008 (C > T; minor allele T, Arg160Trp) in MC1R (Dessinioti et al., 2011; Kohlmeier, 2013). Evolutionary, the south–north gradient of skin pigmentation density reflects a rapid adaptation to the local availability of UV-B light for vitamin D synthesis owing to selective pressure (Berry and Hypponen, 2011), with populations migrating further north, where UV exposure is minimal, obtaining lighter skin (Kohlmeier, 2013).

In the Liver

In the first step of bioactivation taking place primarily in the liver (Bikle, 2014; Dastani et al., 2013), vitamin D is initially hydroxylated to 25(OH)D by vitamin D 25-hydroxylase. Among the genes encoding one mitochondrial and five microsomal cytochrome P450 isoform enzymes with vitamin D 25-hydroxylase activity, evidence suggests that in humans, the enzyme microsomal cytochrome P450 2R1 (encoded by the gene **CYP2R1**) is the major vitamin D 25-hydroxylase involved in the first step of vitamin D bioactivation (Berry and Hypponen, 2011; Bikle, 2014). Another candidate, mitochondrial cytochrome P450 27A1 (encoded by the gene **CYP27A1**, OMIM 606530) seems to be of minor importance and solely metabolizes vitamin D₃ (Bikle, 2014).

CYP2R1 (OMIM: 608713) is a gene encoding the enzyme microsomal cytochrome P450 2R1, also known as vitamin D 25-hydroxylase. In a metaanalysis of GWAS from five cohorts including 4501 persons of European ancestry, rs2060793 in CYP2R1 (which was in complete LD with rs1993116 in CYP2R1 in a replication study) was associated with circulating 25(OH)D levels (Ahn et al., 2010). In another large-scale metaanalysis of GWAS involving 33,996 European individuals from 15 cohorts, rs10741657 in CYP2R1 was associated with circulating 25(OH)D in both the discovery and replication study (Wang et al., 2010). This association of rs10741657 in CYP2R1 with 25(OH)D levels was also replicated in an independent GWAS (Lasky-Su et al., 2012) and another large-scale metaanalysis of GWAS in 79,366 persons of European descent (Jiang et al., 2018). Moreover, the SNP rs10741657 was associated with serum 25(OH)D concentrations in a transmission disequilibrium study in a German population, as well as in another large study including over 10,000 British individuals (Dastani et al., 2013). In addition, associations between two other SNPs (rs12794714 and rs10766197) in CYP2R1 and 25(OH)D blood concentrations were reported (Dastani et al., 2013).

In the Kidney (and Some Other Cells)

The second step of bioactivation takes place primarily in the kidney. A small part of the DBP-bound 25(OH)D in blood is filtered in the renal glomeruli. In the primary filtrate, DBP binds with high affinity to the protein cubilin (encoded by the gene **CUBN**), a receptor protein at the brush border membrane of the proximal renal tubules. In the epithelia of the renal tubules, cubilin mediates the endocytosis and intracellular trafficking of the filtered 25(OH)D-DBP complex together with the protein megalin (encoded by the gene **LRP2**), another

low-density lipoprotein-related receptor protein (Kaseda et al., 2011; Kohlmeier, 2015; McGrath et al., 2010). Here, 25(OH)D can then be hydroxylated to the hormonally active 1,25(OH)₂D catalyzed by the mitochondrial enzyme 1 α -hydroxylase cytochrome P450 27B1 (encoded by the gene **CYP27B1**) (Berry and Hypponen, 2011; Bikle, 2014; Dastani et al., 2013). This enzyme is mainly expressed in the distal tubules of the kidney, but also partially in the brain and other cells including keratinocytes, macrophages, and osteoblasts (Dastani et al., 2013; McGrath et al., 2010). The third important hydroxylation step in vitamin D metabolism responsible for inactivation and catabolism of vitamin D metabolites takes place in the kidney as well, but it also occurs in other cells regulated by 1,25(OH)₂D. In the kidney, in both the proximal and distal tubules, the mitochondrial enzyme 1,25(OH)₂D 24-hydroxylase cytochrome P450 24A1 (encoded by the gene **CYP24A1**), which exhibits both 24- and 23-hydroxylase activity (Berry and Hypponen, 2011; Bikle, 2014), is responsible for the inactivation or catabolism of 25(OH)D and 1,25(OH)₂D to control the levels of metabolically active 1,25(OH)₂D and prevent the accumulation of toxic levels of 1,25(OH)₂D or 25(OH)D (Bikle, 2014). The enzyme 1,25(OH)₂D 24-hydroxylase is found in nearly every cell expressing the nuclear vitamin D receptor (Bikle, 2014). In metabolites with an existing 25-OH group including 25(OH)D and 1,25(OH)₂D, 24-hydroxylation usually results in reduced activity or inactivation and further catabolism. Whereas 1,25(OH)₂D is hydroxylated to 1,24,25-trihydroxyvitamin D (1,24,25(OH)₃D), 25(OH)D is hydroxylated to 24,25-dihydroxyvitamin D (24,25(OH)₃D). However, both of these inactivated metabolites may have their own biological activity or even retain some typical vitamin D activity (Bikle, 2014). In the absence of 25-hydroxylation, 24-hydroxylation may serve to activate vitamin D metabolites (Bikle, 2014). However, next to 24-hydroxylation, hydroxylation at carbon 23 as well as other modifications may occur. Further hydroxylations at the side chain of the inactivated vitamin D metabolites finally produce calcitric acid, the water-soluble 3- and 24-glucuronides of 1,25(OH)₂D that can be excreted via bile (Bikle, 2014; Kohlmeier, 2015).

CUBN (OMIM: 602997) and **LRP2** (OMIM: 600073) are two genes encoding the proteins cubilin (also known as intrinsic factor-cobalamin receptor) and megalin (also known as low-density lipoprotein receptor-related protein 2), respectively. Both are endocytic receptors in proximal tubule cells involved in the reabsorption of DBP-bound 25(OH)D from glomerular filtrates and the subsequent intracellular hydroxylation of 25(OH)D to the biologically active metabolite 1,25(OH)₂D (Kaseda et al., 2011). To date, no genetic variant in **CUBN** or **LRP2** has been related to vitamin D status. However,

because of the involvement of cubilin and megalin in the renal uptake of 25(OH)D and the conversion to 1,25(OH)₂D, homozygous carriers with less active variants of **CUBN** and **LRP2** would be expected to have a limited uptake and conversion rate of 25(OH)D. The gene **CYP27B1** (OMIM: 609506) encodes the mitochondrial enzyme 1 α -hydroxylase CYP27B1. CYP27B1 is the cytochrome P450 gene that has been most strongly associated with vitamin D status (Dastani et al., 2013). Mutations in this gene are responsible for the rare autosomal disease pseudovitamin D deficiency rickets (Bikle, 2014), and heterozygous loss-of-function variants of CYP27B1 that are known to limit 1,25(OH)₂D production have been linked to familial cases of multiple sclerosis (Kohlmeier, 2013). The SNP rs10877012 (G > T, minor allele T) in CYP27B1 has been repeatedly associated with lower levels of 25(OH)D for carriers of the G allele (Dastani et al., 2013). Because CYP27B1 functions downstream of circulating 25(OH)D, the causal SNP captured by rs10877012 could possibly alter the role of CYP27B1 in metabolic feedback loops or adjust the rate at which 25(OH)D is metabolized (Dastani et al., 2013). Associations of two other SNPs (rs4646536 and rs703842) in CYP27B1 with 25(OH)D levels reported in a Canadian multiple sclerosis study were not replicated in Hispanic and African American populations (Dastani et al., 2013). The gene **CYP24A1** (OMIM: 126065) encodes the enzyme 24-hydroxylase CYP24A1, which exhibits both 24-hydroxylase and 23-hydroxylase activity (Bikle, 2014). In a candidate gene study (Dastani et al., 2013), large-scale GWAS metaanalysis (Wang et al., 2010), and study in UK older adults (Jolliffe et al., 2016), rs6013897 (T > A, minor allele A) in CYP24A1 was associated with 25(OH)D status; carriers of the A allele had lower blood levels of 25(OH)D. Moreover, in young African Americans, the G allele at rs2248137 (C > G, minor allele G) in CYP24A1 was associated with reduced 25(OH)D concentration (Kohlmeier, 2013).

In Circulation

In the blood, vitamin D and all of its metabolites are bound to the protein DBP (also referred to as group-specific component (GC)) (85%–88%) or albumin (12%–15%), and only small amounts are free vitamin D metabolites (Dastani et al., 2013). Protein-bound vitamin D metabolites have a longer half-life in the circulation because they are less susceptible to hydroxylation and degradation (Berry and Hypponen, 2011). Almost all of the vitamin D in the circulation is 25(OH)D, and only small amounts (one one-thousandth) are 1,25(OH)₂D (Kohlmeier, 2015). Moreover, 25(OH)D has a longer half-life and is more stable in the circulation. Therefore, total body vitamin D status is usually assessed by circulating blood levels of 25(OH)D (Dastani et al., 2013).

Typical 25(OH)D concentrations of young adults living in a sun-rich region are above 50–100 nmol/L, whereas typical 1,25(OH)₂D concentrations in vitamin D–replete people tend to be around 100 pmol/L (Kohlmeier, 2015). The IOM defines desired serum 25(OH)D levels as 50 nmol/L (20 ng/mL) based on data relative to bone health. Others, however, regard these recommendation as underestimated and propose that the lower threshold for general health should be 75 or even 90–100 nmol/L (Balvers et al., 2015). Levels up to 220 nmol/L reflect the safe upper limit of vitamin D intake (Balvers et al., 2015). Dose–response studies revealed that about 1 µg/d (40 IU/d) of vitamin D intake is required for each 1-nmol/L increase in blood 25(OH)D (Balvers et al., 2015). Because of the importance of sunlight-induced synthesis of vitamin D in the skin, blood levels of 25(OH)D demonstrate a strong seasonal pattern, with the highest concentrations in late summer and the lowest in late winter and early spring (Berry and Hypponen, 2011).

GC (Gc globulin, OMIM: 139200) is a gene encoding the protein DBP, which belongs to the albumin family and transports vitamin D metabolites in the circulation. GC is the most widely studied gene with regard to vitamin D status (Berry and Hypponen, 2011; Dastani et al., 2013). Variants in the GC gene region have been consistently associated with 25(OH)D concentrations in candidate gene studies and were the top hits in a GWAS (Lasky-Su et al., 2012) and three metaanalyses of GWAS (Ahn et al., 2010; Jiang et al., 2018; Wang et al., 2010) in persons of European descent. The most commonly studied GC variants are the nonsynonymous SNPs rs7041 (c.1296T > G; or minor allele G or T, Asp432Glu) and rs4588 (c.1307C > A; minor allele A, Thr436Lys) (Dastani et al., 2013; Wang et al., 2010); rs4588 is also the top hit in a GWAS of 1829 US and Puerto Rican women (O'Brien et al., 2018). The protein sequence variants encoded by rs7041 and rs4588 are separated by only two amino acids and therefore almost always co-occur (Kohlmeier, 2013). Three common GC haplotypes that are derived from these two SNPs encode three common phenotypic DBP isoforms (Gc1s, Gc1f, and Gc2), which differ by amino acid substitutions and/or by the degree of glycosylation (Abbas et al., 2008; Berry and Hypponen, 2011; Dastani et al., 2013; Powe et al., 2013). Caucasians usually carry a combination of four GC haplotypes (Table 32.1).

These four haplotypes suggest that genetic variation in GC may lead to differences in 25(OH)D levels that are associated with ethnicity. Hap1 is most common in Caucasians, less common in Asians, and rare in Africans. By contrast, in both Africans and Asians, Hap2 is the most frequent haplotype. Most Caucasians with the allele G at rs7041 are Hap1, whereas those with the allele A at rs4588 are Hap3. Thus, among Caucasians

no GC haplotype seems to occur that has both minor alleles (rs7041G and rs4588A), probably because such a haplotype would not be viable owing to too-low blood levels of DBP and/or 25(OH)D. DBP levels have been reported to be highest in Gc1S, lowest in Gc1F, and intermediate in Gc2 homozygotes (Powe et al., 2013). Whereas the T allele at rs7041 has been reported to be associated with decreased blood levels of DBP, the A allele at rs4588 has been associated with higher levels of DBP in both blacks and whites, and together, the SNPs at rs7041 and rs4588 appear to have additive effects on the DBP concentration (Powe et al., 2013). However, when genetic variants are considered, ethnicity may explain only less than 0.1% of variation in DBP levels (Powe et al., 2013). In European populations, minor alleles at rs7041 or rs4588 themselves or as part of the three common GC haplotypes derived from these two SNPs have been widely associated with lower 25(OH)D concentrations (Berry and Hypponen, 2011; Dastani et al., 2013; Powe et al., 2013; Wang et al., 2010). In multivariable analysis adjusting for significant lifestyle determinants of vitamin D status, rs7041 was independently associated with serum 25(OH)D concentration among 222 predominantly white older adults; homozygous carriers of the minor allele had on average 10.2 nmol/L lower 25(OH)D blood levels compared with homozygous carriers of the reference allele (Jolliffe et al., 2016). Finally, gene-level principal components analysis of SNPs described an association between the GC gene and 25(OH)D levels (Dastani et al., 2013). All in all, there seems to be a south–north progression of genetic variants associated with lower circulating 25(OH)D levels as part of the adaptation among populations migrating from south to north and obtaining lighter skin. Accordingly, whereas most Europeans carry a haplotype (Hap1 or Hap3) with at least one minor allele for rs7041 or rs4588 associated with lower 25(OH)D concentrations, most Asians and Africans carry a haplotype (Hap2) with no minor allele at rs7041 or rs4588, possibly to compensate for the lower synthesis of vitamin D₃ in the skin owing to the darker skin pigmentation. Moreover, there is a north–south gradient of genetic diversity in GC. Whereas Caucasians usually carry a combination of four haplotypes mainly based on 3 common SNPs, Asians exhibit four haplotypes based on five common SNPs, and Africans six haplotypes based on eight common SNPs.

In Vitamin D Target Cells

Most, but not all actions of the biologically active 1,25(OH)₂D are mediated by the protein nuclear vitamin D receptor (encoded by the gene **VDR**), a transcription factor that mediates the hormonal actions of

TABLE 32.1 Haplotype frequency and allele combinations of single nucleotide polymorphisms (SNPs) in group-specific components.^a

Haplotype	Phenotype	Homozygous (%)	Combination of alleles in SNPs			Blood vitamin D binding protein level of homozygous haplotypes ^b	Blood 25(OH)D level of homozygous haplotypes ^c	Frequency of haplotype in ethnic groups		
			rs76803094 (G > A)	rs7041 (T > G)	rs4588 (C > A)			European (%)	Asian (%)	African (%)
Hap1	Gc1s	15.3	G	G	C	High	Medium	52.6	27.8	9.6
Hap2	Gc1f	3.9	G	T	C	Low	High	14.2	31.8	35.0
Hap3	Gc2	1.7	G	T	A	Intermediate	Low	24.5	27.1	5.5
Hap4		0.09	A	G	C	Not available	↓	3.7	0	0

A, Adenine; C, Cytosine; G, Guanine; Hap, Haplotype; T, Thymine.

^aFrom UCSC Genome Browser: Human Gene GC (uc010iif.3) Description and Page Index, Common Gene Haplotype Alleles.

^bPowe et al. (2013).

^cAbbas et al. (2008).

1,25(OH)₂D (Berry and Hypponen, 2011; Dastani et al., 2013). After binding of 1,25(OH)₂D, the nuclear vitamin D receptor heterodimerizes with other transcription factors such as the nuclear retinoic acid–containing retinoid X receptor α (encoded by the gene **RXRA**) to regulate gene transcription by binding to specific DNA sequences (vitamin D response elements) in the promoter region of vitamin D–regulated genes (Bikle, 2014; Dastani et al., 2013). The nuclear retinoid X receptor α homo- or heterodimerizes with a variety of receptors including the nuclear vitamin D receptor. As a heterodimer partner with the nuclear vitamin D receptor, the binding of 1,25(OH)₂D to the nuclear vitamin D receptor induces allosteric conformational changes in the nuclear retinoid X receptor α that facilitate the recruitment of coregulators. Numerous genes with a variety of physiologic effects are then targeted by the nuclear vitamin D receptor-retinoid X receptor α heterodimer, including those involved in regulating vitamin D metabolite concentrations via a classical endocrine feedback loop (Hibler et al., 2010).

VDR (OMIM: 601769) is a gene encoding the nuclear vitamin D receptor, a ligand-activated transcription factor that is active in nearly every tissue. The nuclear vitamin D receptor binds primarily hormonally active 1,25(OH)₂D, but to a much lesser extent also 25(OH)D or 24,25(OH)₂D (Berry and Hypponen, 2011; Bikle, 2014; Dastani et al., 2013). Whereas there is robust evidence linking polymorphisms in the VDR with increased risk for a range of adverse health outcomes such as risk for fractures, rheumatoid arthritis, tuberculosis, and various types of cancers (Berry and Hypponen, 2011), it is not clear whether and how SNPs in VDR may influence 25(OH)D concentrations (Berry and Hypponen, 2011; Hibler et al., 2010; McGrath et al., 2010). Because of the distance between the nuclear vitamin D receptor and 25(OH)D in the metabolic pathway, strong associations between the two are not expected (Berry and Hypponen, 2011). However, evidence has been found for the minor allele at rs2228570 in VDR gene that was associated with a higher level of 25(OH)D in a longitudinal population-based study. The minor allele introduces a Fok I site (FokI polymorphism) to exon 2, which leads to a vitamin D receptor protein that is three amino acids longer and results in a less active vitamin D receptor protein and less effective transcriptional activity (Berry and Hypponen, 2011; Dastani et al., 2013). This association was replicated in another cohort study (Dastani et al., 2013). In addition, for two adjacent SNPs in VDR, rs7139166 and rs4516035, the G-A haplotype showed higher promoter activity, whereas the C-G haplotype was associated with lower circulating levels of 25(OH)D (Dastani et al., 2013). Moreover, in a cross-sectional study, the minor allele of the intronic SNP rs10783219 (A > T, minor allele T) in

VDR was associated with blood 25(OH)D concentrations (Dastani et al., 2013). **RXRA** (OMIM: 600825) is a gene encoding the protein retinoid X receptor α , a ligand-activated transcription factor activated by 9-cis retinoic acid, and has a role in cell differentiation, cell growth, apoptosis, and vertebrate embryonic development (Hibler et al., 2010). To date, only one study found an association between RXRA and 25(OH)D levels (Berry and Hypponen, 2011; Hibler et al., 2010). This study in 415 patients reported that the minor allele at rs9409929 (G > A, minor allele A) in RXRA was associated with higher blood levels of 1,25(OH)₂D (Hibler et al., 2010).

References

- Abbas, S., Linseisen, J., Slanger, T., Kropp, S., Mutschelknauss, E.J., Flesch-Janys, D., Chang-Claude, J., 2008. The Gc2 allele of the vitamin D binding protein is associated with a decreased postmenopausal breast cancer risk, independent of the vitamin D status. *Cancer Epidemiol Biomark Prev* 17, 1339–1343.
- Ahn, J., Yu, K., Stolzenberg-Solomon, R., Simon, K.C., McCullough, M.L., Gallicchio, L., Jacobs, E.J., Ascherio, A., Helzlsouer, K., Jacobs, K.B., Li, Q., Weinstein, S.J., Purdue, M., Virtamo, J., Horst, R., Wheeler, W., Chanock, S., Hunter, D.J., Hayes, R.B., Kraft, P., Albanes, D., 2010. Genome-wide association study of circulating vitamin D levels. *Hum Mol Genet* 19, 2739–2745.
- Balvers, M.G., Brouwer-Brolsma, E.M., Endenburg, S., de Groot, L.C., Kok, F.J., Gunnewiek, J.K., 2015. Recommended intakes of vitamin D to optimise health, associated circulating 25-hydroxyvitamin D concentrations, and dosing regimens to treat deficiency: workshop report and overview of current literature. *J Nutr Sci* 4, e23.
- Berry, D., Hypponen, E., 2011. Determinants of vitamin D status: focus on genetic variations. *Curr Opin Nephrol Hypertens* 20, 331–336.
- Bikle, D.D., 2014. Vitamin D metabolism, mechanism of action, and clinical applications. *Chem Biol* 21, 319–329.
- Dastani, Z., Li, R., Richards, B., 2013. Genetic regulation of vitamin D levels. *Calcif Tissue Int* 92, 106–117.
- Dessinioti, C., Antoniou, C., Katsambas, A., Stratigos, A.J., 2011. Melanocortin 1 receptor variants: functional role and pigmentary associations. *Photochem Photobiol* 87, 978–987.
- Hibler, E.A., Jurutka, P.W., Egan, J.B., Hu, C., LeRoy, E.C., Martinez, M.E., Thompson, P.A., Jacobs, E.T., 2010. Association between polymorphic variation in VDR and RXRA and circulating levels of vitamin D metabolites. *J Steroid Biochem Mol Biol* 121, 438–441.
- Institute of Medicine, Dietary Reference Intakes for Calcium and Vitamin D, 2011. DC: The National Academies Press, Washington. <https://doi.org/10.17226/13050>.
- Jiang, X., O'Reilly, P.F., Aschard, H., Hsu, Y.H., Richards, J.B., Dupuis, J., Ingelsson, E., Karasik, D., Pilz, S., Berry, D., Kestenbaum, B., Zheng, J., Luan, J., Sofianopoulou, E., Streeten, E.A., Albanes, D., Lutsey, P.L., Yao, L., Tang, W., Econs, M.J., Wallaschofski, H., Volzke, H., Zhou, A., Power, C., McCarthy, M.I., Michos, E.D., Boerwinkle, E., Weinstein, S.J., Freedman, N.D., Huang, W.Y., Van Schoor, N.M., van der Velde, N., Groot, L., Enneman, A., Cupples, L.A., Booth, S.L., Vasan, R.S., Liu, C.T., Zhou, Y., Ripatti, S., Ohlsson, C., Vandenput, L., Lorentzon, M., Eriksson, J.G., Shea, M.K., Houston, D.K., Kritchevsky, S.B., Liu, Y., Lohman, K.K., Ferrucci, L., Peacock, M., Gieger, C., Beekman, M., Slagboom, E., Deelen, J., Heemst, D.V., Kleber, M.E., Marz, W., de Boer, I.H.,

- Wood, A.C., Rotter, J.I., Rich, S.S., Robinson-Cohen, C., den Heijer, M., Jarvelin, M.R., Cavadino, A., Joshi, P.K., Wilson, J.F., Hayward, C., Lind, L., Michaelsson, K., Trompet, S., Zillikens, M.C., Uitterlinden, A.G., Rivadeneira, F., Broer, L., Zgaga, L., Campbell, H., Theodoratou, E., Farrington, S.M., Timofeeva, M., Dunlop, M.G., Valdes, A.M., Tikkanen, E., Lehtimäki, T., Lyytikäinen, L.P., Kahonen, M., Raitakari, O.T., Mikkilä, V., Ikram, M.A., Sattar, N., Jukema, J.W., Wareham, N.J., Langenberg, C., Forouhi, N.G., Gundersen, T.E., Khaw, K.T., Butterworth, A.S., Danesh, J., Spector, T., Wang, T.J., Hyppönen, E., Kraft, P., Kiel, D.P., 2018. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat Commun* 9, 260.
- Jolliffe, D.A., Hanifa, Y., Witt, K.D., Venton, T.R., Rowe, M., Timms, P.M., Hyppönen, E., Walton, R.T., Griffiths, C.J., Martineau, A.R., 2016. Environmental and genetic determinants of vitamin D status among older adults in London, UK. *J Steroid Biochem Mol Biol* 164, 30–35.
- Kaseda, R., Hosojima, M., Sato, H., Saito, A., 2011. Role of megalin and cubilin in the metabolism of vitamin D(3). *Ther Apher Dial* 15 (Suppl. 1), 14–17.
- Kohlmeier, M., 2013. Chapter 4.9.11: Vitamin D, Nutrigenetics: Applying the Science of Personal Nutrition. Academic Press Inc, pp. 191–199.
- Kohlmeier, M., 2015. Chapter 9: Vitamin D, Nutrient Metabolism: Structures, Functions, and Genes. Academic Press Inc, pp. 501–514.
- Lasky-Su, J., Lange, N., Brehm, J.M., Damask, A., Soto-Quiros, M., Avila, L., Celedon, J.C., Canino, G., Cloutier, M.M., Hollis, B.W., Weiss, S.T., Litonjua, A.A., 2012. Genome-wide association analysis of circulating vitamin D levels in children with asthma. *Hum Genet* 131, 1495–1505.
- McGrath, J.J., Saha, S., Burne, T.H., Eyles, D.W., 2010. A systematic review of the association between common single nucleotide polymorphisms and 25-hydroxyvitamin D concentrations. *J Steroid Biochem Mol Biol* 121, 471–477.
- O'Brien, K.M., Sandler, D.P., Shi, M., Harmon, Q.E., Taylor, J.A., Weinberg, C.R., 2018. Genome-wide association study of serum 25-hydroxyvitamin D in US women. *Front Genet* 9, 67.
- Powe, C.E., Evans, M.K., Wenger, J., Zonderman, A.B., Berg, A.H., Nalls, M., Tamez, H., Zhang, D., Bhan, I., Karumanchi, S.A., Powe, N.R., Thadhani, R., 2013. Vitamin D-binding protein and vitamin D status of black Americans and white Americans. *N Engl J Med* 369, 1991–2000.
- Wang, T.J., Zhang, F., Richards, J.B., Kestenbaum, B., van Meurs, J.B., Berry, D., Kiel, D.P., Streeten, E.A., Ohlsson, C., Koller, D.L., Peltonen, L., Cooper, J.D., O'Reilly, P.F., Houston, D.K., Glazer, N.L., Vandenput, L., Peacock, M., Shi, J., Rivadeneira, F., McCarthy, M.I., Anneli, P., de Boer, I.H., Mangino, M., Kato, B., Smyth, D.J., Booth, S.L., Jacques, P.F., Burke, G.L., Goodarzi, M., Cheung, C.L., Wolf, M., Rice, K., Goltzman, D., Hidiroglou, N., Ladouceur, M., Wareham, N.J., Hocking, L.J., Hart, D., Arden, N.K., Cooper, C., Malik, S., Fraser, W.D., Hartikainen, A.L., Zhai, G., Macdonald, H.M., Forouhi, N.G., Loos, R.J., Reid, D.M., Hakim, A., Dennison, E., Liu, Y., Power, C., Stevens, H.E., Jaana, L., Vasani, R.S., Soranzo, N., Bojunga, J., Psaty, B.M., Lorentzon, M., Forouid, T., Harris, T.B., Hofman, A., Jansson, J.O., Cauley, J.A., Uitterlinden, A.G., Gibson, Q., Jarvelin, M.R., Karasik, D., Siscovick, D.S., Econs, M.J., Kritchevsky, S.B., Florez, J.C., Todd, J.A., Dupuis, J., Hyppönen, E., Spector, T.D., 2010. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. *Lancet* 376, 180–188.

Genetic Determinants of Vitamin E Status

Patrick Borel, Charles Desmarchelier

C2VN, Aix Marseille Univ, INRA, INSERM, Marseille, France

VITAMIN E: CHEMISTRY, DIETARY SOURCES, INTAKE, AND RECOMMENDED DIETARY ALLOWANCES

Vitamin E (VE) is the generic term for molecules that possess qualitatively the biological activity of α -tocopherol. Four tocopherols (α , β , γ , and δ) and four tocotrienols (α , β , γ , and δ) occur naturally, with α - and γ -tocopherol the main vitamers present in Western diets. Natural tocopherols are found in foods as unesterified *RRR* stereoisomers. However, synthetic VE, which is widely used as an antioxidant food additive and in dietary supplements, is usually a racemic mixture of eight stereoisomers and is esterified, usually as all-*rac*- α -tocopheryl acetate or succinate, to protect the phenol group against oxidation. Foods naturally rich in VE include some vegetable oils and nuts, but it is also found in other food matrices, e.g., wheat germ and lettuce.

The average VE intake of Americans is still below the U.S. recommended dietary allowances (15 mg/day for people ≥ 14 years of age) and about three-fourths of Americans (19–30 years of age) consume less than 10 mg/day. In Europe, 8% of men and 15% of women fail to meet 67% of the European recommended dietary allowances for VE (12 mg/day). A recent study of global α -tocopherol status has shown that only 21% of participants exhibited serum α -tocopherol concentrations $>30 \mu\text{mol/L}$, which is the proposed threshold for adequate VE status (Peter et al., 2016).

ROLE OF VITAMIN E IN HUMANS

Based on its chemistry and on in vitro experiments, VE is considered to be the primary fat-soluble dietary antioxidant consumed by humans; it acts as a chain-breaking antioxidant, especially against peroxy radicals, and is thus essential in maintaining the integrity of the long-chain polyunsaturated fatty acids found in

cell membranes. Nevertheless, other studies have shown that VE has biological activities independent of its antioxidant properties. Indeed, it can modify gene expression, inhibit cell proliferation, modulate platelet aggregation and monocyte adhesion, and is involved in the regulation of bone mass.

Based on the assumption that free radicals are involved in the etiology of several diseases, e.g., cardiovascular diseases and cancers, the potential benefit of high VE intake from foods or dietary supplements has been extensively studied. In summary, although VE intake or VE status have been consistently found to be inversely associated with the incidence of several diseases, most randomized controlled trials have failed to demonstrate an effect of VE supplementation on the incidence of these diseases. To explain the apparent discrepancy between the results obtained in population studies and randomized controlled trials, it has been suggested that genetic variations in genes involved in VE absorption and metabolism may have interfered with the protective effects of VE supplementation (Zingg et al., 2008). It is therefore crucial to identify such genetic influences and to estimate their biological effect in order to take into account this source of interindividual variability in future randomized controlled trials.

PROTEINS AND CORRESPONDING GENES INVOLVED IN VITAMIN E STATUS

The most widely used and validated marker of VE status is obtained by measuring plasma or serum α -tocopherol concentration after an overnight fast. In order to identify genetic variations that can modulate VE status, two complementary approaches have been used most commonly. Genome-wide association studies (GWAS) display the advantage that they do not make any assumption on the genetic variations that may affect

the studied disease or phenotype, thus allowing researchers to identify unexpected associations. Yet, excessive statistical stringency used in this approach might lead to false negative conclusions, i.e., to reject genetic variants that are marginally associated with the phenotype. Another drawback of GWAS is their cost. Since a very large number of genetic variants are investigated, the sample size required to limit false positive associations is high (typically >10,000 subjects). This is not necessarily an issue if the phenotype of interest is cheap to measure (e.g., with single-point measurements such as plasma/serum VE concentration), but this can become a limiting factor for more complex phenotypes (e.g., VE absorption assessed with several measurement points). Additionally, the studied population needs to be well characterized to account for confounding variables that could decrease the relative effect of genetic variation on the phenotype/disease of interest (e.g., usual VE intake in GWAS of VE status). The second approach used to identify genetic variations that can modulate VE status is candidate gene association studies. By selecting only genes or genetic variations that are assumed to be involved with the phenotype of interest, it allows researchers to use a much smaller study population, improving their capacity to characterize the population and lowering the cost of the study. However, many genes and SNPs that potentially have effects on the bioavailability of fat-soluble vitamins and phytochemicals are left out of the analyses. Moreover, it requires a thorough knowledge of the studied phenotype to appropriately select promising candidate genes. The last part of this chapter therefore lists some of the proteins that are known, assumed, or reasonably hypothesized to modulate, directly or indirectly, the fasting plasma/serum concentration of VE, i.e., VE status. Genes encoding for these proteins are therefore candidate genes for association studies.

The metabolism of vitamin E in the human digestive tract has been reviewed recently (Borel et al., 2013). It might begin in the stomach where VE esters might be partially hydrolyzed by gastric lipase. In the duodenum, VE esters are hydrolyzed, at least partly, by carboxyl ester hydrolase. Moreover, digestive enzymes, i.e., proteases, amylases, and lipases, have been suggested to be involved in the release of VE from food matrices.

After its incorporation into mixed micelles, VE is taken up by enterocytes. Significant advances on this key step of VE absorption have been made in the last 10 years. Indeed, although it was thought that VE uptake takes place only by passive diffusion, studies in cell cultures and in transgenic mice have established that at least three apical membrane proteins are involved in this process, namely SR-BI, NPC1L1, and CD36 (Reboul and Borel, 2011). Genetic variations in genes that encode for these proteins are thus candidates for the modulation

of VE absorption and consequently of VE status. After uptake across the apical side of the enterocyte, VE has to reach the basolateral side to be secreted in the body. Surprisingly, there is no data yet on protein(s) involved in its intracellular transport. Candidates could be sec14p-like proteins (encoded by *TAP1*, 2, and 3 in humans) and Niemann-Pick type C1/C2 (NPC1/2) proteins, which have been shown to transport VE in other cells. However, it is not yet known whether these proteins are expressed in enterocytes (Reboul and Borel, 2011).

Most newly absorbed VE is secreted by enterocytes with chylomicrons (Borel et al., 2013; Reboul and Borel, 2011), but recently, one membrane protein, ABCA1, has been shown to modulate VE efflux at the basolateral side of the enterocyte and another basolateral membrane protein, ATP-binding cassette subfamily G member 1 (ABCG1), has also been suggested to participate in this efflux (Olivier et al., 2014; Borel et al., 2013; Reboul and Borel, 2011). The incorporation of VE into these lipoproteins depends on several proteins, including microsomal transfer protein, which is likely involved in its incorporation in nascent chylomicrons, and ABCA1, which is involved in its secretion in nascent HDL. Chylomicrons and intestinal HDL then carry VE to the liver, which is a key organ for the metabolism of VE, and to other tissues. Apolipoproteins that are constituents of these lipoproteins, e.g., apoB48 and apoAI, as well as proteins that are involved in the blood metabolism of these lipoproteins, e.g., lipoprotein lipase, CETP, PLTP, apoB receptor, are also candidate genes for the modulation of VE status.

Following the hydrolysis of chylomicron triglycerides by lipoprotein lipase, some VE can be taken up by extrahepatic tissues through a SR-BI-dependent mechanism. Chylomicron remnants are then taken up by the liver via endocytic receptors such as the LDL receptor and heparan sulfate proteoglycans (of which syndecan 1 is the primary one). The fate of VE incorporated into intestinal HDL, or transferred to hepatic HDL during its circulation in the blood since VE can exchange in between lipoproteins, is not known. However, it is assumed that most HDL-VE is transported to the liver where it is taken up via SR-BI. After its uptake by the liver, VE can be stored, secreted into bile, resecreted into circulation with VLDL or catabolized. VE catabolism starts with an ω -hydroxylation of the side chain, which is catalyzed by CYP4F2 or CYP3A4, followed by β -oxidation, but the first step is limiting. One enzyme has been shown to metabolize VE in human: the ω -hydroxylase cytochrome P450-4F2 (CYP4F2). These enzymes are therefore good candidates to modulate tissue VE concentration and in turn blood VE status. A key protein involved in these pathways is the α -tocopherol transfer protein (α -TTP). Indeed, it binds VE, preferentially the

RRR- α -tocopherol isomer, with a high specificity and allows its incorporation into very-low-density lipoprotein (VLDL). Loss-of-function variants of the gene that encodes this protein can strongly impair the transfer of tocopherol into VLDL, resulting in a disease called "ataxia with isolated VE deficiency" It is therefore likely that genetic variations in this gene can significantly affect VE status.

After its secretion into VLDL by the liver, VE is distributed to other lipoprotein classes either by their natural conversion to LDL or by PLTP-mediated transfer to HDL. The fate of VE in lipoproteins is not well known. It is assumed that a fraction is degraded by oxidation when VE reacts with free radicals and when it is not regenerated by other antioxidants, e.g., vitamin C or polyphenols. The remaining fraction is either taken up by the liver or other tissues when lipoproteins are taken up by these organ/tissues. VE is then either degraded by oxidation when it reacts with reactive oxygen species or is metabolized to be eliminated.

The last candidate proteins/genes that could modulate blood VE concentration are those involved in the production/neutralization of reactive oxygen species. Indeed, the blood/tissue concentration of these species

can modulate the blood/tissue concentration of VE. This explains why variants of the haptoglobin gene have been associated with VE status.

GENETIC VARIATIONS ASSOCIATED WITH VITAMIN E ABSORPTION

Since VE is not synthesized in humans, all VE present in the human body comes from food and dietary supplements. VE bioavailability has been shown to be variable, with absorption in the range 10%–79%, and in a recent clinical trial where a VE-rich meal was given to a group of healthy subjects, the coefficient of variation of the VE response to the meal was 81% (Borel et al., 2015). It is thus reasonable to hypothesize that genetic variations in genes involved in VE absorption could in turn affect VE status. In the above-mentioned trial, 82% of the variability was explained by a combination of 28 SNPs in or near 11 candidate genes. Seven of these genes were involved in the postprandial chylomicron triacylglycerol response (Desmarchelier et al., 2014). This is not surprising since most newly absorbed VE is carried from the intestine to the liver via

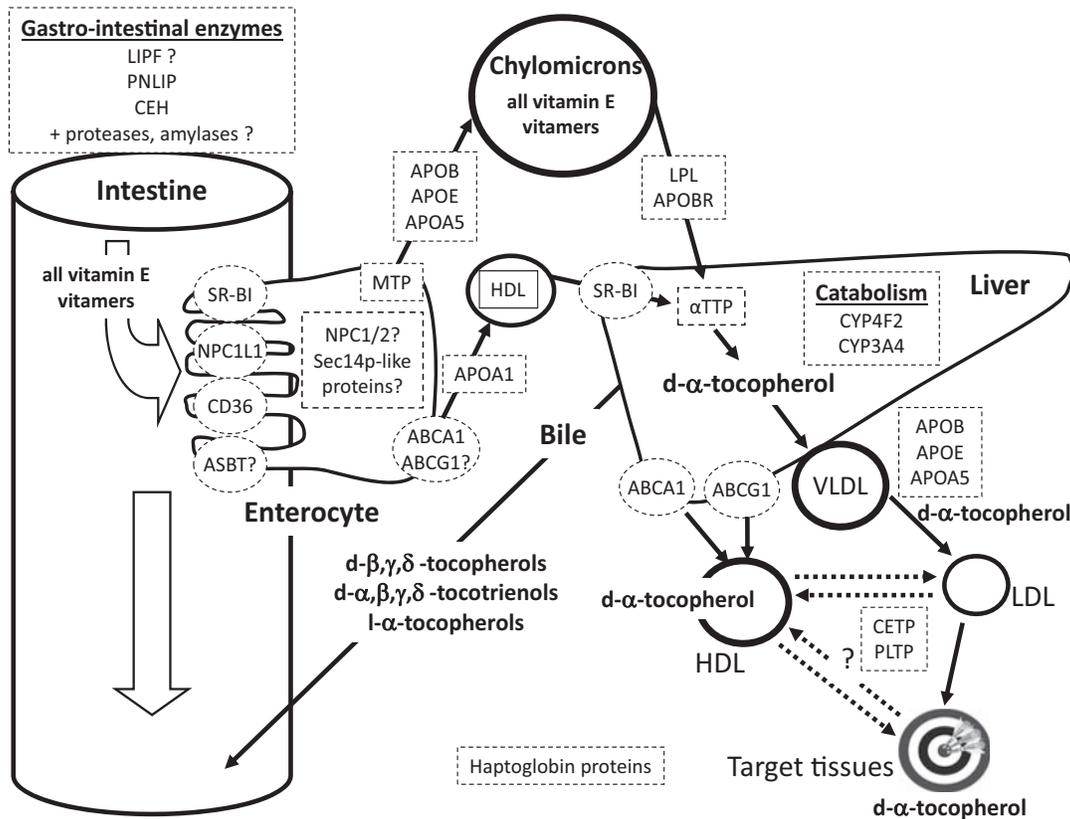


FIGURE 33.1 Summary of the proteins involved in the variability of vitamin E status. Proteins displayed are those encoded by the main candidate genes involved in VE status or by genes for which genetic variations have been associated with the variability of VE status. A question mark next to a name indicates that the proposed function of the gene is still awaiting confirmation.

TABLE 33.1 Summary of the SNPs that have been associated with either fasting blood vitamin E concentration or vitamin E bioavailability.

SNP	Global MAF ^a	Nearest gene	Trait	Reference	Study type
rs12272004	0.085	<i>APOA5</i>	Fasting blood α -tocopherol concentration	Ferrucci et al. (2009)	GWAS
rs964184	0.222	<i>BUD13/ZNF259/APOA5</i>	Fasting blood α -tocopherol concentration following α -tocopherol supplementation	Major et al. (2011)	GWAS
rs2108622	0.237	<i>CYP4F2</i>	Fasting blood α -tocopherol concentration following α -tocopherol supplementation	Major et al. (2011)	GWAS
rs11057830	0.139	<i>SCARB1</i>	Fasting blood α -tocopherol concentration following α -tocopherol supplementation	Major et al. (2011)	GWAS
rs7834588	0.433	<i>NKAIN3</i>	Fasting blood α -tocopherol concentration following α -tocopherol supplementation	Major et al. (2012)	GWAS
rs10401969	0.118	<i>SUGP1</i>	Fasting blood α -tocopherol concentration	Wood et al. (2013)	GWAS
rs58542926	0.067	<i>TM6SF2</i>	Fasting blood α -tocopherol concentration	Wood et al. (2013)	GWAS
Rs675	0.099	<i>APOA4</i>	Fasting blood α - and γ -tocopherol concentration	Borel et al. (2007)	CGAS
E2, E3, E4		<i>APOE</i>	Fasting blood α -tocopherol concentration	Borel et al. (2007)	CGAS
rs4238001	0.064	<i>SCARB1</i>	Fasting blood γ -tocopherol concentration	Borel et al. (2007)	CGAS
rs5888	0.323	<i>SCARB1</i>	Fasting blood α -tocopherol concentration	Borel et al. (2007)	CGAS
rs662799	0.163	<i>APOA5</i>	Fasting blood α -tocopherol concentration in diabetic patients	Girona et al. (2008)	CGAS
rs5128	0.234	<i>APOC3</i>	Fasting blood α -tocopherol concentration	Borel et al. (2009)	CGAS
rs708272	0.378	<i>CETP</i>	Fasting blood α -tocopherol concentration	Borel et al. (2009)	CGAS
rs1800588	0.387	<i>LIPC</i>	Fasting blood γ -tocopherol concentration	Borel et al. (2009)	CGAS
rs1527479	0.349	<i>CD36</i>	Fasting blood α -tocopherol concentration	Lecompte et al. (2011)	CGAS
rs6994076	0.349	<i>TTPA</i>	Fasting blood α -tocopherol concentration	Zanon-Moreno et al. (2013)	CGAS
rs2108622	0.237	<i>CYP4F2</i>	Fasting blood α -tocopherol concentration following α -tocopherol supplementation	Athinarayanan et al. (2014)	CGAS
rs3093105	0.157	<i>CYP4F2</i>	Fasting blood α -tocopherol concentration following α -tocopherol supplementation	Athinarayanan et al. (2014)	CGAS
rs468320	0.234	<i>ABCG1</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs2915775	0.257	<i>PNLIP</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs3010494	0.294	<i>PNLIP</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS

TABLE 33.1 Summary of the SNPs that have been associated with either fasting blood vitamin E concentration or vitamin E bioavailability.—cont'd

SNP	Global MAF ^a	Nearest gene	Trait	Reference	Study type
rs1571513	0.240	<i>SLC10A2</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs9558203	0.198	<i>SLC10A2</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs16961116	0.162	<i>SLC10A2</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs12874168	0.210	<i>SLC10A2</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs2065550	0.160	<i>SLC10A2</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs2839715	0.168	<i>SREBF2</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs4822062	0.153	<i>SREBF2</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs4149314 ^b	0.069	<i>ABCA1</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs11789603 ^b	0.117	<i>ABCA1</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs2274873 ^b	0.082	<i>ABCA1</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs4149297 ^b	0.084	<i>ABCA1</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs4643493 ^b	0.082	<i>APOB</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs1042031 ^b	0.128	<i>APOB</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs1713222 ^b	0.155	<i>APOB</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs10464587 ^b	0.297	<i>BET1</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs1316328 ^b	0.134	<i>IRS1</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs4238329 ^b	0.148	<i>LIPC</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs8041525 ^b	0.086	<i>LIPC</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs7164909 ^b	0.153	<i>LIPC</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs8035357 ^b	0.150	<i>LIPC</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs12591216 ^b	0.084	<i>LIPC</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs12593880 ^b	0.068	<i>LIPC</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs4921920 ^b	0.101	<i>NAT2</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs7296124 ^b	0.107	<i>ZNF664</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS
rs1048497 ^b	0.061	<i>ZNF664</i>	α -tocopherol bioavailability	Borel et al. (2015)	CGAS

^aAbbreviations: CGAS, candidate gene association study; MAF, minor allele frequency.

^bThese SNPs were associated with the variability of α -tocopherol bioavailability, but this association was likely due to their involvement in the postprandial metabolism of chylomicron triacylglycerol.

chylomicrons. Four of these genes were specifically associated with the VE response: solute carrier family 10 (sodium/bile acid cotransporter), member 2 (*SLC10A2*), pancreatic lipase (*PNLIP*), sterol regulatory element binding transcription factor 2 (*SREBF2*), and ATP-binding cassette, subfamily G (WHITE), member 1 (*ABCG1*). The associations observed in this study, which need to be confirmed in other populations, show that several SNPs in genes involved in different pathways can impact VE bioavailability. Although the effect of each SNP was, as expected, relatively low, their additive effect significantly contributed to explain the variability in this phenotype. Furthermore, they can

indirectly affect VE status as well because it was shown in this study that VE bioavailability was correlated with VE status.

GENETIC VARIATIONS ASSOCIATED WITH FASTING BLOOD VITAMIN E CONCENTRATIONS

As previously mentioned, fasting serum/plasma VE concentration is used as a marker for VE status. Thus, several studies have assessed associations between genetic variations and fasting serum/plasma VE

concentration. Three GWAS have shown that a SNP near *ZNF259*, *BUD13*, and *APOA1/C3/A4/A5*, a SNP in *CYP4F2*, and a SNP in *SCARB1* are associated with VE status (Major et al., 2011, 2012; Ferrucci et al., 2009). *SCARB1* encodes for SR-BI, which is a membrane receptor for HDL and has been involved in VE uptake by several tissues. Concerning *CYP4F2*, it encodes for cytochrome P450 4F2, which is an enzyme metabolizing VE in humans. The association with the SNP near *ZNF259*, *BUD13*, and the cluster *APOA1/C3/A4/A5* is likely due to a variation in *APOA5* as suggested by Ferrucci et al. (2009), since *ZNF259* and *BUD13* encode for proteins that have no known role on VE metabolism.

Candidate gene association studies have confirmed the involvement of the genes identified in GWAS (Zanon-Moreno et al., 2013; Athinarayanan et al., 2014; Borel et al., 2007, 2009; Girona et al., 2008) and also suggested that of *CD36* (Lecompte et al., 2011), *CETP* (Borel et al., 2009) and *APOE* (Borel et al., 2007).

Altogether, these association studies show that genetic variants in *CYP4F2*, *SCARB1*, and likely in one of several genes of the *APOA1/C3/A4/A5* cluster can modulate VE status. The associations with *CD36*, *CETP*, and *APOE* remain to be confirmed in other studies.

FUTURE RESEARCH ON GENETIC VARIANTS THAT CAN MODULATE VITAMIN E STATUS

Although associations between SNPs in several genes and VE bioavailability or fasting blood VE concentration have been observed (Fig. 33.1 and Table 33.1), a lot of work remains to be done to identify all the genetic variations involved in VE status and to provide a list of the main genetic variations that govern VE status (Borel and Desmarchelier, 2018). Indeed, GWAS allow us to identify only SNPs that have a relatively high effect on this status and candidate gene studies miss genes not thought to affect this status. Recent results have also shown that it is absolutely necessary to study the additive effects of genetic variants. Indeed, although each variant has generally a very low effect, the additive effect of several variants can have a very significant impact. It should also be noted that SNPs are not the only genetic variations that occur in humans: there are also copy number variants, insertion/deletion variants as well as epigenetic modifications. A genetic score that would aim to predict VE status should therefore take into account all genomic variations that can have a significant impact on VE status. Furthermore, association studies have to be performed in different populations to be sure that the associations are not specific to particular ethnic groups.

In summary, it is now clear that VE status is modulated in part by variants in several genes. Although a lot of work remains to be done to obtain a combination of genetic variations (SNPs as well as other kinds of genetic variations) that will allow us to confidently predict the VE status of an individual by knowing his compound genotype, the potential usefulness of this area of research is exciting with regard to personalized nutrition and for future clinical trials dedicated to assess biological effects of VE. Nevertheless, genetics only represents one variable in the equation, albeit stable over the lifespan, since other factors, such as VE dietary intake, dietary habits, and oxidative stress (through, e.g., smoking), can affect VE status.

Conflicts of Interest

The authors declare no conflict of interest.

References

- Athinarayanan, S., Wei, R., Zhang, M., Bai, S., Traber, M.G., Yates, K., Cummings, O.W., Molleston, J., Liu, W., Chalasani, N., 2014. Genetic polymorphism of cytochrome P450 4F2, vitamin E level and histological response in adults and children with nonalcoholic fatty liver disease who participated in PIVENS and TONIC clinical trials. *PLoS One* 9, e95366.
- Borel, P., Desmarchelier, C., 2018. Bioavailability of fat-soluble vitamins and phytochemicals in humans: effects of genetic variation. *Annu Rev Nutr* 38, 69–96.
- Borel, P., Moussa, M., Reboul, E., Lyan, B., Defoort, C., Vincent-Baudry, S., Maillot, M., Gastaldi, M., Darmon, M., Portugal, H., et al., 2007. Human plasma levels of vitamin E and carotenoids are associated with genetic polymorphisms in genes involved in lipid metabolism. *J Nutr* 137, 2653–2659.
- Borel, P., Moussa, M., Reboul, E., Lyan, B., Defoort, C., Vincent-Baudry, S., Maillot, M., Gastaldi, M., Darmon, M., Portugal, H., et al., 2009. Human fasting plasma concentrations of vitamin E and carotenoids, and their association with genetic variants in apo C-III, cholesteryl ester transfer protein, hepatic lipase, intestinal fatty acid binding protein and microsomal triacylglycerol transfer protein. *Br J Nutr* 101, 680–687.
- Borel, P., Preveraud, D., Desmarchelier, C., 2013. Bioavailability of vitamin E in humans: an update. *Nutr Rev* 71, 319–331.
- Borel, P., Desmarchelier, C., Nowicki, M., Bott, R., Tourniaire, F., 2015. Can genetic variability in alpha-tocopherol bioavailability explain the heterogeneous response to alpha-tocopherol supplements? *Antioxidants Redox Signal* 22 (8), 669–678.
- Desmarchelier, C., Martin, J.C., Planells, R., Gastaldi, M., Nowicki, M., Goncalves, A., Valero, R., Lairon, D., Borel, P., 2014. The postprandial chylomicron triacylglycerol response to dietary fat in healthy male adults is significantly explained by a combination of single nucleotide polymorphisms in genes involved in triacylglycerol metabolism. *J Clin Endocrinol Metab* 99, E484–E488.
- Ferrucci, L., Perry, J.R., Matteini, A., Perola, M., Tanaka, T., Silander, K., Rice, N., Melzer, D., Murray, A., Cluett, C., et al., 2009. Common variation in the beta-carotene 15,15'-monooxygenase 1 gene affects circulating levels of carotenoids: a genome-wide association study. *Am J Hum Genet* 84, 123–133.
- Girona, J., Guardiola, M., Cabre, A., Manzanara, J.M., Heras, M., Ribalta, J., Masana, L., 2008. The apolipoprotein A5 gene -1131T-

- >C polymorphism affects vitamin E plasma concentrations in type 2 diabetic patients. *Clin Chem Lab Med* 46, 453–457.
- Lecompte, S., Szabo de Edelenyi, F., Goumidi, L., Maiani, G., Moschonis, G., Widhalm, K., Molnar, D., Kafatos, A., Spinneker, A., Breidenassel, C., et al., 2011. Polymorphisms in the CD36/FAT gene are associated with plasma vitamin E concentrations in humans. *Am J Clin Nutr* 93, 644–651.
- Major, J.M., Yu, K., Wheeler, W., Zhang, H., Cornelis, M.C., Wright, M.E., Yeager, M., Snyder, K., Weinstein, S.J., Mondul, A., et al., 2011. Genome-wide association study identifies common variants associated with circulating vitamin E levels. *Hum Mol Genet* 20, 3876–3883.
- Major, J.M., Yu, K., Chung, C.C., Weinstein, S.J., Yeager, M., Wheeler, W., Snyder, K., Wright, M.E., Virtamo, J., Chanock, S., et al., 2012. Genome-wide association study identifies three common variants associated with serologic response to vitamin E supplementation in men. *J Nutr* 142, 866–871.
- Olivier, M., Bott, G.R., Frisdal, E., Nowick, M., Plengpanich, W., Desmarchelier, C., Roi, S., Quinn, C.M., Gelissen, I., Jessup, W., et al., 2014. ABCG1 is involved in vitamin E efflux. *Biochim Biophys Acta* 1841, 1741–1751.
- Peter, S., Friedel, A., Roos, F.F., Wyss, A., Eggersdorfer, M., Hoffmann, K., Weber, P., 2016. A systematic review of global alpha-tocopherol status as assessed by nutritional intake levels and blood serum concentrations. *Int J Vitam Nutr Res* 1–21.
- Reboul, E., Borel, P., 2011. Proteins involved in uptake, intracellular transport and basolateral secretion of fat-soluble vitamins and carotenoids by mammalian enterocytes. *Prog Lipid Res* 50, 388–402.
- Wood, A.R., Perry, J.R., Tanaka, T., Hernandez, D.G., Zheng, H.F., Melzer, D., Gibbs, J.R., Nalls, M.A., Weedon, M.N., Spector, T.D., Richards, J.B., Bandinelli, S., Ferrucci, L., Singleton, A.B., Frayling, T.M., 2013. Imputation of variants from the 1000 Genomes Project modestly improves known associations and can identify low-frequency variant-phenotype associations undetected by Hap-Map based imputation. *PLoS One* 8 (5) e64343.
- Zanon-Moreno, V., Asensio-Marquez, E.M., Ciancotti-Oliver, L., Garcia-Medina, J.J., Sanz, P., Ortega-Azorin, C., Pinazo-Duran, M.D., Ordovas, J.M., Corella, D., 2013. Effects of polymorphisms in vitamin E-, vitamin C-, and glutathione peroxidase-related genes on serum biomarkers and associations with glaucoma. *Mol Vis* 19, 231–242.
- Zingg, J.M., Azzi, A., Meydani, M., 2008. Genetic polymorphisms as determinants for disease-preventive effects of vitamin E. *Nutr Rev* 66, 406–414.

This page intentionally left blank

Vitamin K

Jörg Saupe

MVZ Diaverum Erkelenz GmbH, Erkelenz, Germany

Glossary

Gla proteins A small number of functionally important proteins where gamma-glutamyl carboxylase carboxylates glutamyl moieties at specific positions.

Pharmacogenetics The clinical use of genetic information to optimize the selection of medications and the dose of the selected medication.

Undercarboxylated Gla proteins Gla proteins with less than the functionally optimal number of Gla moieties resulting from functional vitamin K deficiency.

Vitamin K-epoxide cycle This cycle refers to the reactivation of oxidized vitamin K-like compounds by VK epoxide reductase and NAD(P)H-dependent quinone oxidoreductases.

contrast to other fat-soluble vitamins, VK is recycled and stored in some tissues. One notable feature of vitamin K metabolism is that the most abundant dietary form, VK1, is poorly retained in the body (Shearer and Newman, 2014).

Function: Humans need VK only as cofactor for a single microsomal enzyme, gamma-carboxyglutamyl carboxylase (GGCX). This single function makes VK among the fat-soluble vitamins. This vitamin K-dependent enzyme converts specific glutamate (Glu) residues by carboxylation into γ -carboxyglutamate (Gla). GGCX itself is a Gla protein. Because Glu is a much weaker calcium chelator than Gla, the vitamin K-dependent step increases the calcium-binding capacity of the modified protein by several orders of magnitude. It is through this mechanism that VKDPs are activated. Complete carboxylation is necessary for full activity of most Gla proteins with known function, and calcium binding is essential for their biological function.

Requirements: Vitamin K requirements in humans are very low because the vitamin is recycled by Vitamin K epoxide reductase (VKOR) with high efficiency. Adequate intake (AI) for VK1 is 120 $\mu\text{g}/\text{day}$ for men and 90 $\mu\text{g}/\text{day}$ for women in the United States. Adult intake recommendations for VK1 in other countries differ over a twofold range. These differences reflect the difficulty to set meaningful guidelines in the absence of suitable endpoints or biomarkers indicating adequacy. The commonly used dosage of VK2 (usually as MK7) in human studies is 180 $\mu\text{g}/\text{day}$.

Deficiency and excessive intake: It is challenging to determine the VK status of a human being. Calculating the VK status from food alone is just one important aspect among others and does not give an answer to the real demand. Of all the fat-soluble vitamins, VK (as VK1) has the largest intra- to inter-individual variation ratios for diet and fasting plasma concentrations, even after accounting for effects of food. An unpredictable intraindividual variability in VK1 bioavailability

INTRODUCTION

Vitamin K (VK) is the umbrella term for a group of fat-soluble substances; among them are phyloquinone (VK1) and the menaquinone (MK) family (VK2). The latter are designated according to the length of its side chain at the naphthoquinone ring starting with four isoprenyl units, numbered MK-4 up to 13 units (MK-13). Biologic activity decreases with length of side chain. VK is important as a cofactor making several proteins active. Among these proteins are clotting factors (II, VII, IX, X) and the clotting inhibiting factor C, the fibrinolytic factor S, and factor Z, an inhibiting factor of the active coagulation factor Xa. During recent decades numerous vitamin K-dependent proteins (VKDPs) have been detected exhibiting very different functions; these are listed in Table 34.1.

Sources: VK1 as the main source of VK is found in green leafy vegetables such as spinach, kale, broccoli, and Brussels sprouts. MKs are found in some cheeses and other fermented foods and Natto, a Japanese soy product cultured with *Bacillus subtilis*. The contribution of MKs from the intestines to VK status is limited and has previously been overestimated. Eggs, pork, fish, and algae also contain very small amounts of MKs. In

TABLE 34.1 Currently confirmed vitamin K-dependent proteins (VKDPs).

Gla-protein name	Gene name	OMIM	Function	Proposed function
Gamma-glutamyl carboxylase	GGCX	137167	Carboxylates specific Glu in a small number of proteins	
Factor II	F2	176930	Supports blood clotting	
Factor VII	F7	613878		
Factor IX	F9	300746		
Factor X	F10			
Protein C	PROC	612283	Inhibits blood clotting	
Protein S	PROS1	176880	1. Fibrinolytic 2. Contributes to cell proliferation	
Protein Z	PROZ	176895	Inhibits activated coagulation factor X	Pro- and anticoagulant properties
Bone gamma-carboxyglutamic acid protein BGP (same as osteocalcin, OC) Undercarboxylated osteocalcin (ucOC)	BGLAP	112260	Regulator of glucose metabolism, insulin production (stimulates), energy expenditure (increases), and insulin sensitivity (increases)	1. γ -carboxylation of OC negatively regulates its endocrine action in humans 2. Energy metabolism-regulating hormone 3. Stimulator of differentiation and mineralization in chondrocytes and vascular smooth muscle cells
Matrix Gla protein (MGP)	MGP	154870	Inhibits calcification of arteries, heart valves, and other tissues	
Gla-rich protein (GRP)				Inhibits calcification of cardiovascular tissues
Growth-arrest specific factor 6 (Gas6)	GAS6	600441	Stimulates proliferation of some cell types	Slows apoptosis
Proline-rich Gla proteins: PRGP1, PRGP2, PRGP3(TM3), PRGP4(TM4)	PRRG1 PRRG2 PRRG3 PRRG4	300935 604429 300685 611690		Expressed in many fetal and adult tissues with uncertain function(s)

has been noted in most absorption studies, suggesting the need to account for nondietary factors. Bioavailability of dietary vitamin K appears to be similar in men and women, whereas age may influence response to intake (Shearer and Newman, 2014). VK1 plasma concentrations are measured very accurately by different HPLC steps. VK1 plasma concentrations in healthy people are around 0.5 nM/L. These concentrations are one to three to four times lower in the blood than for 25-hydroxy-vitamin D, retinol, and tocopherol, respectively. This may explain why a large amount of data in this field is just appearing in recent years with more sophisticated analysis. There still are challenges for methodological laboratory work on MKs because of interfering triglycerides. Consequently, investigations regarding the VK status concentrate on the level of completeness of carboxylation. Undercarboxylated proteins are the hallmark of VK deficiency, and this is called functional VK deficiency. On the other hand,

undercarboxylated forms are found in an increasing number with increasing concentrations of vitamin D in blood. This could be a bias with regard to the estimation of the VK status while measuring concentrations of an undercarboxylated protein. This has been seen with undercarboxylated osteocalcin (ucOC). Consequently, inferences about the presence of an adequate number of sufficient carboxylated proteins are necessarily limited. The currently preferred marker for the assessment of VK status is desphospho-undercarboxylated MGP (dp-ucMGP), which is more sensitive to VK deficiency than the liver-derived coagulation factor PIVKA-II (prothrombin induced in vitamin K absence).

One serious but preventable consequence of VK deficiency is VK-dependent bleeding (VKDB) in exclusively breast-fed newborns, which typically occurs 4 weeks after birth or later and occasionally results in death or permanent brain damage. VK1 contents of milk formulas are typically 50-fold higher than human milk and

therefore prevent most cases (Shearer and Newman, 2014). Functional VK deficiency effects regarding VKDP other than clotting factors are very likely and will be discussed later.

Excessive VK intake does not appear to cause harm, contrary to the unfounded assumption that all fat-soluble vitamins accumulate with high intake and cause hypervitaminosis.

OVERVIEW OF VITAMIN K METABOLISM

The VK-epoxide (see Fig. 34.1) cycle is essential to the function of VK. The cycle involves both the glutamyl gamma-carboxylation step of VK-dependent proteins and the recovery of the cofactor VK quinol that is oxidized to VK epoxide metabolite as a consequence of gamma-carboxylation by GGCX. Reactivation depends on VKOR and various NADH or NADPH-dependent quinone oxidoreductases (see Fig. 34.1) (Shearer and Newman, 2014). Infants born with a defective VKOR often present with severe coagulopathy and/or skeletal defects.

POLYMORPHISMS AFFECTING THE VITAMIN K-EPOXIDE CYCLE AND PHARMACOGENETIC ASPECTS

For therapeutic and prophylactic purposes warfarin (Coumadin) is used to prevent or treat thrombosis and embolism. The effect occurs with a period of latency after 36–72 h. Overdose may lead to harmful side effects such as fatal bleeding. Underdose will fail therapeutic

targets. Clinicians have to choose the doses based on their personal experience since guidelines are lacking. However, it would be highly welcome to know the exact dose at the beginning of treatment. At the intersection of pharmacogenetics and nutrigenetics bioavailability and effectiveness is influenced by both nutrients as VK and drugs as warfarin. The therapeutic target for warfarin is the VKOR. The genetic differences in interindividual response to warfarin dosing are mostly explained by a small number of common polymorphisms.

Several candidate gene studies on polymorphism involved in the VK cycle as meta-analyses of large-scale genome-wide association studies investigating the association of genome-wide SNPs with VK status have been carried out. Resistance and sensitivity to warfarin (Coumadin) treatment can be influenced by variations in several genes, including CYP2A6 (OMIM 122720), VKORC1 (OMIM 608547), CYP2C9 (OMIM 601130), and CYP4F2 (OMIM 604426). In a genome-wide analysis of Caucasians taking warfarin, most significant independent effect of variation in warfarin maintenance dose was conferred by variants in the VKORC1 gene ($P = 6.2 \times 10^{-13}$), and weaker associations with CYP2C9 polymorphisms ($P < 10^{-4}$). These associations were replicated in two separate populations, yielding combined P values of 4.7×10^{-34} and 6.2×10^{-12} for VKORC1 and CYP2C9, respectively. The contribution to the variance in response to warfarin explained by the two genes was estimated to be 25% and 9%, respectively. No significant associations with other polymorphisms were identified. Major contributions to warfarin response by other variants seem unlikely.

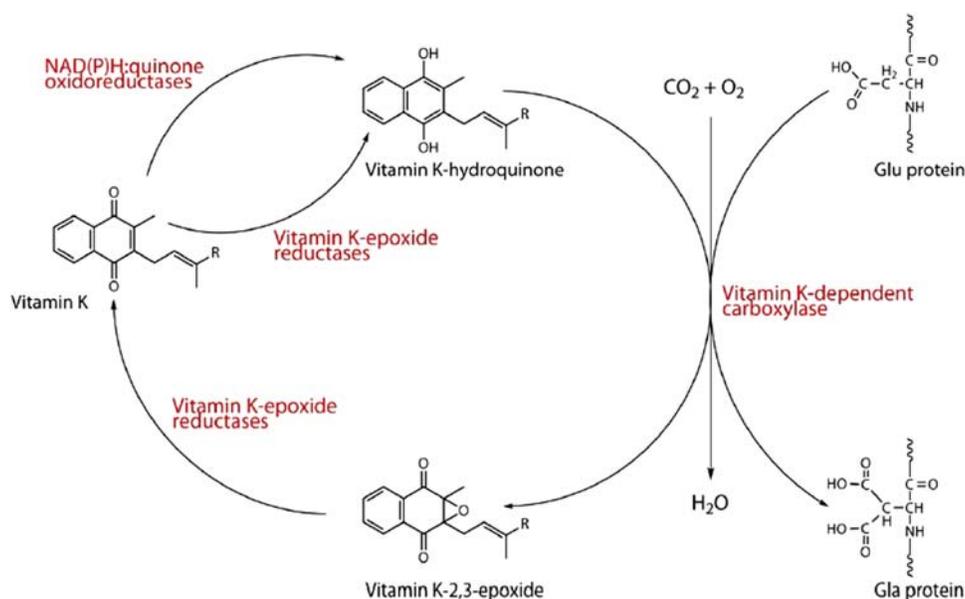


FIGURE 34.1 The metabolically active form, vitamin K-hydroquinone, is oxidized when it acts as a cofactor for Glu carboxylation and the resulting vitamin K-2,3-epoxide can then be reactivated in two successive steps. With permission from: Martin Kohlmeier, *Nutrient metabolism, structures, functions, and genes*, 2nd edition, Academic Press, 2015.

Another study including African Americans and European Americans undergoing warfarin therapy variation in the VKORC1 gene could explain 5% and 18%, respectively, of variability in warfarin dosage. An additive effect was observed when also accounting for polymorphisms in the CYP2C9 gene (8% and 30%, in African Americans and European Americans, respectively). Four common VKORC1 haplotypes were identified in European Americans, and 12 in African Americans, consistent with higher genomic sequence diversity in populations of African descent. African Americans had a lower frequency of the low-dose haplotype compared with European Americans (10.6% vs. 35%, $P < .0001$). The variability in dose explained by VKORC1 haplotype or haplotype groups was similar to that of a single informative polymorphism. Two SNPs in the VKORC1 gene (rs9934438 or rs9923231) were best predictors of warfarin dose in both groups (Johnson et al., 2017).

In the Supplementary Appendix (Johnson et al., 2017; International Warfarin Pharmacogenetics Consortium et al., 2009) of the 3rd International Warfarin Pharmacogenetics Consortium, an algorithm is deposited as an MS Excel file that takes account of the common variant ($-1639G > A$; rs9923231) in the promoter of the VKORC1 gene.

A 65-year-old obese (178 cm; 98 kg) Caucasian male would probably require weekly doses of 47 mg with the GG genotype G/G, 36 mg with the AG genotype, and 26 mg with the AA genotype. Future studies will need to establish algorithms for individuals with different ethnic backgrounds.

Variants in other genes to be taken into account include those in CYP2C9, CYP4F2, and rs12777823, and possibly apolipoprotein E (rs7412 and rs429358) (International Warfarin Pharmacogenetics Consortium et al., 2009).

CYP2C9 (rs1057910) is the cytochrome P450 enzyme responsible for the metabolism of the isomer of warfarin that is principally responsible for the anticoagulant effect of the drug. Two CYP2C9 alleles that produce a phenotype of poor metabolism occur in 11% and 8% of whites but only 3% and 0.8% of blacks. Such persons have impaired metabolism of warfarin and thus increased plasma concentrations of the drug. Persons with the genotype of impaired metabolism require lower doses of warfarin to achieve an anticoagulant effect similar to that in patients with the normal genotype and are more likely to have an excessive anticoagulant response. In addition, bleeding episodes tend to be more common in persons with the genotype of impaired metabolism. Genotype-guided vitamin K antagonist dosing can improve the time in therapeutic range (TTR) and reduce the risk for bleeding episodes, in comparison with standard dosing algorithms (Kheiri et al., 2018).

The apolipoprotein E (APOE; rs7412 and rs429358) allele and genotype frequencies in the northern Han Chinese population appear to differ from racial groups or populations living in other regions of China. The APOE E2 variant was associated with a significantly higher warfarin maintenance dose.

The GGCX enzyme is encoded by the GGCX gene, located on the reverse strand of chromosome 2p11.2. Only two variants, i.e., rs699664 and rs6173310, of the 409 exonic variants (missense and loss-of-function) being identified in the GGCX gene, have an allele frequency of >0.0001 . This enzyme cannot be influenced by warfarin. GGCX polymorphism does play a role in the ratio of ucOC to intact OC as shown below.

Polymorphisms in genes involved in the VK-dependent gamma-carboxylation reaction influence interindividual variation in the activities of protein C and protein S in the general population. The mean activity of protein C in women bearing the GG genotype of GGCX 8016G $>$ A ($130.8\% \pm 1.5\%$, $n = 156$) is significantly greater ($P = 0.002$) than that of individuals with either the AG ($126.8\% \pm 0.7\%$, $n = 728$) or the AA ($125.4\% \pm 0.6\%$, $n = 881$) genotype, after adjusting for confounding factors. The GGCX 8016G $>$ A change leads to the substitution of Gln for Arg at amino acid residue 325 (Arg325Gln). This effect is comparable to that of a previously defined polymorphism in the protein C promoter. Mean protein S activity was influenced by the VKORC1 3730G $>$ A and CALU 20,943T $>$ A genotypes, after adjusting for confounding factors (Johnson et al., 2017; International Warfarin Pharmacogenetics Consortium et al., 2009).

POLYMORPHISMS AFFECTING VITAMIN K CATABOLISM INCLUDING PHARMACOGENETIC ASPECTS

Another enzyme appears to have a role in regulating the rate of catabolism of VK. Searches for genetic polymorphisms in the cytochrome P450 4F2 gene (CYP4F2) that might influence warfarin dosage led to the discovery of a common DNA missense variant (rs2108622 C $>$ T, V433M) with distinct impact on warfarin effectiveness. Individuals with two TT alleles required about 1 mg/d more warfarin than patients with two CC alleles. Recombinant CYP4F2 is known to remove the isoprenyl side chain of VK1 and thereby converts it to an oxidized metabolite. The significance of the rs2108622 polymorphism was highlighted by the fact that liver microsomes from rs2108622 T allele carriers had lower CYP4F2 protein concentration and oxidized VK1 less rapidly. Carriers of the CYP4F2 V433M (rs2108622) allele require higher doses of warfarin to achieve adequate antithrombotic protection. One

explanation might be their lower capacity to metabolize VK, leading to presumed higher hepatic concentrations. One possible demographic impact of this polymorphism to VK requirements is that the minor allele frequency for CYP4F2 is around 30% in Caucasians and Asians, but only 7% in people of African ancestry (Shearer and Newman, 2014).

Control of Calcification

Vascular calcification and osteoporosis are great challenges for the health care system. Among the various compounds with vitamin K-like activity, VK2 has been shown to act on both systems.

Vascular disease: The presence of regulatory proteins with opposing functions is a general feature of calcified matrix formation. Along with excess intracellular calcium, matrix vesicles originating from vascular smooth muscle cells (VSMCs) and chondrocytes are known to contain the mineralization inhibitor MGP and mineralization inducer of alkaline phosphatase. It has been suggested that undercarboxylated form of MGP leads to accelerated vascular matrix calcification. An additional Gla protein, Gas6, and its receptor Axl, were also shown to play a pivotal role in mediating antiapoptotic and antiminerallizing effects of statins in VSMCs (Turgut Cosan et al., 2017). Polymorphic variants of MGP have been investigated in more detail as they are part of the matrix Gla-protein system, which includes MGP, vitamin D receptor (VDR), VKOR, GGCX, and bone morphogenic protein-2 (BMP-2), which is an important factor in vessels protection of ectopic calcification. Polymorphisms of genes, which encode the structure of these proteins, determine their activity and may affect the intensity of calcification and the consequences of acute coronary syndrome (ACS). An association between ACS and polymorphic variants of matrix Gla-protein system genes: MGP (rs1800802, rs1800801, rs4236), VDR (rs2228570, rs1544410, rs7975232, rs731236), GGCX (rs699664), VKORC1 (rs2359612), BMP-2 (rs2273073), was analyzed. There is an association between ACS and some polymorphic variants of matrix Gla-protein system genes: MGP (rs1800801), VDR (rs1544410), GGCX (rs699664), VKORC1 (rs2359612). This indicates a higher risk of complications in the ACS patients with the following genotypes: AA (rs1800801), GG (rs1544410), AA (rs699664), and CC (rs2359612) (Garbuzova et al., 2015).

These pathophysiologic mechanisms are involved in the onset of essential hypertension (EH), identifying people at special risk by diagnosing VKCOR1 gene polymorphism. Vascular calcification increases peripheral arterial resistance a distinguished characteristic of essential EH and results in aortic stiffness and elevation of blood pressure. Regulation of vascular calcification is

physiologically limited by gamma-carboxylated proteins that regulate mineralization. Any deficiencies related to mineralization influence vascular calcification. As a result of VK deficiency or any problem associated with the VK epoxide reductase complex subunit 1 (VKORC1) gene, Glu is not sufficiently transformed to Gla and calcification occurs in the walls of blood vessels, in the myocardium, and in cardiac valves. 3673G/A (rs9923231) and 9041G/A (rs7294) polymorphisms in the VKORC1 gene were determined in hypertensive patients. A significant difference was found between the rs7294 polymorphisms ratios of the case and control groups, but significant differences weren't found in distribution of the rs9923231 alleles. The rs7294 GG genotype is associated with 3.97-fold increased risk for EH compared to the AA genotype. It is reasonable, therefore, to conclude that the VKORC1 rs7294 polymorphism influences the risk of developing EH (Turgut Cosan et al., 2017).

OC may have a role in hypertension due to its close relationship with energy metabolism and insulin sensitivity. The carriers of C allele of rs1800247 were associated with a decreased odds of hypertension compared with the carriers of T allele adjusted for age, sex, and BMI (odds ratio = 0.89, $P = .01$). The carriers of the C allele were also associated with lower diastolic blood pressure in normotensive individuals ($\beta = -0.55$, $P = .02$), but its association with systolic blood pressure levels was statistically not significant. The interaction between rs1800247 and homeostasis model assessment for insulin resistance (HOMA-IR) on hypertension was significant ($P < .001$). In the stratified analysis by the median (1.93) of HOMA-IR, rs1800247 the carriers of the C allele were significantly associated with hypertension in the subgroup with $HOMA-IR \leq 1.93$ (OR = 0.82, $P = .005$), but this association was not significant in the subgroup with $HOMA-IR > 1.93$ (Viegas et al., 2015).

Hemodialysis patients (HD) have markedly functional VK deficiency and increased cardiovascular mortality associated with accelerated vascular calcification (VC). The VKORC1 polymorphism rs9923231 was associated with cardiovascular calcification and clinically overt cardiovascular disease, C1173T polymorphisms with higher risk for disease, and G-1639A with lower risk. A significant association was found between VKORC1 gene polymorphism C1173T (both homozygous [TT] and heterozygous [CT] in comparison to the homozygous wild-type [CC]) and vascular calcification. Moreover, the T allele was also linked with higher vascular calcification risk. It is possible that the uremic environment, present in patients with the chronic kidney disease (CKD), may modify the expression of the VKOR gene (Osman et al., 2016).

More importantly, levels of inactive MGP can be reduced by daily VK2 supplement use. Ongoing end

point studies investigate whether VK1 supplementation affects the progression of coronary and aortic calcification measured by coronary artery calcium score in HD patients. Results should be available within a few years.

Meta-analyses of studies investigating the role of food intake, supplement use, and biomarkers of intake all point to beneficial effects of vitamin K2 intake on chronic heart disease (Juanola-Falgarona et al., 2013).

Gla-rich protein (GRP) is a more recently described VKDP; it appears to affect mineralization activity of extracellular vesicles. GRP activity was found to be dependent on its gamma-carboxylation status, with potential biological relevance (Viegas et al., 2015).

Bone: VK2 stimulates bone formation by promoting osteoblast differentiation and carboxylation of osteocalcin (OC). This action increases the concentrations of alkaline phosphatase, insulin-like growth factor 1, growth-differentiation factor 15, and stanniocalcin 2. On the other hand, VK2 inhibits the proapoptotic proteins Fas and Bax (BCL-2-associated X protein) in osteoblasts, and decreases osteoclast differentiation. Osteoblast-derived OC and bone-specific alkaline phosphatase are products released to the blood during bone formation. The physiologic role of OC in mineralization is highly disputed. A significant association has been reported with VK intake and γ -carboxylation regarding OC. Interestingly, on grouping by the GGCX genotype, (R325Q, 974G > A, dbSNP: rs699664) there was a significant interaction between the ratio of ucOC to intact OC with VK intake in homozygotes (GG-type) and heterozygotes (GA-type). These results suggest that an adequate nutritional strategy is necessary for people with high-risk genotypes (GG or GA-type) (Haraikawa et al., 2013). Identifying the GG or GA type it was thought to be helpful in preventing osteoporosis. As the physiologic relevance of OC may change, genetics still are as important in identifying risk groups for disturbances in glucose metabolism. These data do not exclude a prominent role for VK in bone metabolism.

A common polymorphism rs1800247 in OC gene may affect the risk of osteoporosis and fracture and serum total OC levels in Chinese. Carriers of CC and TC genotypes are associated with a lower odds of fracture compared with the carriers of TT genotype (odds ratio = 0.60, 95%CI 0.40–0.88, $P = .01$) and rs1800247 is associated with serum total OC levels in all models adjusted for age, BMI, menopausal status, and total hip BMD in women (all $P < .001$), with OC levels decreasing across TT, TC, and CC genotypes. There may be gender differences underlying these associations (Ling et al., 2016).

VKORC1 rs8050894 GG homozygotes have significantly higher cross-sectional measures of plasma VK1 than carriers of the CG or CC genotypes (plasma VK1 geometric means: GG 0.874 ± 0.092 vs. CG/CC

0.598 ± 0.044 ; $P = .020$), whereas carriers of VKORC1 rs7294 AA or AG had significantly lower plasma VK1 concentrations compared to GG homozygotes (plasma VK1 geometric means: 0.579 ± 0.045 vs. 0.762 ± 0.057 ; $P = .035$). Cross-sectional analyses also revealed that heterozygous carriers of GGCX rs10187424 and rs7568458 had significantly lower %ucOC relative to either homozygous group. Polymorphisms in genes encoding enzymes involved in VK metabolism may modulate plasma concentrations of VK1 and percent carboxylation of OC (Haraikawa et al., 2013).

In hemodialysis patients, fasting VK1 concentrations in plasma are related to alleles of apolipoprotein E genotype (rs7412 and rs429358) in the order E2 > E3 > E4 (Shearer and Newman, 2014). A greater percentage of patients with the E3/4 and E4/4 genotypes than those with the E2/3 and E2/2 genotypes had a history of bone fractures at the time of recruitment (44% vs. 16%, $P < .005$). The original postulate was that if the APOE4 allele promoted faster hepatic uptake of triglyceride rich lipoproteins (TRL)-VK, there might be a sparing effect on the delivery and therefore availability of VK to bone, with opposite effects for the APOE2 allele. When apoB-48 and apoB-100 were employed as more specific markers of intestinally and hepatically derived TRL, respectively, it was found that rather than enhancing the clearance of chylomicrons, chylomicron remnants, and very-low-density lipoprotein, the APOE4 allele impaired their clearance. These conflicting data may be relevant for the time of blood collection in association studies correlating concentrations with various biomarkers. However, the need for higher warfarin doses by E2 carriers as shown above coincides with their higher VK1 concentrations.

Changes due to small intestinal bacterial overgrowth (SIBO) lead to reduced matrix Gla-protein activation as well as arterial hardening. Both of these observations are important indicators of early atherosclerosis. Hence, screening for SIBO, intestinal decontamination and supplementation with vitamin K2 has the potential to be incorporated into clinical practice as additional preventive measures (Ponziani et al., 2017).

Glucose metabolism: Epidemiological longitudinal studies have associated VK intake with insulin sensitivity, glucose metabolism and thus with diabetes. Those subjects with the highest intakes had lower ghrelin, glucose-dependent insulinotropic peptide, glucagon-like peptide-1, IL-6, leptin, TNF, and visfatin plasma concentrations. Additionally, a lower risk for diabetes mellitus was observed with increased VK intakes (Hsieh et al., 2015). Recent data on OC endocrine regulation are consistent with a mechanism whereby OC from osteoblasts is stored as γ -carboxylated, inactive protein in the bone extracellular matrix. This stored OC can be activated by decarboxylation during bone resorption.

Nevertheless, gain- and loss-of-function mouse models of OC have shown that this protein is not required for normal bone remodeling or mineralization. Cell culture and in vivo studies have shown that OC improves glucose handling by promoting insulin secretion by β -cells on one hand and by favoring insulin sensitivity on the other hand. At least in mice, Glu-OC is the active form of this hormone in vivo and γ -carboxylation negatively regulates both the bioactivity and the bioavailability of osteocalcin. In addition, a novel role for VK-dependent γ -carboxylation in the control of glucose metabolism was shown (Lacombe and Ferron, 2015). The beneficial role of VK in type 2 diabetes prevention observed could be apparently contradictory with the association between uc-OC and improvements of glucose metabolism. An explanation for this apparent contradiction might be that total OC is highly correlated with uc-OC, and therefore, the uc-OC tracks the total OC or vice versa.

Obesity: Gas6 plays a role in hemostasis, atherosclerosis, and the thrombotic process. Gas6 has been shown to be involved in regulating multiple cellular functions through binding to the Tyro-3, Axl, and Mer (TAM) family of receptor tyrosine kinases. Changes in plasma Gas6 levels have been associated with clinical disorders characterized by proinflammatory and prothrombotic events. Recent studies in animal models have shown that Gas6/TAM signaling plays an important role in pathophysiological mechanisms underlying obesity-related inflammation and insulin resistance (IR). In a cohort of Asian hypertensive patients with insulin-resistance plasma, Gas6 levels are associated with glucose intolerance, markers of inflammation, and endothelial dysfunction, both in adults and adolescents. These findings support a likely role for Gas6 in the pathogenesis of obesity, IR, and related complications. Three SNPs (rs8191973, rs8197974, and rs7323932) were found to be associated with steady-state plasma glucose

TABLE 34.2 Predicted genotype effects on VKDP and VK1 concentrations or clinical outcomes (HOMA-IR, homeostasis model assessment for insulin resistance).

VKORC1	rs9923231	3673G/A C1173>G-1639A C1173: TT = CT > CC G/G>G/A>A/A	Influences risk of essential hypertension higher risk of vascular calcification Higher risk of vascular calcification Required warfarin is genotype-dependent
	rs9934438	1173C > T	Required warfarin is genotype-dependent
	rs2359612	C allele	higher risk of coronary artery calcification
	rs7294	9041G/G>G/A>A/A	Risk of essential hypertension depends on genotype
CYP2C9	rs9332238	G/G>G/A>A/A	Required warfarin is genotype-dependent
CYP2C9*2	rs1799853		Required warfarin is genotype-dependent
CYP2C9*3	rs1057910	A/A>A/C>C/C	Required warfarin is genotype-dependent
CYP2C19	rs3814637	C/C>C/T>T/T	Required warfarin is genotype-dependent
CYP4F2	rs2108622	V433M T > C	Required warfarin is genotype-dependent
GGCX	rs699664	A/A>AG>GG GG = GA > AA	Required warfarin is genotype-dependent ucOC/intact OC
MGP	rs1800801	A/A>AG>GG	Acute coronary symptoms are genotype-dependent
	rs4236	C/C>TC>TT	Acute coronary symptoms are genotype-dependent
Gas6	rs8191974 rs7323932 rs7331124 rs8191973	AACG AACG AACG AACG	Effect is statistically significant after adjustment for multiple comparisons
OC	rs1800247	C > T C > T C > T TT < TC < CC	Higher risk of essential hypertension Lower diastolic blood pressure HOMA-IR<1.93 Differential fracture risk in women
Apolipoprotein E	rs7412 plus rs429358	E2>E3>E4	VK1 concentration in blood

(SSPG) levels ($P = .007$, $P = .03$, and $P = .011$, respectively). This association remained significant after multiple testing and showed a significant interaction with physical activity for SNP rs8191973. However, no other significant correlations were observed between Gas6 polymorphisms and other indices of insulin resistance or obesity. A specific haplotype, AACG (from rs8191974, rs7323932, rs7331124, and rs8191973), was associated with SSPG levels ($P < .01$). None of the polymorphisms were associated with an increased risk of T2D development (Hsieh et al., 2015).

THE SIGNIFICANCE OF NUTRIGENETICS IN THE VITAMIN K SYSTEM

VKDP contributes to the regulation of several biological systems and genetic variants influence important health outcomes. In the blood-coagulation system both activating and inhibiting VKDPs contribute to this tightly controlled balance. Another mechanism involving VKDP helps to control tissue calcification. VKDPs also affect glucose homeostasis, tumor proliferation, and autoimmune processes. Well-intentioned interventions by warfarin risks unintended harm including ectopic calcification and worsening of autoimmune disease. On the other hand, excessive VK may be harmful in prooncogenic states. The assignment of nutrigenetics rests in identifying people at risk for certain diseases in preventing them from overdose of VK and VK deficiency, respectively. Therefore, targets of clinical investigation should be in identifying polymorphisms of VCOR1 and GGCX in the VK- epoxide cycle and CYP4F2 in the catabolic pathway of VK in a certain person. In disease states like chronic kidney disease, endogenous disturbances may aggravate imbalances in the VK system and this demands special attention (Table 34.2).

References

Garbuzova, V.Y., Stroy, D.A., Dosenko, V.E., Obukhova, O.A., Ataman, O.V., 2015. Association of allelic polymorphisms of the Matrix Gla-protein system genes with acute coronary syndrome in the Ukrainian population. *Biopolym Cell* 31 (1), 46–56.

Haraikawa, M., Tsugawa, N., Sogabe, N., Tanabe, R., Kawamura, Y., Okano, T., Hosoi, T., Goseki-Sone, M., 2013. Effects of gamma-glutamyl carboxylase gene polymorphism (R325Q) on the association between dietary vitamin K intake and gamma-carboxylation of osteocalcin in young adults. *Asia Pac J Clin Nutr* 22 (4), 646–654.

Hsieh, C.H., Chung, R.H., Lee, W.J., Lin, M.W., Chuang, L.M., Quertermous, T., Assimes, T., Hung, Y.J., Yu, Y.W., 2015. Effect of common genetic variants of growth arrest-specific 6 gene on insulin resistance, obesity and type 2 diabetes in an Asian population. *PLoS One* 10 (8) e0135681.10.

International Warfarin Pharmacogenetics Consortium, Klein, T.E., Altman, R.B., Eriksson, N., Gage, B.F., Kimmel, S.E., Lee, M.T., Limdi, N.A., Page, D., Roden, D.M., Wagner, M.J., Caldwell, M.D., Johnson, J.A., 2009. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N Engl J Med* 360 (8), 753–764.

Johnson, J.A., Caudle, K.E., Gong, L., Whirl-Carrillo, M., Stein, C.M., Scott, S.A., Lee, M.T., Gage, B.F., Kimmel, S.E., Perera, M.A., Anderson, J.L., Pirmohamed, M., Klein, T.E., Limdi, N.A., Cavallari, L.H., Wadelius, M., 2017. Clinical pharmacogenetics implementation consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 update. *Clin Pharmacol Ther* 102 (3), 397–404.

Juanola-Falgarona, M., Salas-Salvad, A.J., Estruch, R., Portillo, M.P., Casas, R., Miranda, J., Martinez-Gonzalez, M.A., Bull, A.M., 2013. Association between dietary phyloquinone intake and peripheral metabolic risk markers related to insulin resistance and diabetes in elderly subjects at high cardiovascular risk. *Cardiovasc Diabetol* 12, 7.

Kheiri, B., Abdalla, A., Haykal, T., Osmo, M., Ahmed, S., Hassan, M., Bachuwa, G., 2018. Meta-analysis of genotype-guided versus standard dosing of vitamin K antagonists. *Am J Cardiol* 121 (7), 879–887.

Lacombe, J., Ferron, M., 2015. Gamma-carboxylation regulates osteocalcin function. *Oncotarget* 6 (24), 19924–19925.

Ling, Y., Gao, X., Lin, H., Ma, H., Pan, B., Gao, J., 2016. A common polymorphism rs1800247 in osteocalcin gene is associated with hypertension and diastolic blood pressure levels: the Shanghai Changfeng study. *J Hum Hypertens* 30 (11), 679–684.

Osman, N.A., El-Abd, N., Nasrallah, M., 2016. VKORC1 gene (vitamin K epoxide reductase) polymorphisms are associated with cardiovascular disease in chronic kidney disease patients on hemodialysis. *Saudi J Kidney Dis Transpl* 27 (5), 908–915.

Ponziani, F.R., Pompili, M., Di Stasio, E., Zocco, M.A., Gasbarrini, A., Flore, R., 2017. Subclinical atherosclerosis is linked to small intestinal bacterial overgrowth via vitamin K2-dependent mechanisms. *World J Gastroenterol* 23 (7), 1241–1249.

Shearer, M.J., Newman, P., 2014. Recent trends in the metabolism and cell biology of vitamin K with special reference to vitamin K cycling and MK-4 biosynthesis. *J Lipid Res* 55 (3), 345–362.

Turgut Cosan, D., Yazıcı, H.U., Colak, E., Soyocak, A., Degirmenci, I., Kurt, H., Birdane, A., Colak, E., Gunes, H.V., 2017. Susceptiveness of vitamin K epoxide reductase complex subunit 1 gene polymorphism in essential hypertension. *Genet Test Mol Biomark* 21 (5), 292–297.

Viegas, C.S., Rafael, M.S., Enriquez, J.L., Teixeira, A., Vitorino, R., Luís, I.M., Costa, R.M., Santos, S., Cavaco, S., Neves, J., Macedo, A.L., Willems, B.A., Vermeer, C., Simes, D.C., 2015. Gla-rich protein acts as a calcification inhibitor in the human cardiovascular system. *Arterioscler Thromb Vasc Biol* 35 (2), 399–408.

Vitamins as Cofactors for Energy Homeostasis and Their Genomic Control, With Special Reference to Biotin, Thiamine, and Pantothenic Acid

Antonio Velazquez-Arellano, Alain de J. Hernandez-Vazquez

Genetics of Nutrition Unit, Institute of Biomedical Investigations, Universidad Nacional Autónoma de México, National Institute of Pediatrics, Ciudad de Mexico, Mexico

List of Acronyms and Abbreviations

- AMPK** Energy sensor, 5' AMP-activated protein kinase or 5' adenosine monophosphate-activated protein kinase
- ATP** Adenosine triphosphate
- CREB** cAMP response element-binding is a cellular transcription factor
- FoxO1** Forkhead box protein O1, is a transcription factor that plays important roles in regulation of gluconeogenesis and glycogenolysis by insulin signaling
- GLUT4** A transporter of glucose in, among other tissues, muscle and adipose
- mTOR** Mammalian target of rapamycin
- PEPCK** Phosphoenolpyruvate carboxykinase is an enzyme in the lyase family used in the metabolic pathway of gluconeogenesis. It converts oxaloacetate into phosphoenolpyruvate and carbon dioxide.
- PDH** Pyruvate dehydrogenase is the first component enzyme of pyruvate dehydrogenase complex.
- PGC-1 α** Peroxisome proliferator-activated receptor gamma coactivator 1-alpha is a transcriptional coactivator that regulates the genes involved in energy metabolism. It is the master regulator of mitochondrial biogenesis.
- SLC19A2** Gene provides instructions for making a protein called thiamine transporter 1. This protein is located on the surface of cells.
- SLC19A3** Gene encodes a ubiquitously expressed transmembrane thiamine transporter that lacks folate transport activity. Mutations in this gene cause biotin-responsive basal ganglia disease.
- SMVT** Sodium-dependent multivitamin transporter
- TCA** Tricarboxylic acid cycle

INTRODUCTION

Thermodynamically, the energy that cells can and must use is free energy. Cells obtain free energy in a chemical form by the catabolism of nutrient molecules, mainly glucose and fatty acids. Cells use that energy to make adenosine triphosphate (ATP) from adenosine diphosphate and phosphate. ATP then donates some of its chemical energy to endergonic processes, such as the synthesis of metabolic intermediates and macromolecules from smaller precursors, the transport of substances across membranes against concentration gradients, and mechanical motion.

Under aerobic conditions, ATP generation is the result of very complex processes, which encompass oxidative phosphorylation, the tricarboxylic acid (TCA) cycle, and the mitochondrial electron transfer chain (Fig. 35.1).

The process of ATP generation has thus multiple components, among which there are indispensable cofactors derived from vitamins, mainly of group B. This chapter will refer to vitamins **biotin** used in carboxylation reactions, **thiamine** in oxidative dehydrogenations, and **pantothenic Acid**, which is the precursor of coenzyme A.

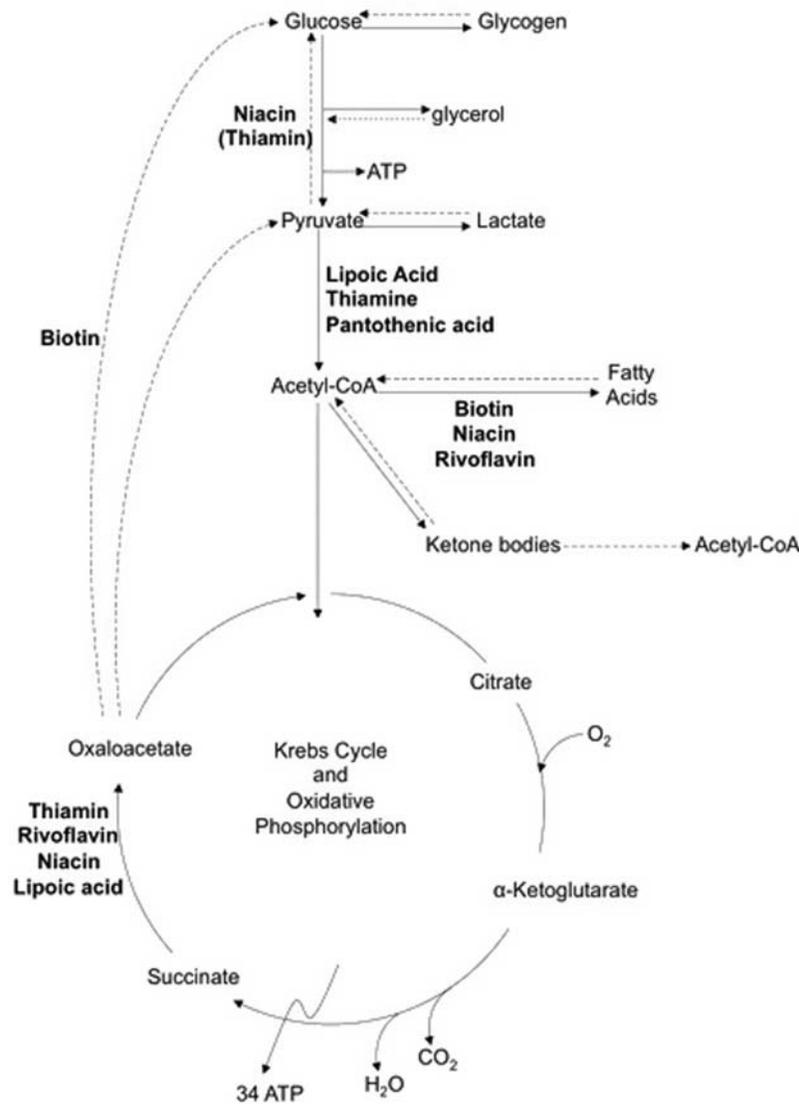


FIGURE 35.1 Cofactors participating in the metabolism of carbohydrates and fatty acids.

Biotin

Biotin is the prosthetic group of five carboxylases that catalyze critical steps in carbohydrate, lipid, and protein metabolism. It is covalently bound to a lysine residue of the apocarboxylases. Pyruvate carboxylase is the main anaplerotic enzyme of the TCA cycle; its product is oxaloacetate, which is also a precursor of glucose in gluconeogenesis, and of fatty acids in lipogenesis. Acetyl CoA carboxylase has two isoforms, I and II, that produce malonyl CoA from acetyl CoA. One is cytosolic and catalyzes the rate-limiting step in fatty acid synthesis; the other is located on the outer mitochondrial membrane and controls fatty acid oxidation through inhibition of carnitine palmitoyltransferase I by malonyl CoA (Fig. 35.2). Propionyl CoA carboxylase participates in the catabolism of several amino acids (isoleucine, valine, methionine, threonine) and in odd-chain fatty

acid oxidation. 3-methylcrotonyl CoA carboxylase catalyzes a step in leucine catabolism. Endogenous biotin is recycled through the biotin cycle. When holocarboxylases are degraded, biotin bound to lysine (biocytin) or lysine residue of peptides is liberated by the action of biotinidase, and most of the vitamin is reused by holocarboxylase synthetase (Fig. 35.3).

Genetic Disorders of Biotin-Dependent Enzymes

The primary transporter for biotin in the intestine and into hepatocytes, and probably other cell types, is coded by the sodium-dependent multivitamin transporter gene, which also transports pantothenic acid (see below). There is a genetic defect of this gene, which courses with acute encephalopathy. Symptoms improve rapidly following biotin supplementation, while the vitamin withdrawal causes recurrent abnormal organic

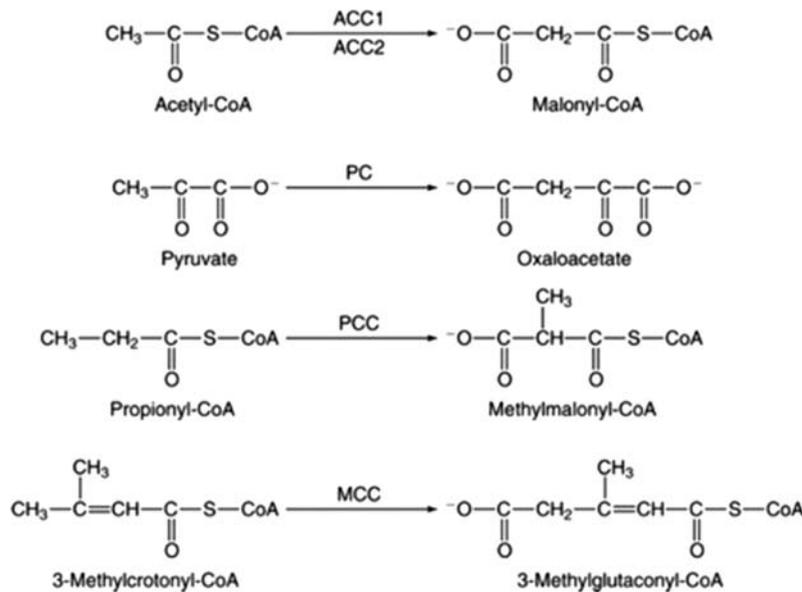


FIGURE 35.2 Biotin-dependent carboxylase reactions.

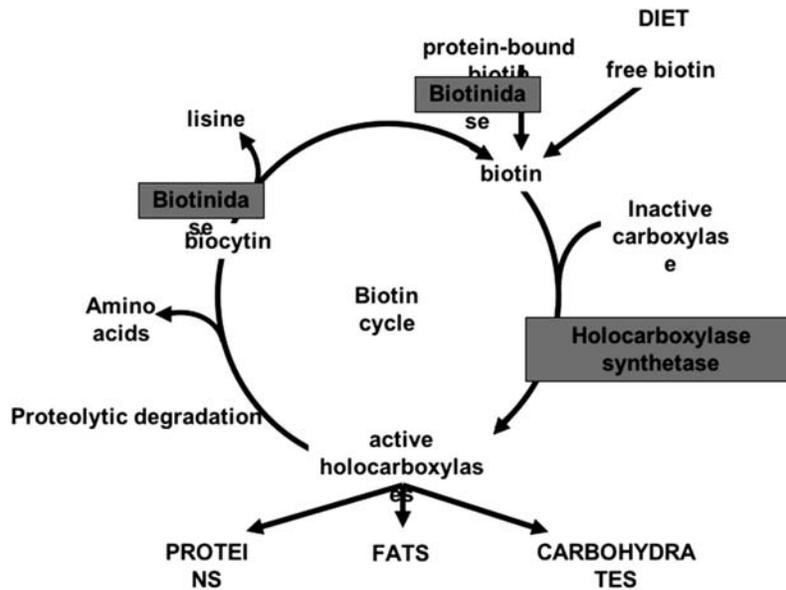


FIGURE 35.3 Biotin recycling.

aciduria, indicating intracellular biotin deficiency. Genetic deficiencies of holocarboxylase synthetase and biotinidase produce early or late onset forms of multiple carboxylase deficiencies. The deficiency of holocarboxylase synthetase generally is the result of mutations with a very high K_m for biotin, and usually leads to a severe neonatal form, frequently with congenital defects, and it is one of the causing defects of Leigh's syndrome. This rare, heterogeneous, neurodegenerative disorder of early childhood is characterized by focal, symmetric necrotic brain lesions, with death occurring generally before the age of five.

The deficiency of biotinidase has a variable clinical phenotype, depending on the mutations of the biotinidase gene. Manifestations include severe lactic acidemia, ataxia, seizures, alopecia, and dermatitis, and—if not treated—may cause death. Treatment requires very high (“pharmacologic”) doses of the vitamin.

Nutrigenomics of Biotin Deficiency

Biotin deficiency diminishes liver glucokinase and elevates phosphoenolpyruvate carboxykinase (PEPCK) mRNAs. Glucokinase reduction can be reversed not only by biotin supplementation but also by insulin and

cyclic GMP. There are reports of effects of biotin on several transcription factors: NF- κ B and Sp1 and Sp3, as well as decreased transcription of the sarco/endoplasmic reticulum ATPase 33 gene. Biotin seems to be necessary for the normal progression of cells through the cell cycle, with deficient cells arresting in the G1 phase.

Velazquez-Arellano and coworkers discovered a new function of biotin as regulator of the genetic expression of carboxylases and of holocarboxylase synthetase, the enzyme that covalently binds biotin to the apocarboxylases. More recently, they found that biotin deprivation has extensive effects on many carbon metabolism genes that follow a coherent pattern, across three very different eukaryotes (yeast *Saccharomyces cerevisiae*, nematode *Caenorhabditis elegans*, and rat *Rattus norvegicus*), whose common ancestor lived approximately 1000 million years ago. In biotin starvation, the expression of several genes for glucose utilization and lipogenesis were lowered, and some for fatty acid β -oxidation and gluconeogenesis were raised. It was also found that biotin deficiency causes an ATP deficit in the three species, which activates the energy sensor AMP kinase (AMPK). These changes were associated with changes in several signal transduction pathways and transcription factors. It seems that most of the genomic effects result from activation of AMPK the energy sensor, although in the case of liver glucokinase it seems to be activated by cyclic GMP.

Genetic biotinidase deficiency in a knockout mouse with a null mutation for this gene leads to genomic and metabolic changes similar to those produced in the biotin-starved rat, effects most likely due to the important role of biotinidase in maintaining the biotin requirements by recycling endogenous biotin. There is ample evidence indicating that neither diet nor intestinal microflora alone satisfies the true demand for biotin. Because part of the required biotin is provided endogenously by biotinidase, human patients with its genetic deficiency resemble those of dietary deprivation of this cofactor.

Biotin and Energy Metabolism

The severe energy shortage in biotin deficiency was found to be multifactorial. Part of this shortage is caused by reduced anaplerosis due to deficient activity of biotin-dependent pyruvate carboxylase, and the associated flux reduction through the TCA cycle, with concomitant decreased mitochondria respiration. Additionally, there was diminished activity of cytochrome *c* oxidase (respiratory chain complex IV) associated with diminished cytochrome activity, whose functional group is heme. Indeed heme—the prosthetic group of cytochromes—has succinyl CoA, an intermediate of the TCA cycle, as one of whose precursors. All these findings explain the reduced oxidative phosphorylation in biotin deprivation, as shown by a

more pronounced decreased oxygen consumption by biotin-deprived, oligomycin-treated mitochondria than in the controls. Furthermore, there was severe damage and reduction in the number and size of mitochondria, due to an accumulation of compounds toxic for mitochondria resulting from deficiencies of biotin-dependent enzymes, such as propionyl CoA carboxylase. In living rats deprived of biotin there is an increase of insulin sensitivity, a consequence of augmented translocation to the plasma membrane of the glucose transporter GLUT4 and consequent glucose uptake by muscle and adipose tissue. These findings are the result of an insulin-independent effect of AMPK activation on GLUT4 membrane translocation.

Biotin Sources and Requirements

Animals derive biotin from de novo synthesis by bacteria, yeast, molds, algae, and some plant species. Biotin is widely distributed in natural foodstuffs, but the absolute content of even the richest sources is low when compared to the content of most other water-soluble vitamins. Foods relatively rich in biotin include milk, egg yolk, liver, cereals, nuts, legumes, and some vegetables. Most foods, except milk, contain the vitamin bound to proteins, such as biocytin, that is, biotin covalently bound to protein lysyl residues. The average adult human daily biotin intake has been estimated to be approximately 30–70 μ g. Biotin synthesized by intestinal bacteria does not seem to contribute significantly to cover its requirements.

Deficiency and Supplementation

One of the causes of biotin deficiency is found in people eating raw egg diets, rich in avidin, a protein that strongly binds biotin and prevents its intestinal absorption. Biotin deprivation is rather frequent in children with protein/energy undernutrition. Studies of biotin status during pregnancy provide evidence that a marginal degree of biotin deficiency develops in at least one-third of women during normal pregnancy. Although the degree of biotin deficiency was not severe enough to produce overt manifestations, it was sufficiently severe to produce metabolic derangements. A similar marginal degree of biotin deficiency causes high rates of cleft palate and limb shortening in some mammals. Moreover, data from a multivitamin supplementation study provide significant, albeit indirect, evidence that the marginal degree of biotin deficiency that occurs spontaneously in normal human gestation is teratogenic. This post hoc analysis indicated that multivitamin supplementation including biotin reduced the combined incidence of cleft palate and limb shortening.

Thiamine

Thiamine functions as prosthetic group of oxidative dehydrogenases, which, as holoenzymes, decarboxylate their respective ketoacid substrates and transfer the resultant acyl groups to CoA, regenerating NADH (nicotinamide adenine dinucleotide reduced form). It becomes a metabolically active coenzyme when it is esterified with two phosphates, yielding thiamine pyrophosphate. Thiamine is an essential cofactor for ATP and NADH generation, being a coenzyme of pyruvate dehydrogenase (PDH), which converts pyruvate into acetyl CoA, and of alpha-ketoglutarate dehydrogenase in the TCA cycle. Thus, thiamine is essential for the complete metabolism of glucose, and its nutritional deprivation slows the flow through the TCA cycle, diminishing the input of electrons to the mitochondrial respiratory chain. Another effect as in the above cycle, like the one in biotin deficiency, through a reduction of succinyl CoA, a precursor of heme formation, might be a diminished activity of the electron transfer chain complex IV, further decreasing oxidative phosphorylation. Thiamine has other different metabolic functions, being a cofactor for branched-chain keto acid dehydrogenase, the enzyme that is affected in maple syrup urine disease, and for transketolase in the pentose phosphate pathway, whose products are NADPH for use in biosynthetic processes (especially lipogenesis), the de novo synthesis of ribose for nucleotide formation and has a key role in oxidant defense, which exerts an important control on the pentose phosphate pathway (Fig. 35.4).

Genetic Disorders of Thiamine-Dependent Enzymes

Genetic disorders leading to a thiamine deficiency-like status may be the result of mutations in thiamine pyrophosphate-dependent enzymes, thiamine transporters, and enzymes of thiamine metabolism, in particular thiamine phosphate kinase. A thiamine responsive alpha-ketoglutarate dehydrogenase deficiency has been reported in human patients, associated with failure to form sufficient succinyl CoA for heme synthesis. Mutations in enzymes of the PDH complex may result in some forms of Leigh's syndrome, as in the deficiency of holocarboxylase synthetase. Administration of thiamine in these patients delays—although does not prevent—the fatal outcome. The thiamine-dependent branched-chain 2-oxoacid dehydrogenase complex has a central role in the degradation of the branched-chain amino acids valine, leucine, and isoleucine. Decreased activity of this enzyme complex owing to genetic defects leads to the accumulation of the toxic 2-oxo acid products, responsible for maple syrup urine disease, and some patients respond to large doses of thiamine. Mutations in the genes coding for thiamine transporters may lead to specific diseases in humans. One of the best known is thiamine-responsive megaloblastic anemia syndrome, characterized, besides the anemia, by sensorineural hearing loss, and diabetes mellitus. It is caused by mutations in the *SLC19A2* gene, coding for the high-affinity transporter. Mutations in *SLC19A3* generally cause encephalopathies.

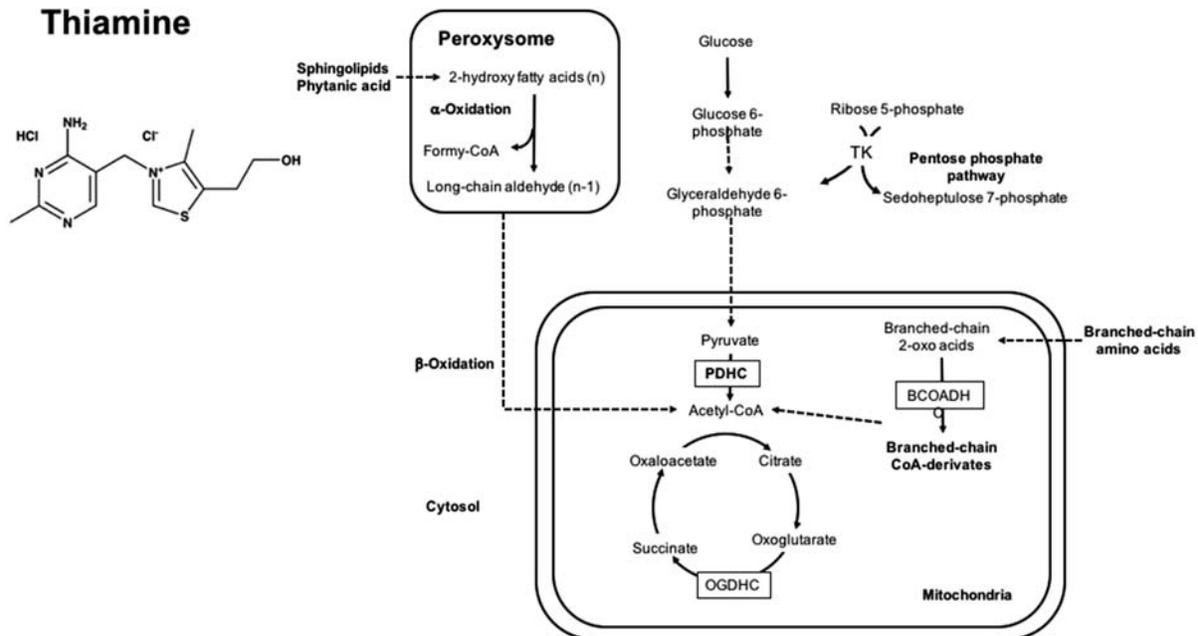


FIGURE 35.4 Thiamine structure and reactions dependent on this vitamin.

Nutrigenomics of Thiamine Deficiency

Similarly to findings in biotin deprivation, Velázquez-Arellano and coworkers found a reduced energy state in the liver and activation of the energy sensor AMPK in thiamine deficiency. These results are a consequence of the key role of thiamine in the ATP production process as a cofactor of PDH and of the TCA cycle enzyme 3-ketoglutarate dehydrogenase. Their deficient activities in thiamine deprivation impinge on mitochondrial oxidative phosphorylation.

Because the activation of the energy sensor AMPK reduces ATP utilizing processes, there is an augmentation of the expression of the gene for liver glucokinase in glycolysis and in fatty acid synthesis, the gene for fatty acid synthetase. In the case of those processes that increase the generation of ATP, in one instance there is an increase of the expression of the gene for carnitine palmitoyltransferase 11 in beta oxidation. An apparent contradiction, the increase of the gluconeogenic phosphoenolpyruvate carboxykinase mRNA can be explained by the difference between an acute AMPK effect, as in a short fast, via disruption of the transcription factor CREB by means of CREB-regulated transcription coactivator 2, and is equivalent to a prolonged fast, such as in thiamine deficiency, when gluconeogenesis would be sustained by other factors, namely the coactivator PGC-1alpha and the transcription factor FOXO1, ensuring continuous supply of enough glucose for the brain. As in biotin deficiency, thiamine-deprived mice are more tolerant to glucose. These changes are parallel to those of two main proteins in the insulin-signaling pathway, AKT and mTORC1, and suggest that the latter may underlie the former results.

Detrimental metabolic consequences of thiamine deficiency (ThDEF) develop more rapidly than in biotin deficiency. It is well known that glucose can be completely oxidized without the participation of biotin, but not thiamine. Both thiamine and biotin function as cofactors of enzymes in amino acid catabolism, thiamine of the branched-chain dehydrogenase and biotin of propionyl CoA carboxylase and of 3-methyl crotonyl CoA carboxylase. In the case of thiamine deprivation, this latter deficient activity most likely results in accumulation of the toxic ketoacids L-alpha-keto-beta-methylvalerate, alpha-ketoisovalerate, and alpha-ketoisocaproate. They may damage mitochondria, as has been observed in maple syrup urine disease, the genetic deficiency of the branched-chain ketoacid dehydrogenase.

Sources and Requirements

Thiamine is widely distributed in foods, but most foods contain only low concentrations of the vitamin. The richest sources are yeast and liver; however, cereal grains comprise the most important dietary sources of

the vitamin in most human diets; however, the greatest concentrations of thiamin in grains are typically found in the scutellum and the germ. Therefore, milling to degerminate grain yields products that result in very low thiamine content. Thiamine recommended daily allowance for adults is 1.2 mg/day.

Deficiency and Supplementation

ThDEF can result in different syndromes, beriberi, or Wernicke's encephalopathy with Korsakoff's psychosis. In experimental animals, it is associated with severe anorexia, thus one of the problems in interpreting its effects is distinguishing between specific ones and those due to general lack of food and inanition. Therefore, thiamine supplementation is advisable for alcohol abusers and might be useful for elderly people, as both might have decreased intestinal thiamine absorption. Therapeutic doses may vary from 10 to 200 mg/day.

Pantothenic Acid

Pantothenic acid is the main component of coenzyme A (CoA), an essential cofactor for 4% of known enzymatic reactions, especially as a prosthetic group in the TCA cycle, the beta oxidation fatty acid pathway, and leucine metabolism. CoA is also required for the oxidation of other fatty acids located in the peroxisomes, and is also required for the first step of cholesterol and fatty acid biosynthesis located in the cytosol. It is also part of heme synthesis, and a derivative from CoA metabolism is incorporated into the prosthetic group of the acyl carrier protein, an essential cofactor in fatty acid synthesis (Fig. 35.5).

Genetic Disorders of Enzymes Involved in CoA Synthesis

Disease states, such as Hallervorden-Spatz syndrome, an autosomal recessive neurodegenerative disorder, are characterized clinically by dystonia and optic atrophy or pigmentary retinopathy with iron deposits in the basal ganglia and globus pallidus. It was mapped to a *PANK2* gene mutation that codes for the mitochondrial kinase involved in initiating intramitochondrial CoA biosynthesis. Other neurodegenerative diseases may have as their base defects in pantothenate kinase.

Pantothenic Acid and Coenzyme A Synthesis and Distribution

Pantothenic acid is present in foods, as CoA. Ingested CoA is hydrolyzed by phosphatases of the intestinal lumen to pantotheine, and is then split, by pantotheinase secreted from the intestinal mucosa, into pantothenic acid. Pantothenic acid is cotransported across the brush-border membrane of the intestinal

Pantothenic acid

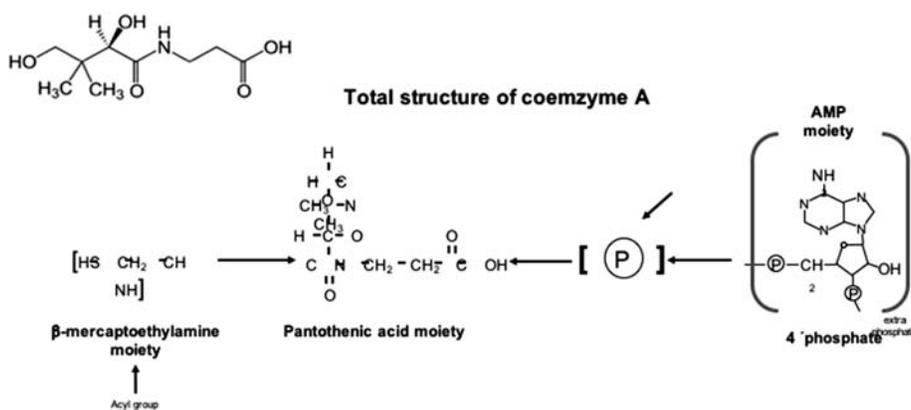


FIGURE 35.5 The structure of coenzyme A. Pantothenic acid is a component of CoA.

enterocyte with sodium via the sodium-dependent multivitamin transporter (SMVT), which transports pantothenate, biotin, and lipoate; therefore, there may be antagonisms between these three cofactors, but there is no evidence for clinical consequences. Pantothenic acid is transported by the portal vein to the liver, where it is again taken up by the SMVT. Cysteine, ATP, and five enzymes are involved in the synthesis of CoA from pantothenic acid in the hepatocyte.

Sources and Requirements

Pantothenic acid is widely distributed in foods, particularly in animal organs, egg yolks, peanuts, and broad beans with smaller amounts in lean meat, milk, potatoes, and green leafy vegetables. Chicken, beef, potatoes, oat cereals, tomatoes, broccoli, and whole grains are major sources of pantothenic acid. Levels, however, are low in highly processed foods, including refined grains, fruit products and meat or fish with added fat or cereals.

A recommended dietary allowance has not been established, but 4–7 mg per day has been suggested. With these intakes, serum levels are maintained in young adults, and no known signs of deficiency are observed. There is little information on its enteric synthesis. The contribution of bacterial synthesis to human pantothenic acid status is not known, and probably is small.

Deficiency and Supplementation

Pantothenic acid deficiency is not normally seen in humans. Pantothenate supplementation increased hepatic glutathione (GSH) levels in Jurkat cells that was

attributed to increased ATP production as a result of increased mitochondrial CoA levels. The cytoprotective effect of pantothenic acid therefore likely resulted from increased CoA and GSH levels.

Suggested Reading

- Brody, T., 1999. *Nutritional Biochemistry*, second ed. Academic Press.
- Hernandez-Vazquez, A., et al., 2017. Thiamine deprivation produces similar metabolic and genomic effects as biotin deficiency and has implications for energy disorders. *J Nutrigenet Nutrigenom* 9, 287–299.
- Labay, V., et al., 1999. Mutations in SLC19A2 cause thiamine-responsive megaloblastic anaemia associated with diabetes mellitus and deafness. *Nat Genet* 22 (3), 300–304.
- Mardach, R., et al., 2002. Biotin dependency due to a defect in biotin transport. *J Clin Invest* 109 (12), 1617–1623.
- Ochoa-Ruiz, E., et al., 2015. Biotin deprivation impairs mitochondrial structure and function and has implications for inherited metabolic disorders. *Mol Genet Metab* 116 (3), 204–214.
- Ortega-Cuellar, D., et al., 2010. Biotin starvation with adequate glucose provision causes paradoxical changes in fuel metabolism gene expression similar in rat (*Rattus norvegicus*), nematode (*Caenorhabditis elegans*) and yeast (*Saccharomyces cerevisiae*). *J Nutrigenet Nutrigenom* 3, 18–30.
- Rodríguez-Meléndez, R., Cano, S., Méndez, S.T., Velázquez, A., 2001. Biotin regulates the genetic expression of holocarboxylase synthetase and mitochondrial carboxylases in rats. *J Nutr* 131, 1909–1913.
- Salvador-Adriano, A., et al., 2014. Insulin sensitivity is inversely related to cellular energy status, as revealed by biotin deprivation. *Am J Physiol Endocrinol Metab* 306 (12), E1442–E1448.
- Valle, D., Beaudet, A.L., Vogelstein, B., et al., 2008. *The Online Metabolic & Molecular Basis of Inherited Disease*. McGraw-Hill. Online edition.
- Velázquez-Arellano, A., et al., 2011. A heuristic model for paradoxical effects of biotin starvation on carbon metabolism genes in the presence of abundant glucose. *Mol Genet Metab* 102 (1), 69–77.

This page intentionally left blank

36

Impact of Genetic Variation on Vitamin C Status

Martin Kohlmeier

Human Research Core and Nutrigenetics Laboratory, UNC Nutrition Research Institute; and Schools of Medicine and Public Health, Department of Nutrition, Kannapolis, NC, United States

FUNCTION OF VITAMIN C IN HEALTH AND DISEASE

The term *vitamin C* covers the fully reduced form L-ascorbic acid (ASC), the oxidized forms semidehydroascorbate (SDA) and dehydroascorbic acid (DHA), as well as a number of other biologically active metabolites and derivatives (Kohlmeier, 2015).

Both ASC and SDA are potent free radical scavengers (antioxidants), particularly in cytosol due to their preferred distribution to this aqueous environment. ASC is also an essential cofactor for several enzyme-catalyzed reactions for protein modifications (such as hydroxylation of proline and lysine moieties of collagen), amino acid, steroid and lipid, iron and sulfate metabolism, DNA modification, and reactivation of oxidized vitamin E.

Humans get most of their vitamin C from fruits and vegetables, and to a much smaller extent from animal-derived foods. Citrus fruits (such as oranges, grapefruits, and lemons) are particularly rich sources. Cooking and prolonged storage decrease the vitamin C content of foods.

Deficiency increases fragility of blood vessels and capillaries, impairs wound healing, and has many less-specific adverse health effects. Manifest deficiency, scurvy, is characterized by bleeding in gums and around hair roots, painful swelling of joints, poor wound healing, impaired immune function, fatigue, confusion, and potential death. A number of other illnesses and conditions have been related to inadequate vitamin C status, but in many cases the causal relationship is insufficiently established.

The concentration of ascorbic acid in blood plasma is an imperfect but nonetheless useful indicator of current vitamin C status.

ABSORPTION, METABOLISM, AND DISPOSITION OF VITAMIN C

Ingested vitamin is absorbed in the small intestine. The modest amounts of vitamin C present in foods are absorbed almost completely, whereas only some of larger quantities, such as in high-dose dietary supplements, are taken up. The sodium-dependent vitamin C transporter 1 (encoded by SLC23A1) mediates the selective and concentrative uptake of luminal ASC, which is critically important during times of low intake (Subramanian et al., 2017). DHA can be taken up from the intestinal lumen through glucose transporters, including GLUT2 (encoded by SLC2A2) and possibly GLUT8 (encoded by SLC2A8), and exit across the basolateral membrane via GLUT2, dependent on the prevailing concentration gradient (Fig. 36.1). DHA can be reduced to ASC by glutathione (nonenzymatically), ascorbate-dependent glutathione reductase (EC1.8.5.1, encoded by GSTO1 as well as GLRX), or by the NADPH-dependent and selenium-containing thioredoxin reductases (EC1.6.4.5, encoded by TXNRD1, TXNRD2, and TXNRD3). SDA can be reduced to ASC by thioredoxin reductases. The sodium-dependent vitamin C transporter 2 (encoded by SLC23A2) facilitates transfer of ASC across the basolateral membrane of the enterocytes. The gene nomenclature of the vitamin C transporters can be confusing because the numbering

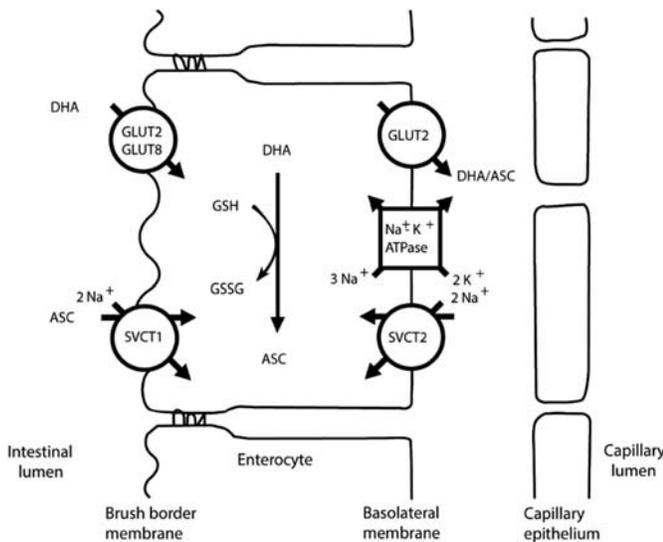


FIGURE 36.1 Pathways for absorption of vitamin C from the small intestines.

of SLC23A1 and SLC23A2 has been switched, and earlier publications have to be read with particular attention to the correct attribution.

Vitamin C circulates with blood through the body almost exclusively as ASC, and is taken up by many different cell types via SLC23A2. The glucose transporters GLUT1, GLUT3 and some others in that family are high-capacity DHA transporters, which can mediate uptake of any available DHA.

Reactivation of DHA depends strongly on reduced glutathione (Fig. 36.2). The glutathione S-transferases, which are important for maintaining redox balance and detoxification of a wide range of endogenous and xenobiotic compounds, compete with DHA for the availability of reduced glutathione.

Both SLC23A1 and SLC23A2 mediate the recovery of vitamin C metabolites from primary filtrate in the proximal tubules of the nephron (Michels et al., 2013). These transporters function very much like in the small intestines, SLC23A1 at the luminal side and SLC23A2 at the basolateral membrane. Losses into urine are smallest at low vitamin C status and much higher with excessive dietary intake.

Irreversible metabolic vitamin C loss occurs through interaction with circulating free hemoglobin, heme and iron. Haptoglobin can reduce the conversion to inactive

metabolites because of its property to bind free hemoglobin released from ruptured red blood cells.

GENETIC VARIATION DIFFERENTIALLY AFFECTING VITAMIN C STATUS

Human populations and individuals within populations can differ greatly in their vulnerability both to inadequate and excessive vitamin C intake. Variants in several genes have been found to modulate individual responses. Among others, carriers of the genotypes SLC23A1 rs4257763 AA, SLC23A2 rs6139591 TT, and Hp2-2 tend to have lower ASC concentration in plasma than noncarriers as recently summarized (Michels et al., 2013).

SLC23A1

In the Toronto Nutrigenomics and Health Study of nonsmoking men and women, concentrations of ASC in blood were distinctly lower in carriers of the SLC23A1 rs4257763 genotypes GG than AA (24.4 vs. 29.7 $\mu\text{mol/L}$; $P = .002$) (Cahill & El-Sohehy, 2010a).

SLC23A2

About 18% of Caucasians carry two copies of the intronic variant rs6139591 T. This genotype appears to be associated with increased vulnerability to low vitamin intake (Shaghghi et al., 2016). Investigators following a cohort of Danish women for more than 6 years found that lower than the median vitamin C intake (<127 mg/day) was associated with more than fivefold increased risk of coronary syndrome in carriers of the SLC23A2 rs6139591 TT genotype, but not in women with the CC genotype (Dalgård et al., 2013). This genotype also appears to predict increased risk of premature delivery (Erichsen et al., 2005) and of some types of cancer (Wright et al., 2009).

Glutathione Transferases

Gene variants of glutathione S-transferase genes encoding enzymes with low or no activity are very common. About half of all Caucasians carry two null alleles of the GSTM1 gene. Individuals with genetically

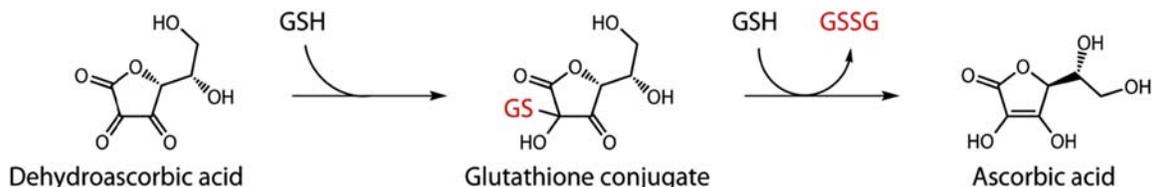


FIGURE 36.2 Reduced glutathione (GSH) helps with the reactivation of dehydroascorbic acid (DHA).

absent M1 or T1 enzyme activities may be slightly less vulnerable to low vitamin C intake and have higher ascorbic acid concentration in blood (Cahill et al., 2010a; Block et al., 2011).

Haptoglobin

A common duplication of a segment in the HP gene gives rise to the Hp1 (without duplication) and Hp2 (with the duplication) isoforms. For simplicity both the genetic alleles and the protein isoforms will be referred to as Hp1 and Hp2, respectively. The haptoglobin protein isoforms differ in their propensity to form polymers. People with two copies of the DNA sequence coding for the Hp1 isoform (Hp1-1 carriers) form mostly dimers, whereas those with the partially duplicated sequence coding for the Hp2 isoform tend to form much larger haptoglobin polymers. The dimers are much more effective hemoglobin binders than the polymers. The prevalence of the Hp2 allele is over 63% in Caucasians, but much less in some other populations including Inuits and Easter Islanders.

Findings differ on the effect of the HP genotypes on vitamin C status. Some investigators report lower ASC concentration with Hp2-2 than with Hp1-1, while others do not confirm this (Guthrie et al., 2014). However, it makes little sense to assess vitamin C status only based on a genotype without taking into account dietary intake. Since vitamin C is an essential nutrient, by definition everybody will become deficient with low intake. The question is just how low that intake threshold is at which vitamin C status becomes compromised. The most consistent finding is that with lower than optimal vitamin C intake Hp2-2 carriers have low ASC concentration in blood and carriers of the other genotypes do not (Cahill & El-Sohehy, 2010b).

It is noteworthy that almost all native Easter Islanders were reported to be Hp1 carriers (Nagel et al., 1964). Since their ancestors came from islands thousands of miles away requiring sea voyages of many months or even years, it seems that only carriers of the HP genotype most resistant to vitamin C loss made it to their new home (Kohlmeier, 2012). Adaptation to a severely vitamin C-poor nutritope (nutritional environment) with daily intakes often below 20 mg (Clow et al., 1975) may similarly be the reason for the high prevalence of the Hp1 allele in native Inuit populations. When individuals from other ethnicities with their high prevalence of the Hp2-2 genotype came to this vitamin C-poor nutritope during the Klondike gold rush, they commonly suffered from scurvy during the winter months with often lethal consequences while the natives survived.

It is safe to conclude that Hp2-2 carriers are dependent on higher vitamin C intakes than people with the genotype producing only Hp1.

Health Significance

The consequences of these genetic vulnerabilities is becoming increasingly evident as the results of an increasing number of high-quality investigations are becoming available.

We may take, for example, the above-mentioned Danish cohort where lower than median vitamin C intake (<127 mg/day) was associated with more than fivefold increased risk of coronary syndrome in carriers of the SLC23A2 rs6139591 TT genotype, but not in women with the CC genotype (Dalgård et al., 2013). This suggests that a sizable number of women with vitamin C intake at recommended levels were at higher risk than carriers of the same genotype with a more generous supply of vitamin C from foods. Uncertainty about the validity of these findings will unavoidably persist, not only until the findings have been replicated in additional cohorts but also until results from randomized placebo-controlled trials are available. The hard question is whether it is reasonable in light of the already available credible information to wait that long before recommending moderately higher vitamin C intakes (from foods) for these genetically vulnerable subgroups.

CONCLUSIONS

The genomes of populations and individuals show clear signs of their ancestors' adaptation to the availability of vitamin C in their food supply. We already know enough about the resulting differences in responses at various levels of vitamin C intake to make risk predictions. Carriers of some genotypes clearly are more likely to have inadequate vitamin C status. Knowing about such greater-than-average vulnerability can help to educate at-risk individuals about good food sources of vitamin C and guide them to avoid deficiency.

References

- Block, G., Shaikh, N., Jensen, C.D., Volberg, V., Holland, N., 2011. Serum vitamin C and other biomarkers differ by genotype of phase 2 enzyme genes GSTM1 and GSTT1. *Am J Clin Nutr* 94 (3), 929–937. PMID:21813807.
- Cahill, L.E., El-Sohehy, A., 2010a. Vitamin C transporter gene polymorphisms, dietary vitamin C and serum ascorbic acid. *J Nutrigenetics Nutrigenomics* 2 (6), 292–301. PMID:20588054.
- Cahill, L.E., El-Sohehy, A., 2010b. Haptoglobin genotype modifies the association between dietary vitamin C and serum ascorbic acid deficiency. *Am J Clin Nutr* 92 (6), 1494–1500. PMID:20926521.

- Clow, C.L., Laberge, C., Scriver, C.R., 1975. Neonatal hypertyrosinemia and evidence for deficiency of ascorbic acid in Arctic and subarctic peoples. *Can Med Assoc J* 113 (7), 624–626. PMID:1181017.
- Dalgård, C., Christiansen, L., Vogel, U., Dethlefsen, C., Tjønneland, A., Overvad, K., 2013. Variation in the sodium-dependent vitamin C transporter 2 gene is associated with risk of acute coronary syndrome among women. *PLoS One* 8 (8), e70421. PMID:23990905.
- Erichsen, H.C., Engel, S.A., Eck, P.K., Welch, R., Yeager, M., Levine, M., Siega-Riz, A.M., Olshan, A.F., Chanock, S.J., 2006. Genetic variation in the sodium-dependent vitamin C transporters, SLC23A1, and SLC23A2 and risk for preterm delivery. *Am J Epidemiol* 163 (3), 245–254. PMID:16357110.
- Guthrie, P.A., Abdollahi, M.R., Gaunt, T., Lawlor, D.A., Ben-Shlomo, Y., Gallacher, J., Davey Smith, G., Day, I.N., Rodriguez, S., 2014. Haptoglobin duplication, hemoglobin, and vitamin C: analyses in the british women's heart and health study and Caerphilly prospective study. *Dis Markers* 2014, 529456. PMID:25525287.
- Kohlmeier, M., 2012. Haptoglobin. In: Kohlmeier, M. (Ed.), *Nutrigenetics: Applying the Science of Personal Nutrition*. Academic Press, San Diego, ISBN 9-7801-2385-9-006, pp. 168–170.
- Kohlmeier, M., 2015. Water-soluble vitamins and non-nutrients: vitamin C. In: Kohlmeier, M. (Ed.), *Nutrient Metabolism: Structures, Functions and Genes*, second ed. Academic Press, San Diego, ISBN 9-7801-2387-7-840, pp. 570–579.
- Michels, A.J., Hagen, T.M., Frei, B., 2013. Human genetic variation influences vitamin C homeostasis by altering vitamin C transport and antioxidant enzyme function. *Annu Rev Nutr* 33, 15.1–15.26 PMID: 23642198.
- Nagel, R., Echeverry, R., Guzman, C., 1964. Haptoglobin types in inhabitants of easter island. *Nature* 201, 216–217. PMID:14118300.
- Shaghghi, M.A., Kloss, O., Eck, P., 2016. Genetic variation in human vitamin C transporter genes in common complex diseases. *Adv Nutr* 7 (2), 287–298. PMID:26980812.
- Subramanian, V.S., Srinivasan, P., Wildman, A.J., Marchant, J.S., Said, H.M., 2017. Molecular mechanism(s) involved in differential expression of vitamin C transporters along the intestinal tract. *Am J Physiol Gastrointest Liver Physiol* 312 (4), G340–G347. PMID:27932501.
- Wright, M.E., Andreotti, G., Lissowska, J., Yeager, M., Zatonski, W., Chanock, S.J., Chow, W.H., Hou, L., 2009. Genetic variation in sodium-dependent ascorbic acid transporters and risk of gastric cancer in Poland. *Eur J Cancer* 45 (10), 1824–1830. PMID: 19243932.

Riboflavin

Annet M. Bosch

Department of Pediatrics, Division of Metabolic Disorders, Amsterdam University Medical Centers, Amsterdam, the Netherlands

INTRODUCTION

Riboflavin is a water-soluble vitamin that must be acquired from the diet through intestinal absorption.

Uptake takes place through carrier-mediated absorption in the small intestine (Powers, 2003). Since 2008, there has been increasing insight into human riboflavin transport and metabolism. Currently, three inborn errors of riboflavin transport and metabolism are known: riboflavin transporter (RFVT) deficiencies (Bosch et al., 2011), deficiency of the mitochondrial flavin adenine dinucleotide (FAD) transporter (Schiff et al., 2016), and mutations in the FLAD1 gene, which encodes FAD synthase (Olsen et al., 2016).

RIBOFLAVIN

Riboflavin (7,8-dimethyl-10-ribityl-isoalloxazine; vitamin B2) is a water-soluble vitamin that cannot be synthesized or stored in the human body but must be acquired from the diet through intestinal absorption. In the human diet, milk and dairy products, cereals, meats, fatty fish, and green vegetables contribute most to riboflavin intake. In these foods, riboflavin can be present as free riboflavin (mostly in eggs and milk) or derivatives (FAD) and flavin mononucleotide (FMN). Besides the major routes of dietary intake, unknown amounts of riboflavin are generated by bacteria and absorbed in the large intestine. The recommended daily allowance of riboflavin varies from 0.4 mg (infants) to 1.6 mg (lactating women) and excess riboflavin (starting at intakes above 1 mg/day) is excreted in the urine. Populations at risk for riboflavin deficiency are pregnant or lactating women, elderly individuals, and people following vegetarian and vegan lifestyles (Powers, 2003).

Dietary Deficiency

Clinical effects of dietary riboflavin deficiency in humans were reported only in the older literature, where it was highly likely that patients experienced multiple (B) vitamin deficiencies; at this time, there is no clearly defined spectrum of clinical symptoms resulting from dietary riboflavin deficiency. Riboflavin depletion studies were conducted in animals, but results were not confirmed in humans. In animal studies, effects on the hematopoietic system and on iron handling were suspected. Poor riboflavin status was implicated as a risk factor for malignancies, and maternal riboflavin depletion during pregnancy has been associated with abnormal gastrointestinal development, neurological abnormalities, and cataracts (Powers, 2003).

Riboflavin Uptake

Riboflavin uptake depends mostly on carrier-mediated absorption in the small intestine. The first RFVT was identified in 2008. Since then, three RFVTs have been characterized: RFVT1 is mostly expressed in the small intestine, RFVT2 in the brain, and RFVT3 in the small intestine (Fig. 37.1) This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>).

Riboflavin Function: Flavin Mononucleotide and Flavin Adenine Dinucleotide

After uptake in the cells, riboflavin is converted by riboflavin kinase (EC 2.7.1.26) to FMN. FMN is further converted to FAD by the enzyme FAD synthase (EC 2.7.7.2). FMN and FAD are flavoenzymes that act as intermediaries in the transfer of electrons in biological oxidation-reduction reactions, and thus are important

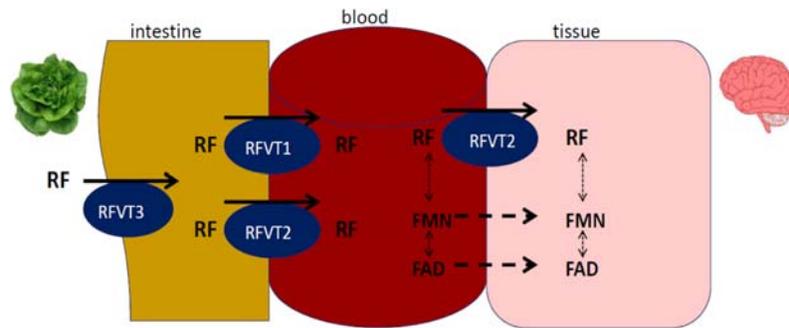


FIG. 37.1 Riboflavin uptake by RFVT 1, 2 and 3. *FAD*, flavin adenine dinucleotide; *FMN*, flavin mononucleotide; *RF*, Riboflavin; *RFVT*, riboflavin transporter. Adapted from Yonezawa, A, Inui, K., 2013. Novel riboflavin transporter family RFVT/SLC52: identification, nomenclature, functional characterization and genetic diseases of RFVT/SLC52. *Mol Asp Med* 34(2–3), 693–701, published in Jaeger, B., Bosch, A.M., 2016. Clinical presentation and outcome of riboflavin transporter deficiency: mini review after five years of experience. *J Inherit Metab Dis* 39(4), 559–564. <https://doi.org/10.1007/s10545-016-9924-2>.

cofactors for carbohydrate, amino acid, and lipid metabolism (Powers, 2003; Olsen et al., 2016).

GENETIC DISORDERS OF RIBOFLAVIN TRANSPORT AND METABOLISM

Three different types of inborn errors of riboflavin transport and metabolism have been reported. First, RFVT deficiencies cause a severe life-threatening neurodegenerative disorder (Bosch, 2012, Jaeger and Bosch, 2016). Second, deficiency of the mitochondrial FAD transporter has been reported to cause exercise intolerance (Schiff et al., 2016). Notably, both transporter deficiencies are treatable with high doses of riboflavin.

Third, mutations in the FLAD1 gene [MIM: 610595], which encodes FAD synthase, cause myopathy and respiratory chain dysfunction and appear to be responsive to riboflavin in some cases (Olsen et al., 2016).

Riboflavin Transporter Deficiency

In 2010, it was demonstrated that the neurodegenerative disorder formerly known as the Brown–Vialeto–Van Laere syndrome or Fazio Londe syndrome is caused by mutations in RFVT genes SLC52A2 (coding for RFVT2 [MIM: 607882]) and SLC52A3 (coding for RFVT3 [MIM: 613350]) (Bosch et al., 2011). Since then, 70 patients with the molecular diagnosis of an RFVT2 or RFVT3 deficiency have been reported in the literature (Jaeger and Bosch, 2016). The inheritance is autosomal recessive.

Clinical Symptoms

Patients may present at all ages; more severely affected patients present at younger ages. In the published cohort, mean age of presentation was 4.1 years (range, 0.25–27 years [SD, 5.0 years]). Clinical symptoms are cranial

nerve deficits (74%); muscle weakness, mostly in the shoulder girdle and neck and upper extremities (73%); hearing loss (73%); sensory symptoms including ataxia (34%); and feeding difficulties (24%) and respiratory symptoms (66%) resulting from muscle weakness and diaphragmatic paralysis (Jaeger and Bosch, 2016). It has been suggested that there may be differences in clinical presentation between RFVT2 and RFVT3 deficiency, with a higher incidence of optic atrophy and ataxia in RFVT2 than in RFVT3 patients (Foley et al., 2014).

Biochemical Abnormalities

The first reported patient was diagnosed with an RFVT deficiency after detection of an abnormal acylcarnitine profile and low plasma flavin levels resulting from riboflavin deficiency. In the reported cohort, untreated acylcarnitine profiles demonstrated abnormalities in 58% of patients whereas profiles in others were fully normal. Also, the plasma flavin levels may be fully normal in affected patients. Therefore, acylcarnitine profiles and flavin levels cannot be used for diagnostic purposes (Jaeger and Bosch, 2016).

Diagnosis

As stated earlier, acylcarnitine profiles and flavin levels demonstrate abnormalities in some but not all patients. The diagnosis can be made only through molecular analysis of all three RFVT genes.

Pathophysiology

Insight into the pathophysiology is limited. Clinical, neurophysiological, and histopathological studies suggest an axonal sensorimotor peripheral neuropathy as well as involvement of the cranial nerves.

Treatment and Prognosis

Untreated, RFVT deficiency is fatal. Of the documented cohort, 31 of 70 patients had not been treated at the time of

publication. All untreated patients demonstrated a deterioration, and 15 patients had died, mostly from respiratory insufficiency. Mean time between onset of symptoms and death was 7.1 years (range, 0.6–19 years), with very rapid progression in the younger age group.

A total of 39 of 70 patients were treated with oral riboflavin. No deaths were reported after the start of treatment and the clinical response to treatment was reported for 30 of 70 patients. Most patients (20 of 30) demonstrated clinical improvement; the other patients remained stable (10 of 30). Clinical improvement occurred in days to months after the start of treatment, and an earlier start of treatment may result in more rapid improvement and a better outcome. Therefore, it is strongly advised to start treatment immediately with oral riboflavin when an RFVT deficiency is suspected, and to continue the treatment while awaiting the results of molecular diagnostics. There is no evidence for an optimal dose of riboflavin, but patients have been reported to recover with doses varying from 7 to 60 mg/kg per day, divided into three doses (Jaeger and Bosch, 2016).

Mitochondrial Flavin Adenine Dinucleotide Transport Deficiency

Schiff et al. (2016) suspected an FAD biosynthesis or transporter deficiency in a 14-year-old child presenting with recurrent exercise intolerance, abnormalities in the acylcarnitine profile, and a strong clinical and biochemical response to oral riboflavin supplementation. The authors demonstrated two heterozygous mutations in the SLC25A32 gene encoding the mitochondrial FAD transporter [MIM: 610815], which imports FAD from cytosol into the mitochondria. In this first reported patient, riboflavin supplementation was a highly effective treatment.

Mutations in Flavin Adenine Dinucleotide Synthase

Olsen et al. (2016) were the first to report on mutations in the FLAD1 gene, which encodes FAD synthase, resulting in myopathy and respiratory chain deficiency. The authors described nine patients from seven unrelated families.

Clinical Symptoms

The age of presentation was variable, from the newborn period to age 45 years, but most patients presented in early infancy. The most frequent symptoms were hypotonia and muscle weakness, as well as swallowing and speech difficulties. Other symptoms were respiratory insufficiency or arrest. One patient presented with cardiomyopathy and two had tachycardia

or cardiac arrest necessitating pacemaker implantation. In contrast to RFVT deficiency, no hearing or visual impairments have been reported.

Biochemical Abnormalities

All nine reported subjects demonstrated acylcarnitine profile abnormalities or ethylmalonic aciduria and/or adipic aciduria. FAD FMN and riboflavin levels in blood and cell extracts were in the control range, but FMN was increased in erythrocytes. Cultured fibroblasts demonstrated mild but significant decreases in FAD synthesis compared with controls.

Diagnosis

All reported subjects were diagnosed through molecular analysis.

Treatment and Prognosis

Four patients died in the first year of life and one died at age 16 years. Three of the nine patients, who had mutations affecting a single amino acid in the expressed protein, demonstrated a milder course and were responsive to riboflavin supplementation. The other patients did not improve after receiving riboflavin.

References

- Bosch, A.M., Abeling, N.G., Ijlst, L., Knoester, H., van der Pol, W.L., Stroomer, A.E., Wanders, R.J., Visser, G., Wijburg, F.A., Duran, M., Waterham, H.R., 2011. Brown-Vialetto-Van Laere and Fazio Londe syndrome is associated with a riboflavin transporter defect mimicking mild MADD: a new inborn error of metabolism with potential treatment. *J Inherit Metab Dis* 34, 159–164.
- Bosch, A.M., Stroek, K., Abeling, N.G., Waterham, H.R., Ijlst, L., Wanders, R.J., 2012. The Brown-Vialetto-Van Laere and Fazio Londe syndrome revisited: natural history, genetics, treatment and future perspectives. *Orphanet J Rare Dis* 7, 83.
- Foley, A.R., Menezes, M.P., Pandraud, A., Gonzalez, M.A., Al-Odaib, A., Abrams, A.J., Sugano, K., Yonezawa, A., Manzur, A.Y., Burns, J., Hughes, I., McCullagh, B.G., Jungbluth, H., Lim, M.J., Lin, J.P., Megarbane, A., Urtizberea, J.A., Shah, A.H., Antony, J., Webster, R., Broomfield, A., Ng, J., Mathew, A.A., O'Byrne, J.J., Forman, E., Scoto, M., Prasad, M., O'Brien, K., Olpin, S., Oppenheim, M., Hargreaves, I., Land, J.M., Wang, M.X., Carpenter, K., Horvath, R., Straub, V., Lek, M., Gold, W., Farrell, M.O., Brandner, S., Phadke, R., Matsubara, K., McGarvey, M.L., Scherer, S.S., Baxter, P.S., King, M.D., Clayton, P., Rahman, S., Reilly, M.M., Ouvrier, R.A., Christodoulou, J., Züchner, S., Muntoni, F., Houlden, H., 2014. Treatable childhood neuropathy caused by mutations in riboflavin transporter RFVT2. *Brain* 137, 44–56.
- Jaeger, B., Bosch, A.M., 2016. Clinical presentation and outcome of riboflavin transporter deficiency: mini review after five years of experience. *J Inherit Metab Dis* 39 (4), 559–564. <https://doi.org/10.1007/s10545-016-9924-2>.
- Olsen, R.K., Koňářková, E., Giancaspero, T.A., Mosegaard, S., Boczonadi, V., Mataković, L., Veauville-Merlii, A., Terrile, C., Schwarzmayer, T., Haack, T.B., Auranen, M., Leone, P., Galluccio, M., Imbard, A., Gutierrez-Rios, P., Palmfeldt, J., Graf, E., Vianey-Saban, C., Oppenheim, M., Schiff, M., Pichard, S.,

- Rigal, O., Pyle, A., Chinnery, P.F., Konstantopoulou, V., Möslinger, D., Feichtinger, R.G., Talim, B., Topaloglu, H., Coskun, T., Gucer, S., Botta, A., Pegoraro, E., Malena, A., Vergani, L., Mazzà, D., Zollino, M., Ghezzi, D., Acquaviva, C., Tyni, T., Boneh, A., Meitinger, T., Strom, T.M., Gregersen, N., Mayr, J.A., Horvath, R., Barile, M., Prokisch, H., 2016. Riboflavin-responsive and -Non-responsive mutations in FAD synthase cause multiple acyl-CoA dehydrogenase and combined respiratory-chain deficiency. *Am J Hum Genet* 98 (6), 1130–1145.
- Powers, H.J., 2003. Riboflavin (vitamin B-2) and health. *Am J Clin Nutr* 77 (6), 1352–1360.
- Schiff, M., Veauville-Merlié, A., Su, C.H., Tzagoloff, A., Rak, M., Ogier de Baulny, H., Boutron, A., Smedts-Walters, H., Romero, N.B., Rigal, O., Rustin, P., Vianey-Saban, C., Acquaviva-Bourdain, C., 2016. SLC25A32 mutations and riboflavin-responsive exercise intolerance. *N Engl J Med* 374 (8), 795–797.
- Yonezawa, A., Inui, K., 2013. Novel riboflavin transporter family RFVT/SLC52: identification, nomenclature, functional characterization and genetic diseases of RFVT/SLC52. *Mol Asp Med* 34 (2–3), 693–701.

Niacin

George Rasti, Nicolas G. Simonet, Alejandro Vaquero

Chromatin Biology Laboratory, Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain

Glossary

Niacin A water-soluble vitamin, also known as nicotinic acid or vitamin B₃.

Nicotinamide A niacin derivative required for NAD⁺ and NADP⁺ biosynthesis.

NAD⁺ and NADP⁺ Nicotinamide adenine dinucleotide (NAD⁺) and nicotinamide adenine dinucleotide phosphate (NADP⁺), coenzymes of the vast majority of redox enzymatic reactions involved in metabolism.

Protein Mono- or Poly-Adenosine Diphosphate ribosylation A post-translational protein modification caused by the transfer of one or more adenosine diphosphate ribose molecules from NAD⁺ to a target protein.

NAD⁺-dependent lysine deacetylation Removal of acetyl group from a lysine residue in a target protein caused by transfer of the acetyl group to an adenosine diphosphate-ribose molecule previously generated from NAD⁺ hydrolysis. As a result, O-acetyl-adenosine diphosphate ribose is generated and released to the aqueous media.

PARPs A superfamily of mono- and/or poly-adenosine diphosphate-ribosyltransferases involved in genome stability through modification of a wide variety of targets, including histones and nonhistone proteins.

Sirtuins Previously classified as class III of histone deacetylases, members of this family catalyze histone and nonhistone NAD⁺-dependent deacetylation and mono-adenosine diphosphate ribosylation.

INTRODUCTION

Niacin or vitamin B₃ is the generic term for nicotinamide (NAM) and nicotinic acid (NA), two related water-soluble organic compounds that are essential for controlling cellular energy and metabolism and the signaling between these processes and all other major cell-regulatory pathways. Although animals can produce niacin in the liver from tryptophan through the kynurenine pathway, the main source of this vitamin is exogenous, from a wide range of foods, particularly meat, grains, and milk-based products (Fig. 38.1). The

European Food Safety Authority recommends that adults consume 5.5–6.6 mg NA per 1000 kcal (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). The uptake of niacin in the intestine is mediated by a specific and high-affinity acidic pH-dependent, carrier-mediated process. Once in the bloodstream, NAM and NA cross cell membranes by simple diffusion, whereas transit in erythrocytes and kidney tubules is mediated by a carrier. Niacin supplementation during gestation prevents congenital malformations caused by NAD⁺ deficiency (Shi et al., 2017). Although it is rare in industrialized countries, niacin or tryptophan deficiency can lead to the development of the human disease pellagra, which is characterized by dermatitis, diarrhea, dementia, depression, high sensitivity to sunlight, and tongue and mouth soreness.

NIACIN IN METABOLISM REGULATION AND SIGNALING

NA is an essential component of metabolism regulation because it is the precursor of NAM adenosine dinucleotide (NAD⁺) and NAM adenine dinucleotide phosphate (NADP⁺), which, together with their reduced forms NADH and NADPH, have an essential coenzymatic role as electron donors (reduced form) and acceptors (oxidized form) of the vast majority of cellular oxidation and reduction reactions involved in catabolic and anabolic processes. The NAD⁺-dependent dehydrogenases are among the most significant NAD⁺-consuming enzymes. They participate in almost all steps of intermediary metabolism that require electron transport, such as glycolysis, the tricarboxylic acid cycle, and the respiratory chain. In contrast to the general role of NAD⁺, NADP⁺ has a more restricted role in specific pathways such as the pentose-phosphate pathway,

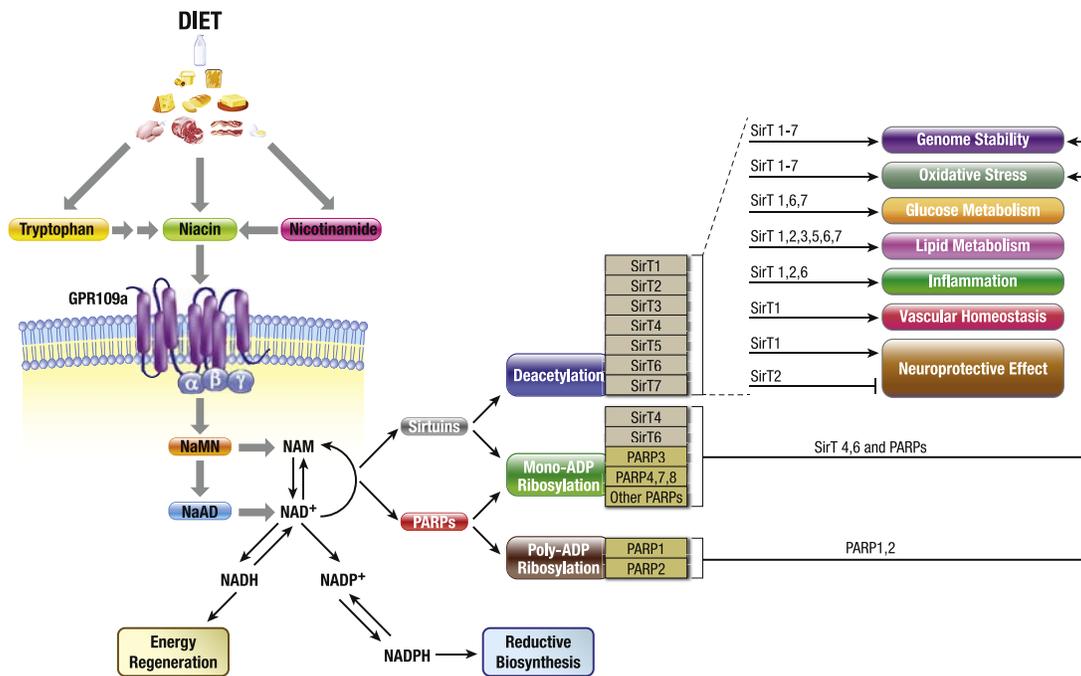


FIGURE 38.1 Schematic representation of niacin metabolism and nicotinamide adenine dinucleotide (NAD⁺) signaling mediated by sirtuins and poly(ADP-ribose) polymerases (PARPs). Niacin and its derivative nicotinamide (NAM) are dietary precursors of NAD⁺/NAM adenosine dinucleotide reduced (NADH) and NADP⁺/NAM adenine dinucleotide phosphate reduced (NADPH), important factors in oxidation and reduction (redox) reactions in metabolic homeostasis (NaMN, nicotinic acid mononucleotide; NaAD, nicotinic acid adenine dinucleotide). *Right*, NAD⁺ signaling regulates a wide range of cellular processes through two families of NAD⁺-dependent enzymes, sirtuins (SirT1–7) and PARPs (PARP1–17). Activation of these NAD⁺-dependent deacetylase (sirtuins and PARPs), and PARP activities ensures optimal adaptation to these stress conditions, in part by regulating stress-related gene expression programs and genome stability.

and fatty acids and steroid biosynthesis (Fig. 38.1). Upon ingestion, niacin is rapidly metabolized through two pathways; glycine conjugation to nicotinic acid, or NAM production directly involved in the synthesis of NAD⁺ and NADP⁺. In turn, NAD⁺ or NADP⁺ catabolism produces NAM once again, which can be further transformed into pyrimidine metabolites such as *N*-methyl-2-pyridone-5-carboxamide or reused in NAD⁺/NADP⁺ synthesis.

Niacin and Metabolic Stress Response Signaling

In addition to this essential function in metabolism, niacin has a key role through NAD⁺ in signaling metabolic status to the other basic functions of the cell. NAD⁺ is required for the catalysis of posttranslational modifications of key specific protein factors that participate in a wide range of cellular processes from metabolism to DNA repair, genomic structure, and transcriptional activity. In all cases, these reactions are based on hydrolysis of the NAD⁺ molecule into adenosine diphosphate (ADP)-ribose and NAM. Because this signaling involves NAD⁺ but not NADH, these activities are triggered under conditions of energy scarcity, which are reflected by an increase in the NAD⁺/NADH ratio.

In particular, NAD⁺ acts a cofactor in two types of enzymatic activities, ADP-ribosyltransferases and NAD⁺-dependent deacetylation. In addition to metabolic and oxidative stress, which induce alterations in the NAD⁺/NADH ratio, these activities are activated upon genotoxic stress, reflecting direct cross-talk between the modulation of metabolic and energetic status and the maintenance of genome integrity. The most significant families of NAD⁺-dependent enzymes involved in this response are poly(ADP-ribose) polymerases (PARPs) and sirtuins (SIRTs) (Fig. 38.1).

Poly(adenosine diphosphate-ribose) polymerases

The superfamily of PARPs is formed of 17 members in humans. The enzymatic reaction involves hydrolysis of the NAD⁺ molecule into NAM and ADP-ribose, which in turn can be transferred to an unmodified target protein (mono-ADP-ribosylation) or to other ADP-ribose molecules already attached to the target residues (poly-ADP-ribosylation). Although most PARPs catalyze poly-ADP-ribosylation of proteins, some are mono-ADP-ribosyltransferases. In general, PARPs have been linked to DNA damage signaling and repair, as well as to genome organization through telomere length

regulation, telomerase activity, and chromosome end protection. Their key roles in genome stability mean that some PARPs are also involved in cell death arising from chronic genotoxic stress.

A few PARP family members, such as PARP1, have been extensively studied, but the vast majority remain relatively uncharacterized. PARP1 is activated upon binding to damaged or abnormal DNA and catalyzes the formation of poly(ADP-ribose) polymers (PARs) on different acceptor proteins, including PARP1 itself (auto-PARYlation), and histone and nonhistone proteins. It also regulates telomere integrity through interplay with telomeric-repeat binding factor-2. Indirectly, PARP-1 has also been linked to metabolism because its hyperactivation can deplete cellular NAD^+ levels. It has also been shown that PARPs and other NAD^+ -dependent enzymes (such as SIRT6) compete for the same intracellular NAD^+ pool, which indirectly connects PARPs with metabolism (Fig. 38.1). Thus, deletion of PARP1 and PARP2 in mice activates SIRT1, leading to mitochondrial biogenesis, increased energy expenditure, and protection from metabolic diseases. PARP3 (also known as ARTD3) has an important role in the cellular response to DNA double-strand breaks (DSBs) through poly-ADP-ribosylation of factors Ku70/Ku80 to limit DNA end resection and control the choice between homologous recombination and non-homologous end-joining in the DSB repair pathway.

Sirtuins

SIRT6s are probably the NAD^+ -dependent enzymes that best exemplify the wide range of processes regulated by NAD^+ signaling. They have an important role in the cross-talk between cellular metabolic and energetic status and the genome by which they coordinate the response to metabolic, oxidative, and genotoxic stress. This broad functional diversity probably explains why they are at the crossroads of many signaling pathways. At the organismal level, they are closely related to endocrine signaling, in which they perform tissue-specific functions to promote a global response to organismal metabolic changes. The global role of SIRT6s has led to the assumption that some of the direct effects promoted by niacin administration might be related to NAD^+ -induced activation of SIRT6s. The seven mammalian members of the SIRT6 family are highly diverse with respect to their enzymatic mechanism, cell localization, and function. Whereas all mammalian SIRT6s have been shown to be NAD^+ -dependent deacetylases, some also display mono-ADP-ribosyltransferase activity. These reactions stem from an initial enzymatic reaction similar to that of PARPs. Once NAD^+ is broken into NAM and ADP-ribose, SIRT6s could either transfer the acetyl group from the target, releasing to the media O-acetyl ADP-ribose

(deacetylase), or simply transfer the ADP-ribose to a target protein (mono-ADP-ribosyltransferase). More general deacetylase activity has also been described in some members that are mainly involved in metabolism. The range of targets of SIRT6s is also wide, including histones, enzymes, transcription factors, and coregulators. SIRT6s regulate the stress response at many levels: first, they directly participate in metabolic regulation by modulating key specific enzymes such as acetyl-coenzyme A synthase, glutamate dehydrogenase (GDH), and long-chain acyl coenzyme A dehydrogenase, among others. Three of the seven mammalian SIRT6s, SIRT3–5, are mainly located in the mitochondria, where they modulate key events. SIRT3 is a general regulator of mitochondrial function, SIRT4 modulates glutamine metabolism in the pancreas by inhibiting mitochondrial GDH, and SIRT5 is involved in the urea cycle. Second, the non-mitochondrial members of the family, SIRT1, SIRT2, SIRT6, and SIRT7, participate in gene expression programs regulated by the most important master regulators of stress, including p53, forkhead transcription factor (FOXO) proteins, hypoxia-inducible factor 1 α (HIF1 α), and nuclear factor- κ B (NF- κ B). This mechanism is partially based on their ability to modulate epigenetic information through direct regulation of key histone marks (e.g., H4K16ac, H3K9ac, and H3K18ac) or by modulating chromatin-associated enzymes or transcription regulators. This in turn confers on SIRT6s the ability to promote cell survival or induce apoptosis if the levels of stress become too high. They also participate in controlling the major metabolic pathways by interacting with key transcription factors such as peroxisome proliferator gamma coactivator 1- α (PGC1- α) or peroxisome proliferator-activated receptor γ (PPAR γ). Third, together with PARPs, they perform direct functions maintaining genome stability in the face of stress conditions through DNA repair (SIRT1, SIRT6, and SIRT7), cell cycle control (SIRT2), and genome structure and organization (SIRT1, SIRT6, and SIRT7).

Niacin and the Oxidative Stress Response

In addition to the signaling role of SIRT6s and PARPs in response to stress, niacin is directly involved in oxidative stress. This type of stress arises from an imbalance between the production of free radical reactive oxygen species (ROS) and antioxidant defenses that detoxify the reactive intermediate to repair the resulting damage. Thus, oxidative stress is implicated in the pathogenesis of a wide variety of chronic diseases. Niacin deficiency increases oxidative stress in several ways. First, a reduction in synthesis of NAD^+ and NADP^+ alters the redox balance of the cell, disrupting electron transport in the respiratory chain. Second, a decrease in NAD^+ production also directly affects the response to stress mediated

by SIRT-dependent regulation of the oxidative stress response expression program. This regulatory event is complex, because high levels of NAM have been shown to inhibit SIRT activity *in vitro* and in cell culture. In any case, increased levels of antioxidant enzymes, such as superoxide dismutase, catalase, and glutathione peroxidase observed after niacin uptake can be explained, at least in part, by the activation of SIRT signaling, because SIRTs have been directly implicated in the control of these genes (Ilkhani and Saedisomeolia, 2016). Third, NADPH is directly involved in regenerating reduced glutathione (GSH) by the enzyme glutathione reductase. Later, glutathione participates in deactivating peroxides such as hydrogen peroxide. In fact, a significant increase in GSH levels has been reported in response to stress. Furthermore, niacin administration could reduce the production of lipid peroxidation (LPO) biomarkers such as malondialdehyde in animal and human studies.

NIACIN IN PHYSIOLOGY

Glucose Metabolism and Insulin Signaling

Although a direct effect of niacin intake on the global alteration of glucose metabolism has not been reported, it has been shown to increase insulin sensitivity. However, in some specific conditions such as obesity, some reports suggest that it may induce insulin resistance (Montastier et al., 2018). Nevertheless, it must obviously have an important role through the production of NAD⁺ and NADP⁺ molecules, because the most relevant glucose metabolic pathways, such as glycolysis or gluconeogenesis, involve redox reactions with NAD⁺/NADH or NADP⁺/NADPH. Moreover, NAD⁺ also has an essential role in controlling glucose-associated pathways and insulin signaling through SIRTs. SIRT1, SIRT3, and SIRT6 inhibit glycolysis and stimulate mitochondrial oxidation of fatty acids. This has a particular effect on tumorigenesis, because cancer cells are highly dependent on glycolysis, a phenomenon known as the Warburg effect. Accordingly, part of the tumor suppressor activity of most SIRTs has been associated with controlling the switch from glycolysis to oxidative metabolism.

SIRT1 and SIRT6 have been shown to inhibit glycolysis by directly repressing glycolytic enzyme phosphoglycerate mutase-1 activity or inhibiting the whole glycolysis expression program by acting as a corepressor for HIF1 α . In the liver, SIRT1 promotes the expression of genes involved in gluconeogenesis by deacetylating and activating PPAR γ and PGC-1 α , whereas overexpression of SIRT1 in mice pancreatic β cells induces insulin release. In addition, SIRT1 and SIRT6 both act as negative regulators of insulin-like growth factor-I and downstream Akt—mammalian target of rapamycin signaling

in skeletal muscle cells. SIRT7 is also involved in glucose homeostasis in a 5'-adenosine monophosphate—activated protein kinase—dependent manner, although the mechanism has not been characterized. Consistent with this, fasting blood glucose and insulin levels are significantly lower in *Sirt*^{-/-} mice than in wild-type animals.

Adipogenesis and Lipid Metabolism

The therapeutic use of niacin to treat atherosclerosis was first considered decades ago owing to its favorable ability to increase high-density lipoprotein (HDL) cholesterol levels in the plasma and reduce levels of all proatherogenic lipids and lipoprotein particles, such as triglycerides, total cholesterol, very-low-density lipoprotein, intermediate-density lipoprotein, low-density lipoprotein, and lipoprotein (a). Niacin is a ligand for a G protein—coupled receptor named GPR109A, which is expressed on the surface of several cell types, such as adipocytes, neutrophils, macrophages, keratinocytes, and epidermal Langerhans cells. The antilipolytic activity of niacin has been demonstrated in adipocytes, in which, upon GPR109A activation, the level of cyclic adenosine monophosphate (cAMP) is reduced by adenylyl cyclase inhibition, resulting in the reduced hydrolysis of triacylglycerols to free fatty acids (Offermanns, 2006). In addition, the interaction between niacin and GPR109A receptor in macrophages induces the adenosine triphosphate—binding cassette transporters ABCA1 and ABCG1, which are both transmembrane proteins responsible for the efflux of cholesterol to the serum and of HDL from macrophage foam cells (Gaidarov et al., 2013). One of the specific effects of niacin is that it regulates the level of serum adiponectin under obesity-induced conditions; it is known to attenuate obesity-induced adipose tissue inflammation through increased adiponectin and antiinflammatory cytokine expression. Niacin also reduces the plasma levels of retinol-binding protein 4 (RBP4), a member of the lipocalin superfamily synthesized in liver and adipose tissue involved in vitamin A (retinol) transport from the liver to peripheral tissues. Thus, niacin administration leads to a decrease in serum RBP4 in wild-type mice but not in niacin receptor GPR109A knockout mice (69th Scientific Sessions, 2009). The key role of niacin in lipid metabolism is also reflected in the beneficial effect of niacin supplementation on the progression of nonalcoholic fatty liver disease. This disease occurs when an abnormal amount of fat accumulates within the liver from causes other than excessive alcohol consumption (nonalcoholic hepatic steatosis [NASH]).

Consistent with this role of niacin, SIRTs are involved in controlling adipogenesis and lipid metabolism, both as a direct effect on the metabolic pathway and by

regulating the expression of the key genes involved. In keeping with the protective role for promoting the efficient use of the limited resource, SIRT1 and SIRT2 have been shown to inhibit adipogenesis. SIRT1 promotes adipogenesis inhibition as well as a switch from white fat adipocytes to metabolically active brown fat by inhibiting the transcription activity of the master regulator of adipogenesis PPAR γ . This mechanism involves direct deacetylation of PPAR γ and binding of SIRT1 with the PPAR γ corepressors nuclear receptor corepressor and silencing mediator of retinoid and thyroid hormone receptors. SIRT1 has also been implicated in NASH, because *Sirt1*^{-/-} mice were found to be more prone to developing hepatic steatosis than were wild-type animals. The involvement of SIRT2 has been associated with the modulation of transcription factor FOXO1 activity, altering the expression of regulators PPAR γ , CCAAT-enhancer-binding protein α , and causing the repression of markers of terminal adipocyte differentiation, such as Glut4, aP2, and fatty acid synthase. In contrast, mitochondrial SIRT3 and SIRT5 promote fatty acid β -oxidation by regulating pyruvate dehydrogenase. SIRT6 has also been linked to fatty acid synthesis inhibition, because fatty acid transport and lipogenesis are induced in *Sirt6*^{-/-} livers in mice. Hepatic SIRT7 has been proposed to have a role opposite the other SIRTs in lipid metabolism, in which SIRT7 promotes the expression of genes involved in fatty acid uptake and synthesis of triglycerides by inhibiting DCAF1/DDB1/CUL4B E3-dependent polyubiquitination of transcription factor TR4. Consistently, *Sirt7*^{-/-} mice accumulate less body fat and are more resistant to developing high-fat diet-induced obesity.

NIACIN IN ENDOTHELIAL DYSFUNCTION

Endothelial dysfunction is a condition characterized by an imbalance between vasodilating and vasoconstricting substances produced by endothelial cells, which may result in raised blood pressure, reduced vasodilation, and proinflammatory and prothrombotic states. Niacin can improve endothelial function under vascular injury and endothelial tube formation under lipotoxic and hypoxic conditions, which are stress conditions that can coexist in skeletal muscle during peripheral ischemic vascular disease associated with obesity and metabolic syndrome. A major risk factor for endothelial dysfunction is oxidative stress and increased formation of ROS, which quench endothelium-derived nitric oxide (NO) and inhibit its vasorelaxing effects. Niacin administration reduces endothelial oxidative stress by increasing the cellular content of NADPH and reducing GSH, whereas ROS are generated in endothelial cells.

SIRTs have a key role in regulating endothelial and vascular homeostasis. In the case of SIRT1, it has been demonstrated that increased NO production promotes endothelial-dependent vasodilation by targeting endothelial NO synthase and regulates the expression of many genes involved in vascular endothelial homeostasis, angiogenesis signaling, and remodeling through the deacetylation of FOXO1.

NIACIN AND VASCULAR HOMEOSTASIS

As a consequence of the hypolipidemic effects of niacin, reduced thrombocyte aggregation and blood viscosity have also been observed. Niacin treatment using HepG2 cells lowers plasminogen activator inhibitor-1, a factor possibly involved in atherogenesis and thrombosis, and intercellular adhesion molecule 1 levels, a cell surface glycoprotein important for the adhesion of white blood cells to vascular surfaces, and which has significant properties as a predictor of future coronary events (Tavintharan et al., 2007). Interestingly, among the SIRT family, SIRT1 is the only one reported to control angiogenesis signaling, by modulating postnatal vascular growth through FOXO1 deacetylation.

NIACIN AND ISCHEMIA

An interesting link was established between niacin administration and the attenuation of reperfusion injury; tissue damage occurs when blood flows into the tissue after a period of ischemia. Niacin treatment leads to improved kidney ischemia-reperfusion (I/R)-induced cardiac dysfunction and the severity of myocardial LPO through sustained myocardial PGC-1 α expression. I/R injury can also be suppressed by SIRT1 or SIRT6, whose functions either induce autophagy in liver or are essential to ensure the sodium sulfide-mediated cytoprotective effect in brain endothelial cells, respectively. Whereas SIRT1 and SIRT6 suppress I/R injury, SIRT2 inhibition and/or depletion in mice promote protection from ischemic injury.

NIACIN AND INFLAMMATION

In addition to its beneficial lipid-regulating effect, niacin reduces the levels of C-reactive protein, one of the most important inflammatory markers, detected in 15% of patients with coronary artery disease. Increasing evidence supports the anti-inflammatory property of niacin. Activation of GPR109A induces macrophage and dendritic cells in the colon to differentiate into T-regulatory lymphocytes and interleukin-10(IL-10)-producing

T cells, whereas in vascular smooth muscle cells, niacin suppresses the activity of the pro-inflammatory NF- κ B signaling pathway. Consistent with this, SIRT1, SIRT2, and SIRT6 have been shown to regulate the NF- κ B pathway through different mechanisms. SIRT1 and SIRT2 directly deacetylate the NF- κ B subunit p65, inducing repression of the target genes and thereby promoting an anti-inflammatory response. In turn, the main role of SIRT6 in the pathway is to prevent NF- κ B hyperactivation. SIRT6 binds to p65 after tumor necrosis factor α (TNF α)-induced activation of the pathway repressing target genes by promoter deacetylation of the active mark H3K9ac. In keeping with this, a report showed that by inducing local inflammation with intravitreal injections of soluble murine TNF α , NAM treatment protects against the loss of pattern electroretinogram amplitude and cell loss (Williams et al., 2017).

Pleiotropic Effect of Niacin on Chronic Kidney Disease

Studies showed that low-dose niacin supplementation (500 mg/day) improves the main risk parameters for chronic kidney disease (CKD), including dyslipidemia and reduced serum phosphorus levels (Kang et al., 2013). Niacin effectively binds phosphate by blocking the NaPi-IIb transporter in the intestine, thereby inhibiting absorption of phosphate into the blood circulation. This disease is also associated with cardiovascular disease and CKD—mineral and bone disorder.

SIRT1, SIRT2, and SIRT3 have been shown to have important protective roles in the kidney. In podocytes, reduced levels of SIRT1 lead to the accumulation of acetylated FOXO4 and expression of the proapoptotic gene Bcl2l11 (also known as Bim), resulting in apoptosis. Moreover, mouse models of diabetic nephropathy showed SIRT1 downregulation in proximal tubules before kidney failure, as monitored by albuminuria. The protective role of SIRT2 in kidney is directly related to its role as an anti-inflammatory factor. SIRT2 deacetylates mitogen-activated protein kinase phosphatase-1 and the NF- κ B p65 subunit in a lipopolysaccharide-induced acute kidney injury model.

Neuroprotective Effects of Niacin

The nervous system is sensitive to metabolic fluctuations and oxidative stress. NAM prevents oxidant-induced apoptotic neuronal injury. NAM administration over 5.0–25.0 mmol/L significantly protects neurons during oxidative stress injuries. In animal studies, NAM improves cognitive function and cell survival and reduces edema after cortical trauma (El Atrash et al., 2015). Conversely, NAM N-methyltransferase in

the brain could methylate NAM excess to form 1-methylnicotinamide (MNA) in the cytoplasm, and superoxides formed by MNA via mitochondria could damage nigro-neostriatal dopaminergic neurons that may be involved in the pathogenicity of Parkinson disease (PD) (Fukushima, 2005). In PD, niacin and its receptor GPR109A have also been shown to induce an antiinflammatory effect through macrophage polarization (Wakade et al., 2018). At the signaling level, SIRT1 were also implicated in neurodegenerative diseases, where they perform general protective roles. SIRT1 participates in neuron differentiation and protects neurons under stress conditions in a number of neurodegenerative diseases including Parkinson, Alzheimer, and Huntington disease (HD), and amyotrophic lateral sclerosis. In Alzheimer disease, SIRT1 deacetylates retinoic acid receptor β and activates the transcription of the secretase ADAM metalloproteinase domain 10, reducing A β production and accumulation. Interestingly, SIRT1 and SIRT2 appear to act in mutual opposition in several neurodegenerative diseases such as PD and HD. In PD, whereas SIRT1 protects against α -synuclein aggregation-induced stress via the deacetylation of heat shock factor 1 (HSF1), promoting an increase in the HSP70 chaperone protein, SIRT2 deacetylates and promotes α -synuclein aggregation and neuronal apoptosis through deacetylation of FOXO3. In HD, SIRT1 promotes expression of brain-derived neurotrophic factor together with TORC1, which maintains dopamine- and cAMP-regulated phosphoprotein DARPP32 levels in the striatum. In HD, as in PD, SIRT2 activity induces polyglutamine repeat accumulation through a mechanism that probably involves protein misfolding response defects and microtubule transport.

References

- El Atrash, A., Dawood, L., Tousson, E., Salama, A., 2015. Neuroprotective role of vitamin B3 in experimentally induced oxidative stress. *IJCEN* 3, 21–25. <https://doi.org/10.12691/ijcen-3-1-4>.
- EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014. Scientific opinion on dietary reference values for niacin. *EFSA J* 12, 3759–42.
- Fukushima, T., 2005. Niacin metabolism and Parkinson's disease. *Environ Health Prev Med* 10 (1), 3–8. <https://doi.org/10.1265/ehpm.10.3>. PMID: 21432157.
- Gaidarov, I., Chen, X., Anthony, T., Maciejewski-Lenoir, D., Liaw, C., Unett, D.J., 2013. Differential tissue and ligand-dependent signaling of GPR109A receptor: implications for anti-atherosclerotic therapeutic potential. *Cell Signal* 25, 2003–2016.
- Ilkhani, F., Saedisomeolia, B.H.A.A., 2016. Niacin and oxidative stress: a mini-review. *J Nutr Med Diet Care* 2 (1), 1–5.
- Kang, G.Y., Parks, J.R., Fileta, B., Chang, A., Abdel-Rahim, M.M., Burch, H.B., et al., 2013. Thyroxine and triiodothyronine content in commercially available thyroid health supplements. *Thyroid* 23, 1233–1237. <https://doi.org/10.1089/thy.2013.0101>.

- Montastier, E., Beuzelin, D., Martins, F., Mir, L., Marqués, M.A., Thalamas, C., Iacovoni, J., Langin, D., Viguier, N., 2018. Niacin induces miR-502-3p expression which impairs insulin sensitivity in human adipocytes. *Int J Obes.* <https://doi.org/10.1038/s41366-018-0260-5> (Epub ahead of print).
- Offermanns, S., 2006. The nicotinic acid receptor GPR109A (HM74A or PUMA-G) as a new therapeutic target. *Trends Pharmacol Sci* 27, 384–390.
- Shi, H., Enriquez, A., Rapadas, M., Martin, E.M.M.A., Wang, R., Moreau, J., Lim, C.K., JO, S., Ip, E., Hughes, J.N., Sugimoto, K., Humphreys, D.T., McNerney-Leo, A.M., Leo, P.J., Maghzal, G.J., Halliday, J., Smith, J., Colley, A., Mark, P.R., Collins, F., Sillence, D.O., Winlaw, D.S., Ho, J.W.K., Guillemin, G.J., Brown, M.A., Kikuchi, K., Thomas, P.Q., Stocker, R., Giannoulatou, E., Chapman, G., Duncan, E.L., Sparrow, D.B., Dunwoodie, S.L., 2017. NAD deficiency, congenital malformations, and niacin supplementation. *N Engl J Med* 377 (6), 544–552.
- Tavintharan, S., Sivakumar, M., Lim, S.C., Sum, C.F., 2007. Niacin affects cell adhesion molecules and plasminogen activator inhibitor-1 in HepG2 cells. *Clin Chim Acta* 376, 41–44.
- Wakade, C., Giri, B., Malik, A., Khodadadi, H., Morgan, J.C., Chong, R.K., Baban, B., 2018. Niacin modulates macrophage polarization in Parkinson's disease. *J Neuroimmunol* 320, 76–79.
- Williams, P.A., Harder, J.M., Foxworth, N.E., Cochran, K.E., Philip, V.M., Porciatti, V., Smithies, O., John, S.W.M., 2017. Vitamin B3 modulates mitochondrial vulnerability and prevents glaucoma in aged mice. *Science* 355, 756–760.

This page intentionally left blank

Folate and Vitamins B₆ and B₁₂

Natalia I. Krupenko

Department of Nutrition, Nutrition Research Institute, University of North Carolina Chapel Hill, Chapel Hill, NC, United States

Glossary Terms

CBS Cystathionine β-synthase
Del Deletions
DHFR Dihydrofolate reductase
Ins Insertions
MMA Methylmalonic aciduria
MS Methionine synthase
MTHFD1 Trifunctional C1-synthase
MTHFR Methylene tetrahydrofolate reductase
MTR Methionine synthase gene
MTRR Methionine synthase reductase gene
NTDs Neural tube defects
OCM One-carbon metabolism
PLP Pyridoxal phosphate
RBC Red blood cells
SAM S-adenosylmethionine
SHMT Serine hydroxymethyltransferase
SNPs Single nucleotide polymorphisms
TC II Transcobalamin II
TCbIR Transcobalamin receptor
tHcy Total homocysteine
THF Tetrahydrofolate

OVERVIEW

One-carbon metabolism (OCM) (Fig. 39.1) encompasses a group of biochemical reactions involving the folate coenzymes, vitamin B₆ (B₆), and vitamin B₁₂ (B₁₂). As with other vitamins, folate is an essential nutrient because humans cannot synthesize this molecule and strictly depend on diet for its proper supply. The folate coenzymes are critical for fundamental cellular processes such as nucleic acid biosynthesis, DNA repair, amino acid biogenesis, and cellular methylation. Therefore, insufficient folate intake, as well as deficiency of folate pathways, are associated with increased risk for certain diseases, most notably neural tube defects (NTDs). The status of one-carbon metabolic pathways is commonly assessed by measurements of plasma levels of folate and homocysteine,

as well as the S-adenosylmethionine (SAM)–S-adenosylhomocysteine ratio. Although plasma metabolites are well-established biomarkers of the vitamin status, accurate disease diagnosis and appropriate prevention and treatment should consider that the intracellular vitamin status is modified depending on individual genetic polymorphisms. Because the OCM pathways display numerous genetic polymorphisms, optimal intakes of folate and associated vitamins within populations will vary owing to variations within relevant genes. Individuals with single nucleotide polymorphisms (SNPs), insertions (ins), or deletions (del) in genes related to folate metabolism may have increased susceptibility to certain diseases including cancer, cardiovascular disease, Alzheimer disease, and dementia.

One-Carbon Metabolism and B Vitamins

Vitamins B₉ (folate), B₆ (pyridoxine), and B₁₂ (cobalamin) (Figs. 39.2–39.4) are important constituents of metabolic pathways in humans. Derivatives of these vitamins often function as coenzymes and therefore are critical elements of numerous biochemical pathways. In some cases, complex reactions require participation of more than one coenzyme. In this regard, the biochemistry of folate-dependent pathways is inseparable from metabolism of B₁₂ and B₆. Likewise, the status of each of the three vitamins should be considered to understand the basis for certain clinical outcomes. This is exemplified by studies that implicated the inadequacy of B₁₂ and choline in the etiology of NTDs in folate-fortified populations.

Folate is a common name for a group of coenzymes involved in the biosynthesis of purine nucleotides and thymidylate, and amino acid metabolism (Fig. 39.2). Folate-dependent amino acid biogenesis includes reactions of serine and glycine interconversion, degradation

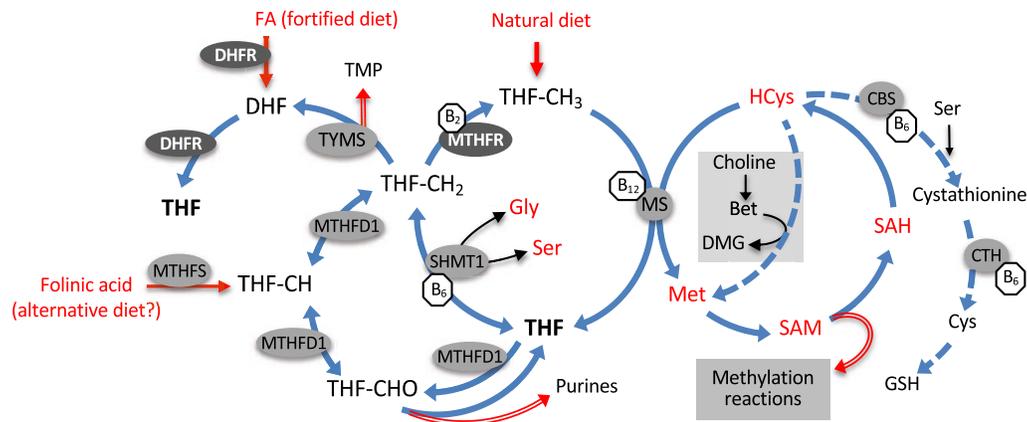


FIGURE 39.1 Pathways of one-carbon metabolism (OCM). Enzymes catalyzing conversions of folate coenzymes are dihydrofolate reductase (DHFR), serine hydroxymethyltransferase (SHMT), 5,10-methylenetetrahydrofolate reductase (MTHFR), methionine synthase (MS), trifunctional C1-synthase (MTHFD1), 5,10-methenyltetrahydrofolate synthetase (MTHFS); and thymidylate synthase (TS). Enzymes related to OCM are cystathionine β -synthase (CBS), cystathionine γ -lyase (CTH). Tetrahydrofolic acid (THF), 5,10-methylenetetrahydrofolic acid (THF-CH₂), 5-methyltetrahydrofolic acid (THF-CH₃), 10-formyltetrahydrofolic acid (THF-CHO), 5,10-methenyltetrahydrofolic acid (THF-CH), dihydrofolic acid (DHF), folic acid (FA), glycine (Gly), serine (Ser), cysteine (Cys), glutathione (GSH), homocysteine (HCys), methionine (Met), S-adenosyl-methionine (SAM), S-adenosylhomocysteine (SAH), betaine (Bet), dimethylglycine (DMG). Octagons designate corresponding B vitamins required for the enzyme activity; red arrows indicate the dietary input of the specific folate form, whereas red and white arrows indicate the biosynthetic pathways using folate bound one-carbon groups. Dashed blue arrows indicate alternative pathways of Hcy use.

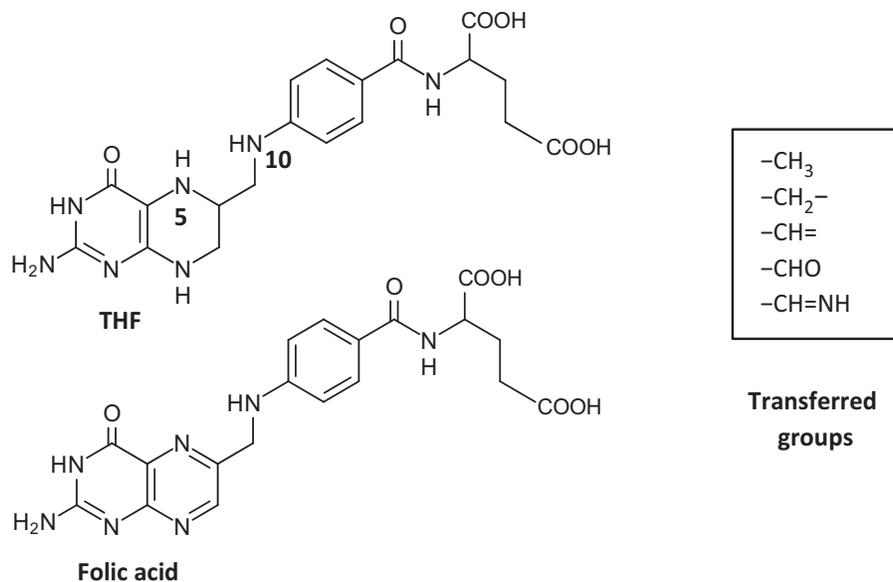


FIGURE 39.2 Folate, vitamin B₉. Structures of the synthetic provitamin, folic acid, and an active form of the coenzyme, tetrahydrofolic acid (THF). Positions, where one-carbon groups can be attached to THF, are indicated in bold. Inset indicates one-carbon groups that can be attached to the tetrahydrofolate form only: -CH₃, methyl; -CH₂-, methylene; -CH=, methenyl; -CHO, formyl; and -CH=NH, formimino.

of glycine and histidine, and the biosynthesis of methionine from homocysteine. In these reactions, folate coenzymes function as carriers of chemical moieties called one-carbon groups (OCG) (Fig. 39.1). Accordingly, the set of folate-dependent reactions is referred to as OCM, a term commonly synonymous to folate metabolism. OCG associated with folate metabolism differ in the oxidation state of the carbon atom and include methyl (CH₃-), methylene (-CH₂-), methenyl

(-CH=), formyl (HCO-), and formimino (NHCH-) groups (Fig. 39.2). In folate-dependent biochemical reactions, the folate coenzyme accepts or donates OCG and enables oxidation or reduction of the carbon that is bound to the folate molecule.

Coenzymes derived from vitamin B₁₂ (Fig. 39.3) participate in only two biochemical reactions in humans: (1) the transfer of a methyl group in the folate-dependent reaction of remethylation of homocysteine

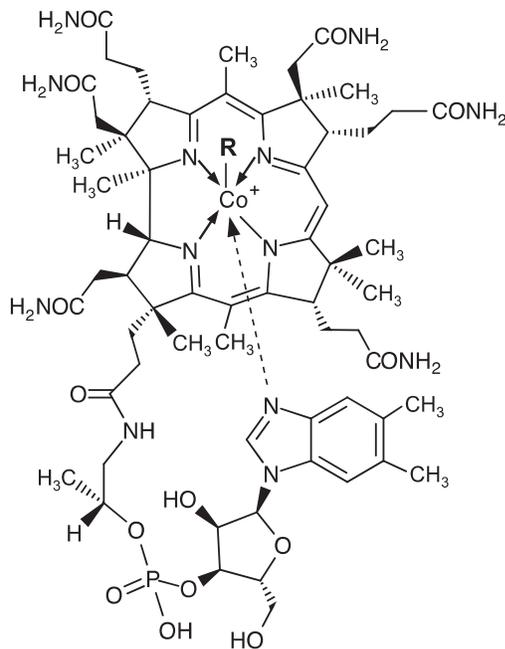


FIGURE 39.3 Vitamin B₁₂. Structure of cobalamin. R is the upper axial ligand, which can be hydroxy- (–OH), cyano- (–CN), methyl- (–CH₃), or adenosyl- (Ado). Methylcobalamin and adenosylcobalamin are active coenzyme forms of the vitamin.

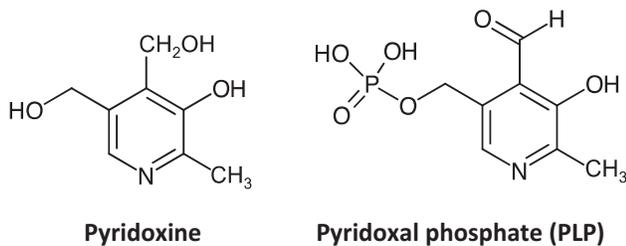


FIGURE 39.4 Vitamin B₆. Structures of the vitamin pyridoxine and the active coenzyme form, pyridoxal phosphate.

to methionine (Fig. 39.1), and (2) the conversion of methylmalonyl–coenzyme A (CoA) to succinyl–CoA. The active forms of this vitamin acting as coenzymes are methylcobalamin and adenosylcobalamin. Deficiency of B₁₂ will prevent both reactions, causing increased levels of circulating total homocysteine (tHcy) and methylmalonic acid. These conditions are associated with specific types of metabolic disorders, hyperhomocysteinemia, and methylmalonic acidemia (MMA, also called methylmalonic aciduria), respectively.

Pyridoxal phosphate (PLP), the derivative of vitamin B₆ (Fig. 39.4), is one of the most common coenzymes in the mammalian cell and is involved in numerous biochemical reactions. Among these reactions is folate-dependent reversible conversion of serine to glycine, catalyzed by the enzyme serine hydroxymethyltransferase (SHMT) (Fig. 39.1). The SHMT reaction is central to

folate metabolism because it loads OCG on tetrahydrofolate, activating it for subsequent reactions (Fig. 39.1). Thus, the three vitamins, B₉, B₁₂, and B₆, are closely related through OCM. Of note, another B₆-dependent pathway, the biosynthesis of cysteine (Fig. 39.1), competes with OCM for homocysteine and serine, providing an additional link between these vitamins. Additional coenzymes (flavine adenine dinucleotide, derivative of B₂; nicotinamide adenine dinucleotide, derivative of B₃; and pantothenic acid, or B₅) also participate in OCM, but they will not be discussed here.

Because of requirements of B₉, B₁₂, and B₆ for fundamental cellular processes, deficiencies in these vitamins will have dramatic effects in humans, causing severe disorders or diseases. Specifically, folate deficiency is associated with megaloblastic anemia and neural tube defects and is also implicated in the increased risk for cardiovascular disease, age-related neurodegeneration, and certain types of cancer. Mandatory food fortification with the synthetic form of the vitamin, folic acid, implemented in the United States and several other countries has led to a significant drop in NTD occurrence. B₁₂ deficiency, which is more common in vegetarians and elderly people, can be caused by insufficient dietary intake owing to malnutrition, as well as by malabsorption. The latter mechanism is associated with insufficient function of intrinsic factor, the protein responsible for B₁₂ absorption, and is the cause of a classic autoimmune disorder, pernicious anemia. However, this type of deficiency can be overcome by intramuscular injection of the vitamin, a common clinical approach to treating patients with insufficient B₁₂ status, as well as by high oral doses of cobalamin. Clinically evident deficiency of B₆ per se is uncommon. However, studies suggest that low plasma PLP levels inversely correlate with major markers of inflammation and are associated with increased risk of coronary artery disease and chronic kidney disease. Suboptimal B₆ status is also associated with impaired cognitive function, Alzheimer disease, cardiovascular disease, and different types of cancer in elderly people. Moreover, inadequacy in B₆ combined with deficiencies of other B vitamins could be a much stronger contributor to etiology of certain diseases.

Genetic Variations in Folate Enzymes

It has become clear that the physiological effects of nutrients cannot be attributed exclusively to their intake since they are absorbed and also modified in different ways depending on an individual genotypic pattern. Numerous studies in the past two decades investigating nutrient–gene interactions and their phenotypic consequences provided clues of how individual variations in genes modulate dietary input. The field of nutrigenomics underwent a rapid expansion after the

development of novel technologies of genetic analysis and metabolomics. Although complete comprehensive analyses of these complex “omics” and epidemiologic data as well as ontology and annotation of the interactions are still missing, studies of B vitamins provided several important links to physiological and metabolic consequences of genes variation and highlighted possible need for alterations in nutritional requirements depending on genetic variation.

POLYMORPHISMS IN VITAMIN B₁₂ PATHWAYS

Transcobalamin II

Insufficient dietary intake of B₁₂ (Fig. 39.3) and the impairment of intrinsic factor secretion are not the only mechanisms causing B₁₂ deficiency. The B₁₂ transporting protein transcobalamin II (TC II), which delivers the vitamin from the ileum to the tissues, has a common polymorphism 776C > G in Caucasian populations leading to the replacement of a proline with an arginine (P259R). The 776C > G genotype significantly influences tissue B₁₂ delivery and functional B₁₂ status. Although the effect of this polymorphism on circulating levels of tHcy is not clear, healthy older adults with two mutant alleles (RR) have significantly higher levels of methylmalonic acid than do those with PP or RP genotypes. Methylmalonic acid is commonly elevated in patients who have a deficiency of the enzyme methylmalonyl-CoA mutase. Methylmalonyl-CoA mutase requires the B₁₂ derivative adenosylcobalamin to catalyze the conversion of methylmalonyl-CoA to succinyl-CoA. This is why insufficient B₁₂ causes the impairment of this reaction. Thus, the elevation of MMA in individuals who carry only the R259 type of TC II is indicative of the deficiency of B₁₂. Another TC II polymorphism causing the S348F substitution was shown to correlate with plasma Hcy levels reflecting effects of B₁₂ deficiency on homocysteine remethylation. It has been hypothesized that these conditions could be a predisease state with regard to B₁₂ status. Direct correlation of TC II polymorphisms with cobalamin levels has also been observed. Thus, the I23V variant of the protein is associated with lower B₁₂ levels and with pernicious anemia. To further complicate the picture, the variant of TC II receptor (TCbIR) with a G220R substitution (owing to C > T polymorphism) was associated with higher serum cobalamin levels. Interestingly, this SNP is strongly linked to a neighboring SNP on the same chromosome that has been shown to have an even stronger influence. TCbIR, which is found in the plasma membrane of most cell types, binds the plasma TC II-cobalamin complex and internalizes it via receptor-mediated endocytosis. Alterations in the

receptor sequence apparently impair the cellular internalization of cobalamin and increase its serum concentration. Thus, the cells will experience B₁₂ deficiency despite the normal or elevated serum cobalamin levels.

Methionine Synthase and Methionine Synthase Reductase

The second enzyme, which requires vitamin B₁₂ for its function in humans, is methionine synthase (MS). This enzyme, which has a complex organization, catalyzes remethylation of homocysteine to methionine using the methyl group from 5-methyl-tetrahydrofolate (THF-CH₃). Under conditions of adequate dietary methionine, approximately 40% of homocysteine is remethylated to methionine through this pathway. In the enzyme mechanism, the cobalamin cofactor first accepts the methyl group from methyl-THF and then transfers it to homocysteine. The enzyme also binds SAM, which is the mechanism of its activation. Catalytically inactive oxidized cobalamin cofactor, which is formed every 200–1000 catalytic cycles, is reduced by the action of a specialized enzyme, MS reductase (MTRR). Overall, MS catalysis requires the B₁₂ cofactor, the folate coenzyme, and SAM and MTRR as activators. In mice, disruption of *Mtr*, the gene encoding MS, results in embryonic lethality, which cannot be rescued by nutritional supplements. These findings indicate that methionine biosynthesis is an absolutely critical process that cannot be bypassed by dietary methionine supplementation. On the flip side, methionine biosynthesis by MS clears homocysteine produced by the action of adenosylhomocysteine hydrolase on S-adenosylhomocysteine, the reaction product of numerous cellular methyltransferases. Lower MS activity results in the accumulation of homocysteine. Significant associations have been found between elevated plasma homocysteine and cardiovascular disease, cerebrovascular disease, cognitive decline, bone dysfunction, pregnancy complications, and other conditions, which underscores the central role of MS in homocysteine detoxification.

The number of natural variants identified for the *MTR* gene in humans that are associated with MS deficiency and cause methylcobalamin deficiency disorder, is relatively small. Patients with these conditions exhibit homocysteinemia, homocystinuria, and hypermethioninemia, and in many cases have megaloblastic anemia associated with neural dysfunction and mental retardation. One mutation in the *MTR* gene produces the enzyme with the substitution of leucine for proline (P1173L). This mutation, which affects the interaction of the enzyme with SAM, is detrimental to MS activity and causes severe cases of cobalamin deficiency. In contrast, a common SNP in *MTR*, 2756A > G, which replaces aspartate 919 with glycine, has a mild phenotype.

This SNP involves the MTRR binding domain of MS, thus affecting methylation and activation of the B₁₂ cofactor. Although this SNP has been implicated as a risk factor of certain diseases (such as systemic lupus erythematosus, hypertension, and gastric cancers), its effect on circulating levels of homocysteine and clinical outcomes still is unclear. MS function can be also affected by genetic variants of the MTRR gene. A genetic polymorphism in this gene (66 A>G), causing the substitution of isoleucine with methionine (I22M) in the flavin mononucleotide-binding domain of the enzyme, influences circulating levels of tHcy. Case-control and prospective studies indicate that this polymorphism is associated with increased risk for NTDs. Such risk could be further increased in individuals who have genetic variants in other folate pathway genes or low vitamin B₁₂ status. For example, the double heterozygosity MTR 2756 A>G/MTRR 66 A>G was a significant risk factor for having Down syndrome.

In addition, it should be recognized that there is no pool of free B₁₂ in the body, because B₁₂ is sequentially bound to the proteins that assist in its transport, delivery, and formation of active coenzyme forms and enzymes. Currently, more than half a dozen proteins (cblA-cblG) that make active MS and malonyl-CoA mutase have been identified owing to studies of patients with severe inborn disturbances in B₁₂ metabolism. There is little information about the function of these proteins and even less about their polymorphisms, which could modulate the B₁₂ function.

Folate Trap

The reaction catalyzed by MS uses the methyl group donated by 5-methyl-THF, which is converted to THF in this reaction. This is the only pathway using 5-methyl-THF because the reaction of the 5,10-methylene-THF to 5-methyl-THF reduction (Fig. 39.1) is irreversible. Because B₁₂ is required for MS activity, the vitamin deficiency impairs 5-methyl-THF use and results in its accumulation at the expense of other folate forms. This metabolic derangement is called methyl trap or folate trap and can be associated with B₁₂ deficiency as well as low activities of MS or MTRR enzymes. Whether mutations or common genetic variants in corresponding genes can trap methyl groups has not yet been experimentally tested. However, the methyl trap, can be at least partially compensated for by increased folate intake.

ENZYMES OF FOLATE METABOLISM AND COMMON GENETIC VARIANTS

In contrast to B₁₂, folate malabsorption rarely causes disease. Instead, folate metabolism is affected by

polymorphisms in genes coding for folate-metabolizing enzymes. Genetic variants in these genes are common and numerous and are found in all key genes involved in folate metabolism. Epidemiologic studies attempted to link these variants to NTDs, different types of cancer, and cardiovascular and neurologic disease. So far, the most characterized genetic variants in folate metabolism are those found in methylenetetrahydrofolate reductase (MTHFR) and dihydrofolate reductase (DHFR). Genetic variants in MTHFR mostly affect methylation-dependent reactions whereas variants in DHFR affect the whole OCM owing to the strategic position of the enzyme in this pathway (Fig. 39.1).

Methylenetetrahydrofolate Reductase

MTHFR (5,10-methylene-THF reductase) catalyzes the conversion of 5,10-methylene-THF to 5-methyl-THF (Fig. 39.1). To better understand the disposition of this reaction in folate metabolism, it should be pointed out that it is strictly irreversible. Therefore, the only way to use 5-methyl-THF produced by MTHFR is by the B₁₂-dependent remethylation of homocysteine into methionine. As discussed earlier, this reaction catalyzed by MS has a dual significance. First, it is an important source of methionine in the cell. Although methionine has several functions, a large portion of this amino acid is used to biosynthesize SAM, a universal methyl donor involved in more than 100 methylation reactions in the cell. Second, folate-dependent methionine biosynthesis contributes significantly to the removal of homocysteine. This could be an important function because homocysteine accumulation generally has a negative effect in humans and is associated with increased risk for cardiovascular diseases. In agreement with this phenomenon, a common MTHFR polymorphism, 677C > T, is associated with mild hyperhomocysteinemia and increased risk for coronary heart disease and stroke. These conditions are exacerbated in individuals with low folate status. The 677C > T polymorphism, which is the most common inborn error of folate metabolism, results in an amino acid alteration in the catalytic domain of the enzyme. The mutant MTHFR is a less stable protein that results in an overall decrease in MTHFR activity in affected individuals. Importantly, 5-methyl-THF was shown to have a protective effect on enzyme stability.

The 677C>T MTHFR variant is common in European and American populations; the homozygous form is found in 5%–20% of individuals. Its clinical manifestation is not limited to hyperhomocysteinemia but includes lower levels of folate in plasma and red blood cells (RBC), a likely contributing factor to the associated pathologies, which could be diverse. Thus, in addition to cardiovascular disease, this MTHFR SNP was implicated in the

increased risk for neural tube defects, cleft lip and palate, Down syndrome, thrombosis, schizophrenia, depression, adverse pregnancy outcomes, and cancer. Interestingly, this mutation may have a protective effect against some cancers, although mechanisms underlying this phenomenon remain elusive. Numerous studies have also shown that mild MTHFR deficiency resulting from the 677C>T polymorphism may be corrected by dietary folic acid. At the same time, severe deficiency of the enzyme can be rescued by 5-methyl-THF, but not by folic acid. In line with this finding, 5-methyl-THF was shown to increase plasma folate more efficiently than folic acid in women of either 677CC or 677TT genotype. Overall, over 50 naturally occurring mutations have been identified in the *MTHFR* gene, as well as nine common polymorphisms, but 677C>T remains most studied and recognized. Another relatively common *MTHFR* polymorphism, 1298A>C, is found in a homozygous state in 4%–12% of individuals in different populations. This variant has reduced activity, but not as dramatically as the 677C>T mutant (only to 68% of wild-type [WT] activity), and it is not thermolabile. Individuals homozygous for this SNP have normal tHcy, not different from subjects with the WT variant. The compound heterozygotes for the 1298C and 677T variants have biochemical profiles similar to those of 677T homozygotes as well as increased homocysteine. This is the why numerous studies are testing both 677C>T and 1298A>C polymorphisms for association with diseases.

Dihydrofolate Reductase

DHFR is a key enzyme in folate metabolism that incorporates ingested folic acid into the reduced folate pool. Folic acid is a synthetic provitamin that requires enzymatic activation to become an active cofactor. DHFR catalyzes the two-step reduction of folic acid, first to dihydrofolate and then to the active coenzyme THF (Fig. 39.1). The enzyme is also responsible for the conversion of dihydrofolate, which is produced in the reaction of thymidine monophosphate biosynthesis, to THF (Fig. 39.1). Homozygous knockout of *Dhfr* in mice was shown to be embryonic-lethal. In humans, DHFR activity can vary at least sixfold between individuals, and generally is lower than in rodents. Low activity of the enzyme results in a low rate of dietary folic acid incorporation into folate pool and higher levels of unmetabolized folic acid in serum. Although negative effects of unmetabolized folic acid can be diverse and indirect, the low rate of incorporation of the vitamin to the active folate pool creates symptoms of folate deficiency even under conditions of high folic acid intake.

A well-established *DHFR* polymorphism is a 19–base pair (bp) deletion in the first intron, 60 nucleotides downstream from the splice donor site. This

polymorphism is highly prevalent with the reported frequency of the del/del genotype reaching 48% in some populations (20% del/del homozygosity in the US population). Because this is an intronic mutation, its likely effect is associated with transcriptional or posttranscriptional regulation of DHFR expression, which could change the protein level. Although this mechanism has not been explored, it has been demonstrated that the del/del genotype is associated with increased unmetabolized folic acid in plasma for intakes exceeding 500 µg/d (high folate consumption), and with decreased RBC folate at intakes below 250 µg/d (low folate consumption). The effect of this polymorphism on NTDs, carcinogenesis, and cognitive function has also been investigated. Data on its role in NTDs and cognitive function are conflicting or incomplete, but there is compelling evidence that this mutation is associated with an increased cancer risk, which is further modified by the folic acid intake: (1) the *DHFR* 19del allele is associated with greater breast cancer risk among multivitamin users; and (2) del/del homozygous mothers taking folic acid supplements are at a higher risk for having a child with early childhood retinoblastoma. Although the molecular mechanisms underlying the biological effects of this genetic variant are not yet known, studies to date suggest that this is a functional polymorphism that alters the incorporation of folic acid into the intracellular pool of reduced folate. Most likely, the deletion diminishes the capacity of affected individuals to make active folate coenzymes from provitamin folic acid. This brings a paradoxical conclusion that a significant part of the population with the del/del genotype who have low folate status will not benefit from the increased intake of folic acid. These findings emphasize another important issue with dietary folate: the form of the vitamin used for supplementation.

Forms of Dietary Folate

Natural folate obtained from a nonfortified diet is presented mostly as 5-methyl-THF. In contrast, the component of multivitamin supplements and food fortification, folic acid, is a synthetic compound not found in the nature. It is a provitamin, which requires two rounds of enzymatic reduction to THF to become an active cofactor. Because folic acid and 5-methyl-THF are functionally different, the efficiency of the natural sources of folate versus folic acid should be considered. Cross-sectional analysis of numerous studies underscores the importance of the natural food folate. The advantage of 5-methyl-THF is its higher absorption rate and higher efficiency in increasing the RBC and plasma folate than folic acid. Specific concerns of supplementation with folic acid include masking of B₁₂ deficiency, promotion of cancer and metastasis, epigenetic changes, and

presence of unmetabolized folic acid in the circulation. So, should the supplementation with 5-methyl-THF be considered instead of folic acid? Two aspects have to be considered. Folic acid is proven to reduce the incidence of NTDs, however there are no such data for 5-methyl-THF. Furthermore, as discussed in this chapter, the incorporation of 5-methyl-THF into the folate pool requires the B₁₂-dependent remethylation of homocysteine. If this process is impaired owing to B₁₂ deficiency, or by polymorphisms in the *MTR/MTRR* or some other genes, supplementation will be ineffective. Another viable alternative to folic acid could be supplementation with folinic acid (5-formyl-THF), which is used to treat cancer patients in combination with 5-fluorouracil. Folinic acid rapidly incorporates into the reduced folate pool upstream of MTHFR, thus avoiding both the folate trap and the necessity to be activated by DHFR. Folinic acid has been shown to bypass the DHFR deficiency effectively by correcting hematological abnormalities, normalizing the cerebrospinal fluid folate levels, and improving neurological symptoms in patients homozygous for L80F or D153V DHFR mutations, both of which result in reduced enzyme activity.

VITAMIN B₆-RELATED FOLATE PATHWAY ENZYMES

Serine Hydroxymethyltransferase

Serine hydroxymethyltransferase (SHMT) synthesizes glycine from serine. This is a reversible reaction, but thermodynamics favors the biosynthesis of glycine (Fig. 39.1). The reaction uses two coenzymes: THF, which accepts OCG from serine, and PLP (a derivative of vitamin B₆, Fig. 39.4), which is absolutely required for enzyme catalysis. This reaction is important for bringing OCG to folate metabolism and also as the source of glycine in the cell. In fact, up to 70% of glycine could come from this reaction and not from dietary sources. There are two identical SHMT reactions in the cell, one in cytoplasm and another in mitochondria, both catalyzed by highly similar enzymes, cytosolic SHMT1 and mitochondrial SHMT2. Interestingly, vitamin B₆ restriction decreases the activity and stability of SHMT, but the cytoplasmic isozyme is more sensitive to vitamin B₆ deficiency than the mitochondrial isozyme. A polymorphism in SHMT1, 1420 C > T, has been described. This is an exonic mutation that results in the substitution of leucine by phenylalanine at position 474 of the protein (L474F). This is a common SNP; the reported frequency for the TT genotype is about 12%. Although this mutation is not believed to alter enzyme activity, it prevents the enzyme from sumoylation and thereby prevents its transport into the nuclei. It has been observed that individuals with the 1420TT

genotype have increased RBC and plasma folate levels. The 1420TT genotype is also associated with the reduction of risk of several cancers and some other diseases. This effect was linked to decreased DNA damage, observed in the TT genotype. Polymorphisms in *SHMT2* are rarer and are not yet linked with the risk for any disease.

Cystathionine β-Synthase

Cystathionine β-synthase (CBS) is a key enzyme in the two-step biosynthesis of cysteine from homocysteine and serine and requires vitamin B₆ for its catalysis (Fig. 39.1). This is the only pathway in humans that leads to cysteine production. This pathway competes for homocysteine with homocysteine remethylation by MS in the methyl cycle pathway. Over 100 variations have been identified in the *CBS* gene. Many of them cause decreased enzyme activity, whereas for about a quarter of mutations an effect has not yet been established. Deficiency of CBS leads to homocystinuria, a rare autosomal recessive inborn error of sulfur amino acid metabolism. Some of the mutations can be corrected by pyridoxine (pyridoxine-responsive) whereas others are pyridoxine nonresponsive. There is at least one mutation in the *CBS* gene that produces hyperactive enzyme (up to 10-fold higher activity compared with the WT enzyme). In addition to rare mutations in *CBS*, several common polymorphisms in this gene were identified. One of these polymorphisms, a 68-bp insertion in exon 8, has a high prevalence (12% in the healthy US population, but varies between 8% and 26% depending on race, whereas the prevalence among healthy men and women in Northern Ireland is 18%). By itself, this mutation might not have a significant effect on circulating homocysteine levels. Nevertheless, it might have an effect in combination with the other common variant genotypes in folate pathways, *MTHFR* 677TT and *MS* 2756 AA. The 68ins polymorphism is also associated with decreased schizophrenia risk. Finally, CBS has two binding sites for SAM, and binding of SAM activates the enzyme. Because SAM levels are linked to folate metabolism (as discussed earlier), this provides an additional regulatory circuit between B vitamin-related pathways and adds to the overall complexity of dietary interventions. Deleterious mutations within SAM regulatory sites will impair CBS activation, thus altering folate-dependent responses.

CONCLUSIONS

Correction of metabolic folate deficiencies caused by specific gene variations through enhanced supplementation is an attractive and feasible approach to the treatment or prevention of diseases. Mandatory food fortification

with folic acid has proven to decrease the incidence of NTDs. However, the outcome of supplementation approach will depend on many factors including the ingested form of the vitamin, the status of other B vitamins, and the overall complex interaction between enzymes in the pathway, as well as between pathways. Considering that functional polymorphisms in folate enzyme genes are rather common, their combinations can produce multiple effects. It could be expected that some genetic variants synergize to exacerbate metabolic derangement, whereas others perhaps can compensate for each other. The situation is complicated further as currently, the effects of a number of polymorphisms on the corresponding protein function have not been established. The entire field is still in its infancy and future studies will provide more precise knowledge of how genotypes affect the human body's ability to process the nutrients and how deficient genotypes can be corrected by dietary manipulations.

Suggested Readings

- Cario, H., Smith, D.E., Blom, H., Blau, N., Bode, H., Holzmann, K., Pannicke, U., Hopfner, K.P., Rump, E.M., Ayric, Z., Kohne, E., Debatin, K.M., Smulders, Y., Schwarz, K., 2011. Dihydrofolate reductase deficiency due to a homozygous DHFR mutation causes megaloblastic anemia and cerebral folate deficiency leading to severe neurologic disease. *Am J Hum Genet* 88 (2), 226–231.
- Cheng, T.Y., Makar, K.W., Neuhauser, M.L., Miller, J.W., Song, X., Brown, E.C., Beresford, S.A., Zheng, Y., Poole, E.M., Galbraith, R.L., Duggan, D.J., Habermann, N., Bailey, L.B., Maneval, D.R., Caudill, M.A., Toriola, A.T., Green, R., Ulrich, C.M., 2015. Folate-mediated one-carbon metabolism genes and interactions with nutritional factors on colorectal cancer risk: women's health initiative observational study. *Cancer* 121 (20), 3684–3691.
- Green, R., Miller, J.W., 2005. Vitamin B₁₂ deficiency is the dominant nutritional cause of hyperhomocysteinemia in a folic acid –fortified population. *Clin Chem Lab Med* 43 (10), 1048–1051.
- Kozich, V., Sokolová, J., Klatovská, V., Krijt, J., Janosík, M., Jelínek, K., Kraus, J.P., 2010. Cystathionine beta-synthase mutations: effect of mutation topology on folding and activity. *Hum Mutat* 31 (7), 809–819.
- Kurnat-Thoma, E.L., Pangilinan, F., Matteini, A.M., Wong, B., Pepper, G.A., Stabler, S.P., Guralnik, J.M., Brody, L.C., 2015. *Biol Res Nurs* 17 (4), 444–454.
- Leclerc, D., Sibani, S., Rozen, R. Molecular biology of methylenetetrahydrofolate reductase (MTHFR) and overview of mutations/polymorphisms. In: *Madame Curie Bioscience Database [Internet]*, Landes Bioscience, Austin (TX), 2000-2013.
- Molloy, A.M., 2012. Genetic aspects of folate metabolism. *Subcell Biochem* 56, 105–130.
- Philip, D., Buch, A., Moorthy, D., Scott, T.M., Parnell, L.D., Lai, C.Q., Ordovás, J.M., Selhub, J., Rosenberg, I.H., Tucker, K.L., Troen, A.M., 2015. Dihydrofolate reductase 19-bp deletion polymorphism modifies the association of folate status with memory in a cross-sectional multi-ethnic study of adults. *Am J Clin Nutr* 102 (5), 1279–1288.
- Scaglione, F., Panzavolta, G., 2014. Folate, folic acid and 5-methyltetrahydrofolate are not the same thing. *Xenobiotica* 44 (5), 480–488.
- Selhub, J., Rosenberg, I.H., 2016. Excessive folic acid intake and relation to adverse health outcome. *Biochimie* 126, 71–78.
- Stover, P.J., 2012. Polymorphisms in 1-carbon metabolism, epigenetics and folate-related pathologies. *J Nutrigenetics Nutrigenomics* 4 (5), 293–305.
- Tibbetts, A.S., Appling, D.R., 2010. Compartmentalization of Mammalian folate-mediated one-carbon metabolism. *Annu Rev Nutr* 30, 57–81.
- Watkins, D., Ru, M., Hwang, H.Y., Kim, C.D., Murray, A., Philip, N.S., Kim, W., Legakis, H., Wai, T., Hilton, J.F., Ge, B., Doré, C., Hosack, A., Wilson, A., Gravel, R.A., Shane, B., Hudson, T.J., Rosenblatt, D.S., 2002. Hyperhomocysteinemia due to methionine synthase deficiency, cblG: structure of the MTR gene, genotype diversity, and recognition of a common mutation, P1173L. *Am J Hum Genet* 71 (1), 143–153.
- Woeller, C.F., Anderson, D.D., Szebenyi, D.M., Stover, P.J., 2007. Evidence for small ubiquitin-like modifier-dependent nuclear import of the thymidylate biosynthesis pathway. *J Biol Chem* 282 (24), 17623–17631.

The Nutrigenetics of Choline

Steven H. Zeisel

Nutrition Research Institute, School of Public Health and School of Medicine, The University of North Carolina at Chapel Hill, Kannapolis, NC, United States

Glossary

Adequate intake (AI)
 Creatine phosphokinase (CPK)
 Methylene tetrahydrofolate reductase (*MTHFR*)
 5,10-Methylenetetrahydrofolate dehydrogenase (*MTHFD1*)
 Nonalcoholic fatty liver disease (NAFLD)
 Single nucleotide polymorphisms (SNPs)
 Tolerable upper limit (UL)
 US Food and Drug Administration (FDA)
 US National Health and Nutrition Examination Survey (NHANES)

INTRODUCTION

Biochemists and nutritional scientists normally think about metabolism as if all people are the same. Dietary recommendations consider age- and gender-related differences in nutrient metabolism but assume that within these subgroups there is a single bell-shaped requirement/response curve that describes metabolism and requirements/responses to nutrients. For many nutrients, we now realize that a better characterization of requirement for a nutrient can be achieved by using multiple curves, each describing subgroups with significant and common sources of metabolic variation. There are many reasons that explain why a person might have metabolic alterations that make them less efficient (or possibly more efficient) in the metabolism or utilization of a specific nutrient: genetic, epigenetic, and microbiome variations being among the most common. In this chapter, the focus is on the genetic mechanisms whereby metabolic variation in the requirements for the nutrient choline occurs.

NUTRIGENETICS AND METABOLIC VARIATION

The study of how gene variations influence metabolism (nutrigenetics) is a rapidly developing new

discipline within nutrition. Each of us has approximately 45,000 variant nucleotides (single nucleotide polymorphisms, SNPs; simply put, there are common variant spellings of our genetic code) and 107,000 small insertions or deletions in our genetic code (sometimes inserting extra copies of genes; copy number variations). These gene variants are inherited from our ancient ancestors. There is an official catalog of SNPs (dbSNP database) that uniquely identifies each location on chromosomes where a variant occurs and designates each with a numerical identifier “rs” followed by a number (reference SNP cluster ID). Using this nomenclature, one can avoid confusion about which gene variation is being discussed. If these SNPs occur within key genes of metabolism, or in the regions of chromosomes that regulate the expression of these key genes, then metabolic alterations can occur that change the requirements or responses to a specific nutrient. It is not uncommon that we find a nutrient-disease association in one study, and subsequently do not observe this relationship in another study. This variability in conclusions is not only due to random biological noise but also to inclusion of study subjects who metabolically come from genetically definable subpopulations with widely differing responses to the nutrient. Genome-wide-association studies (GWAS), in which genetic variations are associated with a risk for a disease in large populations, have identified very few such metabolic roadblock SNPs. This is because, for genes that exert effects on a metabolic pathway, a gene variant that adds only a small amount of metabolic inefficiency is often difficult to distinguish from background variation. In addition, most GWAS collect limited dietary intake information for the persons studied. Without such dietary data, people eating diets high in a nutrient are lumped with people who eat diets with low intake of the nutrient. If there is a metabolic inefficiency caused by an SNP, it may only be important in people who are challenged by low

intake of the nutrient, while people eating diets rich in the nutrient manage to compensate by eating enough of the nutrient to push enough substrate through the metabolic inefficiency. When people eating diets rich in the nutrient are combined with people eating little of the nutrient, the effects of the SNP are averaged out and may not appear to be statistically significant; however, if the diet groups can be separated, the effects of the SNP may be very significant in the diet-challenged group.

Studies on how SNPs influence the dietary requirement for choline provide a good example of how SNPs can modify nutritional phenotypes.

CHOLINE

Choline has several important functions. It is (1) a source of methyl groups needed to make the important methyl donor, *S*-adenosylmethionine; (2) a part of the neurotransmitter acetylcholine; and (3) a component of several major phospholipids in membranes (phosphatidylcholine and sphingomyelin) (Zeisel, 2011). Choline is a required nutrient for people, and in 1998, an adequate intake (AI) level and a tolerable upper limit (UL) level for choline were established (Zeisel, 2011). In 2016, the US Food and Drug Administration (FDA) set a Recommended Daily Intake (RDI) level for choline as part of the new Nutrition Facts label for packaged foods (published in the Federal Register on May 27, 2016; FDA-2012-N-1210-0875, Federal Register Number:2016-11867). The AI/RDI varies by age and gender, but is 550 mg/d in adult men and 425 mg/d in adult women (more in pregnant and lactating women).

There is an approximately 2.5-fold variation in choline intake in the diet (Zeisel, 2011); the US National Health and Nutrition Examination Survey (NHANES, 2009–12) found that only 11% of adult Americans achieved the AI level of choline, with mean intake being 300 mg/day (10% being 200 mg/d; 90% being 500 mg/d). Similar ranges of intake were reported in the Framingham Offspring Study, the Atherosclerosis Risk in Communities study, and the Nurse's Health Study. Intake of choline is even lower in low-income countries such as Jamaica and The Gambia, in West Africa.

Diets low in choline result in fatty liver, liver damage, and muscle damage in people (Zeisel, 2011). Risk for developing nonalcoholic fatty liver disease (NAFLD), that occurs as people become obese, can be predicted by a genetic signature (pattern of SNPs) in choline-related genes (Corbin et al., 2013) (see later discussion), and higher dietary intake of choline reduces the risk of developing fatty liver (Zeisel, 2011). Adults and infants receiving total parenteral nutrition can have low plasma choline concentrations and often develop hepatic

abnormalities, including NAFLD (Buchman et al., 2001). At present, choline is not routinely added to commercial parenteral solutions for infants and adults. In men, dietary intakes in excess of the choline RDI-AI are needed to optimize homocysteine disposal after a methionine load as well as for the removal of fat from liver (Veenema et al., 2008). Women eating diets low in choline during pregnancy (about 200 mg/d) are more likely to give birth to a child with birth defects than are women eating 500 mg/d (Shaw et al., 2009). However, other studies have found no relationship between plasma or serum choline concentrations during pregnancy and neural tube defects in offspring (likely because plasma choline is not an adequate marker of choline status).

Higher choline intake in pregnant women in Massachusetts was associated with modestly (but significantly) better visual memory in the children (born of the pregnancy) at age 7 years (Boeke et al., 2013). Beyond the period of pregnancy, low dietary intake of choline (about 150 mg/d) was associated with decreased cognitive function in the Framingham Offspring Cohort (Poly et al., 2011). In Swedish teenagers, plasma choline concentration was significantly and positively associated with academic achievement independent of socioeconomic status (SES) factors (paternal education and income, maternal education and income, smoking, school) and of folate intake (Nilsson et al., 2016). In Norwegian septuagenarians, people with plasma choline concentrations lower than 8.4 $\mu\text{mol/L}$ (20th percentile for concentrations in the study population) had poorer sensorimotor speed, perceptual speed, executive function, and global cognition than did those with plasma choline concentrations higher than 8.4 $\mu\text{mol/L}$ (Nurk et al., 2013).

Diets low in choline increased the overall relative risk for developing cancer in a quantitative meta-analysis of 11 papers published on this topic (Sun et al., 2016), with the largest reported effects found for lung (30% increase), nasopharyngeal (58% increase), and breast cancer (60% increase). An increment of 100 mg/day of choline and betaine (a metabolite derived from choline) intake helped reduce cancer incidence by 11% (Sun et al., 2016). Other studies also report that lower dietary intake of choline was associated with an increased risk of breast cancer and colorectal cancer. At the same time, diets high in choline may be associated with an increased risk for prostate cancer progression, for colorectal adenomas, and for heart disease (Tang et al., 2015). Analysis of data on 80,978 women from the Nurses' Health Study and 39,434 men from the Health Professionals Follow-up Study found an increased risk of mortality in those consuming higher levels of choline (Zheng et al., 2016). Thus, it appears that diets at the lower end of normal intake for choline have adverse health consequences,

while diets at the higher end of normal intake also have adverse consequences. This U-shaped risk-benefit curve defines a relatively narrow range for optimal choline intake, making it important that we can more precisely identify people who have differing requirements for choline using nutrigenetic measures.

CHOLINE NUTRIGENETICS

A large number of genes code for enzymes important in choline metabolism (Zeisel, 2011). Choline is transported into cells (and mitochondria) by a bidirectional choline transporter (gene *SLC44A1*) that traffics choline across the cell membrane and across the mitochondrial inner membrane. Once in the cell, choline can be phosphorylated to form phosphocholine (catalyzed by choline kinase; genes *CHKA* and *CHKB*), which, after a series of reactions is eventually converted to phosphatidylcholine. There is an alternative pathway whereby phosphatidylcholine is formed, involving the triple methylation of phosphatidylethanolamine, catalyzed by phosphatidylethanolamine-*N*-methyltransferase (gene *PEMT*). It is important that this is the only pathway where new choline moiety (as part of phosphatidylcholine) can be formed—the other pathway uses preformed choline to make phosphatidylcholine. This preformed choline moiety, to a large extent, comes from the diet. Choline can be oxidized to form betaine, catalyzed by choline dehydrogenase (gene *CHDH*) located in the mitochondria. A methyl group of betaine is then used to methylate homocysteine to form methionine, catalyzed by betaine-homocysteine methyltransferase (gene *BHMT*) in the cytosol. In a parallel pathway, methionine synthase, a 5-methyltetrahydrofolate-requiring enzyme, can methylate homocysteine, and thereby reduce the demand for betaine (and thereby choline) for use as a methyl donor. Thus, genes of folate metabolism that modulate the generation of 5-methyltetrahydrofolate can be important determinants of choline demand (genes such as 5,10-methylenetetrahydrofolate dehydrogenase [*MTHFD1*] and methylene tetrahydrofolate reductase [*MTHFR*]).

A number of common SNPs in genes related to choline metabolism alter the requirement for dietary choline (Zeisel, 2011). Thus, depending on genotype, similar dietary intakes may lead to very different choline status (the amount of choline needed for optimal organ function). The dietary requirement for choline is highest for men and postmenopausal women, whereas premenopausal women may have a reduced dietary requirement for choline. Men have little estrogen, postmenopausal women make small amounts of estrogen, while premenopausal women, especially pregnant women, make much more estrogen. *PEMT* expression (important for

de novo formation of choline moiety that does not have to be derived from diet) is increased by estrogen; with maximum induction at the concentrations achieved in women during pregnancy. Women with the SNPs *PEMT* rs12325817-C (meaning they have a cytosine [C] at this location on the gene) or *PEMT* rs4646365-A (adenine [A] at this location) are more susceptible to liver damage when consuming a low-choline diet because these SNPs prevent the estrogen response element in the promoter of the gene from responding to estrogen, thereby blocking estrogen-induced increased expression of the gene. Other *PEMT* SNPs also increase choline requirement; rs3760188-A in all women, rs1531100-A and rs7946-A in postmenopausal women. The three *PEMT* SNP alleles that were associated with liver dysfunction when women consumed a low-choline diet (rs12325817-C, rs4646343-A, and rs3760188-A) are in linkage disequilibrium (these SNPs are often inherited together). These 3 *PEMT* SNPs are located within a 5 Kb region of the first intron of the gene, in the likely promoter region that regulates expression of the gene.

The prevalence of these *PEMT* SNPs varies with ancestry: about 24% of the European Americans (EA), 11% of the Mexican Americans (MA), and 1% and 0% of the Asian Americans (AsA) and the group of African descendants (AD), respectively, were homozygous for *PEMT* rs12325817-C. *PEMT* rs7946-A has a very different distribution in the population groups. In the AsA and AD groups, the A allele was present in 51% and 30% of people, respectively. However, over 80% of the EA and MA individuals had the A allele—with 60% and 30%, respectively, being homozygous for this allele of the SNP.

As noted earlier, there are other SNPs in genes of choline and folate metabolism that modify the dietary need for choline. Women homozygous for one of three SNPs in *CHKA* (product catalyzes the phosphorylation of choline) rs10791957-C, rs7928739-C, and rs2512612-G (guanine [G] at this location)) were protected when they were fed a low-choline diet, while postmenopausal women who were carriers of *CHKA* rs6591331-T (thymine [T] at this location) were more likely to develop organ dysfunction when fed a low-choline diet (da Costa et al., 2014). *CHKA* rs10791957-C and rs6591331-T are located in the first intron of the gene, in a region that may play a regulatory role as an enhancer, increasing the transcription levels of the gene (UCSC Genome browser, <http://genome.ucsc.edu>, H3K27Ac tracks).

SNPs in the choline dehydrogenase gene (*CHDH*; needed to convert choline to betaine) alter the dietary choline requirement; an SNP in the coding region of *CHDH* (rs9001-C) had a protective effect on susceptibility to choline deficiency, while a second *CHDH* variant (rs12676-T) was associated with increased susceptibility

to choline deficiency. It is possible that one of these SNPs results in less *CHDH* activity, while the other results in more.

Men and women who were carriers of the very common SNP rs2236225-A in *MTHFD1* (a gene for an enzyme that directs folate either toward use in DNA synthesis or toward use in methylation reactions) were more likely than noncarriers to develop signs of choline deficiency on a low-choline diet unless they were also treated with a folic acid supplement (Kohlmeier et al., 2005); individuals who were homozygous carriers of rs1801133-T in *MTHFR* (a gene for an enzyme that is needed to make 5-methyltetrahydrofolate) had increased dietary choline requirements to maintain normal plasma homocysteine concentrations.

Genotype also modifies the organ most affected by a low-choline diet. In one study of people that developed organ dysfunction on a low-choline diet, 13% presented first with muscle damage (elevated creatine phosphokinase) while 87% presented first with liver damage (elevated aspartate transaminase); both types of organ damage resolved when choline was returned to the diet (Zeisel, 2011). People who presented with muscle rather than liver dysfunction had SNPs in genes that influenced choline availability and metabolism in muscle. They were more likely to have the SNP rs7873937-G in *SLC44A1* (codes for the choline transporter), and 9 of 10 of these individuals also had *SLC44A1* rs2771040-G and *SLC44A1* rs6479313-G (these two SNPs are in linkage disequilibrium). People with *SLC44A1* rs440290-G or rs3199966-C also were more likely to develop muscle damage when fed a low-choline diet (da Costa et al., 2014). Only *SLC44A1* rs3199966-C is exonic, causing an amino acid change from serine to alanine that can affect the function of the transporter. *SLC44A1* rs2771040-G is located in another important site for regulating gene expression, the 3'-UTR or untranslated region of the gene. The C allele of rs3199966 only occurs together with rs2771040 G allele on the same strand. The prevalence of *SLC44A1* SNPs depends on ancestry: *SLC44A1* rs7873937-G SNP is present in ~24% of the EA, MA, and AD groups, but is not present in the AsA group; in fact, none of the *SLC44A1* SNPs associated with low-choline-related muscle damage were present in the AsA population (da Costa et al., 2014). 71% of those from African lineage had the *SLC44A1* rs2771040-G SNP compared with 24% of the EA group and 31% of the MA group (da Costa et al., 2014).

In addition, 9 of 10 people who developed muscle damage when fed a low-choline diet were heterozygous or homozygous carriers of both *SLC44A1* rs2771040-G and another SNP, rs1557502-A in *CHKB* (needed for

choline phosphorylation, which traps choline as a charged molecule within the cell). These two polymorphisms were not correlated in people who did not develop muscle damage when fed a low-choline diet (da Costa et al., 2014). It is interesting that these two SNPs in *SLC44A1* and *CHKB* (on two different chromosomes) are linked; perhaps a variant at one locus manifests its effect only when the variant at the other locus is present. The prevalence of *CHKB* SNPs depends on ancestry: 88% of the AD group had at least one variant allele for the *CHKB* rs1557502-A SNP (38% were homozygous), as well as did more than half of the AsA and MA groups, while only 27% of the EA group had one variant A allele of this SNP (da Costa et al., 2014).

As noted previously, many of the SNPs in genes of choline metabolism are less frequent in people of African descent when compared to people of European ancestry. Could it be that traditional diets in some areas of Africa delivered little choline and resulted in evolutionary pressure that eliminated SNPs, causing inefficient choline utilization, while traditional diets in Europe delivered high choline (every mother fed her children eggs, a good source of choline) and exerted no such negative selection pressure? We tested the hypothesis that three SNPs that increase dependence on dietary choline (*CHDH* rs12676-T, *MTHFD1* rs2236225-A, and *PEMT* rs12325817-C) would be under such negative selection pressure in The Gambia, where there is historic evidence of a choline-poor diet, versus in the United States (European descent) where there is a comparatively choline-rich diet. We found that frequencies of SNPs known to increase dependence on dietary choline are significantly reduced in the low-choline setting of The Gambia, while SNPs located nearby on the chromosome that are known to have no effect on choline requirements were not reduced compared to that seen in Americans (Silver et al., 2015). Also, we found that the Maasai (an East African population with similar genetic background to that of Gambians), who ate a traditional diet (including milk and blood) high in choline, did not have a reduced frequency of these functional SNPs, and in fact, had an SNP prevalence that was similar to that in people of European descent (Silver et al., 2015). Thus, the SNPs of interest disappeared in regions of Africa eating a low-choline diet, rather than spontaneously appeared in Europe. These interesting studies suggest that the study of populations with traditional diets low in a specific nutrient (e.g., low vitamin B12 in some Indian populations) can be useful for identifying SNPs that are functional (cause metabolic inefficiencies that change diet requirements) as they should be subject to negative selection pressure.

NEXT-GENERATION NUTRIGENETICS

All of the above examples of choline nutrigenetics associated one (or a few) functional SNPs with diet requirement changes, rather than examined the additive effects of multiple SNPs across many pathways. If the effects of individual SNPs are relatively small, it may take multiple metabolic inefficiencies in a pathway that are caused by many SNPs to result in a significant observable effect on nutrient requirements or responses. Given enhanced gene sequencing and bioinformatic methods, the next generation of nutrigenetics needs to move on from identifying single gene variants to working with patterns of many gene variations that are associated with altered nutrient requirements. An example of how grouping individuals by shared patterns of SNPs in pathways functionally linked to the phenotype of interest, rather than using traditional single SNP, is provided by our work on genetic signatures for NAFLD (Corbin et al., 2013). We defined SNP signatures using an unsupervised hierarchical clustering approach, based on the number of SNP alleles carried that were associated with severity of steatosis (fatty liver) in specific groups of people. The genetic signatures that we identified involved 260 SNPs across 21 genes of pathways that are known to be mechanistically related to NAFLD—choline metabolism, lipid transport, etc.—including genes such as *CHDH*, *PEMT*, *SLC44A* and *PNPLA3* (patatin-like phospholipase domain containing 3) but also genes that we would not have expected to be associated with fatty liver such as *ABCB4* (ATP-binding cassette, subfamily B member 4). *ABCB4* codes for an enzyme that flips phosphatidylcholine from the inner to the outer hepatocyte canalicular membrane, so that this phospholipid is available for incorporation into bile (Corbin et al., 2013).

When SNPs were examined individually, using logistic regression, we only identified a single SNP (*PNPLA3* rs738409-G) that was significantly associated with severity of hepatic steatosis after adjusting for confounders and multiple comparisons. However, when groupings of SNPs in similar metabolic pathways were defined using unsupervised hierarchical clustering, we identified groups of subjects with shared SNP signatures that were significantly correlated with steatosis burden ($P = .0002$). This is because the interaction of these genetic variants results in metabolic inefficiency due to modest alterations in function of multiple enzymes that cumulatively alter flux through metabolic pathways and add up to a significant effect, even though each individual SNP does not make a large enough contribution such that alone it would be recognized as significant. Sometimes, a genetic signature approach can reveal effects of SNPs that are different than those

seen using analysis of individual SNPs. An example is the effect of *APOC3* apolipoprotein C-III (*APOC3*; a major constituent of VLDL that inhibits lipoprotein lipase); the rs2854117-T SNP has previously been associated with hepatic steatosis (Petersen et al., 2010), although subsequent publications were unable to replicate the original finding. In our study using a genetic signature approach (Corbin et al., 2013), an association with increased NAFLD was seen for a genetic signature that included the rs2854117-C, and not the T allele, which is different from the original finding and suggests a protective effect of the rs2854117-T SNP.

The understanding of genetic polymorphisms and their effects on responses to, and requirements for, nutrients is rapidly developing. Nutrigenetics helps us to gain additional insight into why there is metabolic variation between people. The approach may bring us one step closer to personalized nutrition by allowing people to be stratified by SNP patterns (genetic signatures) to better separate the metabolically inefficient from the metabolically efficient for specific nutrients.

Grant support

This work was supported by a grant from the NIH (DK056350).

References

- Boeke, C.E., Gillman, M.W., Hughes, M.D., Rifas-Shiman, S.L., Villamor, E., Oken, E., 2013. Choline intake during pregnancy and child cognition at age 7 years. *Am J Epidemiol* 177, 1338–1347.
- Buchman, A.L., Sohel, M., Brown, M., Jenden, D.J., Ahn, C., Roch, M., Brawley, T.L., 2001. Verbal and visual memory improve after choline supplementation in long-term total parenteral nutrition: a pilot study. *JPEN - J Parenter Enter Nutr* 25, 30–35.
- Corbin, K.D., Abdelmalek, M.F., Spencer, M.D., da Costa, K.A., Galanko, J.A., Sha, W., Suzuki, A., Guy, C.D., Cardona, D.M., Torquati, A., Diehl, A.M., Zeisel, S.H., 2013. Genetic signatures in choline and 1-carbon metabolism are associated with the severity of hepatic steatosis. *FASEB J* 27, 1674–1689.
- da Costa, K.A., Corbin, K.D., Niculescu, M.D., Galanko, J.A., Zeisel, S.H., 2014. Identification of new genetic polymorphisms that alter the dietary requirement for choline and vary in their distribution across ethnic and racial groups. *FASEB J* 28, 2970–2978.
- Kohlmeier, M., da Costa, K.A., Fischer, L.M., Zeisel, S.H., 2005. Genetic variation of folate-mediated one-carbon transfer pathway predicts susceptibility to choline deficiency in humans. *Proc Natl Acad Sci U S A* 102, 16025–16030.
- Nilsson, T.K., Hurtig-Wennlof, A., Sjostrom, M., Herrmann, W., Obeid, R., Owen, J.R., Zeisel, S., 2016. Plasma 1-carbon metabolites and academic achievement in 15-yr-old adolescents. *FASEB J* 30, 1683–1688.
- Nurk, E., Refsum, H., Bjelland, I., Drevon, C.A., Tell, G.S., Ueland, P.M., Vollset, S.E., Engedal, K., Nygaard, H.A., Smith, D.A., 2013. Plasma free choline, betaine and cognitive performance: the Hordaland health study. *Br J Nutr* 109, 511–519.
- Petersen, K.F., Dufour, S., Hariri, A., Nelson-Williams, C., Foo, J.N., Zhang, X.M., Dziura, J., Lifton, R.P., Shulman, G.I., 2010. Apolipoprotein C3 gene variants in nonalcoholic fatty liver disease. *N Engl J Med* 362, 1082–1089.

- Poly, C., Massaro, J.M., Seshadri, S., Wolf, P.A., Cho, E., Krall, E., Jacques, P.F., Au, R., 2011. The relation of dietary choline to cognitive performance and white-matter hyperintensity in the Framingham offspring cohort. *Am J Clin Nutr* 94, 1584–1591.
- Shaw, G.M., Finnell, R.H., Blom, H.J., Carmichael, S.L., Vollset, S.E., Yang, W., Ueland, P.M., 2009. Choline and risk of neural tube defects in a folate-fortified population. *Epidemiology* 20, 714–719.
- Silver, M.J., Corbin, K.D., Hellenthal, G., da Costa, K.A., Dominguez-Salas, P., Moore, S.E., Owen, J., Prentice, A.M., Hennig, B.J., Zeisel, S.H., 2015. Evidence for negative selection of gene variants that increase dependence on dietary choline in a Gambian cohort. *FASEB J* 29, 3426–3435.
- Sun, S., Li, X., Ren, A., Du, M., Du, H., Shu, Y., Zhu, L., Wang, W., 2016. Choline and betaine consumption lowers cancer risk: a meta-analysis of epidemiologic studies. *Sci Rep* 6, 35547.
- Tang, W.H., Wang, Z., Shrestha, K., Borowski, A.G., Wu, Y., Troughton, R.W., Klein, A.L., Hazen, S.L., 2015. Intestinal microbiota-dependent phosphatidylcholine metabolites, diastolic dysfunction, and adverse clinical outcomes in chronic systolic heart failure. *J Card Fail* 21, 91–96.
- Veenema, K., Solis, C., Li, R., Wang, W., Maletz, C.V., Abratte, C.M., Caudill, M.A., 2008. Adequate intake levels of choline are sufficient for preventing elevations in serum markers of liver dysfunction in Mexican American men but are not optimal for minimizing plasma total homocysteine increases after a methionine load. *Am J Clin Nutr* 88, 685–692.
- Zeisel, S.H., 2011. Nutritional genomics: defining the dietary requirement and effects of choline. *J Nutr* 141, 531–534.
- Zheng, Y., Li, Y., Rimm, E.B., Hu, F.B., Albert, C.M., Rexrode, K.M., Manson, J.E., Qi, L., 2016. Dietary phosphatidylcholine and risk of all-cause and cardiovascular-specific mortality among US women and men. *Am J Clin Nutr* 104, 173–180.

41

Electrolytes

Olivia Dong¹, Tim Wiltshire²¹Duke Center for Applied Genomics and Precision Medicine, Duke University School of Medicine; ²UNC Eshelman School of Pharmacy and UNC Center for Pharmacogenomics and Individualized Therapy, Kannapolis, NC, United States

INTRODUCTION

Electrolytes, such as salts, acids, and bases, are important minerals in the body that aid in numerous physiological processes vital for life. Electrolytes are defined as inorganic compounds able to conduct electrical currents when dissociated in their ionic state. When dissolved uniformly in a polar solvent, the solution is neutral, and when electrolytes are dissociated into ions—cations and anions—they are able to form an electrical charge. An electrical charge or signal is formed when these electrolytes move against their concentration gradient.

The main electrolytes relevant for physiology include sodium, potassium, magnesium, calcium, chloride, hydrogen carbonate, and hydrogen phosphate. Precise control over these osmotic gradients among blood and the intracellular and extracellular environments in the

body is critical in maintaining blood pH, muscle contraction, nerve function, and fluid balance, among many other processes. Hormones, such as antidiuretic hormone, parathyroid hormone, and aldosterone, regulate homeostasis of these ions. Abnormal regulation of these electrolytes can become fatal as they lead to neurological and cardiac complications.

Dietary recommendations help to ensure cellular levels remain within an appropriate range. [Table 41.1](#) summarizes the function and dietary recommendations for sodium, potassium, and chloride.

Many factors, including genetics, contribute to the overall homeostasis of these electrolytes. The focus of this chapter will be on key genetics that influence sodium, potassium, and chloride homeostasis. Of these electrolytes, the genetics of sodium homeostasis has been studied most extensively, as it is linked with the salt-sensitivity phenotype.

TABLE 41.1 A Summary of the Function and Dietary Recommendations for Sodium, Potassium, and Chloride.

Ion	Function	Dietary recommendations
Sodium	Plays a key role in controlling fluid balance in the body, muscle and nerve function, balancing electrolytes, and blood pressure.	Dietary Guidelines: <2300 mg/day for general population <1500 mg/day if >51 years old, black, or presence of hypertension, diabetes, chronic kidney disease
Potassium	Plays a key role in balancing electrolytes, transmitting nerve impulses, muscle contractions, regulating heart and blood pressure, and bone health	Recommended Daily Allowances: 4700 mg/day for >13 years old 5100 mg/day for nursing women
Chloride	Plays a key role in balancing electrolytes, aiding in digestion, balancing acidity and alkalinity, affecting pH	Adequate intake: 0–6 mo: 0.18 g/day 7–12 mo: 0.57 g/day 1–3 y: 1.5 g/day 4–8 y: 1.9 g/day 9–13 y: 2.3 g/day 14–50 y: 2.3 g/day 51–70 y: 2.0 g/day >70 y: 2.3 g/day

SODIUM

Monogenic Disorders Affecting Sodium Homeostasis

A number of rare monogenic variants cause sodium dysregulation. Liddle syndrome (caused by defective epithelial sodium channel, ENaC) is an inheritable salt-sensitive hypertension resulting in gain-of-function variants in ENaC. Glucocorticoid-remediable aldosteronism (GRA, caused by genetic variants of aldosterone synthase) is another inheritable disorder that causes salt-sensitive hypertension by increasing aldosterone secretion. The syndromes of Bartter (caused by defective Na–K–Cl cotransporter function) and Gitelman (caused by defective thiazide-sensitive sodium chloride symporter) lead to salt and water losses in the kidney. EAST syndrome is also genetically based and causes electrolyte imbalances and renal tubule abnormalities.

Sodium: The Salt Sensitivity Phenotype (Felder et al., 2013; Sanada et al., 2011)

Salt sensitivity is defined by a change in blood pressure in response to a change in sodium chloride intake. The exact numeric change in these two measurements to classify an individual as salt sensitive varies within the field. A 5%–10% blood pressure change or a >4 mmHg in mean blood pressure are two accepted cut-offs in response to a change in sodium chloride intake.

Salt sensitivity has also been defined by a >10 mmHg increase in mean blood pressure as a result of a normal 2 L saline (0.9%) infusion over a 4-h period in relation to a low-sodium diet (10 mmol) and loop diuretic administered the next morning after the saline infusion. Despite the numerous ways salt sensitivity can be determined, the most reliable method of assessing salt sensitivity is a crossover study design where a 5–7 d intervention period is given for a normal-, low-, and high-dietary sodium intake level. When a change in blood pressure is not observed despite changes in sodium chloride intake, the individual is classified as salt resistant.

Current standard of practice in the nutrition field intervenes on an individual's sodium intake when blood pressure is elevated. This standard of practice misses an opportunity to intervene and reduce sodium intake for approximately 14% of the population who are salt sensitive but normotensive. In addition, this standard of practice is also ineffective for 16% of individuals who are hypertensive and salt resistant (Felder et al., 2013). Current intervening strategies for elevated blood pressure are appropriate for 11.8% of adults who are both hypertensive and salt sensitive. Excessive sodium intake can cause hypertension, but this threshold for what is considered excessive differs depending on how the body is able to handle sodium homeostasis. This varied response to sodium intake highlights the need to move away from the current “one-size-fits-all” approach in sodium recommendations to a paradigm that is more precise in characterizing sodium response (Fig. 41.1).

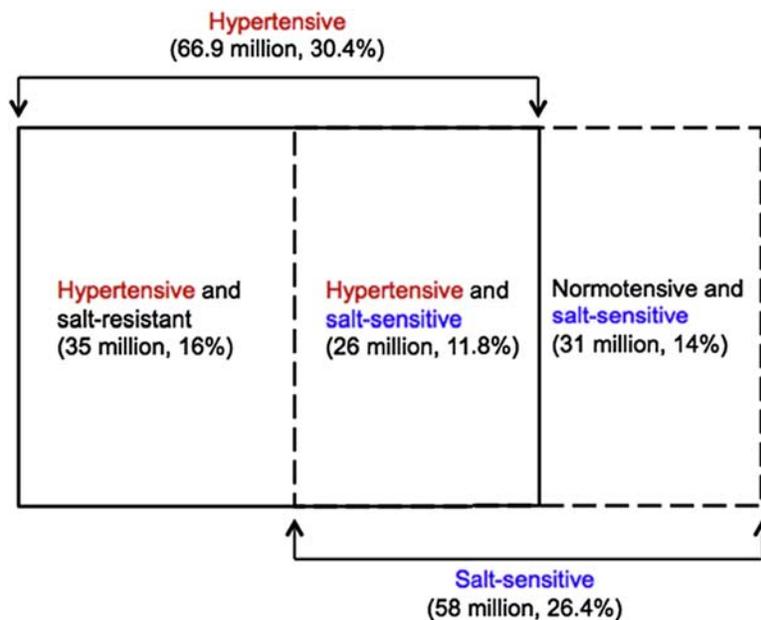


FIGURE 41.1 A breakdown of the adult population (≥ 18 years and older) who are hypertensive, salt sensitive, or both using CDC prevalence data on hypertension from 2003 to 2010 (Felder et al., 2013; CDC, 2012).

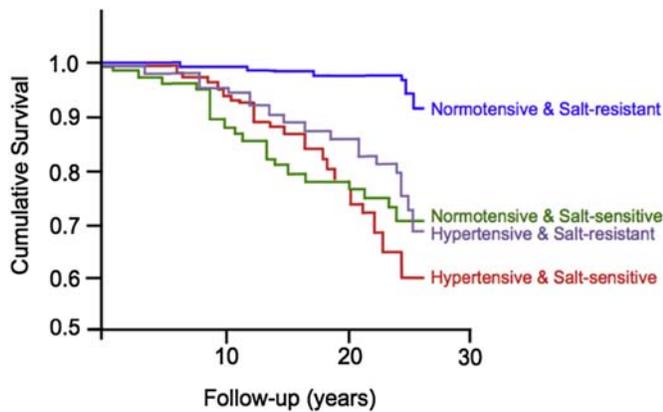


FIGURE 41.2 Cumulative survival over a 27-year period stratified by blood pressure (hypertensive or normotensive and response to salt—salt sensitive or salt resistant) (Weinberger et al., 2001).

Being salt sensitive was found in a 27-year observational study to be an important significant risk factor for mortality independent of blood pressure level that mimics mortality risk observed in hypertensive and salt-resistant individuals (Weinberger et al., 2001). Fig. 41.2 below summarizes the cumulative survival for this cohort, which included 278 hypertensive and 338 salt sensitive individuals. The cumulative survival curve indicates the best survival over the 27-year period was observed in individuals who were normotensive and salt resistant and the worst in those who were hypertensive and salt sensitive. Survival over this 27-year period was not significantly different for those normotensive and salt sensitive from those who were hypertensive. Individuals who were normotensive and salt resistant had significantly better survival ($P < .001$). An odds ratio of 1.73 (CI 1.02 to 2.94, $P = .042$) for death was observed for individuals who were salt sensitive.

How are dietary recommendations impacted given this data on salt sensitivity? Current practice is appropriate for the subset of the population who are salt resistant, but how would dietary recommendations change if we knew how individuals responded to salt? Knowing that salt sensitivity is an independent risk for mortality, it may be beneficial for salt-sensitive individuals to follow a low-salt diet regardless of current blood pressure. For those who are hypertensive, it would be important to know whether blood pressure will be affected by sodium intake. For those who are salt resistant and hypertensive, focusing more on other dietary and lifestyle interventions or on pharmaceutical therapies may result in better control over blood pressure. Being able to tease out someone's response to sodium can potentially better individualize dietary recommendations for optimal health outcomes. Following a restricted sodium diet can be burdensome, so identifying those who would

most benefit from such drastic dietary restrictions can focus health care resources to ensure this subpopulation is identified and managed appropriately. With current sodium intake well above the recommended amount, it would be imperative for salt-sensitive individuals to pay particular attention to this recommendation given their increased chances of mortality.

Genes Associated With the Salt Sensitivity Phenotype

Currently, there are no practical methods to assess this phenotype in the clinical setting as it involves weeks of careful dietary monitoring using a crossover study design. The genetics that characterize this phenotype have been extensively studied in an attempt to use genetic screening tools to predict this phenotype for more precise dietary and pharmaceutical interventions. Numerous genes have been identified and associated with the salt-sensitivity phenotype. This section highlights a small selection of genes that have been investigated for association with the salt-sensitivity phenotype.

The Renin-Angiotensin System

When sodium levels are low, renin will convert angiotensinogen into angiotensin I, which will be subsequently converted into angiotensin II by angiotensin-converting enzyme (ACE). Angiotensin II will cause arterioles to constrict, causing arterial blood pressure to increase. Angiotensin II will also increase aldosterone, a hormone that causes the kidney to increase sodium ion reabsorption (Fig. 41.3).

Mineralocorticoids and the Mineralocorticoids Receptor Cascade

Activation of the mineralocorticoid receptor helps with sodium homeostasis in the body by regulating epithelial sodium channels (ENaC), serum and glucocorticoid induced kinase (SGK1), and Na^+/K^+ pumps. This receptor cascade plays a key role in reabsorbing sodium. Both aldosterone and cortisol activate the mineralocorticoid receptor cascade, which is expressed in various locations within the body, including the kidney and colon (Fig. 41.4).

Beyond the list outlined in Table 41.1, there are many additional genes that have been linked to the salt-sensitivity phenotype. All of these genes and variants are linked to salt sensitivity with varying levels of conclusive evidence. **The strongest and most consistent genetic links to salt sensitivity are variants in the GRK4 and SLC4A5 genes.**

A randomized crossover study studying blood pressure in response to low- and high-salt diets was

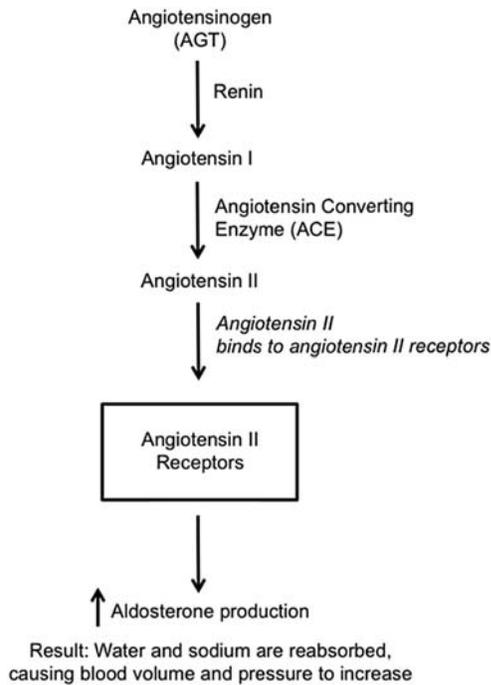


FIGURE 41.3 The renin-angiotensin system (RAS) system pathway. When blood pressure is low or when sodium levels are lower than normal, renin is secreted from the kidneys to convert angiotensinogen into angiotensin I. Then, ACE converts angiotensin I into angiotensin II, which binds to angiotensin II receptors to increase aldosterone secretion. Aldosterone increases the reabsorption of water and sodium to increase blood pressure.

conducted in 185 Caucasians (Carey et al., 2012). Out of all the polymorphisms studied in 17 genes, two single-nucleotide polymorphisms in SLC4A5, rs7571842, and rs10177833, were found to be associated with salt sensitivity (P -value 1.0×10^{-4} and 3.1×10^{-4} , respectively). This association was replicated in a second hypertensive population. G alleles for rs7571842 and rs10177833 protect against salt sensitivity (OR = 0.221 and 0.221, respectively) (Carey et al., 2012). One limitation to note with the SLC4A5 variants is the generalizability since the two populations studied were Caucasian.

The GRK4 variants R65L, A142V, and A486V have been studied in several ethnicities and remain associated with salt sensitivity. In a crossover study, 184 Japanese subjects followed either a normal-, low-, or high-sodium diet to assess salt sensitivity. It was found that Japanese subjects carrying three or more variants at any of the three sites were more likely to be salt sensitive than salt resistant. Using this criterion, 94% of the cohort was correctly classified by their response to sodium (Sanada et al., 2006). The GRK4 A142V variant was the most accurate predictor of salt sensitivity, with a 78.4% accuracy rate among hypertensive Japanese. Replication of these findings in other populations is needed to understand the extent to which these conclusions can be generalized.

It is important to note that variants that are associated with the salt-sensitivity phenotype in one ethnic

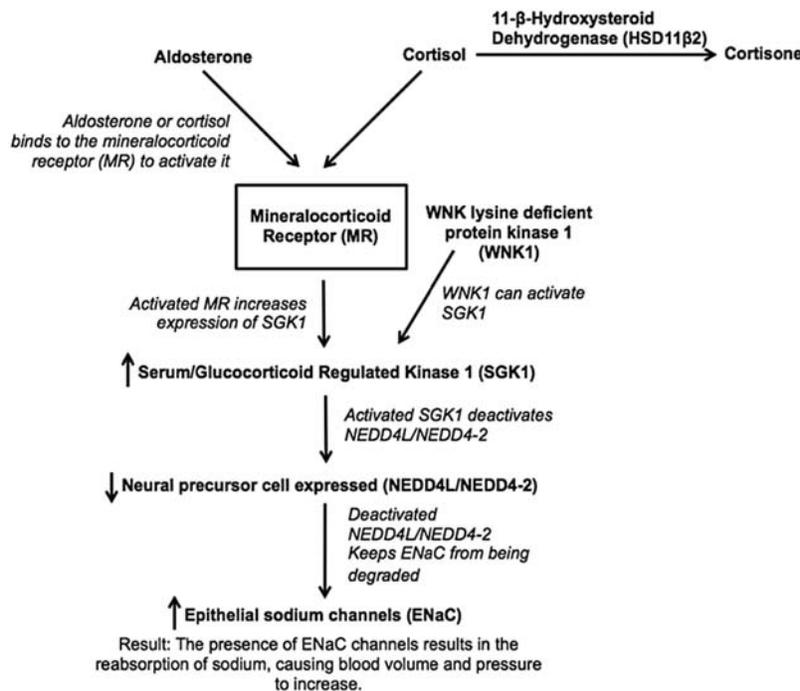


FIGURE 41.4 The mineralocorticoid receptor (MR) pathway. When blood pressure is low or when sodium levels are lower than normal, aldosterone secretion increases. Aldosterone or cortisol can bind to the mineralocorticoid receptor, activating it and activating SGK1. Activated SGK1 leads to a deactivation of NEDD4L/NEDD4-2, which inhibits ENaC degradation. An increase in ENaC will increase blood pressure.

TABLE 41.2 Select Genes Associated With Salt Sensitivity (Felder et al., 2013; Sanada et al., 2006, 2011; Carey et al., 2012; Armando et al., 2015).

Gene	Function	Impact in body	Genetic Position(s)
<i>The RAS system</i>			
Angiotensin-converting enzyme (ACE)	Converts angiotensin I (hormone) to active angiotensin II (vasoconstrictor).	Blood pressure increases by constricting blood vessels.	rs1799752 (insertion/deletion in intron 15)
Angiotensinogen (AGT)	Mainly released by the liver into circulation. Renin converts it to angiotensin I (rate-limiting step of RAS).	Thyroid hormone, angiotensin II, estrogen, and corticosteroid increase angiotensinogen levels to eventually increase blood pressure by acting as a vasoconstrictor. Increasing the number of copies of AGT has been associated with an increase in blood pressure.	rs699 (M235T) rs4762 (T174M) rs5051 (G6A)
Angiotensin II receptor type 1 (AGTR1)	Regulates aldosterone secretion. Activated by angiotensin II, a peptide that is a vasoconstrictor.	More than 50% of sodium reabsorption is due to AGTR1 in the kidney. Vasoconstriction, aldosterone synthesis and secretion, decreased renal blood flow, renal tubular sodium reuptake.	rs4524238 (G > A)
<i>Mineralocorticoids and the mineralocorticoid receptor cascade</i>			
Aldosterone synthase (CYP11B2)	Enzyme is regulated by the RAS system and produces aldosterone.	Aldosterone binds to mineralocorticoid receptors to increase sodium and water retention in the kidney nephron. Aldosterone also increases sodium absorption in the collecting duct by upregulating epithelial sodium channels (ENaCs). Sodium-chloride symporter (NCC) expression in the distal convoluted tubule is also upregulated with aldosterone.	rs1799998 (T-344C)
Serum/Glucocorticoid regulated kinase 1	Involved with renal sodium excretion by activating specific potassium, sodium, and chloride channels. Upregulated by aldosterone/mineralocorticoid receptor cascade.	Regulates epithelial Na ⁺ channel (ENaC) to control sodium reabsorption.	rs2758151 (C/T) rs9402571 (T/G)
11-B-Hydroxysteroid dehydrogenase (HSD11B2)	Converts cortisol (active) to cortisone (inert). Cortisol binds to mineralocorticoid receptors with greater affinity than aldosterone.	Downregulates mineralocorticoid receptor cascade.	G-209A (rs45598932)
Cytochrome P450 3A5 (CYP3A5)	Metabolizes endogenous steroids, such as cortisol.	Downregulates mineralocorticoid receptor cascade.	CYP3A5*3 allele
Epithelial Na ⁺ channel (ENaC)	A membrane bound ion channel with selective permeability to sodium ions.	Involved with the absorption of sodium ions in the collecting ducts in the kidney.	rs1799979 ENaC beta subunit ENaC beta subunit rs149868979
Neural precursor cell expressed, (NEDD4L/NEDD4-2)	Downregulates the epithelial sodium channel (ENaC) in the kidney.	Increases salt excretion through downregulating ENaC in the kidney. It has the opposite effect of aldosterone.	rs4149601

Continued

TABLE 41.2 Select Genes Associated With Salt Sensitivity (Felder et al., 2013; Sanada et al., 2006, 2011; Carey et al., 2012; Armando et al., 2015).—cont'd

Gene	Function	Impact in body	Genetic Position(s)
WNK lysine-deficient protein kinase 1 (WNK1)	Expressed in the kidney and activates SGK1, which activates ENaC.	Activating ENaC will increase the absorption of sodium ions in the collecting ducts in the kidney.	rs1159744 rs880054
<i>Solute carrier family members</i>			
Solute Carrier Family 4 Member 5 (SLC4A5)	Encodes a sodium bicarbonate cotransporter. A membrane protein involved with intracellular pH regulation.	Encodes a protein that transports sodium and bicarbonate ions in renal tubule cells back into circulation.	rs7571842 rs10177833
Solute Carrier Family 8 Member A1 (SLC8A1)	A sodium-calcium exchanger that removes calcium from cells and imports sodium into cells.	Regulates intracellular calcium concentrations, affects peripheral vascular resistance, and affects blood pressure regulation in response to salt load.	rs434082
Solute Carrier Family 24, Member 3 (SLC24A3)	A potassium-dependent sodium-calcium exchanger that transports 1 calcium and 1 potassium ion for 4 extracellular sodium ions.	Regulates intracellular calcium concentrations and electrical conduction.	rs3790261
<i>Additional proteins</i>			
Uromodulin (UMOD)	Expressed in the thick ascending limb of Henle and is excreted in urine.	Regulates water/electrolyte balance through sodium excretion and reabsorption.	rs13333226 rs4293393
G protein-coupled receptor kinase 4 (GRK4)	This protein deactivates activated G protein-coupled receptors via phosphorylation.	GRK4 regulates the dopamine D ₁ receptor. The D ₁ receptor plays a role in regulating the renal tubule ion transport.	rs296036 rs1024323 rs1801058
Adducin 1 (ADD1)	A cytoskeleton protein.	Polymorphisms in this gene have been associated with sodium retention via an increase in Na + K + -ATPase activity.	rs4961 rs4963

group may not be the same markers in another group. A specific set of genetic markers tends to be inherited together by ethnicity, so how salt sensitivity is defined genetically for one group may not be held true when looking at another group. This is the reason why replication of these genetic studies in another group is critical in being able to understand the extent to which genetic findings can be generalized to other ethnic groups. One critical point to remember is that diversity among African Americans is greater than the diversity across all other groups. To put this into perspective, this means there is more of a chance of genetic homology between Asians and Caucasians than there is between two African Americans. This is critical to point out because when clinical screening tools are created, it is important to take into consideration which populations were used for the development of these tools, and whether it can be safely extrapolated to the general population.

POTASSIUM (JAIN ET AL., 2013)

Disorders Leading to Hypokalemia

A number of genetic disorders result in a dysregulation of potassium homeostasis. Liddle syndrome and GRA both result in early onset hypertension and hypokalemia. Bartter syndrome and Gitelman syndrome result in normal to low blood pressure and hypokalemia.

Disorders Leading to Hyperkalemia

Genetic variants can cause retention of potassium. Pseudohypoaldosteronism (PHA) Type I is one disorder, although not all the genetic variants have been identified. Variants in the ENaC and regulators of the ENaC have been identified. PHA II is another disorder thought to be caused by variants in serine/threonine protein

kinase 1 (WNK1) and 4 (WNK4). Congenital isolated hypoaldosteronism has been connected to variants in the aldosterone synthase enzyme.

CHLORIDE (PULJAK AND KILIC, 2006)

Genetic variants resulting in a loss-of-chloride-channel function have been linked to myotonia congenita, dystrophia myotonica, cystic fibrosis, osteopetrosis, and epilepsy. On the other hand, genetic variants resulting in chloride channel activation have been associated with brain gliomas and accelerated growth of malaria parasite in red blood cells.

SUMMARY

There are numerous genetic disorders that affect sodium, potassium, and chloride homeostasis. Sodium is the most extensively studied electrolyte because of its association with the salt-sensitive phenotype and the role it has with hypertension. Many genes have been identified and investigated because of the resulting benefits of identifying the salt-sensitive phenotype in clinical practice. There is great potential to tailor dietary and pharmacological interventions based on whether individuals are sensitive or resistant to salt. While further research is needed to clarify this relationship, there are promising genetic variants that have been identified to predict this phenotype. As heart disease is one of the top public health burdens in the United States, identifying more effective treatment strategies in this therapeutic area is greatly needed and utilizing the salt-sensitive classification is one approach.

References

- Armando, I., Villar, V.A.M., Jose, P.A., 2015. Genomics and pharmacogenomics of salt-sensitive hypertension minireview. *Curr Hypertens Rev* 11 (1), 49–56.
- Carey, R.M., Schoeffel, C.D., Gildea, J.J., et al., 2012. Salt-sensitivity of blood pressure is associated with polymorphisms in the sodium-bicarbonate co-transporter. *Hypertension* 60 (5), 1359–1366. <https://doi.org/10.1161/HYPERTENSIONAHA.112.196071>.
- Centers for Disease Control and Prevention (CDC), 2012. Vital signs: awareness and treatment of uncontrolled hypertension among adults—United States, 2003–2010. *MMWR Morb Mortal Wkly Rep* 61, 703–709. PubMed PMID: 22951452.
- Felder, R.A., White, M.J., Williams, S.M., Jose, P.A., 2013. *Curr Opin Nephrol Hypertens* 22 (1), 65–76. <https://doi.org/10.1097/MNH.0b013e32835b3693> PMID: PMC3724405. Author manuscript; available in PMC 2013 September 1. Published in final edited form as: *Curr Opin Nephrol Hypertens*.
- Jain, G., Ong, S., Warnock, D.G., 2013. Genetic disorders of potassium homeostasis. *Semin Nephrol* 33 (3), 300–309. <https://doi.org/10.1016/j.semnephrol.2013.04.010>. Review. PubMed PMID: 23953807.
- Puljak, L., Kilic, G., 2006. Emerging roles of chloride channels in human diseases. *Biochim Biophys Acta* 1762 (4), 404–413. Review. PubMed PMID: 16457993.
- Sanada, H., Yatabe, J., Midorikawa, S., Hashimoto, S., Watanabe, T., Moore, J.H., Ritchie, M.D., Williams, S.M., Pezzullo, J.C., Sasaki, M., Eisner, G.M., Jose, P.A., Felder, R.A., 2006. Single-nucleotide polymorphisms for diagnosis of salt-sensitive hypertension. *Clin Chem* 52 (3), 352–360. PubMed PMID: 16439609.
- Sanada, H., Jones, J.E., Jose, P.A., 2011. Genetics of salt-sensitive hypertension. *Curr Hypertens Rep* 13 (1), 55–66. <https://doi.org/10.1007/s11906-010-0167-6>. Review. PubMed PMID: 21058046; PubMed Central PMCID: PMC4019234.
- Weinberger, M.H., Fineberg, N.S., Fineberg, S.E., Weinberger, M., 2001. Salt sensitivity, pulse pressure, and death in normal and hypertensive humans. *Hypertension* 37 (2 Pt 2), 429–432. PubMed PMID: 11230313.

This page intentionally left blank

42

Iron

Martin Kohlmeier

Human Research Core and Nutrigenetics Laboratory, UNC Nutrition Research Institute; and Schools of Medicine and Public Health, Department of Nutrition, Kannapolis, NC, United States

Glossary

Ferroportin This iron transporter (encoded by SLC40A1) is located at the basolateral side of small intestinal enterocytes and on macrophages.

Bronze diabetes Deposits of excess iron in patients with severe hemochromatosis damage pancreas and decrease insulin production.

Nutritope Specific nutritional environment that historically contributed to the shaping of the genome of modern human populations.

Selective advantage Variants may give carriers a slight edge over non-carriers, the variant may spread by raising more offspring over many generations, but not so much that everybody becomes a carrier.

INTRODUCTION

Absorption of dietary iron and iron-containing compounds (mainly myoglobin, heme, and cytochromes), metabolism, and delivery to tissues has been described elsewhere in detail (Kohlmeier, 2015). Key steps of absorption of elemental iron from foods include uptake by the luminal transporter DMT1, export by ferroportin, and oxidation by hephaestin. Heme-containing food compounds are mostly taken up via the HCP1 transporter (which also takes up foliate), and their ferrous (Fe^{3+}) iron released and converted to ferric iron (Fe^{++}) by heme oxidase (Fig. 42.1).

The key regulator of iron homeostasis is the hormone hepcidin, encoded by the HAMP gene and expressed mainly in the liver. Expression of hepcidin in the liver is stimulated when transferrin receptor 1 binds diferric transferrin and thereby alters the steric configuration of transferrin receptor 2 and hemochromatosis protein (encoded by the HFE gene). The overall regulator of hepcidin expression is circulating bone morphogenic protein 6 (BMP6), which acts on a receptor complex at the hepatocyte membrane and thereby communicates iron status to the nucleus through a signaling cascade involving hemojuvelin (encoded by HJV) and several

SMAD proteins. The membrane complex is inactivated by matriptase-2, a membrane-associated protease. Hepcidin itself inhibits expression of the iron exporter (ferroportin) in the basolateral membrane of enterocytes in the small intestine and in macrophages.

RISK OF DEFICIENCY

Iron is the micronutrient that is most commonly deficient worldwide, affecting particularly children and young women in the less affluent regions of the world. Periodic bleeding in young women and intestinal blood loss due to parasitic infections in children often cannot be compensated by new red blood cell production due to low iron supplies. The main dietary factors contributing to iron deficiency are low intake of highly available heme iron, the lack of absorption-promoting organic acids, and concurrent consumption of absorption inhibitors like black tea with nonheme iron-containing foods.

Relatively common variants in several genes, including HFE, TF, TNF- α , and TMPRSS6, have been found to slightly decrease the risk of iron deficiency in vulnerable populations. This selective advantage of being able to survive with somewhat lower iron supply is probably the reason why their frequency has historically risen in some groups to what is found in contemporary populations.

Prompted by a previous observation in a genome-wide association study, investigators examined the role of the common nonsynonymous variant TMPRSS6 (transmembrane protease serine 6, encoding matriptase 2) rs855791 C in a group of women. The matriptase 2 enzyme inhibits hepcidin production in the liver. Women with iron-deficiency anemia were significantly less likely (OR = 0.4) to carry the previously identified risk genotype rs855791 CC than women with

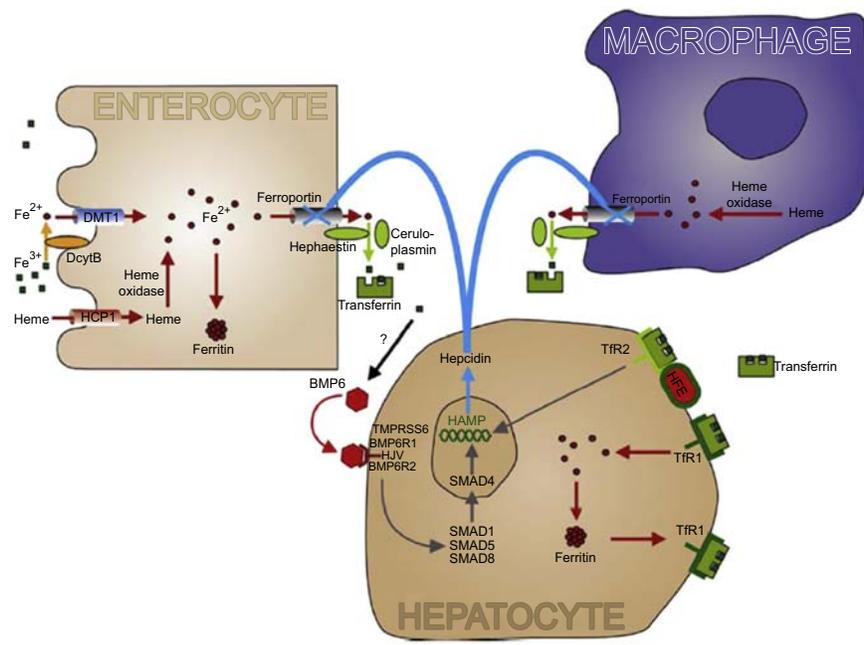


FIGURE 42.1 Overview of the steps in the absorption, transport, storage, and excretion of iron. From Kohlmeier, M. 2013. *Nutrigenetics. Applying the Science of Personal Nutrition*. Academic Press, London.

hemoglobin concentration in the normal range (Pei et al., 2014). They further found that the CC genotype abolished the usually observed inverse relationship of hemoglobin concentration and estimated menstrual blood loss. This appears to indicate that this allele abolishes much of the matriptase 2 action and allows hepcidin production to proceed unhindered. The investigators assumed, therefore, that this variant provided some protection against iron deficiency in menstruating women. Two different variants in *TMPRSS6*, rs228918 and rs228921, apparently afforded similar protective effects in South African women (Gichohi-Wainaina et al., 2015), indicating that independently arisen polymorphisms have provided the same type of selective advantage to geographically separated populations.

Variants in additional genes, including *TF* (rs1799852) and *TNF- α* (rs1799964 and rs1800629), have been found to have similar effects (Gichohi-Wainaina et al., 2015).

Carriers of the *HFE* rs1800562A and rs1799945G alleles appear to be less likely to have iron deficiency and anemia (Sørensen et al., 2019). These variants are known to promote iron retention. It is likely that they gave people in Ireland and other historically iron-poor environments in Western Europe a selective edge, which would explain their relatively high prevalence of several percent or more in some populations there. As will be explored in more detail, some carriers of these variants have an increased risk of excessively accumulated iron, turning the historic advantage into a potential health risk with the currently prevalent nutritional environment (nutritope) with plentiful availability of meats and other iron-rich foods.

RISK OF IRON ACCUMULATION AND HEMOCHROMATOSIS

Pathophysiology and Health Consequences

Clinically manifest hemochromatosis with significant iron accumulation in tissues is most common in people with Northwest European ancestry, but also occurs in other populations. Premenopausal women are less affected than men as long as they keep losing blood during menstruation. Contraceptive practices and optimal health measures tend to limit blood losses, which means that the women lose some of their gender-specific protection. Hemochromatosis affects particularly tissues of the liver, pancreas, heart, brain, and other tissues due to iron accumulation, much of which is in the form of irreversible hemosiderin deposits. This is one of the reasons why detection and treatment need to be as timely as possible. Early nonspecific symptoms include lethargy, fatigue, abdominal pain, weight loss, and joint stiffness or arthritis. A slate-blue tinge of the skin in severely affected patients is typical. Over the long term, excess iron accumulation will cause significantly adverse outcomes. Liver cirrhosis is very common and may lead to liver failure or liver cancer. Deposits in the pancreas can cause failure of this organ's endocrine function and is sometimes called bronze diabetes due to the associated skin color. Heart failure and cognitive decline are also common consequences.

Another harmful feature in many affected individuals is the inordinately high concentration of unbound iron in blood. The free iron increases susceptibility to septicemia

from consumption or handling of raw oysters, which commonly contain the gram-negative marine bacterium *Vibrio vulnificus*. Vulnerability to infection also extends to various other pathogens, particularly involving gram-negative bacteria, more commonly than in patients without hemochromatosis. Affected patients have a distinctly increased risk of ending up with septicemia. A recent widely reported case concerned a microbiologist with hemochromatosis who died after a lab-acquired infection with an attenuated and particularly iron-dependent strain of *Pasteurella pestis* for vaccination (Queen et al., 2012).

Iron accumulation in patients with hemochromatosis results from a combination of genetic disposition and iron intake. This makes it the classic prototype of a nutrigenetic disease that depends on the interaction of inherited vulnerability and precipitating nutritional exposure. Many carry the risk alleles, but long-term iron intake is too low in many allele carriers to cause significant harm. Some may even benefit from their condition by avoiding iron deficiency as mentioned above.

Genetic Risk Factors

The classical HFE gene variants associated with manifest hemochromatosis are rs1800562A (845G > A, Cys282Tyr, about 2% in West-European populations), less closely rs1799945G (187C > G, His63Asp, about 7% in West-European populations) or rs1800730T (193A > T, Ser65Cys). Clinical manifestation typically requires either two copies of rs1800562 (Cys282Tyr) or three of any of the risk variants. The risk of developing hemochromatosis is several thousand times higher in carriers with two Cys282Tyr copies than in those without a variant allele. A lesser dose of risk variants in conjunction with high iron and alcohol exposure can also cause significant iron accumulation with harmful health outcomes. Increased transferrin saturation and massively increased concentration of unbound iron is typical.

Pathogenic variants in the SLC40A1 gene (encoding ferroportin) cause a slightly different form of iron accumulation, called ferroportin disease or hemochromatosis type 4. A large number of SLC40A1 variants reduce to different degree the ability of hepcidin to interact and thereby inactivate ferroportin (Fig. 42.2). The resulting lack of ferroportin capacity causes iron accumulation in the liver and in macrophages as indicated by greatly elevated ferritin concentration in circulating blood. Unlike in HFE-related hemochromatosis, transferrin saturation is not always increased. The key problem is that the feedback regulation of iron balance is disrupted.

Variants of several other genes including HAMP (encoding hepcidin), HJV (encoding hemojuvelin), and TFR2 (encoding transferrin receptor 2) also cause clinically relevant iron accumulation in rare cases.

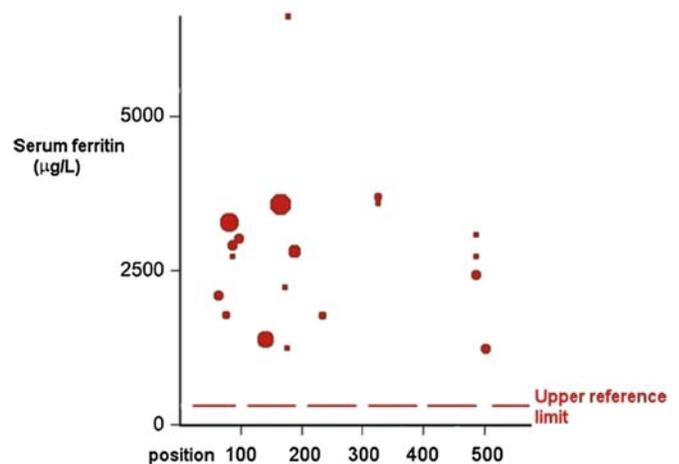


FIGURE 42.2 Several specific SLC40A1 variants cause ferroportin disease with the ferritin concentration in blood reflecting the severity of the disturbance of iron balance (Mayr et al., 2010). The size of the dots indicates the frequency of a variant at the indicated gene position. From Kohlmeier, M. 2013. *Nutrigenetics. Applying the Science of Personal Nutrition*. Academic Press, London.

Interventions

The most common disease-promoting exposures are high iron intake (particularly red meat, foods made with iron-fortified flour, and highly fortified breakfast cereals), alcohol abuse, smoking, and obesity. Their avoidance is often sufficient to halt progression of the disease progress, and improvement of overall outcomes has been observed since diagnostic and therapeutic capabilities have improved (Deugnier et al., 2019). When allele carriers become aware of their vulnerability to excessive iron intake, they can stop taking iron-containing supplements and reduce their exposure to iron-supplemented foods. The iron content of many boxed breakfast cereals is an easy target because so many alternatives with lower iron content are available.

Carefully controlled blood draws are also helpful for the treatment of hemochromatosis, regardless of the underlying genetic cause. It is now legally possible in many jurisdictions to do this as regular blood donations several times a year, an economical and effective solution for the millions of risk-allele carriers.

Liver transplantation is an option for severely affected patients with the HFE rs1800562AA genotype that corrects the genetic defect in liver (Pericleous and Kelly, 2018).

References

- Deugnier, Y., Morcet, J., Lainé, F., Hamdi-Roze, H., Bollard, A.S., Guyader, D., Moirand, R., Bardou-Jacquet, E., 2019. Reduced phenotypic expression in genetic hemochromatosis with time: role of exposure to non-genetic modifiers. *J Hepatol* 70 (1), 118–125.
- Gichohi-Wainaina, W.N., Melse-Boonstra, A., Swinkels, D.W., Zimmermann, M.B., Feskens, E.J., Towers, G.W., 2015. Common

- variants and Haplotypes in the TF, TNF- α , and Tmprss6 genes are associated with iron status in a female black South African population. *J Nutr* 145 (5), 945–953.
- Kohlmeier, M., 2013. *Nutrigenetics. Applying the Science of Personal Nutrition*. Academic Press, London.
- Kohlmeier, M., 2015. *Nutrient Metabolism. Structures, Functions and Genes*, second ed. Academic Press, London.
- Mayr, R., Janecke, A.R., Schranz, M., Griffiths, W.J., Vogel, W., Pietrangelo, A., Zoller, H., 2010. Ferroportin disease: a systematic metaanalysis of clinical and molecular findings. *J Hepatol* 53 (5), 941–949.
- Pei, S.N., Ma, M.C., You, H.L., Fu, H.C., Kuo, C.Y., Rau, K.M., Wang, M.C., Lee, C.T., 2014. Tmprss6 rs855791 polymorphism influences the susceptibility to iron deficiency anemia in women at reproductive age. *Int J Med Sci* 11 (6), 614–619.
- Pericleous, M., Kelly, C., 2018. The clinical management of hereditary haemochromatosis. *Frontline Gastroenterol* 9 (2), 110–114.
- Quenee, L.E., Hermanas, T.M., Ciletti, N., Louvel, H., Miller, N.C., Elli, D., Blaylock, B., Mitchell, A., Schroeder, J., Krausz, T., Kanabrocki, J., Schneewind, O., 2012. Hereditary hemochromatosis restores the virulence of plague vaccine strains. *J Infect Dis* 206 (7), 1050–1058.
- Sørensen, E., Rigas, A.S., Didriksen, M., Burgdorf, K.S., Thørner, L.W., Pedersen, O.B., Hjalgrim, H., Petersen, M.S., Erikstrup, C., Ullum, H., 2019. Genetic factors influencing hemoglobin levels in 15,567 blood donors: results from the Danish Blood Donor Study. *Transfusion* 59 (1), 226–231.

Trace Elements

Martin Kohlmeier

Human Research Core and Nutrigenetics Laboratory, UNC Nutrition Research Institute; and Schools of Medicine and Public Health, Department of Nutrition, Kannapolis, NC, United States

Glossary

Chaperone Intracellular proteins that guide compounds to targets for synthesis and other needed interactions.

Founder effect High frequency of a genetic variant from a shared ancestor in a population that has expanded from a few individuals.

Goiter Abnormal enlargement of the thyroid gland, often in people with iodine deficiency.

Metallothioneins Small cysteine-rich proteins that bind heavy metals, including zinc and copper, facilitating their reverse transport and excretion.

Nutritope Nutrition environment to which local groups are exposed to and which over time can shape their genome.

INTRODUCTION

Proper functioning of human structures and metabolism depends on at least 26 elements. We call about half of them trace elements because daily intakes of a few milligrams or even micrograms are sufficient for good health. Many of them, including iron, copper, zinc, manganese, selenium, and molybdenum, function as part of catalytic complexes. A few more have narrowly specialized functions, such as iodine as a constituent of thyroid hormones and bromine as the halogenating oxidant hypobromous acid deployed by eosinophils against parasites and microbes (Kohlmeier, 2015).

Although the needed amounts are very minimal, it is often difficult for humans to obtain the needed quantities. Deficiencies of some trace elements are exceedingly common in many developing countries and also continue to be a problem for many individuals in affluent countries. In other instances, harm is caused by excess intake. While everybody will become deficient with very low intakes or become ill from getting too much, individuals with particular genetic vulnerabilities are the ones most commonly affected (Kohlmeier, 2013).

What is fairly unique about humans is their prehistoric presence on all continents except Antarctica. Our

invasive species has survived and even thrived in all but the most barren environments. Humans and their near-human hominid ancestors and cousins have lived in rain forests, hot and cold deserts, at the sea shores and on high mountains, on frozen tundras, and even arctic ice long before the development of modern survival technologies. Many of these diverse nutritional environments (nutritopes) confronted inhabitants with critical shortfalls or excesses of trace elements. A well-understood example is the life in mountain regions with low iodine availability in the air, water, or foods. Iodine deficiency causes severe developmental and other disorders (cretinism, goiter) and limits metabolic, physiological, and mental function. As a consequence, individuals with genetic variants most suitable to tolerate the local trace element exposures were the ones most likely to have more offspring and perpetuate their nutritional idiosyncrasies. Modern humans carry a jumble of these ancestral adaptations whether or not they fit their current nutritional exposures.

The following sections will give a brief overview and introduction into the genetic diversity that governs human requirements and tolerances to excess of key trace elements. In each case the focus will be on common variants (polymorphisms) and not on rare disease-cause gene defects.

Zinc

Zinc is an essential cofactor for many hundreds of numerous enzymes and regulatory proteins, accounting for nearly 10% of all expressed proteins (Kimura and Kambe, 2016). The zinc in enzymes may have a structural role, such as binding the neighboring amino acids Cys97, Cys100, Cys103, and Cys111 in alcohol dehydrogenases and thereby maintaining the necessary three-dimensional structure of the active enzyme. In the alkaline phosphatase dimer, on the other hand, four

zinc atoms are part of the catalytic center and directly involved in the reaction. Zinc in the numerous zinc finger proteins ensures the tight binding of these transcription factors to specific DNA or RNA sequences and thereby plays a critical role in the regulation of gene expression, often with great tissue specificity.

Intestinal absorption and transfer to specific organs and tissues depends on a series of Zrt- and Irt-like proteins (ZIP), which move zinc into cells but out of cell storage organelles, and zinc transporters (ZnT), which move zinc out of cells but into cell storage organelles. Each of these transporters is encoded by a different gene and follows a cell-specific expression profile in response to regulatory signals. Zinc enters small intestinal enterocytes from the luminal side via ZIP4 (SLC39A4) and moves across the basolateral membrane into the blood stream by the zinc transporters ZnT1 (SLC30A1) and ZnT2 (SLC30A2). Transfer into target tissues uses an array of specialized transporters, close to 20 by current count. The pancreas is particularly dependent on a plentiful zinc supply because insulin can only be secreted into blood as a zinc complex. The metallothioneins are small cysteine-rich peptides that bind zinc, copper, and other heavy metals in circulation. A high load of zinc, for instance from a dietary supplement, induces the expression of metallothioneins and promotes reverse transport from cells toward excretion with bile. Since zinc and copper share the metallothioneins, their increased expression will increase losses for both of them and may lead to copper deficiency over time. Metallothioneins also contribute greatly to zinc storage in cells.

It is obvious that insufficient zinc status has numerous detrimental health consequences, including weakened immune function, decreased production of blood cells, poor regeneration of skin, and many other signs of failing biological functions. A very large number of people, particularly in less affluent environments, do not get as much zinc as would be needed for optimal health. Common diarrheal infections and parasite infestations increase zinc losses and thereby weaken resistance to another cycle of illness. Typical symptoms of zinc deficiency include dry and scaling skin, seborrhea (acne), poor wound healing, diarrhea, anemia, and infections. Particularly young children and older people are very vulnerable to potentially life-threatening infections. High-dose zinc supplementation is known to greatly decrease mortality in older people, for instance, by 17% after 10 years in a randomized placebo-controlled trial with more than 2000 participants (Chew et al., 2013).

What may be less known is that zinc deficiency is often associated with a declining sense of taste and smell.

Meats and organ meats, legumes, and nuts are zinc-rich foods (Kohlmeier, 2015). Oysters are often mentioned as a particularly rich source, but it is much

more likely that on a daily basis people get a good dose from beans.

When using zinc supplements, it is critical to understand that high doses promote copper excretion, which can lead to severe copper deficiency. It is now considered good practice to combine zinc with a suitable amount of copper (commonly 50 mg zinc plus 2 mg copper).

Interaction of Genetic Variation With Zinc Nutrition

Not everybody is equally vulnerable to low zinc intake or increased losses. Rare variants occur in most known genes contributing to normal function, and often the severe pathologies caused by the variants have led to the discovery of the gene. The very rare condition acrodermatitis enteropathica (OMIM 201100) is due to loss-of-function variants in the ZIP4 (SLC39A4) gene coding for the zinc transporter that moves zinc from the intestinal lumen into the enterocyte. This type of unfortunate natural experiment demonstrates the critical importance of this gene. Patients typically have diarrhea, hair loss, scaling, and oozing dermatitis at perioral, perianal, and other sites, starting from earliest infancy. High vulnerability to infections is another common consequence with severe impact. Sustained high doses of zinc can largely compensate for the defective zinc transport (Hammersen et al., 2018).

Several polymorphic variants, particularly of zinc transporters, impact the risk of common disease (Kohlmeier, 2013). The genotype-specific differences in the prevalence of chronic diseases are remarkable.

An important example is the common variant rs13266634 in ZIP8 (SLC30A8), which encodes a zinc transporter in pancreas beta-cells. Ablation of ZIP8 in an animal knockout model demonstrated that the encoded transporter pumps zinc into the storage granules, where two zinc atoms become coordinated to six proinsulin molecules. This form of storage ensures availability of insulin for secretion when needed. The zinc transporter 8 protein in people with the T allele contains in position 325 a tryptophan instead of the R in people with the C allele. Carriers of the rs13266634 T allele tend to have an increased risk of type 2 diabetes, particularly when they are obese, and need higher insulin reserves to overcome their insulin resistance. Use of high-dosed zinc supplements and high dietary zinc-to-iron ratio decreases diabetes risk in people with T allele, but less so in those without a T allele (Drake et al., 2017). The authors suggest to take both genotype and weight status into account when considering the use of high-dosed zinc supplements.

The same variant is also involved in a faster rate of progression in children who eventually develop type 1 diabetes. A majority of children with type 1 diabetes

have autoantibodies against the zinc transport protein 8. Remarkably, about 80% of the ZnT8 autoantibody-positive children with either two copies of the rs13266634 T allele or two copies of the rs13266634 C allele developed type 1 diabetes within 10 years, but less than 30% of those with both alleles (Achenbach et al., 2009). This difference in progression rate was consistent throughout the 10 years of observation.

Preliminary study findings in Korean adults suggest a combination of variants in ZIP11 (SLC39A11) coding for yet another zinc transporter (Ha and Bae, 2018). Individuals with the rs17183225T/rs17780814C haplotype combination were more likely to suffer from chronic gastritis than those with the rs17183225C/rs17780814A/rs17780820A haplotype (odds ratio, 2.620; 95% confidence interval, 1.207–5.689). ZIP11 has previously been shown to be located to the nucleus. A number of reports have linked variants in this gene with cancer of the stomach and other organs.

The common variant allele G in the promoter region of the MT2A gene decreases transcription efficiency (Hattori et al., 2016). An association with increased risk of diabetes and kidney disease has been observed. Whether this genotype-associated risk is related to reduced zinc status or can be ameliorated by higher zinc intake remains to be demonstrated.

The impact of zinc supplementation on the progression of age-related macular degeneration (AMD) differs by a common variant in the CFH gene, coding for the complement factor H. Complement factor H is a soluble glycoprotein in circulation that together with the other components of the complement complex protect against infectious agents and aberrant cells. The variant CFH rs1061170 T > C (Phe402His) appears to predict AMD progression with zinc plus antioxidant supplementation (Seddon et al., 2016). The likelihood for progressing to a more advanced AMD stage in the patients with the TT genotype was only half as much with supplementation than with placebo (HR: 0.55, 95% CI 0.32 to 0.95, $P = .033$), while little difference between supplementation and placebo was found for the others.

Copper

Copper plays critical roles in numerous biology functions, including mitochondrial energy generation, antioxidant defense, collagen maturation, blood coagulation, nutrient metabolism, and hormone synthesis (Kohlmeier, 2015).

Copper distribution to and from tissues is closely regulated, and disruption of this balance will cause specific pathologies in the liver, central nervous system, and other organs. Several of these disorders will be explored in more detail below. Excessive accumulation of copper in liver, pancreas, and brain is most often a matter of

excessive intake in combination with specific vulnerabilities due to genetic variants.

Major sources of copper are organ meats, oysters, some mushrooms, and dark chocolate. Drinking water can be high in copper, usually from exposure to copper-containing water pipes. Water from private water wells can also occasionally contain large amounts of copper.

Genetic Variants Affecting Copper Transport

Inherited variants in two ATP-dependent copper transporters cause severe disease. What is not well understood due to a lack of representative data is to what extent single copies of risk variants in combination with environmental and nutritional exposure may impact health. In some endemic areas, the frequency of the risk alleles is significant.

The Wilson protein (ATP7B) is important for reverse transport out of cells and moves excess copper into bile for excretion. Wilson's disease (OMIM 606882) is a rare condition (about one in 30,000) due to one or two of hundreds of known defect variants in ATP7B (Gomes and Dedoussis, 2016). The condition is more common in East Asia than elsewhere, often due to the rs28942074 (2333G > T, Arg778Leu) variant, present in more than 10% of the healthy population. Patients with European ancestry often carry the rs76151636 (3207C > A, H1069Q) variant. Another variant, a 15 bp deletion from position -441 to position -427 upstream of the translation start site, is relatively common in people with Sardinian ancestry, apparently due to a founder effect (Loudianos et al., 1999).

The inability to properly remove excess copper from the body leads to accumulation in liver, pancreas, brain, cornea, and other tissues. Copper deposits in brain cause ataxia and tremor. Hepatitis and cirrhosis are typical outcomes in the liver. Excess copper in the eyes causes the pathognomonic Kayser-Fleischer rings of the cornea, and also a typical sunflower cataract. Cardiomyopathy and arrhythmia are typical findings in the heart. Additional presentations may include hypoparathyroidism, acromegaly, unusual by height, osteoarthritis, increased bone fracture risk, pancreatitis, kidney stones, and nephrocalcinosis. Despite all this severe morbidity, the affected patients still tend to have a nearly average life expectancy. Patients are best treated with an individually tailored therapy with a combination of penicillamine, trientine, tetrathiomolybdate, or other chelators (Horn et al., 2019), which needs to be continued lifelong.

Another important active transporter is the Menkes disease protein (ATP7A), which pumps copper in conjunction with specific chaperones toward the intracellular sites of copper-protein synthesis. Very rare defect variants cause Menkes disease (previously named

Menkes kinky hair syndrome, OMIM 309400), which results in severe functional copper deficiency due to inability to absorb copper from the intestinal lumen, but also the inability to transport copper across the blood-brain barrier and to the many different intracellular locations for functional use (Horn et al., 2019). Since the ATP7A gene is located on the X chromosome, the disease affects mainly men. More than 500 variants causing some type of Menkes disease are known. Because the loss of function caused by individual variants ranges from complete to significant residual activity, the clinical presentations are quite diverse. Carriers of two nonfunctional variants typically have severe dysfunction across most organ systems. Impaired brain function causes seizures and mental retardation. Aortic aneurysm and fragile blood vessels, endocrine disorders, emphysema, urogenital abnormalities, and bone and joint deformities pose severe health challenges that usually cause early death. Typical is the kinky, coarse, and easily breakable hair. Treatment options are very limited. Some types of chelating agents can bypass the defective transporter and act as artificial ionophoric channels.

Very rare mutations of the ceruloplasmin gene (CP) disrupt normal copper transport in blood (OMIM 117700). Progressive degeneration of the basal ganglia in brain and of the retina are typical. Iron status tends to be low due to diminished action of the copper protein ferroxidase (EC 1.16.3.1), which is needed for oxidizing the intraepithelial diferric iron to the ferrous form transported in blood with transferrin. The condition was found to be associated in some cases with high ferritin concentration in blood and liver siderosis due to impaired iron transport. Iron depositions in the cerebellum and other organs have been found in other patients.

Selenium

Selenium contributes to defense against free radicals, is essential for the synthesis of thyroid hormones and insulin, and supports fertility and cell growth. All these functions rely on the site-specific incorporation by tRNA^{Ser,Sec} of selenocysteine.

The selenium content of foods differs greatly between regions; for example, there are areas in China and elsewhere with very low-selenium content of foods due to selenium-poor soil. In those regions, symptomatic selenium deficiency is common, causing a specific form of osteoarthritis (Kashin–Beck disease) due to increased loss of chondrocytes. Many of the consequences of selenium deficiency are subtle and not easily measured. Poor selenium status, for instance, has the potential to make influenza and other viruses more virulent.

A few foods with unusually large amounts of selenium include Brazil nut (contains several times the required amount), liver, clams, and some dietary supplements. Long-term consumption of excessive amounts (more than 400 µg per day) can cause diarrhea, dermatitis, and peripheral neuropathy (Kohlmeier, 2015).

Interaction of Genetic Variation With Selenium Nutrition

A common genetic variant in the GPX1 gene, encoding glutathione peroxidase 1 (EC 1.11.1.9), modulates the effect of dietary intake. Individuals with two copies of the rs1050450 T allele (679C > T, Pro198Leu) have about 7% lower selenium concentration in blood comparable intakes. The glutathione peroxidase activity is affected to even greater extent, with a difference of about 15%. While such differences may not appear to be much, they shift the entire population distribution toward the lower end of the curve and increase the frequency of both modest inadequacy and severe deficiency. Since the exposure difference is lifelong, it can translate into greater rates of serious chronic disease such as cancer. Thus, the risk of osteoarthritis in a region of China with low-selenium soil was found to be twice as high in carriers of the TT genotype than in people without the T allele (Xiong et al., 2010).

Another nearby GPX1 variant, rs1050450 (Pro200-Leu), was also associated with lower blood selenium concentration and enzyme activity (Miller et al., 2012). A meta-analysis with a fully adjusted model found a statistically significant higher cancer risk associated with the low-activity genotype (Hong et al., 2013). While recognizing the complexity of these relationships, the findings appear to converge on the notion that these loss-of-function GPX1 variants are associated with slightly higher selenium requirement. This higher need can easily be met with various strategies, from growing enriched foods on selenium-fortified soil, to counseling for consumption of naturally selenium-rich foods, to use of supplements with suitable selenium compounds.

Iodine

The volatile halogen iodine is only used for the synthesis of thyroid hormones. To get iodine to the site of hormone synthesis, it has to be absorbed as iodide (ionized iodine in oxidation state –1) from the intestines (active uptake through the sodium-iodide symporter, NIS), travel with blood to the thyroid gland for uptake via the sodium iodide symporter, and finally flow into the colloid via pendrin. SLC26A4 produces the iodine and chloride transporter pendrin. The selenoenzyme thyroid peroxidase converts the ionized iodine into the neutral atomic form that reacts with specific tyrosine residues in thyroglobulin (Fig. 43.1).

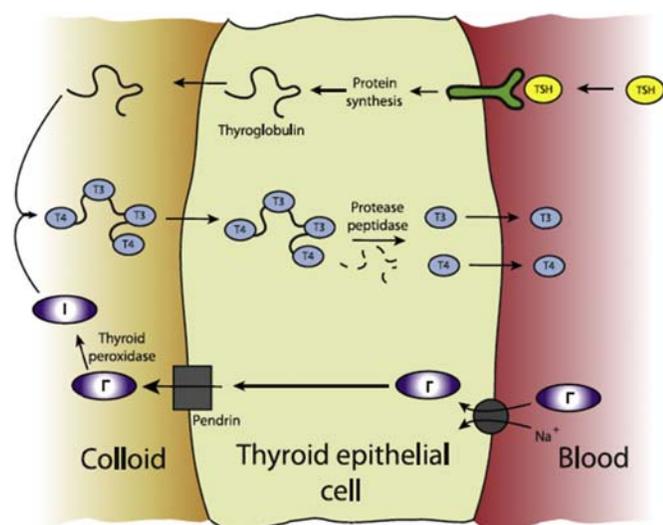


FIGURE 43.1 Pendrin mediates transfer of iodide into the colloid where thyroid peroxidase converts it into the neutral atomic form, which in turn reacts with specific tyrosines in thyroglobulin.

Interaction of Genetic Variation with Iodine Nutrition

Congenital hypothyroidism, which affects about one in 3000 newborn children, is characterized by an insufficient level of thyroid hormone at birth. Of the many possible defects, variants in the thyroid peroxidase gene (TPO) are the most common cause. Loss-of-function variants of the maturation factors DUOX1 and DUOX2, the dehalogenase DEHAL1, the gene TG encoding thyroglobulin, and the gene SLC26A4 encoding pendrin also are found to be responsible.

Endemic goiter with a significant enlargement of the thyroid gland to insufficient availability of dietary and inhaled iodine is common in regions with low-iodine content in soil and water. Since only some individuals are affected, a genetic connection has long been understood to be important. Pendred syndrome (affecting about one in 13,000 Britons) with goiter among its typical presentations is due to defect variants of SLC26A4. This syndrome mirrors endemic goiter to some extent because adequate iodine intake will prevent or mitigate the thyroid gland enlargement. Defect variants in a closely related transporter, SLC26A7, which is also mainly expressed in the thyroid gland, also cause hypothyroid goiter that responds well to iodine supplementation (Cangul et al., 2018). How SLC26A7 participates in the hormone synthesis sequence is as yet unknown, but it is obviously a previously unrecognized necessary step.

Since thyroid peroxidase is a selenoprotein, as pointed out earlier, variants impairing selenium availability will also affect thyroid function. This interaction remains largely unexplored in sufficient detail and needs some attention.

References

- Achenbach, P., Lampasona, V., Landherr, U., Koczwara, K., Krause, S., Grallert, H., Winkler, C., Pflüger, M., Illig, T., Bonifacio, E., Ziegler, A.G., 2009. Autoantibodies to zinc transporter 8 and SLC30A8 genotype stratify type 1 diabetes risk. *Diabetologia* 52 (9), 1881–1888. PMID:19590848.
- Cangul, H., Liao, X.H., Schoenmakers, E., Kero, J., Barone, S., Srichomkwun, P., Iwayama, H., Serra, E.G., Saglam, H., Eren, E., Tarim, O., Nicholas, A.K., Zvetkova, I., Anderson, C.A., Frankl, F.E.K., Boelaert, K., Ojaniemi, M., Jääskeläinen, J., Patyra, K., Löf, C., Williams, E.D., UK10K Consortium, Soleimani, M., Barrett, T., Maher, E.R., Chatterjee, V.K., Refetoff, S., Schoenmakers, N., 2018. Homozygous loss-of-function mutations in SLC26A7 cause goitrous congenital hypothyroidism. *JCI Insight* 3 (20) pii: 99631.
- Chew, E.Y., Clemons, T.E., Agrón, E., Sperduto, R.D., Sangiovanni, J.P., Kurinij, N., Davis, M.D., Age-Related Eye Disease Study Research Group, 2013. Long-term effects of vitamins C and E, β -carotene, and zinc on age-related macular degeneration: AREDS report no. 35. *Ophthalmology* 120 (8), 1604–1611.e4.
- Drake, I., Hindy, G., Ericson, U., Orho-Melander, M., 2017. A prospective study of dietary and supplemental zinc intake and risk of type 2 diabetes depending on genetic variation in SLC30A8. *Genes Nutr* 12, 30. PMID: 29093761.
- Gomes, A., Dedoussis, G.V., 2016. Geographic distribution of ATP7B mutations in Wilson disease. *Ann Hum Biol* 43 (1), 1–8.
- Ha, E., Bae, J.H., 2018. Zinc transporter SLC39A11 polymorphisms are associated with chronic gastritis in the Korean population: the possible effect on spicy food intake. *Nutr Res* 57, 78–85.
- Hammersen, J., Has, C., Galiano, M., Lindner, M., Rossi, R., Kohlhase, J., Schneider, H., 2018. Sustained need for high-dose Zinc supplementation in children with Acrodermatitis Enteropathica. *Clin Pediatr* 57 (1), 99–102.
- Hattori, Y., Naito, M., Satoh, M., Nakatochi, M., Naito, H., Kato, M., Takagi, S., Matsunaga, T., Seiki, T., Sasakabe, T., Suma, S., Kawai, S., Okada, R., Hishida, A., Hamajima, N., Wakai, K., 2016. Metallothionein MT2A A-5G polymorphism as a risk factor for chronic kidney disease and diabetes: Cross-Sectional and Cohort Studies. *Toxicol Sci* 152 (1), 181–193.
- Hong, Z., Tian, C., Zhang, X., 2013. GPX1 gene Pro200Leu polymorphism, erythrocyte GPX activity, and cancer risk. *Mol Biol Rep* 40 (2), 1801–1812.
- Horn, N., Møller, L.B., Nurchi, V.M., Aaseth, J., 2019. Chelating principles in Menkes and Wilson diseases: choosing the right compounds in the right combinations at the right time. *J Inorg Biochem* 190, 98–112.
- Kimura, T., Kambe, T., 2016. The functions of metallothionein and ZIP and ZnT transporters: an overview and perspective. *Int J Mol Sci* 17 (3), 336.
- Kohlmeier, M., 2013. *Nutrigenetics. Applying the Science of Personal Nutrition*. Academic Press, London.
- Kohlmeier, M., 2015. *Nutrient Metabolism. Structures, Functions and Genes*, second ed. Academic Press, London.
- Loudianos, G., Dessi, V., Lovicu, M., Angius, A., Figus, A., Lilliu, F., De Virgiliis, S., Nurchi, A.M., Deplano, A., Moi, P., Pirastu, M., Cao, A., 1999. Molecular characterization of Wilson disease in the Sardinian population – evidence of a founder effect. *Hum Mutat* 14 (4), 294–303.
- Miller, J.C., Thomson, C.D., Williams, S.M., vanHavre, N., Wilkins, G.T., Morison, I.M., Ludgate, J.L., Skeaff, C.M., 2012. Influence of the glutathione peroxidase 1 Pro200Leu polymorphism on the response of glutathione peroxidase activity to selenium supplementation: a randomized controlled trial. *Am J Clin Nutr* 96 (4), 923–931. PMID: 22952184.

- Seddon, J.M., Silver, R.E., Rosner, B., 2016. Response to AREDS supplements according to genetic factors: survival analysis approach using the eye as the unit of analysis. *Br J Ophthalmol* 100 (12), 1731–1737.
- Xiong, Y.M., Mo, X.Y., Zou, X.Z., Song, R.X., Sun, W.Y., Lu, W., Chen, Q., Yu, Y.X., Zang, W.J., 2010. Association study between polymorphisms in selenoprotein genes and susceptibility to Kashin-Beck disease. *Osteoarthritis Cartilage* 18 (6), 817–824.

Polyphenols and Nutrigenetic/Nutrigenomic Associations With Obesity-Related Metabolic Diseases

Eun-Young Kwon, Myung-Sook Choi

Center for Food and Nutritional Genomics Research, Department of Food and Nutrition, Kyungpook National University, San-Kyuk Dong Puk-Ku, Daegu, Korea

List of Abbreviations

ACC acetyl-CoA carboxylase
 ADIPOQ adiponectin
 ADRB2 beta-adrenergic receptor 2
 AMPK activating adenosine-monophosphate-activated protein kinase
 BMI body mass index
 C/EBP α CCAAT/enhancer binding protein α
 COMT catechol-O-methyl transferase
 CPT carnitine palmitoyltransferase
 EC epicatechin
 EGCG epigallocatechin gallate
 FTO fat mass and obesity
 GLUT4 glucose transporter 4
 GTCs green tea catechins
 HNF4 α hepatocyte nuclear factor 4 α
 HOMA-IR homeostatic model assessment for insulin resistance
 IL interleukin
 Irs2 insulin receptor substrate 2
 LTA lymphotoxin alpha
 MC4R melanocortin-4 receptor
 MCP-1 monocyte chemoattractant protein 1
 MedDiet Mediterranean diet
 MTHFR methylenetetrahydrofolate reductase
 PGC PPAR γ coactivator
 PLIN perilipins
 PPAR γ peroxisome proliferator activator receptor γ
 SIRT1 sirtuin 1
 SNPs single-nucleotide polymorphisms
 SREBP-1c sterol regulatory element-binding protein 1c
 T2DM type 2 diabetes mellitus
 TCA tricarboxylic acid
 TFAM mitochondrial transcription factor A
 TG triglycerides
 TNF tumor necrosis factor
 UCP uncoupling protein

INTRODUCTION

Considerable attention has been focused on a thorough understanding of which dietary substances influence diseases induced or favored by specific nutritional imbalances through changes in the structure or expression of particular genes, and on the development of appropriate diagnostic tools to monitor the onset, incidence, progression, and/or severity of metabolic diseases, including obesity.

The discipline that studies the effects of genetic variation on cellular and organism responses to dietary intervention is known as nutritional genomics. Nutritional genomics includes two parts: nutrigenomics involves the study of interactions between the nutritional environment and the genome, as well as the regulation of changes in proteins and other metabolic processes; nutrigenetics examines how single-nucleotide polymorphisms (SNPs) or other gene variants alter the bioactivity of important metabolic pathways and result in distinct nutritional requirements or specificities in the response to nutrients. Studies suggested that various SNPs interact with environmental factors to influence physiological parameters including lipid metabolism, inflammatory responses, and obesity-related diseases.

Among dietary components, a wide variety of dietary chemical substances derived from a plant-based diet, collectively known as polyphenols, are effective and beneficial modulators of cellular function and contribute to the prevention of obesity and its complications.

Polyphenols are secondary metabolites of plants and comprise a large group of compounds with aromatic ring(s), characterized by the presence of one or more hydroxyl groups and varying structural complexities. The most widely distributed group of plant phenolics is flavonoids, which contains the subclasses flavonols, flavones, isoflavones, flavanones, anthocyanidins, flavanols, and others. Gene polymorphisms and polyphenol intake have been studied independently in association with lipid profiles and susceptibility to obesity or other metabolic diseases. Observational nutrigenetic studies investigated whether interactions between polyphenol intake and some genetic variants can modulate target markers of obesity in specific populations.

Protective roles of variants in some genes affecting the obesity risk under high polyphenol consumption have been reported. Gene variants may represent novel biomarkers for classifying individuals who might benefit from targeted dietary recommendations for health promotion and preventive strategies. The individual susceptibility to obesity depends on the genetic predisposition of energy balance regulation. Dietary polyphenols can affect polymorphisms in genes encoding taste receptors, peripheral signaling peptides such as insulin, leptin, ghrelin, and corresponding receptors, and a number of proteins related to lipid metabolism. Here, we provide an overview of the metabolic action of polyphenols on intracellular molecular mechanisms from a nutritional genetics and genomics perspective.

GENETIC POLYMORPHISMS IN OBESITY-RELATED METABOLIC DISEASES

Like many other medical conditions, obesity and its complications such as insulin resistance and cardiovascular disease result from the interplay between genetic and environmental factors. Among the different mechanisms that can lead to obesity, polymorphisms in various genes controlling appetite and biological processes are involved, which in turn influence body composition (among which are fat metabolism and adipocyte differentiation). SNPs are the most common form of genetic variation, and specific genomic variations can explain some of the risk for and individual susceptibility to human obesity, inflammation, dyslipidemia, and metabolic diseases.

Peroxisome proliferator activator receptor γ (PPAR γ) is a master regulator of genes contributing to adipocyte differentiation, obesity susceptibility, and insulin sensitivity. A genetic polymorphism in PPAR γ , Pro12Ala (a substitution of proline to alanine at codon 12 [P12A], rs1801282), is associated with obesity and metabolic syndrome, but the frequency of this polymorphism varies according to the ethnic group. Studies showed that the

presence of the Ala allele predict higher body mass index (BMI) in Iranian individuals (Mirzaei et al., 2009), and the Ala allele interacts with gender and contributes to the susceptibility to obesity in white Italian men (Morini et al., 2008). However, the P12A SNP was not associated with obesity, but with the genetic risk for type 2 diabetes mellitus (T2DM) in subjects with obesity in a French Caucasian population (Ghoussaini et al., 2005).

Human obesity is associated with decreased lipolysis and increased adipogenesis in human fat cells. Acetyl-coenzyme A (CoA) carboxylase (ACC) has an essential role in regulating fatty acid synthesis and degradation, and catalyzes the committed step in fatty acid synthesis, namely the production of malonyl-CoA (activated two-carbon donor). Carnitine palmitoyltransferase (CPT) has been implicated in mitochondrial oxidation, and perilipins (PLIN), proteins encoded by the PLIN genes, clearly have a role in adipocyte metabolism and lipolysis, in turn influencing obesity and insulin resistance (IR). Moreover, gene polymorphisms in ACC2 (rs4766587 AA genotype), CPT1A (rs28936372), CPT1B (rs470017), and PLIN (A allele carriers [AA and AG] of the rs894160) are associated with obesity in several populations. In addition, among obesity-associated gene polymorphisms, those in β -adrenergic receptor 2 (ADRB2) and 3 (ADRB3) lead to amino acid changes associated with lipolysis in human fat cells and transfected cells. SNPs in the ADRB2 gene (alleles at nonsynonymous SNP) and ADRB3 gene (coding allele, Arg64) are associated with weight gain and reduced basal metabolic rate, as well as obesity and reduced lipolysis. Hormone-sensitive lipase is the rate-limiting step in lipolysis, hydrolyzing triglycerides (TG) to diglycerides and subsequently monoglycerides, and hormone-sensitive lipase gene polymorphisms are associated with low lipolysis in human fat cells.

Phillips et al. also found a positive relationship between polymorphisms in lymphotoxin α (LTA) (rs915654 A allele genotype), tumor necrosis factor (TNF)- α (rs1800629 GG genotype), and interleukin (IL)-6 (rs1800795 CC genotype), and obesity (Phillips et al., 2010). TNF- α and IL-6 are secreted by a wide variety of cells, such as endothelial cells, osteoblasts, myocytes, adipocytes, and macrophages, as well as numerous other tissues. Obesity leads to increased production of TNF- α and IL-6, which are crucial mediators of the IR state and other related comorbidities. LTA, like TNF- α and IL-6, encodes for a cytokine involved in inflammatory responses and has been implicated in several metabolic disorders including obesity and IR. In addition, a metaanalysis suggests that adiponectin (ADIPOQ) rs17300539 and rs1501299 are associated with obesity risk in Caucasian people, and rs266729 is associated with obesity risk in Asian people (Lu et al., 2014). ADIPOQ is a hormone exclusively expressed in

differentiated adipocytes; it has been primarily implicated in insulin sensitivity, but it is also a candidate for obesity. A consequence of obesity is the development of cardiovascular disease. Polymorphic genes involved in regulating energy balance can also affect the development of cardiovascular disease.

LINKS BETWEEN NUTRITION AND GENOMICS

Nutrition is one of the most important factors contributing to obesity susceptibility, and food consumption can modify the patterns of gene expression thus influencing the phenotype. In addition, the genetic makeup of an individual may coordinate the response to diet. Many human studies have reported that SNP analysis can be used to investigate the role of nutrition in human health and diseases and the identification of optimal diets. SNPs in various genes can alter the metabolic response and influence nutrient interactions.

Ortega-Azorin et al. suggested that the association of the fat mass and obesity-associated (FTO)-rs9939609 and melanocortin-4 receptor (MC4R)-rs17782313 polymorphisms with type 2 diabetes depends on diet, and that a high adherence to a Mediterranean diet (MedDiet), which is rich in folate, counteracts genetic predisposition. When adherence to the MedDiet was low, carriers of variant alleles had a higher risk for T2DM. Conversely, when adherence to MedDiet was high, these associations disappeared (Ortega-Azorín et al., 2012). Numerous studies also investigated the relation between the FTO-rs9939609 polymorphism and BMI, and whether this association is modulated by total fat intake and low-fiber consumption (Razquin et al., 2011; Phillips et al., 2012). Obesity susceptibility genes FAIM2, FLJ35779, FTO, LRRN6C, RBJ, and SEC16B were found to interact with dietary carbohydrates to increase BMI. Other gene–diet interaction studies related to obesity demonstrated that homozygotes for the P allele in the PPAR γ gene (P12A, rs1801282) who consumed a high-fat diet were at higher risk for central obesity, whereas the risk was lower when the intake of monounsaturated fatty acids was high.

Methylenetetrahydrofolate reductase (MTHFR) catalyzes the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, the major form of circulating folate in plasma. The C677T polymorphism in the MTHFR gene is associated with reduced MTHFR enzyme activity and impaired folate accumulation, which may increase homocysteine levels in plasma, increasing the risk of venous thromboembolic and arterial ischemic heart disease (Subbiah, 2007).

Examples of common polymorphisms in gene–diet interactions in relation to obesity-related metabolic diseases are listed in Table 44.1.

POLYPHENOLS STUDIED WITH RESPECT TO NUTRIGENETICS AND NUTRIGENOMICS

Nutrients and bioactive compounds such as polyphenols can interfere with the genome in highly complex forms and may in turn be influenced by genetic polymorphisms that lead to important interindividual differences in the capacity to metabolize polyphenols.

CATECHINS

Catechins, the major active components of green tea, are monomeric flavanols composed of chemically similar compounds such as catechin, epicatechin, epigallocatechin, epicatechin gallate, and epigallocatechin gallate (EGCG). EGCG predominates among the green tea catechins (GTCs) and accounts for most of the therapeutic effect of green tea. A routine intake of green tea exerts antiobesity effects by suppressing adipocyte differentiation and proliferation, inhibiting fat absorption from the gut and lipogenesis, and promoting lipolysis and fatty acid β -oxidation (Meydani and Hasan, 2010). The mechanism of the antiobesity action of GTCs, particularly EGCG, was previously reviewed (Meydani and Hasan, 2010). EGCG has been shown to inhibit adipocyte differentiation by activating adenosine monophosphate–activated protein kinase (AMPK), a suppressor of lipid accumulation, by modulating downstream-signaling components. AMPK downregulates the expression of PPAR γ and CCAAT/enhancer binding protein α (C/EBP α) at the messenger RNA (mRNA) and protein levels. These transcription factors orchestrate the expression of adipogenic and lipogenic factors, including ACC and the transcriptional factor sterol regulatory element-binding protein 1c (SREBP-1c). Therefore, EGCG and other GTCs contribute to the reduction of adipogenesis and prevention of the growth and expansion of adipose tissue by activating AMPK and inhibiting PPAR γ and C/EBP α .

Complementing the effect of EGCG on adipogenesis is the ability of this molecule to inhibit pancreatic lipase activity, resulting in reduced fat absorption from the gut and enhanced fat oxidation and thermogenesis (Meydani and Hasan, 2010). EGCG increased glycerol release and the expression of hormone-sensitive lipase and CPT-1, which enhance lipolysis and fatty acid β -oxidation. In addition, the mRNA levels of uncoupling protein (UCP)-2, a key protein in fat-supported thermogenesis, were increased by EGCG, and the activities of catechol-*O*-methyl transferase (COMT), an enzyme that inhibits fatty acid oxidation in brown adipose tissue, was suppressed by EGCG. Moreover,

TABLE 44.1 Examples of Interaction Between Diet and Variants of Genes in Relation to Obesity-Related Metabolic Diseases.

Gene	Variant	Population	Main Findings
Fat mass and obesity	rs9939609	7052 subjects from Spain	When adherence to mediterranean diet was low, type 2 diabetes risk was high in subjects carrying risk allele.
		13,000 subjects from France	When SFA intake was high, waist circumference was low in A allele carriers.
M melanocortin-4 receptor	rs17782313	7052 subjects from Spain	When adherence to mediterranean diet was low, type 2 diabetes risk was high in subjects carrying risk allele.
PPARG	rs1801282	1628 subjects from French Canada	When total intake fat was high, waist circumference was high in P12/P12 homozygotes.
		538 subjects from Spain	When MUFA intake was low, HOMA-IR index was high in obese Ala carriers.
		926 subjects from Italy	Alcohol consumption was associated with higher HDL levels in Ala carriers
		313 subjects from Spain	When intake of carbohydrates was high, obesity risk was high in Ala carriers
PLIN	rs894160	920 subjects from Puerto Rico	When complex carbohydrates intake was low, waist circumference was high in A allele carriers.
ACC2	rs4766587	464 subjects from France	When PUFA intake was high, waist circumference was high in AA homozygotes.
TNF	rs1800629	138 subjects from South Africa	When total fat intake was high, obesity risk was high in women with GA + AA genotype.

HDL, high-density lipoprotein; *HOMA-IR*, homeostasis model assessment of insulin resistance; *MUFA*, monosaturated fatty acids; *SFA*, saturated fatty acid.

COMT enzyme activity differs between Asian and Caucasian populations. Asian populations have a higher frequency of the thermostable, high-activity enzyme, encoded by the COMTH allele (Val/Val polymorphism), compared with Caucasian populations, which have a

higher frequency of the thermolabile, low-activity enzyme, encoded by the COMTL allele (Met/Met polymorphism). These genetic variations may be responsible for the differential effects of tea consumption on suppressing COMT activity and lipid metabolism. Green

tea or catechins intake is regarded as an effective method for reducing and maintaining body weight by increasing fat oxidation, thermogenesis, and energy expenditure.

In addition to these effects described, GTCs decreases the risk for IR and T2DM by increasing glucose uptake into the skeletal muscles through an upregulation of glucose transporter 4 (GLUT4) and by reducing the translocation of GLUT4 from the cell membrane and insulin levels in adipose tissues (Meydani and Hasan, 2010). Moreover, reduced expression of resistin and increased expression of adiponectin may contribute to the EGCG-mediated antiinflammatory response and decrease the risk for IR and type 2 diabetes.

RESVERATROL

Resveratrol (3,4',5-trihydroxystilbene), a naturally occurring polyphenolic compound, is a well-studied stilbenoid present in red grapes, red wine, peanuts, and ground nuts. Many in vitro and in vivo studies demonstrated that resveratrol has antiobesity potential via its ability to increase the phosphorylation and activation of AMPK, the master regulator of energy metabolism (Meydani and Hasan, 2010; Martins et al., 2012; Aguirre et al., 2014). Similar to EGCG, AMPK activated by resveratrol inhibits ACC and consequently blocks the production of malonyl-CoA, a stimulator of lipogenesis. Furthermore, upregulation of sirtuin 1 (SIRT1) and AMPK causes resveratrol to increase PPAR γ coactivator (PGC)-1 α , a cofactor in mitochondrial biogenesis and function, thus contributing to the suppression of lipid accumulation. Genetic studies revealed a strong correlation between obesity and SIRT1 gene polymorphisms. Studies showed that Japanese subjects with the A allele of SIRT1 polymorphism rs7895833, the G allele of rs7069102, and the T allele of rs2273773 had a higher risk for obesity. In French Caucasian adults, a strong association was detected between high BMI and the SIRT1 SNPs rs3395786 and rs11599176 (Martins et al., 2012). Moreover, a common SNP in a novel p53-binding sequence in the human SIRT-1 promoter was found to affect nutrients, including polyphenols such as resveratrol, and thus may significantly affect SIRT1-mediated changes in human metabolism and physiology.

Resveratrol also inhibits adipocyte differentiation by decreasing intracellular TG accumulation with concomitant suppression of expression of PPAR γ , C/EBP, and SREBP-1c. Downregulation of PPAR γ , C/EBP, and SREBP-1c expression is associated not only with downregulation of lipogenic genes, including fatty acid synthase, lipoprotein lipase, fatty acid binding protein, stearoyl-CoA desaturase-1, malic enzyme, and glucose-6 phosphate dehydrogenase, but also with resveratrol-mediated activation of AMPK and SIRT1. Microarray

analysis of genes in the liver revealed that resveratrol supplementation downregulates the expression of genes involved in lipogenesis (Aguirre et al., 2014).

The antiadipogenic and antilipogenic activities of resveratrol are further enhanced by resveratrol-stimulated lipolysis. Studies demonstrated that resveratrol enhances both adipose triglyceride lipase and SIRT1 gene and protein expression, associated with enhanced free fatty acid release (Aguirre et al., 2014). These effects are reversed by AMPK inhibitors, which suggests that the antiobesity action of resveratrol is mediated by AMPK. Resveratrol also increases fatty acid β -oxidation, mitochondrial biogenesis and activity, and the expression of thermogenesis-related genes including UCP-1, UCP-3, PGC-1 α , SIRT1, and mitochondrial transcription factor A (TFAM) (Aguirre et al., 2014). Resveratrol activates SIRT1, which in turn activates PGC-1 α by deacetylation; PGC-1 α is involved in regulating mitochondrial biogenesis via increased expression of TFAM. In the skeletal muscle, resveratrol led to increased gene expression of TFAM or cyclooxygenase-2, which has a key role in the oxidative phosphorylation pathway, suggesting increased mitochondriogenesis. In addition, resveratrol increases UCP-1 in the brown adipose tissue and UCP-3 in the skeletal muscle, which is mediated by increased expression of SIRT1. Increased levels of UCP-1 and UCP-3 by resveratrol, as well as increased mitochondriogenesis, were likely related to increased PGC-1 α and SIRT1. In the liver of obese mice treated with resveratrol, activities of CPT and acyl CoA oxidase, enzymes involved in mitochondrial β -oxidation and fatty acid oxidation, were significantly increased with increased expression of genes related to PPAR α . Consequently, resveratrol appears to increase fatty acid consumption in oxidative tissues (liver and skeletal muscle) by increasing mitochondriogenesis.

Regarding glucose metabolism, resveratrol has been suggested to increase glucose uptake by upregulating estrogen receptor- α , which in turn increases 44 expression through the phosphatidylinositol-3 kinase and AKT pathway. Resveratrol also prevents oxidative damage resulting from impaired glucose metabolism. These metabolic actions of resveratrol lead to reduced insulin secretion and hyperglycemia, and thus may prevent the onset of diabetic complications.

CURCUMIN

Curcumin, a yellow-colored hydrophobic polyphenol, is a member of the curcuminoid family present in the spice turmeric obtained from the rhizome of the perennial herb *Curcuma longa*. Curcumin is one of the most extensively studied polyphenols; it is known

to possess antiinflammatory, anticarcinogenesis, anti-obesity, antiangiogenesis, and antioxidant properties (Meydani and Hasan, 2010).

Studies demonstrated that, similar to GTCs and resveratrol, curcumin has a significant effect on adiposity through several mechanisms, including the inhibition of adipocyte differentiation, suppression of lipogenesis, promotion of fatty acid oxidation, and modulation of energy metabolism and inflammation (Meydani and Hasan, 2010). Curcumin significantly inhibited adipocyte differentiation with concomitant decreases in the expression of fatty acid binding protein 4 and C/EBP α , adipocyte differentiation biomarkers, and decreased fat accumulation by decreasing CD antigen 36 mRNA expression, a fatty acid transporter in adipocytes. This effect of curcumin may have been mediated by suppression in the expression of the PPAR γ transcription factor. In addition, curcumin increases the phosphorylation and activation of AMPK and downregulates ACC activity; consequently, it downregulates the flow of acetyl-CoA to malonyl-CoA. Through activation of AMPK, curcumin can also decrease the expression of fatty acid synthase and glycerol-3-phosphate acyl transferase-1, which in turn inhibits the synthesis of fatty acid and glycerol lipids.

Curcumin reduces adipogenesis and attenuates liver fatty acid synthesis by enhancing fatty acid β -oxidation in adipose tissue and the liver, with a concomitant increase in the expression of CPT-1, a key enzyme that transfers cytosolic long-chain fatty acyl CoA into the mitochondria for β -oxidation, and acetyl CoA oxidase, the first catalytic enzyme in fatty acid β -oxidation (Meydani and Hasan, 2010). Moreover, curcumin enhances dyslipidemia by increasing cholesterol efflux by upregulating the expression of the liver X receptor and adenosine triphosphate-binding cassette A1 in adipocytes, suppressing lipogenesis, and promoting fatty acid oxidation.

In addition to the antiobesity mechanisms described earlier, curcumin prevents obesity by suppressing angiogenesis. Angiogenesis has pivotal roles in the growth and expansion of the adipose tissue because angiogenesis is mediated by adipokine secretion in the adipose tissue. Curcumin suppresses angiogenesis in the adipose tissue by downregulating several factors, including vascular endothelial growth factor, basic fibroblast growth factor, epidermal growth factor, and hypoxia-inducible factor-1 α , thus preventing the growth of adipose tissue and obesity. Curcumin supplementation decreased serum glucose, increased glucose disposal, decreased homeostatic model assessment for insulin resistance (HOMA-IR), and ameliorated inflammation by lowering the expression of nuclear factor- κ B. Curcumin also reduced the infiltration of macrophages in the adipose tissue by inhibiting monocyte

chemoattractant protein 1 (MCP-1) release from adipocytes, and may partially contribute to its antiinflammatory activity.

LUTEOLIN

Luteolin (3',4',5,7-tetrahydroxyflavone) is a yellow crystal in its pure form and is a food-derived flavone typically present in its glycosylated form in celery, green pepper, perilla leaf, and chamomile tea. Luteolin possesses a variety of antitumor, antiinflammatory, antimicrobial, and antidiabetic properties. Luteolin inhibits the interaction between adipocytes and macrophages to suppress the production of inflammatory mediators such as TNF- α , MCP-1, and IL-6 from macrophages by inhibiting the phosphorylation of c-Jun N-terminal kinases. Particularly in humans, higher circulating IL-6 levels have been associated with obesity and T2DM, and the CC genotype (rs1800797) of the IL-6 gene was related to a higher risk for obesity (Möhlig et al., 2004). Barth et al. (Barth et al., 2012) reported that intake of polyphenol-rich apple juice containing luteolin in obese subjects with the C/C variants of the IL-6 gene significantly reduced body fat after 4 weeks of intervention. Furthermore, luteolin exerts hypolipidemic effects by suppressing the transcription factor hepatocyte nuclear factor 4 α (HNF4 α). HNF4 α is a nuclear receptor that regulates the expression of genes involved in secreting apolipoprotein B-containing lipoproteins, such as microsomal triglyceride transfer protein. Treatment of mice with luteolin significantly suppressed the expression of HNF4 α target genes in the liver of animals with diet-induced obesity and improved serum glucose and lipid parameters (serum very low-density lipoprotein and low-density lipoprotein cholesterol and apolipoprotein B protein levels), which was not accompanied by hepatic fat accumulation.

Luteolin is a PPAR γ ligand with antioxidant and anti-inflammatory properties. However, unlike thiazolidinediones, it does not promote adipocyte differentiation (Kwon et al., 2015). Luteolin can function as a PPAR agonist by enhancing the expression of PPAR γ -dependent genes in adipocytes. Kwon et al. (2015) reported that luteolin increased PPAR γ protein expression, thus attenuating hepatic steatosis by redistributing fatty acids from the liver to adipose tissue and enhancing the expression of fatty acid uptake genes (fatty acid binding protein, CD36, and lipoprotein lipase) and lipogenic proteins (SREBP1, FAS, and ACC) in adipose tissue. Although Kwon et al. (2015) showed that luteolin activates lipogenesis, thus preventing hepatic lipotoxicity, it simultaneously increases lipolysis and the tricarboxylic acid (TCA) cycle, which contributes to reduced adiposity. The antiobesity effect of luteolin is

accompanied by increasing energy expenditure and activation of browning and thermogenesis through an AMPK/PGC1 α pathway-mediated mechanisms.

The striking improvement in hepatic steatosis coupled with decreased adiposity by luteolin was associated with prevention of insulin resistance (Kwon et al., 2015). Luteolin increased hepatic phosphatidylcholine metabolite and the expression of insulin receptor substrate 2 (Irs2) mRNA. When membrane phosphatidylcholine concentration is low, the transcription of SREBP1 is activated. In the liver, SREBP1 modulates Irs2 expression through its negative feedback at an insulin response element on the Irs2 promoter; decreased Irs2 expression can lead to insulin resistance. In addition, luteolin enhances 4 through the phosphatidylinositol-3K signaling pathway and attenuates inflammatory responses by suppressing nuclear factor- κ B and mitogen-activated protein kinase activation

in adipocytes. This observation is linked to the prevention of IR and glucose intolerance.

Major mechanisms involved in the antiobesogenic effect of polyphenols, particularly catechins, resveratrol, curcumin, and luteolin, are given in Fig. 44.1.

CONCLUSIONS AND LIMITATIONS

Nutrigenomic and nutrigenetic research in obesity and the cardiometabolic syndrome has provided insights into how particular nutrients and nonnutritive food components such as polyphenols have an important role in detecting and modulating the cellular sensor system influencing individual susceptibility to developing these conditions. Understanding the relations within the diversity in human genetics, genome function, and dietary components should enable more

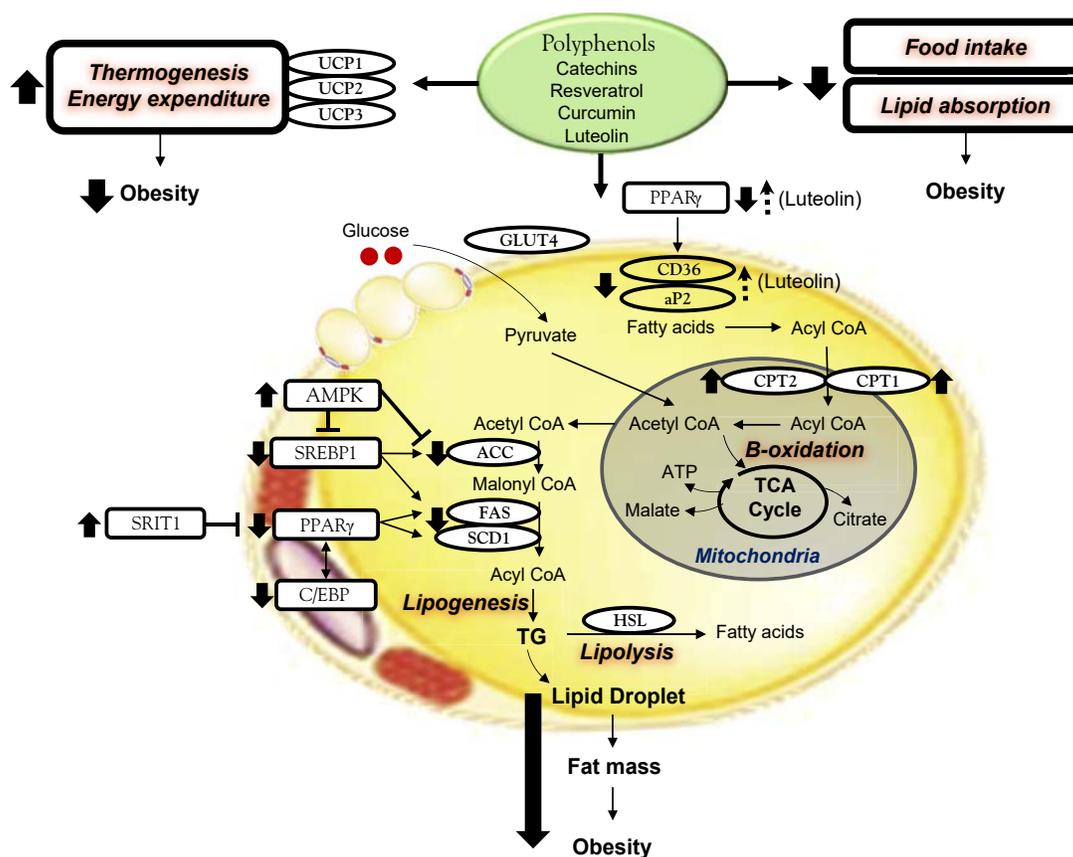


FIGURE 44.1 Major mechanisms involved in antiobesogenic effect of polyphenols, particularly catechins, resveratrol, curcumin, and luteolin. These polyphenols decrease food intake and lipid absorption while increasing energy expenditure by upregulating uncoupled proteins. They also inhibit adipogenesis and lipogenesis and activate lipolysis and fatty acid β -oxidation. However, luteolin activates lipogenesis to prevent hepatic lipotoxicity; it also simultaneously increases lipolysis and the tricarboxylic acid (TCA) cycle. These actions would lead to decreased adipocyte size and number and lowered fat mass, which should contribute to reduced risk for obesity. AMPK, adenosine monophosphate-activated protein kinase; ATP, adenosine triphosphate; CoA, coenzyme A; CPT, carnitine palmitoyltransferase; FAS, fatty acid synthase; GLUT4, glucose transporter 4; HSL, hormone-sensitive lipase; PPAR γ , peroxisome proliferator activator receptor γ ; SCD1, stearoyl-CoA desaturase; SIRT1, sirtuin 1; SREBP1, sterol regulatory element-binding protein 1; TG, triglycerides; UCP, uncoupling protein.

tailored dietary recommendations for optimal human health and preventing obesity-induced comorbidities. So far, however, few studies have investigated differences in the effects of nutrient and nonnutritive food components such as polyphenols on obesity and related metabolic diseases according to individual gene haplotypes and SNPs selected because of their possible influence on human obesity. Moreover, the relevance and magnitude of nutrient–gene interactions require further analysis. It is important to evaluate specific polyphenol functions on nutrigenetics to advise different lifestyles. Advances in understanding the complex interplay among the genotype, phenotype, and polyphenols will contribute to a healthy lifestyle to prevent obesity-related metabolic diseases.

Acknowledgments

This work was supported by grants funded by the National Research Foundation of Korea (NRF) by the Korea government (MSIP) (2015R1A5A6001906; NRF-2012M3A9C4048818; NRF-2015R1C1A2A01051533).

References

- Aguirre, L., Fernández-Quintela, A., Arias, N., Portillo, M.P., 2014. Resveratrol: anti-obesity mechanisms of action. *Molecules* 19, 18632–18655.
- Barth, S.W., Koch, T.C., Watzl, B., Dietrich, H., Will, F., Bub, A., 2012. Moderate effects of apple juice consumption on obesity-related markers in obese men: impact of diet-gene interaction on body fat content. *Eur J Nutr* 51, 841–850.
- Ghousaini, M., Meyre, D., Lobbens, S., Charpentier, G., Clément, K., Charles, M.A., Tauber, M., Weill, J., Froguel, P., 2005. Implication of the Pro12Ala polymorphism of the PPAR-gamma 2 gene in type 2 diabetes and obesity in the French population. *BMC Med Genet* 6, 11.
- Kwon, E.Y., Jung, U.J., Park, T., Yun, J.W., Choi, M.S., 2015. Luteolin attenuates hepatic steatosis and insulin resistance through the interplay between the liver and adipose tissue in mice with diet-induced obesity. *Diabetes* 64, 1658–1669.
- Lu, J.F., Zhou, Y., Huang, G.H., Jiang, H.X., Hu, B.L., Qin, S.Y., 2014. Association of ADIPOQ polymorphisms with obesity risk: a meta-analysis. *Hum Immunol* 75, 1062–1068.
- Martins, I., Wilson, A., Lim, W.L., Laws, S., Fuller, S., Martins, R., 2012. Sirtuin-1 mediates the obesity induced risk of common degenerative diseases: Alzheimer’s disease, coronary artery disease and type 2 diabetes. *Health* 4, 1448–1456.
- Meydani, M., Hasan, S.T., 2010. Dietary polyphenols and obesity. *Nutrients* 2, 737–751.
- Mirzaei, H., Akrami, S.M., Golmohammadi, T., Doosti, M., Heshmat, R., Nakhjavani, M., Amiri, P., 2009. Polymorphism of Pro12Ala in the peroxisome proliferator-activated receptor gamma2 gene in Iranian diabetic and obese subjects. *Metab Syndr Disord* 7, 453–458.
- Möhlig, M., Boeing, H., Spranger, J., Osterhoff, M., Kroke, A., Fisher, E., Bergmann, M.M., Ristow, M., Hoffmann, K., Pfeiffer, A.F., 2004. Body mass index and C-174G interleukin-6 promoter polymorphism interact in predicting type 2 diabetes. *J Clin Endocrinol Metab* 89, 1885–1890.
- Morini, E., Tassi, V., Capponi, D., Ludovico, O., Dallapiccola, B., Trischitta, V., Prudente, S., 2008. Interaction between PPARgamma2 variants and gender on the modulation of body weight. *Obesity* 16, 1467–1470.
- Ortega-Azorín, C., Sorlí, J.V., Asensio, E.M., Coltell, O., Martínez-González, M.Á., Salas-Salvadó, J., Covas, M.I., Arós, F., Lapetra, J., Serra-Majem, L., Gómez-Gracia, E., Fiol, M., Sáez-Tormo, G., Pintó, X., Muñoz, M.A., Ros, E., Ordovas, J.M., Estruch, R., Corella, D., 2012. Associations of the FTO rs9939609 and the MC4R rs17782313 polymorphisms with type 2 diabetes are modulated by diet, being higher when adherence to the Mediterranean diet pattern is low. *Cardiovasc Diabetol* 11, 137.
- Phillips, C.M., Goumidi, L., Bertrais, S., Ferguson, J.F., Field, M.R., Kelly, E.D., Mehegan, J., Peloso, G.M., Cupples, L.A., Shen, J., Ordovas, J.M., McManus, R., Hercberg, S., Portugal, H., Lairon, D., Planells, R., Roche, H.M., 2010. Additive effect of polymorphisms in the IL-6, LTA, and TNF- α genes and plasma fatty acid level modulate risk for the metabolic syndrome and its components. *J Clin Endocrinol Metab* 95, 1386–1394.
- Phillips, C.M., Kesse-Guyot, E., McManus, R., Hercberg, S., Lairon, D., Planells, R., Roche, H.M., 2012. High dietary saturated fat intake accentuates obesity risk associated with the fat mass and obesity-associated gene in adults. *J Nutr* 142, 824–831.
- Razquin, C., Marti, A., Martinez, J.A., 2011. Evidences on three relevant obesogenes: MC4R, FTO and PPAR γ . Approaches for personalized nutrition. *Mol Nutr Food Res* 55, 136–149.
- Subbiah, M.T., 2007. Nutrigenetics and nutraceuticals: the next wave riding on personalized medicine. *Transl Res* 149, 55–61.

Caffeine

Giulia Renda¹, Raffaele De Caterina²

¹G. d'Annunzio University of Chieti and Center of Excellence on Aging CeSI-Met, Chieti, Italy ²Cattedra di Malattie dell'Apparato Cardiovascolare, Università degli Studi di Pisa, Italy

INTRODUCTION

Caffeine is a central nervous system (CNS) stimulant of the methylxanthine class. Caffeine is indeed a methylxanthine alkaloid (1,3,7-trimethyl xanthine) (Fig. 45.1), and is chemically related to the purine adenine and guanine bases of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). It is contained in the seeds, nuts, and leaves of a number of plants native to South America and East Asia. The most abundant source of caffeine is the seed of *Coffea* plants, belonging to the botanical family of Rubiaceae, which includes about 80 species. The two main species are *Coffea arabica* (Arabica coffee) and *Coffea canephora* (Robusta coffee), accounting for 75% and 20% of the global coffee production, respectively. Caffeine-containing drinks, such as coffee, tea, and cola, are very popular. Particularly, coffee, after water, is probably the most frequently used beverage worldwide. Beverages containing caffeine are assumed to relieve or prevent drowsiness and to improve performance. Caffeine is indeed the most widely consumed psychoactive drug, legal and unregulated in nearly all parts of the world. To make these drinks, caffeine is extracted by steeping the plant product in water, a process called infusion, or by steam extraction at higher temperatures.

MECHANISM OF ACTION

Caffeine is quickly absorbed through the gastrointestinal tract and metabolized in the liver by cytochrome P450 1A2 (CYP1A2), resulting in three metabolites: paraxanthine, theophylline, and theobromine (Fig. 45.1). Elevated levels appear in the bloodstream within 15–45 min of consumption, and peak concentrations are evident 1 h post ingestion, also crossing the blood-brain barrier due to its lipid solubility. Caffeine and its metabolites are excreted by the kidneys, and concentrations decrease by 50%–75% within 3–6 h of consumption.

The main mechanism of action of caffeine is adenosine receptor antagonism. A secondary effect is the inhibition of phosphodiesterases, with the subsequent accumulation of cyclic AMP, and an intensification of the effects of catecholamines, the levels of which increase after caffeine (Fig. 45.2). Such properties translate, in most people, in a cognitive response, including increased alertness and attention, and in a complex cardiovascular response, mainly consisting of increased blood pressure (Tofalo et al., 2016).

Many caffeine psychoactive effects are thought to be mediated by the two adenosine receptor subtypes A1 and A2A, both expressed and differently distributed in the human brain, and thought to play an important role in the sleep–wake regulation. Since adenosine slows down neural activity, caffeine receptor antagonism attenuates these inhibitory effects, resulting in wake-promoting effects and performance benefits (Svenningsson et al., 1997).

Controlled studies have found conflicting results concerning psychoactive responses to caffeine, and both positive and null effects of caffeine on cognitive performance have been reported. A recent randomized controlled study has indicated that caffeine improved the attention level, and has confirmed that tasks involving alerting, orienting, and the (verbal) executive control components of attention are positively affected by caffeine consumption (Renda et al., 2015). Furthermore, research studies involving caffeine supplementation and physical performance overall have indicated a combined effect on both the central and peripheral systems. Therefore, it is possible that caffeine acts on the CNS as an adenosine antagonist, but may also have an effect on substrate metabolism and neuromuscular function (Goldstein et al., 2010).

HEALTH EFFECTS OF CAFFEINE

Health effects of caffeine intake are wide-ranging. Cardiovascular effects have been extensively

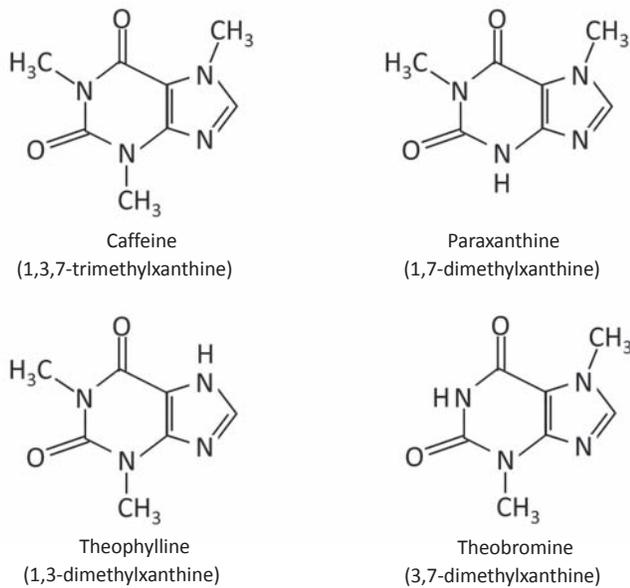


FIGURE 45.1 Chemical structure of caffeine (1,3,7-trimethylxanthine) and its three metabolites: paraxanthine (1,7-dimethylxanthine), theophylline (1,3-dimethylxanthine), and theobromine (3,7-dimethylxanthine).

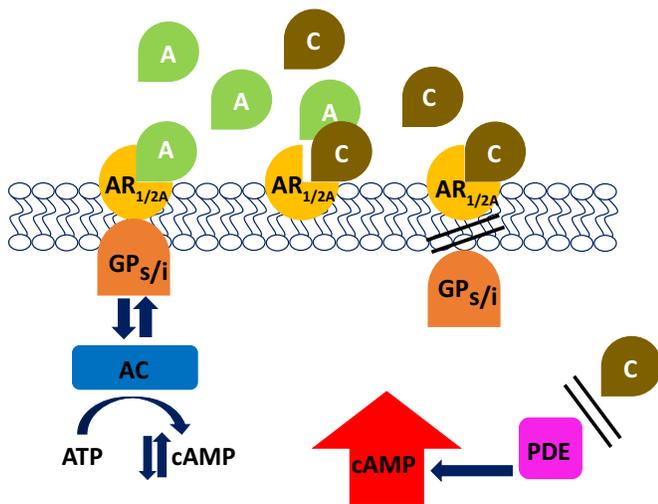


FIGURE 45.2 Main mechanisms of action of caffeine. The A₁ or A_{2A} adenosine receptor antagonism prevents the activation of the G protein signaling and the consequent increase or decrease of adenylyl cyclase activity, according to receptor and G protein subtype. On the other hand, the inhibition of phosphodiesterases causes the accumulation of cyclic AMP, and an intensification of the effects of catecholamines. A, adenosine; AC, adenylyl cyclase; AR_{1/2A}, adenosine receptors 1 or 2A; ATP, adenosine triphosphate; C, caffeine; cAMP, cyclic adenosine monophosphate; GP_{s/i}, G protein stimulatory or inhibitory; PDE, phosphodiesterase.

investigated, but are still debated, because underlying mechanisms are complex and incompletely understood, and because of the inconsistency of study results and the considerable variability in individual responses.

Most studies have aimed at evaluating the effects of coffee, because this is the most consumed beverage containing caffeine. Experimental data from short-term and animal studies indicated coffee as a potential cardiovascular risk factor, explained by unfavorable effects of caffeine on blood pressure, insulin resistance, and arrhythmias (Rebello and van Dam, 2013). Moreover, evidence from the older case-control studies suggested that coffee consumption is associated with a higher cardiovascular risk (Sofi et al., 2007). On the other hand, prospective cohort studies generally did not support the existence of an association between coffee consumption and cardiovascular risk, indicating that on the average, for most healthy people, moderate coffee consumption is unlikely to adversely affect cardiovascular outcome, including coronary heart disease, congestive heart failure, arrhythmias, and stroke (Rebello and van Dam, 2013).

The cardiovascular risk factor more extensively studied in relation to coffee is blood pressure, a risk factor for coronary heart disease, congestive heart failure, stroke, kidney disease, and all-cause death. Since caffeine is an antagonist of adenosine receptors and increases catecholamine levels, it can induce vasoconstriction, and thus elevates systolic and diastolic blood pressure acutely. However, whether coffee consumption has chronic effects on blood pressure and cardiovascular disease remains controversial.

Randomized controlled trials, which are mostly of short duration, have shown that coffee intake is associated with a small increase in blood pressure when compared with abstinence or with the use of decaffeinated coffee. Conversely, observational studies have indicated that the risk of a sustained increase in blood pressure may be lower both in abstainers and in subjects with a relatively high intake, although this was not found in all studies (Renda and De Caterina, 2015). A recent meta-analysis of 20 randomized, controlled trials and five cohort studies has shown no clinically important effects of long-term coffee consumption on blood pressure or the risk of hypertension in coffee consumers (Steffen et al., 2012).

Besides blood pressure, other cardiovascular risk factors may be affected by coffee consumption. Recent evidence from prospective cohort studies has suggested an inverse association between coffee and the risk of type 2 diabetes mellitus, but this was attributed to some effects on insulin sensitivity and/or insulin secretion modulated by different minerals, antioxidants, and phytochemical compounds found in coffee (Rebello and van Dam, 2013). On the other hand, the role of caffeine in increasing or decreasing the risk of type 2 diabetes is still poorly understood. In short-term metabolic studies, caffeine intake acutely reduced insulin sensitivity, exaggerating the blood glucose response to glucose loads. This action was related to the adenosine receptor antagonism of caffeine and/or to its ability in increasing adrenaline levels (Rebello and van Dam, 2013).

The effects of coffee on lipid profile depend on how the beverage is prepared. It was observed that the consumption of boiled coffee dose-dependently increases serum total and LDL cholesterol concentrations, whereas the consumption of filtered coffee resulted in very little change in serum cholesterol, according to the time of contact with hot water. In fact, cholesterol-increasing compounds classified as diterpenes (cafestol and kahweol) are extracted from the coffee beans by prolonged contact with hot water, while the much shorter contact required for brewed/filtered preparation and the retention by filter paper generates a much lower concentration of these compounds (Rebello and van Dam, 2013).

Inconsistent evidence is available about the association between coffee consumption and the risk of stroke. Two recent meta-analyses have concluded that moderate coffee consumption has a preventive effect on stroke, probably mediated by the healthy effect of coffee on main risk factors for stroke, including hypertension, cardiovascular diseases, and diabetes (Kim et al., 2012; Larsson and Orsini, 2011).

Similarly, the beneficial effects on blood pressure and other cardiovascular risk factors may contribute to positively affect the risk of heart failure. Indeed, another meta-analysis has indicated that there is a modest inverse association between moderate coffee consumption and the incidence of heart failure (Mostofsky et al., 2012). Current heart failure prevention guidelines have thus suggested a revision of the warning about coffee consumption (Schocken et al., 2008).

The relation between coffee consumption and the risk of arrhythmias has also been investigated, and recent studies have suggested that coffee does not increase arrhythmias. To the contrary, long-term coffee drinking might actually reduce the risk of abnormal cardiac rhythms, including atrial fibrillation.

Controlled interventional studies have shown that, in normal adults, even acute high-dose caffeine does not affect cardiac rhythm and rate, and does not cause clinically significant ventricular or supraventricular arrhythmias. Moreover, prospective cohort studies did not find significant associations between coffee consumption and the risk of atrial fibrillation. The mechanisms conferring potential protection against arrhythmias are still largely unknown, but it has been hypothesized that caffeine attenuates negative effects of endogenous adenosine on cardiac electrical conduction (Pelchovitz and Goldberger, 2011).

In addition, large epidemiological studies suggest that regular coffee drinkers may have reduced risk of death, both from cardiovascular and other causes.

Potential benefits of chronic coffee intake also include protection against neurodegenerative diseases:

Coffee consumption has indeed been inversely associated with the risk of Alzheimer disease, due to a

caffeine-related increase in plasma cytokines. Coffee and caffeine intake was also associated with reduced risk of Parkinson disease, due to the antagonism on the adenosine 2A (A2A) receptors in the brain (Tofalo et al., 2016).

Coffee is also known as a dietary source of antioxidants and free radical scavengers, such as caffeine, together with other compounds, having antimicrobial, anticariogenic, antiinflammatory, antihypertensive, and antiglycation activities, also derived from the roasting process. These properties might also explain a protective effect with cancers, particularly hepatocellular and kidney cancers, although caffeine can both stimulate and suppress tumors, depending on the species and phase of administration (Tofalo et al., 2016).

One has to recognize that most data on coffee health effects are based on observational data, with very few randomized, controlled studies; furthermore, associations found do not prove causation. An additional problem is that considerable interindividual variability in responses to coffee drinking has been observed. This is in part due to tolerance, but there is now evidence that some variability may have a genetic basis.

DETERMINANTS OF THE INTERINDIVIDUAL VARIABILITY IN RESPONSE TO CAFFEINE

Large interindividual variability in responses to caffeine have been reported, and a wide inconsistency in the results of numerous studies aimed at evaluating the effect of coffee/caffeine on health have been observed. The inconsistency in the evidence for several of its effects is probably influenced by various confounding factors.

First of all, although many effects of coffee have been attributed to caffeine, hundreds of other chemical substances—many of which are pharmacologically active—contained in coffee could have affected the observed results.

A second cause of variability is the fact that populations examined (differing in age, sex, usual frequency of coffee drinking, and smoking), as well as types of coffee blends and types of preparations, could all have influenced the obtained data. Importantly, differences in the study design (observational vs. cohort studies vs. randomized clinical trials, these latter with coffee or caffeine in tablets vs. a control treatment of decaffeinated coffee or placebo) could all have affected findings about the relation between coffee and cardiovascular risk.

Additionally, the variability observed in the results of coffee studies may be in part explained by the development of tolerance. Tolerance to the caffeine-induced

pressor effect quickly develops in habitual coffee drinkers, possibly because the complex set of counterregulatory hormones maintaining blood pressure may induce tolerance to the hemodynamic effects of caffeine, or also because the caffeine pressor effects might be counterbalanced by other coffee ingredients, including chlorogenic acid, flavonoids, melanoidins, quinide, magnesium, cafestol, and kahweol, or potassium (Tofalo et al., 2016).

Finally, one recently unraveled source of variability is the different genetic backgrounds of coffee drinkers.

THE NUTRIGENETICS OF CAFFEINE

An important explanation for the heterogeneity in the responses to caffeine is the interaction between individual genetic traits and caffeine itself, the so-called nutrigenetics of caffeine. Genetic interindividual differences in caffeine metabolism or caffeine effectors, such as the adenosine or adrenergic receptors, could be responsible for different reactions to caffeine intake.

Genetic polymorphisms related to caffeine metabolism may affect various cardiovascular responses to coffee drinking. Caffeine is mainly metabolized by cytochrome P450 1A2 (CYP1A2) in the liver. An A → C substitution at position 734 (subjects with the CYP1A2*1F variant) identifies slow caffeine metabolizers, with decreased enzyme inducibility, in turn resulting in impaired caffeine metabolism compared with rapid caffeine metabolizers, carrying the wild-type A2*1A allele (Fig. 45.3) (Sachse et al., 1999). The CYP1A2*1F variant seems to be associated with an increased risk of hypertension (Palatini et al., 2009), with carriers of the slow-metabolism *1F allele at increased risk with higher coffee intake compared with individuals with the fast-metabolism *1A/*1A genotype. The same slow CYP1A2 *1F allele was related to an increased risk of impaired fasting glucose in hypertension, with the highest risk in heavy drinkers (Palatini et al., 2015). As a matter of fact, the intake of coffee was associated with an increased risk of nonfatal myocardial infarction only among individuals with slow caffeine metabolism (Cornelis et al., 2006). For these reasons, carriers of CYP1A2 *1F variant might in principle derive a benefit from abstaining from coffee, whereas rapid caffeine metabolizers might safely drink coffee.

One study also examined whether the relation between coffee intake and incident coronary heart disease is dependent on the metabolism of catecholamines, specifically polymorphisms of the catechol-O-methyltransferase (COMT) gene. In either homozygotes or heterozygotes for the high-activity COMT allele, a substantial coffee intake did not increase the incidence of acute coronary events, while in homozygotes for the low-activity COMT allele, heavy coffee consumption was associated with a

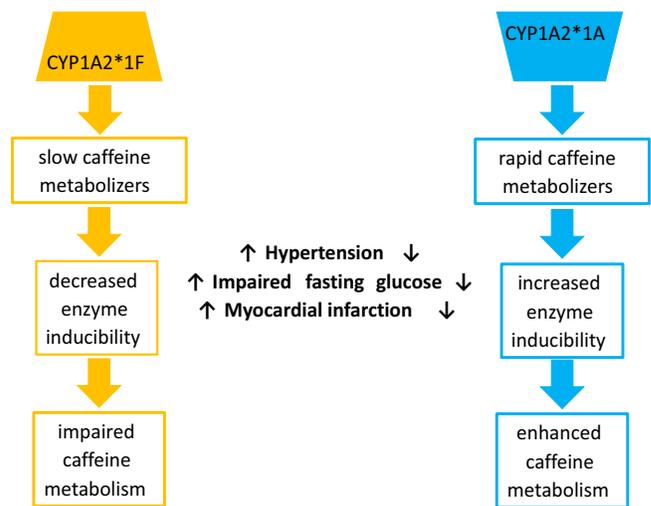


FIGURE 45.3 Caffeine metabolism by cytochrome P450 1A2 (CYP1A2) in the liver. Slow caffeine metabolizers (subjects with the CYP1A2*1F variant) have decreased enzyme inducibility, in turn resulting in impaired caffeine metabolism compared with rapid caffeine metabolizers (subjects with the CYP1A2*1A variant). The CYP1A2*1F variant seems to be associated with an increased risk of hypertension, impaired fasting glucose, and myocardial infarction compared with CYP1A2*1A variant.

higher incidence of acute coronary events after multivariable adjustment (Happonen et al., 2006).

A recent randomized trial of ours, aimed at evaluating the acute cardiovascular effects of coffee on blood pressure, and exploring whether these are influenced by the genetic asset related to caffeine metabolism, of the adenosinergic system, or of the adrenergic α - and β -receptors (Renda et al., 2012). Blood pressure was measured by automated ambulatory monitoring for 2 h after coffee drinking in over 100 healthy young men. An increase in blood pressure after the consumption of caffeinated coffee compared with a decaffeinated coffee preparation was found, and was consistent with an increase in plasma caffeine and adrenaline concentrations (Renda et al., 2012). However, a wide variability in blood pressure response to caffeine was observed, and was explained by a relationship of blood pressure changes with polymorphisms of the adenosine and catecholamine receptors. Specifically, homozygotes for the T variant of the A2A adenosine receptor (ADORA2A), as well as homozygotes for the insertion (I) allelic variant of the α 2B-adrenergic receptor insertion/deletion (ADRA2B I/D) polymorphism, showed a susceptibility to higher blood pressure responses to caffeine (Renda et al., 2012) (Fig. 45.4). It is thus plausible that the exposure to higher blood pressure values as the result of such nutrigenetic interactions in some genetically predisposed individuals may expose such individuals to a higher coffee-related cardiovascular risk.

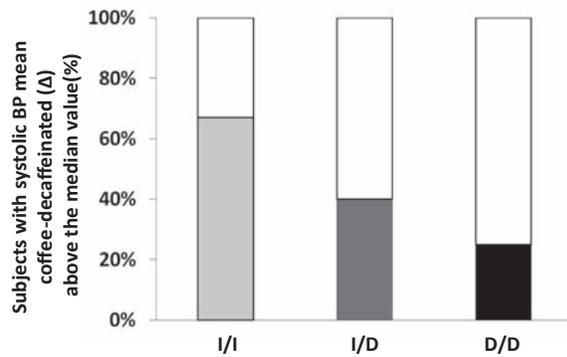


FIGURE 45.4 Relationship of blood pressure changes with polymorphisms of catecholamine receptors. Homozygotes for the insertion (I) allelic variant of the α 2B-adrenergic receptor insertion/deletion (ADRA2B I/D) polymorphism show a susceptibility to higher blood pressure responses to caffeine. The graph shows the percentage of subjects with systolic blood pressure (BP) mean above the median value after consumption of decaffeinated coffee with or without the addition of caffeine (Δ coffee-decaffeinated coffee). Modified from Renda, G., Zimarino, M., Antonucci, I., et al., 2012. Genetic determinants of blood pressure responses to caffeine drinking. *Am J Clin Nutr* 95, 241–248.

Furthermore, different genetic variants have been described for A1 and A2A receptors in the brain.

In particular, the 1976 C \rightarrow T variant of the adenosine A2A receptors seems to mediate the cognitive effects of caffeine. A recent study has demonstrated that CC homozygotes for the adenosine receptor A2A did not improve the orientation component of attention after caffeine intake compared with the other genotypes, and the same variant was previously observed to confer a tendency to disturbed sleep after caffeine intake. On the other hand, TT homozygotes for the adenosine receptor A2A even worsened their motor executive

control, in agreement with previous evidence indicating that TT homozygotes had a greater increase in anxiety after caffeine administration. This might explain the decreased performance in motor executive control observed in TT homozygotes (Renda et al., 2015).

CONCLUSIONS

The currently available overall evidence on cardiovascular effects related to habitual coffee consumption is largely reassuring. Moreover, coffee seems to have a protective effect with cancers and neurodegenerative disease. Many benefits of coffee probably derive from its caffeine content, although other substances may have an important role in health protection. However, genetic interindividual differences in caffeine metabolism or caffeine effectors, such as the adenosine or adrenergic receptors, are likely responsible for different reactions to caffeine intake. Nutrigenetic interactions may expose some genetically predisposed individuals to a higher coffee-related risk of cardiovascular disease. This hypothesis should be tested in large population-based studies.

For all these reasons, the old admonition by physicians to refrain from coffee drinking is not warranted both for healthy subjects and for patients with cardiovascular disease. However, caution has to be exerted for vulnerable people, particularly taking into account the interindividual variability in response to caffeine intake. A summary statement of recommendations based on these findings has been issued by a group of experts within the International Society of Nutrigenetics/Nutrigenomics (De Caterina and El-Sohemy, 2016) (Table 45.1).

TABLE 45.1 Possible Nutrigenetic Recommendations on the Consumption of Caffeine.

Evidence	Recommendations	Classes of Recommendation and LOE
Average increase in BP response after acute caffeine intake (double espresso cup) in homozygotes for ADORA2A TT and ADRA2B II genotypes; intermediate-graded risk for heterozygotes	It would appear prudent that individuals who have the ADORA2A TT and ADRA2B II genotypes, especially if hypertensive, refrain from amounts of acute coffee drinking (double espresso cup)	Grade IIb, LOE B
Increased risk of myocardial infarction in homozygotes or heterozygotes for the CYP1A2*1F variant (slow metabolizers), already at a level of habitual intake of 2 cups per day	It would appear prudent that individuals who have CYP1A2*1F genotypes (slow metabolizers) refrain from drinking more than one cup of coffee per day	Grade IIa, LOE B
Higher incidence of acute coronary events after heavy coffee consumption (>814 mL) in homozygotes for the low-activity COMT AA allele	Heavy coffee drinkers (>814 mL) with the low-activity COMT AA genotype should be advised to limit their coffee drinking	Grade IIb, LOE B

ADORA2A, A2A adenosine receptor; ADRA2B, α 2B-adrenergic receptor; BP, blood pressure; COMT, catechol-O-methyltransferase; CYP1A2, cytochrome P450 1A2; LOE, level of evidence.

Modified from De Caterina, R., El-Sohemy, A., 2016. Moving towards specific nutrigenetic recommendation algorithms: caffeine, genetic variation and cardiovascular risk. *J Nutrigenetics Nutrigenomics* 9, 106–115.

References

- De Caterina, R., El-Soheby, A., 2016. Moving towards specific nutrigenetic recommendation algorithms: caffeine, genetic variation and cardiovascular risk. *J Nutrigenetics Nutrigenomics* 9, 106–115.
- Cornelis, M.C., El-Soheby, A., Kabagambe, E.K., Campos, H., 2006. Coffee, CYP1A2 genotype, and risk of myocardial infarction. *J Am Med Assoc* 295, 1135–1141.
- Goldstein, E.R., Ziegenfuss, T., Kalman, D., et al., 2010. International society of sports nutrition position stand: caffeine and performance. *J Int Soc Sports Nutr* 7, 5.
- Happonen, P., Voutilainen, S., Tuomainen, T.P., Salonen, J.T., 2006. Catechol-o-methyltransferase gene polymorphism modifies the effect of coffee intake on incidence of acute coronary events. *PLoS One* 1, e117.
- Kim, B., Nam, Y., Kim, J., Choi, H., Won, C., 2012. Coffee consumption and stroke risk: ameta-analysis of epidemiologic studies. *Korean J Fam Med* 33, 356–365.
- Larsson, S.C., Orsini, N., 2011. Coffee consumption and risk of stroke: a dose-response meta-analysis of prospective studies. *Am J Epidemiol* 174, 993–1001.
- Mostofsky, E., Rice, M.S., Levitan, E.B., Mittleman, M.A., 2012. Habitual coffee consumption and risk of heart failure: a dose-response meta-analysis. *Circ Heart Fail* 5, 401–405.
- Palatini, P., Ceolotto, G., Ragazzo, F., et al., 2009. CYP1A2 genotype modifies the association between coffee intake and the risk of hypertension. *J Hypertens* 27, 1594–1601.
- Palatini, P., Benetti, E., Mos, L., et al., 2015. Association of coffee consumption and CYP1A2 polymorphism with risk of impaired fasting glucose in hypertensive patients. *Eur J Epidemiol* 30, 209–217.
- Pelchovitz, D.J., Goldberger, J.J., 2011. Caffeine and cardiac arrhythmias: a review of the evidence. *Am J Med* 124, 284–289.
- Rebello, S.A., van Dam, R.M., 2013. Coffee consumption and cardiovascular health: getting to the heart of the matter. *Curr Cardiol Rep* 15, 403.
- Renda, G., De Caterina, R., 2015. Coffee and hypertension: a focus on contrasting acute and chronic effects and nutrigenetics. In: *Coffee in Health and Disease Prevention*. Elsevier, pp. 395–402.
- Renda, G., Zimarino, M., Antonucci, I., et al., 2012. Genetic determinants of blood pressure responses to caffeine drinking. *Am J Clin Nutr* 95, 241–248.
- Renda, G., Comitteri, G., Zimarino, M., et al., 2015. Genetic determinants of cognitive responses to caffeine drinking identified from a double-blind, randomized, controlled trial. *Eur Neuropsychopharmacol* 25, 798–807.
- Sachse, C., Brockmoller, J., Bauer, S., Roots, I., 1999. Functional significance of a C→A polymorphism in intron 1 of the cytochrome P450 CYP1A2 gene tested with caffeine. *Br J Clin Pharmacol* 47, 445–449.
- Schocken, D.D., Benjamin, E.J., Fonarow, G.C., et al., 2008. Prevention of heart failure: a scientific statement from the American Heart Association Councils on Epidemiology and Prevention, Clinical Cardiology, Cardiovascular Nursing, and High Blood Pressure Research; Quality of Care and Outcomes Research Interdisciplinary Working Group; and Functional Genomics and Translational Biology Interdisciplinary Working Group. *Circulation* 117, 2544–2565.
- Sofi, F., Conti, A.A., Gori, A.M., et al., 2007. Coffee consumption and risk of coronary heart disease: a meta-analysis. *Nutr Metab Cardiovasc Dis* 17, 209–223.
- Steffen, M., Kuhle, C., Hensrud, D., Erwin, P.J., Murad, M.H., 2012. The effect of coffee consumption on blood pressure and the development of hypertension: a systematic review and meta-analysis. *J Hypertens* 30, 2245–2254.
- Svenningsson, P., Nomikos, G.G., Ongini, E., Fredholm, B.B., 1997. Antagonism of adenosine A2A receptors underlies the behavioural activating effect of caffeine and is associated with reduced expression of messenger RNA for NGFI-A and NGFI-B in caudate-putamen and nucleus accumbens. *Neuroscience* 79, 753–764.
- Tofalo, R., Renda, G., De Caterina, R., Suzzi, G., 2016. Coffee: health effects. In: Caballero, B., Finglas, P., Toldrá, F. (Eds.), *The Encyclopedia of Food and Health*. Academic Press, Oxford, pp. 237–243.

Phytanic Acid Metabolism in Health and Disease

Ronald J.A. Wanders

Laboratory Genetic Metabolic Diseases, Departments of Paediatrics and Clinical Chemistry, Emma Children's Hospital, Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands

Glossary

- Phytanic acid is one of the most common branched-chain fatty acids in the human diet.
- Phytanic acid cannot be synthesized *in vivo* in humans and is derived only from exogenous sources.
- Phytanic acid requires a special oxidation system localized in peroxisomes, called fatty acid α -oxidation to become degraded.
- Phytanic acid has been associated with prostate cancer and other cancers but the association is weak.
- Phytanic acid is a ligand for different nuclear receptors including PPAR- α and receptor retinoid-X receptor, which may be the basis for its proclaimed health-promoting effects.

Peroxisomes; Peroxisomal Disorders; Fatty Acids; Human Metabolism; Refsum Disease; Fatty Acid Oxidation; Nuclear Hotmones

INTRODUCTION

In the early 1950s, the identification of a multibranching, C₂₀ saturated fatty acid was reported in the butterfat of bovine milk. Full characterization of this compound as phytanic acid (PA) (3,7,11,15-tetramethylhexadecanoic acid) required sophisticated analytical techniques including gas-liquid chromatography and nuclear magnetic resonance spectroscopy. Chemical preparation of PA from phytol (3,7,11,15-tetramethylhexadec-*trans*-2-en-1-ol) had already been reported in 1911. Given the polyisoprenoid-like structure, it was originally thought that PA would be synthesized from four molecules of mevalonic acid analogous to the synthesis of isoprenoids such as farnesoic acid. However, studies in both humans and experimental animals revealed that neither labeled

acetate nor mevalonate was incorporated into PA. After the demonstration that PA cannot be synthesized endogenously, at least in humans and rats, it became clear that free phytol and especially PA itself, but not chlorophyll-bound phytol, were the true sources of PA as derived from dietary sources.

PA in human plasma naturally occurs in the form of two different diastereomers because the carbon-3 atom in phytol can have an *R*- and *S*-configuration (3*R*/3*S*,7*R*,11*R*,15-tetramethylhexadec-2-en-1-ol). The two PA diastereomers produced from phytol are 3*R*,7*R*,11*R*,15-tetramethylhexadecanoic acid (RRR) and 3*S*,7*R*,11*R*,15-tetramethylhexadecanoic acid (SRR).

Interest in PA increased considerably in the early 1960s, when the massive accumulation of PA in tissues from a patient affected by Refsum disease was reported. Since those early years, much has happened and many intricate details of the metabolism of PA have been resolved, although much remains to be learned about the pathophysiology associated with the accumulation of PA. This chapter reviews the current state of knowledge about PA in health and disease.

ORIGIN OF PHYTANIC ACID IN HUMANS

PA is one of the most common branched chain fatty acids in the human diet. Common knowledge holds that PA is derived from phytol, which is a constituent of the chlorophyll molecule. Because almost all photosynthetic organisms use chlorophyll, phytol is abundantly present in nature. As a consequence of the consumption of large quantities of grasses, ruminant animals take in lots of chlorophyll, which undergoes effective degradation by resident bacteria in the rumen, releasing phytol from the chlorophyll molecule followed

by oxidation to PA. The pathway involved in the conversion of phytol into PA has been resolved (Wanders et al., 2011). Although PA and phytol occur as individual molecules and are readily absorbed by humans and subsequently metabolized after oral administration, they contribute only little to the total burden of PA in humans. PA predominantly occurs in esterified form in different lipid species including cholesterol esters, di- and triglycerides, and phospholipids present in fat-containing foods of animal origin, although PA is also present in marine foods. PA is derived from phytoplankton species such as algae and microalgae, which also contain chlorophyll. Subsequently, the phytoplankton chlorophyll molecule is degraded in the gut of zooplankton, allowing phytol to enter the marine system, which explains the occurrence of substantial amounts of PA in animals higher up the food chain, such as mollusks, fish, and whales, in particular, the oils, fat, and milk derived from these species. Although data on the PA content of meat are scarce, available information holds that this may vary considerably. For instance, it has been reported that the PA content is low in lean (organic) beef, amounting to 4 mg/100 g, but it may be greater than 300 mg/100 g in (organic) beef fat (Brown et al., 1993). PA levels in meat also depend very much on the feeding regiment. Feeding with grass silage or red clover silage led to a content of 30–70 mg PA in 100 g milk fat, whereas values as high as 450 mg/100g milk fat were obtained when feeding was entirely based on fresh grass.

PHYTANIC ACID IN HUMAN PLASMA

Because it is derived only from exogenous sources, it is clear that the concentration of PA in human plasma strictly depends on the balance between dietary intake and breakdown. It has been shown that there is a huge variation in the concentration of PA in human serum. In serum from 250 healthy human subjects, PA levels ranged from 0.04 to 11.5 $\mu\text{mol/L}$, with a medium of 1.6 $\mu\text{mol/L}$ (Al-Dirbashi et al., 2008). Earlier, the important role of the diet in determining plasma PA levels was shown. When plasma PA levels were determined in a group of British meat eaters compared with lacto-ovo-vegetarians and vegans, geometric mean plasma PA concentrations were 5.77, 3.93, and 0.86 $\mu\text{mol/L}$, respectively (Allen et al., 2008). As discussed earlier, PA may not come only from dairy products; it is also high in certain marine fats, which indicates that traditional marine based diets in arctic regions may contain much higher levels of PA than in Western, meat-based diets (Hellgren, 2010).

Apart from the fact that the PA levels in plasma depend highly on the diet of each individual person,

the ratio between the two stereoisomers SRR and RRR may also vary among individuals. Che and coworkers reported that the percentage of the different RR isomers in total PA in milk fat derived from milk samples from five organic herds during the grazing season differed considerably and depended on the share of grazed clovers in dry matter intake (Che et al., 2013a).

METABOLISM OF PHYTANIC ACID

The realization that PA like any 3-methyl fatty acid cannot undergo β -oxidation, at least not directly, prompted Steinberg and coworkers to study its mechanism of oxidation both in vivo in humans and rats and in vitro, notably in fibroblasts, using radiolabeled PA. These studies clearly showed that PA first undergoes oxidative decarboxylation to produce pristanic acid (PR) and CO_2 . It became equally clear that PR, which is a 2-methyl fatty acid that can be β -oxidized, undergoes β -oxidation, because the expected β -oxidation products were readily identified after in vivo loading of rats with [$\text{U}-^{14}\text{C}_{20}$]-PA.

The exact enzymatic machinery involved in the α -oxidation of PA to PR has long remained a mystery, but it has now been resolved to a large extent. Because this has been the subject of several reviews, this will not be discussed in great detail here (Van Veldhoven, 2010; Wanders et al., 2011). The α -oxidation pathway is composed of five different enzymatic steps (Fig. 46.1), including (1) activation of PA to phytanoyl-coenzyme A (CoA); (2) hydroxylation of phytanoyl-CoA to 2-hydroxyphytanoyl-CoA; (3) cleavage of 2-hydroxyphytanoyl-CoA to pristanal and CO_2 ; (4) oxidation of pristanal to PR, and (5) activation of PR to pristanoyl-CoA. Some of the enzymes involved in the pathway still have not been identified with certainty, including aldehyde dehydrogenase, which catalyzes the conversion of pristanal into PR.

INBORN ERRORS OF PHYTANIC ACID METABOLISM

To date, only a single defect in the PA α -oxidation pathway has been identified: phytanoyl-CoA 2-hydroxylase deficiency. Refsum disease (RD) is associated with this enzyme deficiency; it was originally described as Heredopathia atactica polyneuritiformis in 1946 by Sigvald Refsum, a Norwegian neurologist. Patients with RD may have a large variety of clinical signs and symptoms, some of which are observed in virtually all patients, including retinitis pigmentosa and anosmia; other features are not universally observed in all patients. Wierzbicki and collaborators

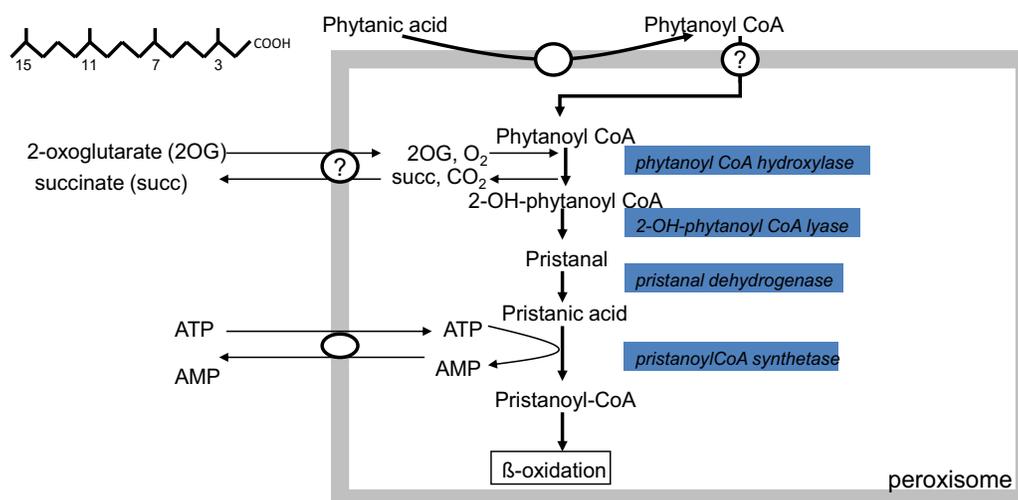


FIGURE 46.1 Schematic representation of mechanism of phytanic acid α -oxidation in peroxisomes. AMP, adenosine monophosphate; ATP, adenosine triphosphate; CoA, coenzyme A.

described the cumulative incidence of a range of features over many decades in a group of 15 established RD patients which revealed retinitis pigmentosa in all 15 patients, followed by anosmia in 14, neuropathy in 11, deafness in 10, ataxia in 8, and ichthyosis in 4 (Wierzbicki et al., 2002). Polyneuropathy is a mixed motor and sensory neuropathy that is asymmetric, chronic, and progressive in untreated individuals. It may not be clinically apparent at the start of the illness. Initially, symptoms often wax and wane. Later, the distal lower limbs are affected with resulting muscular atrophy and weakness. Over the course of years, muscular weakness can become widespread and disabling, involving not only the limbs but also the trunk. Almost without exception, individuals with RD have peripheral sensory disturbances, most often impairment of deep sensation, particularly perception of vibration and position-motion in the distal legs.

In addition to these features, RD patients may have a number of abnormalities including skeletal abnormalities with short metacarpals and metatarsals present in about 30% of affected individuals, and cardiac abnormalities, including cardiac arrhythmia and heart failure resulting from cardiomyopathy, which are frequent causes of death in RD patients.

PA is also elevated in a number of other genetic diseases, including the Zellweger spectrum disorders (ZSDs), and Rhizomelic chondrodysplasia punctata type 1 (RCDP type 1). In these diseases, phytanoyl-CoA hydroxylase is also deficient, similar to RD, not because of mutations in the structural gene coding for phytanoyl-CoA hydroxylase, but because of mutations in other genes, the products of which are involved in peroxisome biogenesis. Because phytanoyl-CoA hydroxylase is a peroxisomal enzyme, the inability to synthesize peroxisomes as in the ZSDs causes a deficiency

of phytanoyl-CoA hydroxylase, because this enzyme is unstable when it is present in the cytosol of the cell. In RCDP type 1, peroxisome biogenesis is not completely impaired but only partially, because only one of the two import routes for peroxisomal proteins, i.e., the PTS2 import route, is deficient. Phytanoyl-CoA hydroxylase requires a fully functional PTS2 import pathway to achieve proper localization within the peroxisome, which explains why the hydroxylase is deficient if the PTS2 pathway is not functioning properly (Van Veldhoven, 2010; Wanders et al., 2011).

PHYTANIC ACID AND GENE EXPRESSION

PA has been claimed to have health-promoting effects, although concerns have arisen because of its association with different cancers. The health-promoting effects of PA are based on the initial observation (Kitarawan et al., 1996) that PA was capable of activating the nuclear-receptor retinoid-X receptor (RXR) at physiological concentrations and the suggestion that PA might be important in coordinating cell metabolism via RXR. A similar observation was made independently by another group (Lemotte et al., 1996). These observations received extra attention when it was reported that PA also activates another nuclear receptor involved in metabolic regulation: peroxisome proliferator-activated receptor- α (PPAR- α) (Ellinghaus et al., 1999). RXR and PPAR- α form a heterodimer upon ligand binding, followed by binding to specific promoter elements in selected target genes, called PPRE, and thereby achieve selective activation of transcription of PPAR- α target genes. In many of the original studies, relatively high physiological concentrations of PA were used and the conclusion was that

PA is a relatively low-affinity ligand for both RXR and PPAR- α (Ellinghaus et al., 1999; Zomer et al., 2000; Schluter et al., 2002a). However, work by Schroeder and coworkers shed a different light on this issue (Hostetler et al., 2006). Indeed, work by those authors on the affinity of different fatty acids and their CoA-esters toward recombinant PPAR- α revealed that PA is a high-affinity ligand with a dissociation constant as low as 34 nmol/L for free PA and 11 nmol/L for phytanoyl-CoA. How these effects would work in an in vivo situation that requires the participation of many additional factors, including coactivators and corepressors, for final activation of gene expression remains to be established (see subsequent discussion). Some in vivo studies have been conducted in this area, although these are usually done with phytol as the precursor of PA rather than with PA itself, because PA is expensive. A major concern with such in vivo studies is that PA as derived from phytol produces PR, which has been claimed to be an even better ligand for PPAR- α than PA. The availability of a mouse deficient in phytanoyl-CoA 2-hydroxylase, which prohibits conversion of PA to PR, together with the availability of a PPAR- α knockout mouse model, would make such studies feasible. Work done in our laboratory suggests that phytol and PA are both PPAR- α dependent and independent (Gloerich et al., 2005).

PHYSIOLOGICAL EFFECTS OF PHYTANIC ACID

PA has been shown to have multiple effects on a variety of different physiological processes including cell differentiation and glucose and lipid metabolism. After the finding that PA is an activator of RXR, it was discovered that PA is a novel activator of both uncoupling protein 1 (UCP-1) gene transcription and brown-adipocyte differentiation (Schluter et al., 2002b). That PA induces UCP-1 expression in differentiated brown-adipocytes at physiological concentrations is important, especially because UCP-1 has a crucial role in nonshivering thermogenesis and thus has a major role in energy balance. In mouse models in which PA oxidation was blocked at the level of sterol carrier protein X (Seedorf et al., 1998) and phytanoyl-CoA 2-hydroxylase (Ferdinandusse et al., 2008), it was found that there was an imbalance between body weight and feeding. *Scp2* ($-/-$) mice had body weights similar to those of control mice whereas food intake was significantly (30%) higher in *Scp2* ($-/-$) mice. Furthermore, in phytanoyl-CoA 2-hydroxylase-deficient mice (*Phyh* [$-/-$]), the 0.25% phytol diet that led to plasma PA levels of 1200 $\mu\text{mol/L}$ caused a decrease in body weight of $25\% \pm 5\%$ in the *Phyh* ($-/-$) males and $16\% \pm 10\%$ in the *Phyh* ($-/-$)

females (Ferdinandusse et al., 2008). Upon dissection, almost no white fat could be observed in the *Phyh* ($-/-$) mice with the 0.25 phytol diet. Although energy balance has not been studied in RD patients, they usually have a low body weight with reduced white adipose tissue mass.

Because PPAR- α is a key regulator of fatty acid metabolism, at least in rodents, PA may also have an effect on whole-body lipid homeostasis. In a study by Hellgren, hepatic triacylglyceride levels dropped considerably with a 0.2% phytol diet (Hellgren, 2010). With respect to the possible effects of PA on glucose metabolism, earlier work (Heim et al., 2002) showed that PA is able to enhance glucose uptake in hepatocytes without strongly promoting adipogenic differentiation by upregulating both GLUT-1 and GLUT-2, and glucokinase ultimately resulted in increased glucose uptake in primary hepatocytes. The same authors reported that PA is a transactivator not only of PPAR- α but also of PPAR- β PPAR- γ . More recent work (Che et al., 2013b) in primary porcine myotubes revealed that PA stimulated glucose uptake, especially at low insulin concentrations. Taken together, although it is clear that PA has major effects on multiple physiological processes, much remains to be learned how these effects, which were often obtained in isolated cells, cell homogenates, and/or mitochondria, work in vivo. The availability of a mouse model deficient in phytanoyl-CoA 2-hydroxylase (Ferdinandusse et al., 2008) would make such studies feasible.

PHYTANIC ACID AND CANCER

In recent years, concerns have been raised about PA owing to its association with different cancers, including prostate cancer. Xu and coworkers reported on the association between plasma PA levels and prostate cancer risk, although the sample size was small (Xu et al., 2005). Earlier work (Luo et al., 2002; Rubin et al., 2002) revealed that the enzyme α -methylacyl-CoA racemase, which has a crucial role in PA metabolism, is markedly overexpressed in high-grade prostatic intraepithelial neoplasias and in most prostate cancers, thus providing a link with PA in plasma. In 2010, results were published of the first prospective investigation of the relation between plasma PA levels and subsequent risk for cancer in a series of 566 cases (Price et al., 2010). The conclusion was that plasma PA levels are significantly associated with intake of dietary fat but not with the overall risk for prostate cancer, at least in the population studied with prostate cancer patients from Germany, Greece, Italy, The Netherlands, Spain, and the United Kingdom. Taken together, although the association between PA levels and prostate cancer remains unclear, current evidence holds that the association is weak.

OTHER PATHOPHYSIOLOGICAL MECHANISMS OF PHYTANIC ACID, NOTABLY ON MITOCHONDRIA

Apart from the pathophysiological effects of PA described earlier, many additional reports on PA have appeared, especially regarding mitochondria. First, it was reported that PA affects the dynamics of phospholipids in membranes as well as the physical state of membrane proteins, which prompted the authors to propose a novel mechanism for the protonophoric effect of PA (Schonfeld and Struy, 1999). In agreement with its protonophoric action, PA was subsequently found to deenergize rat brain mitochondria respiring under state 4 conditions, which occurred at low concentrations (Schonfeld et al., 2004). Interestingly, similar low concentrations of PA turned out to inhibit rather than stimulate mitochondrial respiration under active conditions using adenosine diphosphate (ADP) to induce respiration maximally (state 3). That the inhibition of the respiratory chain by PA was less under uncoupler-stimulated versus ADP-stimulated conditions was explained by the observation that PA also turned out to inhibit the adenine nucleotide carrier, which exchanges extramitochondrial ADP for intramitochondrial adenosine triphosphate (ATP). Finally, at higher concentrations, PA also appeared to inhibit complex I of the respiratory chain, which explains the inhibition of uncoupler-stimulated respiration by PA at concentrations greater than 10 $\mu\text{mol/L}$. Subsequent studies in hippocampal neurons, astrocytes, and oligodendrocytes revealed that PA and PR induce a complex array of toxic activities including mitochondrial dysfunction and calcium-deregulation via involvement of the $\text{InsP}_3\text{-Ca}^{++}$ signaling pathway (Ronicke et al., 2009). Subsequent work revealed that the PA/PR-induced deregulation of cytosolic calcium is mediated through activation of the free fatty acid receptor GPR40 (Kruska and Reiser, 2011). Along similar lines, Wajner and coworkers showed that PA induces oxidative damage and reduces antioxidant defense systems in the cerebellum and cerebral cortex of rats (Leipnitz et al., 2010) and also compromises Na^+/K^+ ATPase and the functional capacity of the respiratory chain in the brain cortex of rats (Busanello et al., 2013a, 2013b). Finally, work (Nagai, 2015) pointed to yet another effect of PA: on histone deacetylation. PA was first found to enhance histone deacetylase activity *in vitro*. In neuro2a cells, PA significantly reduced histone acetylation and induced cell death. These effects were inhibited by Hdac inhibitors. The authors suggested that the PA-induced mitochondrial abnormalities and cell death in this neuronal cell line via activation of Hdac2 and Hdac3 may be the basis for the pathophysiology of RD.

CONCLUSIONS

Much has been learned about the metabolism of PA and the exact mechanisms through which PA is degraded. Furthermore, there is much information on the sources of PA in different dietary products. In addition, the content of PA in foodstuffs may vary considerably, depending on the feeding regiment of ruminants, among other factors. However, much less is known about the physiological and pathophysiological consequences of PA in normal human beings; PA levels range from 0.1 to 10 $\mu\text{mol/L}$, and in patients with a defect in PA α -oxidation, levels may be as high as 1000 $\mu\text{mol/L}$. It is especially important to study to what extent the many abnormalities induced by PA in isolated systems, including isolated mitochondria, cells, and/or cell homogenates, can be extrapolated to the *in vivo* situation. One aspect often overlooked is that the level of free unesterified PA as used in many studies is actually low in human plasma, because virtually all PA is esterified and present in the form of cholesterol esters, phospholipids, triglycerides, and so on. An additional problem is that in the few *in vivo* studies conducted in mice and rats, phytol was administered as a substitute for PA and PA underwent rapid α -oxidation to PR, which makes it difficult to resolve the true effects of PA. Some of these problems can be circumvented by using the mouse model in which PA α -oxidation is disrupted, such as the phytanoyl-CoA 2-hydroxylase-deficient mouse. Construction of a double knockout in which both the gene coding for phytanoyl-CoA 2-hydroxylase and PPAR- α are knocked out may open the way to study the true effects of PA to resolve to what extent at least PPAR- α is involved. An additional point to consider is that PA should be used instead of phytol, because phytol itself may also have toxic effects *per se*. Studies along these lines are in progress.

Acknowledgments

The author gratefully acknowledges the expert help of Mrs. Maddy Festen for preparing the manuscript and Mr. Jos Ruiten for drawing Fig. 46.1.

References

- Al-Dirbashi, O.Y., Santa, T., Rashed, M.S., Al-Hassnan, Z., Shimozaawa, N., Chedrawi, A., Jacob, M., Al-Mokhadab, M., 2008. Rapid UPLC-MS/MS method for routine analysis of plasma pristanic, phytanic, and very long chain fatty acid markers of peroxisomal disorders. *J Lipid Res* 49, 1855–1862.
- Allen, N.E., Grace, P.B., Ginn, A., Travis, R.C., Roddam, A.W., Appleby, P.N., Key, T., 2008. Phytanic acid: measurement of plasma concentrations by gas-liquid chromatography-mass spectrometry

- analysis and associations with diet and other plasma fatty acids. *Br J Nutr* 99, 653–659.
- Brown, P., Guam Mei, F.B., Gibberd, D., Burston, D., Mayne, P.D., McClinchy, J.E., Sidey, M., 1993. Diet and Refsum's disease. The determination of phytanic acid and phytol in certain foods and the application of this knowledge to the choice of suitable convenience foods for patients with Refsum's disease. *J Hum Nutr Diet* 6, 295–305.
- Busanello, E.N., Amaral, A.U., Tonin, A.M., Zanatta, A., Viegas, C.M., Vargas, C.R., Wajner, M., 2013. Disruption of mitochondrial homeostasis by phytanic acid in cerebellum of young rats. *Cerebellum* 12, 362–369.
- Busanello, E.N., Zanatta, A., Tonin, A.M., Viegas, C.M., Vargas, C.R., Leipnitz, G., Ribeiro, C.A., Wajner, M., 2013. Marked inhibition of Na⁺, K(+) -ATPase activity and the respiratory chain by phytanic acid in cerebellum from young rats: possible underlying mechanisms of cerebellar ataxia in Refsum disease. *J Bioenerg Biomembr* 45, 137–144.
- Che, B.N., Kristensen, T., Nebel, C., Dalsgaard, T.K., Hellgren, L.I., Young, J.F., Larsen, M.K., 2013. Content and distribution of phytanic acid diastereomers in organic milk as affected by feed composition. *J Agric Food Chem* 61, 225–230.
- Che, B.N., Oksbjerg, N., Hellgren, L.I., Nielsen, J.H., Young, J.F., 2013. Phytanic acid stimulates glucose uptake in a model of skeletal muscles, the primary porcine myotubes. *Lipids Health Dis* 12, 14.
- Ellinghaus, P., Wolfrum, C., Assmann, G., Spener, F., Seedorf, U., 1999. Phytanic acid activates the peroxisome proliferator-activated receptor alpha (PPARalpha) in sterol carrier protein 2-/sterol carrier protein x-deficient mice. *J Biol Chem* 274, 2766–2772.
- Ferdinandusse, S., Zomer, A.W., Komen, J.C., Van den Brink, C.E., Thanos, M., Hamers, F.P., Wanders, R.J.A., Van Der Saag, P.T., Poll-Thé, B.T., Brites, P.M., 2008. Ataxia with loss of Purkinje cells in a mouse model for Refsum disease. *Proc Natl Acad Sci USA* 105, 17712–17717.
- Gloerich, J., van Vlies, N., Jansen, G.A., Denis, S., Ruiters, J.P.N., van Werkhoven, M.A., Duran, M., Vaz, F.M., Wanders, R.J.A., Ferdinandusse, S., 2005. A phytol-enriched diet induces changes in fatty acid metabolism in mice both via PPARalpha-dependent and -independent pathways. *J Lipid Res* 46, 716–726.
- Heim, M., Johnson, J., Boess, F., Bendik, I., Weber, P., Hunziker, W., Fluhmann, B., 2002. Phytanic acid, a natural peroxisome proliferator-activated receptor (PPAR) agonist, regulates glucose metabolism in rat primary hepatocytes. *FASEB J* 16, 718–720.
- Hellgren, L.I., 2010. Phytanic acid – an overlooked bioactive fatty acid in dairy fat? *Ann N Y Acad Sci* 1190, 42–49.
- Hostetler, H.A., Kier, A.B., Schroeder, F., 2006. Very-long-chain and branched-chain fatty acyl-CoAs are high affinity ligands for the peroxisome proliferator-activated receptor alpha (PPARalpha). *Biochemistry* 45, 7669–7681.
- Kitareewan, S., Burka, L.T., Tomer, K.B., Parker, C.E., Deterding, L.J., Stevens, R.D., Forman, B.M., Mais, D.E., Heyman, R.A., McMorris, T., Weinberger, C., 1996. Phytol metabolites are circulating dietary factors that activate the nuclear receptor RXR. *Mol Biol Cell* 7, 1153–1166.
- Kruska, N., Reiser, G., 2011. Phytanic acid and pristanic acid, branched-chain fatty acids associated with Refsum disease and other inherited peroxisomal disorders, mediate intracellular Ca²⁺ signaling through activation of free fatty acid receptor GPR40. *Neurobiol Dis* 43, 465–472.
- Leipnitz, G., Amaral, A.U., Zanatta, A., Seminotti, B., Fernandes, C.G., Knebel, L.A., Vargas, C.R., Wajner, M., 2010. Neurochemical evidence that phytanic acid induces oxidative damage and reduces the antioxidant defenses in cerebellum and cerebral cortex of rats. *Life Sci* 87, 275–280.
- Lemotte, P.K., Keidel, S., Apfel, C.M., 1996. Phytanic acid is a retinoid X receptor ligand. *Eur J Biochem* 236, 328–333.
- Luo, J., Zha, S., Gage, W.R., Dunn, T.M., Hicks, J.L., Bennet, C.J., Ewing, C.M., Platz, E.A., Ferdinandusse, S., Wanders, R.J.A., Trent, J.M., Isaacs, W.B., De Marzo, A.M., 2002. Alpha-methylacyl-CoA racemase: a new molecular marker for prostate cancer. *Cancer Res* 62, 2220–2226.
- Nagai, K., 2015. Phytanic acid induces Neuro2a cell death via histone deacetylase activation and mitochondrial dysfunction. *Neurotoxicol Teratol* 48, 33–39.
- Price, A.J., Allen, N.E., Appleby, P.N., Crowe, F.L., Jenab, M., Rinaldi, S., Slimani, N., Kaaks, R., Rohrmann, S., Boeing, H., Pischon, T., Benetou, V., Naska, A., Trichopoulou, A., Palli, D., Sieri, S., Tumino, R., Vineis, P., Bueno-de-Mesquita, H.B., Donate, I., Gonzalez, C.A., Sanchez, M.J., Chirlaque, M.D., Ardanaz, E., Larranaga, N., Khaw, K.T., Rodwell, S., Gallo, V., Michaud, D.S., Riboli, E., Key, T.J., 2010. Plasma phytanic acid concentration and risk of prostate cancer: results from the European Prospective Investigation into Cancer and Nutrition. *Am J Clin Nutr* 91, 1769–1776.
- Ronicko, S., Kruska, N., Kahlert, S., Reiser, G., 2009. The influence of the branched-chain fatty acids pristanic acid and Refsum disease-associated phytanic acid on mitochondrial functions and calcium regulation of hippocampal neurons, astrocytes, and oligodendrocytes. *Neurobiol Dis* 36, 401–410.
- Rubin, M.A., Zhou, M., Dhanasekaran, S.M., Varambally, S., Barrette, T.R., Sanda, M.G., Pienta, K.J., Ghosh, D., Chinnaiyan, A.M., 2002. alpha-Methylacyl coenzyme A racemase as a tissue biomarker for prostate cancer. *J Am Med Assoc* 287, 1662–1670.
- Schluter, A., Barbera, M.J., Iglesias, R., Giral, M., Villarroya, F., 2002. Phytanic acid, a novel activator of uncoupling protein-1 gene transcription and brown adipocyte differentiation. *Biochem J* 362, 61–69.
- Schluter, A., Giral, M., Iglesias, R., Villarroya, F., 2002. Phytanic acid, but not pristanic acid, mediates the positive effects of phytol derivatives on brown adipocyte differentiation. *FEBS Lett* 517, 83–86.
- Schonfeld, P., Kahlert, S., Reiser, G., 2004. In brain mitochondria the branched-chain fatty acid phytanic acid impairs energy transduction and sensitizes for permeability transition. *Biochem J* 383, 121–128.
- Schonfeld, P., Struy, H., 1999. Refsum disease diagnostic marker phytanic acid alters the physical state of membrane proteins of liver mitochondria. *FEBS Lett* 457, 179–183.
- Seedorf, U., Raabe, M., Ellinghaus, P., Kannenberg, F., Fobker, M., Engel, T., Denis, S., Wouters, F., Wirtz, K.W.A., Wanders, R.J.A., Maeda, N., Assmann, G., 1998. Defective peroxisomal catabolism of branched fatty acyl coenzyme A in mice lacking the sterol carrier protein-2/sterol carrier protein-x gene function. *Genes Dev* 12, 1189–1201.
- Van Veldhoven, P.P., 2010. Biochemistry and genetics of inherited disorders of peroxisomal fatty acid metabolism. *J Lipid Res* 51, 2863–2895.
- Wanders, R.J.A., Komen, J.C., Ferdinandusse, S., 2011. Phytanic acid metabolism in health and disease. *Biochim Biophys Acta* 1811, 498–507.
- Wierzbicki, A.S., Lloyd, M.D., Schofield, C.J., Feher, M.D., Gibberd, F.B., 2002. Refsum's disease: a peroxisomal disorder affecting phytanic acid alpha-oxidation. *J Neurochem* 80, 727–735.
- Xu, J., Thornburg, T., Turner, A.R., Vitols, M., Case, D., Shadle, J., Hinson, L., Sun, J., Liu, W., Chang, B., Adams, T.S., Zheng, S.L., Torti, F.M., 2005. Serum levels of phytanic acid are associated with prostate cancer risk. *Prostate* 63, 209–214.
- Zomer, A.W.M., van der Burg, B., Jansen, G.A., Wanders, R.J.A., Poll-The, B.T., Van Der Saag, P.T., 2000. Pristanic acid and phytanic acid. Naturally occurring ligands for the nuclear receptor peroxisome proliferator-activated receptor alpha. *J Lipid Res* 41, 1801–1807.

Nutrigenetics, Fatty Acids, and Cognition

Carol L. Cheatham

Department of Psychology & Neuroscience, Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, NC, United States

Glossary

- Alpha-linolenic acid** An essential omega-3 fatty acid (18:3n-3) found in plants.
- Docosahexaenoic acid** An omega-3 fatty acid (22:6n-3) that is found in animal products like fish and grass-fed beef.
- Electrophysiology** The study and recording of the electrical properties of neurons and other cells in the body.
- Essential fatty acid** A fatty acid that cannot be synthesized endogenously in the body and therefore, must be obtained through the diet.
- Event-related potentials** A measured brain response that is the direct result of external stimuli such as visual or auditory input.
- Hippocampus** A brain structure that is located bilaterally in the temporal lobes and supports spatial memory and the consolidation of explicit memories.
- Long-term potentiation** A type of synaptic plasticity first discovered in the hippocampus that strengthens the connection between two neurons that are frequently in communication.
- Oddball paradigm** An experimental design used in cognitive testing in which a series of frequently occurring stimuli are interspersed with randomly occurring novel stimuli.

INTRODUCTION

Even though Hippocrates (c.460–377 B.C.) proposed the brain as the seat of the mind, it was not until the second half of the 20th century that scientists studying behavior began to explore the concept. Prior to Karl Spencer Lashley's 1948 suggestion of the addition of neuroanatomy to the study of cognition, behaviorists led the field with their premise that all behavior could be explained in terms of environmental influences, as in operant conditioning. The mechanism behind conditioning was relegated to the "black box." What ensued was a revolution of sorts that led to the "birth" of cognitive psychology as a field in the late 1960s. In the decades that followed, the debate over nature versus nurture was dominant. Any phenomenon studied seemed to require the answer to the question does "it"

(the behavior of interest) have its roots in nature (biology) or nurture (environment). By the 21st century, it had become clear that no behavior could be explained wholly by one or the other, and the concept of nature *and* nurture working synergistically began to be accepted.

Nutrition (nurture) and genetics (nature) have, therefore, only recently begun to be studied for their synergistic effects on cognitive function. In fact, only a few scientists are studying the effects of nutrition on the brain and its functionality, with even fewer of those scientists including molecular data. Nonetheless, the data that are available are very promising. In this chapter, you will discover why one would want to consider the interplay between genetics and nutrition when studying cognition, how individual differences in cognitive performance may be explained in terms of responders versus nonresponders, and why the materno-fetal dyadic relationship must be considered when attempting to understand the nutrigenetics of cognition.

CONFOUNDERS IN NUTRITION AND COGNITION RESEARCH

The most studied nutrigenetic phenomenon with relation to brain development and function is that of fatty acids and their genetic underpinnings. Fatty acids are essential to cell functioning, including neurons. One such fatty acid, docosahexaenoic acid (22:6n-3; DHA), is the most prevalent n-3 fatty acid in the brain, and it plays a critical role in neural function. DHA improves synaptic transmission and neuronal fluidity, two processes that are necessary for the movement of information from neuron to neuron. The ability to move information through the brain (aka processing) underlies all cognitive abilities.

The main source of DHA in the diet is fatty fish, and DHA intake is low in countries such as the United States

and Australia relative to the more fish-focused cultures such as Japan or the Nordic countries. Thus, supplementation with DHA has become a ubiquitous recommendation among scientists, medical professionals, and agencies despite a paucity of evidence; scientists focusing on the effects of increasing postnatal exogenous DHA (e.g., DHA-supplemented infant formula) on brain development have found an effect on cognition in fewer than 40% of trials. Those conducting maternal supplementation trials during gestation and lactation also report mixed results. Positive effects of supplementation (maternal or infant) have been found on infant problem-solving, preschool-age processing abilities, and in the reduction in risk of neurological disorders, language disorders, and developmental delays. Conversely, no effects have been found on global cognitive function, recognition memory, visual acuity, language, attention, or working memory/inhibitory control. Indeed, reviewers of the literature conclude that there is insufficient evidence to support the use of DHA supplements in pregnancy and lactation to improve neural function.

Importantly, significantly smaller head circumferences, lower language abilities, and lower working memory/inhibitory control abilities have been reported in the offspring of fish-eating women given prenatal DHA supplements, which may indicate an inverted-U-shaped dose response with respect to the offspring when supplementing with DHA those who already have sufficient omega-3 dietary intake.

However, the sufficiency of dietary intake has not been established, and the premise in this chapter is that dietary requirements of fatty acids may be based in genetics. Those participants who responded to DHA supplementation with an improvement in cognitive abilities may have typical requirements for intake. Participants who did not respond to the treatment may have background genetics that result in suboptimal fatty acid metabolism, and as such, they may require a higher intake of fatty acids. The effects of fatty acids in studies conducted without accounting for genetics would surely be masked by these differences—the effects would be confounded by the inclusion of both responders and nonresponders in the sample.

Before moving to a discussion of fatty acid metabolism and the genes that control it, it is important to understand the background on the importance of fatty acids to the brain structure and function. The next section provides a cursory review of contemporary understanding of the role of DHA in the brain. Other fatty acids certainly provide support to the brain; however, scientists have studied DHA more extensively given its theoretical import to child development.

FATTY ACIDS IN BRAIN

DHA accumulates most rapidly in the brain during the third trimester of gestation and the first two postnatal

years. During this period, DHA is deposited in the phospholipid membranes of the brain where it serves a structural role that contributes to the functioning of neurons, glia, and synapses. Animal models developed to explore the effects of DHA-deficient maternal diets have revealed neuronal anomalies in areas of pup brain that subservise cognition. Relative to DHA-replete maternal models, smaller neuron size in the hippocampus (memory) and parietal lobe (attention) has been found in pups with low brain DHA. DHA deficiency during rodent brain development has also been shown to delay neurogenesis in the cerebral cortex and dentate gyrus of the hippocampus resulting in smaller overall size for these regions. Humans are born with a rudimentary hippocampus as neural progenitor cell proliferation and differentiation are maximal from the 24th week of gestation through at least 15 months after birth, with the dentate gyrus developing between the 36th week of gestation to the third postnatal month. Thus, the DHA requirements would be high from mid-gestation through the first 15–18 months of life for optimal hippocampal development, and subsequent optimal cognitive development. Indeed, it has been found that the accretion of DHA into the developing human brain is highest from the third trimester through the second year of life.

Within the hippocampus, DHA is directly implicated in the neural processes underlying learning and memory. Long-term potentiation (LTP) is a form of synaptic plasticity that neurons utilize in the formation of a memory trace. LTP relies on molecular processes at the *N*-methyl-*D*-aspartate (NMDA) channels. Both the process of LTP and activity at the NMDA channels are dependent on the presence of DHA. In addition, a fatty acid-binding protein integral to the extension of glial processes in response to neural signals exhibits an unprecedented high affinity for DHA, which indicates DHA also has a role in glial-neural interactions. DHA's involvement in cell signal transduction, through both primary and secondary messenger systems, has also been documented, and high DHA levels are associated with increased speed of cellular processes at the membrane and across the synapse. Biomarkers of DHA at birth evidence the importance of in utero DHA levels. Neonates with higher umbilical DHA status at birth reportedly have more mature EEG at 2 days of age, which suggests more efficient synaptic activity in the cortex. Fatty acid levels in umbilical cord blood have also been related to overall neurologic status soon after birth as well as at infant and toddler age; those with higher DHA and lower trans-fatty acids in venous umbilical cord blood are more likely to have favorable neurological ratings.

As a whole, the evidence illustrates that DHA is essential to brain health and in the processes that underlie cognition, specifically processing and storage of memories in humans. Most importantly, maternal (prenatal) as well as individual (postnatal) factors

have an effect on the development of the brain and subsequent cognitive abilities.

GENETICS OF FATTY ACID METABOLISM

As mentioned previously, the reports of DHA supplementation studies contain very mixed results with some scientists finding effects and others finding no effects or even negative effects. It is concerning that no definitive conclusions arise from this rich collection of DHA studies, given the seeming importance of DHA to information processing as well as the integrity of the neural substrates that subserve attention and memory. Along with processing, attention and memory underlie all other cognitive abilities, and as such, their development early on determines the trajectory of subsequent cognitive development and school readiness. Many scientists have called for standardization of methodology in DHA trials as a way to elucidate the effects, which would be very beneficial. However, the wide range of results in the supplementation studies are indicative also of unobserved individual differences. Notably, background genetics and diet have not been included in these trials historically.

Evidence has accumulated that genetic variation has an integral role in individual fatty acid requirements in that the enzymes required in the fatty acid biosynthesis pathways are under genetic control (see Fig. 47.1). Fatty acid metabolism is limited by the delta-5 and delta-6 desaturase coded by the *FADS* genes. The polyunsaturated fatty acids (PUFA) rely on delta-5 and delta-6 desaturase to synthesize the long-chain fatty acids (LC-PUFA) DHA and arachidonic acid, 20:4n-6

(AA) from their precursors alpha-linolenic acid, 18:3n-3 (LNA) and linoleic acid, 18:2n-6 (LA), respectively (see Fig. 47.1). LNA and LA are essential nutrients—they must be obtained from the diet. The n-3 and n-6 (also n-9, not discussed here) pathways compete for a limited supply of delta-5 and delta-6 desaturase (center column, Fig. 47.1) encoded by the *FADS* genes and of elongases encoded by the *ELOVL* genes. Beyond the competition for substrate, the supply of the delta-5 and delta-6 desaturase (and possibly the elongase) is further limited by an individual's genotype; the supply of enzymes is directly related to the ability of the genes to code for them. In fact, single nucleotide polymorphisms (SNPs) in the *FADS* gene complex are related to lower levels of fatty acids in plasma and in milk. Thus, the need for exogenous DHA is dependent, in part, on an individual's ability to synthesize endogenous long-chain fatty acids from the essential parent fatty acids. That is, if an individual carries certain *FADS* SNPs, consumption of exogenous fatty acids would be crucial as the supply of endogenous fatty acids would be limited.

INFLUENCE OF GENETIC VARIANTS ON DHA STATUS AND REQUIREMENTS

Scientists considering the *FADS* gene complex have found individual differences in fatty acid synthesis. The relation between SNPs of the *FADS* cluster and levels of serum phospholipid fatty acids has been documented. It follows that the clinical significance of dietary intake of fatty acids would vary widely between individuals. Indeed, pregnant women who are homozygous for minor alleles in a *FADS* haplotype have higher levels

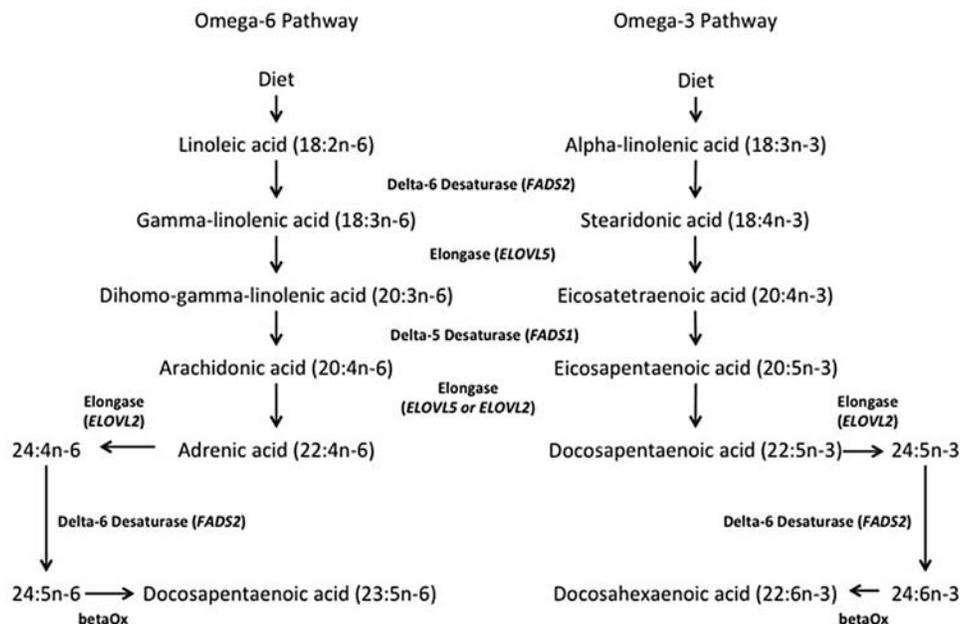


FIGURE 47.1 The metabolic pathways of n-3 and n-6 fatty acids. Adapted from Sheppard and Cheatham (2013).

of substrates and lower levels of products of the desaturases in their plasma phospholipids and red blood cells relative to those homozygous for the major alleles. This difference suggests a lower level of activity in the metabolic pathway presumably due to lower levels of delta-5 and delta-6 desaturase. In addition, relative to carriers of the major *FADS* alleles, women who are homozygous for the minor alleles produce milk containing lower levels of DHA, AA, and eicosapentaenoic acid [20:5 (*n*-3)] (EPA), all of which are fatty acids with potential to positively affect the brain. With the exception of AA, the differences in fatty acids by genotype are twofold and thus nontrivial.

Collectively, the data suggest those with the minor *FADS* SNP may not synthesize DHA at a sufficient rate, and thus, they would have a higher requirement for exogenous DHA than other genotypes. Indeed, a gene-diet interaction exists such that fish/fish oil intake and DHA content of milk in carriers of the major allele at location rs174575 on the *FADS2* gene were positively and significantly correlated, whereas no such relation existed for women who were homozygous for the minor allele. That is, the data indicate that fish intake at the level recommended (3/week) is not sufficient to increase the DHA content of milk in women who are homozygous for the minor allele at *FADS2* rs174575. Those who are homozygous for the minor SNP may, therefore, have a much higher requirement for exogenous fatty acids than those who are homozygous for the major alleles.

Given the evidence that DHA is important to the integrity of the neural substrates underlying cognitive abilities and the individual differences in fatty acid metabolism conferred by genetic status, research in which the correlation between cognition, diet, and genetic status is assessed should yield important information regarding the essentiality of and requirements for exogenous DHA. Only a few scientists have attempted to explicate the relations between DHA, *FADS* genotypes, and cognition. In what follows, details of those studies are presented along with details of work from the Cheatham Nutrition and Cognition Laboratory at the UNC-Chapel Hill Nutrition Research Institute. The evidence is sparse, but promising results are beginning to emerge from studies of early development when the brain is still forming.

FADS AND COGNITIVE DEVELOPMENT

Scientists who have studied the relation between an individual's *FADS2* status and cognitive function also report mixed results. In the seminal study of *FADS* gene-environment interaction, Caspi et al. found that the genotype at the rs174575 locus on the *FADS2* gene

was related to higher mean childhood IQ scores in carriers of the major allele (CC or CG) relative to the homozygous minor group (GG). It has long been known that children who were breastfed have IQ scores higher than those of formula-fed children. However, Caspi et al. found that this expected relation between breastfeeding and IQ was not present in those who were homozygous for the minor allele at rs174575 when comparing breastfed and bottle-fed participants. However, an attempt to replicate these findings was not successful. Whereas Caspi et al. did not find evidence of the cognitive advantage usually conferred by breastfeeding in those who were homozygous for the minor allele, Steer et al. reported that the breastfed homozygous minor allele group had the highest IQ of all the groups, albeit not significantly so. In a similar design, Martin et al. analyzed the IQs of adolescents with *FADS* SNPs (rs174575, rs1535, rs174583) against their mode of infant feeding. Martin et al. reported no significant main effects or interactions of breastfeeding and *FADS2* genotypes on IQ. It is important to note that these studies were conducted retrospectively and the formula that the participants received did not contain supplemental DHA.

One explanation for the mixed results may be sex differences. Interestingly, the influence of the *FADS* gene complex on the structure and function of brain substrates may be stronger in males than females. In the Lauritzen laboratory in Denmark, work to determine the association among *FADS1/2* SNPs and academic performance in 8–11-year-old Danish school children led them to conclude that the abilities of males, but not females, are influenced by the SNP rs17448. It is unclear whether this finding would replicate in a nonfish-eating country; the Danes have a background diet that is very high in omega-3s and, as such, are not representative of much of the world population. Nonetheless, the findings of sex differences in the influence of *FADS* SNP on brain function is intriguing.

More likely, the inconsistent findings reported with regard to the relation between genotype and IQ may be explained by maternal effects. That is, in the studies described, the results are based on the participants' genotype. However, the maternal genotype may be just as, if not more, important. The development of the brain not only relies on an individual's ability to metabolize fatty acids but also relies on the ability of the placenta to deliver sufficient fatty acids in utero when important structural DHA is accruing. In the case of the breastfed infant, the ability of the mammary glands to produce milk that has sufficient fatty acids is integral to continued optimization of brain development. Therefore, since maternal genetics are central to placental transfer and milk production, the genotype of the mother is an important variable in the consideration of fatty acid effects on development of the brain.

In sum, unobserved differences in genetically controlled maternal fatty acid metabolism may underlie the mixed results of DHA supplementation trials designed to explicate the relation between fatty acids and cognitive abilities. Moreover, for women with a certain genetic makeup, consumption of exogenous, preformed DHA may be integral to fetal and infant brain development and subsequent cognitive development of offspring, whereas for others, an increase in DHA intake may not be necessary, and as previously mentioned, may even be detrimental to infant outcomes. Optimal infant development may be achieved, in part, by individualized nutritional guidance based on the characterization of not only infant and maternal genotypes but also their interaction.

MATERNAL FADS GENOTYPES AND COGNITIVE DEVELOPMENT

Two things are clear: maternal stores of fatty acids are essential for optimal fetal brain development, and in the exclusively breastfed infant, mother's milk is the only source of exogenous fatty acids. Thus, the fetus/infant is dependent on maternal fatty acids. Few studies have been conducted to explore the idea that maternal genetics are a correlate of the effects of fatty acids on fetal and infant brain development. In a prospective study (participants aged 0–18 months of age) of the relation between fatty acids in colostrum (first milk), genetic status, and infant cognition, Morales et al. found, when

considering the individuals' *FADS1* and *ELOVL5* SNP, that breastfeeding was related to higher cognitive abilities regardless of SNPs, but formula-fed infants had lower cognitive abilities that were related to their own genotypes. The variants in the *FADS* and *ELOVL* genes resulted in formula feeding differentially affecting an infant's cognitive development. Importantly, they also found that maternal SNPs were related to cognitive abilities. Maternal *FADS2* SNPs that were related to high enzymatic activity in the n-3 pathway (higher output of products, e.g., DHA) were related to higher cognitive abilities in the infants, relative to infants of mothers with lower enzymatic activity. Infants of mothers whose *ELOVL5* SNPs were associated with higher EPA/AA and DHA/AA ratios in colostrum also had higher cognitive abilities, relative to others. This seminal prospective study illustrates the importance of considering the genetics of the dyad rather than only the individual.

In a study designed to determine the relation between maternal *FADS2* genotype, fatty acid content of milk, and infant recognition memory abilities, milk samples from mothers who were not taking any fatty acid supplements were assayed for fatty acid content when the infants were 3 months of age. As shown in Fig. 47.2, DHA was lower ($P < .05$) in the milk of those who were homozygous for the minor allele (GG) relative to those who were homozygous for the major alleles (AA) at *FADS2* rs174553. Also, there was no difference between the milk of mothers who were homozygous for the major alleles and those who were heterozygous (rs174553). Similar results were found when the sample

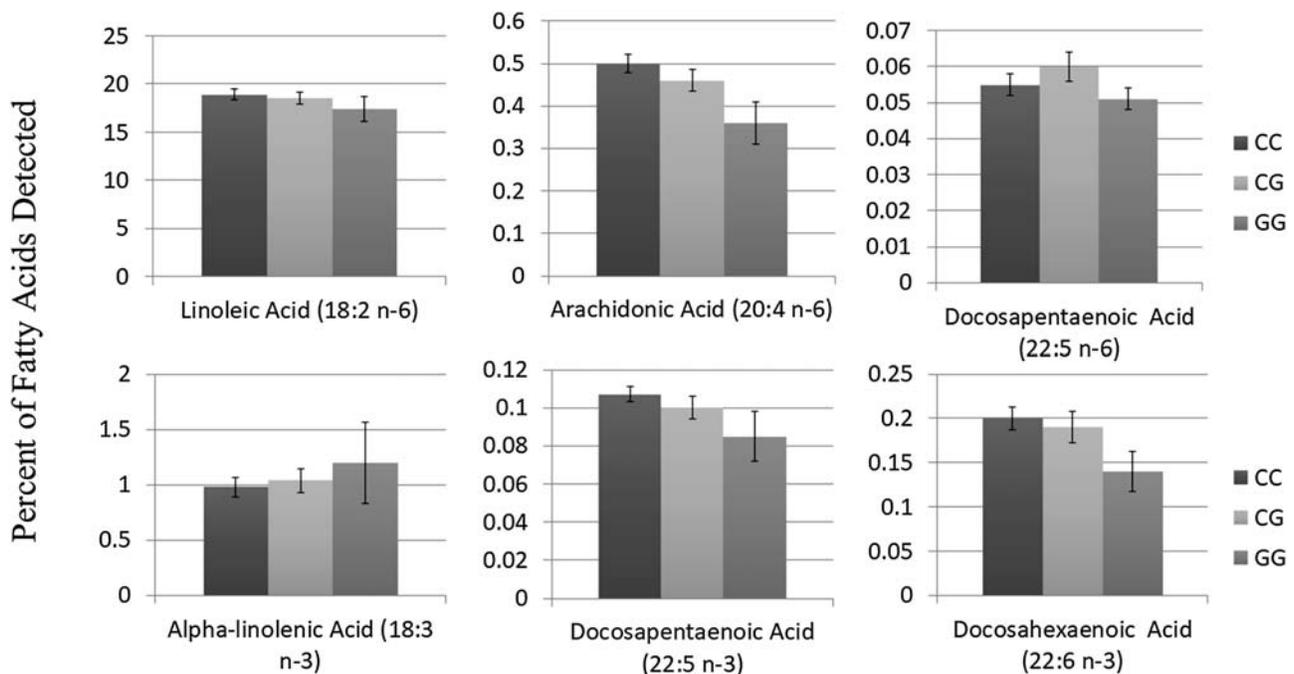


FIGURE 47.2 Percent of fatty acids detected in precursor and product n-3 and n-6 in human milk 3 months postpartum by maternal genotype at *FADS2* rs174553 (AA: black bar, AG: light-grey bar, GG: dark-grey bar).

was divided by maternal *FADS2* rs174575 (the SNP explored in the retrospective studies detailed previously). Because the incidence of rs174553 is higher in the cohort than rs174575 (18% and 9%, respectively), data when the cohort is divided by rs174553 are reported here to improve the ability to detect differences. Importantly, while there was more variability in intake of LNA (essential precursor), there was no statistically significant difference in precursor content. There was, however, a difference between the ratio of products to precursors, by maternal genotype, as shown in Fig. 47.3, which suggests that the issue may have been a shortage of desaturase.

Infants of GG mothers were therefore consuming milk that contained lower levels of fatty acids relative to their counterparts whose mothers were AA and AG. To test how this difference translated into brain function, when the infants were 6 months old, they returned for an electrophysiology session in which recognition memory was tested in an event-related potential paradigm. After being habituated to a picture of a wooden toy, they participated in a 70-30 oddball paradigm that included a random presentation of 70 of the familiar picture and 30 trial-unique pictures. At 6 months of age, brain activity of the typical infant will indicate whether she or he differentiated between the novel and familiar pictures in that less processing of the familiar stimulus will be needed (smaller absolute amplitude in the negative deflection waveform) relative to brain activity when processing a novel stimulus (larger absolute amplitude in the negative deflection waveform). This difference is taken as recognition of the familiar stimulus or recognition memory. In this study, the infants whose mothers were GG, and thus were consuming milk with lower levels of fatty acids, did not differentiate between novel and familiar pictures, as evidenced by their brain activity (Fig. 47.4, right two bars), relative to infants of the

mothers who were homozygous for the major allele who exhibited typical brain activity (left two bars on Fig. 47.4). The inability to recognize stimuli as familiar could be due to behavioral (e.g., infant not fully examining stimuli due to an attentional issue) and/or neurological (e.g., slowed or dysfunctional parahippocampal synapse firing) differences. It should be noted that this difference did not emerge when the analyses were conducted on infant genotype.

It is not surprising that maternal genetics are primary at 6 months of age when the mother exclusively has provided both prenatal and postnatal fatty acids based on her genetic background. When does the individual's ability to biosynthesize fatty acids become important? Scientists who study prenatal effects of fatty acids might argue that the structural DHA accrued in the third trimester of gestation remains central to brain function for an extended period. Indeed, some have reported effects of prenatal supplementation in children as old as 7 years. Thus, to a certain extent, maternal *FADS* status could be considered a variable for an extended period of time. In fact, in a reanalysis of a clinical trial conducted by Cheatham and colleagues, data for 16-month-olds revealed that their ability to perform in a declarative memory task was not related to their own *FADS2* status but rather was related to maternal status: toddlers whose mothers were GG at rs174575 could not recall the steps of a memory test, whereas others in the cohort performed as expected. These results suggest that the fatty acids controlled by maternal genetics and conferred by the placenta (and possibly in the breastmilk, but not all of these toddlers were breastfed) continue to support brain function in the first 2 years of life, at a minimum.

The retrospective studies of individuals' *FADS* status and their cognitive abilities when they were children were very generative. It is clear that scientists should

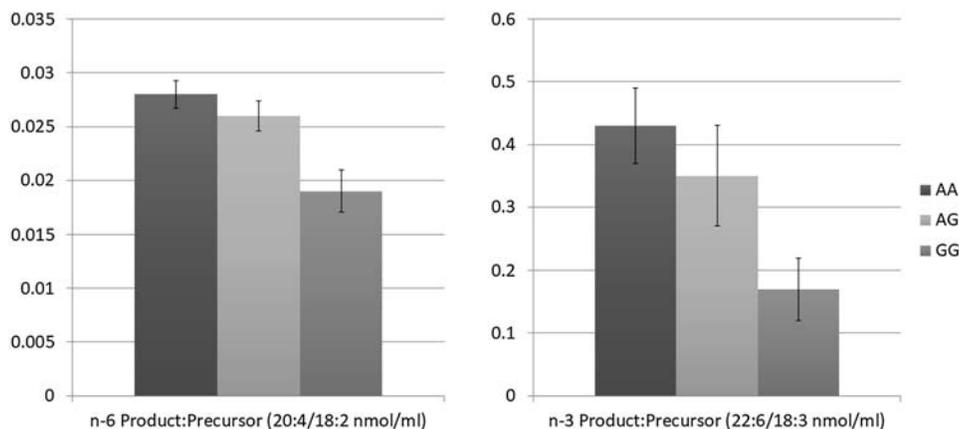


FIGURE 47.3 n-6 and n-3 ratio of the percent of total product:precursor detected in human milk by maternal *FADS2* rs174553 (AA: black bar, AG: light-grey bar, GG: dark-grey bar).

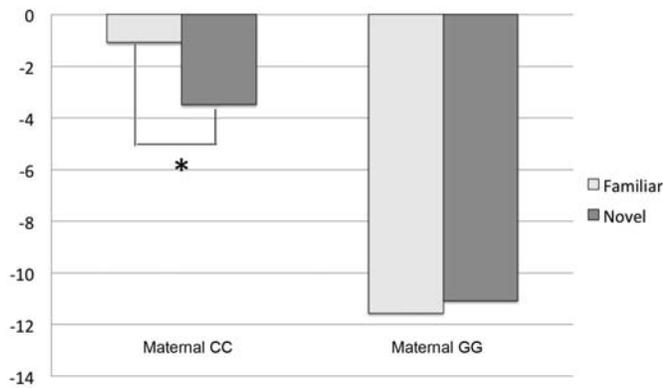


FIGURE 47.4 Negative amplitude of midline brain activity in response to novel and familiar stimuli (lower negative deflection is indicative of recognition). Bars on the left represent brain activity of infants whose mothers were CC at rs174575 and the bars on the right represent the brain activity of infants whose mothers were GG at rs174575. * $p < 0.01$.

consider maternal genetics in future studies. Additionally, much can be learned from prospective studies as cognitive abilities concurrent with the metabolic differences exacted by the differences in genetic activity lend themselves more readily to conclusions of an effect.

References

Sheppard, K.W., Cheatham, C.L., 2013. Omega-6 to omega-3 fatty acid ratio and higher order cognitive functions in 7- to 9-year-olds : a cross-sectional study. *Am J Clin Nutr* 98, 1–9.

Suggested Reading

Caspi, A., Williams, B., Kim-Cohen, J., Craig, I.W., Milne, B.J., Poulton, R., Schalkwyk, L.C., Taylor, A., Werts, H., Moffitt, T.E.,

2007. Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism. *Proc Natl Acad Sci USA* 104 (47), 18860–18865.

Cheatham, C.L., Lupu, D.S., Niculescu, M.D., 2015. Genetic and epigenetic transgenerational implications related to omega-3 fatty acids. Part II : maternal FADS2 rs174575 genotype and DNA methylation predict cognitive performance. *Nutr Res* 35, 948–955.

Delgado-Noguera, M.F., Calvache, J.A., Bonfill Cosp, X., Kotanidou, E.P., Galli-Tsinopoulou, A., 2015. Supplementation with long chain polyunsaturated fatty acids (LCPUFA) to breastfeeding mothers for improving child growth and development. *Cochrane Database Syst Rev* 7, CD007901.

Gould, J.F., Smithers, L.G., Makrides, M., 2013. The effect of maternal omega-3 (n-3) LCPUFA supplementation during pregnancy on early childhood cognitive and visual development: a systematic review and meta-analysis of randomized controlled trials. *Am J Clin Nutr* 97 (3), 531–544.

Lauritzen, L., Sorensen, L.B., Harslof, L.B., Ritz, C., Stark, K.D., Astrup, A., Dyssegaard, C.B., Egelund, N., Michaelson, K.F., Damsgaard, C.T., 2017. Mendelian randomization shows sex-specific associations between long-chain PUFA-related genotypes and cognitive performance in Danish schoolchildren. *Am J Clin Nutr* 106, 88–95.

Lupu, D.S., Cheatham, C.L., Corbin, K.D., Niculescu, M.D., 2015. Genetic and epigenetic transgenerational implications related to omega-3 fatty acids. Part I: maternal FADS2 genotype and DNA methylation correlate with polyunsaturated fatty acid status in toddlers: an exploratory analysis. *Nutr Res* 35, 939–947.

Martin, N.W., Benyamin, B., Hansell, N.K., Montgomery, G.W., Martin, N.G., Wright, M.J., Bates, T.C., 2011. Cognitive function in adolescence: testing for interactions between breast-feeding and FADS2 polymorphisms. *J Am Acad Child Adolesc Psychiatry* 50 (1), 55–62.

Morales, E., Bustamante, M., Gonzalez, J.R., Guxens, M., Torrent, M., Mendez, M., Garcia-Esteban, R., Julvez, J., Forns, J., Vrijheid, M., et al., 2011. Genetic variants of the FADS gene cluster and ELOVL gene family, colostrums LC-PUFA levels, breastfeeding, and child cognition. *PLoS One* 6 (2), e17181.

Steer, C.D., Davey Smith, G., Emmett, P.M., Hibbeln, J.R., Golding, J., 2010. FADS2 polymorphisms modify the effect of breastfeeding on child IQ. *PLoS One* 5 (7), e11570.

This page intentionally left blank

The Nutrigenetics of Cardiovascular Disease

Rosalinda Madonna¹, Raffaele De Caterina²

¹Center of Aging Sciences and Translational Medicine, CESI-MeT, “G. D’Annunzio” University, Chieti, Italy ²Cattedra di Malattie dell’Apparato Cardiovascolare, Università degli Studi di Pisa, Italy

Glossary

5-LO 5-lipoxygenase
 APOE apolipoprotein E
 CVDs cardiovascular disease
 DHA docosahexaenoic acid
 EPA eicosapentaenoic acid
 FADS fatty acid desaturase
 MTHFR methylenetetrahydrofolate reductase
 PPARs peroxisome proliferator-activated receptors
 PUFA polyunsaturated fatty acids
 SNPs single nucleotide polymorphisms

NUTRIENTS, GENES, AND CARDIOVASCULAR DISEASES: DEFINING THE PROBLEM

Cardiovascular disease (CVD) here refers to diseases of the cardiovascular system with atherosclerosis as the background pathology and with thrombosis as a main mechanism of disease instability: atherothrombotic CVD. Its manifestations affect the heart (acute coronary syndromes and stable coronary heart disease), the brain (stroke), or the peripheral circulation (peripheral arterial disease) and is the main cause of death worldwide. Epidemiologic studies identified several risk factors involved in its pathophysiology, such as smoking, hypertension, diabetes mellitus, obesity, dietary habits, physical inactivity, alcohol consumption, dyslipidemias, and psychosocial factors. According to the INTERHEART Study, these modifiable risk factors may explain 90% of the population-attributable risk in men and 94% in women (Yusuf et al., 2004). However, most of these factors have a genetic background and are either strictly dietary factors (e.g., alcohol consumption, consumption of fruits and vegetables [see related

chapters]) or modifiable by the diet (hypertension, diabetes, obesity, dyslipidemias [see related chapters]). The interaction of diet, risk factors, and the individual genotype is the main determinant of the intensity and extent to which CVD surfaces in an individual. Gene–diet interaction (nutrigenetics) thus has an important role in determining the structural background, function, and homeostasis of the cardiovascular system, conditioning the onset and progression of CVD (De Caterina and Madonna, 2004). Understanding gene–diet interactions may thus lead to a better understanding of the individual susceptibility to CVD.

From a pathophysiological standpoint, nutrients and genes interact, affecting mechanisms of cardiovascular repair and regeneration. Abnormalities in the ability of the cardiovascular system to undergo repair and regeneration can induce or exacerbate ischemic conditions such as myocardial infarction or stroke. On the other hand, excessive tissue regeneration (or rather new tissue generation) via neoangiogenesis can induce hypertension, diabetes, inflammation, atherosclerotic plaque instability, and thrombosis (De Caterina, 2007).

The age of a person determines for how long the exposure of an individual to environmental factors such as dietary components has occurred, and therefore has a key role in the transition from health to disease. The study of endothelial dysfunction and of related cellular and molecular mechanisms has attracted the interest of a large number of research groups, in accordance with the assumption that endothelial dysfunction is the first pathological condition associated with most age-associated disease, including CVD. Here, endothelial cells, which are present in all tissues and organs, also act as a filter against pathogenic noxae, some of which, derived from the diet, come in contact with the blood (De Caterina, 2007). An imbalance between

endothelial death and regeneration may thus favor the onset and progression of CVD (De Caterina, 2007; Ramos-Lopez et al., 2017).

This chapter describes a few key players in genetic susceptibility to the effect of nutrients as related to the cardiovascular system, with a special focus on genetic variants that have an important role in the onset or clinical manifestations of CVD. We will thus review a few selected examples of regulation of gene–nutrient interaction. In particular, we will focus on polymorphisms of the following genes associated with different types of responsiveness to dietary factors: apolipoprotein E (*APOE*), fatty acid desaturase (*FADS*), 5-lipoxygenase (*5-LO*), peroxisome proliferator-activated receptors (*PPARs*), and methylenetetrahydrofolate reductase (*MTHFR*).

APOLIPOPROTEIN E

(ApoE, encoded by the *APOE* gene) is a 34-kDa protein of 299 amino acids, with a single glycosylation site at threonine-194. ApoE has a central role in lipoprotein metabolism. It is involved in the metabolism of chylomicrons, the synthesis and secretion of very–low density lipoproteins (VLDL), and the removal of cellular debris from the circulation. ApoE also has a role in blood coagulation, oxidative processes, macrophage, glial cell and neuronal cell homeostasis, adrenal function, central nervous system physiology, inflammation, and cell proliferation.

First described in the early 1970s as a minor apolipoprotein in VLDL, ApoE was subsequently identified as a major apolipoprotein in cholesterol-rich VLDL (β -VLDL) and low density lipoprotein (LDL) in cholesterol-fed animals. ApoE has two structural domains separated by a hinge region. The amino-terminal domain (amino acids 1–191) contains the LDL receptor binding region (amino acids 136–150). Two missense single nucleotide polymorphisms (SNPs) in the *APOE* gene on chromosome 19 result in three allelic isoforms, known as *E2*, *E3*, and *E4*. These different *APOE* isoforms differ from each other, with *APOE2* containing a Cys at amino acids 112 and 158, *APOE3* containing Cys at amino acid residue 112 and Arg at residue 158, and *APOE4* containing Arg at both positions. These amino acid changes influence salt bridge formation between the *N*- and *C*-terminal domains of the protein, which have an impact on receptor binding activities, lipoprotein preference, and ApoE stability, and ultimately on tissue protein concentrations. The prevalence of these SNPs varies in different populations (Bennet et al., 2007). It has been reported that approximately 65% of Caucasians are *E3/E3* homozygous, 21% are *E4* carriers (*E3/E4* or *E4/E4*), 11% are *E2* carriers (*E2/E2* or *E2/E3*), and the remaining 4% have

the *E2/E4* genotype. The *E4* isoform is the most atherogenic (Dwyer et al., 2004), and *E4* carriers have concentrations of LDL cholesterol by 31% higher than *E2* carriers. An average higher 40%–50% risk for coronary heart disease is observed in male *E4* carriers, because they have significantly higher levels of LDL cholesterol (Bennet et al., 2007). Women with the *E4* allele have been reported to have premature coronary heart disease and increased risk for a heart attack. Individuals with the *APOE4* genotype have significantly higher systolic and diastolic blood pressure. As to nutrigenetic interactions, *APOE4* carriers have been shown to be associated with variable response to fish oil intervention and a higher risk for CVD (Bennet et al., 2007). Nevertheless, the variable penetrance of *APOE4* carriers in the general population can influence the interindividual response to modifiers such as intake of saturated or unsaturated fat and cholesterol, and smoking status. Several studies showed that individuals with the *APOE4* genotype have significantly greater reduction in triglycerides after eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) supplementation (Bennet et al., 2007). In the SAT-GENE study, 44 *E3/E3* and 44 *E3/E4* carriers were recruited and subjected to a sequential dietary intervention, by which they were assigned to a low-fat diet, a high-fat diet, and a high-fat diet supplemented with 3.45 g DHA, each for 8 weeks. The study showed greater sensitivity of fasting triglycerides to dietary fat manipulation in *APOE3/E4* genotype carriers (25% of the population), with no effect of this allelic profile on cholesterol (Carvalho-Wells et al., 2012). Nevertheless, in a study of 38 healthy normolipidemic males prospectively recruited on the basis of *APOE* genotype, high-dose DHA supplementation was associated with increases in total cholesterol and LDL-C in *E4* carriers, which appeared to negate the cardioprotective action of DHA in this population subgroup (Olano-Martin et al., 2010). The triglyceride response to EPA and DHA dietary supplementation is also highly variable. In the FIN-GEN trial, no triglyceride lowering was observed in 118 of 312 participants in response to the higher dose of EPA and DHA tested (Caslake et al., 2008). However, the greatest triglyceride-lowering responses were evident in *APOE4* homozygotes (Caslake et al., 2008).

FATTY ACID DESATURASE

Fatty acid desaturases (encoded by the *FADS1* and *FADS2* genes) are enzymes involved in the synthesis of polyunsaturated fatty acids (n-3 PUFA and n-6 PUFA), namely in the conversion of linoleic acid (LA) into arachidonic acid (AA), and of α -linolenic acid (ALA) into EPA. Several studies showed that plasma lipid response to n-3 PUFA supplementation can be influenced by

genetic variations in the *FADS* gene cluster. Individuals with the C allele of rs174546 *FADS1* genotype and with high intake of n-6 PUFAs also have higher total cholesterol and higher high-density lipoprotein (HDL) cholesterol than those with a low intake. However, Dumont et al. showed that the C allele of rs174546 is associated with decreased plasma triglycerides and non-HDL-C levels, independent of PUFA intake (Dumont et al., 2011). Studies have also shown that *FADS1* haplotypes, including the C allele rs174546, are strongly associated with lower long-chain PUFA synthesis and transcript levels (Ameur et al., 2012). Reduced long-chain PUFA substrate availability leads to a reduction in VLDL and triglyceride synthesis. Therefore, the *FADS* gene cluster, especially rs174546, is associated with variable levels of plasma lipids, which makes it a significant SNP in terms of associations between plasma lipids and metabolism of fatty acids.

5-LIPOXYGENASE

5-*LO* is an enzyme responsible for the production of proinflammatory leukotrienes from arachidonic acid (De Caterina and Zampolli, 2004) (Fig. 48.1). Individuals

carrying two variant alleles of 5-*LO* have increased carotid intima-media thickness compared with carriers of the common allele (Dwyer et al., 2004) (Fig. 48.2). C-Reactive protein is also twofold higher in carriers of two variant alleles (Dwyer et al., 2004). Higher dietary intakes of n-6 PUFAs (AA and LA) significantly increase the atherogenic effect of two variant alleles of 5-*LO*. Increased intake of DHA and EPA blunts this effect, along with decreased risk of myocardial infarction. These genetic variants are responsible for a greater expression of 5-*LO* in the presence of an AA-rich diet, and this may increase the risk of CVD (Dwyer et al., 2004).

METHYLENETETRAHYDROFOLATE REDUCTASE

MTHFR is an enzyme encoded by the *MTHFR* gene; it catalyzes the conversion of homocysteine to methionine (Fig. 48.3). High levels of homocysteine (hyperhomocysteinemia), often resulting from *MTHFR* gene polymorphisms and/or low intake of cofactors such as folate, vitamin B₆, and vitamin B₁₂, can themselves determine MTHFR dysfunction. Hyperhomocysteinemia is potentially an independent risk factor for myocardial

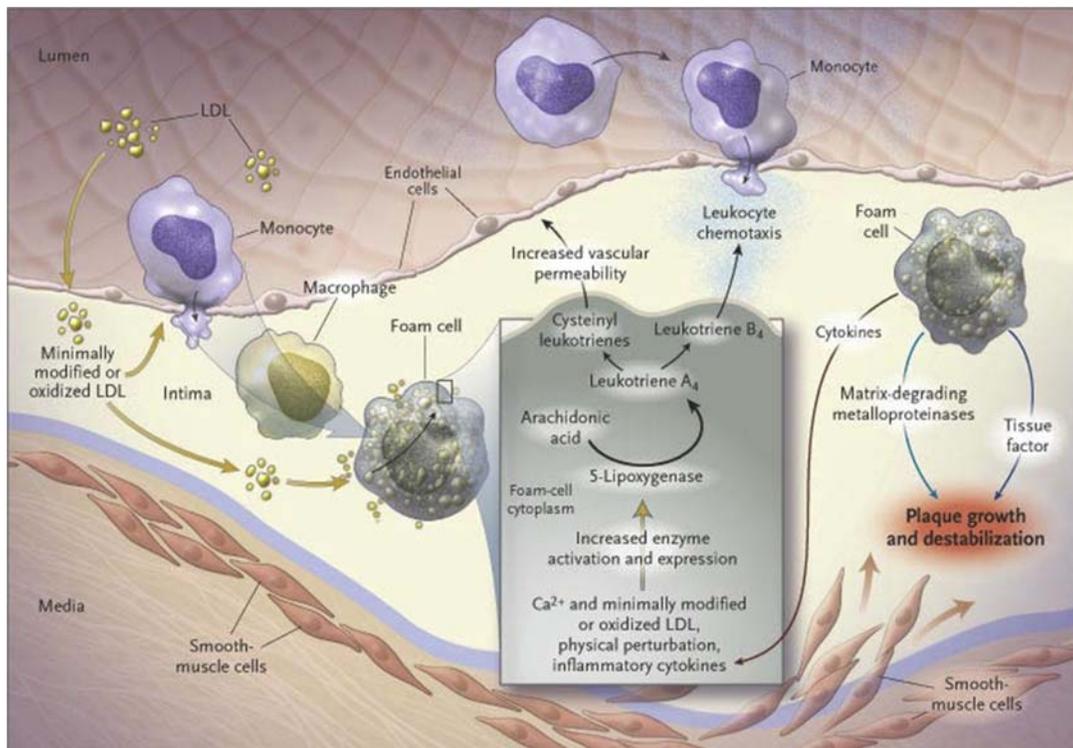


FIGURE 48.1 Role of 5-lipoxygenase in the early development of atherosclerosis. 5-Lipoxygenase is responsible for leukotriene biosynthesis from arachidonic acid. Leukotrienes can contribute to atherosclerosis by promoting nonspecific leukocyte chemotaxis (leukotriene B₄) and by increasing vascular permeability (cysteinyl leukotrienes C₄, D₄, and E₄). The cytokines released during inflammation induce the expression and activity of 5-lipoxygenase, thus initiating a vicious circle that perpetuates and amplifies the proatherosclerotic process. *LDL*, low density lipoprotein. Reproduced from De Caterina, R., Zampolli, A., 2004. *N Engl J Med* 350, 4–7, with permission.

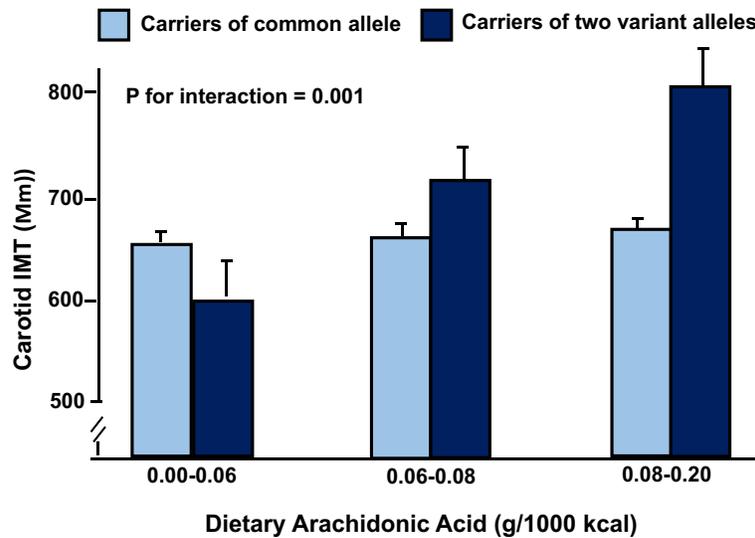


FIGURE 48.2 Effect of 5-lipoxygenase promoter genotype on carotid intima–media thickness (IMT). Columns and bars represent means \pm standard error. Redrawn from Dwyer, J.H., et al., 2004. *N Engl J Med* 350, 29–37, with permission.

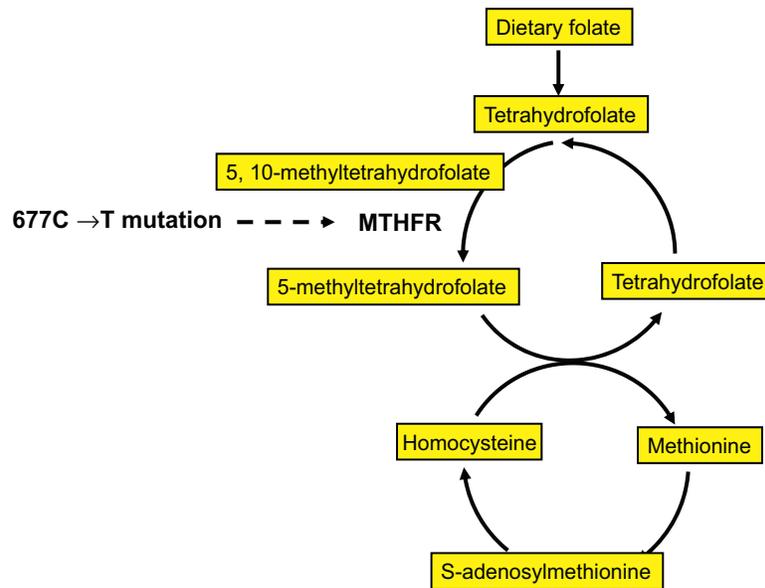


FIGURE 48.3 The metabolism of homocysteine. Methylene tetrahydrofolate reductase (MTHFR) converts the inactive form of folate into active 5-methyltetrahydrofolate, which provides the methyl group to homocysteine. Homocysteine is methylated to produce methionine. Methionine is further converted into S-adenosylmethionine, which acts as a donor of the methyl group in methyltransferase-catalyzed reactions and is converted to S-adenosylhomocysteine, further metabolized back into homocysteine. The 677C \rightarrow T variant encodes a thermolabile enzyme with reduced activity, which is responsible for hyperhomocysteinemia.

infarction, hypertension, stroke, and glaucoma. The *MTHFR* C677T mutation is the polymorphism mostly associated with CVD. It may occur in the heterozygous (CT) and homozygous genotype (TT). The 677C \rightarrow T variant encodes a thermolabile enzyme with reduced activity. Homocysteine levels are higher in individuals carrying the TT and the CT genotypes. Those individuals undergo significant homocysteine reduction after a Mediterranean diet intake. However, a large-scale prospective study enrolling 24,968 participants and examining

healthy white American women for 10 years did not confirm the effect of the Mediterranean diet on the incidence of CVD (de Lorgeril and Salen, 2011). In fact, although the TT genotype determined higher homocysteine levels and higher intakes of folate and B vitamins reduce such homocysteine levels, neither the *MTHFR* 677C \rightarrow T mutation nor the intake of folate or B vitamins significantly affected the incidence of CVD. However, a metaanalysis on cardiovascular outcomes related to the *MTHFR* C677T polymorphism showed that populations

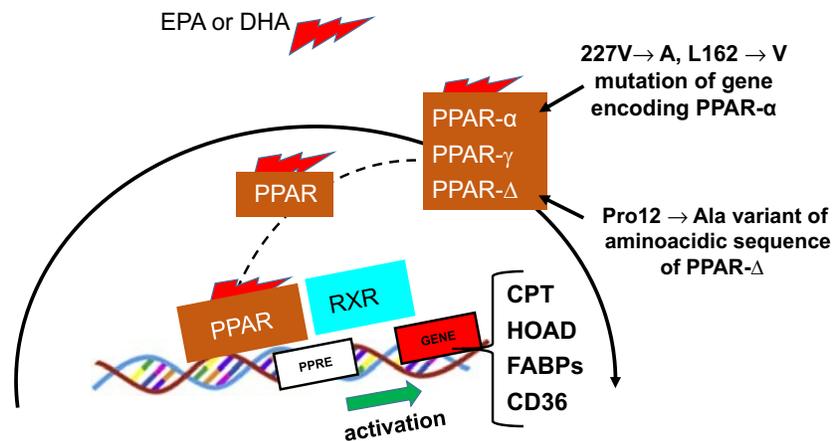


FIGURE 48.4 Binding of dietary n-3 fatty acids to transcription factors peroxisome proliferator-activated receptors α and β (PPAR α and β) activates genes regulating lipid metabolism. n-3 eicosapentaenoic acid (EPA) and n-3 docosahexaenoic acid (DHA) are natural ligands for PPARs. Binding to the ligand causes the dimerization of PPAR with the retinoid X receptor (RXR). The PPAR-RXR complex activates target genes by recognizing the regions of the promoter called peroxisome proliferator response elements (PPRE). Target genes include the β -oxidation enzymes carnitine palmitoyl transferase (CPT) and hydroxyacyl dehydrogenase (HOAD), fatty acid binding proteins (FABPs), and the transmembrane fatty acid transporters CD36 and FAT. The figure shows the main mutations and variants of the PPARs.

carrying the TT genotype experienced a higher incidence of stroke in response to low folate intake.

PEROXISOME PROLIFERATOR-ACTIVATED RECEPTORS

PPARs are nuclear receptors binding specific DNA sequences (PPAR response elements [PPREs]), located in the promoter regions of several genes involved in the lipid and glucose metabolism, and affecting cell differentiation and development, as well as glucose and lipid metabolism (Fig. 48.4). n-3 and n-6 PUFAs bind to PPARs after they are transported into the nucleus through the binding of a fatty acid binding protein (Fig. 48.4). Three different subtypes of PPARs are known: PPAR- α , γ , and δ , all of which are involved in lipid and glucose metabolism. PPAR- δ has an important role in development and is the least involved in lipid or glucose metabolism and in CVD. Conversely, genetic variants of PPAR- α and γ are highly associated with lipid metabolism and CVD risk. PPAR- α is highly expressed in organs with high lipid metabolism, such as the liver, heart, and skeletal muscle. PPAR- α has a key role in lipid oxidation and inflammation by decreasing plasma triglycerides and increasing plasma HDL, resulting in a less atherogenic lipid profile. In the Atherosclerosis Risk in Communities study, homozygous carriers of the V227A gene variant of PPAR- α were associated with lower total and LDL cholesterol after higher intake of n-6 fatty acids (Volcik et al., 2008). In addition, the L162V variant was associated with higher

fasting total and LDL cholesterol in individuals who switched from a low polyunsaturated to a saturated and high-fat diet. PPAR- γ is highly expressed in the adipose tissue, where it regulates adipocyte differentiation, glucose and lipid storage, and inflammation (Volcik et al., 2008). PPAR- γ is also expressed in the skeletal muscle, monocytes, and endothelium, where it has a role in controlling insulin sensitivity, blood pressure, and inflammation. Homozygous carriers of the P12A variant of PPAR- γ are associated with an increased risk for adiposity and insulin resistance. Obese subjects with the P12A variant who consumed fewer monounsaturated fatty acids (MUFA) are more insulin resistant. In a fish oil n-3 PUFA supplementation study, carriers of the A12 allele were reported to have greater decrease in plasma triglycerides after n-3 PUFA supplementation when total fat intake was below 37% energy or the intake of saturated fatty acids was below 10% energy.

CONCLUSIONS

Available data on gene variants capable of influencing lipid metabolism and responsiveness to PUFA supplementation are relatively limited, although there are clear demonstrations of the importance of the gene variants of *APOE* and *FADS*. Understanding the genetic modulators of the PUFA response may allow targeted preventive intervention and the formulation of personalized recommendations for the intake of fatty acids, based on the genotype of these key enzymes involved in the lipid metabolism.

References

- Ameur, A., Enroth, S., Johansson, A., et al., 2012. Genetic adaptation of fatty-acid metabolism: a human-specific haplotype increasing the biosynthesis of long-chain omega-3 and omega-6 fatty acids. *Am J Hum Genet* 90, 809–820.
- Bennet, A.M., Di Angelantonio, E., Ye, Z., et al., 2007. Association of apolipoprotein E genotypes with lipid levels and coronary risk. *J Am Med Assoc* 298, 1300–1311.
- Carvalho-Wells, A.L., Jackson, K.G., Lockyer, S., Lovegrove, J.A., Minihane, A.M., 2012. APOE genotype influences triglyceride and C-reactive protein responses to altered dietary fat intake in UK adults. *Am J Clin Nutr* 96, 1447–1453.
- Caslake, M.J., Miles, E.A., Kofler, B.M., et al., 2008. Effect of sex and genotype on cardiovascular biomarker response to fish oils: the FIN-GEN Study. *Am J Clin Nutr* 88, 618–629.
- De Caterina, R., 2007. *Endothelial Dysfunctions in Vascular Disease*. Blackwell Publishing, Oxford, UK.
- De Caterina, R., Madonna, R., 2004. Nutrients and gene expression. *World Rev Nutr Diet* 93, 99–133.
- De Caterina, R., Zampolli, A., 2004. From asthma to atherosclerosis—5-lipoxygenase, leukotrienes, and inflammation. *N Engl J Med* 350, 4–7.
- de Lorgeril, M., Salen, P., 2011. Mediterranean diet in secondary prevention of CHD. *Publ Health Nutr* 14, 2333–2337.
- Dumont, J., Huybrechts, I., Spinneker, A., et al., 2011. FADS1 genetic variability interacts with dietary alpha-linolenic acid intake to affect serum non-HDL-cholesterol concentrations in European adolescents. *J Nutr* 141, 1247–1253.
- Dwyer, J.H., Allayee, H., Dwyer, K.M., et al., 2004. Arachidonate 5-lipoxygenase promoter genotype, dietary arachidonic acid, and atherosclerosis. *N Engl J Med* 350, 29–37.
- Olano-Martin, E., Anil, E., Caslake, M.J., et al., 2010. Contribution of apolipoprotein E genotype and docosahexaenoic acid to the LDL-cholesterol response to fish oil. *Atherosclerosis* 209, 104–110.
- Ramos-Lopez, O., Milagro, F.I., Allayee, H., et al., 2017. Guide for current nutrigenetic, nutrigenomic, and nutriepigenetic approaches for precision nutrition involving the prevention and management of chronic diseases associated with obesity. *J Nutrigenetics Nutrigenomics* 10, 43–62.
- Volcik, K.A., Nettleton, J.A., Ballantyne, C.M., Boerwinkle, E., 2008. Peroxisome proliferator-activated receptor [alpha] genetic variation interacts with n-6 and long-chain n-3 fatty acid intake to affect total cholesterol and LDL-cholesterol concentrations in the Atherosclerosis Risk in Communities Study. *Am J Clin Nutr* 87, 1926–1931.
- Yusuf, S., Hawken, S., Ounpuu, S., et al., 2004. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* 364, 937–952.

Hemostasis and Thrombosis

Antonio Capurso¹, Cristiano Capurso²

¹Geriatrics and Internal Medicine, University of Bari Medical School, Bari, Italy ²Internal Medicine, University of Foggia, Foggia, Italy

HEMOSTASIS AND COAGULATION

Coagulation of the blood safeguards against death by hemorrhage or against the undue loss of blood.

Primary and Secondary Hemostasis

Hemostasis is the normal physiological response that prevents significant blood loss after vascular injury, whereas thrombosis is the pathological intravascular phenomenon that leads to the formation of a clot along the wall of a blood vessel, frequently causing occlusion of the vessel. Hemostasis depends on a complex series of events involving platelets and the activation of specific blood proteins, known as coagulation factors (Figs. 49.1–49.3). These coagulation factors are generally serine proteases (enzymes), with the exception of Factor V (FV) and FVIII, which are glycoproteins, and FXIII, which is a transglutaminase. They circulate as inactive zymogens (inactive enzyme precursors) and act by cleaving downstream proteins so that they become active enzymes.

Coagulation begins almost instantly after injury to a blood vessel has damaged the endothelium lining the vessel. The exposure of blood to the space under the endothelium initiates two processes:

- (1) *Primary hemostasis*, in which platelets change their status and adhere to the site of injury to form a primary cloth. When the endothelium is damaged, the normally isolated, underlying collagen becomes exposed to circulating platelets that bind directly to collagen with collagen-specific glycoprotein Ia/IIa surface receptors. This adhesion is further strengthened by von Willebrand factor (vWF) released from the endothelium and platelets; vWF forms additional links between the platelets' glycoprotein Ib/IX/V and the collagen fibrils. Through these mechanisms, platelets become

activated and release the contents of stored granules into the blood plasma. The granules include adenosine diphosphate (ADP), serotonin, platelet-activating factor (PAF), vWF, platelet factor 4, and thromboxane A₂ (TxA₂), which in turn activate additional platelets that adhere to the site of injury.

- (2) *Secondary hemostasis* is a complex phenomenon consisting of the formation of insoluble, cross-linked fibrin through the action of activated coagulation factors, particularly thrombin. Fibrin stabilizes the primary platelet plug, particularly in larger blood vessels in which the platelet plug is insufficient in itself to stop hemorrhage.

The Secondary Hemostasis Pathways

To stabilize the primary cloth, a secondary insoluble cloth made of cross-linked fibrin is created. This phenomenon represents the core of secondary hemostasis.

Traditionally, secondary hemostasis has been divided into three distinct pathways: intrinsic, extrinsic, and common pathway (Fig. 49.3).

Intrinsic pathway. The intrinsic pathway (Fig. 49.3) starts after a surface contact phenomenon and is sustained by coagulation factors FXII, FXI, and FIX, cofactor FVIII, Ca, and phosphatidylserine. The ultimate product of the intrinsic pathway is activated Factor IX (FIXa), which, with the aid of activated cofactor FVIIIa, activates FX.

Extrinsic pathway. The extrinsic pathway (Fig. 49.3) is the coagulation phenomenon that starts when tissue damage takes place; it is called the *tissue factor (TF)-activated extrinsic pathway*. It starts when, in damaged blood vessels, circulating FVII comes into contact with subendothelial cell membranes with exposed TF, forming a TF-FVII complex (Fig. 49.1). The monomer FVII is then activated to the active dimer FVIIa in the presence of FXIIa, FXa, FIXa, and thrombin. The role of TF is to

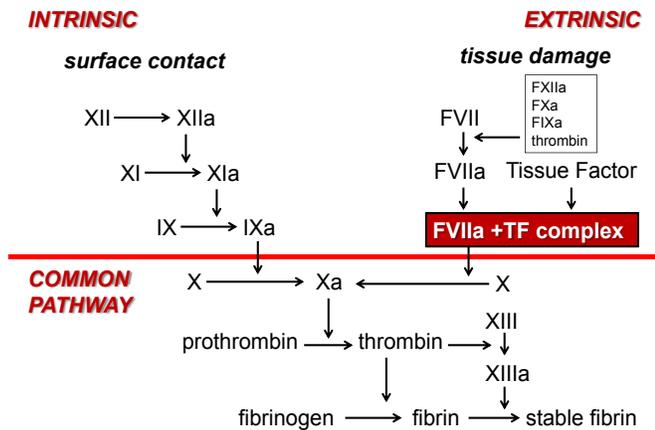


FIGURE 49.1 The coagulation cascade of secondary hemostasis. Both intrinsic and extrinsic pathways converge in the common pathway, where they activate FX which in turn activate the subsequent coagulation cascade leading to the thrombin formation and the fibrin cloth. *F*, factor; *TF*, tissue factor.

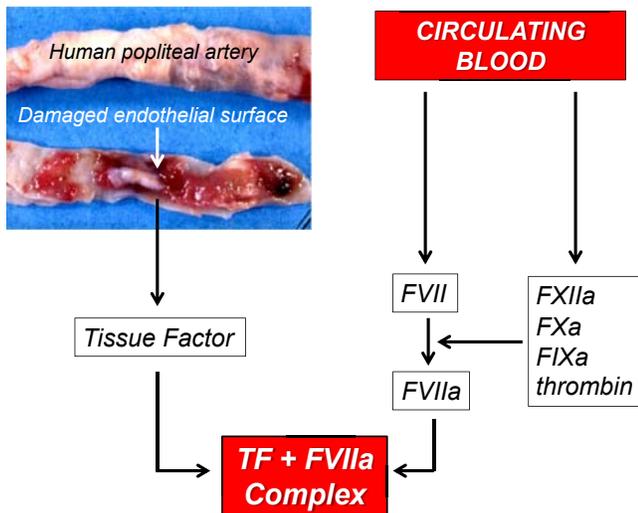


FIGURE 49.2 A schematic representation of the extrinsic pathway of blood coagulation. The extrinsic pathway is activated when blood comes in contact with cell membranes with exposed tissue factor (TF). The monomer FVII is activated to the active dimer FVIIa in the presence of FXIIa, FXa, FIXa and thrombin. These sequential proteolytic activation take place on cell membrane surfaces of damaged vessels and leads to the formation of TF + FVIIa complex. *F*, factor; *TF*, tissue factor.

enhance the activity of FVIIa, making it an efficient catalyst of FIX and FX activation. These sequential proteolytic activations take place on cell membrane surfaces and depend on phospholipids, mostly provided by activated platelets, but also endothelial cells and leukocytes. Platelets, which are attracted to the vessel wall by collagen, become activated and release several of their constituents, including fibrinogen, causing thrombosis in concert with activated coagulation and fibrin formation (Figs. 49.2 and 49.3).

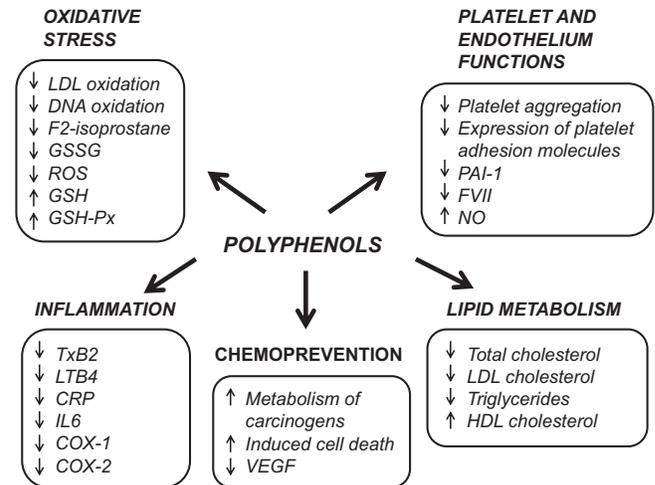


FIGURE 49.3 COX-1, cyclooxygenase-1; CRP, C-reactive protein; FVII, Factor VII; GSH, glutathione; GSH-Px, glutathione peroxidase; GSSG, glutathione disulfide; HDL, high-density lipoprotein; IL6, interleukin 6; LDL, low-density lipoprotein; LTB4, leukotriene B4; NO, nitric oxide; PAI-1, plasminogen activator inhibitor 1; ROS, reactive oxygen species; TxB2, thromboxane B₂; VEGF, vascular endothelial growth factor.

Common pathway. The extrinsic and intrinsic coagulation pathways converge on FX (Fig. 49.3), which is activated to FXa. FXa starts the final part of the coagulation process shared by *intrinsic* and *extrinsic pathways* (Fig. 49.3, area under the red line), leading to fibrin (clot) formation.

The Prominent Role of the Tissue Factor–Activated Extrinsic Pathway in the Coagulation Process and Human Pathology

It was previously thought that the two pathways of coagulation cascade, the intrinsic (platelets) pathway and the extrinsic (FVII) pathway were equally important, but it is now known that the primary pathway for initiating blood coagulation is the TF-activated-extrinsic pathway. FVII has a pivotal role in activating the common pathway of the clotting cascade and thrombin formation. The main role of the TF pathway is to generate a *thrombin burst*, a process by which thrombin, an important constituent of the coagulation cascade, is rapidly released.

The role of the TF pathway, and specifically of FVII, is particularly relevant in the thrombotic process associated with coronary heart disease (CHD). Numerous studies have determined the significant association between FVII and cardiovascular mortality. In a prospective cohort study (De Stavola and Meade, 2007) reporting mortality data after a follow-up of nearly 30 years, FVIIc was confirmed to be significantly related to CHD mortality as an independent risk factor. High FVII levels increase the risk of fatal CHD by increasing the probability of a life-threatening

coronary thrombosis. The increase in circulating FVII thus appears to be one of the most important risk factors for CHD, equal to high cholesterol levels, hypertension, smoke, and diabetes.

The Circulating Level of Factor VII Is Influenced by Diet

Coagulation FVII is a 50-kDa single-chain, vitamin K-dependent protease that has an important role in the extrinsic pathway of blood coagulation. FVII is synthesized principally in the liver and is secreted as an inactive single-chain glycoprotein. In the presence of TF, inactive FVII is converted by limited proteolysis to its fully activated two-chain form. Activation can be affected by a number of activated coagulation factors, including Xa, IXa, XIIa, and thrombin. After activation, FVIIa rapidly converts FIX and FX into their active forms, initiating the generation of thrombin and fibrin clot formation.

Numerous studies demonstrated that (i) the levels of circulation FVII are influenced to some extent by diet; (ii) there is a substantial increase in FVII circulating levels in the postprandial phase; (iii) the total intake of dietary fat appears to be the main determinant of postprandial FVII plasma levels; and (iv) the ratio of saturated fatty acids (SFAs) to monounsaturated fatty acids (MUFAs) in the diet is crucial to postprandial levels of FVII.

Postprandial Increase in Factor VII

FVIIc increases consistently in the postprandial phase. This FVII increase has been shown to be associated with postprandial triglyceride levels and takes place within 2–3 h after the intake of a fatty meal, persisting for several hours afterward. The maximum activation of FVIIa takes place 8 h after eating. Fat intake, rather than dietary energy intake, has been shown to be the primary determinant of the postprandial increase in FVIIc.

NUTRACEUTICAL EFFECTS OF DIET ON HEMOSTASIS AND THROMBOSIS

The impact of diet and its components on blood coagulation and thrombosis have been known for a long time.

Factor VII: The Impact of Extra Virgin Olive Oil and Monounsaturated Fatty Acids on Postprandial Factor VII Levels

Alimentary fats are involved in the activity of FVII. Although the results on FVII and SFA–unsaturated fatty acids have given conflicting results, more recent data

showed that olive oil and its main fatty acid, MUFA oleic acid, have a favorable impact on the activity of FVII, particularly in the postprandial phase.

Numerous studies demonstrated that during the postprandial state, a procoagulatory situation exists (an increase in thromboxanes and D-dimer and a decrease in tissue plasminogen activator). The type of fat consumed, both in an acute meal and during the previous weeks, is the main determinant of these changes.

A sustained and remarkable increase in FVII antigen (FVIIag) and activity (FVIIa) also takes place in the postprandial phase. Postprandial activation of FVIIa is mainly driven by a diet rich in long-chain saturated fatty acids. A single fat-rich meal, irrespective of whether it is rich in SFAs or polyunsaturated fatty acids (PUFAs), induces an increase in FVIIa only in individuals with a background diet rich in long-chain SFAs, not in those whose usual diet is rich in unsaturated fatty acids (Roche et al., 1998).

Of interest are data on the different postprandial responses to dietary fat in North Europe compared with South Europe. FVIIc was shown to be significantly greater 8 h after eating in Northern Europeans compared with Southern Europeans, two populations that follow different habitual diets; the northern population is rich in SFAs and the southern one is rich in MUFAs, particularly extravirgin olive oil (EVOO) (Zampelas et al., 1998). These data were further confirmed in a comparative study in which 40% of SFAs was replaced with MUFAs (Silva et al., 2003). Postprandial FVIIa and FVIIag were significantly lower after an MUFA-rich diet than an SFA-rich diet. Also, long-term dietary intervention studies (Roche et al., 1998) confirmed the favorable effect of MUFAs on FVII circulating levels. Compared with an SFA-rich diet, consumption of an MUFA-rich diet for 16 weeks was associated with significantly lower postprandial FVIIc and FVIIa levels. The beneficial effects of the MUFA-rich diet were sustained in the long term with no attenuation through adaptation.

Taken together, these data demonstrate that diets rich in MUFAs, particularly from olive oil, are associated with a lower postprandial peak level of FVII, and likely explain the lower rates of CHD in countries where the diet is habitually rich in MUFAs, such as in Southern European countries.

Factor VII: Mechanism of Action of Monounsaturated Fatty Acids on Factor VII. The Role of Triglycerides

The mechanism by which lipoproteins and fatty acids support the increase or reduction of FVII is not completely clear. However, some studies partially clarified these mechanisms.

The particle size of triglyceride-rich lipoproteins (TRLs) (chylomicrons and very low-density lipoproteins) has been suggested to be a determinant of postprandial FVII activation.

It is known that an SFA-rich diet determines the postprandial formation of a high number of triglyceride-rich chylomicrons and activation of FVII. Different from SFA, an MUFA-rich diet results in the postprandial formation of a smaller number of larger chylomicron particles and attenuation of postprandial FVII activation.

Chronic exposure to an MUFA-rich diet increases the capacity of large TRL particles to transport lipids during the postprandial phase while reducing the absolute number of TRL particles (Duttaroy, 2005). FVII binds to the protein moiety of TRLs, prolonging the length of its stay in the bloodstream. By producing a smaller number of large TRL particles, the MUFA-rich diet makes less FVII bind to TRL particles. In turn, this determines less activation of FVII. Furthermore, the large TRL particles rich in MUFAs or n-3 fatty acids are more easily cleared from the plasma than are particles rich in SFAs, because of their conformational structure (Sirtori et al., 1986). The SFAs are located primarily in position s2 of the triacylglycerols, which interchange with greater difficulty from the surface of the chylomicrons than do fatty acids located in position s1 and s3 of the acylglycerols, occupied preferentially by MUFAs and n-3 fatty acids (Karupaiah and Sundram, 2007). The combination of these two factors (i.e., the smaller number of large TRL particles and the shorter stay in the bloodstream of MUFA-rich particles compared with SFA-rich particles) may explain the lower concentration of FVII after an MUFA-rich diet as well as differences between MUFA- and SFA-rich diets in the postprandial phase.

Another suggested mechanism implicated in postprandial levels of FVII is the different affinity of fatty acids for peroxisome proliferator activated receptor α (PPAR α), a nuclear hormone receptor that has a critical role in regulating lipid metabolism. PPAR α activity could be part of a more complex mechanism responsible for the FVII activation, at least in the acute test meal situation.

Factor VII: High Carbohydrate Diet Affects Postprandial Factor VII

Studies that compared the effects of a high-fat meal with those of a high-carbohydrate meal in healthy subjects showed that both meals increase postprandial levels of FVII and impaired the antiplatelet functions of the endothelium. However, the postprandial absolute increase in FVIIa was significantly lower after the high-carbohydrate meal than after the high-fat meal. In the high-carbohydrate meal, FVIIa returned to normal values after an L-arginine intravenous infusion increased NO bioavailability; this was different from the high-fat

meal, in which the normalizing effect of L-arginine did not occur. Thus, increased NO availability from L-arginine or other substrates, such as vegetable inorganic nitrate, resets the activated hemostasis. This putative effect appears to be seriously impaired after a high-fat meal.

Platelets: Effects of Extravirgin Olive Oil, Monounsaturated Fatty Acids, and n-3 Polyunsaturated Fatty Acids

Platelets are small cell fragments produced by the breakdown of megakaryocytes, the large precursor cells found in bone marrow. Upon release into the circulation, they circulate for approximately 9–12 days. The normal platelet count varies from 140 to 400 $\times 10^9$ /L. As discussed, platelets are involved in primary hemostasis, with the formation of a *white* thrombus.

Dietary fats, which are expected to modify the composition of the platelet membrane, would affect their function. In principle, diets high in saturated fat, particularly long-chain fatty acids, are associated with a greater incidence of thrombosis compared with when a diet high in monounsaturated or polyunsaturated fat is eaten. Studies in vitro and in vivo showed that long-chain saturated fatty acids increase platelet aggregation, contrary to unsaturated fatty acids, which inhibit it.

Studies in humans largely demonstrated the antiplatelet aggregation effect of EVOO and MUFAs. In a study (Barradas et al., 1990), a 21-g/day supplementation of olive oil for 8 weeks reduced platelet aggregation induced by ADP and collagen. In a population study on young adults, consumption of an MUFA-rich diet, compared with an SFA-rich diet, resulted in a significant decrease in platelet aggregation in response to ADP, collagen, and arachidonic acid at 8 weeks; reduced platelet aggregation was maintained at 16 weeks (Smith et al., 2003). Despite the contrasting results of other studies, the weight of evidence suggests the significant beneficial effect of an MUFA-rich diet on platelet aggregation.

Studies in rabbits also showed that oleic acid is a potent inhibitor of PAF-induced platelet aggregation. PAF is a platelet agonist, a powerful endogenous mediator of platelet aggregation made of a mixture of unsaturated free fatty acids (FFAs). The PAF effect on platelets results from its interaction with a specific membrane receptor that induces the degradation of platelet plasma membrane phosphatidylinositol. Oleic acid has been demonstrated to induce the inhibition of phosphatidylinositol synthesis. The effect of oleic acid on platelet activation was attributed to the beneficial effects of oleic acid and other unsaturated FFAs in thrombotic disease prevention (Nunez et al., 1990; Delgado-Lista et al., 2008).

ω 3 (n-3) PUFA eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) extracted from fish oil were

demonstrated to reduce platelet aggregation actively. This effect results from several mechanisms: (i) competition with arachidonic acid, replacing active TxA_2 with TxA_3 ; (ii) inhibition of cyclooxygenase; and (iii) a direct antagonistic effect on TxA_2 -prostaglandin H_2 receptor in human platelets.

The EPA and DHA metabolic activity, however, is not confined to the platelet antiaggregation effects. Numerous studies and extensive reports showed the cardioprotective effects of dietary fish, fish oil, or a combination of EPA and DHA used in nutritional supplementation. Cardioprotection conferred by these PUFA species has been attributed to different mechanisms, including regulation of lipid metabolism and blood pressure, stabilization of atherosclerotic plaques, and antiarrhythmic and antiinflammatory actions. Individual gene variations in apolipoprotein (apo)AI, apoA5, apoE, tumor necrosis factor α ($\text{TNF}\alpha$), PPAR α , nitric oxide synthase 3 (NOS3), and arachidonate 5-lipoxygenase interact with n-3 PUFA intake modulating lipid metabolism and cardiovascular outcomes.

EPA and DHA also have a lowering effect on triglyceride blood levels. n-3 PUFA interacts with triglyceride metabolism modulated by NOS and PPAR α gene polymorphisms. Carriers of the minor allele for rs1799983 single nucleotide polymorphism (SNP) of NOS3 gene have shown a negative correlation between plasma TG concentrations and plasma n-3 PUFA levels. After n-3 PUFA supplementation, subjects with the minor allele had a better response to the change in plasma n-3 PUFA in reducing serum triglyceride concentration than did major allele homozygous carriers.

Concerning PPAR α gene polymorphisms, the minor allele Leu162Val variant was associated with higher triglycerides and apoC-III blood levels only in subjects consuming a low-PUFA diet. Conversely, high consumption of PUFAs with diet modulates the effect of this SNP on lipid metabolism and triglyceride blood levels.

Platelets: Polyphenols and Platelet Aggregation

Polyphenols are potent antioxidants present in several foods, particularly in fresh fruit, wine, and EVOO. The EVOO is particularly rich in phenolic compounds. The richness of EVOO in polyphenols depends on the mechanical procedures employed to obtain it, without the use of chemical solvents. Other olive oils (i.e., *refined* olive oils), obtained with solvents by pomace, contain no polyphenols because solvents do not allow any phenolic compound to be recovered by the pomace. The refining process serves to remove color, odor, and flavor from low-quality olive oil and leaves behind a pure form of olive oil that is tasteless, colorless, and odorless, with no bioactive compound. EVOO, on

the contrary, contains large amounts of polyphenols, which, however, are only a tenth of the polyphenols contained in olives; in fact, most of these olives' bioactive compounds are lost during the various processing stages, particularly in the mill wastewater.

The major phenolic compounds in EVOO are oleuropein, tyrosol, hydroxytyrosol, and luteolin. These phenols, which represent almost 90% of the EVOO polyphenols, are only four of the almost 30 phenolic compounds present in EVOO.

The polyphenols have been mostly investigated in CHD, in which they have demonstrated to reduce numerous cardiovascular risk factors significantly. EVOO polyphenols possess antioxidant properties and influence many biological activities that may account, at least partly, for the observed effects of olive oil on the cardiovascular system. Some of these effects include (i) inhibition of low-density lipoprotein oxidation, (ii) production of nitric oxide, and (iii) downregulation of intercellular adhesion molecule-1 and vascular cell adhesion molecule-1 expression in endothelial cells (Carluccio et al., 2003).

EVOO with a high phenolic compound content (400 mg/kg) has also been demonstrated to inhibit plasminogen activator inhibitor-1 and FVII significantly.

As far as platelet aggregation is concerned, EVOO polyphenols have been demonstrated to reduce platelet aggregability. The EVOO phenol component (2-[3,4-dihydroxyphenyl]-ethanol) significantly decreases platelet aggregation induced in vitro by ADP and collagen, and thromboxane B_2 (TxB_2) production by collagen and thrombin-stimulated platelet-rich plasma (Carluccio et al., 2003). Two isochroman polyphenols present in EVOO (1-[3'-cyclic adenosine monophosphate (cAMP)-methoxy-4'-hydroxy-phenyl]-6,7-dihydroxy-isochroman and 1-phenyl-6,7-dihydroxy-isochroman), inhibit in vitro platelet aggregation and thromboxane release induced by arachidonic acid and collagen (Togna et al., 2003). In vivo beneficial effects of EVOO polyphenols on hemostasis were also reported. In a study on volunteers, a significant reduction in plasma TxB_2 concentration after an EVOO-rich diet, compared with a high oleic-sunflower diet, was observed; this effect was attributed to the greater number of polyphenols in the EVOO-rich diet (Oubiña et al., 2001).

The molecular mechanisms of EVOO polyphenol antiplatelet activity were studied. Platelet activation is regulated by a number of physiological activators (TxA_2 , vasopressin, ADP, thrombin, and serotonin) and inhibitors (endothelium-derived relaxing factor and prostaglandin inhibitor-2). Platelet antagonists inhibit platelet function by increasing the intracellular levels of cyclic nucleotides cAMP and cyclic guanosine monophosphate by activating the respective cyclases. Cyclic nucleotide levels are downregulated by degradation

through phosphodiesterases (PDE). Platelets contain mainly PDE3, which preferentially hydrolyzes cAMP as substrate. EVOO phenols, particularly luteolin, were demonstrated to act on the cAMP-PDE complex by inhibiting PDE activity, increasing intracellular levels of cyclic nucleotide cAMP (Dell'Agli et al., 2008).

3,3',4',5,7-Pentamethylquercetin (PMQ), a member of the polymethoxylated flavone family, is the methylated form of quercetin, a flavonol polyphenol found in many fruits and vegetables, teas, grains, and olive oil. PMQ is a potent antioxidant with anticarcinogenic activity and cardioprotective properties. In vitro studies showed that PMQ exerts a potent inhibitory effect on platelet function; in vivo, it inhibits thrombus formation in an acute animal model. PMQ has been shown to protect mice from death in the acute lung thromboembolism model, and from carotid artery injury induced by ferric chloride. Moreover, PMQ inhibits platelet aggregation induced by several agonists and regulates the functional response of platelets, including the release of adenosine triphosphate and P-selectin. The molecular mechanisms of the inhibitory effects of PMQ on platelet function were shown to be the suppression of the PI3K/Akt-GSK3 β and Syk-PLC γ 2-Erk signaling cascades (Liang et al., 2015).

Finally, the peculiar combination of biologically active compounds in EVOO produces a milder activation of mechanisms of inflammation and coagulation during the postprandial phase, a situation that leads to reduced postprandial activation of nuclear factor- κ B, an important cellular regulator that initiates the formation of procoagulant and proinflammatory signal peptides (Nunez et al., 1990).

NUTRIGENETICS, NUTRIGENOMICS, AND COAGULATION SYSTEM: THE ROLE OF DIET AND ITS COMPONENTS

Polymorphisms of Factor VII Gene

Plasma FVII levels vary significantly in the general population; they are influenced by environmental and genetic factors. Studies linked the presence of mutations in FVII gene with plasma FVII levels and activity.

The FVII gene is 12.8 kb long; the mature FVII protein is encoded by exons 2–8. Several insertion-deletion and single-point mutations in the FVII gene were described in the promoter region (–323P0/10), within the transcribed space (R353Q point mutation), and within intronic sequences (IVS7). These FVII gene polymorphisms may have opposite effects on FVII levels and were shown to modulate the risk for myocardial infarction (MI) in males with advanced coronary artery disease.

R353Q is a polymorphism in the FVII gene closely associated with variations in plasma levels of FVII. It is

a single nucleotide polymorphism (point mutation in the transcribed region) characterized by the substitution of an arginine (R allele) with a glutamine (Q allele) in codon 353 of the transcribed protein. Lower postprandial levels of FVII associated with hypertriglyceridemia were found in carriers of the Q allele compared with individuals who were homozygous for the more common R allele. Compared with subjects with the R allele (Arg353), subjects with one or two Q alleles (Gln353) had levels of FVIIc that were 20% and 78% lower, respectively. FVII gene polymorphisms may explain up to about one-third of FVII-level variation in plasma. This indicates that to a major extent, plasma FVII levels are determined by genetic influences.

Other FVII gene polymorphisms have been described: –323A1/A2, –401G/T, and –402G/A.

–323A1/A2 is a polymorphism characterized by a 10–base pair insertion in the FVII promoter region at position –323. Carriers of the A2 allele of the 323A1/A2 polymorphism have been demonstrated to have a significant decrease in FVII circulating level (by 36% in males and 39% in females) and a lower risk for MI than noncarriers (odds ratio [OR] = 0.40; 95% confidence interval [CI] = 0.20–0.80).

–401G/T and –402G/A are two novel polymorphisms found in the promoter region of FVII gene. They are associated with altered plasma concentrations of FVII in vivo in population-based healthy human volunteers. Together, the –401G/T and –402G/A polymorphisms account for 18% and 28% of variation in the plasma concentrations of total FVII and fully activated factor FVII (FVIIa). These two polymorphisms affect the binding of two hepatic nuclear proteins to the promoter of the FVII gene, with accompanying changes of FVII expression in hepatocytes and FVII secretion by the liver. This, in turn, results in decreased (–401T allele) or increased (–402A allele) plasma levels of FVII, respectively. Male carriers of the –402A allele were associated with a significantly increased risk for MI (OR = 1.79; 95% CI = 1.15–2.80).

Among these FVII gene polymorphisms, R353Q SNP has been described as dominant over the other polymorphisms and as the major contributor to circulating levels of FVII.

Polymorphisms of Genes Involved in Homocysteine-Methionine Metabolism: Example of Gene–Environment Interaction

Elevated levels of homocysteine in the blood represent a potential risk factor for cerebrovascular diseases such as stroke and a predisposition to thrombosis and other pathological conditions. The reason for an elevated homocysteine level in the blood may be inadequate nutrition or genetic variants that reduce the

activity of enzymes necessary to remove homocysteine efficiently from the blood. A diet containing vitamins B12 and B6 and folic acid helps lower homocysteine level in the blood.

Several polymorphisms have been described in 5,10-methylene tetrahydrofolate reductase (MTHFR) and methionine synthase (MS) genes. Some of these polymorphisms were found to have a crucial role in the thrombotic risk associated with homocysteine blood level and nutritional status.

MTHFR is an enzyme involved in converting the amino acid homocysteine to another amino acid, methionine, providing the folate derivative for converting homocysteine to methionine. Another enzyme, MS, is responsible for regenerating methionine from homocysteine. Two polymorphisms of these genes are crucial for thrombotic risk:

1. The C677T polymorphism in exon 4 of the *MTHFR* gene, which leads to the amino acid substitution of alanine by valine at codon 222 (A222V), causing a thermolabile enzyme with lower activity; and
2. 2756 A → G of MS, which leads to amino acid substitution of aspartate by glycine.

These two polymorphisms are associated with (i) decreased activity of the enzyme MTHFR; (ii) a relative deficiency in the remethylation process of homocysteine into methionine; and (iii) mild to moderate hyperhomocysteinemia, a condition recognized as an independent risk factor for atherosclerosis and thrombosis. In studies on the interaction between these polymorphisms and B vitamin nutritional status in individuals affected by thromboembolic events, the genetic influence of the MTHFR polymorphism on homocysteine levels was not significant in individuals at high risk for thrombotic events who exhibited serum levels of folate and/or vitamin B12 above the 50th percentile of distribution in the general population.

Transcriptomic Effects of Adherence to Mediterranean Diet

The TCF7L2 gene has been shown to have the strongest association with type 2 diabetes. The effect of adhering to the Mediterranean diet (MeD) on the incidence of diabetes and cardiovascular risk in subjects with the TCF7L2 gene polymorphism was investigated in the PREDIMED study (Corella and Ordovas, 2014). The rs7903146 (C > T) SNP in intron 4 of the TCF7L2 gene was the most important one associated with type 2 diabetes. In the PREDIMED study, the strong association between risk allele (T) and the incidence of type 2 diabetes mellitus, high glucose, high cholesterol, high triglycerides, and high risk for cardiovascular disease was completely reversed in subjects with a high level

of adherence to the MeD. These effects appeared to result from a synergistic action of different components of the MeD, particularly MUFA, PUFA, and polyphenols. The intake of polyphenols was calculated as 1 g/day, 54% of which were flavonoids, 37% phenolic acids, and 9% other polyphenols. These high polyphenol intakes derived mostly from EVOO and fresh fruit.

When the transcriptomic profile in peripheral blood mononuclear cells was evaluated in the PREDIMED study, a weak thrombin signaling pathway modulation was found in subjects following the MeD supplemented with EVOO or nuts. The authors concluded that one mechanism by which MeD supplemented with EVOO can exert health benefits is through changes in the transcriptomic response of genes related to cardiovascular risk.

In a meta-analysis summarizing current knowledge concerning nutrigenomic studies on MeD and olive oil interventions, significant changes in proatherothrombotic (TF and TFP1), inflammation (TNF α and monocyte chemoattractant protein 1), and oxidative stress-related gene expression were found in subjects following a MeD pattern compared with control subjects (Konstantidou et al., 2013). TF and TF pathway inhibitor (TFPI) genes were downregulated and upregulated, respectively, by the MeD.

The TNF α gene acts as an activator of a cascade of inflammatory cytokine production; TNF α itself is considered to be a crucial proinflammatory cytokine. The TNF α gene is upregulated after consumption of a butter-rich meal, but not after an EVOO-rich meal or diet.

SUMMARY

Hemostasis is a complex phenomenon involving numerous coagulation factors. These factors circulate as inactive zymogens (inactive enzyme precursor) and act by cleaving downstream proteins so that they become active enzymes (Figs. 49.1 and 49.2).

The role of FVII in secondary hemostasis while forming a complex with TF is of primary importance. It was thought that the two pathways of coagulation cascade, the intrinsic (platelet) pathway and the extrinsic (FVII) pathway were equally important, but it is now known that the primary pathway for the initiation of blood coagulation is the TF-activated extrinsic pathway. FVII has a pivotal role in activating the common pathway of the clotting cascade and thrombin formation (Fig. 49.3).

The MeD, EVOO, and polyphenols have been shown to affect several coagulation factors.

The circulating level of FVII is deeply influenced by diet. There is a substantial increase in FVII circulating levels in the postprandial phase. The intake of dietary fat is the main determinant of the postprandial FVII plasma level. The ratio of SFAs to MUFAs is crucial to

postprandial levels of FVII. Diets rich in MUFA (i.e., olive oil, are associated with a significantly lower postprandial peak level of FVII and likely explain the lower rates of CHD in countries in which the diet is habitually rich in MUFAs, such as the Southern European countries.

Gene polymorphisms in the FVII promoter region modulate FVII circulating levels. Some of these polymorphisms are associated with lower levels of circulating FVII.

Also, platelet activity is influenced by diet. The n-3 PUFAs EPA and DHA actively reduce platelet aggregation. This effect results from several mechanisms (i) competition with arachidonic acid, replacing active TxA₂ with TxA₃; (ii) inhibition of cyclooxygenase; and

(iii) a direct antagonistic effect on the TxA₂–prostaglandin H₂ receptor in human platelets.

Polyphenols of EVOO, particularly luteolin, also reduce platelet aggregability, acting as an inhibitor of platelet PDE₃ (Table 49.1).

Probably the most exciting data are those concerning the effects of the MeD in modulating gene expression. Dietary interventions have been demonstrated to modulate the expression of pro-atherothrombotic and inflammation genes actively even in high-risk populations. SFAs upregulate both proinflammatory and proatherothrombotic genes, whereas the MeD, EVOO, and polyphenols downregulate the expression of these genes. The MeD, which is rich in olive oil, MUFAs, and polyphenols were demonstrated to exert a modulatory effect toward a protective mode on genes related to chronic degenerative diseases, oxidation, inflammation, and thrombosis (modulating the activity of TF, TFP1, and thrombin). The phenolic compounds present in EVOO appear to be responsible for the transcriptomic effects, as demonstrated in randomized, controlled human studies in which similar olive oils, but with different phenolic contents, were tested (Table 49.2).

TABLE 49.1 Protective Effects of Eicosapentaenoic Acid/Docosahexaenoic Acid on Cardiovascular Disease Risk Factors.

Effect	Proposed Mechanism
Reduced platelet aggregation	Reduction in prothrombotic prostanoids through competition with arachidonic acid
Antiarrhythmic effect	Modulation of electrophysiological properties of cardiac myocytes
Serum triglycerides reduction	Reduction in hepatic triglyceride production and lipoprotein assembly
Lowered blood pressure	Improved endothelial function, vascular relaxation, and arterial compliance
Decreasing inflammation	Reduced leukotriene production and signaling through competition with arachidonic acid and leukotriene receptor antagonism

TABLE 49.2 The most important phenolic compounds of the extra virgin olive oil (EVOO).

Phenolic Acids	Secoiridoids
Verbascoside	Oleuropein glucoside
Caffeic acid	Oleuropein aglycone
p-Hydroxybenzoic acid	Ligstroside
Protocatechuic acid	
Vanillic acid	
Coumaric acid	
Ferulic acid	
Anthocyanin	Simple Phenols
Cyanidin glucoside	Hydroxytyrosol
Cyanidin rutinoside	Tyrosol
Glucosyl rutinoside	
Delphinidin rhamnosyl	
	Flavanols
	Quercetin rutinoside
	Flavones
	Luteolin 5-glucoside
	apigenin 5-glucoside

References

- Barradas, M.A., Chirtofides, J.A., Jeremy, J.Y., Mikhailidis, D.P., Fry, D.E., Dandona, P., 1990. The effect of olive oil supplementation on human platelet function, serum cholesterol related variables and plasma fibrinogen concentration: a pilot study. *Nutr Res* 10, 403–411.
- Carluccio, M.A., Siculella, L., Ancora, M.A., Massaro, M., Scoditti, E., Storelli, C., Visioli, F., Distanti, A., De Caterina, R., 2003. Olive oil and red wine antioxidant polyphenols inhibit endothelial activation: antiatherogenic properties of Mediterranean diet phytochemicals. *Arterioscl Thromb Vasc Biol* 23, 622–629.
- Corella, D., Ordovas, J.M., 2014. How does Mediterranean diet promote cardiovascular health? Current progress toward molecular mechanism gene-diet interactions at the genomic, transcriptomic, and epigenomic levels provide novel insight into new mechanisms. *Bioessays* 36, 526–537.
- De Stavola, B.L., Meade, T.W., 2007. Long term effects of hemostatic variables of fatal coronary heart disease. 30-years results from the first prospective Northwick Park Heart Study (NPHS-I). *J Thromb Haemost* 5, 461–471.
- Delgado-Lista, J., Lopez-Miranda, J., Cortes, B., Perez-Martinez, P., Lozano, A., Gomez-Luna, R., Gomez, P., Gomez, M.J., Fuentes, F., Perez-Jimenez, F., 2008. Chronic dietary fat intake modifies the postprandial response of hemostatic markers to a single fatty test meal. *Am J Clin Nutr* 87, 317–322.
- Dell'Agli, M., Maschi, O., Galli, G.V., Fagnani, R., Dal Cero, E., Caruso, D., Bosisio, E., 2008. Inhibition of platelet function by olive oil phenols via cAMP-phosphodiesterase. *Br J Nutr* 99, 945–951.
- Duttaroy, A.K., 2005. Postprandial activation of hemostatic factors: role of dietary fatty acids. *Prostaglandins Leucot Essent Fatty Acids* 72, 381–391.
- Karupaiah, T., Sundram, K., 2007. Effects of stereospecific positioning of fatty acids in triacylglycerol structures in native and randomized fats: a review of their nutritional implications. *Nutr Metab* 4, 16.

- Konstantinidou, V., Covas, M.I., Sola, R., Fitò, M., 2013. Up-to date knowledge on the in vivo transcriptomic effect of the Mediterranean diet in humans. *Mol Nutr Food Res* 57, 772–783.
- Liang, M.L., Da, X.W., He, A.D., Yao, G.Q., Xie, W., Liu, G., Xiang, J.Z., Ming, Z.Y., 2015. Pentamethylquercetin (PMQ) reduces thrombus formation by inhibiting platelet function. *Sci Rep* 5, 11142.
- Nunez, D., Randon, J., Gandhi, C., Siafaka-Kapadai, A., Olson, M.S., Hanahan, D.J., 1990. The inhibition of platelet-activating factor-induced platelet activation by oleic acid is associated with a decrease in polyphosphoinositide metabolism. *J Biol Chem* 265, 18330–18338.
- Oubiña, P., Sánchez-Muniz, F.J., Ródenas, S., Cuesta, C., 2001. Eicosanoid production, Thrombogenic ratio, and serum LDL peroxides in Normo- and hypercholesterolaemic post-menopausal women consuming two oleic acid-rich diets with different content of minor components. *Br J Nutr* 85, 41–47.
- Roche, H.M., Zampelas, A., Knapper, J.M., Webb, D., Brooks, C., Jackson, K.G., Wright, J.W., Gould, B.J., Kafatos, A., Gibney, M.J., Williams, C.M., 1998. Effect of long-term olive oil dietary intervention on postprandial triacylglycerol and factor VII metabolism. *Am J Clin Nutr* 68, 552–560.
- Silva, K.D., Kelly, C.N., Jones, A.E., Smith, R.D., Wootton, S.A., Miller, G.J., Williams, C.M., 2003. Chylomicron particle size and number, factor VII activation and dietary monounsaturated fatty acids. *Atherosclerosis* 166, 73–84.
- Sirtori, C.R., Tremoli, E., Gatti, E., Montanari, G., Sirtori, M., Colli, S., Gianfranceschi, G., Maderna, P., Dentone, C.Z., et al., 1986. Controlled evaluation of fat intake in the Mediterranean diet: comparative activities of olive oil on plasma lipids and platelets in high-risk patients. *Am J Clin Nutr* 44, 635–642.
- Smith, R.D., Kelly, C.N., Fielding, B.A., Hauton, D., Silva, K.D., Nydahl, Miller, G.J., Williams, C.M., 2003. Long-term monounsaturated fatty acids reduce platelet aggregation in healthy young subjects. *Br J Nutr* 90, 597–606.
- Togna, G.I., Togna, A.R., Franconi, M., Marra, C., Guiso, M., 2003. Olive oil isochromans inhibit human platelet reactivity. *J Nutr* 23, 622–629.
- Zampelas, A., Roche, H., Knapper, J.M., Jackson, K.G., Tornaritis, M., Gibney, M.J., Kafatos, A., Gould, B.J., Wright, J., Williams, C.M., 1998. Differences in postprandial lipaemic response between Northern and Southern Europeans. *Atherosclerosis* 139, 83–93.

This page intentionally left blank

Diet–Gene Interactions: Perspectives in Insulin Resistance and Type 2 Diabetes

Jessica C. Ralston¹, Helen M. Roche^{1,2}

¹Nutrigenomics Research Group, UCD Conway Institute of Biomolecular Research and School of Public Health, Physiotherapy, and Sports Science, University College Dublin, Dublin, Republic of Ireland; ²Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, Belfast, Northern Ireland, United Kingdom

Glossary

Adipogenesis Differentiation of preadipocytes into mature adipocytes (fat cells)
Adipokines Secreted proteins from adipose tissue; adipocytokines
AMPK Adenosine monophosphate-activated protein kinase
BAT Brown adipose tissue
C3 Complement component 3
GOLDN Genetics of Lipid Lowering Drugs and Diet Network
GWAS Genome-wide association studies
HFD High-fat diet
Hyperplasia Increase in cell number
Hypertrophy Increase in cell size
IL-1 β Interleukin 1 β ; proinflammatory cytokine
IL-6 Interleukin 6
LTA Lymphotoxin- α
MUFA Monounsaturated fatty acid
Phenotypic switch Antiinflammatory M2 macrophages become proinflammatory M1 macrophages
SAT Subcutaneous adipose tissue; located underneath the skin outside of the abdominal cavity
SCD1 Stearoyl-coenzyme A desaturase 1
SFA Saturated fatty acid
SNP Single nucleotide polymorphism
SVF Stromal vascular fraction
T2D Type 2 diabetes
TAG Triacylglycerol
TLR4 Toll-like receptor 4
TNF α Tumor necrosis factor α
VAT Visceral adipose tissue; located within the abdominal cavity
WAT White adipose tissue

PROGRESSIVE PHENOTYPE LEADING TO TYPE 2 DIABETES

Within the context of nutrigenetics and nutrigenomics pertaining to type 2 diabetes (T2D), it is important to appreciate that T2D represents a progressive

phenotype. Thus, the initial triggers and molecular processes that promote the pathogenesis of this condition are highly important. Fig. 50.1 illustrates how the progressive phenotype is initiated by obesity with insulin resistance, which further develops, with multiple metabolic organ dysfunction, to culminate in T2D. Therefore, the primary focus of this chapter is on reviewing the evidence with respect to genetic versus environmental determinants of T2D and the molecular processes that underlie metabolic organ dysfunction, with particular focus on adipose tissue. We will pay special attention to the molecular mechanisms in which nutritional states and stresses can adversely affect adipose tissue function to promote progression toward T2D.

GENES VERSUS ENVIRONMENT AS CAUSAL DETERMINANTS OF TYPE 2 DIABETES

Genetic factors have an important role in the onset of T2D. Several elegant genome-wide association studies (GWAS) have identified more than 130 genetic variants associated with T2D, β -cell function, glucose regulation, or insulin action (Morris et al., 2012; Gaulton et al., 2015). However, these variants explain less than 15% of disease heritability. Also, it has been difficult to infer biological mechanisms from the genetic variants that would explain the pathogenesis and progression of T2D. A real challenge has been defining the impact of genetics plus the environment, such as processes leading from energy excess to obesity and from that to the risk for T2D. GWAS approaches have also identified important genetic determinants of obesity, adipose biology, and

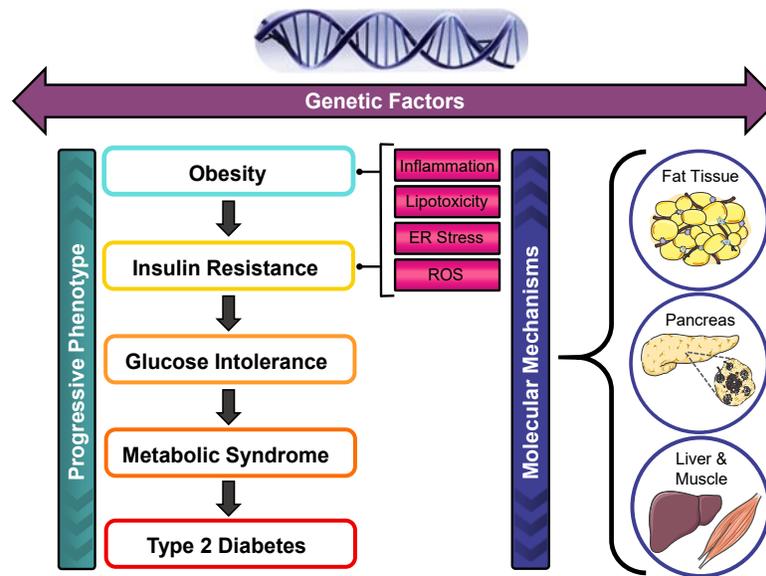


FIGURE 50.1 Progressive phenotype leading to type 2 diabetes (T2D). ER, endoplasmic reticulum; ROS, reactive oxygen species. Adapted from Phillips, C.M., Tierney, A.C., Roche, H.M., 2008. Gene–nutrient interactions in the metabolic syndrome. *J Nutrigenetics Nutrigenomics* 1(3), 136–151.

insulin resistance (Shungin et al., 2015). Given the causal interrelationship between obesity and adipose biology with pancreatic dysfunction, these also precipitate risk toward T2D. Key targets included genetic variants involved in processes relating to synaptic function, glutamate signaling, insulin secretion and insulin action, energy metabolism, lipid biology, and adipogenesis. Nevertheless, it needs to be acknowledged that the increasing prevalence of T2D is outpacing genetic variation. Thus, the impact of environmental factors, which no doubt interact with genetic susceptibility, have a key role in the pathogenesis and progression of T2D. Common environmental factors include diet (excess energy, saturated fatty acids, simple sugars, etc.), physical inactivity, metabolic/endocrine disruptors, chronobiology, the gut microbiome, and so on, some of which will be discussed further in terms of molecular mechanisms later in the chapter.

GENE–NUTRIENT INTERACTIONS VERSUS METABOTYPES RELATING TO TYPE 2 DIABETES

The field of gene–nutrient interactions evolved based on the fundamental question of interindividual variation in response to different diets. This research area is based on the premise that the impact of a protective or risk genotype may be different, either enhanced or negated, according to the dietary environment and/or diet-derived biomarkers (Smith and Ordovás, 2010). For example, our work has shown that different

inflammatory genotypes can be further enhanced by dietary saturated fatty acids to augment risk associated with phenotypic traits associated with T2D (Phillips et al., 2010). Several other groups completed comprehensive gene–nutrient interaction studies (Ortega et al., 2017; Bernstein et al., 2017). Gene–nutrient research will continue as GWAS uncover new genetic variants associated with T2D. This will further define mechanistic relations between genetic and nutritional factors. However, there are limitations to this approach. False-positive results are a common potential pitfall owing to multiple testing of metabolic outcomes or variants. Therefore, in addition to defining new variants from GWAS, continuing efforts must focus on confirming previously reported genetic associations and interactions in different population groups. Moreover, given the significant influence of epigenetic inheritance (i.e., the passing of changes in gene regulation to offspring) on T2D heritability, future studies will need to consider these additional aspects as the field advances (Ortega et al., 2017; Bernstein et al., 2017).

Defining the role of gene–nutrient interactions in the risk for T2D seeks to develop more effective personalized nutrition approaches, by which a nutrigenetic-based framework would allow for tailoring diets and develop preventive nutritional strategies to minimize disease risk. The extent to which lifestyle factors can offset genetic risk is largely unknown, particularly relating to T2D. However, a large GWAS and lifestyle assessment study demonstrated that a favorable lifestyle, which included dietary components (as well as smoking, lack of physical activity, and obesity) was

associated with nearly a 50% reduction in the relative risk for cardiovascular disease in individuals at high risk (Khera et al., 2016). More recent approaches have progressed the personalized nutrition approach, wherein a more comprehensive metabolome (a combined metabolic, genetic, transcriptomic, proteomic, and/or metabolomic phenotype) may be a comprehensive tool to define the interaction between dietary components and the human genome with a view to understanding the impact of diet and genes on the development of T2D, as well as developing personalized nutrition approaches (Brennan, 2017). For example, such approaches revealed the potential impact of previously unknown factors, such as the gut microbiome or branch chain amino acid metabolism, on the risk for T2D (Newgard, 2012; Zeevi et al., 2015). In addition to the genetic determinants of T2D, nutrigenomics has helped us advance our understanding in relation to how diet affects gene and protein expression in key metabolic organs. The remainder of this chapter focuses on these complementary molecular processes.

NUTRIGENOMIC INSIGHTS RELATING TO DIET-INDUCED MOLECULAR PROCESSES IN ADIPOSE TISSUE DYSFUNCTION THAT PROMOTE TYPE 2 DIABETES

Dietary components have the ability to influence gene expression in numerous metabolic tissues throughout the body, which can ultimately affect the progression of T2D (Berná et al., 2014). This is partly because of the capacity of certain nutrients and dietary components to act as ligands for transcription factors, thereby inducing changes in the expression of genes implicated in the development of T2D. Certain nutrients and dietary factors can also alter the abundance of precursors, intermediates, or products in cellular signaling pathways responsible for altering gene expression, in addition to directly or indirectly changing the activity of metabolic pathways. These interactions between nutrients and gene expression have the potential to increase T2D risk as well as the development of numerous disease complications (Berná et al., 2014).

In particular, diet–gene interactions in adipose tissue (fat) can have key roles in the progression of T2D; the remaining sections of the chapter will explore this tissue in more detail. Indeed, the liver, pancreas, and skeletal muscle are also important metabolic tissues in the context of T2D development. Several excellent reviews

(Ralston et al., 2017; Lyons et al., 2016; Bleau et al., 2015) have addressed this comprehensively.

ADIPOSE TISSUE

Two types of adipose tissue are found in mammals: brown adipose tissue (BAT) and white adipose tissue (WAT). WAT is the main storage sink for excess energy which predominantly exists as triacylglycerol in adipocytes. In addition to its important role as a passive storage tissue, WAT is incredibly dynamic as an endocrine organ. Through the secretion of WAT-derived proteins (adipocytokines [adipokines]), WAT can dramatically influence tissue-to-tissue communication, gene expression, and whole-body metabolism. Interestingly, the functions of WAT in relation to weight gain and disease progression may differ by anatomical location. Subcutaneous WAT (located underneath the skin outside the abdominal cavity) has been proposed to act as a protective sink for extra calories in the early stages of WAT expansion and may be linked to a more insulin-sensitive profile in relation to T2D risk (Virtue and Vidal-Puig, 2008). On the other hand, visceral WAT (located within the abdominal cavity) has been associated with increased insulin resistance and is often described as being more metabolically active (partially owing to drainage of proinflammatory cytokines into the portal vein) (Virtue and Vidal-Puig, 2008). Furthermore, all WAT is composed of two distinct fractions: (1) adipocyte, and (2) stromal vascular fraction (SVF), which includes preadipocytes, macrophages, mesenchymal stem cells, and endothelial cells. Both of these cell populations and fractions can have a role in the relationships among diet, genes, and the risk for T2D (Ralston et al., 2017).

When caloric intake consistently exceeds demand, WAT depots expand and eventually promote the progression of disease states (e.g., obesity and T2D). The process of WAT expansion occurs through coordinated changes in gene expression, which result in both adipocyte hypertrophy (increase in cell size) and hyperplasia (increase in cell number) (Fig. 50.2). Generally speaking, hyperplastic WAT expansion is preferable, because hypertrophic WAT tends to be proinflammatory, insulin resistant, and metabolically dysfunctional (Virtue and Vidal-Puig, 2008). Therefore, hypertrophic WAT has a greater propensity to promote T2D-related complications. However, different diets and nutrients (e.g., saturated versus unsaturated fat) can support differing types of WAT expansion, changing the propensity for complications such as T2D. For instance, a diet high in monounsaturated fat (monounsaturated fatty acids [MUFAs]) seems to promote

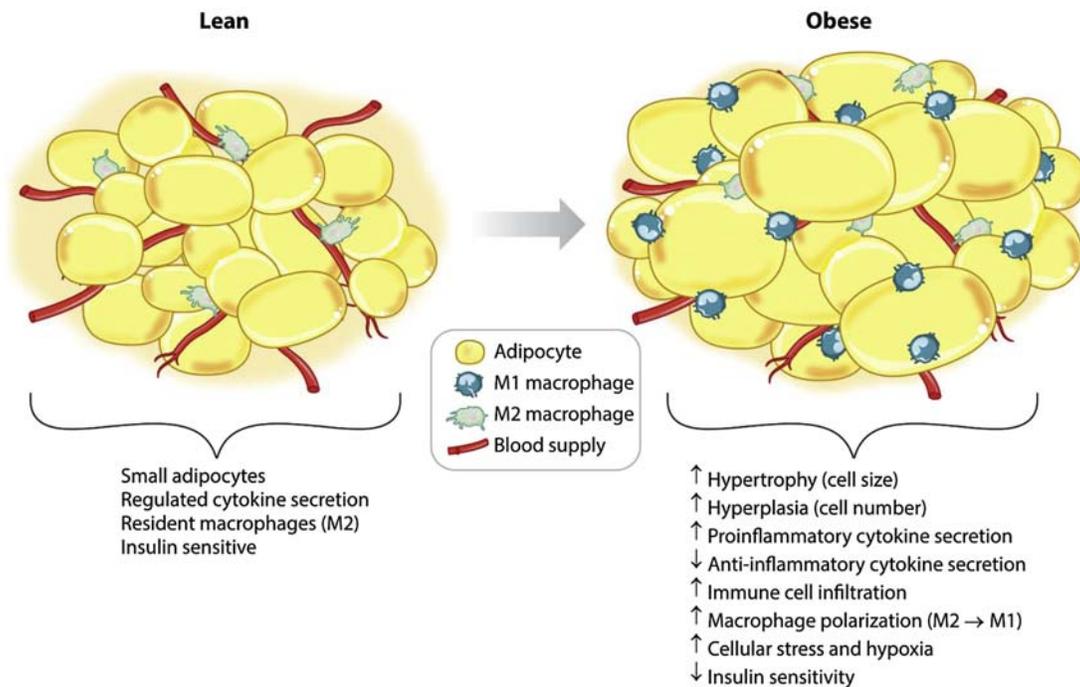


FIGURE 50.2 Expansion of white adipose tissue (WAT) in obesity, which promotes type 2 diabetes (T2D). *Published in Ralston, J.C., et al., 2017. Fatty acids and NLRP3 inflammasome-mediated inflammation in metabolic tissues. Annu Rev Nutr 37(1), 77–102.*

more hyperplastic WAT expansion compared with a diet high in saturated fat (saturated fatty acids [SFAs]) (Finucane et al., 2015). Compared with dietary SFA-induced WAT expansion, MUFA-induced expansion seems to coincide with reduced inflammatory gene expression, less priming of proinflammatory pro-interleukin (IL)-1 β , and increased adipogenic gene expression (i.e., genes that promote differentiation and healthy growth of adipocytes) (Finucane et al., 2015). Importantly, these changes in WAT gene expression on an MUFA-enriched diet occurred alongside attenuations in insulin resistance and hyperinsulinemia.

In addition to phenotypic changes in adipocytes (i.e., hyperplasia and hypertrophy), the expansion of WAT in obesity involves the infiltration of immune cells (e.g., macrophages, dendritic cells), which can further propagate inflammation. Furthermore, a phenotypic switch also occurs in the macrophages of obesogenic WAT, in which antiinflammatory M2 macrophages become proinflammatory M1 macrophages (Fig. 50.2). Moreover, SFAs seem to exacerbate these metabolic processes (Ralston et al., 2017). Together, it appears that nutrigenomic influences can significantly affect WAT functionality, which in turn can have ramifications for the progression of T2D.

The relevance of SFAs versus MUFAs in the study of obesity and T2D stems from the fact that high-fat diets are typically derived from SFAs (e.g., palmitate). It has been well-established that SFAs have a detrimental role in inflammation, obesity, and T2D (Ralston et al., 2017). Moreover, evidence in humans with respect to the

negative effects of SFAs on inflammation and insulin resistance is well-established (Frouhi et al., 2014). This is partly because SFAs can activate Toll-like receptor (TLR)4 signaling, which subsequently results in increased nuclear factor- κ B activity, a transcription factor controlling the gene expression of several proinflammatory signaling molecules as a master switch. In turn, numerous other proinflammatory pathways are upregulated by SFAs and promote the release of inflammatory cytokines (e.g., tumor necrosis factor α [TNF α]) in addition to the impairment of normal metabolic signaling. Conversely, it has been suggested that a diet high in MUFAs may be less proinflammatory compared with SFA-enriched diets (Finucane et al., 2015). Evidence from Brehm et al. (2009) also demonstrated in an intervention study that an MUFA-enriched diet in T2D could reduce markers of insulin resistance. However, weight loss was also a factor in that study, and therefore additional work is needed. Nonetheless, the Mediterranean diet, which is high in MUFAs (derived from olive oil), has been associated with reduced hypertriglyceridemia and inflammation, as well as improved lipids and markers of the metabolic syndrome (Chrysohoou et al., 2004). One specific MUFA, oleate, has also specifically been shown to inhibit SFA-induced IL-1 β secretion from macrophages (Bradley et al., 2008). The mechanisms through which MUFAs exert beneficial effects have not been fully elucidated; however, evidence suggests that MUFAs may maintain adenosine monophosphate-activated protein kinase (AMPK) activation (Finucane

et al., 2015). AMPK is a key regulatory hub known to control cellular energy homeostasis and influence metabolic inflammation in addition to its association with T2D (Ralston et al., 2017; Lyons et al., 2016). It is therefore likely that in obese and/or diabetic adipose tissue, a lack of AMPK hinders the ability of the tissue to adapt to nutritional or environmental challenges.

Diet has an important role in altering gene expression and the risk for T2D. However, it is now known that genetic diversity can also influence this relationship significantly. For instance, single nucleotide polymorphisms (SNPs) in key genes involved in nutrient metabolism can promote or reduce the activity of certain genes (e.g., higher or lower converters or metabolizers for particular nutrients). For example, SFAs, which we have described as having detrimental roles in inflammation and T2D, can be converted to MUFAs (i.e., less inflammatory fatty acids) through stearoyl-coenzyme A desaturase 1 (SCD1). Previous work showed that certain genotypes for SCD1 SNPs are associated with improved insulin sensitivity, as well as altered body fat distribution (e.g., reduced waist circumference) (Warensjö et al., 2007). Moreover, genetic variation may augment the impact of inflammation and/or cytokine signaling (Stryjecki and Mutch, 2011), which can affect adipose tissue and pancreatic function and, again, alter progression to T2D. For example, the Genetics of Lipid Lowering Drugs and Diet Network (GOLDN) study demonstrated important interactions between IL-1 β genetic variants and polyunsaturated fats to decrease the risk for the metabolic syndrome (Shen et al., 2007). A European Union program, LIPGENE, specifically addressed the interaction between inflammation and dietary fat composition and showed potential effects of several inflammatory components. Complement component 3 (C3) is an innate immune biomarker with distinct sensitivity to SFAs. High dietary SFA intake further accentuates the inflammatory impact of both circulating C3 and C3 genotype to augment insulin resistance (Phillips et al., 2012). In addition, the combination of IL-6, lymphotoxin- α , and TNF α inflammatory risk genotypes interacted with plasma fatty acid status to increase the risk for metabolic syndrome (Phillips et al., 2010). Nevertheless, it must be acknowledged that the genetic determinants of T2D linked to pancreatic function are stronger candidates. Within this context, more work is needed with respect to their interactions with diet.

CONCLUSIONS

Genetic predisposition is important to the pathogenesis of T2D but also to the dietary environment. This relationship exists within the context of genetic determinants and molecular processes in which dietary

components modulate gene expression and metabolism to alter the risk for T2D.

References

- Berná, G., et al., 2014. Nutrigenetics and nutrigenomics insights into diabetes etiopathogenesis. *Nutrients* 6 (11), 5338–5369.
- Bernstein, D., Golson, M.L., Kaestner, K.H., 2017. Epigenetic control of β -cell function and failure. *Diabetes Res Clin Pract* 123 (Suppl. C), 24–36.
- Bleau, C., et al., 2015. Crosstalk between intestinal microbiota, adipose tissue and skeletal muscle as an early event in systemic low-grade inflammation and the development of obesity and diabetes. *Diabetes Metabol Res Rev* 31 (6), 545–561.
- Bradley, R.L., Fisher, F.M., Maratos-Flier, E., 2008. Dietary fatty acids differentially regulate production of TNF- α and IL-10 by Murine 3T3-L1 adipocytes. *Obesity* 16 (5), 938–944.
- Brehm, B.J., Lattin, B.L., Summer, S.S., Boback, J.A., Gilchrist, G.M., Jandacek, R.J., D'Alessio, D.A., 2009. One-year comparison of a high-monounsaturated fat diet with a high-carbohydrate diet in Type 2 diabetes. *Diabetes Care* 32 (2), 215–220. Pubmed ID: 18957534.
- Brennan, L., 2017. Use of metabolotyping for optimal nutrition. *Curr Opin Biotechnol* 44 (Suppl. C), 35–38.
- Chrysohoou, C., et al., 2004. Adherence to the Mediterranean diet attenuates inflammation and coagulation process in healthy adults: the Attica study. *J Am Coll Cardiol* 44 (1), 152–158.
- Finucane, O.M., et al., 2015. Monounsaturated fatty acid-enriched high-fat diets impede adipose NLRP3 inflammasome-mediated IL-1 β secretion and insulin resistance despite obesity. *Diabetes* 64 (6), 2116–2128.
- Forouhi, N.G., et al., 2014. Differences in the prospective association between individual plasma phospholipid saturated fatty acids and incident type 2 diabetes: the EPIC-InterAct case-cohort study. *Lancet Diabetes Endocrinol* 2 (10), 810–818.
- Gaulton, K.J., et al., 2015. Genetic fine-mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nat Genet* 47 (12), 1415–1425.
- Khera, A.V., et al., 2016. Genetic risk, adherence to a healthy lifestyle, and coronary disease. *N Engl J Med* 375 (24), 2349–2358.
- Lyons, C.L., Kennedy, E.B., Roche, H.M., 2016. Metabolic inflammation-differential modulation by dietary constituents. *Nutrients* 8 (5), 1–21.
- Morris, A.P., et al., 2012. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 44 (9), 981–990.
- Newgard, C.B., 2012. Interplay between lipids and branched-chain amino acids in development of insulin resistance. *Cell Metabolism* 15 (5), 606–614.
- Ortega, Á., et al., 2017. Gene-diet interactions in type 2 diabetes: the chicken and egg debate. *Int J Mol Sci* 18 (6), 1188.
- Phillips, C.M., Tierney, A.C., Roche, H.M., 2008. Gene-nutrient interactions in the metabolic syndrome. *J Nutrigenetics Nutrigenomics* 1 (3), 136–151.
- Phillips, C.M., et al., 2010. Additive effect of polymorphisms in the IL-6, LTA, and TNF- α genes and plasma fatty acid level modulate risk for the metabolic syndrome and its components. *J Clin Endocrinol Metab* 95 (3), 1386–1394.
- Phillips, C.M., et al., 2012. Dietary fat, abdominal obesity and smoking modulate the relationship between plasma complement component 3 concentrations and metabolic syndrome risk. *Atherosclerosis* 220 (2), 513–519.

- Ralston, J.C., et al., 2017. Fatty acids and NLRP3 inflammasome–mediated inflammation in metabolic tissues. *Annu Rev Nutr* 37 (1), 77–102.
- Shen, J., et al., 2007. Interleukin1 β genetic polymorphisms interact with polyunsaturated fatty acids to modulate risk of the metabolic syndrome. *J Nutr* 137 (8), 1846–1851.
- Shungin, D., et al., 2015. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 518 (7538), 187–196.
- Smith, C.E., Ordovás, J.M., 2010. Fatty acid interactions with genetic polymorphisms for cardiovascular disease. *Curr Opin Clin Nutr Metab Care* 13 (2), 139–144.
- Stryjecki, C., Mutch, D.M., 2011. Fatty acid-gene interactions, adipokines and obesity. *Eur J Clin Nutr* 65 (3), 285–297.
- Virtue, S., Vidal-Puig, A., 2008. It's not how fat you are, it's what you do with it that counts. *PLoS Biol* 6 (9), e237.
- Warensjö, E., et al., 2007. Polymorphisms in the SCD1 gene: associations with body fat distribution and insulin sensitivity. *Obesity* 15 (7), 1732–1740.
- Zeevi, D., et al., 2015. Personalized nutrition by prediction of glycemic responses. *Cell* 163 (5), 1079–1094.

Nutrigenetics of Bone Health

José-Luis Pérez-Castrillon¹, Marta Ruiz-Mambrilla², José A. Riancho³

¹Center of Investigation of Endocrinology and Nutrition, Medical School and Department of Internal Medicine, Hospital Universitario Río Hortega, University of Valladolid, Valladolid, Spain; ²Medical Center of Rehabilitation and Speech Disorders and School of Medicine, University of Valladolid, Valladolid, Spain; ³Department of Medicine, School of Medicine, Department of Internal Medicine, Hospital Universitario Marques de Valdecilla University of Cantabria, IDIVAL, Santander, Spain

Abbreviations

ApoE Apolipoprotein E
ATP Adenosine triphosphate
BMD Bone mineral density
CBG Intracellular β -glucosidase
FGF23 Fibroblast growth factor 23
LDL Low-density lipoprotein
LRP5 Low-density lipoprotein receptor-related protein 5
MTHFR Methylene tetrahydrofolate reductase
PON1 Paraoxonase 1
PTH Parathyroid hormone
PUFA Polyunsaturated fatty acid
SNP Single nucleotide polymorphism
VDR Vitamin D receptor

SENTENCE

Bone mass, the main determinant of bone health, is genetically determined, with an interaction between key nutrients and vitamin D receptor polymorphisms.

INTRODUCTION

A correct bone health status is accomplished through two mechanisms, modeling and remodeling, which are differently involved at different stages of life. Modeling is the process by which bone tissue is formed and constantly renewed until the end of growth through two processes: membranous ossification, in which embryonic mesenchymal cells differentiate into osteoblasts, and endochondral ossification, in which mesenchymal cells subsequently differentiate into cartilage and bone. Bone remodeling is a physiological process whose functions are the permanent renewal of

the skeleton to ensure biomechanically correct bone and provide the minerals (calcium, phosphorus, magnesium) required for the body needs.

Bone remodeling comprises an initial phase of osteoclast-mediated bone resorption, followed by a phase of formation through the osteoblasts. Both phases are regulated by general, endocrine factors, and local, paracrine factors. Of the endocrine factors, the calcitropic hormones (parathyroid hormone [PTH], also called parathormone or parathyrin; and vitamin D), and the sex hormones, especially estrogen and, to a lesser extent, androgen, play a determining role. Other hormones, such as thyroid hormones, growth hormone, and leptin play a minor role (Zaidi et al., 2018). The local factors include various cytokines and growth factors that regulate the process; a key role is played by the inflammatory cytokines, interleukin (IL)-1, IL-6 and tumor necrosis factor (TNF)- α (Tella and Gallagher, 2014). There is a final common pathway in this process, the RANK/RANKL/OPG system, which is involved in regulating bone resorption.

There is a balance between bone formation and resorption until the third decade of life, with the same amount of bone destroyed by osteoclasts as is formed by osteoblasts. From then on, osteoclast-mediated bone resorption predominates, and increases in the perimenopausal and early menopausal period, coinciding with the fall in estrogen levels. Bone remodeling is genetically regulated, with between 50% and 85% of bone mass being determined by genetic factors. Numerous genes involved in maintaining correct bone health have been reported, with the influence of each being relatively small and with results obtained from different, heterogeneous populations. These differences may be determined by the interaction with environmental factors,

with a key role being played by nutrition. In this scenario, calcium and vitamin D are essential elements in the development and maintenance of bone mass.

VITAMIN D AND CALCIUM

Interaction Between Calcium and the Vitamin D Receptor Genotype

Calcium is the most abundant mineral in the human body. An adult body contains 1–1.5 kg of calcium, mostly in the bones and teeth. Calcium participates in numerous metabolic functions that are essential for a correct body functioning. Calcium regulation is determined by the interaction between two hormonal systems: PTH and vitamin D. Dietary calcium is predominantly absorbed in the duodenum through two processes, one passive, through the paracellular pathway, and one active, dependent on 1,25 dihydroxycholecalciferol or calcitriol (Bouillon et al., 2018). The duodenum has the greatest concentration of vitamin D receptors (VDRs). The binding of calcitriol to VDRs results in the synthesis of a protein, calbindin D-28K. Calcium passes into the duodenal cells, where it binds with calbindin and is transported to the basolateral membrane and from there into the bloodstream through an active adenosine triphosphate (ATP) ase-dependent process (Fig. 51.1).

Vitamin D is a hormone derived from a steroids group backbone. It is synthesized endogenously from 7-dehydrocholesterol (provitamin D₃), which is converted into 7-dihydrocholecalciferol in the skin by ultraviolet light. Vitamin D is also derived from the diet of animal origin (D₃ cholecalciferol) or vegetable origin

(D₂ ergocalciferol). 7-dihydrocholecalciferol binds with a transporter protein (D-binding protein or vitamin D-transporting protein, DBP) and reaches the liver, where it is metabolized by 25-hydroxylase (microsomal CYP2R1, mitochondrial CYP27R1) into 25-hydroxycholecalciferol (25OHD₃). This metabolite is the element by which the levels of vitamin D in the body are measured, due to its long half-life of 3 weeks. 25-hydroxycholecalciferol binds with DBP and is transported to the kidney. Megalin is a receptor that mediates DBP-25OHD absorption in the proximal tubule, via endocytosis. This process is facilitated by cubilin. In the kidney it is metabolized to 1,25-dihydroxycholecalciferol by α 1-hydroxylase (CYP27B1), or to 24,25-dihydroxycholecalciferol by 24-hydroxylase (CYP24A1). The active metabolite of vitamin D is 1,25-dihydroxycholecalciferol, the synthesis of which is stimulated by PTH and inhibited by fibroblast growth factor 23 (FGF23). A mirror situation occurs with 24,25-dihydroxycholecalciferol, whose production is inhibited by PTH production and stimulated by FGF23 (Deluca, 2014).

Vitamin D exerts its effects by binding to the VDR, a nuclear receptor that regulates blood levels of calcium and phosphorus by increasing intestinal absorption, reducing tubular reabsorption, and inhibiting PTH secretion. VDR, a transcription factor, is widely distributed throughout the body. Characteristically, it is located in the nucleus, even in the absence of the ligand. It is encoded by a gene located on the long arm of chromosome 12, and comprises a region of approximately 100 kb of DNA, although only 4.6 Kb encode the protein. At this locus, more than 900 allelic variants have been described. Early studies used restriction enzymes for

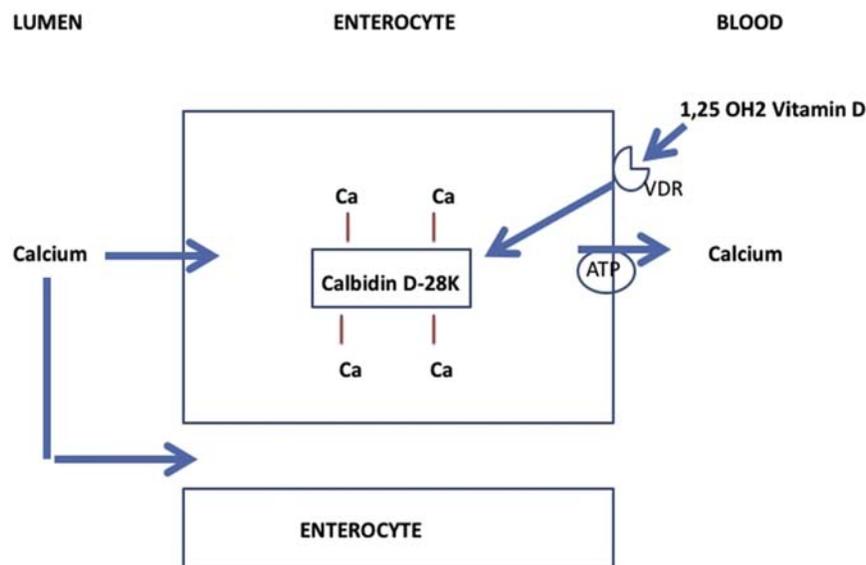


FIGURE 51.1 Vitamin D and intestinal absorption of calcium.

analysis. The most studied variants have been those determined by the Apal (rs7975232), BsmI (rs1544410), TaqI (rs731236), and FokI (rs10735810) (Uitterlinden et al., 2004) enzymes.

Most of these polymorphisms do not result in changes in the protein sequence, but can influence the stability of the receptor mRNA. The only exception is FokI, whose restriction site is located on exon 2 of the 5' coding region. This polymorphism results in a shorter protein with three fewer amino acids, which increases the stability of the receptor and has greater activity.

Studies in populations of different ages have assessed the effect of calcium supplementation on intestinal absorption and bone mass according to VDR polymorphisms. Ferrari et al. (1998a) examined the interaction between calcium intake, bone mass, and VDR polymorphisms and found that, in prepubescent girls, homozygous FF subjects gained more bone mass than homozygous ff subjects. This interaction was not observed in fertile women or men. Analysis of the BsmI polymorphism in prepubescent girls showed a better BMD response in heterozygous Bb subjects and a nonsignificant trend in homozygous BB subjects, without any effect in homozygous bb subjects, who spontaneously presented higher values of bone mass without responding to calcium supplements (Ferrari et al., 1998b). In males with the BsmI polymorphism, no differences in bone mass were observed, but there were differences in phosphorus levels. With a diet low in calcium and phosphorus, homozygous BB subjects showed a decrease in tubular reabsorption of phosphorus and associated hypophosphatemia (Ferrari et al., 1999). In a population of young men and women, the effect of calcium supplements on bone mass measured by ultrasound according to two polymorphisms, FokI and Apal, was analyzed, but no association was found (Correa-Rodriguez et al., 2015).

Another way to assess the interaction is based on the fraction of intestinal calcium absorption. The relationship between calcium intake, VDR BsmI polymorphisms, and calcium absorption was studied in 60 females. In subjects with low-calcium intake, there was greater absorption in homozygous bb subjects than in homozygous BB individuals. However, in individuals with high-calcium intake (800–1000 mg/24 h), no differences in absorption were observed according to the genotype (Dawson-Hughes et al., 1995). The hypothesis was that in high-intake subjects, the absorption pathway differed from that mediated by the VDR, and was probably the paracellular pathway of passive diffusion. Similar results were obtained in obese women, in whom the BB genotype was associated with reduced absorption calcium in women with low-dietary-calcium intake. The authors assessed the haplotypes of three polymorphisms—BsmI, Apal, and TaqI—and found

that BsmI modified calcium absorption in women with a lower calcium intake (Chang et al., 2015). The mechanism through which the BsmI polymorphism and this haplotype modify intestinal calcium absorption is unknown. It is speculated that BsmI alters the transcription and stability of the receptor and may reduce the number of VDR in the duodenum. Gennari et al. (1997) found lower intestinal calcium absorption in European subjects with the BB genotype (single nucleotide polymorphism [SNP] BsmI) and tt genotype (SNP TaqI).

Studies have analyzed the interaction between VDR polymorphisms and calcium intake. Likewise, the interaction with polymorphisms of the Wnt pathway has been evaluated. The effect appears only with low-calcium intake; a high-calcium intake may be governed by a different, non-VDR, absorption pathway. Thus, a polymorphism associated with low-density lipoprotein receptor-related protein 5 (LRP5), a coreceptor of the Wnt pathway, has been associated with calcium absorption. The SNP evaluated was rs4988321, and it was observed that carriers of the A allele, when compared with homozygous GG subjects, had a lower bone mass, although this only occurred in women with low-calcium intake, while individuals with high-calcium intakes showed no genotype-related differences (Stathopoulou et al., 2010).

Interaction Between Vitamin D and the Vitamin D Receptor Genotype

Studies assessing the nutrigenetics of vitamin D supplements according to the VDR have heterogeneous designs. Studies are few, with varying initial levels of vitamin D, and they analyze different genotypes. A large number of studies have assessed the effect of VDR genotypes on bone mass and fractures.

Barry et al. (2014) studied the effect of calcium (1000 mg/day) and vitamin D (1000 IU/day) supplements on vitamin D levels according to different SNPs in 1887 white, non-Hispanic males and females and a placebo group. The follow-up was 1 year, and therefore there were no seasonal variations. They analyzed 41 SNPs related to the metabolic pathways of vitamin D and calcium, and included nine VDR SNPs, of which only rs7968585 was associated with increased vitamin D levels after supplementation, although the statistical significance disappeared after adjustment for treatment adherence. SNPs of the gene encoding DBP, of CYP2R1, which regulates the synthesis of calcitriol, and CYP24A1, which determines the metabolism of calcitriol to a less active metabolite, 24,25-hydroxycholecalciferol, were associated with baseline levels of vitamin D. These results are consistent with a recent genome-wide association study (Wang et al., 2010) that found that three SNPs—rs10766197 of the CYP2R1 gene, rs6013897 of

the CYP24A1 gene, and rs7768585 of the VDR gene—determined the response to vitamin D supplements. Two other studies, with fewer patients, found similar results but with different SNPs, some linked to genes that regulate DBP, and with higher doses of vitamin D.

A small 3-month case-control study by [Elnenaei et al. \(2011\)](#) evaluated the effect of calcium (1200 mg/day) and vitamin D (800 IU/day) supplements on vitamin D levels in osteopenic or osteoporotic postmenopausal women. The study population was divided into responders and nonresponders according to the PTH response, and found a different frequency of the alleles of the VDR polymorphisms determined by FokI in nonresponders and responders. In other polymorphisms (BsmI and TaqI) no differences were observed. In postmenopausal African American women, [Nieves et al. \(Nieves et al., 2007\)](#) also found an association between the response to vitamin D (measured by reduced PTH levels) and the FokI polymorphism. Other authors found a better response with the B allele of the BsmI polymorphism.

The effect of vitamin D supplementation (high and low doses) on bone mass according to VDR polymorphisms was assessed in adolescents. A better response was found for the bb and tt genotypes, with no variations according to the dose ([Arabi et al., 2009](#)).

[Perez-Alonso et al. \(2019\)](#) conducted a prospective study in which 151 women were randomized to two groups: one with 1000 mg of calcium and 800 IU vitamin D supplementation (102 women); and a placebo group with neither calcium nor vitamin D supplements (49 women). The follow-up was from May to September. There were no differences in baseline and 12-week levels of vitamin D in terms of genotypes and haplotypes, except for the Bat haplotype, where baseline values were lower (25OHD: 21 ± 10 , $P = .038$). The Bat haplotype is associated with lower baseline levels of vitamin D and a worse response to supplementation, and may therefore be a risk factor for vitamin D deficiency.

Therefore, although it is difficult to draw definitive conclusions given the heterogeneity of the studies, such studies support the idea that VDR polymorphisms may influence the response to the administration of vitamin supplements.

Interaction Between Soy and the Vitamin D Receptor Genotype

There are no studies showing a synergistic effect of the coadministration of soy and vitamin D, although this is theoretically possible. In vitro studies have shown that isoflavones increase VDR expression on the cell surface, which could potentiate the action of vitamin D. It is speculated that isoflavones increase CYP27B1 activity and reduce CYP24R gene activity, resulting in an

increase in calcitriol levels and a reduction in 24,25-dihydroxycholecalciferol levels. Furthermore, the coadministration of vitamin D and isoflavones increases the maximum plasma concentration of isoflavones, which might produce a greater effect. Therefore, these data suggest that there may be a synergistic effect of the two elements that could enhance the individual action of each ([Swami et al., 2005](#)).

There are few nutrigenetics studies of soy supplements. [Serrano et al. \(Serrano et al., 2013\)](#), in a study where the primary objective was to assess the effect of soy and vitamin D supplements on the cardiovascular risk profile, analyzed the total population and subsequently divided it according to the BsmI restriction enzyme of the VDR. In the overall population, increases in vitamin D were related to the dose administered, but soy supplements had no additional effect. There was a different response according to VDR polymorphisms. The G allele was associated with a greater increase in 25-hydroxycholecalciferol and 1,25-hydroxycholecalciferol. The addition of genistein to vitamin D supplementation has a beneficial additive effect on levels of vitamin D and markers of bone remodeling, an indirect marker of bone quality, and whether this effect is mediated by genes that regulate isoflavone metabolism. Only the heterozygous AT of the SNP rs358231 showed no reduction in bone remodeling markers. Intracellular β -glucosidases (CBG) are expressed in the intestine, where they hydrolyze genistein and daidzein. The lack of suppression of bone-remodeling markers in heterozygous AT patients could indicate a lack of response in women who obtained no benefits from supplementation ([Pérez-Alonso et al., 2017](#)).

Interaction Between Caffeine and the Vitamin D Receptor Genotype

An adverse effect of excess caffeine consumption on bone mineral density and fracture risk has been described, although the effect is weak and disappeared after adjusting for several variables. An intake >300 mg/24 h of caffeine increases bone loss in the elderly, but this effect is modified by the VDR TaqI polymorphism. Homozygous tt subjects had a greater reduction in lumbar bone mineral density than subjects with the T allele (TT and Tt) ([Rapuri et al., 2001](#)).

Interaction Between the Vitamin B Complex and the Methylene-tetrahydrofolate Reductase Genotype

Homocysteine is a modifiable risk factor for the occurrence of osteoporotic fracture. Homocysteine may interfere with the collagen bridges (cross-linking),

producing a deterioration in bone quality. Methylene-tetrahydrofolate reductase (MTHFR) is a flavoprotein that catalyzes the reduction of N⁵,N¹⁰-methylene-tetrahydrofolate to N⁵-methyltetrahydrofolate, which is necessary for homocysteine methylation. There is a thermolabile variant of MTHFR, with a 50% reduction of activity. This variation is caused by an alteration in the MTHFR gene due to the substitution of a cytosine with thymidine, in the transition of the 677 nucleotide (C677T). The T allele results in decreased affinity of MTHFR with a cofactor necessary for its activity. This binding may be stabilized by riboflavin (vitamin B2) or folic acid (vitamin B11). Homozygous subjects (TT) have higher levels of homocysteine and an increased risk of fracture than CC individuals. The effect can be observed only with low levels of riboflavin, which can be reversed by dietary administration. Low folic acid plasma concentrations are associated with high homocysteine levels, but not with an increased risk of fracture (Jacques et al., 2002).

Interaction Between Vitamin K and the ApoE Genotype

Apolipoprotein E (ApoE) is a ligand of the low-density lipoprotein receptor. Analysis of the relationship between the ApoE gene and osteoporosis showed an association between E4 allele carriers and the risk of fracture. The possible mechanism is an accumulation of oxidized lipids in the subendothelial space of the bone vessels that inhibits osteoblast differentiation. In addition, ApoE is essential for the transport of vitamin K, which is involved in the carboxylation of the glutamic acid residues of osteocalcin, a key protein in bone formation. Analysis of the relationship between vitamin K intake, the risk of hip fracture, and the E4 allele found no interaction between the three elements, although women with high-vitamin K intake showed a decreased risk of hip fracture (Peter et al., 2011).

Small studies have analyzed other polymorphisms, but may be considered anecdotal. The effect of lycopene, a carotenoid present in vegetables, on bone remodeling according to polymorphisms of the gene encoding the enzyme human paraoxonase 1 (PON1) was analyzed in a study that found that lycopene administration reduced bone remodeling when two polymorphisms (TT of SNP 172 TA and GG of SNP 584AG) were present. The effect was mediated by their antioxidant activity (Mackinnon et al., 2010).

Interaction Between Fat Intake-Genotype Visfatin

We analyzed the relationship between the visfatin genotype (rs2110385, an SNP in the gene promoter region)

and dietary fat intake in obese people. Visfatin is an adipocyte peptide that has a variable effect on bone density. Subjects with the GG genotype and high polyunsaturated fatty acid diets had lower hip bone mineral density (Khorrami-Nezhad et al., 2018).

CONCLUSION

Studies have found an association between polymorphisms of some molecules related to bone metabolism, especially the VDR, and the response to calcium and vitamin D supplementation. The studies are heterogeneous, with designs that make them prone to bias, and the effects are generally small. However, they suggest that allelic variations may influence the response to calcium and vitamin D supplements and, therefore, affect mineral and bone homeostasis.

References

- Arabi, A., Zahed, L., Mahfoud, Z., El-Onsi, L., Nabulsi, M., Maalouf, J., El-Hajj Fuleihan, G., 2009. Vitamin D receptor gene polymorphisms modulate the skeletal response to vitamin D supplementation in healthy girls. *Bone* 45, 1091–1097.
- Barry, E.L., Rees, J.R., Peacock, J.L., Mott, L.A., Amos, C.I., Bostick, R.M., Figueiredo, J.C., Ahnen, D.J., Bresalier, R.S., Burke, C.A., Baron, J.A., 2014. Genetic variants in CYP2R1, CYP24A1, and VDR modify the efficacy of vitamin D3 supplementation for increasing serum 25-hydroxyvitamin D levels in a randomized controlled trial. *J Clin Endocrinol Metab* 99, E2133–E2137.
- Bouillon, R., Marcocci, C., Carmeliet, G., et al., 2018. Skeletal and extra-skeletal actions of vitamin D: current evidence and outstanding questions. *Endocr Rev*. <https://doi.org/10.1210/er.2018-00126>.
- Chang, B., Schlussek, Y., Sukumar, D., Schneider, S.H., Shapses, S.A., 2015. Influence of vitamin D and estrogen receptor gene polymorphisms on calcium absorption: Bsm1 predicts a greater decrease during energy restriction. *Bone* 81, 138–144.
- Correa-Rodriguez, M., Schmidt Rio-Valle, J., Gonzalez-Jimenez, E., Rueda-Medina, B., 2015. A cross-sectional study of the association of VDR gene, calcium intake and heel ultrasound measures in early adulthood. *Calcif Tissue Int*. <https://doi.org/10.1007/s00223-015-0086-2>.
- Dawson-Hughes, B., Harris, S.S., Finneran, S., 1995. Calcium absorption on high and low calcium intakes in relation to vitamin D receptor genotype. *J Clin Endocrinol Metab* 80, 3657–3661.
- Deluca, H.F., 2014. History of the discovery of vitamin D and its active metabolites. *BoneKey Rep* 3, 8–15.
- Elnenaï, M.O., Chandra, R., Mangion, T., Moniz, C., 2011. Genomic and metabolomic patterns segregate with response to calcium and vitamin D supplementation. *Br J Nutr* 105, 71–79.
- Ferrari, S., Rizzoli, R., Manen, D., Slosman, D., Bonjour, J.P., 1998. Vitamin D receptor gene start codon polymorphisms (FokI) and bone mineral density: interaction with age, dietary calcium and 3'-end region polymorphisms. *J Bone Miner Res* 13, 925–930.
- Ferrari, S.L., Rizzoli, R., Slosman, D.O., Bonjour, J.P., 1998. Do dietary calcium and age explain the controversy surrounding the relationship between bone mineral density and vitamin D receptor gene polymorphisms? *J Bone Miner Res* 13, 363–370.
- Ferrari, S., Manen, D., Bonjour, J.P., Slosman, D., Rizzoli, R., 1999. Bone mineral mass and calcium and phosphate metabolism in young

- men: relationships with vitamin D receptor allelic polymorphisms. *J Clin Endocrinol Metab* 84, 2043–2048.
- Gennari, C., Adami, S., Agnusdei, D., Bufalino, L., Cervetti, R., Crepaldi, G., Di Marco, C., Di Munno, O., Fantasia, L., Isaia, G.C., Mazzuoli, G.F., Ortolani, S., Passeri, M., Serni, U., Vecchiet, L., 1997. Effect of chronic treatment with ipriflavone in postmenopausal with low bone mass. *Calcif Tissue Int* 61 (Suppl. 1), S19–S22.
- Jacques, P.F., Kalmbach, R., Bagley, P.J., Russo, G.T., Rogers, G., Wilson, P.W.F., Rosenberg, I.H., Selhub, J., 2002. The relation between riboflavin and plasma total homocysteine in the Framingham Offspring cohort is influenced by folate status and the C677T transition in the methylenetetrahydrofolate reductase gene. *J Nutr* 132, 283–288.
- Khorrami-Nezhad, L., Mirzaei, K., Maghbooli, Z., Keshavarz, S.A., 2018. Dietary fat intake associated with bone mineral density among visfatin genotype in obese people. *Br J Nutr* 119, 3–11.
- Mackinnon, E.S., El-Soheby, A., Rao, A.V., Rao, L.G., 2010. Paraoxonase 1 polymorphisms 172TA and 584AG modify the association between serum concentrations of the antioxidant lycopene and bone turnover markers and oxidative stress parameters in women 25–70 years of age. *J Nutrigenetics Nutrigenomics* 3, 1–8.
- Nieves, J., Ralston, S.H., Vasquez, E., Ambrose, B., Cosman, F., Lindsay, R e, 2007. Vitamin D receptor FokI polymorphism influences response to vitamin D supplementation in postmenopausal African-American women. In: Burckhardt, P., Heaney, R., Dawson-Hughes, B. (Eds.), *Nutritional Aspects of Osteoporosis*. Academic Press, San Diego, pp. 126–132.
- Pérez-Alonso, M., Briongos, L., Ruiz-Mambrilla, M., et al., 2017. The effect of genistein supplementation on vitamin D levels and bone turnover markers during the summer in healthy postmenopausal women. Role of genotypes of isoflavone metabolism. *J Nutrigenetics Nutrigenomics* 10, 139–145.
- Perez-Alonso, M., Briongos, S., Ruiz-Mambrilla, M., et al., 2019 Feb. Association between Bat Vitamin D receptor 3' haplotypes with vitamin D levels at baseline and worse response after increase of vitamin D by supplements and exposure to sunlight. *Int J Vitam Nutr Res* 21, 1–5.
- Peter, I., Crosier, M.D., Yoshida, M., Booth, S.L., Cupples, L.A., Dawson-Hughes, B., Karasik, D., Kiel, D.P., Ordovas, J.M., Trikalinos, T.A., 2011. Association of APOE gene polymorphisms with bone mineral density and fracture risk: a meta-analysis. *Osteoporos Int* 22, 1199–1209.
- Rapuri, P.B., Gallagher, J.C., Kinyamu, H.K., Ryschon, K.L., 2001. Caffeine intake increases the rate of bone loss in elderly women and interacts with vitamin D receptor genotypes. *Am J Clin Nutr* 74, 694–700.
- Serrano, J.C., De Lorenzo, D., Cassanye, A., Martin-Gari, M., Espinel, A., Delgado, M.A., Pamplona, R., Portero-Otin, M., 2013. Vitamin D receptor BsmI polymorphism modulates soy intake and 25-hydroxyvitamin D supplementation benefits in cardiovascular disease risk factors profile. *Genes Nutr* 8, 561–569.
- Stathopoulou, M.G., Dedoussis, G.V., Trovas, G., Katsalira, A., Hammond, N., Deloukas, P., Lyritis, G.P., 2010. Low-density lipoprotein receptor-related protein 5 polymorphisms are associated with bone mineral density in Greek postmenopausal women: an interaction with calcium intake. *J Am Diet Assoc* 110, 1078–1083.
- Swami, S., Krishnan, A.V., Peehl, D.M., Feldman, D., 2005. Genistein potentiates the growth inhibitory effects of 1,25-dihydroxyvitamin D (3) in DU145 human prostate cancer cells: role of the direct inhibition of CYP24 enzyme activity. *Mol Cell Endocrinol* 241, 49–63.
- Tella, S.H., Gallagher, J.C., 2014. Prevention and treatment of postmenopausal osteoporosis. *J Steroid Biochem Mol Biol* 142, 155–170.
- Uitterlinden, A.G., Fang, Y., Van Meurs, J.B.J., Pols, H.A.P., Van Leuw, J.P.T., 2004. Genetic and biology of vitamin D receptor gene polymorphisms. *Gene* 338, 143–156.
- Wang, T.J., Zhang, F., Kestenbaum, B., Van Meurs, J.B., Berry, D., Kiel, D.P., et al., 2010. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. *Lancet* 376, 180–188.
- Zaidi, M., Yuen, T., Rosen, C.J., 2018. Regulation of skeletal homeostasis. *Endocr Rev* 39, 701–718.

Nutrients and Genes in the Liver: The Central Role of Liver X Receptor in Lipogenesis

Marica Cariello, Elena Piccinin, Oihane Garcia-Irigoyen, Antonio Moschetta

Department of Interdisciplinary Medicine, “Aldo Moro” University of Bari, Bari, Italy

Glossary Terms and Keywords

The conversion of excess of dietary carbohydrates into fatty acids in the liver is orchestrated by the coordinate function of liver X receptor (LXR), sterol regulatory element-binding protein 1c (SREBP1c), and carbohydrate response element-binding protein (ChREBP).

NUTRIGENOMICS IN THE LIVER: THE CENTRAL ROLE OF LIVER X RECEPTOR IN LIPOGENESIS

During food abundance periods, nutrients in excess can be stored as fat to provide safeguard mechanism in case of food shortage. Evolution supports this adaptation with the aim of keeping energy equilibrium under changeable nutritional status. A major control on metabolic homeostasis is carried out by liver, responsible for the synthesis or degradation, storage, and redistributions of nutrients and macromolecules, as carbohydrates and lipids.

Lipids and glucose are two major energy sources of the organism. Dietary fats are necessary for membrane constitution and act as a precursor for important molecules, such as bile acids and hormones. Similarly, glucose not only serves as a fuel substrate but also provides building blocks for the synthesis of numerous macromolecules. Fatty acid synthesis and glucose metabolism are tightly regulated pathways in the liver, critical for the whole-body energy homeostasis. Changes in diet composition can deeply influence normal hepatic metabolism. For instance, high-carbohydrate intake as well as an overload of dietary lipids are able to influence biological processes, modifying gene expression profile through the induction of specific transcription factors. The liver X receptor (LXR) plays a key role in metabolic liver homeostasis via sterol regulatory element-binding

protein 1c (SREBP1c) and carbohydrate response element-binding protein (ChREBP). Specifically, these transcription factors coordinately regulate hepatic de novo lipogenesis in response to carbohydrate overload, in order to convert excess of glucose into body fats, providing protection against intermittent famine. Modifications of these transcription factors may alter normal metabolism, thus increasing the susceptibility to a plethora of diseases.

THE NUCLEAR RECEPTOR LXR: DEFINITION AND MECHANISM OF ACTION

The liver is an important regulator of metabolic homeostasis able to control body energy balance, enabling synthesis or degradation, storage, and redistributions of different nutrients and macromolecules. The liver plays a central role in carbohydrate, lipid, and amino acid metabolism, and it is responsible for the synthesis of plasma proteins and detoxification of xenobiotics. Liver metabolic processes are tightly regulated by neuronal and hormonal systems. Furthermore, numerous transcription factors and coactivators are involved in controlling the expression of enzymes catalyzing the rate-limiting steps of liver metabolic processes, thus controlling liver and finally whole-body energy metabolism.

Nuclear receptors (NRs) are master transcriptional regulators of liver development, differentiation, and function. They act as sensors of dietary and endogenous molecules, thus linking the nutritional and hormonal changes to the transcriptional modifications. Thus, NRs can be considered crucial for the whole-body physiology as they are able to integrate the metabolic environmental stimuli to the gene expression responses.

NR structure consists of a DNA-binding domain and ligand-binding domain, which interact with specific ligands. Common metabolites, as bile acids, nutrients, and xenobiotics, are usually involved in the activation of NRs. Upon binding to specific ligands, NRs undergo conformational changes that cause the consequent released of the corepressor complex and the concomitant recruitment of the coactivator. In this way, NRs can activate the expression of target genes.

The LXR is one of the members of the NRs super family that is encoded by the gene NR1H3; LXRs exist as two isoforms, LXR α and LXR β , that bind to DNA as obligate heterodimer with the retinoid X receptor (RXR) and can be activated either by LXR or RXR ligands. LXR/RXR heterodimer binds to LXR responsive elements, a specific DNA sequence found in LXR target genes, composed by two 5'-AGGTCA-3' hexameric half-sites separated by four bases (DR4 motif). Studies of sequence comparison revealed a 77% sequence homology between the two isoforms of LXR genes. However, whereas the LXR β isoform is ubiquitously expressed, LXR α is mainly expressed in high metabolically active tissues, including liver, intestine, adipose tissue, and macrophages (Janowski et al., 1999).

Originally identified as an orphan receptor, it is now known that LXRs respond to sterol concentration, being activated by cholesterol derivatives, including oxysterols and 24(S), 25-epoxycholesterol. Intracellular high-sterol concentration can activate LXRs, that in turn implement different strategies aimed to the reduction of the cellular sterol load. Specifically, once activated, LXRs mobilize cholesterol from the peripheral organs to the liver, where it is converted and excreted as bile acids into the intestinal lumen. Concomitantly, LXRs block cholesterol synthesis and limit its uptake at cellular level, finally resulting in a net cholesterol removal from the body and in a better lipoprotein profile. All of this evidence contributes to the definition of LXR as the whole body transcriptional cholesterol sensor (Chen et al., 2007).

One of the first LXR target genes to be identified is the cholesterol 7 α -hydroxylase (Cyp7a1), a rate-limiting enzyme in the synthesis of bile acids, also involved in the fecal sterol excretion. In mice fed with high cholesterol diet, LXR is activated and promotes a net increase of CYP7A1 gene expression that results in the upregulation of cholesterol conversion to bile acids for excretion into the intestinal lumen.

After the identification of CYP7A1, numerous other genes have been described as targets of LXR activity. Many of them play a pivotal role in the reverse cholesterol transport (RCT), a process finely tuned by LXRs, where the excess of cholesterol is transported to the liver, undergoing different processes that determine timely elimination through the feces and the restoring of

physiological cholesterol levels. Among LXR target genes involved in RCT, the ATP-binding cassette transporters ABCA1 and ABCG1 concur to mobilize cholesterol from peripheral organs, and from macrophages in particular. Macrophages are not only fundamental in the immune response but they also collect cholesterol on blood vessel walls, thus contributing to atherosclerotic plaque formation. The LXR's activation induces both the ABC transporters, which actively promote the cholesterol efflux from macrophages to HDL and apolipoproteins, ultimately leading to cholesterol transport to the liver. Concomitantly, LXRs may influence lipoprotein profiles via the induction of several genes (inducible degrader of the LDL receptor, Idol; cholesteryl ester transfer protein, Cept; lipoprotein lipase, Lpl) that regulate lipoprotein composition, along with the cholesterol transport and uptake (Repa et al., 2000a).

Moreover, LXRs promote the expression of ATP-binding ABCG5 and ABCG8 transporters at both hepatic and intestinal levels, which modulate cholesterol excretion, consequently diminishing liver cholesterol content and intestinal cholesterol absorption, while increasing the fecal sterol excretion.

Besides the fundamental role in cholesterol homeostasis, LXRs exert a main function in liver fatty acid metabolism. Most of these effects are mediated by the LXR's induction of sterol regulatory element-binding protein 1c (SREBP-1c), the principal regulator of triglycerides and fatty acids synthesis. The expression of lipogenic genes is LXR isoform-mediated; whereas the LXR β deletion does not show any effect on the transcription of SREBP-1c, mice LXR α knockout display a great reduction on liver fatty acids synthesis due to the downregulation of SREBP1c and associated target genes, stearoyl-CoA desaturase 1 (SCD1) and fatty acids synthase (FASN). Moreover, recently it has been described that LXR activation can also modulate glycemia, promoting the conversion in fatty acids. Indeed, LXRs increase ChREBP expression, a glucose-sensitive transcription factor that, increasing lipogenic enzymes, mediates the transformation of carbohydrates into lipids (Repa et al., 2000b; Cha and Repa, 2007; Wang and Tontonoz, 2018). Furthermore, the significant LXR function in glucose metabolism has been confirmed by several studies, suggesting that LXR expression and some LXR target genes can be upregulated by insulin. Interestingly, LXR knockout mice fed with a diet rich in fat and cholesterol are resistant to obesity and display defects in hepatic lipids metabolism. These observations highlight the fine-tuned function of LXRs, where tissue-specific effects are well interconnected in order to promote body homeostasis. Indeed, in conditions characterized by excess of hepatic cholesterol, LXR is able to promote hepatic cholesterol clearance while permitting triglycerides' normal uptake and storage in other tissues (Fig. 52.1).

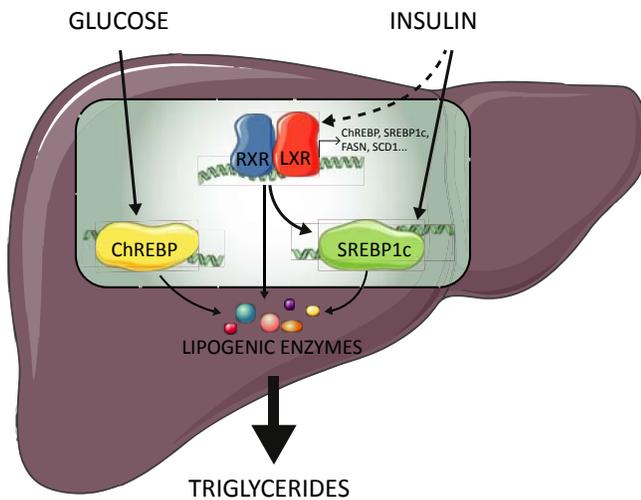


FIGURE 52.1 The central role of LXR in hepatic lipogenesis. The nuclear receptor LXR plays a central role in liver lipogenesis. Upon carbohydrate overload, the rise of glucose and insulin levels in the blood contributes to LXR activation, that in turn specifically coordinates the upregulation of the transcription factor SREBP-1c. The functions of SREBP-1c and ChREBP, another transcription factor, can increase due to the stimulation of insulin and glucose, respectively. Thus, during times of carbohydrate abundance, LXR, ChREBP, and SREBP-1c tend to increase in the liver, finally orchestrating the expression of different lipogenic genes in order to convert the excess of glucose into fatty acids. The conversion of excess of dietary carbohydrates into triglycerides that will be stored as fat by the organism is a fundamental safeguard mechanism against food-shortage periods. *ChREBP*, carbohydrate-responsive element-binding protein; *LXR*, liver X Receptor; *SREBP-1c*, sterol regulatory element-binding transcription factor 1c.

LXR AND HEPATIC LIPOGENESIS: THE ROLE OF SREBP1C

Fatty acid and fat synthesis in the liver is a tightly regulated metabolic pathway critical for energy utilization. Lipogenic genes are coordinately regulated at the transcriptional level by the nuclear receptor LXRs that play a critical role in cholesterol homeostasis and bile acids metabolism.

In mice, oral administration of LXR agonists results in elevated hepatic fatty acids synthesis and steatosis, increased triglyceride-rich very-low-density lipoprotein secretion that leads to hypertriglyceridemia. LXR upregulates the CYP7A1 gene expression to promote production and elimination of bile acids. In macrophages and enterocytes, LXR-mediated expression of ABCA1 promotes free cholesterol efflux from the cells. Furthermore, LXR stimulates hepatic lipogenesis via the upregulation of SREBP-1c (Shimano and Sato, 2017).

SREBP-1c is a transcription factor that belongs to the basic helix-loop-helix leucine zipper (bHLH-Zip) family. The mammalian genome encodes three SREBP isoforms: SREBP-1a, SREBP-1c, and SREBP-2. SREBP-1a mediates

the synthesis of cholesterol, fatty acids, and triglycerides. SREBP-1c activates the expression of genes dedicated to fatty acids synthesis, but not cholesterol synthesis. Finally, the last isoform SREBP-2 enhances cholesterologenesis.

The structure of SREBPs consists of three domains: (1) NH₂-terminal domain that contains the bHLH-Zip region for the DNA binding; (2) two hydrophobic transmembrane segments, and (3) COOH-terminal domain with regulatory functions. All SREBPs are assembled as long inactive precursors that are bound to membranes of endoplasmic reticulum. In order to act as a transcription factor, the NH₂-terminal domain of SREBPs must be released from the membrane through a proteolytic cleavage. SREBPs form complexes with SREBP cleavage-activating protein (SCAP), an oxysterol sensor, and an escort of SREBPs. When cells are depleted of sterols, SCAP transports SREBPs from endoplasmic reticulum to the Golgi reticulum, where two proteases, site 1 (S1P) and site 2 (S2P) proteases, act sequentially to release the NH₂-terminal domain from the membrane. In this way, the bHLH-Zip domain comes into the nucleus and, subsequently, binds to sterol responsive element (SRE) in the enhancer/promoter region of target genes starting their transcription. When cellular cholesterol concentration raises, the SCAP/SREBPs complex is no longer incorporated into endoplasmic reticulum transport vesicles and NH₂-terminal domain cannot be released from the membrane leading to the inhibition of target genes transcription. SREBP-1c activates enzymes that play a crucial role in fatty acids biosynthesis including ATP citrate lyase, acetyl-CoA carboxylase, and fatty acid synthase (essential to palmitate production), the rate limiting enzyme of the fatty acid elongase complex that converts palmitate to stearate, stearoyl-CoA desaturase that converts stearate to oleate and glycerol-3-phosphate acyltransferase, a key enzyme in triglyceride and phospholipids synthesis. Finally, SREBP-1c induces three genes for NADPH production, a molecule employed in the lipid biosynthetic pathways (Horton et al., 2002).

The transcriptional regulation of SREBP-1c is selectively regulated by insulin, glucagon, and LXRs. In the liver, following carbohydrate overload, insulin stimulatory effects on fatty acids synthesis is mediated via SREBP-1c activation. In the liver and adipose tissue, the total amount of SREBP-1c is reduced by fasting, leading to insulin levels suppression and glucagon levels increase. In line with these phenomena, rats treated with streptozotocin, which abolish insulin secretion, exhibited low SREBP-1c RNA levels. Insulin stimulates SREBP-1c gene transcription in the liver, and this gene mediates insulin-stimulated fatty acids synthesis. The insulin effect on SREBP-1c is abolished in presence of high cAMP levels (due to elevated glucagon levels);

this event is associated with acetyl-CoA carboxylase (ACC) and FASN inhibition. Moreover, it has been demonstrated that in *ob/ob* mice SREBP-1c mRNA levels are elevated as a result of hyperinsulinemia due to leptin deficiency. Leptin treatment normalized insulin levels with a subsequent SREBP-1c downregulation. Insulin works by increasing the ability of LXR to activate SREBP-1c promoter. This process is specific for SREBP-1c promoter, because insulin does not induce other LXR-responsive genes transcription.

LXR-mediated induction of SREBP-1c expression represents the mechanism to stimulate oleate synthesis in presence of high cellular sterols concentration. Mice lacking LXR genes present reduced levels of SREBP-1c mRNA in parallel with the downregulation of the lipogenic enzymes mRNA levels. Moreover, in wild-type mice, but not in LXR knockout mice, fed with high-cholesterol-content diets show the overexpression of SREBP-1c. Cholesterol feeding arises in LXR-mediated upregulation of SREBP-1c and subsequent induction of stearoyl CoA desaturase (SCD-1), an enzyme involved in *cis* desaturation of stearoyl-CoA and palmitoyl-CoA into oleoyl-CoA and palmitoyl-CoA, respectively. SCD-1 upregulation increases oleoyl-CoA concentration, a substrate for ACAT-mediated cholesterol esterification. In this way, LXR is indirectly able to promote the esterification of free cholesterol to protect the cell from its deleterious effects. LXR-mediated SREBP-1c upregulation raises lipid synthesis to coordinate a homeostatic balance between fatty acids and sterols concentrations.

LXR AND HEPATIC GLUCOSE SIGNALING: THE ROLE OF CHREBP

In mammals, the liver has a central role in glucose homeostasis being in charge of the conversion of excess dietary carbohydrates into triglycerides (TG), a process mediated by *de novo* lipogenesis. Glucose, in addition to being an energetic substrate and a basic anabolic building block for the synthesis of several macromolecules, acts as a signaling molecule in the liver. Glucose influx in hepatocytes stimulates the expression of numerous genes encoding glycolytic and lipogenic enzymes, such as liver-pyruvate kinase (L-PK), ACC, FASN, and SCD1, in order to suppress endogenous glucose production and convert excess glucose into fat. Over recent years, the transcription factor ChREBP has emerged as a crucial mediator of glucose-mediated induction of both glycolysis and lipogenesis in liver. ChREBP (also known as Mondo B or MLXIPL) belongs to the Mondo family of bHLH-Zip transcription factors, whose expression is most abundant in liver, small intestine, and white and brown adipose tissue, which are the most active sites of *de novo* fatty acid synthesis in the

body. In accordance with a role as glucose-regulated transcription factor, hepatic ChREBP expression is induced in response to a high-carbohydrate dietary intake, but not in response to polyunsaturated fatty acid diet or fasting.

Soon after the identification of ChREBP as an important mediator of glucose transcriptional effect in 2001, the Max-like protein X (Mlx), its functional obligatory partner, was identified. ChREBP-Mlx heterodimers bind the promoters of glucose-responsive genes, activating their transcription in a glucose-dependent manner. This transcriptional effect is dictated by a conserved consensus DNA sequence found within the promoter regions of glucose-regulated genes, named the carbohydrate response element (ChoRE). The ChoRE sequence is composed of two CACGTG E box elements separated by five base pairs (5'-CACGTGnnnnn-CACGTG-3'), where the two ChREBP-Mlx heterodimers bind (Filhoulaud et al., 2013).

ChREBP is a large transcription factor (864 amino acids) that contains several functional domains, including two nuclear export signals (NES1 and NS2) and a nuclear import signal (NLS) in the N-terminal region, a polyproline domain, and in the c-terminal region a DNA-binding bHLH-Zip domain. Systematic serial deletions of the ChREBP sequence led to the identification of a glucose-sensing module, an evolutionally conserved domain in Mondo proteins (also known as Mondo conserved region), which contains a low glucose inhibitory domain (LID) and a glucose response conserved element (GRACE). The transactivative activity of GRACE is retained by LID and high glucose releases this inhibition (Filhoulaud et al., 2013; Beaven et al., 2013).

The regulation of ChREBP function relies on subcellular localization, DNA binding, and transcriptional activity. Interestingly, the activation of this transcription factor in response to glucose involves multiple post-translational modifications, such as dephosphorylation, O-GlcNAcylation, and acetylation. During fasting or hypoglycemia conditions, glucagon-induced protein kinase A (PKA) phosphorylates ChREBP on serine 196 (Ser196) leading to the 14.3.3 protein binding and cytosolic retention. Threonine 666 (Thr666), located in the DNA-binding domain, is also phosphorylated by PKA avoiding ChREBP DNA-binding activity. Under high blood glucose concentrations, ChREBP undergoes dephosphorylation on Ser196 and Thr666 allowing nuclear translocation, heterodimerization with Mlx, and binding to the ChoRE of its target genes. Nevertheless, the fact that mutations in Ser196 and Thr666 do not produce a constitutively active ChREBP led to the search for additional regulatory mechanism. In response to glucose, the histone acetyltransferase transcriptional coactivator P300 mediates the acetylation on lysine 672, located in the DNA-binding domain of ChREBP

protein, enhancing its binding to promoters. High glucose also triggers O-linked β -N-acetylglucosamine transferase (OGT)-mediated O-GlcNacylation resulting in an increased ChREBP DNA binding and protein stability. At the transcriptional level, ChREBP is regulated by glucose and by LXR. ChREBP is a target gene of LXR as demonstrated by the presence of two LXR-responsive elements in the promoter of the ChREBP gene and the increased of ChREBP mRNA in response to LXR agonists (Cha and Repa, 2007). However, in the absence of an increased glucose flux, ChREBP is not transcriptionally active, suggesting that LXR is not able to overcome the posttranslational modifications induced by glucose.

ChREBP plays a central role in the regulation of glycolysis and de novo lipogenesis in liver acting in synergy with SREBP-1c in the conversion of excess dietary carbohydrates into TGs. Several ChREBP target genes identified to date are also directly regulated by LXR or SREBP-1c, such as FASN, ACC, and SCD-1, whereas the induction of L-PK, a rate-limiting enzyme of glycolysis, is exclusively dependent on ChREBP. Hepatic overexpression of ChREBP in mice leads to liver steatosis, while ChREBP deficient mice display reduced hepatic glycolytic and lipogenic gene expression, larger glycogen-laden livers, smaller adipose depots, decreased plasma free fatty acids, and impaired glucose tolerance. Importantly, liver specific inhibition of ChREBP in obese and insulin-resistant ob/ob mice leads to reversal of hepatic steatosis and improved insulin resistance pointing to a controversial connection between ChREBP-induced steatosis and insulin sensitivity. However, it has been recently shown by adenoviral-mediated ChREBP overexpression in mice that ChREBP-induced hepatic steatosis is dissociated from insulin resistance (Benhamed et al., 2012). In this study, lipidomic analysis revealed that in ChREBP-mediated hepatic steatosis beneficial lipid species, including monounsaturated fatty acids (MUFAs), were preferentially enriched, showing that ChREBP overexpression, elevating SCD-1 protein content, was able to reverse palmitate-induced insulin resistance by increasing the MUFA/saturated fatty acid ratio.

Altogether, ChREBP has emerged as the main mediator of the metabolic adaptation to changing glucose levels by regulating target genes involved in glycolysis and lipogenesis.

CONCLUDING REMARKS AND NUTRIGENETIC IMPLICATION OF LXR

Evolution has favored the capacity of storing nutrients as fat during food abundance periods as a safeguard against the intermittent famine, maintaining energy balance under varying nutritional conditions.

Nevertheless, in humans, this adaptation is now implicated in the emergence of hepatic steatosis and metabolic disorders due to the dramatic changes in feeding behavior and lifestyle (Hong and Tontonoz, 2014). Given the prominent role of LXR in controlling numerous aspect of whole-body physiology, it is not surprising that alterations on this nuclear receptor may alter normal metabolism and concur to the pathogenesis of various diseases. LXR, as a sensor of dietary components able to orchestrate the physiological response to nutrients regulating fatty acid and cholesterol metabolism in the liver, represents a target of interest for obesity-related diseases such as nonalcoholic fatty liver disease (NAFLD). Excessive accumulation of lipid droplets within hepatocytes is the hallmark of NAFLD, a lipotoxic disease, where liver fat accumulation may lead to the alteration of the organ function. It has been estimated in patients with NAFLD that around 30% of the TG content in liver came from de novo lipogenesis, highlighting the importance of LXR signaling in this disease. Moreover, it has been described that LXR activation alters the composition of atherosclerotic plaques, limiting inflammation and promoting changes in fibrous cap thickness, finally attenuating atherosclerosis. Finally, in vivo experiments using LXRs ligands provide evidence on the ability of this nuclear receptor in decreasing cell growth and promoting cell death; thus, LXRs negatively affect liver regeneration as well as tumor progression in different types of tissue.

The identification of the nuclear receptor LXR together with the incessant effort made to deepen our understanding on LXR mechanism in the regulation of lipid homeostasis contributes to the develop of new therapeutic approaches. Multiple evidence proving LXR's implication in different type of disease led to the development of specific LXR modulators. The first LXR synthetic ligands (T0901317 and GW3965) constituted invaluable tools in scientific research; however, they failed a clinical application due to their pleiotropic effect. Indeed, it has been demonstrated that systemic LXR activation could result in desirable as well as undesirable outcomes that discourage implementation in clinical practice. Nowadays, new compounds have entered into clinical trials, but most of them do not display any success. Studies focused on new drug discovery strategies are therefore desired to develop original safe LXR target to efficiently treat metabolic alterations.

References

- Beaven, S.W., et al., 2013. Reciprocal regulation of hepatic and adipose lipogenesis by liver X receptors in obesity and insulin resistance. *Cell Metabol.*
- Benhamed, F., Denechaud, P.D., Lemoine, M., Robichon, C., Moldes, M., Bertrand-Michel, J., Ratziu, V., Serfaty, L.,

- Housset, C., Capeau, J., Girard, J., Guillou, H., Postic, C., 2012. The lipogenic transcription factor ChREBP dissociates hepatic steatosis from insulin resistance in mice and humans. *J Clin Investig.*
- Cha, J.Y., Repa, J.J., 2007. The liver X receptor (LXR) and hepatic lipogenesis. The carbohydrate-response element-binding protein is a target gene of LXR. *J Biol Chem.*
- Chen, W., Chen, G., Head, D.L., Mangelsdorf, D.J., Russell, D.W., 2007. Enzymatic reduction of oxysterols impairs LXR signaling in cultured cells and the livers of mice. *Cell Metabol.*
- Filhoulaud, G., Guilmeau, S., Dentin, R., Girard, J., Postic, C., 2013. Novel insights into ChREBP regulation and function. *Trends Endocrinol Metabol.*
- Hong, C., Tontonoz, P., 2014. Liver X receptors in lipid metabolism: opportunities for drug discovery. *Nat Rev Drug Discov.*
- Horton, J.D., Goldstein, J.L., Brown, M.S., 2002. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J Clin Investig.*
- Janowski, B.A., Grogan, M.J., Jones, S.A., Wisely, G.B., Kliewer, S.A., Corey, E.J., Mangelsdorf, D.J., 1999. Structural requirements of ligands for the oxysterol liver X receptors LXRalpha and LXRbeta. *Proc Natl Acad Sci USA.*
- Repa, J.J., Turley, S.D., Lobaccaro, J.A., Medina, J., Li, L., Lustig, K., Shan, B., Heyman, R.A., Dietschy, J.M., Mangelsdorf, D.J., 2000. Regulation of absorption and ABC1-mediated efflux of cholesterol by RXR heterodimers. *Science.*
- Repa, J.J., Liang, G., Ou, J., Bashmakov, Y., Lobaccaro, J.M., Shimomura, I., Shan, B., Brown, M.S., Goldstein, J.L., Mangelsdorf, D.J., 2000. Regulation of mouse sterol regulatory element-binding protein-1c gene (SREBP-1c) by oxysterol receptors, LXRalpha and LXRbeta. *Genes Dev.*
- Shimano, H., Sato, R., 2017. SREBP-regulated lipid metabolism: convergent physiology - divergent pathophysiology. *Nat Rev Endocrinol.*
- Wang, B., Tontonoz, P., 2018. Liver X receptors in lipid signalling and membrane homeostasis. *Nat Rev Endocrinol.*

Hyperuricemia and Gout

Geetha Chittoor, Venkata Saroja Voruganti

Department of Nutrition and UNC Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, NC, United States

INTRODUCTION

This chapter concerns the biomedical importance of uric acid and health implications of increased concentrations of uric acid in serum (hyperuricemia).

Uric acid is a product of purine metabolism. In most animals, uric acid is converted to allantoin, which is readily excreted with urine. This conversion is catalyzed by the enzyme urate oxidase (uricase). However, in humans and higher-order primates the activity of uricase has been lost, thus resulting in a tendency for higher levels of serum uric acid (SUA) than in other mammals (Sautin and Johnson, 2008). Elevated SUA concentrations (hyperuricemia defined by SUA >7 mg/dL in men and SUA >6 mg/dL in women), or elevated urinary uric acid concentrations (hyperuricosuria) can lead to gout and uric acid nephrolithiasis and increase the risk for hypertension, chronic kidney disease (CKD), and cardiovascular disease (CVD) (Sautin and Johnson, 2008; Kutzing and Firestein, 2008; Kang and Ha, 2014; Maiuolo et al., 2016). About 70% of total uric acid excretion occurs through the kidney and the remainder through the intestines. Imbalance in the production and renal excretion of uric acid is commonly involved in the development of either hyperuricemia or hypouricemia, which further leads to the metabolic diseases mentioned above (Maiuolo et al., 2016).

With the exception of gout, which is triggered by precipitation of urate crystals, direct mechanistic connections between circulating uric acid levels and associated pathologies have been difficult to establish. Uric acid is a powerful scavenger of reactive oxygen species (singlet oxygen, peroxy, and hydroxyl radicals) and is considered a key antioxidant in humans (Sautin and Johnson, 2008). However, high SUA is also associated with several diseases, including several linked to

increased oxidative stress. Thus, the paradoxical role of uric acid as anti- and/or prooxidant has been the subject of an ongoing debate (Kang and Ha, 2014).

The concentration of uric acid in serum reflects four aspects of uric acid metabolism: dietary purine intake, endogenous purine metabolism, renal urate excretion, and intestinal urate excretion (Anzai et al., 2007). Of these, intake is the major modifiable risk factor for the development of hyperuricemia. Major food sources of purines include organ meats, other meats, seafood, and beer. Notably, high-fructose intake may also contribute to hyperuricemia. Phosphorylation of fructose to fructose 1-phosphate increases the adenosine monophosphate (AMP) pool at the expense of adenosine triphosphate (ATP). The AMP then enters the purine metabolic pathway and results in formation of uric acid (Fig. 53.1). Phosphorylation of fructose is catalyzed by fructokinase, which differs from other hexokinases in that it is not subject to product inhibition. Consequently, fructose intake causes rapid depletion of ATP and increased production of uric acid in the liver (Kutzing and Firestein, 2008).

The consumption of added sugars in the United States has increased considerably in the past few decades. This increase parallels the increase in hyperuricemia and cardiometabolic diseases. Added sugars, particularly of fructose-containing sucrose and of high-fructose corn syrup, have been a subject of concern due to their adverse metabolic effects on various diseases such as gout, type 2 diabetes, and CVD (Choi et al., 2005). Importantly, fructose not only rapidly elevates concentrations of SUA, as described earlier, but also appears to compete with uric acid excretion (Maiuolo et al., 2016). Other dietary modifiers of SUA include alcohol, vitamin C, and caffeine. Alcohol (ethanol) is known to decrease uric acid excretion and increase SUA. High-dose vitamin C has a uricosuric

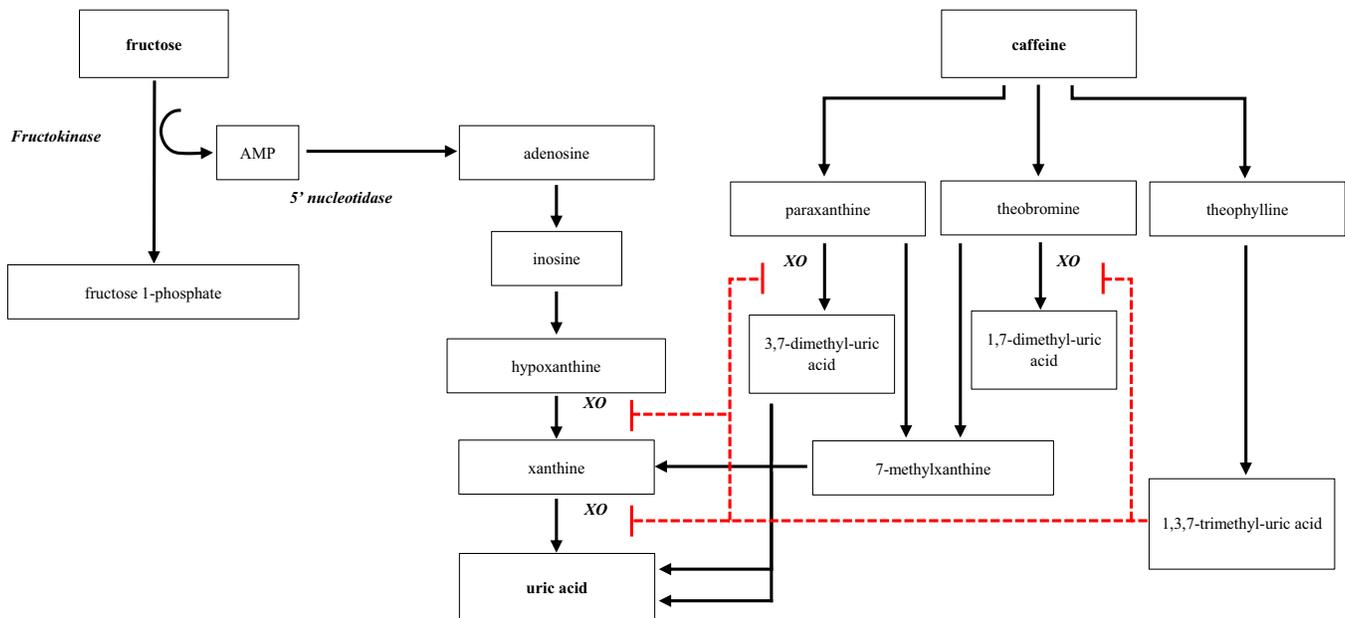


FIGURE 53.1 Key nutrients affecting hyperuricemia. Enzymes are bolded and italicized. Red lines indicate inhibition. XO, xanthine oxidase.

effect and reduces SUA levels (Choi et al., 2005). Caffeine inhibits xanthine oxidase activity (Fig. 53.1), which decreases SUA concentration.

Endogenous Purine Metabolism

Several enzymes are involved in the conversion of purines to uric acid. AMP (both endogenous and dietary) is deaminated and dephosphorylated to form inosine. Inosine is deribosylated to hypoxanthine, which is oxidized to xanthine. Xanthine is then oxidized to form uric acid. The key enzyme that oxidizes hypoxanthine to xanthine and xanthine to uric acid is xanthine oxidase (XO) (Fig. 53.1). The xanthine dehydrogenase (XDH) gene encodes the XDH enzyme, which can then be converted to XO through either reversible sulfhydryl oxidation or irreversible proteolysis. Most of the XDH product in the liver exists in the XDH form, which prefers NAD^+ as an electron acceptor for the oxidation of xanthine to urate. In contrast, XO prefers O_2 as the electron acceptor for oxidation of hypoxanthine to xanthine and of xanthine to urate. Thus, XO, but not XDH, is known to produce peroxide reactive oxygen species, causing oxidative stress in various tissues (Maiuolo et al., 2016). Hence, knowledge of these intermediary metabolites and XO-related pathways is necessary for the understanding of the role of uric acid and the risks and diseases its high concentration can cause in humans.

Renal Urate Excretion

The kidney is the major site for excretion of uric acid and accounts for two-thirds to three-fourths of daily losses. The proximal tubule is the site of secretion and reabsorption of uric acid. Renal urate handling can be categorized into four processes: filtration, reabsorption, secretion and postsecretory absorption (Fig. 53.2). Most of the reabsorption takes place at the S1 segment whereas most of the secretion occurs at the S2 segment of the proximal tubule. This whole process requires a set of urate transporters as shown in Fig. 53.3. These specialized transporters are key to the excretion of uric acid. Anzai et al. (2007) proposed the existence of a urate-transporting molecular complex (the urate transportome) as a model for renal urate transport. According to this model, renal urate transport should be evaluated not only based on a single transporter but considering the functional unit comprised of urate transporters and other molecules requiring specific protein-protein interactions of the molecular complex at the plasma membrane. This complex includes transporters of solute carrier channels SLC2A9, SLC5A8, SLC5A12, SLC13A3, SLC16A9, SLC17A1, SLC17A3, SLC22A6, SLC22A8, SLC22A11, and SLC22A12, ATP-binding cassette, subfamily G, member 2 (ABCG2), sodium phosphate protein (NTP1) and ATP-binding cassette, subfamily C, and member 4 (ABCC4) (Fig. 53.3).

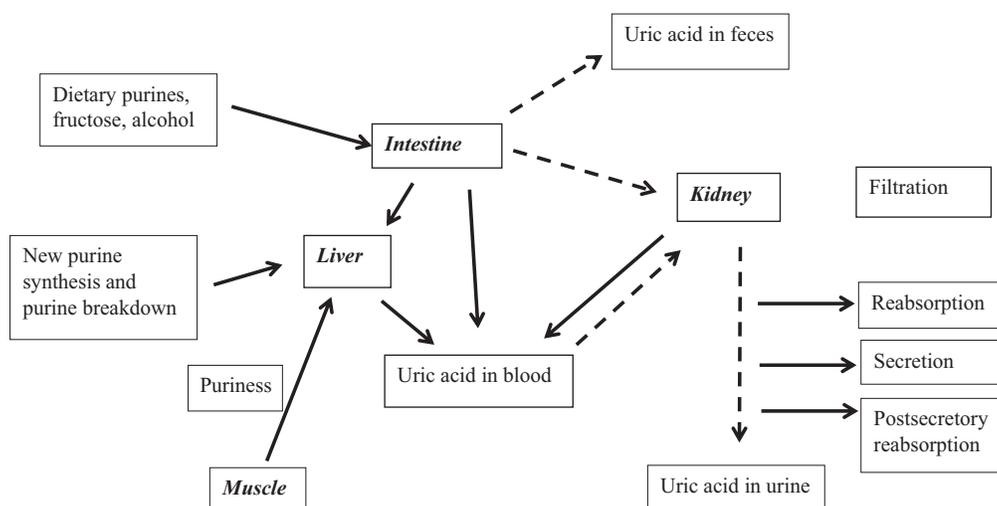


FIGURE 53.2 **Uric acid metabolism.** Major organs are bolded and italicized. *Straight lines* indicate contribution of uric acid to the serum pool. *Dotted lines* indicate uric acid excretion.

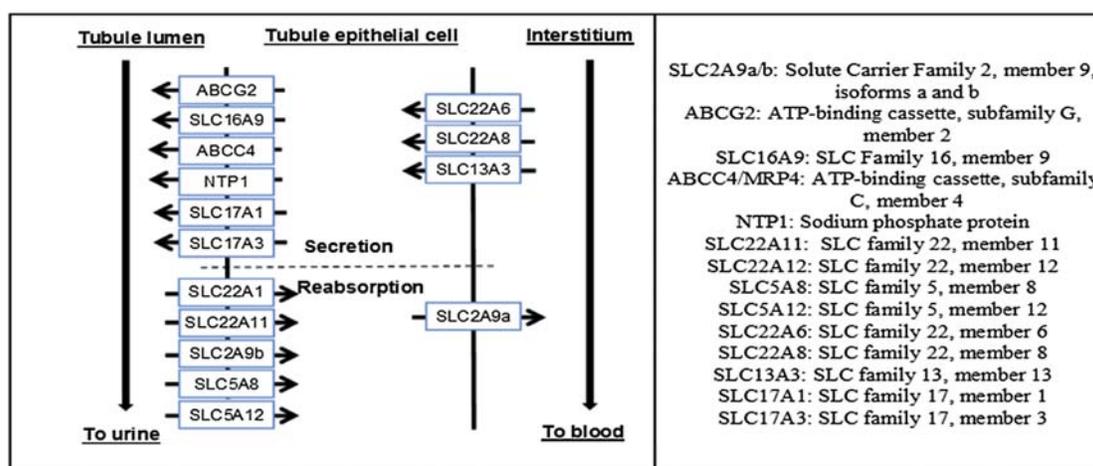


FIGURE 53.3 **Urate handling in renal proximal tubules.** Adopted from Anzai, N., Kanai, Y., Endou, H., 2007. *New insights into renal transport of urate.* *Curr Opin Rheumatol* 19, 151-157.

Intestinal Urate Excretion

Normally, ~30% of the uric acid is excreted through the intestine. This pathway becomes more significant in cases of renal insufficiency and gout as it has been observed that individuals with these conditions tend to increase their enteric uric acid disposal (Xu et al., 2016). As in renal urate transport, several urate transporters operate in the intestines as well. The key transporters that are active in both renal and intestinal transport are ABCG2, SLC2A9, SLC16A9, ABCC4 and SLC22A11, and SLC22A12. The intestines are the main site of uric acid degradation with an important contribution by bacterial uric acid transporters. The main group of bacterial transporters, nucleobase-ascorbate transporters (NAT/NCS2), includes more than 2000 members

of which 15 have been characterized experimentally (Xu et al., 2016).

CLINICAL SIGNIFICANCE

Hyperuricemia

Hyperuricemia is characterized by SUA >7 mg/dL in men and SUA >6 mg/dL in women. Hyperuricemia is determined by two important physiological mechanisms: (1) increased production of uric acid through endogenous and/or dietary sources and (2) reduced renal and intestinal excretion of uric acid (Maiuolo et al., 2016; Merriman, 2015). Studies have linked hyperuricemia to a number of diseases and conditions

including gout, kidney stones, CKD, hypertension, and CVD, including myocardial infarction and stroke. Although the causal relationship of hyperuricemia with the development of gout and kidney stones is clear, the way that hyperuricemia contributes to the development of other diseases/conditions is less certain. Hyperuricemia is believed by some to be just a marker or a consequence of disease but not a causal factor. However, recent studies have shown hyperuricemia to be an independent risk factor for renal and cardiovascular morbidity and mortality (Kutzing and Firestein, 2008). Hyperuricemia in children is known to be associated with hypertension in adult life. Animal studies have also implicated hyperuricemia in the development of hypertension, endothelial dysfunction, and kidney injury (Kang and Ha, 2014).

Gout

Gout occurs in about a quarter of people with elevated serum uric acid concentrations (hyperuricemia). Gout is characterized by the accumulation and crystallization of uric acid in the joints with subsequent attacks of acute inflammatory arthritis. Gout usually affects all joints, but the metatarsal-phalangeal joint of the big toe is most often affected. Although it is well known that the hyperuricemia results in the formation of monosodium urate crystals in the synovial fluid and the ensuing inflammatory response, the underlying mechanisms are not well understood (Riches et al., 2009; Merriman, 2015). However, some studies have indicated that the central pathway may be the one involving the activation of Toll-like receptors and inflammasome-mediated release of proinflammatory cytokine interleukin-1 β (Merriman, 2015). Gout has also been linked to increased incidence of oxidative stress-related diseases. Hyperuricemia and gout, like any other complex phenotype, result from the interplay between inherited genetic risk variants and environmental exposures.

Other Diseases/Disorders

Because uric acid is a metabolic end product, it is not always apparent if an association between SUA levels and a particular phenotype (especially CVD and CKD) is causative or symptomatic. While earlier epidemiological analysis suggested that SUA level was a covariate of CVD risk, subsequent studies identified hyperuricemia as an independent risk factor in carotid plaque formation and cardiovascular mortality. A recent study found that soluble monosodium urate enhances NLRP3 and IL-1 production through a TLR4-dependent pathway (Sautin and Johnson, 2008; Kang and Ha, 2014), thus providing a potential mechanistic link between

hyperuricemia and renal mesangial injury and supporting hyperuricemia as a causative factor in diseases beyond gout.

Similarly, human and animal studies have shown the role of hyperuricemia in hypertension, atherosclerosis, CVD, initiation and progression of renal disease, and metabolic syndrome. In fact, drugs that decrease SUA levels, mainly allopurinol, have been shown to improve survival in chronic heart failure patients, improve endothelial function in patients with chronic kidney disease, and reduce oxidative stress and improve endothelial function in patients with coronary artery disease (Kutzing and Firestein, 2008; Riches et al., 2009), supporting the idea that uric acid plays a causal role in disease risk.

The role of hyperuricemia in the development of kidney stones or nephrolithiasis and gout is well recognized. The relationship between uric acid and kidney function seems to be two-sided. On the one hand, decline in glomerular filtration rate (kidney function parameter) may lead to elevation of uric acid; on the other hand, increase in uric acid seems to alter glomerular function through renal vasoconstriction and increased rennin expression (Sautin and Johnson, 2008). Animal studies, mainly in rodent models, have shown hyperuricemia to inhibit nitric oxide system in the kidney, resulting in endothelial dysfunction and kidney injury (Kang and Ha, 2014). It has been strongly advocated that hyperuricemia be included as an additional component of metabolic syndrome since individuals with metabolic syndrome tend to have higher levels of SUA (Sautin and Johnson, 2008).

Benefits of Treating Asymptomatic Hyperuricemia

Hyperuricemia is not a disease by itself. But, sustained hyperuricemia, even if not symptomatic, increases risk for disease previously mentioned in the chapter. In general practice, asymptomatic hyperuricemia is not treated aggressively until complications arise. However, treatment of hyperuricemia helps in multiple ways. Reduction of uric acid concentrations not only helps alleviate gout and nephrolithiasis symptoms but also helps improve several cardiorenal disease risk markers. Usually, the first line of therapy for gout is lifestyle modification through changes in diet and physical activity. The main focus of the medical treatment has been reducing uric acid production by inhibiting XO using drugs such as allopurinol and febuxostat, which are XO inhibitors. These drugs have been shown to reduce SUA levels and results in improvement of associated diseases such as survival in chronic heart failure patients, endothelial function in patients with CKD and coronary artery disease (Kutzing and Firestein, 2008;

Riches et al., 2009), and treating end-stage renal disease in patients with diabetes (Doria and Krolewski, 2011).

Genetics and Environmental Factors Affecting Hyperuricemia and Gout

There is strong evidence that genetic factors play a key role in the regulation of SUA.

Several genetic studies have been conducted to understand the genetic contribution to the variation in SUA concentrations and gout. They have shown that SUA concentrations are significantly heritable (Kottgen et al., 2013; Voruganti et al. 2013, 2014). Kottgen et al. (2013) reported 28 genetic loci to be associated with SUA levels. This study is considered to be significant as it was conducted in more than 140,000 individuals. It confirmed 10 previously reported loci and identified 18 novel loci. The 10 loci include urate transporters, SLC2A9, ABCG2, SLC22A11, SLC22A12, SLC17A1, and SLC16A9, which have been previously reported to be involved in renal urate transport. The 18 new loci showed weaker effects on SUA as compared to the 10 established ones, and their role in renal urate transport is not yet clear.

There are few Mendelian disorders that cause hyperuricemia and gout. Mutations in hypoxanthine-guanine phosphoribosyl transferase, an enzyme involved in the purine salvage pathway, cause an X-linked recessive disorder, Lesch-Nyhan syndrome, which is characterized by hyperuricemia and gout; and secondly, mutations in phosphoribosyl pyrophosphatase synthetase, an enzyme involved in urate synthesis, causes another X-linked but dominant disorder of primary hyperuricemia and gout (Riches et al., 2009). Other disorders such as glycogen storage disorders and mutations in uromodulin gene are also known to cause hyperuricemia and gout (Riches et al., 2009).

Major Urate Transporters

1. Solute carrier channel 2, member 9 (*SLC2A9*) encodes for a transporter that reabsorbs urate by transporting it from tubular cells to the peritubular interstitium. Located on chromosome 4p16, *SLC2A9* exists in two isoforms (long and short) and is expressed in kidney, liver and gut. Both isoforms are involved in urate transport, with the short isoform expressed on the apical side of the collecting duct and the long isoform expressed on the basolateral side. The *SLC2A9* also transports glucose and fructose, and, thus, it is hypothesized that the sharing of the transporter interferes with excretion of uric acid (Merriman, 2015). The genetic effect of *SLC2A9* is very strong with several of its gene single-nucleotide polymorphisms

having strong associations with SUA with significant effect sizes. Besides Kottgen et al. (2013) who reported these findings in Europeans, other studies have replicated the strong effects in Mexican Americans (Voruganti et al., 2013), African Americans (Tin et al., 2011), Asians (Yang et al., 2014), Native Americans (Voruganti et al., 2014), and Hispanic children (Voruganti et al., 2015).

2. ATP-binding cassette, subfamily G, member 2 (*ABCG2*) encodes a membrane transporter that is expressed in kidney, liver, and gut. It was initially found to be a xenobiotic transporter involved in multidrug resistance for a specific human breast cancer. However, its role in urate transport has been strongly supported with molecular evidence to show that *ABCG2* mediates both renal as well as intestinal excretion (Merriman, 2015). Like *SLC2A9*, several genetic association studies have shown strong association of *ABCG2* with SUA concentrations in various populations (Kottgen et al., 2013; Tin et al., 2011; Yang et al., 2014).
3. Solute carrier channel 22, member 12 (*SLC22A12*) encodes a transporter URAT1 that is a renal urate-anion exchanger. Localized to chromosome 11q13, its mutations have been associated with increased excretion of uric acid and lower SUA concentrations. The protein, URAT1, was found to be located in epithelial cells of the proximal tubule and known to exist in multiple transcript variants. Molecular studies in humans and rodent models have shown that mutations in *SLC22A12* lead to idiopathic renal hypouricemia (Riches et al., 2009).

Identification of genetic determinants of uric acid transporters will provide insight into the physiological role of SUA transport in the kidney and open new avenues for developing therapeutic targets. Although many genetic variants affecting SUA levels have been identified through genome-wide or candidate-gene association studies, very few variants have been found to affect either gene expression or the predicted amino acid sequence of the encoded protein making interpretation of their functional significance difficult.

Most common variants in urate transporter genes have been associated with serum uric acid across populations but seem to differ between populations based on their risk allele, frequencies, effect sizes and the direction of effects, indicating differential patterns of genetic contribution to SUA variation based on ancestries or ethnicities. For example, *SLC2A9* and *ABCG2* have been associated with SUA across populations of various ancestries. However, the variants, risk alleles, minor allele frequencies, and the effect sizes were different in each population (Kottgen et al., 2013; Voruganti et al. 2013,

2014; Yang et al., 2014). Therefore, detection of complete variation for disease risk necessitates that both rare as well as common variants are identified. Additionally, serum uric acid levels are affected by numerous nongenetic factors, including diet, age, gender, BMI, and other lifestyle factors, and differences in these factors between populations may also contribute to the observed differences in association.

CONCLUSION

Uric acid is a unique molecule with contradictory roles as antioxidant and prooxidant in the body. Variation in serum uric acid concentrations is affected by both genetic and environmental factors and their interactions. Regardless of the debate related to the beneficial and detrimental effects of uric acid, it is very likely that uric acid has an important causal role in the pathogenesis of many complex diseases including CVD and CKD, and thus, warrants further investigation.

References

- Anzai, N., Kanai, Y., Endou, H., 2007. New insights into renal transport of urate. *Curr Opin Rheumatol* 19, 151–157.
- Choi, H.K., Liu, S., Curhan, G., 2005. Intake of purine-rich foods, protein, and dairy products and relationship to serum levels of uric acid: the Third National Health and Nutrition Examination Survey. *Arthritis Rheum* 52, 283–289.
- Doria, A., Krolewski, A.S., 2011. Lowering serum uric acid levels to prevent kidney failure. *Nat Rev Nephrol* 7, 495–496.
- Kang, D.H., Ha, S.K., 2014. Uric acid puzzle: dual role as anti-oxidant and pro-oxidant. *Electrolyte Blood Res* 12, 1–6.
- Kutzing, A.K., Firestein, B.L., 2008. Altered uric acid levels and disease status. *J Pharmacol Exp Ther* 324, 1–7.
- Kottgen, A., Albrecht, E., Teumer, A., Vitart, V., et al., 2013. Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nat Genet* 45, 145–154.
- Maiuolo, J., Oppedisano, F., Gratteri, S., Muscoli, C., Mollace, V., 2016. Regulation of uric acid metabolism and excretion. *Int J Cardiol* 213, 8–14.
- Merriman, T., 2015. An update on the genetic architecture of hyperuricemia and gout. *Arthritis Res Ther* 17, 98.
- Riches, P.L., Wright, A.F., Ralston, S.H., 2009. Recent insights into the pathogenesis of hyperuricemia and gout. *Hum Mol Genet* 18, R177–R184.
- Sautin, Y.Y., Johnson, R.J., 2008. Uric acid: the oxidant-antioxidant paradox. *Nucleosides Nucleotides Nucl Acids* 27, 608–619.
- Tin, A., Woodward, O.M., Kao, W.H., Liu, C.T., Lu, X., Nalls, M.A., Shriner, D., Semmo, M., Akylbekova, E.L., Wyatt, S.B., Hwang, S.J., Yang, Q., Zonderman, A.B., Adeyemo, A.A., Palmer, C., Meng, Y., Reilly, M., Shilpak, M.G., Siscovick, D., Evans, M.K., Rotimi, C.N., Flessner, M.F., Kottgen, M., Cupples, L.A., Fox, C.S., Kottgen, A., 2011. CARE and CHARGE consortia. *Hum Mol Genet* 20, 4056–4068.
- Voruganti, V.S., Kent Jr., J.W., Debnath, S., Cole, S.A., Haack, K., Göring, H.H.H., Carless, M.A., Curran, J.E., Johnson, M.P., Almasy, L., et al., 2013. Genome-wide association analysis confirms and extends the association of SLC2A9 with serum uric acid levels to Mexican Americans. *Front Genet* 4, 279.
- Voruganti, V.S., Franceschini, N., Haack, K., Laston, S., MacCluer, J.W., Umans, J.G., et al., 2014. Replication of the effect of SLC2A9 genetic variation on serum uric acid levels in American Indians. *Eur J Hum Genet* 22, 938–943.
- Voruganti, V.S., Laston, S., Haack, K., Mehta, N.R., Cole, S.A., Butte, N.F., Comuzzie, A.G., 2015. Serum uric acid concentrations and SLC2A9 genetic variation in Hispanic children: the Viva La Familia study. *Am J Clin Nutr* 101 (4), 725–732.
- Xu, X., Li, C., Zhou, P., Jiang, T., 2016. Uric acid transporters hiding in the intestine. *Pharmaceut Biol* 54, 3151–3155.
- Yang, B., Mo, Z., Wu, C., Yang, H., He, Y., Gui, L., Zhou, L., Guo, H., Zhang, X., Yuan, J., Dai, X., Li, J., Qiu, G., Huang, S., Deng, Q., Feng, Y., Guan, L., Hu, D., Zhang, X., Wang, T., Zhu, J., Min, X., Lang, M., Li, D., Hu, F.B., Lin, D., Wu, T., He, M., 2014. A genome-wide association study identifies common variants influencing serum uric acid concentrations in a Chinese population. *BMC Med Genom* 11, 7.

Aging: Mechanisms, Signaling and Dietary Intervention

Deyang Yu^{1,2,3}, Luigi Fontana^{4,5,6}, Dudley W. Lamming^{1,2,3}

¹Department of Medicine, University of Wisconsin–Madison, Madison, WI, United States; ²William S. Middleton Memorial Veterans Hospital, Madison, WI, United States; ³Molecular & Environmental Toxicology Center, University of Wisconsin–Madison, Madison, WI, United States; ⁴Division of Geriatrics and Nutritional Sciences, Washington University in St. Louis, MO, United States; ⁵Department of Clinical and Experimental Sciences, Brescia University Medical School, Brescia, Italy; ⁶CEINGE Biotechnologie Avanzate, Napoli, Italy

INTRODUCTION TO CALORIE RESTRICTION

Calorie restriction (CR) is a dietary intervention in which caloric intake is limited while micronutrients are increased to ensure adequate nutrition, and which in the 1930s was discovered to significantly increase the life span and health of rodents. Since that time, there has been immense interest in understanding the physiological and molecular basis by which CR extends health and longevity. While a plethora of mechanisms have been advanced to explain the powerful benefits of CR, it is now generally believed that the “mechanisms underlying the beneficial effects of a CR diet are complex, and no single pathway likely accounts for the full range of effects” (Lamming and Anderson, 2014).

We recently reviewed the metabolic effects of CR in some detail (Fontana and Partridge, 2015; Lamming and Anderson, 2014), but in brief, CR has powerful beneficial effects on metabolic and molecular health in mammals ranging from mice and rats to nonhuman primates and humans. As might be expected in animals in which food intake is restricted, animals and humans placed on a CR diet are lean, having lower body weight and a dramatic reduction in adipose mass. Less apparent to the naked eye, one of the most dramatic effects of a CR diet is the impact on glucose metabolism; mammals including humans placed on a CR diet have significantly decreased inflammatory markers, fasting blood glucose levels, and increased insulin sensitivity. This effect has attracted significant attention as a

possible mediator of longevity; while insulin sensitivity itself is not clearly associated with longevity, inflammation, insulin resistance, and high blood glucose levels are very clearly associated with negative health consequences, including not only the development of diabetes but also cardiovascular disease, Alzheimer disease, and some of the most common types of cancer.

Consumption of a CR diet results in alterations in many hormones that are associated with health and longevity. Most notably, CR in rodents decreases levels of insulin-like growth factor 1 (IGF-1), which a variety of evidence has linked to the regulation of longevity. Several strains of mutant mice, including the Ames and Snell dwarf mice, have naturally low levels of IGF-1 and are extremely long lived. Higher levels of IGF-1 are associated with decreased life span in both rodents and dogs, and IGF-1 levels are also associated with cancer risk. In humans, however, CR lowers serum IGF-1 concentration only when protein intake is restricted as well. CR also significantly alters the secretion of many adipose-derived hormones such as leptin and adiponectin, which are involved in central regulation of feeding behavior, appetite suppression, and weight maintenance. Several of these hormones are also involved in the regulation of insulin sensitivity in peripheral tissues.

Other proposed physiological mediators of the effects of CR include alterations in the metabolic rate, decreased oxidative stress and DNA damage, and alterations in sex hormones and reproduction. We have covered these mechanisms at some length (e.g. Fontana and Partridge, 2015; Lamming and Anderson, 2014) and

thus will not discuss them in-depth here. However, it is clear that one of the first proposed models—that CR passively lowers the rate of metabolism and thus the rate of cellular damage due to reactive oxygen species (ROS)—is most likely not correct. After a period of adaptation, mice placed on CR do not generally have a decreased metabolic rate. Further, genetically increasing ROS defense mechanisms usually does not significantly extend mouse life span, and indeed inhibiting ROS acts to block the effects of CR on life span in *Caenorhabditis elegans*. It is now generally believed that CR is instead an active process, in which protective responses are triggered in response to the challenge of decreased calories.

At the molecular level, CR causes multiple beneficial modifications. Both in rodents and humans, CR upregulates transcripts involved in energy metabolism (e.g., TCA cycle, mitochondrial ETC, and glycolysis), and downregulates the expression of genes involved in inflammation, fibrosis, antioxidant responses, cholesterol metabolism, and cell proliferation. In particular, CR in humans inhibits the insulin/IGF-1 signaling pathway, and increases chaperone and autophagy genes and proteins involved in protein quality control and organelle homeostasis that promote the removal of dysfunctional proteins and organelles from cell (Yang et al., 2016).

DIETARY PROTEIN IS A CRITICAL REGULATOR OF THE EFFECTS OF CR

Since the discovery of CR, there has been significant interest in understanding if the effects of CR are mediated by particular macronutrients. Not only would this potentially highlight less onerous dietary regimens that are easier to adhere to, but understanding which dietary nutrients are responsible for the beneficial effects of CR might provide a clue as to the biological mechanisms behind these effects. CR typically involves the restriction of all major macronutrient categories, including glucose and other carbohydrates, amino acids, and lipids, and by testing the effects of carbohydrate-, protein-, and fat-restricted diets, clues to the molecular mechanisms that mediate the effects of CR may be gleaned. However, testing the effect of individual dietary components on longevity has long been complicated by the fact that any change in the macronutrient composition of a diet inevitably leads to alterations in either calorie content or the amounts of other macronutrients.

Recently, a new approach—termed nutritional geometry—has been utilized to uncover the impact of different dietary macronutrients on health and longevity. Essentially, this approach tests many diets, each with a distinctive ratio of different dietary macronutrients, in parallel. Using a nutritional geometry approach, Dr. Stephen Simpson and colleagues have

analyzed the effect of distinct dietary macronutrient ratios on the life span of both *Drosophila* and mice. Analyzing the life span of a large cohort of mice fed one of 25 different diets with distinct ratios of protein, carbohydrates, and fat, Dr. Simpson observed that mice fed low-protein diets lived the longest (Solon-Biet et al., 2014). This result, building on previous data from the Simpson laboratory as well as the laboratory of Dr. Linda Partridge and others on the important role of amino acids in *Drosophila* life span, has highlighted the critical role of dietary protein in the regulation of life span—and perhaps, in the response to calorie restriction (Green and Lamming, 2019).

Many of the metabolic benefits that result from CR are reproduced by protein restriction (PR) in both mice and humans (Fontana et al., 2016; Fontana and Partridge, 2015; Solon-Biet et al., 2015). PR improves metabolic health, promoting both leanness and the regulation of blood sugar in both humans and mice (Fontana et al., 2016). Critically, CR utilizing a low-protein diet does not result in a further increase in metabolic health, suggesting that reduced dietary protein is critical to the beneficial effects of CR (Solon-Biet et al., 2015). The next step in the nutrigenomic study of CR and aging is to determine the precise mechanisms and nutrients responsible for the effects of a PR diet.

SPECIFIC DIETARY AMINO ACIDS MEDIATE THE IMPACT OF PROTEIN RESTRICTION

The building blocks of protein are amino acids; traditionally, we reference the 20 dietary amino acids that are directly encoded in DNA, nine of which are essential. It has recently been realized that many dietary amino acids have important metabolic roles in addition to their requirement for protein synthesis. The most studied of these are the sulfur-containing amino acids methionine and cysteine; restriction of dietary methionine and cysteine extends the life span of both mice and rats.

Several studies have reported that mice fed on methionine-restricted diet exhibit positive biometric measures including decreased visceral fat deposition, improved insulin sensitivity, and increased medium and maximum life span. While young mice were used for methionine restriction in the majority of studies, one study found that MR could reverse age-induced obesity and insulin resistance in adult animals. Twelve-month-old mice on methionine-restricted diet were shown to have decreased body weight and fat mass, and increased physical activity, and improved glucose tolerance to the level measured in healthy 2-month-old control-fed mice. Mechanistically, these beneficial effects were associated with decreased hepatic

lipogenic gene expression, a remodeled lipid metabolism in white adipose tissue, and improved insulin sensitivity in peripheral tissues as well as an increase in circulating and hepatic gene expression levels of the insulin-sensitizing hormone FGF21. We observed similar effects when we deprived diet-induced obese mice of methionine, which rapidly reduced adiposity and improved glucose tolerance (Yu et al., 2018).

The contribution of other amino acids to the longevity-promoting effects PR and CR have not been the subject of significant investigation, although dietary restriction of tryptophan has been reported to have positive effects on life span and age-related pathologies in rodents; long-lived mammals also tend to have reduced levels of tryptophan metabolites (Brown-Borg and Buffenstein, 2016; Ma et al., 2015). However, over the past decade evidence has mounted suggesting that the three-branched chain amino acids (BCAAs) leucine, isoleucine, and valine play a critical role in the regulation of insulin sensitivity. The BCAAs are elevated in insulin-resistant humans as well as mouse and rat models of diet-induced obesity, and circulating levels of BCAAs correlate well with the likelihood of developing type 2 diabetes. Studies in high-fat-diet-fed rats have also demonstrated that BCAA supplementation impairs insulin sensitivity, implicating a causal role for BCAAs in the pathogenesis of diabetes. Notably, dietary protein intake correlates with blood levels of BCAAs in mice, and short-term PR decreases blood levels of BCAAs in humans, suggesting that BCAAs may mediate some of the effects of a PR diet in both species (Fontana et al., 2016; Solon-Biet et al., 2014). Using a novel dietary paradigm in which we utilized amino acid-defined diets to reduce, rather than eliminate, levels of the three BCAAs consumed by mice, we recently determined that restriction of dietary BCAA is as effective as reduction of dietary protein by the same amount with regard to metabolic health, improving glucose tolerance, alleviating β cell metabolic stress, and reducing adiposity (Fontana et al., 2016). Restriction of BCAAs to diet-induced obese mice rapidly restores metabolic health, reducing adiposity and restoring glucose tolerance and insulin sensitivity (Cummings et al., 2018).

These findings suggest that the effects of PR and DR on metabolism, and perhaps even on longevity, may be mediated in part by decreased consumption of methionine and/or BCAAs. Identifying the precise contribution of each amino acid to the regulation of metabolic health and longevity will be critical in appreciating how dietary quality—the source and precise composition of the calories we consume—and not simply dietary quantity regulates health. Identifying the mechanisms by which the restriction of specific amino acids, including methionine or the BCAAs, improves

metabolic health is of great significance and may provide valuable guidance in aiding the development of pharmacological agents to harness these mechanisms to promote metabolic health and to combat metabolic diseases.

AMINO ACID SENSING LINKS SPECIFIC AMINO ACIDS TO METABOLISM

Coordinating the availability of amino acids with growth and protein translation is critical to the maintenance of cellular homeostasis. Two evolutionarily conserved protein kinases, general control nonderepressible 2 (GCN2) and the mechanistic target of rapamycin (mTOR) tightly couple the availability of amino acids to anabolic processes such as protein translation and catabolic cellular processes such as autophagy (Green and Lamming, 2019). These kinases are likely key in coordinating the metabolic response to protein and calorie restriction, conditions in which amino acids derived from dietary proteins are less abundant.

As amino acids are transported and incorporated to a nascent peptide by covalently binding to transfer RNA during protein synthesis, the accumulating uncharged tRNAs resulting from amino acid starvation/deficiency bind to and activate GCN2, resulting in the phosphorylation of eukaryotic translation initiation factor 2 alpha (eIF2 α) at Ser51 by GCN2. Phosphorylation of this key residue of eIF2 α has two effects on protein translation: it decreases global translation in order to reduce the utilization of amino acids, while selectively promoting the translation of mRNAs containing short upstream open reading frames, most notably activating transcription factor 4 (ATF4).

ATF is a member of the ATF subfamily of the basic leucine zipper (bZIP) transcription factor superfamily, and it controls a large quantity of genes involved in promoting recovery from stress. In the context of amino acid scarcity, this enables the induction and upregulation of genes coding for amino acid biosynthetic enzymes and amino acid transporters. Therefore, the GCN2-eIF2 α -ATF4 axis serves as a key pathway regulating the response to decreased amino acid availability. It remains to be determined if this axis, which is strongly activated by short-term protein-free diets, plays an important role in the response to CR and PR.

THE MECHANISTIC TARGET OF RAPAMYCIN IS REGULATED BY DIETARY PROTEIN

One target of ATF4 is Sestrin2, a stress response protein well characterized as a regulator of metabolism and

AMPK. Recently, it was discovered that Sestrin2 has another role: linking GCN2 activity to the regulation of the second evolutionarily amino acid-sensing kinase: the mechanistic target of rapamycin (mTOR). mTOR is a serine/threonine protein kinase that belongs to the phosphatidylinositol-3-OH kinase (PI (3) K)-related family and functions as a master regulator of growth and metabolism in the cell. mTOR forms two distinct protein complexes, mTOR complex 1 (mTORC1) and mTOR complex 2 (mTORC2). Amino acids, particularly leucine but also the other BCAAs and methionine, are potent agonists of mTORC1 signaling. We and others have reviewed recent advances in understanding how mTORC1 activity is regulated elsewhere (Kennedy and Lamming, 2016; Goberdhan et al., 2016); briefly, mTORC1 is activated by recruitment to the lysosome, a process mediated by the GTP-loading status of the Rag family of small GTPases. Once at the lysosome, mTORC1 may be activated by interaction with GTP-loaded Rheb.

GATOR1, a GTPase activating protein (GAP) for Rag A and B, inhibits mTORC1 activity; GATOR2, a negative regulator of GATOR1, promotes mTORC1 activity. In the absence of amino acids, Sestrin2 binds to GATOR2 and blocks its ability to regulate GATOR1 function. Leucine binds to a specific pocket on Sestrin2, relieving Sestrin2-mediated inhibition of GATOR2 activity and activating mTORC1. Other BCAAs as well as methionine may also bind to this same pocket and activate mTORC1 *in vitro*, but the *in vivo* relevance of these interactions is not clear.

While mTORC1 is inactivated following withdrawal of amino acids (in cell culture) or feeding of a protein-free diet, the effect of translatable, low-protein diets on mTORC1 signaling has been investigated only recently. Dietary protein intake was recently shown to correlate with branched-chain amino acid levels as well as hepatic mTORC1 activity in C57BL/6J mice (Solon-Biet et al., 2014). We recently found that a low-protein diet significantly decreases mTORC1 signaling in numerous tissues as well as xenografted tumors of tumor-bearing mice (Lamming et al., 2015).

Evidence gathered over the last decade clearly indicates that mTOR pathway is involved in regulating aging and age-related diseases. Genetic as well as pharmacological inhibition of mTORC1 signaling extends the life span of model organisms ranging from yeast to mice. Treatment with rapamycin not only extends life span but also prevents or delays age-related diseases including cancer and Alzheimer disease. Unfortunately, rapamycin has a number of side effects, including immunosuppression, hyperlipidemia, glucose intolerance, insulin resistance, and an increased risk of new-onset diabetes (reviewed in Dumas and Lamming, 2019).

We and others have demonstrated that rapamycin, while an acute inhibitor of mTORC1, also inhibits a

second mTOR-containing complex, mTORC2, *in vivo* in most mouse tissues. Many side effects of rapamycin, particularly with regard to glucose homeostasis, are due to the disruption of mTORC2. With regard to aging, genetic experiments in model organisms including mice suggest that inhibition of mTORC1 is sufficient to extend life span, and that inhibition of mTORC2 may even have negative consequences for males. The negative effects of rapamycin on longevity resulting from inhibition of mTORC2 may preclude its use as a general therapy for the prophylactic treatment of age-related diseases. However, genetic experiments as well as data from mice treated with alternative rapamycin treatment regimens suggest that selective inhibition of mTORC1 could be a powerful antiaging therapy.

While pharmaceutical companies scramble to generate such selective molecules (Schreiber et al., 2019; Mahoney et al., 2018), dietary strategies may be able to target mTORC1. Both CR and PR inhibit mTORC1 signaling *in vivo*, and thus PR, as a less-onerous dietary regimen, may offer an intriguing alternative to caloric restriction. Defining how physiological and dietary relevant amino acids regulate mTORC1 signaling *in vivo*, and clarifying the mechanisms by which dietary quality regulates mTORC1, remains a challenging task.

THE REGULATION OF AGING BY AMPK

The maintenance of energy homeostasis and improved stress resistance are the hallmarks of healthy aging. As the master regulator of energy metabolism in the cell, 5' adenosine monophosphate-activated protein kinase (AMPK) plays a critical role in regulating health and longevity. AMPK is a serine/threonine kinase that is activated by energy deficiency as a result of elevated energy utilization such as exercise and/or decreased energy production such as nutrient starvation. Once activated, AMPK phosphorylates and regulates the activity of an array of substrates to regulate carbohydrate, lipid, and protein metabolism to stimulate ATP production and inhibit ATP consumption to restore energy homeostasis. In addition to regulating energy homeostasis, AMPK has also been shown to cross talk to other signaling pathways that regulate a number of biological processes including autophagy, stress resistance, and inflammation.

Most notably, AMPK promotes autophagy via two mechanisms: one by directly phosphorylating ULK1 (unc-51-like autophagy-activating kinase 1) to stimulate the initiation of autophagosome formation, the other via negatively regulating the activity of mTORC1. Enhanced autophagy resulting from AMPK activation constitutes an important protein quality control mechanism to remove defective proteins and organelles in the cell

and thus promotes protein homeostasis. It was recently revealed that AMPK directly phosphorylates the mammalian FOXO family of Forkhead transcription factors 3 (FoxO3), which induces the transcription of a number of genes involved in mediating cellular stress resistance. Moreover, AMPK suppresses inflammatory responses via inhibition of NF- κ B signaling. As chronic inflammation associated with increased NF- κ B signaling has been observed in several age-related metabolic disorders, reduced NF- κ B signaling resulting from AMPK activation constitutes another mechanism by which AMPK promotes longevity. Several studies have demonstrated the role of AMPK in mediating life-extending effects of calorie restriction and the efficacy of AMPK activators in extending life span in lower organisms such as *C. elegans*. Intriguingly, it has been observed that the sensitivity of AMPK to cellular stress declines with aging, implicating a causal role of decreased AMPK activity in aging.

THE NUTRIGENETICS OF CALORIE RESTRICTION

Until recently, the effects of CR on health and longevity have largely been believed to be universal. Recently, emerging evidence in mice as well as in nonhuman primates has challenged the notion that CR universally improves health and extends life span, and has implicated experimental variables such as species, sex, diet composition, and genetic background as important in the response to CR. Indeed, some studies have found that CR can even be detrimental for certain strains of mice. Using a panel of 41 recombinant inbred strains of mice placed on either an ad libitum or 40% CR regimen, it was recently shown that the majority of these strains did not have an extended life span in response to CR. Moreover, CR shortened the life span of a number of strains. These results highlight that genetic variation is an important determinant for the effects of CR on life span.

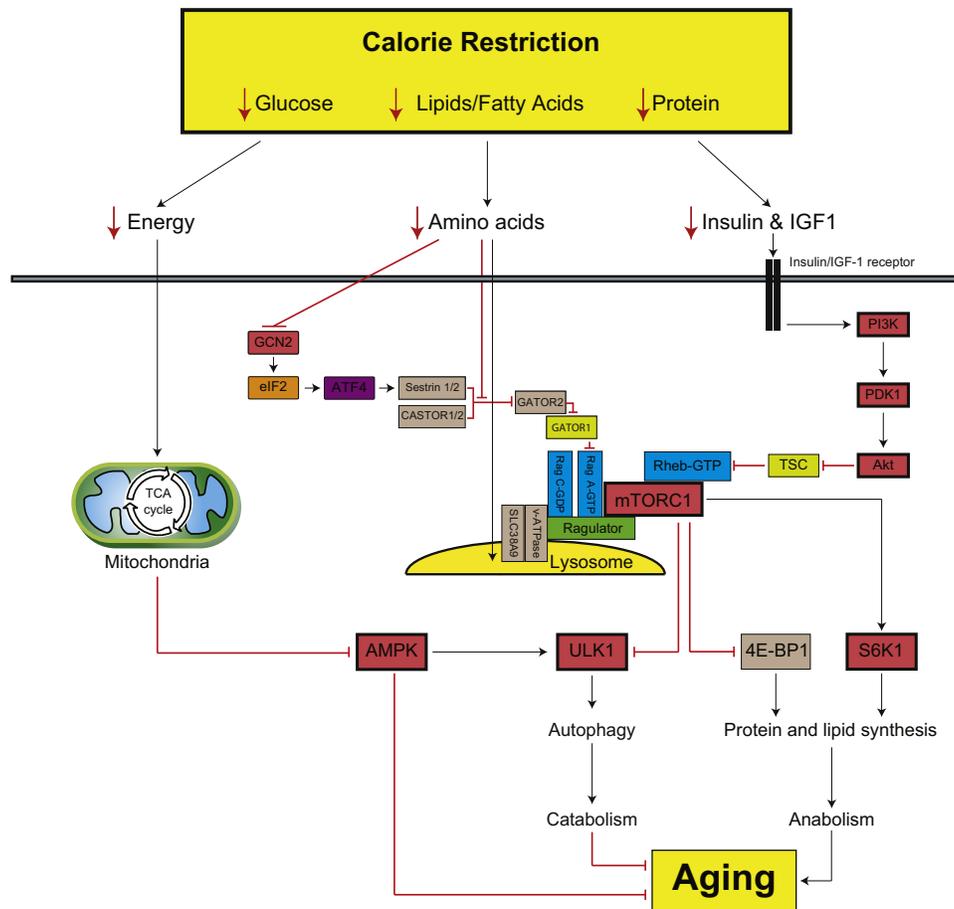


FIGURE 54.1 Key signaling pathways mediating the response to calorie and protein restriction. Dietary consumption of macronutrients (carbohydrates, lipids, and proteins) is reduced during calorie or protein restriction. CR or PR diets reduce the energy and amino acids available to cells, while reducing circulating levels of insulin and IGF-1. CR diets inhibit anabolic processes through inhibition of PI3K/Akt/mTORC1 signaling, while promoting autophagy through coordinated inhibition of mTORC1 and activation of AMPK. PR mimics many of the metabolic effects of CR, inhibiting mTORC1 signaling. Inhibition of mTORC1, activation of AMPK, and increased autophagy promote longevity and health.

To test this further, a recent study by Rafael de Cabo and colleagues compared the effects of 20% and 40% calorie restriction on C57BL/6J and DBA/2J mice of both sexes (Mitchell et al., 2016). Surprisingly, while both levels of restriction improved health span in all strains and sexes, life span was not similarly improved. Indeed, the “standard” 40% CR dietary regimen was less efficacious than a 20% CR regimen in C57BL/6J mice, with the 40% CR regimen failing to extend the life span of female mice at all. The efficacy of CR as a life span–extending dietary regimen is thus clearly dependent upon genetic background, and cannot clearly be assessed from improved health span.

The relevance of these results is further supported by recent life span and health span studies using nonhuman primates. CR studies in rhesus monkeys have consistently demonstrated decreased risk of developing age-related diseases and improved overall health in CR animals. However, while a study at UW-Madison reported a significant improvement in age-related and all-cause survival in monkeys on a CR diet, a similar study conducted at the National Institute on Aging (NIA) observed no difference in survival between groups. While many factors may have contributed to these results, including the CR paradigm used (Mitchell et al., 2019), the monkeys used in the NIA research came from a different population (subgroup) of rhesus monkeys than those at UW-Madison. In light of the mouse studies mentioned above, the substantial variation in genetic background may contribute to the altered response to a CR diet.

CONCLUSIONS

The optimal response to CR is determined by the intricate interaction between nutrition and genetics. Although the effects of CR on aging are not universal, accumulating evidence supports that conserved signaling pathways and mechanisms exist to regulate the response to CR. New discoveries suggest that the composition of diet, particularly, protein, plays important roles in determining the response to CR. Protein restriction, or specific amino acid restriction, have been shown to be as effective as CR in improving metabolic health. Mechanistically, GCN2, mTOR signaling and AMPK signaling have been demonstrated as regulators of metabolism with critical roles in the response to CR and PR. Ongoing research on CR, PR, and meal pattern distribution will help to reveal how genetic background regulates the response to alterations in dietary energy and macronutrient content, help to further illuminate the mechanisms by which diet regulates health and longevity, and suggest novel therapeutic targets for the treatment of age-related diseases (Figs. 54.1 and 54.2).

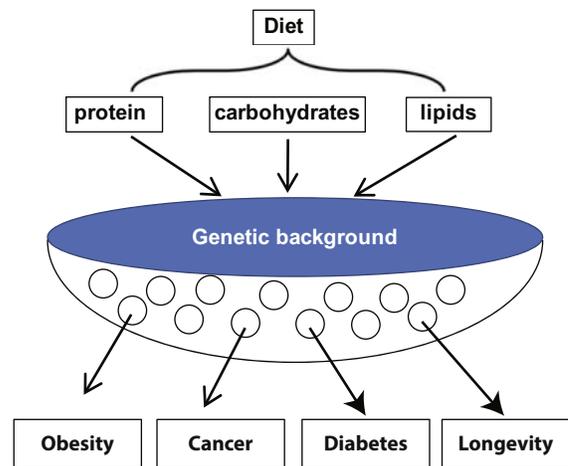


FIGURE 54.2 Genetic background mediates the effects of diet on metabolic health. The effects of dietary macronutrients on aging and age-related disease are mediated by genetic background, and thus while broad generalizations about the effects of dietary macronutrients and energy content on health and longevity can be made, there is likely to be considerable individual variability. Just as some individuals are more or less susceptible to weight gain, some individuals may benefit greatly, not benefit, or suffer negative consequences from dietary interventions such as calorie or protein restriction with regard to longevity and age-associated diseases.

References

- Brown-Borg, H.M., Buffenstein, R., 2017. Cutting back on the essentials: can manipulating intake of specific amino acids modulate health and lifespan? *Ageing Res Rev* 39, 87–95. <https://doi.org/10.1016/j.arr.2016.08.007>. Epub 2016 Aug 26.
- Cummings, et al., 2018. *J Physiol* 596 (4), 623–645. <https://doi.org/10.1113/JP275075>. Epub 2017 Dec 27.
- Dumas, S.N., Lamming, D.W., 2019. *J Gerontol A Biol Sci Med Sci* pii: glz056. doi: 10.1093/gerona/glz056. [Epub ahead of print].
- Fontana, L., Cummings, N.E., Arriola Apelo, S.L., Neuman, J.C., Kasza, I., Schmidt, B.A., Cava, E., Spelta, F., Tosti, V., Syed, F.A., Baar, E.L., Veronese, N., Cottrell, S.E., Fenske, R.J., Bertozzi, B., Brar, H.K., Pietka, T., Bullock, A.D., Figenshau, R.S., Andriole, G.L., Merrins, M.J., Alexander, C.M., Kimple, M.E., Lamming, D.W., 2016. Decreased consumption of branched-chain amino acids improves metabolic health. *Cell Rep* 16 (2), 520–530. <https://doi.org/10.1016/j.celrep.2016.05.092>. Epub 2016 Jun 23.
- Fontana, L., Partridge, L., 2015. Promoting health and longevity through diet: from model organisms to humans. *Cell* 161, 106–118.
- Goberdhan, D.C., Wilson, C., Harris, A.L., 2016 Apr 12. *Cell Metab* 23 (4), 580–589. <https://doi.org/10.1016/j.cmet.2016.03.013>.
- Green, C.L., Lamming, D.W., 2019. *Mech Ageing Dev* 177, 186–200. <https://doi.org/10.1016/j.mad.2018.07.004>. Epub 2018 Jul 22.
- Kennedy and Lamming, 2016. *Cell Metab* 23 (6), 990–1003. <https://doi.org/10.1016/j.cmet.2016.05.009>. (Accessed 14 June 2016).
- Lamming, D.W., Anderson, R.M., 2014. *Metabolic Effects of Caloric Restriction*, eLS. John Wiley & Sons, Ltd, Chichester.
- Lamming, D.W., Cummings, N.E., Rastelli, A.L., Gao, F., Cava, E., Bertozzi, B., Spelta, F., Pili, R., Fontana, L., 2015. Restriction of dietary protein decreases mTORC1 in tumors and somatic tissues of a tumor-bearing mouse xenograft model. *Oncotarget* 6, 31233–31240.
- Ma, S., Yim, S.H., Lee, S.G., Kim, E.B., Lee, S.R., Chang, K.T., Buffenstein, R., Lewis, K.N., Park, T.J., Miller, R.A., Clish, C.B., Gladyshev, V.N., 2015. Organization of the mammalian

- metabolome according to organ function, lineage specialization, and longevity. *Cell Metabol* 22, 332–343.
- Mahoney, et al., 2018. *Nat Commun* 9 (1), 548. <https://doi.org/10.1038/s41467-018-03035-z>.
- Mitchell, S.J., Madrigal-Matute, J., Scheibye-Knudsen, M., Fang, E., Aon, M., Gonzalez-Reyes, J.A., Cortassa, S., Kaushik, S., Gonzalez-Freire, M., Patel, B., Wahl, D., Ali, A., Calvo-Rubio, M., Buron, M.I., Guitierrez, V., Ward, T.M., Palacios, H.H., Cai, H., Frederick, D.W., Hine, C., Broeskamp, F., Habering, L., Dawson, J., Beasley, T.M., Wan, J., Ikeno, Y., Hubbard, G., Becker, K.G., Zhang, Y., Bohr, V.A., Longo, D.L., Navas, P., Ferrucci, L., Sinclair, D.A., Cohen, P., Egan, J.M., Mitchell, J.R., Baur, J.A., Allison, D.B., Anson, R.M., Villalba, J.M., Madeo, F., Cuervo, A.M., Pearson, K.J., Ingram, D.K., Bernier, M., de Cabo, R., 2016. Effects of sex, strain, and energy intake on hallmarks of aging in mice. *Cell Metabol* 23, 1093–1112.
- Mitchell, et al., 2019. *Cell Metab* 29 (1), 221–228. <https://doi.org/10.1016/j.cmet.2018.08.011> e3.
- Schreiber, et al., 2019. *Nat Commun* 10 (1), 3194. <https://doi.org/10.1038/s41467-019-11174-0>.
- Solon-Biet, S.M., McMahon, A.C., Ballard, J.W.O., Ruohonen, K., Wu, L.E., Cogger, V.C., Warren, A., Huang, X., Pichaud, N., Melvin, R.G., Gokarn, R., Khalil, M., Turner, N., Cooney, G.J., Sinclair, D.A., Raubenheimer, D., Le Couteur, D.G., Simpson, S.J., 2014. The ratio of macronutrients, not caloric intake, dictates cardiometabolic health, aging, and longevity in ad libitum-fed mice. *Cell Metabolism* 19, 418–430.
- Solon-Biet, S.M., Mitchell, S.J., Coogan, S.C., Cogger, V.C., Gokarn, R., McMahon, A.C., Raubenheimer, D., de Cabo, R., Simpson, S.J., Le Couteur, D.G., 2015. Dietary protein to carbohydrate ratio and caloric restriction: comparing metabolic outcomes in mice. *Cell Rep* 11, 1529–1534.
- Yang, L., Licastro, D., Cava, E., Veronese, N., Spelta, F., Rizza, W., Bertozzi, B., Villareal, D.T., Hotamisligil, G.S., Holloszy, J.O., Fontana, L., 2016. Long-term calorie restriction enhances cellular quality-control processes in human skeletal muscle. *Cell Rep* 14, 422–428.
- Yu, et al., 2018. *FASEB J* 32 (6), 3471–3482. <https://doi.org/10.1096/fj.201701211R>. Epub 2018 Jan 30.

This page intentionally left blank

Epigenetics in Food Allergies: The Missing Piece of the Puzzle

Shimrit Bar-El Dadon, Ram Reifen

The Center for Nutrigenomics and Functional Foods, The Robert H. Smith Faculty of Agriculture, Food and Environment,
The Hebrew University of Jerusalem, Israel

BACKGROUND

The “first wave” of allergic diseases (asthma and allergic rhinitis) started to gain momentum more than 50 years ago. Another wave of allergies—*food allergies*—has been slowly building behind it. This “second wave” of allergies is now emerging as an epidemic with vast and considerable implications, concerning not only the patients but also their families, schools, the food industry, clinicians, and governmental offices. There are an estimated 90,000 emergency room visits related to food allergies every year in the United States, and more than \$25 billion are spent on food allergy care. At the same time, the prevalence of food allergies continues to rise at alarmingly rapid rates, particularly in Westernized and economically developed countries. Currently, it is estimated that food allergies affect nearly 5% of all adults and 8%–10% of all children (Sicherer and Sampson, 2014). However, while there is solid evidence for its increasing prevalence, there are few clues as to *why* this is happening. A possible answer to this intriguing question might be found in the complex interplay between environmental factors and genetic risk, also termed “gene-by-environment” (G×E) interactions, which are partially mediated by epigenetic mechanisms.

In this chapter, we highlight the latest advances pertaining to the interrelationships of genetics, environment, and epigenomic alterations in the development of food allergies, which may offer future perspectives for diagnosis, prevention, and management of this new epidemic.

FOOD ALLERGY AND FOOD INTOLERANCE

Ingestion of food represents the greatest foreign antigenic load faced by the human immune system. Most individuals develop a tolerance to food antigens, which are constantly gaining access to the body. However, when tolerance fails to develop, the immune system responds with a hypersensitivity reaction (Valenta et al., 2015).

Food allergy is an immunologically mediated, reproducible adverse reaction to a food substance, generally involving two classic mechanisms—immunoglobulin (IgE) and non-IgE-mediated reactions—with symptoms involving the skin, gastrointestinal tract, respiratory tract, and cardiovascular system (Berin and Sampson, 2013). Virtually any food can cause an allergic reaction, yet 90% of the recognized food allergies are thought to be caused by eight food groups, referred to as “the big eight” including eggs, milk, fish, shellfish, peanuts, soybeans, tree nuts, and wheat. Remarkably, some of the big eight, as well as some other food allergens, are known to cause food allergies frequently in particular geographic regions (Sicherer and Sampson, 2014; McClain et al., 2014). Peanut allergy, for instance, is more frequent in North America than in other parts of the world (Sicherer, 2011), while allergy to chickpea is more prevalent in the Mediterranean region (Bar-El Dadon et al., 2013).

The cross-reactivity phenomenon is yet another factor contributing to the complexity of food allergies. IgE cross-reactivity is the ability of IgEs raised against one allergen to bind or recognize a similar or another homologous allergen originating from another source; some allergens may share common epitopes or similar

sequences and/or structural features, resulting in a possible cross-reactivity. The clinical relevance of cross-reacting antibodies varies due to the individual reactivity of the patient and due to the structural similarity of the allergens in question with the original sensitizing allergen. For example, chickpea allergy is associated with lentil and/or pea allergy, but evidently may not present independently. As previously described, chickpeas may not be the trigger for allergy per se but rather bind to IgEs originally raised against “incriminating” molecules that appear in pea and lentil.

Some reports confirm that food allergy rates are now higher than ever before (Sicherer and Sampson, 2014). Another intriguing observation is the change in the disease profile over the last 10 years, which is expressed by an earlier and a more severe failure of oral tolerance during the first year of life, as well as a decreased likelihood of outgrowing food allergies (Savage et al., 2007; Skripak et al., 2007). These changes in allergy pathogenesis raise a series of critical questions: Why are we not all allergic? How might propensity to allergies be modified across generations? What is the role of environmental factors in food allergy development, and are they more likely to affect those who are genetically at risk?

Various environmental, nutritional, and lifestyle factors contribute to allergy development, including declining microbial exposure, Western diet, and air pollutants (Campbell et al., 2015; Harb and Renz, 2015; Martino and Prescott, 2011). Yet, only a small number of studies have linked these factors to the development of a food allergy (Hong et al., 2009; Lack, 2012), and in any case, they fail to explain the changing patterns of food allergy severity and persistence. Thus, the significant questions remain unanswered.

GxE INTERACTIONS AND THE DEVELOPMENT OF FOOD ALLERGIES

While genetics may play an important role in food allergy development, as revealed through familial aggregation studies, heritability estimates (ranging from 15% to 82%), and candidate genes (Hong et al., 2009; Alexander et al., 2014; Tsai et al., 2009), the rapidly growing prevalence of allergic diseases cannot be explained in terms of Mendelian inheritance. This means that the rising incidence is more rapid than changes to the genome sequence would allow (Neeland et al., 2015; Tan et al., 2012). In addition, the prevalence of food allergies is assumed to be significantly lower in developing countries; however, migrants from those countries arriving to developed countries are not protected from developing food allergies (Boye, 2012; Leung et al., 1994). Taken together, the causes of food allergies are multifactorial, and are likely to involve many genes in combination

with modern lifestyle and environmental factors (Neeland et al., 2015; Tan et al., 2012). These GxE interactions reflect the complex interplay between genetic predisposition and environmental exposure.

EPIGENETICS—MEDIATORS OF THE GXE INTERACTION

Nowadays, there has been a growing recognition of epigenetic mechanisms being potential mediators of GxE interactions, as they are critical for normal human development, and specifically regulate gene expression during *immune* development (Paparo et al., 2014; Feinberg, 2007; Feng et al., 2010). Epigenetic mechanisms are all potentially affected by environmental changes (including microbial exposure, nutrition, tobacco smoke, and air pollutants), and can remain very stable across life spans (Feil and Fraga, 2011). Importantly, epigenetic marks were originally thought to be erased upon passage through the germline. However, a huge paradigm shift occurred more than a decade ago, leading to the concept of *transgenerational epigenetic inheritance* (Skinner and Guerrero-Bosagna, 2009). This concept provides a mechanism by which parents can transfer information about the environmental conditions they have encountered to their offspring (Uller and Tollefsbol, 2014). In this way, the epigenetic code provides plasticity of gene expression in response to environmental changes, allowing more rapid phenotypic adaptations across generations than those associated with Darwinian selection. Accordingly, when maladaptive changes, due to adverse environmental exposures, for example, occur and are inherited to by future generations, the risk of developing diseases increases, as demonstrated in animal models of allergic diseases (Hollingsworth et al., 2008).

Given its involvement in GxE interactions, the epigenome may represent the “missing piece” in the etiological puzzle for allergic diseases in general, and food allergies in particular (Hong and Wang, 2014).

EPIGENETICS IN ALLERGIC DISEASES AND FOOD ALLERGIES

Epigenetics refers to the study of changes in gene-expression patterns independent of changes in the underlying sequence. The epigenetic program is primarily governed by DNA methylation and specific histone-modification patterns, which determine the degree of DNA compaction and thus the accessibility of genes to transcription (Bird, 2007). Specifically, DNA methylation occurs predominantly at CpG islands and in repetitive genomic sequence regions, usually repressing transcription (Chen and Riggs, 2011). Histones undergo

posttranslational modifications that alter their interaction with DNA and nuclear proteins (Bonifer and Cockrill, 2011). In addition, microRNAs are also considered to be regulators of gene expression (Sato et al., 2011).

Convincing experimental evidence suggests that epigenetic marks serve as a memory of exposure in early life to inadequate or inappropriate nutritional factors (which will be discussed later). Moreover, epigenetic alterations occur not only prenatally or shortly after birth, but also during later developmental periods (Simmons, 2011). Together, these marks induce long-term changes in gene expression throughout life, potentially leading to disease development.

There is now conclusive data for epigenetic regulation of immune system development and function (Fernandez-Morera et al., 2010; Lal and Bromberg, 2009). These mechanisms are known to control differentiation of both T helper type 1 and 2 (Th1 and Th2, respectively) cells, and are also prerequisites for forkhead box P3 (FoxP3) expression and regulatory T cell (Treg) differentiation. For example, the Th1/Th2 balance associated with allergic diseases is maintained by epigenetics (Fields et al., 2002; White et al., 2002). During pregnancy, the maternal immune system normally adapts to a “Th2 state” concomitant with downregulation of the Th1 response to prevent a reaction to fetal antigens (Wegmann et al., 1993). The neonate reflects these maternal events with underlying epigenetic changes that reduce Th1 and increase Th2 function (Zaghouani et al., 2009). Postnatally, Th2 dominance is normally suppressed along with progressive Th1 maturation, accompanied by progressive demethylation within the interferon gamma (IFN γ) promoter and upregulated IFN γ transcription (Prescott et al., 1999; White et al., 2006). In allergic diseases, however, differences in neonatal T-cell function as well as postnatal maturation patterns are present, including immaturity of Th1 function (Prescott, 2003) and attenuated Treg function (Smith et al., 2008). These abnormalities lead to an uncontrolled Th2 response (Th2 bias) associated with attenuated Th1 maturation in early childhood. Thus, altered epigenetic mechanisms may disrupt normal development and lead to dysregulation of the immune system associated with allergic disease phenotypes (Hong and Wang, 2014; Lovinsky-Desir and Miller, 2012; Prescott and Saffery, 2011; Song et al., 2014). In light of the well-observed phenomenon of “atopic march” and comorbidity of food allergy and other atopic diseases, studies exploring DNA methylation in allergic diseases might shed light on how these mechanisms are involved in the development of food allergies as well.

A *direct* link between food allergy and epigenetics has so far been described in a small number of studies. In a mouse model of peanut allergy, Song et al. (2014) demonstrated a predisposition of offspring of allergic

mothers to developing peanut allergy, which is associated with epigenetic alteration of the interleukin (IL) 4 gene promoter. Later, evidence of the involvement of epigenetics in food allergies emerged in human studies as well. Martino et al. (2012) demonstrated that a dysregulated neonatal CD4+ T-cell activation response is a feature of pediatric food allergy, involving altered expression signaling via the TNF–NF– κ B pathway. Hence, CD4+ T cells are key drivers of the allergic response, and may therefore harbor epigenetic profiles. In their follow-up study, Martino et al. (Martino et al., 2014) examined whether pre-existing differences in DNA methylation might underscore suboptimal CD4+ T-cell gene expression. Indeed, they identified stable changes in DNA methylation at genes involved in MAP kinase signaling (a pro-inflammatory pathway), which were present in CD4+ T cells at birth, prior to development of a food allergy. In addition, as current diagnostic tests are excellent markers of sensitization but poor predictors of clinical reactivity, the same group of researchers sought to discover genomic biomarkers of clinical food allergies that might be useful for predicting food-challenge outcomes. Using genome-wide DNA methylation profiling, they found an association between clinical food allergy and DNA methylation in MAP kinase signaling, and provided proof of principle for the use of DNA methylation biomarkers in novel diagnostic testing for patients with food allergies (Martino et al., 2015).

The mechanism of sustained tolerance probably involves the development of Tregs and a shift from the Th2 response toward a Th1 response with increased IFN γ production (Berin and Mayer, 2013). Syed et al. (2014) studied epigenetic changes associated with clinical immune tolerance in patients undergoing oral immunotherapy (OIT) for peanut allergy. In their study, DNA-methylation levels of *FoxP3* in Tregs of patients who remained unresponsive to peanuts following 3 months of OIT were significantly lower than in patients who were re-sensitized (Syed et al., 2014). In addition, patients in the control group who did not undergo OIT had higher DNA-methylation levels than in the experimental groups. These findings suggest that epigenetic modifications of an antigen-induced T-cell subset might predict clinical immune tolerance during peanut OIT (Syed et al., 2014). In a study from 2015, acquisition of tolerance in children with an allergy to cow’s milk was investigated by Berni Canani et al. (2015). The analysis revealed that tolerance acquisition characterized by a distinct DNA-methylation pattern in Th1 (IL-10, IFN γ) and Th2 (IL-4, IL-5) cytokine genes (Berni Canani et al., 2015). Specifically, DNA methylation of IL-4 and IL-5 was lower, and that of IL-10 and IFN γ was higher in patients who were actively allergic to cow’s milk than in those who had acquired tolerance (Berni Canani et al., 2015).

Epigenetic mechanisms by which genetic variants affect food allergies are largely unexplored. The recent development of high-throughput genotyping platforms has allowed the discovery of genetic variants associated with specific traits in a blind, systematic yet cost-effective approach, by genome-scan or genome-wide association study (GWAS). [Hong et al. \(2015\)](#) conducted the first GWAS of food allergies in a U.S. cohort of children and their biological parents. First, the researchers detected genetic associations with peanut allergy, and identified peanut allergy-specific single nucleotide polymorphisms (SNPs) in the human leukocyte antigen (HLA)-DQ and DR regions at 6p21.32 (rs7192 and rs9275596). Next, they found that both SNPs are significantly associated with differential DNA methylation at multiple CpG sites, and that differential DNA methylation of the HLA-DQB1 and HLA-DRB1 genes partially mediated the identified SNP–peanut allergy association. These results were the first to demonstrate the key role of differential DNA methylation in mediating identified genetic risk factors for peanut allergy ([Hong et al., 2015](#)).

Collectively, these studies suggest that epigenetic factors, and specifically DNA methylation, regulate the expression of key immune genes, and thus play an important role in food allergy pathogenesis ([Fig. 55.1](#)).

EARLY LIFE, ALLERGIES AND EPIGENETICS

Early life is clearly a time when immune deviation and atopy are determined. As such, it is a significant window for the development of food allergies as well, supported by the fact that most cases of food allergy develop in the first few years of life. The prenatal period is also believed to be a critical window for programming, since the placenta and the fetus are both vulnerable to exogenous and endogenous maternal influences. This notion is supported by recent findings on the relationships between maternal environmental exposures, such as smoking, diet, and microbial exposure during pregnancy, and the risk of food allergies in offspring ([Furuhjelm et al., 2011](#); [Granel et al., 2008](#); [Kulig et al., 1999](#)). Epigenetics, and DNA methylation in particular, have been proposed as mechanisms for understanding the “developmental origins” of complex modern diseases such as allergies ([Shaheen and Adcock, 2009](#); [Waterland and Michels, 2007](#)). Although most studies focus on allergic reactions in general (asthma, eczema), inference can be made from them to food allergies due to the strong association between these processes.

Recent studies have shown that prenatal environmental exposures can epigenetically activate or silence

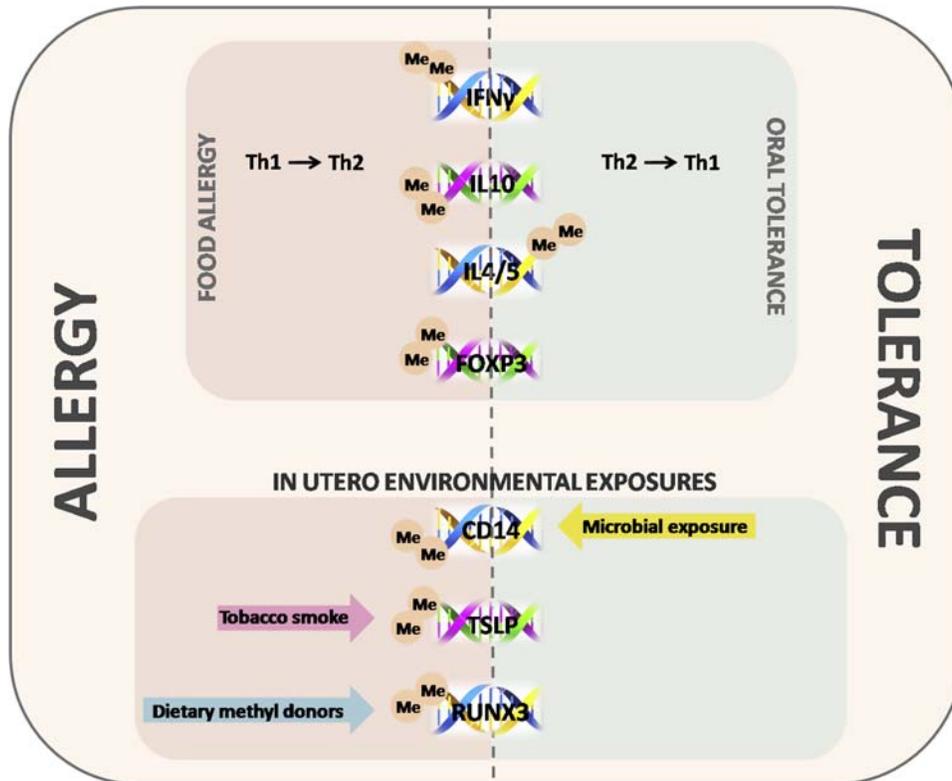


FIGURE 55.1 Regulation of gene expression of key immune genes by epigenetic factors, and its role in food allergy pathogenesis.

immune-related genes with substantial effects on immune programming, which could explain the early life origin of allergic diseases. For example, TSLP 5'-CGI methylation level was significantly associated with prenatal smoking exposure and may mediate the effects of prenatal tobacco smoke exposure on atopic dermatitis (Wang et al., 2013). In humans, exposure to a high-microbial burden in rural farming environments is protective against allergic diseases (Depner et al., 2013). This may be partly mediated by DNA methylation, as demonstrated by Slaats et al. (2012), who suggested that the protective effect of the farm environment against allergy development may be mediated in utero via hypomethylation of CD14 in the placenta. Dietary methyl donors and cofactors, such as folic acid, vitamin B12, vitamin B6, and zinc, are necessary for one-carbon metabolism, which provides the methyl group for DNA methylation. In utero supplementation of mice with a methyl donor altered DNA methylation of various genes, among them runt-related transcription factor 3, which may negatively regulate allergic airway disease (Hollingsworth et al., 2008). Dietary fat could also play a role, as suggested by Lee et al. (2013), who showed that docosahexaenoic acid supplementation for pregnant women from midgestation until delivery could significantly affect the methylation levels of genes involved in immune function in cord blood mononuclear cells.

Although most food allergies develop during critical periods in the first few years of life, most food allergy studies have focused on older age groups, after the allergy has developed. Identification of critical epigenetic marks associated with environmental exposures implicated in food allergy at critical developmental time points—preconception, in utero, and early childhood—would provide strong evidence for clinical and public health preventive measures against food allergies, such as dietary intake and breast feeding (Hong and Wang, 2014). A better understanding of epigenetic mechanisms in the development of food allergies during critical time windows offers the potential to gain new mechanistic insight and identify new diagnostic, prognostic, and therapeutic targets for food allergies.

CONCLUDING REMARKS AND FUTURE DIRECTIONS

The new insights into the importance of epigenetically regulated allergy-related pathways open up exciting new avenues of investigation for the causes of, and mechanisms governing, food allergies. Moreover, they open doors for new approaches to the treatment and prevention of allergic diseases. However, most epigenetic studies on allergic diseases have focused on

limited pathways and several candidate genes. Thus, there is a need for more GWAS that could reveal truly novel genes, especially those associated with moderate disease risks. These GWAS would not only shed light on the future genetic research of food allergy but also highlight their promise in dissecting the genetics of food allergy.

In addition, randomized, controlled intervention studies to determine the relative causal effects of environmental risk factors on the incidence of food allergies should be performed. However, given the complexity of food allergies, addressing one or two factors at a time will not provide a full understanding of the etiology and biological mechanisms underlying the development and progression of those allergies. Therefore, there is a dire need for large prospective birth cohorts that will enable a simultaneous study of genetic susceptibility and environmental exposures, and their involvement in the development of food allergies through the epigenome. This could eventually lead to early identification of allergic individuals as well as prediction of clinical tolerance versus reactivity.

In the same way, any epigenetic intervention might have a different effect on allergy outcome, depending on both genetic and environmental factors. As is the case with other complex conditions, the long route to allergy resolution must pass through a series of studies examining different populations using different regimens under different conditions, hopefully ending in an “individualized” prevention scheme for food allergies.

The prevalence of food allergies has risen substantially over the past decade, following in the footsteps of rises in other atopic conditions. This growing clinical and public health problem is thought to be reaching epidemic proportions in light of its increasing prevalence and potential fatality.

Although our understanding of food allergies has progressed, knowledge about the underlying mechanisms and causes of this puzzling epidemic remains limited. Few specific genes have been conclusively associated with food allergy, leaving its heritability largely unexplained. In addition, several lifestyle risk factors have been implicated in the development of food allergies. However, genes or environmental factors alone cannot explain the dramatic rise in these allergies, which is likely a result of the complex interplay of GxE interactions.

Epigenetic modification acts as a mediator for environmental influences on gene expression and a modulator of disease risk associated with genetic variation. As such, epigenetics is an exciting new field in food allergy research. New studies, reviewed here, on epigenetics in food allergies have shed new light on the pathogenesis of this complex disease, with regard to not only GxE interactions but also the model of

inheritance and its epidemiological implications. Therefore, epigenetic mechanisms, such as DNA methylation, have been proposed to understand the “developmental origins” of complex modern diseases such as allergies.

Unraveling genetic and epigenetic involvement in GxE interactions is crucial not only for understanding mechanisms of food allergy but also as an opportunity to provide a more personalized approach to public health recommendations. As some epigenetic alterations are reversible, this holds promise for novel interventions that may restore the “normal epigenome” by activating or silencing disease-related genes, thus reaching the ultimate goal of arresting, or even preventing, the allergy epidemic.

References

- Alexander, E.S., Martin, L.J., Collins, M.H., Kottyan, L.C., Sucharew, H., He, H., et al., 2014. Twin and family studies reveal strong environmental and weaker genetic cues explaining heritability of eosinophilic esophagitis. *J Allergy Clin Immunol* 134 (5), 1084–1092 e1081.
- Bar-El Dadon, S., Pascual, C.Y., Eshel, D., Teper-Bamnolker, P., Ibanez, M.D., Reifen, R., 2013. Vicilin and the basic subunit of legumin are putative chickpea allergens. *Food Chem* 138 (1), 13–18.
- Berin, M.C., Mayer, L., 2013. Can we produce true tolerance in patients with food allergy? *J Allergy Clin Immunol* 131 (1), 14–22.
- Berin, M.C., Sampson, H.A., 2013. Food allergy: an enigmatic epidemic. *Trends Immunol* 34 (8), 390–397.
- Berni Canani, R., Paparo, L., Nocerino, R., Cosenza, L., Pezzella, V., Di Costanzo, M., et al., 2015. Differences in DNA methylation profile of Th1 and Th2 cytokine genes are associated with tolerance acquisition in children with IgE-mediated cow’s milk allergy. *Clin Epigenet* 7 (1), 38.
- Bird, A., 2007. Perceptions of epigenetics. *Nature* 447 (7143), 396–398.
- Bonifer, C., Cockerill, P.N., 2011. Chromatin mechanisms regulating gene expression in health and disease. *Adv Exp Med Biol* 711, 12–25.
- Boye, J.I., 2012. Food allergies in developing and emerging economies: need for comprehensive data on prevalence rates. *Clin Transl Allergy* 2 (1), 25.
- Campbell, D.E., Boyle, R.J., Thornton, C.A., Prescott, S.L., 2015. Mechanisms of allergic disease – environmental and genetic determinants for the development of allergy. *Clin Exp Allergy* 45 (5), 844–858.
- Chen, Z.X., Riggs, A.D., 2011. DNA methylation and demethylation in mammals. *J Biol Chem* 286 (21), 18347–18353.
- Depner, M., Ege, M.J., Genuneit, J., Pekkanen, J., Roponen, M., Hirvonen, M.R., et al., 2013. Atopic sensitization in the first year of life. *J Allergy Clin Immunol* 131 (3), 781–788.
- Feil, R., Fraga, M.F., 2011. Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet* 13 (2), 97–109.
- Feinberg, A.P., 2007. Phenotypic plasticity and the epigenetics of human disease. *Nature* 447 (7143), 433–440.
- Feng, S., Jacobsen, S.E., Reik, W., 2010. Epigenetic reprogramming in plant and animal development. *Science* 330 (6004), 622–627.
- Fernandez-Morera, J.L., Calvanese, V., Rodriguez-Rodero, S., Menendez-Torre, E., Fraga, M.F., 2010. Epigenetic regulation of the immune system in health and disease. *Tissue Antigens* 76 (6), 431–439.
- Fields, P.E., Kim, S.T., Flavell, R.A., 2002. Cutting edge: changes in histone acetylation at the IL-4 and IFN-gamma loci accompany Th1/Th2 differentiation. *J Immunol* 169 (2), 647–650.
- Furuhjelm, C., Warstedt, K., Fageras, M., Falth-Magnusson, K., Larsson, J., Fredriksson, M., et al., 2011. Allergic disease in infants up to 2 years of age in relation to plasma omega-3 fatty acids and maternal fish oil supplementation in pregnancy and lactation. *Pediatr Allergy Immunol* 22 (5), 505–514.
- Granell, R., Heron, J., Lewis, S., Davey Smith, G., Sterne, J.A., Henderson, J., 2008. The association between mother and child MTHFR C677T polymorphisms, dietary folate intake and childhood atopy in a population-based, longitudinal birth cohort. *Clin Exp Allergy* 38 (2), 320–328.
- Harb, H., Renz, H., 2015. Update on epigenetics in allergic disease. *J Allergy Clin Immunol* 135 (1), 15–24.
- Hollingsworth, J.W., Maruoka, S., Boon, K., Garantziotis, S., Li, Z., Tomfohr, J., et al., 2008. In utero supplementation with methyl donors enhances allergic airway disease in mice. *J Clin Invest* 118 (10), 3462–3469.
- Hong, X., Wang, X., 2014. Epigenetics and development of food allergy (FA) in early childhood. *Curr Allergy Asthma Rep* 14 (9), 460.
- Hong, X., Tsai, H.J., Wang, X., 2009. Genetics of food allergy. *Curr Opin Pediatr* 21 (6), 770–776.
- Hong, X., Hao, K., Ladd-Acosta, C., Hansen, K.D., Tsai, H.J., Liu, X., et al., 2015. Genome-wide association study identifies peanut allergy-specific loci and evidence of epigenetic mediation in US children. *Nat Commun* 6, 6304.
- Kulig, M., Luck, W., Lau, S., Niggemann, B., Bergmann, R., Klettke, U., et al., 1999. Effect of pre- and postnatal tobacco smoke exposure on specific sensitization to food and inhalant allergens during the first 3 years of life. Multicenter Allergy Study Group, Germany. *Allergy* 54 (3), 220–228.
- Lack, G., 2012. Update on risk factors for food allergy. *J Allergy Clin Immunol* 129 (5), 1187–1197.
- Lal, G., Bromberg, J.S., 2009. Epigenetic mechanisms of regulation of Foxp3 expression. *Blood* 114 (18), 3727–3735.
- Lee, H.S., Barraza-Villarreal, A., Hernandez-Vargas, H., Sly, P.D., Biessy, C., Ramakrishnan, U., et al., 2013. Modulation of DNA methylation states and infant immune system by dietary supplementation with omega-3 PUFA during pregnancy in an intervention study. *Am J Clin Nutr* 98 (2), 480–487.
- Leung, R.C., Carlin, J.B., Burdon, J.G., Czarny, D., 1994. Asthma, allergy and atopy in Asian immigrants in Melbourne. *Med J Aust* 161 (7), 418–425.
- Lovinsky-Desir, S., Miller, R.L., 2012. Epigenetics, asthma, and allergic diseases: a review of the latest advancements. *Curr Allergy Asthma Rep* 12 (3), 211–220.
- Martino, D., Prescott, S., 2011. Epigenetics and prenatal influences on asthma and allergic airways disease. *Chest* 139 (3), 640–647.
- Martino, D.J., Bosco, A., McKenna, K.L., Hollams, E., Mok, D., Holt, P.G., et al., 2012. T-cell activation genes differentially expressed at birth in CD4⁺ T-cells from children who develop IgE food allergy. *Allergy* 67 (2), 191–200.
- Martino, D., Joo, J.E., Sexton-Oates, A., Dang, T., Allen, K., Saffery, R., et al., 2014. Epigenome-wide association study reveals longitudinally stable DNA methylation differences in CD4⁺ T cells from children with IgE-mediated food allergy. *Epigenetics* 9 (7), 998–1006.
- Martino, D., Dang, T., Sexton-Oates, A., Prescott, S., Tang, M.L., Dharmage, S., et al., 2015. Blood DNA methylation biomarkers predict clinical reactivity in food-sensitized infants. *J Allergy Clin Immunol* 135 (5), 1319–1328 e1311-1312.
- McClain, S., Bowman, C., Fernandez-Rivas, M., Ladics, G.S., Ree, R., 2014. Allergic sensitization: food- and protein-related factors. *Clin Transl Allergy* 4 (1), 11.
- Neeland, M.R., Martino, D.J., Allen, K.J., 2015. The role of gene-environment interactions in the development of food allergy. *Expert Rev Gastroenterol Hepatol* 9 (11), 1371–1378.
- Paparo, L., di Costanzo, M., di Scala, C., Cosenza, L., Leone, L., Nocerino, R., et al., 2014. The influence of early life nutrition on

- epigenetic regulatory mechanisms of the immune system. *Nutrients* 6 (11), 4706–4719.
- Prescott, S.L., 2003. Early origins of allergic disease: a review of processes and influences during early immune development. *Curr Opin Allergy Clin Immunol* 3 (2), 125–132.
- Prescott, S., Saffery, R., 2011. The role of epigenetic dysregulation in the epidemic of allergic disease. *Clin Epigenet* 2 (2), 223–232.
- Prescott, S.L., Macaubas, C., Smallacombe, T., Holt, B.J., Sly, P.D., Holt, P.G., 1999. Development of allergen-specific T-cell memory in atopic and normal children. *Lancet* 353 (9148), 196–200.
- Sato, F., Tsuchiya, S., Meltzer, S.J., Shimizu, K., 2011. MicroRNAs and epigenetics. *FEBS J* 278 (10), 1598–1609.
- Savage, J.H., Limb, S.L., Brereton, N.H., Wood, R.A., 2007. The natural history of peanut allergy: extending our knowledge beyond childhood. *J Allergy Clin Immunol* 120 (3), 717–719.
- Shaheen, S.O., Adcock, I.M., 2009. The developmental origins of asthma: does epigenetics hold the key? *Am J Respir Crit Care Med* 180 (8), 690–691.
- Sicherer, S.H., 2011. Epidemiology of food allergy. *J Allergy Clin Immunol* 127 (3), 594–602.
- Sicherer, S.H., Sampson, H.A., 2014. Food allergy: epidemiology, pathogenesis, diagnosis, and treatment. *J Allergy Clin Immunol* 133 (2), 291–307 quiz 308.
- Simmons, R., 2011. Epigenetics and maternal nutrition: nature v. nurture. *Proc Nutr Soc* 70 (1), 73–81.
- Skinner, M.K., Guerrero-Bosagna, C., 2009. Environmental signals and transgenerational epigenetics. *Epigenomics* 1 (1), 111–117.
- Skripak, J.M., Matsui, E.C., Mudd, K., Wood, R.A., 2007. The natural history of IgE-mediated cow's milk allergy. *J Allergy Clin Immunol* 120 (5), 1172–1177.
- Slaats, G.G., Reinius, L.E., Alm, J., Kere, J., Scheynius, A., Joerink, M., 2012. DNA methylation levels within the CD14 promoter region are lower in placentas of mothers living on a farm. *Allergy* 67 (7), 895–903.
- Smith, M., Tourigny, M.R., Noakes, P., Thornton, C.A., Tulic, M.K., Prescott, S.L., 2008. Children with egg allergy have evidence of reduced neonatal CD4(+)CD25(+)CD127(lo/-) regulatory T cell function. *J Allergy Clin Immunol* 121 (6), 1460–1466, 1466 e1461-1467.
- Song, Y., Liu, C., Hui, Y., Srivastava, K., Zhou, Z., Chen, J., et al., 2014. Maternal allergy increases susceptibility to offspring allergy in association with TH2-biased epigenetic alterations in a mouse model of peanut allergy. *J Allergy Clin Immunol* 134 (6), 1339–1345 e1337.
- Syed, A., Garcia, M.A., Lyu, S.C., Bucayu, R., Kohli, A., Ishida, S., et al., 2014. Peanut oral immunotherapy results in increased antigen-induced regulatory T-cell function and hypomethylation of forkhead box protein 3 (FOXP3). *J Allergy Clin Immunol* 133 (2), 500–510.
- Tan, T.H., Ellis, J.A., Saffery, R., Allen, K.J., 2012. The role of genetics and environment in the rise of childhood food allergy. *Clin Exp Allergy* 42 (1), 20–29.
- Tsai, H.J., Kumar, R., Pongratic, J., Liu, X., Story, R., Yu, Y., et al., 2009. Familial aggregation of food allergy and sensitization to food allergens: a family-based study. *Clin Exp Allergy* 39 (1), 101–109.
- Uller, T., 2014. Evolutionary perspectives on transgenerational epigenetics. In: Tollefsbol, T. (Ed.), *Transgenerational Epigenetics Evidence and Debate*, 1 ed. San Diego Academic Press (Elsevier), London, Oxford, Boston, New York, pp. 175–185.
- Valenta, R., Hochwallner, H., Linhart, B., Pahr, S., 2015. Food allergies: the basics. *Gastroenterology* 148 (6), 1120–1131 e1124.
- Wang, I.J., Chen, S.L., Lu, T.P., Chuang, E.Y., Chen, P.C., 2013. Prenatal smoke exposure, DNA methylation, and childhood atopic dermatitis. *Clin Exp Allergy* 43 (5), 535–543.
- Waterland, R.A., Michels, K.B., 2007. Epigenetic epidemiology of the developmental origins hypothesis. *Annu Rev Nutr* 27, 363–388.
- Wegmann, T.G., Lin, H., Guilbert, L., Mosmann, T.R., 1993. Bidirectional cytokine interactions in the maternal-fetal relationship: is successful pregnancy a TH2 phenomenon? *Immunol Today* 14 (7), 353–356.
- White, G.P., Watt, P.M., Holt, B.J., Holt, P.G., 2002. Differential patterns of methylation of the IFN-gamma promoter at CpG and non-CpG sites underlie differences in IFN-gamma gene expression between human neonatal and adult CD45RO- T cells. *J Immunol* 168 (6), 2820–2827.
- White, G.P., Hollams, E.M., Yerkovich, S.T., Bosco, A., Holt, B.J., Bassami, M.R., et al., 2006. CpG methylation patterns in the IFN-gamma promoter in naive T cells: variations during Th1 and Th2 differentiation and between atopics and non-atopics. *Pediatr Allergy Immunol* 17 (8), 557–564.
- Zaghouani, H., Hoeman, C.M., Adkins, B., 2009. Neonatal immunity: faulty T-helpers and the shortcomings of dendritic cells. *Trends Immunol* 30 (12), 585–591.

This page intentionally left blank

Gut Microbiota and Their Influence on Response to Food

Clarisse Marotz¹, Steven Brown², Amir Zarrinpar^{2,3}

¹Department of Pediatrics, UC San Diego, La Jolla, CA, United States; ²Division of Gastroenterology and Hepatology, UC San Diego, La Jolla, CA, United States; ³VA San Diego Health Sciences, La Jolla, CA, United States

Glossary

- Antimicrobial peptides (AMPs)** a class of natural and synthetic peptides with a wide spectrum of microbial targets
- Chyme** partially digested food bolus passing from the stomach into the small intestine
- Germ-free mouse model** mice with no detectable microbiota
- Gnotobiotic mouse model** mice with a known microbiota
- Inflammasome** cytoplasmic protein complexes that activate proinflammatory cytokines in response to PAMPs
- Microbiome** the collective genomes of the microbiota
- Microbiota** a collection of microorganisms native to a particular environment (e.g., human gut)
- Pathogen-associated molecular patterns (PAMPs)** molecular moieties distinctive to specific pathogens
- Secondary bile acid** microbially modified (i.e., hydrolyzed) primary bile acids
- Short chain fatty acid (SCFA)** fatty acids with less than six carbon atoms (e.g., butyric acid, propionic acid, acetic acid) which, in the intestines, result from microbial fermentation of complex carbohydrates

INTRODUCTION

Gut microbiota are involved in nearly all aspects of human health. The relation between a microbe and its host depends on their shared food sources (Wu et al., 2015; Kau et al., 2011). When describing the bacterial members that reside in the mammalian lumen, the term “microbiota” refers to nonpathogenic commensal organisms, and the term “microbiome” refers not only to the microbiota but also their collective genomes. The gut microbiota includes organisms from the bacterial, protozoan, fungal, and other eukaryotic (e.g., amoeba) species as well as viruses and archaea species. Bacteria dominate the biomass and genes of the microbiome and have been the main focus of most studies.

Microbes inhabit food boluses (or chyme), intestinal crypts, and the mucosal layer that separates the lumen from the epithelium, all of which harbor communities with unique member profiles adapted to their particular niche. Infectious agents of all these microbiota, particularly bacteriophages, are also present (Dutilh et al., 2014). Besides inducing lysis of bacteria and releasing inflammatory molecules (e.g., lipopolysaccharides) and absorbable nutrients, bacteriophage genomes can carry genes between bacteria, including the most well-known toxins associated with *Vibrio cholerae* and *Escherichia coli*.

Most studies of microbiota are performed by examining either mice luminal content or feces or human feces, because it is difficult to sample the microbiota at internal sites in humans. In humans, there is no standard collection method of stool, and hence there is significant variability in the composition of the gut microbiome based on how specimens are collected (Allaband et al., 2018). Because of this, the most convincing studies of the function of the gut microbiota are from experiments performed in murine models. These studies often take advantage of the plethora of different breeds and knockout models publicly available. Furthermore, because mice can be bred or grown in a sterile environment, they can be germ-free, or once exposed to specific bacteria or microbiomes, become “gnotobiotic” mice.

The most compelling studies of the gut microbiota’s effects on the response to food come from observations in germ-free mice and gnotobiotic mice who have received stool transplantation from mice with different metabolic phenotypes. Germ-free mice are lean, with little adipose tissue. They excrete a significant number of their consumed calories in their stool. Transplanting the microbiota from mice with genetically induced

obesity (e.g., *ob/ob* mice), diet-induced obese mice, or obese humans into germ-free mice induces an obese phenotype and an increase in adipose tissue compared with germ-free mice that received stool from lean control mice. Similarly, germ-free mice that receive microbiota transplants from malnourished Malawian children develop a further malnourished phenotype.

The gut microbiome is highly dynamic, with cyclical fluctuations in absolute and relative organism abundances. These cyclical fluctuations are affected by diet (by both nutritional quality and quantity), feeding patterns, gender, circadian behavior disruptions, and genetics. Hence, it is difficult to determine whether compositional differences observed in some early studies are the result of one of these factors or related to the metabolic phenotype.

Our understanding of how the microbiota can affect host metabolism through the production of secondary metabolites has advanced significantly with improvement in the use of metabolomics to characterize the luminal environment. Microbial fermentation of carbohydrates to **short-chain fatty acids** (SCFAs) and microbial hydrolysis of primary bile acids into **secondary bile acids** are major drivers of the host response to food consumption (Nieuwdorp et al., 2014) (Fig. 56.1).

SCFAs (mostly C2–C5) account for roughly 10% of the calories extracted from food. Most of these are consumed by the gut epithelia and surrounding tissue, but a minor fraction (~1%) reach the systemic circulation. SCFAs also acidify the chyme, which inhibits the growth of certain pathogenic organisms. Derived mainly from plant sources, SCFAs are metabolites of

saccharide fermentation. Certain prominent members of the microbiota, such as *Bacteroides thetaiotaomicron*, produce glycolytic enzymes covering nearly all known saccharide structures.

Bile acids are derivatives of cholesterol classically produced by the liver and used for the digestion of luminal fatty acids. They can be converted by the microbiota into potent signaling molecules (Li and Chiang, 2014). Biochemically, the long carbon chain at C17 of the D ring is the site of taurine and glycine conjugation, and oxidation of C7 to an alcohol improves the surfactant properties of the acids. Microbiota convert these primary bile acids into secondary bile acids by hydrolyzing the amino acid amide and optionally reducing the C7 alcohol to an aliphatic carbon atom. These secondary bile acids drive changes in metabolism through various bile acid signaling receptors.

Nutrient production by the microbiota is not limited to SCFAs and secondary bile acids. Microbiota also produce essential compounds and vitamins. For example, *Bifidobacteria* can produce vitamin B12, which is essential in humans, and modulate vitamin K before absorption.

In this chapter, we will discuss the methods used to study gut microbiota and its effects on the host. We will also highlight the results from studies that link host genetics, the gut microbiome, and nutrition to provide insight into their relation with host physiology and nutrigenetics.

Design and Interpretation of Microbiota Studies

In the past, microbiologists painstakingly characterized luminal microbes by isolating individual species

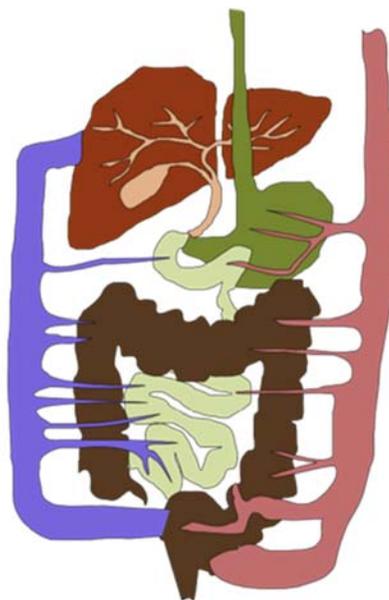
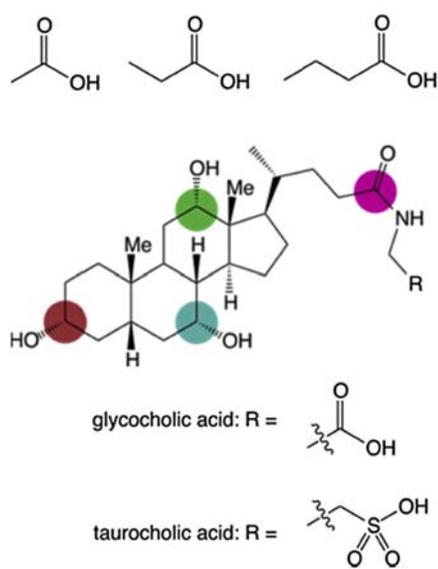


FIGURE 56.1 Chemical structures of short-chain fatty acids and bile acids in enterohepatic circulation. Highlighted atoms indicate positions of structural diversity.

and strains from feces. Bacteria were characterized by colony shape characteristics, growth media requirements, the need for oxygen, and the staining methods to visualize them best. Sequencing technology showed greater promise in identifying different species. However, because many of the microbial genomes were not known, often sequences would lead to unknown genes from undiscovered microbes. There was a need to develop a bacterial marker that would categorize organisms into taxonomic units. Preferably, the more related the taxonomic unit for two organisms, the more similar this bacterial marker would be.

Microbiome research was greatly advanced with the discovery that the 16S ribosomal RNA gene can be used to determine bacterial phylogeny. The 16S rRNA gene has conserved regions where the DNA sequence is the same for nearly all known bacteria, and nine variable regions where sequence similarity is correlated with evolutionary closeness. In many cases, these variable regions can distinguish bacterial strains. Hence, the first step of all experiments that characterize the composition of the gut microbiome is DNA extraction from either stool or luminal content. When 16S rRNA gene sequencing (commonly referred to as 16S sequencing) is used to characterize the composition of the microbiome, one of the variable regions of the 16S rRNA gene is amplified using primers directed at conserved regions flanking it. By convention, sequences that are $\geq 97\%$ identical are grouped together into an operational taxonomic unit (OTU). OTUs allow the data to “speak for themselves” by avoiding assumptions about relations between and within microorganisms but require further interpretation to produce convenient genus-species descriptions of the members of the microbiota. Representative sequences in each OTU are then matched to databases (e.g., myRDP, SILVA) to identify the OTU phylogeny. Hence, 16S sequencing determines which bacteria are present by giving the OTU phylogenies and the relative abundance of each OTU in the community.

In general, 16S sequencing provides the highest-resolution, most easily interpreted information of the general structure of a microbial community and answers the question “Who is there?” This type of analysis has become fast and inexpensive; thus, it is commonly used to characterize the gut microbiome. However, even single species can have significant variation in their genomic content, including ribosomes, which is not revealed by sequencing only 16S amplicons. More information about the gut microbiome can be attained with the shotgun metagenomic approach, which answers what the community can potentially do and what genes are available in the community. In this costlier approach, all DNA extracted from samples are sequenced. Using large reference databases (e.g., UniProt, NCBI nt), the sequences are then matched to genes or gene pathways of

available bacterial genomes. Not only does this form of analysis provide phylogeny, it shows the relative abundance of gene pathways in a community and allows a comparison of the sequences to reference genomes and the identification of variant sequences.

Other forms of sequencing and -omic approaches can further characterize the functional role of the gut microbiome. For example, extracting RNA instead of DNA from samples and removing polyadenylated messenger RNA (mRNA) and rRNA enriches the sample for prokaryotic mRNA, which can then be sequenced for metatranscriptomic analysis. This process is even costlier than metagenomic analysis and requires far more sophisticated and experimental computational analysis, but it provides an answer to the prescient question of the community is doing now. Metatranscriptome sequencing can inform us about the metabolic state of the microbiota. Finally, mass spectrometry can reveal metabolic (i.e., thorough metabolomics) and proteomic information on markers of microbiota and microbial metabolism, particularly SCFAs and bile acids and can potentially answer the question “What are they changing?”

Microbiota experiments frequently employ murine models because of their speed of reproduction and defined genetic backgrounds. However, the microbiota of any specific animal is largely determined by the microbiota inherited from its mother. Differences in the room, bedding, or cage in which an animal is housed starting from birth can also lead to unexpected interindividual variation. This can produce significant variation within an experimental group and confound the analysis of variation between groups. Furthermore, microbiota transmission between cohoused mice may not be uniform, depending on the composition of each unique gut’s ecosystem. Ideally, microbiome studies in mice should be performed in sterile, pathogen-free, or “barrier” facilities, which stabilize the microbiota of mice housed in them.

Human microbiota studies are more challenging, because it is difficult to control for known factors that affect the gut microbiome (e.g., diet, environment). Hence, there is a great deal of variation between samples collected from the same person at different times or using different methods. However, subtle differences between groups of individuals can be identified by minimizing technical variation between samples and using a large number of samples and individuals. Variations in primer pairs for 16S sequencing, methods of collecting the sample, DNA extraction procedures, or analysis techniques can easily obscure subtle differences between groups or produce erroneous results. These variations may also make it difficult to compare results of one study with another because there are not yet standardized techniques for microbiome experimentation. Because financial resources are often

limited, a difficult choice arises in sampling fewer individuals more than once or sampling more individuals a single time.

Despite these challenges, some variables are known and can be controlled. Individuals with recent exposure to antibiotic or proton pump inhibitor drugs should be excluded. In addition to controlling for standard clinical parameters such as age, ethnicity, and pregnancy, it is ideal if a time series for each subject can be collected to serve as its own control. Case-matched controls are preferable to “normal” or random controls. Investigators should be careful not to use control samples that are collected at a different time or geographical region or that have diets or medical practices different from those of their experimental comparisons. Monozygotic twins are particularly useful in specific analyses of environmental effects on microbiota community structure, although larger studies have made discoveries that smaller twin studies have not, mainly because of a lack of statistical power. Investigators and critical readers of microbiome studies should be aware that the gut microbiome is a highly dynamic system that is easily influenced by the environment, diet, collection, and analysis techniques, among other variables, and provides only a single frame of a highly cinematic movie.

Dietary Effects on the Composition of the Gut Microbiota

Diet is the primary determinant in the taxonomic and functional structure of the gut microbiota. The nutritional composition of food can bias which microbial species flourish and in some cases selectively deplete certain taxa. All mammalian gut microbiomes share a core set of genes covering essential metabolic functions. However, the relative abundances of these genes and the specific taxa that carry them clearly distinguish carnivores, omnivores, and herbivores. This emphasizes the observation that diet drives convergence of gut microbiomes across mammalian species.

One of the best examples of how diet can directly influence the gut microbiota is the infant gut. Breast milk-fed infants have a distinct gut microbiota that is directly modulated both positively and negatively by the molecular properties specific to breast milk. Human breast milk contains unique oligosaccharides not found in any other mammalian milk. Many of these sugars are recalcitrant to host digestion but can be directly used as a primary carbon source by *Bifidobacterium* and, to a lesser extent, some *Bacteroides* species. This is directly reflected in the distinct infant gut microbiota, which is dominated by *Bifidobacterium*. Furthermore, certain breast milk glycans can inhibit adhesion of pathogenic bacteria, including *Campylobacter* species and enterotoxigenic *E. coli*, to host epithelial ligands. This provides an

explanation for the finding that breast-fed infants have lower morbidity caused by diarrhea compared with formula-fed infants.

In addition to modulating the population structure of the existing gut microbiota, diet can also introduce microbial diversity. Much of the food we consume contains live microbes. Transcriptomic and culture-based analyses have shown that some of these microbes can survive transit through the gastrointestinal tract and remain metabolically active (David et al., 2014). However, whether these ingested microbes can colonize the host for any significant amount of time in the absence of newly consumed relatives remain unclear.

Broadly speaking, human gut microbiota can be divided into two major diet-driven enterotypes, or microbial organism profiles (Fig. 56.2). The first enterotype is associated with a diet high in fiber but low in fat and protein. It is characterized by an abundance of *Prevotella* species, which are specialized in metabolizing carbohydrates and polysaccharides from plants. The microbiome of this enterotype contains the metabolic potential to harness energy from macronutrients indigestible to humans. The second enterotype is associated with an almost zero-fiber diet that is high in fat and protein. In this enterotype, *Bacteroides* species are most prevalent. High fat intake increases bile acid secretion, and these organisms are characterized by bile acid tolerance. Although this can be a useful framework for visualizing dietary effects on gut microbiota, in reality enterotypes are more complex and gradual in nature.

Enterotypes are linked with long-term diet and are generally stable over months and years. However, significant alterations of gut microbiota in response to acute diet changes can be detected within 24 hours. For example, individuals consuming a strict 5-day, animal-based, high fat, high protein diet had increased bile-tolerant microbes, including *Bacteroides* and *Bilophila* species, that returned to baseline after resuming their regular diet. Similar diet-induced changes in microbial gene expression (metatranscriptomics) in all individuals

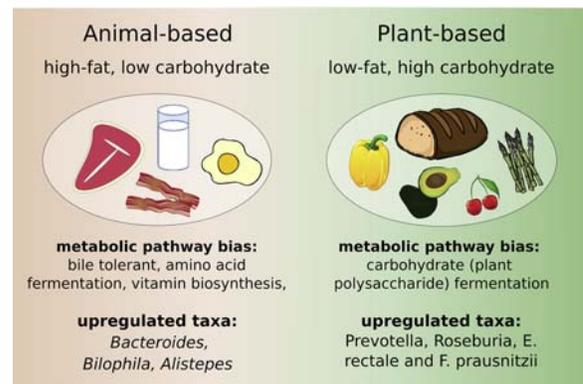


FIGURE 56.2 Diets associated with human gut enterotypes.

on the same diet superseded their initial interindividual microbiome differences. Therefore, the gut microbiota can rapidly and dramatically respond to dietary alterations.

Interactions among host diet, microbiota structure, and nutrient harvest are complex and influenced by a myriad of environmental and genetic factors. To study more clearly how diet affects gut microbiota and host nutrient harvest, researchers have employed microbially humanized mouse models. Gnotobiotic mice colonized with human gut microbiota can create stable, heritable gut microbiota that can be experimentally manipulated. These models have demonstrated that whereas initial colonization affects the structure of the gut microbiota, it can rapidly be altered by diet. For example, humanized mice fed a high-fat, high-sugar Western diet become obese, and this phenotype can be transmitted by fecal microbiota transplantation.

Diurnal Cycling of the Microbiome

Variability in the gut microbiome caused by diet led many investigators to examine how quickly and often the luminal milieu changes. Their experiments show that the gut is a highly dynamic environment; the luminal environment has fluctuations in nutrient content, pH, and secondary metabolites. These cyclical fluctuations appear to result from how each member of the gut microbiota has adapted to specific niches within the gut. With few exceptions, most mammals have periods of feeding and fasting, often with a rhythm that fluctuates considerably within a 24-hour period and coincides with host circadian transcriptional and biochemical oscillations. For example, nocturnal mice are active during dark periods, which is when they consume more than 80% of their calories with a relative fast during light periods. Although the content of a diet has an important role in shaping the composition of the gut microbiome, the daily rhythm of feeding and fasting causes cyclical fluctuations in the luminal environment. This leads to cyclical changes in the composition of the gut microbiome within a 24-hour period. This cycling is even evident at the phylum level, in which wild-type mice consuming a normal chow diet have an increase in *Firmicutes* species during the dark phase, when the animal is consuming most of its diet. During the light phase, the gut microbiome is more populated by species from the *Bacteroidetes* and *Verrucomicrobia* species.

These cyclical fluctuations are heavily influenced by diet. For example, when wild-type mice are fed a high-fat diet, a condition known as diet-induced obesity, they lose their diurnal feeding pattern and start consuming approximately 40% of their diet during the light phase. This loss of natural feeding rhythms in

diet-induced obesity mice is accompanied by a loss of cyclical fluctuation in the composition of the gut microbiome, especially at the phylum level, in which the composition of the gut microbiome is dominated by *Firmicutes* species during both the dark and light phase.

Cyclical fluctuations in the gut microbiome are influenced by diet, gender, feeding pattern, and the expression of circadian genes in the gut. Mice lacking circadian genes (e.g., *Bmal1*, *Per1/2*, *Cry1/2*) have an alteration of the cyclical fluctuation of their gut microflora. The relation between the gut microbiome and circadian genes in the peripheral tissue deserves special mention. As with other organs, the transcriptional activity in intestinal epithelial cells (IECs) is highly dynamic and circadian. These are driven by circadian clock genes that are master regulators and drive cellular machinery with approximately 24-hour periodicity. They have a vital role in nearly every aspect of cellular and whole-body physiology, especially in metabolic homeostasis. Interestingly, the gut microbiome is required for normal circadian expression of genes in the IECs. Without the commensal bacteria (as in germ-free or antibiotic-induced microbiome depletion mice), studies show that the normal circadian intestinal homeostasis is perturbed. The mechanism for this is poorly understood, and it is unclear whether the microflora or its secondary metabolites entrain circadian genes. Furthermore, germ-free mice have disruption of circadian rhythms in the liver and the hypothalamus (the central clock). Hence, these results suggest a relation between diet and feeding rhythms, the cyclical fluctuation of the gut microbiome, and host circadian gene expression. A perturbation in any one of these three can affect the other two and lead to dysmetabolism.

The Inflammasome

In addition to host circadian gene expression and the fluctuations in meals, the gut microbiota are in a dynamic equilibrium with the host immune system. The disruption of this equilibrium can lead to various diseases. Interactions between gut microbiota and the host immune system is modulated by host inflammasomes. Inflammasomes are cytoplasmic protein complexes that cleave proinflammatory cytokines into their active state in response to pathogen-associated molecular patterns (PAMPs). Upon activation, inflammasomes can induce downstream production of antimicrobial peptides (AMPs). These AMPs directly modulate gut microbiota. Interestingly, gut microbiota can influence inflammasome signaling via specific microbiota-modulated metabolites (Levy et al., 2015). Thus, there is a feedback loop among resident microbiota, diet, microbiota-modulated metabolites, inflammasomes, and AMPs.

Therefore, disruption of this interspecies feedback loop at any point can have wide-reaching effects. For example, the NLRP6 inflammasomes can regulate colonic microbiota through interleukin 18–induced AMP secretion. Deficiencies in certain subunits of the inflammasome lead to altered AMP production and can result in microbial communities that can induce colitis. Generally stated, this is an example of host genetics having a crucial role in shaping gut microbiota composition. The inflammasome status has been shown to affect susceptibility to nonalcoholic fatty liver disease and nonalcoholic steatohepatitis directly, diseases historically associated with diet, and is being investigated for other metabolic diseases. This finding highlights the importance of inflammasomes as key regulators between the microbiota and host tissue, with significant impacts on disease susceptibility.

MICROBIOME-BASED PERSONALIZED NUTRITION

Whereas diet affects the microbiota, in turn, the microbiota can affect the host's response to diet (Zeevi et al., 2015). The composition of the gut microbiome has an important role in human blood glucose homeostasis and glycemic response to food. After a meal, a healthy individual's blood glucose levels spike and then return to baseline within a small amount of time. This rise and fall in blood glucose levels in response to a meal is known as the postprandial glycemic response. However, there is a large variability in the magnitude and duration of the postprandial glycemic response across individuals, even in response to identical meals. This variability is multifactorial and includes genetics, lifestyle, and insulin sensitivity. The microbiota composition can partially predict postprandial glycemic response. Several bacterial taxa and microbial functional pathways have been identified that either positively or negatively correlate with postprandial glycemic response. By integrating behavioral, genetic, and microbial metadata, an individual's postprandial glycemic response to a specific meal may one day be predicted accurately.

Microbiota line the intestine and interact with all food entering the human body before the host tissue in the intestine. Manipulating the nutrients available to the microbiota (e.g., with diet or prebiotic supplementation) directly influences which taxa will be able to thrive. At the same time, the metabolic capabilities of the microbiota dictate which nutrients will be available to host tissue, thereby potentially directly affecting blood glucose levels. Studies that have investigated the role of microbiota in individual glycemic response have looked only at nondiabetic, healthy participants. However,

blood glucose levels are an important predictor for the onset of metabolic disorders. Two large studies conducted in Europe and China found that although there was no compositional signature of patients with diabetes, metagenomic markers could distinguish which patients were diabetic. However, it is still unclear whether diabetes, or the diets that diabetic subjects have, that drive the altered microbiome or to what extent it is the microbiome that drives the disease.

In the era of personalized medicine, there is not yet a good genetic predictor of which patients will respond best to different therapies. Because the microbiota appear to be directly related to blood glucose levels, they might have an important role in precision medicine, in which certain microbiome compositional characteristics could direct to which dietary or medical therapies an individual would have the best response. Although there is great interest in determining whether modifying an individual's microbiota can influence their postprandial glycemic response, few therapies have been successful in changing the composition of the gut microbiota. Probiotics have not shown significant or persistent changes in the gut microbiome. Research in prebiotics has been more promising. Fecal microbiome transplantation is the only known method of changing gut microbial composition, and several small studies investigating its potential utility in patients with diabetes demonstrated modest, transient improvements. However, this is still at an investigational stage. Nevertheless, it is likely that the interplay of host genetics, diet, and the microbiome will have an important role in our ability to treat obesity, diabetes, or other recalcitrant diseases effectively.

L-CARNITINE METABOLISM

Studies show that the microbiome is a contributor to atherosclerosis and cardiovascular disease (CVD) through its role in metabolizing compounds in meat (Koeth et al., 2013). Previous explanations for red meat consumption as a risk factor for these diseases have focused on the effects of saturated fats, cholesterol, and aromatic carbon compounds. Prospective studies testing these particular components of prepared red meat failed to account for the total effects of red meat consumption on major adverse cardiovascular events (i.e., heart attack, stroke, and death) associated with CVD. However, there is a direct link between red meat consumption and CVD through the gut microbiota and the small molecule trimethylamine *N*-oxide (TMAO). TMAO is produced by the liberation of trimethylamine (TMA) from carnitine, choline, or betaine (Fig. 56.3) by gut microbiota, followed by oxidation of TMA to TMAO in the liver. TMAO is absent from germ-free

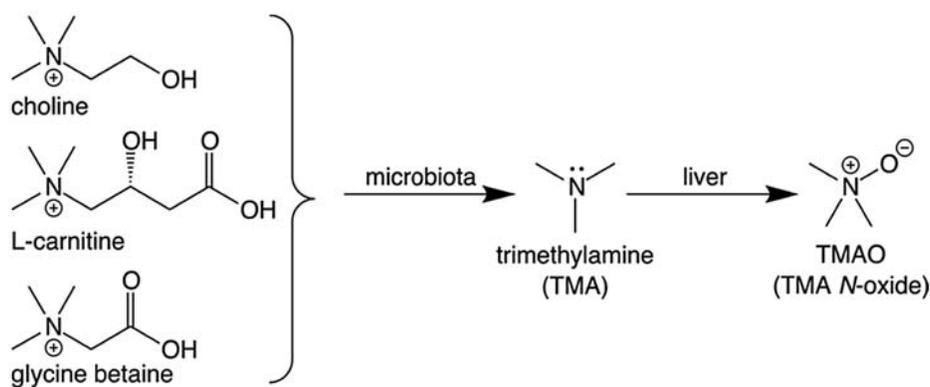


FIGURE 56.3 Catabolism of quaternary ammonium by microbiota produces trimethylamine, which is absorbed and oxidized to trimethylamine *N*-oxide in the liver.

mice and humans whose microbiomes have been depleted by antibiotic treatment.

The identification of a single small molecule, TMAO, as a causative agent gives researchers the unique opportunity to test these observed retrospective correlations prospectively. In humans, plasma TMAO levels were strongly correlated with the subsequent incidence of major adverse cardiovascular events. An analog of choline that inhibits the production of TMA from carnitine and choline by some microbiota, 3,3-dimethyl-1-butanol, also reduced the volume of atherosclerotic plaques in mice on a choline-rich diet.

Despite these encouraging studies, important questions remain open. Many foods known to be sources of choline have been studied for their effect on CVD-related mortality with limited positive correlations, if any. Furthermore, patients may present in the clinic with carnitine or choline deficiency syndromes, which highlights the importance of these molecules despite their suggested risk. It is unclear how much the dietary source of TMA affects the conversion of TMA into TMAO. Investigations in patients with deficiency of the enzymes necessary to convert TMA to TMAO (trimethylaminuria or “fish odor syndrome”) are ongoing to determine whether they have a reduced risk for atherosclerotic plaques. Furthermore, more investigation is needed to determine whether there is an interaction between the genetic predisposition to atherosclerosis and the presence of TMA-producing bacteria.

HOST GENETICS AND THE GUT MICROBIOME

Gut microbiota have coevolved with humans and were subjected to the same environmental selective pressures as their hosts. Therefore, it is interesting to examine potential correlations between a host’s genetic background and the microbiome (e.g., an inflammasome

deficiency described earlier). For example, the differences observed among individuals’ response to a high-fat diet (HFD) are likely influenced by their microbiota, but to what extent is the microbiota determined by the host’s genetic background?

Mouse models have proven to be a useful tool to examine the influence of host genetics on the composition of the gut microbiota. In one experiment, the microbiome of three mouse strains with different genetic backgrounds from two commercial animal facilities were compared. Initially, all three had distinct microbiomes and their metabolic responses to HFD were distinct. To control for the effect of environmental variables on the microbiome, the mice were bred for at least three generations under common conditions. This environmental normalization led to more similar fecal compositional (16S sequencing) profiles among the strains, which remained distinct but with smaller differences. Interestingly, the differences in the metabolic response to an HFD among the mice strains was nearly eliminated. However, environmental normalization had less of an impact on the one of the three mouse strains, which suggests that genetic backgrounds allow for different responses to microbiome-related metabolic manipulation. Although diet has the greatest impact on shaping the gut microbiota, the host’s genetic background continues to influence the metabolic response to diet. This highlights the care that must be taken to control for both genetic background and environmental variables when comparing mice from different animal facilities.

The best evidence for genetic influence on the microbiota in humans comes from twin studies comparing monozygotic twins (which come from the same embryo and therefore share nearly 100% of their genes) and dizygotic twins (which come from separate embryos and share ~50% of their genes). This allows researchers to compare the effect of genetic contribution while controlling for the known impact of early environmental

variables (e.g., birthing and infant care). In one study, investigators compared the adult gut microbiota of 31 monozygotic twin pairs and 23 dizygotic twin pairs at two different time points (Turnbaugh et al., 2009). There were relatively few differences in 16S sequencing results when the gut microbiota of monozygotic and dizygotic twin were compared. However, the microbiota of each twin was more similar to the twin than to the mother or an unrelated adult. More recent studies provide evidence that the link between host genetics and gut microbiota may be more nuanced, and targeted approaches (e.g., evaluating concordance of specific taxa abundance) have yielded significant results.

A far larger, and thus better-powered twin study showed that specific taxa were significantly associated with host genetic background. In this study, over 1000 fecal samples from 171 monozygotic and 245 dizygotic twins were analyzed by 16S sequencing (Goodrich et al., 2014). One particular bacterial family, the gram-positive Christensenellaceae, was the most heritable taxon overall. Furthermore, this bacterial family was the most frequent taxon to co-occur with other heritable taxa. Importantly, these findings were validated in another large twin cohort.

The presence of this heritable family, Christensenellaceae, is associated with a lean body mass index. Experiments in mouse models strongly suggest a functional role of this bacterial family in diet-induced weight gain. Germ-free mice gained significantly more weight when receiving fecal microbiota transplants from obese human donors lacking Christensenellaceae compared with lean human donors containing Christensenellaceae. Furthermore, the addition of live *Christensenella minuta* to a fecal microbiota preparation from an obese human donor lacking Christensenellaceae taxa significantly ameliorated the increased adiposity of the germ-free mouse recipient. These results demonstrate that specific taxa are associated with host genetic background and can influence the host's response to diet. However, it is unclear whether the *C. minuta* has similar antiobesity effects in humans.

THE RELATIONSHIP OF HOST GENETIC POLYMORPHISMS AND THE GUT MICROBIOME

Thus far, we have discussed several different ways in which the gut microbiome and the host reciprocally interact. Understanding these relations has been instrumental to understand the role of the gut microbiome better and its interaction with host genetics in the pathophysiology of disease.

Helicobacter pylori and CYP2C19

Helicobacter pylori, which is acquired early in life and persists unless eradicated, is the main bacteria inhabiting the stomach and interacts intimately with gastric epithelial cells. Since its discovery, *H. pylori* has been mainly characterized by its pathogenic interactions with the host, which can produce chronic inflammation and lead to a variety of gastrointestinal disorders. These include dyspepsia, peptic ulcer disease, and even mucosa-associated lymphoid tissue lymphoma. Current clinical practice is to eradicate the organism when it is found, because it can be responsible for catastrophic gastrointestinal bleeds from peptic ulcer disease and is viewed as a carcinogenic agent. Nevertheless, there is debate about whether *H. pylori* should be characterized as a commensal organism that has coevolved with humans for at least 58,000 years. *H. pylori* in children is thought to be protective against the development of asthma, and there is a suggestion that eradicating the bacterium is associated with obesity, increased blood glucose levels, and worsening acid reflux disease. Because the bacterium is so commonly found and often is not pathogenic, some researchers classify it as a commensal organism that can become pathogenic.

As the approach to *H. pylori* eradication has become more aggressive, the bacterium has become more resistant to antibiotic treatment; annual cure rates have gradually decreased. Thus, individualized treatments by better characterizing either *H. pylori* or the host may be more successful. An important factor for eradicating *H. pylori* is acid suppression with a proton pump inhibitor (PPI). Cytochrome P450 2C19 (CYP2C19) is the principal enzyme necessary for metabolizing many common PPIs. The effectiveness of PPI suppression depends on the host genotype. Those who have the "rapid metabolizer" genotype of CYP2C19 do not have sufficient acid suppression of many common PPIs, which is necessary for the efficacy of *H. pylori* treatment. These patients are usually assigned to the few PPIs that are either longer-lasting in the blood or are metabolized differently for the acid suppression necessary for treating *H. pylori*.

Patients who have the "poor metabolizer" phenotype of CYP2C19 thus can gain sufficient acid suppression for therapy for *H. pylori*, but prolonged PPI therapy can lead to long-term adverse events such as susceptibility to vitamin B12 deficiency and iron deficiency. CYP2C19 does not have a physiological role in acid production in the stomach; hence, these variant genotypes are not associated with different susceptibility to *H. pylori* or acid-related diseases (e.g., acid reflux). Rather, the genotype at CYP2C19 links the metabolism of a drug to the treatment of *H. pylori* and to nutrient absorption.

Celiac Disease and Human Leukocyte Antigen DQ2

Another link among genotype, microbiome, and nutrition lies in celiac disease. Individuals with expression of the human leukocyte antigen (HLA) class II molecules DQ2 and DQ8 are strongly susceptible to celiac disease, a chronic inflammation of the intestine caused by an aberrant immune response to gluten. Gluten is a cereal protein found in wheat, barley, and rye. Individuals with the HLA-DQA1*05:01 and DQB1*02:01 alleles forming the DQ2.5 haplotype have the highest risk for developing celiac disease, and families with these alleles have a particularly high susceptibility to celiac disease. Those with celiac disease have an altered gut microbiome, although it is unclear whether these microbiome differences are dysbiotic and contribute to the etiology of the disease or are the result of an altered (i.e., gluten-free) diet.

More recent studies show that the HLA-DQ2 genotype influences the composition of the gut microbiome in infants who have a strong familial risk for celiac disease. These changes include higher *Firmicutes/Clostridium* and *Proteobacteria/Enterobacteriaceae* species and lower *Actinobacteria/Bifidobacterium* species. Furthermore, this study suggests that higher *Bifidobacteria* species are associated with decreased celiac disease species. Whether HLA-DQ2 predisposes infants to a dysbiotic microbiome that makes them more susceptible to celiac disease remains an open question.

CONCLUSION

Until only a couple of decades ago, the role of the gut microbiome in host physiology and disease was completely unknown. The advent of next-generation sequencing technology has catapulted forward our understanding of the role of the microbiome in health and disease. Although the composition of the gut microbiome is affected by diet, gender, and a multitude of other factors, the concept that it is affected by our own

genome is fascinating and under rigorous study. What agents from the host are mediating the effects of our genes on the composition of the gut microbiome? How are these changes affecting the luminal response and absorption of nutrients? These relations are complex but important to understanding the metaorganism. These questions will be answered by careful experimentation and standardized use of multiple -omic approaches (e.g., host transcriptomics and genomics, luminal metagenomics, and metabolomics).

References

- Allaband, C., McDonald, D., Vazquez-Baeza, Y., et al., 2018. Microbiome 101: studying, analyzing, and interpreting gut microbiome data for clinicians. *Clin Gastroenterol Hepatol* 17 (2), 218–230.
- David, L.A., Maurice, C.F., Carmody, R.N., et al., 2014. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505 (7484), 559–563.
- Dutilh, B.E., Cassman, N., McNair, K., et al., 2014. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat Commun* 5, 1–11.
- Goodrich, J.K., Waters, J.L., Poole, A.C., et al., 2014. Human genetics shape the gut microbiome. *Cell* 159 (4), 789–799.
- Kau, A.L., Ahern, P.P., Griffin, N.W., Goodman, A.L., Gordon, J.I., 2011. Human nutrition, the gut microbiome and the immune system. *Nature* 474 (7351), 327–336.
- Koeth, R.A., Wang, Z., Levison, B.S., et al., 2013. Intestinal microbiota metabolism of l-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med* 19 (5), 576–585.
- Levy, M., Thaïss, C.A., Zeevi, D., et al., 2015. Microbiota-modulated metabolites shape the intestinal microenvironment by regulating NLRP6 inflammasome signaling. *Cell* 163 (6), 1428–1443.
- Li, T., Chiang, J.Y.L., 2014. Bile acid signaling in metabolic disease and drug therapy. *Pharmacol Rev* 66 (4), 948–983.
- Nieuwdorp, M., Gilijamse, P.W., Pai, N., Kaplan, L.M., 2014. Role of the microbiome in energy regulation and metabolism. *Gastroenterology* 146 (6), 1525–1533.
- Turnbaugh, P.J., Hamady, M., Yatsunenkov, T., et al., 2009. A core gut microbiome in obese and lean twins. *Nature* 457 (7228), 480–484.
- Wu, H., Tremaroli, V., Bäckhed, F., 2015. Linking microbiota to human diseases: a systems biology perspective. *Trends Endocrinol Metab* 26 (12), 758–770.
- Zeevi, D., Korem, T., Zmora, N., et al., 2015. Personalized nutrition by prediction of glycemic responses. *Cell* 163 (5), 1079–1094.

This page intentionally left blank

S E C T I O N I I I

Nutrigenomics

This page intentionally left blank

Nutrients and Gene Expression in Development

Dolores Busso, Nicolás Santander, Francisca Salas-Pérez, José Luis Santos

Department of Nutrition, Diabetes, and Metabolism, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile

Glossary

CpG site DNA sequence where a cytosine sits next to a guanine.

CpG islands Clusters of CpG sites.

Developmental Origins of Health and Disease Concept linking the rise in the prevalence of noncommunicable diseases to intrauterine programming.

DHA Polyunsaturated fatty acid highly concentrated in the brain that might promote neural development.

DNA methylation Epigenetic change involving the addition of methyl groups to DNA.

Epigenetic Potentially heritable changes regulating gene expression that do not involve changes in DNA sequence.

Epigenome Total epigenetic modifications to the genome.

Extraembryonic tissues Tissues for protection and nutrition of the developing embryo, such as the placenta, yolk sac, and amnios.

Gene promoter DNA sequence that defines the initiation of gene transcription by RNA polymerase.

Genomic imprinting Epigenetic mechanism that involves DNA and histone methylation in a parent of origin-specific manner.

Histones Family of basic proteins that associate with DNA in the nucleus and help it condense into chromatin.

Intrauterine programming Group of fetal adaptations to intrauterine environment that lead to diseases in adulthood.

microRNA Small noncoding RNA molecules that block RNA transcription.

Neural tube Embryonic structure that develops into the brain and spine.

ROS (reactive oxygen species) Molecules generated by oxidative metabolism that act in cell signaling in physiological concentrations but can be harmful to cells in high concentrations.

Royal jelly Honeybee secretion given exclusively to certain larvae in a beehive to produce a queen.

Teratogenic Any substance capable of disrupting normal development and generating malformations.

Trophoblast Cells of embryonic origin derived from the blastocyst trophoderm that differentiate to form extraembryonic tissues (i.e., the yolk sac and placenta).

Vitamin B₉ (folic acid) Water-soluble vitamin that promotes DNA methylation and protects against neural tube closure defects.

Window of susceptibility Time frame during development when the embryo is more likely to be programmed by environmental factors.

STAGES OF EMBRYONIC AND FETAL DEVELOPMENT

Embryonic and fetal intrauterine development encompasses a series of complex, well-concerted events of cell, tissue and organ differentiation and growth, starting immediately after sperm–egg fertilization and ending at birth. Although prenatal development is a continuum, it can be subjectively categorized into three main stages. The first stage, **preimplantation development**, begins with the unicellular zygote and includes embryonic cleavage and the differentiation of the totipotent blastomeres forming morulae into the two cell types in the blastocyst: trophoblast and inner mass cells. The second stage, **embryonic development**, starts with attachment of the blastocyst to the uterine wall, in a process known as implantation. This stage is also composed of gastrulation, involving cell migration concomitant to differentiation, resulting in the three germ cell layers, which later combine and continue differentiating to form all of the tissues and organs in the body. The first organs that are formed in the embryo are the neural tube, which later originates the brain and the spine, and the heart. During this stage of development, extraembryonic tissues are also formed. The amniotic membrane, the yolk sac, and the early placenta are essential during early development, because they prevent dehydration and provide a suitable milieu for the embryo, allowing nutrient and gas exchange with the mother. The human yolk sac disappears before the 12th week of pregnancy, whereas in rodents this organ is functional until the end of gestation. Finally, the third stage, known as **fetal development**, involves mainly organ maturation and growth until parturition. The placenta matures and becomes fully functional during this stage. The timing of the different stages varies among different

species: in humans, preimplantation development takes around 5 days after conception, embryonic development ends at around week 8, and fetal development lasts until the end of gestation, around 7 months.

EPIGENETICS IN DEVELOPMENT

The formation of a complex organism, starting from a unicellular zygote, and the coordinated physiological responses that take place during development require massive changes in gene expression across several cells and tissues. Besides being primarily determined by the DNA sequence, gene expression is regulated by the epigenome through a combination of stable and dynamic epigenetic modifications that make the DNA more or less accessible to transcriptional regulators. One of the best-described epigenetic mechanisms is the addition of a methyl group to the 5' position of a cytosine sitting next to a guanine in DNA (CpG site), yielding 5-methylcytosine (5mC) through the process of **DNA methylation**. This is established in mammals during embryogenesis and maintained by replication of cell division in somatic cells. DNA methylation has important roles in gene expression regulation, genomic imprinting, X-chromosome inactivation, and retrotransposon silencing. In general, it is thought that methylation of CpG sites by DNA methyltransferases (Dnmts) renders the DNA less accessible to transcription factors, resulting in repression of gene expression. Three main Dnmts have been described in mammals: Dnmt3a and Dnmt3b function primarily to establish *de novo* DNA methylation patterns during early embryogenesis, whereas Dnmt1 is the major maintenance enzyme that reproduces CpG methylation patterns after DNA replication. A relevant percentage of mammalian gene promoters contains clusters of CpG sites, denominated CpG islands, which potentially regulate gene expression. However, many gene promoters remain scarcely methylated, especially those from ubiquitously housekeeping genes.

Two waves of global DNA demethylation occur during specific stages of embryonic development. One wave occurs during germ cell differentiation, when nearly complete demethylation and subsequent remethylation take place. The methylation pattern is again reprogrammed shortly after fertilization, in the zygote. After fertilization, paternal 5mC is oxidized to 5-hydroxymethylcytosine (5hmC) by the Teneleven translocation (Tet3 member) dioxygenase in such a way that the resulting paternal 5hmC together with maternal 5mC, gradually declines during subsequent divisions. A new pattern of DNA methylation, known as *de novo* methylation, is established some time later, during the blastocyst stage. Appropriate reprogramming of

methylation at this stage is crucial for the development, survival, and even postnatal health and behavior of individuals. Once lineage-specific methylation patterns are established, aging and different environmental factors such as dietary components, toxins, drugs, or diseases may regulate the DNA methylation processes in different periods of life.

In some genes, DNA methylation allows the differential expression of alleles, depending on whether they come from the spermatozoa or the egg. Around 1% of autosomal genes undergo this maternal or paternal imprinting in mammals, which is settled during gametogenesis. If an allele is maternally imprinted, marks are imposed on both the maternally and paternally inherited genes during oogenesis and demethylation takes place in both maternally and paternally inherited during spermatogenesis. Thus, after fertilization, the offspring inherits a silenced (imprinted) allele from the mothers and functional (nonimprinted) allele from the father. The fact that most imprinted genes have roles in the control of embryonic growth and development, including development of the placenta, has led researchers to suggest that they function as an adaptive mechanism to obtain the best nutritional provision to the embryo without affecting the mother's nutritional status. In fact, several methylated genes code for growth factors. Those paternally expressed genes generally enhance embryonic growth, whereas those that are maternally expressed appear to suppress embryonic growth. Besides being related to embryonic growth, imprinting has been linked to some human diseases. On the one hand, defective imprinting mechanisms and aberrant gene expression during development have been shown to explain congenital syndromes such as Beckwith-Wiedemann, Prader-Willi, and Angelman syndromes. On the other hand, the phenotypic consequences of a mutation can remain inactive if the mutated allele is silenced by imprinting, which explains the parent-of-origin effect of certain diseases such as schizophrenia and epilepsy. In the rest of the genes that are not subjected to imprinting, methylation is also important, because some genes need to be expressed specifically during defined windows during embryonic development and then have to be inhibited.

Besides DNA, histones can also be subject to modifications that confer genomic plasticity. These proteins pack DNA into chromatin, the structural component of chromosomes. Posttranslational modifications such as methylation and acetylation can change the affinity of histones for DNA, modifying chromatin compaction and modulating the accessibility of transcription factors to interact with genes. Histone deacetylases, acetyltransferases, and methyltransferases have been described in different cell types. Whereas **histone acetylation** makes chromatin more active and promotes gene expression,

histone methylation can render chromatin more or less active, depending on the modified amino acid. Histone methylation and deacetylation, together with DNA methylation, are involved in inactivating one of the X chromosomes in females, to maintain gene dosage. In addition to DNA methylation and histone modifications, **microRNAs (miRNAs)**, have more recently been shown to be involved in the epigenetic regulation of gene expression. These small noncoding RNA molecules are able to silence gene expression at the translational level, preventing messenger RNA (mRNA) from being transcribed or accelerating mRNA degradation. It has been shown that the expression of several genes coding for miRNAs can be regulated either by DNA methylation or histone modifications, which suggests that different components of the cellular epigenome are interdependent. Making the epigenetics scenario even more complex, several miRNAs directly target DNA methyltransferases and histone-modifying enzymes. Several examples of macronutrients and micronutrients and bioactive food components influencing development through epigenetic marks have been described. Mechanisms explaining the effect of nutrients on epigenetics are diverse and include directly modifying the activity of enzymes that catalyze DNA methylation or histone modifications, altering the availability of substrates necessary for those enzymatic reactions, and regulating the expression of miRNAs.

NUTRITION AND DEVELOPMENT: DEMANDS, SOURCES, AND STAGES OF SUSCEPTIBILITY

The epigenetic status of the developing organism is modulated by the intrauterine environment, which in turn is defined both by maternal characteristics (e.g., metabolism) and gestational environmental exposures (e.g., diet). Nutrients not only influence the maternal health status but are transported across maternal fetal barriers to reach the embryo or fetus, directly affecting their development. Both nutrient deficiency and excess have been shown to hinder development. The overall effect for each nutrient on the embryo or fetus depends on two main aspects: the difference between its demand and availability and the timing of exposure. In general terms, macronutrients reaching the extraembryonic tissues and the embryo or fetus serve as sources of energy for cell proliferation and function, to sustain proper development and growth. Micronutrients, in turn, are crucial to regulate diverse biological processes, i.e., DNA repair, gene expression, hormone secretion, vascular tone, and protection against cell oxidative damage. The nutritional demand of a developing organism varies qualitatively and quantitatively as it matures

from an embryo into a fetus, and eventually into a newborn. Fetal nutrient availability is determined by the type and amount of nutrients consumed by the mother, but also by her body composition, her placental function, and adaptations to cope with her own nutritional requirements and those of the developing embryo or fetus. For example, a teenager mother who has not yet achieved her adult body composition will have a significantly higher nutrient demand than will an adult woman. So will be the case of a woman doing intense physical work or coping with disease. In those cases, there will be competition between maternal needs and those of the fetus. As well, multiple pregnancies reduce the nutrient availability of each fetus, and short intervals between pregnancies hinder the appropriate repletion of maternal nutritional reserves.

The source of available nutrients also changes along development. During the preimplantation stage, nutrients are first obtained from the newly fertilized egg cytoplasm and later from the maternal oviductal and uterine fluids, by phagocytosis. After implantation, the early embryo starts taking up nutrients from maternal sources. Even before the establishment of a mature placenta, the blastocyst trophoblasts and yolk sac endoderm cells phagocyte secretions from endometrial and uterine glands. This mechanism of nutrient uptake in the early conceptus is called **histiotrophic nutrition**. Once a mature, functional placenta is established **hemotrophic nutrition** takes place. This kind of nutrition corresponds to the exchange of blood-borne materials between the maternal and fetal circulations. Whereas histiotrophic nutrition provides for most of the embryonic stage of development, hemotrophic nutrition becomes predominant during fetal development. However, in certain species, such as in rodents, these two pathways may coexist for much of the gestational period.

The placenta is formed by embryonic and maternally derived tissues; it regulates the transport of nutrients, excretions, and gases between the fetal and maternal circulations and modulates the maternal immune system to prevent the rejection of the embryo. Inside the placenta, maternal blood lacunae are separated from fetal capillaries by multinucleated cells called syncytiotrophoblasts. These intriguing cells originated by the fusion of numerous trophoblasts allow the well-controlled, bidirectional transport of molecules. Whereas gases such as oxygen and carbon dioxide are permeable and can cross syncytiotrophoblasts by passive diffusion, other substrates require channels and transporters. Glucose, for example, crosses the placenta by facilitated diffusion using specific transmembrane proteins. Other nutrients such as calcium, amino acids, and lactate cross the placenta by active transport using energy from adenosine triphosphate (ATP) or ion gradients. Fatty acids are an example of molecules that can

either diffuse passively or be actively transported by specialized fatty acid transport proteins. There is significant evidence showing that certain maternal conditions such as diabetes and obesity can alter the expression and/or activity of nutrient receptors and transport molecules in syncytiotrophoblasts, influencing nutrient transport across the placental barrier and in some cases hindering fetal development and growth. In addition, structural abnormalities in the placenta that affect placental blood flow and available surface exchange area can interfere with nutrient transport to the embryo and fetus.

Periods when genes are more actively predisposed to epigenetic modifications are considered windows of susceptibility (Fig. 57.1). These are stages when developing individuals are particularly sensitive to different environmental cues. One of these is preimplantation development, when reprogramming of methylation takes place. This stage is one of the most important windows of potential environmental influence over the embryo. Once implantation takes place, a high DNA synthesis rate and elaborate DNA epigenetic patterning are required for appropriate cell differentiation and specification, as well as normal tissue and organ

development. Thus, environmental exposures affecting gene expression during embryonic development can result in **changes in phenotype**, such as in color or behavior, or in more severe defects ranging from **malformations to embryonic lethality**, depending on the organ affected and the severity of the effect. One of the first and best described changes in phenotype is observed in the agouti mouse, whose coat color distribution is shifted from yellow to brown, depending on the maternal consumption of methyl-donor nutrients during pregnancy. In humans, maternal nutritional deficiencies during the first trimester have been linked to developmental syndromes. One clear example is severe maternal iodine deficiency, which results in fetal iodine insufficiency and different degrees of mental retardation. The most extreme manifestation of prenatal iodine deficiency is cretinism, a syndrome characterized by severe mental retardation, deaf-mutism, and strabismus, among other defects. Another well-known example is folic acid deficiency, which has been linked to an increased risk for gestating an infant with neural tube defects (NTD). On the other hand, the excessive ingestion of nutrients such as vitamin A (as retinol/retinyl esters) during pregnancy has also been suggested to result in birth defects.

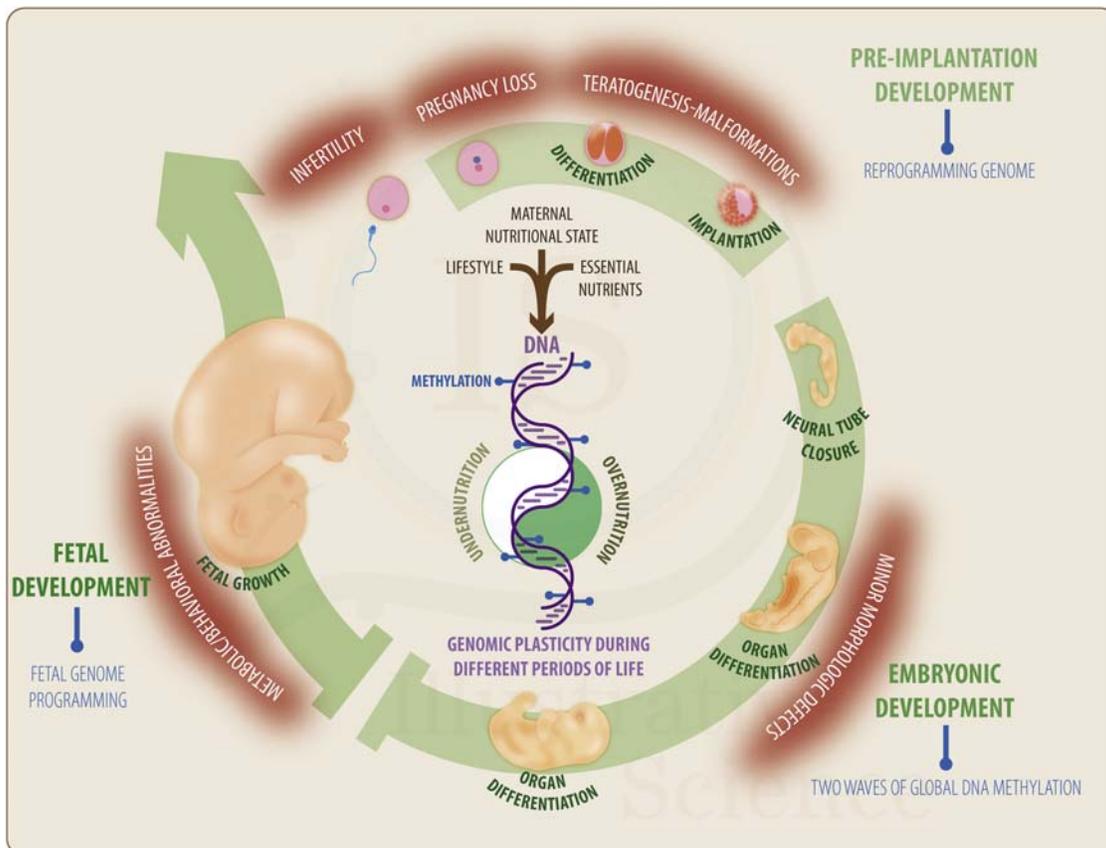


FIGURE 57.1 Windows of susceptibility to nutritional exposures during intrauterine development and consequences in the offspring.

Although there are only a few cases of retinol excess in humans, experimental models have supported the teratogenic effects of this vitamin.

Environmental exposures can also take place during the fetal stage once the vital organs are differentiated. Epigenetic changes in the fetal stage of development are usually not lethal but may affect fetal growth, organ maturation, and/or the postnatal phenotype of the offspring. The process by which insults at critical stages of fetal development lead to permanent changes in the offspring (mostly postnatally) is known as **intrauterine programming**. Growing evidence shows that both maternal undernutrition and overnutrition have a significant impact on the offspring's postnatal susceptibility to disease. It has been shown that fetuses, in particular males, can adapt to malnutrition using the limited nutrients available for survival of vital organs, such as the brain, at the expense of growth of other organs, such as the liver or pancreas. It has also been hypothesized that in undernourished fetuses, epigenetic modifications in key organs remodel the metabolism of the fetus to adapt to this nutrient-deprived environment. These fetal adaptive epigenetic responses, however, can render individuals susceptible to metabolic imbalances when exposed postnatally to a nutrient-rich environment. Overnutrition, on the other hand, has also been linked by both epidemiological and experimental data to programming of metabolic syndrome, diabetes, and cardiovascular disease (CVD). That the prevalence of chronic diseases has increased substantially in most developing countries, as well as in countries undergoing economic transitions, has led many researchers around the world to reconsider how early-life environment can affect later health in an approach known as Developmental Origins of Health and Disease.

Among a significant number of nutritional exposures on different developmental stages and diverse effects via interdependent epigenetic mechanisms, this chapter focuses on some illustrative cases describing how nutrigenomics can shape development.

To Be or Not to Be: Phenotype Can Be Determined by Nutrients That Modulate Methylation

There are two remarkable examples in the animal kingdom of the interaction between diet and epigenome affecting phenotype. In the honeybee (*Apis mellifera*), queens and workers are genetically identical in such a way that larvae may develop into a queen or a worker, depending on the environmental conditions. In a striking case of phenotypic plasticity, only queens are reproductively competent and have an increased life span and other important differences in morphology and

physiology compared with workers. Epigenetic modifications in DNA methylation are induced by the provision of larvae with royal jelly or worker jelly, resulting in a nutrient-derived caste differentiation that leads to different gene and protein expression patterns in queens compared with workers. The second example, extensively studied, on the effect of diet in reversible epigenetic marks and subsequent phenotype determination corresponds to the agouti viable yellow (A^{vy}) mouse. This mouse carries an insertion known as intracisternal A-particle (IAP) retrotransposon within the *agouti* gene, producing stochastic differences in DNA methylation that lead to variable ectopic expression of this gene, resulting in variability among individuals in traits such as coat color, energy intake, and obesity. Both coat color and body adipose tissue content can be modified by supplementing pregnant mothers' diets with compounds favoring methylation (e.g., folic acid, vitamin B₁₂, choline, and betaine). Increased DNA methylation upstream the intracisternal A particle element in the *agouti* gene results in partial reversion of the phenotypic features of A^{vy} mice, so methyl-donor fed dams give birth to pseudoagouti (brown) slim mice instead of yellow obese mice. Dietary supplementation of pregnant A^{vy} mice with the polyphenol genistein, present in soybeans, also increases DNA methylation and shifts the offspring phenotype to pseudoagouti. Apart from this striking example in mice, maternal nutrition and phenotype determination have been studied in women from rural Gambia, because they are exposed to seasonal variations in methyl-donor nutrients. This natural experiment has led to the idea that maternal methyl donor consumption during specific stages of pregnancy may modulate DNA methylation of metastable epialleles in the offspring. However, observational studies measuring methyl-donor intakes in relation to global DNA methylation have not yet yielded conclusive evidence for a causal association among diet, methylation marks in the DNA, and phenotype in this population.

Nutrients and Gene Expression in the Brain: Preventing Neural Tube Defects and Promoting Brain Cognitive Function

NTDs constitute the second most common malformation in humans. NTD are determined by maternal and embryonic genes and have a great influence on the environment. Significant prevention of NTD in humans and in animal models is achieved by supplementation with vitamin B₉ (also known as folic acid). The success of this intervention to decrease the risk for NTD in the general population led several countries, particularly in the Americas, to implement the mandatory folic acid fortification of flour. Maintaining this policy resulted in the

reduction of the NTD prevalence in 50%–70%, despite controversial evidence linking folic acid fortification with increased risk for colorectal cancer, epilepsy, or twin birth. Work with animal models has been invaluable in trying to unravel mechanisms linking folic acid consumption and NTD. Several murine models with natural or artificial mutations recapitulate different aspects of human NTD. Interestingly, some of the mouse mutants are completely sensitive to folic acid supplementation whereas others are totally resistant to this treatment. Folic acid, as well as other molecules involved in one-carbon metabolism such as vitamin B₁₂ and choline, serves as donors of methyl groups necessary for DNA synthesis and methylation. In agreement with this role for folic acid, maternal folate status has been associated with the methylation level and expression level of certain genes in the embryo, including those involved in neural tube closure (e.g., *Alx3*, *Pax3*). Folic acid has also been shown to act as an antioxidant by regulating the activity of the nicotinamide adenine dinucleotide phosphate oxidase, an enzyme that produces large quantities of reactive oxygen species (ROS). In rodent models of NTD, overt ROS production has been extensively documented. Besides folate, other antioxidants have been successfully used to prevent NTD in experimental models. In neural cells, oxidative imbalance has been shown not only to be toxic but also to regulate the expression of genes involved in migration and differentiation. For example, exposing mouse embryos to oxidative stress *in vitro* or *in vivo* leads to a reduction in *Pax3* expression, whereas mutations in this gene are associated with NTD in mice and humans. After neural tube closure, the ventricles expand and the different areas of the brain form by the concerted migration and differentiation of neural progenitors. Nutrients such as long-chain polyunsaturated fatty acids (PUFA) have been suggested to modulate these processes. Maternal supplementation with the PUFA docosahexanoic acid (DHA) during pregnancy has been associated with improvements in specific cognitive functions in children aged 9 years or younger. Observational studies comparing high versus low DHA status have also supported the beneficial effect of DHA on cognitive tasks in infants. However, the long-term effects of DHA and the epigenetic mechanisms explaining its effects are still unknown.

Ying and Yang: Maternal Undernutrition and Overnutrition and the Epigenetic Programming of Chronic Diseases

The idea that abnormal fetal growth resulting from undernutrition caused the 20th century epidemic of coronary heart disease in Western countries was pioneered by the English epidemiologist David Barker in his

studies on newborns in England around 1940. His hypothesis was strongly supported by posterior data provided by the Dutch Famine in 1944, when a military-occupied region of The Netherlands was subjected to limited rations of food for nearly 6 months. In those pregnant women, undernourishment from mid to late gestation was associated with reduced birth weight, whereas exposure to famine during early gestation did not affect birth weight but increased susceptibility to developing CVD, obesity, renal dysfunction, and type 2 diabetes. Studies on the methylation pattern of genes from individuals conceived during the Dutch famine demonstrated that embryonic undernutrition caused several epigenetic changes that persisted throughout life, including reduced DNA methylation of the insulin-like growth factor (IGF)2 (*IGF2*) gene and increased DNA methylation in interleukin-10, *leptin*, and ATP-binding cassette subfamily A member 1 genes. Similarly, individuals exposed to the Chinese Great Famine (1959–61) during the first trimester of pregnancy have higher risk for type 2 diabetes and hypertension in adulthood, although the epigenetic mechanisms explaining this phenomenon remain largely underexplored. Another important human study showing the influence of maternal nutrient consumption on perinatal outcomes examined pregnancies in Gambia, where seasonal food availability determines seasonal increase in the incidence of intrauterine growth restriction and prematurity. Studying methylation patterns in tissues derived from the three germ lines demonstrated that the embryonic epigenome is altered before gastrulation by periconceptual undernutrition in these women.

Animal models, primarily rodents and sheep, revealed some information about the complex mechanisms of the programming of CVD by intrauterine nutrient deficiency. Gestational caloric restriction in different animal models has linked nutrient deficiency with altered expression of IGFs and abnormal fetal growth. In sheep, maternal global nutrient restriction during early pregnancy affected the offspring's hypothalamus–pituitary–adrenal axis response through downregulation of the adrenal glucocorticoid receptor (GR) expression. In pregnant rats, consumption of reduced protein diets leads to promoter hypomethylation in fetal liver genes such as peroxisome proliferator-activated receptor α and GR, key regulators of carbohydrate and lipid metabolism. Folic acid supplementation in that model can prevent some of the metabolic abnormalities in the offspring, which supports the idea that interventions with nutrients can reduce the long-term risk for diseases associated with undernutrition. Besides affecting gene methylation, prenatal undernutrition can induce epigenetic modifications in histones and in genes coding for miRNAs regulating metabolism.

The population in most developing countries faces the opposite of malnutrition, caused by consumption of diets with low-quality nutrients and excessive calories. Maternal overweight and obesity are independent risk factors for maternal and fetal pregnancy complications, including gestational diabetes mellitus, embryonic malformations, and fetal macrosomia. Women who are obese before conception or gain excess weight during pregnancy are more likely to have babies who develop obesity, insulin resistance and early markers of CVD during their infancy or adult life, including higher blood pressure and inflammation. Patients with obesity and metabolic syndrome exhibit different global DNA methylation, histone acetylation, and miRNA expression marks than do healthy patients, which suggests that epigenetic modifications on key genes for metabolism could be a mechanism explaining metabolic abnormalities in the offspring of obese mothers. A major target of epigenetic influence is peroxisome proliferator-activated receptor gamma coactivator 1 α (PGC-1 α), a transcriptional coactivator that regulates mitochondrial biogenesis and integrates different metabolic signals. Hypermethylation of the PGC-1 α promoter has been linked to a reduction in PGC-1 α expression in skeletal muscle and pancreatic islets from patients with type 2 diabetes. A variety of animal models have been established to study developmental programming by overnutrition. Pregnant rodents fed highly palatable, hypercaloric diets are being used as experimental models of gestational obesity. Their offspring exhibit hyperphagia, insulin resistance, hypertension, and glucose intolerance during adulthood. Several different, yet mutually interacting epigenetic mechanisms have been proposed to explain intrauterine programming by maternal obesity. Different metabolically relevant fetal tissues are epigenetically modified in response to obesity. In the hypothalamus, specific neuropeptides involved in energy homeostasis (e.g., Neuropeptide Y and proopiomelanocortin) and receptors for metabolites such as insulin and estrogen are differentially expressed in offspring from obese mice compared with controls. In livers of mice exposed to maternal obesity, high triglyceride content is associated with changes in histone marks at the liver X receptor, an important regulator of cholesterol, fatty acid, and glucose metabolism.

THE ROLE OF THE FATHER

Evidence described previously showed how different nutrients consumed by females can affect embryonic and fetal development and modulate the offspring's health. Besides the intrauterine influence on embryos and fetuses, studies in both humans and

animal models support the idea that maternal prepregnancy and even paternal nutrition can also contribute to reprogramming of the offspring. Germ cells resulting in sperm and eggs start differentiation during fetal life, so DNA damage, mutations, and epigenetic marks on those embryonic germ cells can also affect the following generation. Existing evidence demonstrates that feeding male mice a suboptimal diet (e.g., low-protein, high-fat, or folate deficient) can alter the global methylation and miRNA content in germ cells, which persist in sperm and may influence the offspring phenotype. Thus, epigenetic changes resulting from malnutrition in both mothers and fathers might have important roles in the transgenerational persistence of metabolic diseases.

Different studies show that exposure to a high-fat diet (HFD) in utero causes metabolic disturbances in the offspring and that epigenetic modifications in certain genes, i.e., those coding for adiponectin and leptin, may persist for several generations. Paternal body mass index (BMI) also affects the offspring's BMI via epigenetic marks, and these effects are additive to the influence of maternal BMI. Interestingly, the transgenerational reprogramming effects are reversible: a study showed that consumption of a normal diet by the offspring during subsequent generations after maternal or paternal HFD exposure first diminished and then completely abolished the effect of HFD exposure on the metabolic traits and the epigenetic changes in the offspring.

CONCLUDING REMARKS: UNDERSTANDING NUTRIGENOMICS TO IMPROVE PREGNANCY OUTCOME

Nutrients encompass a large, diverse, and complex family of molecules with many known roles, and probably more yet to be discovered. These molecules can serve as energy sources, substrates, or cofactors of other physiologically relevant molecules, antioxidants, methyl donors, and enzymes, among many functions. In this chapter, we focused on nutrients affecting epigenetics and gene expression during intrauterine development. Pregnancy is a critical period when maternal and paternal nutrition choices can influence the embryonic, fetal, and postnatal health of the offspring. Inadequate levels of nutrients during critical stages of intrauterine development can lead to birth defects, suboptimal fetal development, and reprogramming of fetal tissues, predisposing the individual to chronic conditions later in life. In adolescent and young women, consuming a healthy, balanced diet and keeping a BMI within the recommended range for the gestational stage are required to achieve a healthy pregnancy. Consumption of insufficient,

excessive, or poor-quality diets, or increased body requirements (e.g., caused by maternal infections, intense physical activity, or growth in adolescents) may promote malnutrition and negatively affect the offspring's health. Appropriate and timely nutrient supplementation policies, nutritional counseling, and education can aid in preventing some consequences of malnutrition. Our knowledge of nutrigenomics during pregnancy will help us understand how the environment tailors diversity and susceptibility to disease in human beings and provide us with the necessary information to improve the health of future generations globally.

TAKE-HOME MESSAGES

- Embryonic development is a complex and dynamic process requiring the well-controlled regulation of gene expression to sustain cell division and differentiation.
- Embryonic nutrition can be histotrophic, hemotrophic, or both. In humans, histotrophic nutrition takes place before the establishment of a mature placenta, the specialized organ that allows the maternal fetal exchange of blood-borne nutrients.
- Altered nutrient availability (by defect or excess) can induce embryonic lethality, teratogenesis, or organ malformations. In some cases, maternal nutrient consumption determines the phenotype of the offspring via epigenetic marks. These can influence fetal growth, organ maturation, and the postnatal susceptibility to disease.
- Epigenetic marks in the embryo (DNA methylation, histone methylation, or acetylation) provide mechanisms to ensure the spatial and temporal organization of gene expression and confer adaptability to the intrauterine environment.
- Maternal nutrient intake during pregnancy can modify (enhance or reduce) epigenetic marks in the embryo, changing the expression of critical genes associated with development.
- Both maternal preconceptional and paternal nutrient inadequacies may affect the epigenetic status of germ cells and their derived gametes, explaining the transgenerational effect of nutrition.

Suggested Reading

- Fleming, T.P., Watkins, A.J., Velazquez, M.A., Mathers, J.C., Prentice, A.M., Stephenson, J., Barker, M., Saffery, R., Yajnik, C.S., Eckert, J.J., Hanson, M.A., Forrester, T., et al., 2018. Origins of lifetime health around the time of conception: causes and consequences. *Lancet* 391, 1842–1852.
- Gilbert, S.F., 2006. *Developmental Biology*, eighth ed. Sinauer Associates, Sunderland, MA.
- Greene, N.D., Copp, A.J., 2014. Neural tube defects. *Annu Rev Neurosci* 37, 221–242.
- Hanson, M.A., Bardsley, A., De-Regil, L.M., Moore, S.E., Oken, E., Poston, L., et al., 2015. The international federation of gynecology and obstetrics (FIGO) recommendations on adolescent, preconception, and maternal nutrition: "think nutrition first". *Int J Gynaecol Obstet* 131 (Suppl. 4), S213–S253.
- Jirtle, R.L., Skinner, M.K., 2007. Environmental epigenomics and disease susceptibility. *Nat Rev Genet* 8 (4), 253–262.
- Masuyama, H., Mitsui, T., Eguchi, T., Tamada, S., Hiramatsu, Y., 2016. The effects of paternal high-fat diet exposure on offspring metabolism with epigenetic changes in the mouse adiponectin and leptin gene promoters. *Am J Physiol Endocrinol Metab* 311, E236–E245.
- Masuyama, H., Mitsui, T., Nobumoto, E., Hiramatsu, Y., 2015. The effects of high-fat diet exposure in utero on the obesogenic and diabetogenic traits through epigenetic changes in adiponectin and leptin gene expression for multiple generations in female mice. *Endocrinology* 156, 2482–2491.
- Nilsson, E.E., Sadler-Riggelman, I., Skinner, M.K., 2018. Environmentally induced epigenetic transgenerational inheritance of disease. *Environ Epigenet* 4, dvy016.
- Rasmussen, E.M., Amdam, G.V., 2015. Cytosine modifications in the honey bee (*Apis mellifera*) worker genome. *Front Genet* 6, 8.
- Vanhees, K., Vonhogen, I.G., van Schooten, F.J., Godschalk, R.W., 2014. You are what you eat, and so are your children: the impact of micronutrients on the epigenetic programming of offspring. *Cell Mol Life Sci* 71 (2), 271–285.
- Warner, M.J., Ozanne, S.E., 2010. Mechanisms involved in the developmental programming of adulthood disease. *Biochem J* 427 (3), 333–347.

Nutrients, Obesity and Gene Expression

Fermín I. Milagro^{1,2}, María J. Moreno-Aliaga^{1,2}, J. Alfredo Martínez¹

¹Department of Nutrition, Food Science, and Physiology, Center for Nutrition Research, University of Navarra, Pamplona, Spain ²CIBERobn, Centro de Investigación Biomédica en Red de la Fisiopatología de la Obesidad y Nutrición, Carlos III Health Institute, Madrid, Spain

Glossary

Epigenetics The study of heritable changes in phenotype that do not involve changes in the underlying DNA sequence.

Genome-wide association study Screening of a genome-wide set of genetic variants in many individuals to identify variants associated with a trait.

Metabolomics The large-scale study of small-molecule metabolite profiles in a cell, tissue, organ, or organism.

miRNA microRNAs (miRNA) are small noncoding RNA molecules that have a role in RNA silencing and posttranscriptional regulation of gene expression.

mRNA Messenger RNA (mRNA) is a kind of RNA that conveys genetic information from DNA to protein by way of the processes of transcription and translation.

Noncoding RNA An RNA molecule that is not translated into a protein.

Nutrigenomics The branch of nutritional genomics that studies the effects of foods and food constituents on gene expression, tries to understand the response of the human body to a food via systems biology, and analyzes single gene–single food compound relations. In some definitions, nutrigenomics includes the effects of nutrients on human health through alterations in the genome, epigenome, proteome, metabolome, and resulting physiological and metabolic outcomes.

Promoter A region of DNA that initiates transcription of a particular gene; it is located near the transcription start site of the gene.

Proteomics The large-scale study of proteins.

Single-nucleotide polymorphism (SNP) A variation in a single nucleotide that occurs at a specific position in the genome.

Transcriptome The ensemble of all RNA molecules in one cell or a population of cells. This term is used to refer to all RNAs, or just mRNA, depending on the particular study. Transcriptomics is the science that studies the transcriptome using microarrays and next-generation sequencing technologies.

OBESITY, POLYMORPHISMS, AND GENE EXPRESSION

Obesity is a chronic condition in which excessive fat accumulates in the body as a result of a chronic

imbalance between energy consumption and expenditure. The pathogenesis of obesity is complex, involving interactions between genetics and epigenetics, environmental, and behavioral factors. Changes in adiposity and metabolic responses to low-calorie diets and different nutrients may be influenced by genetic variants related to obesity, metabolic status, and preference to nutrients. Because human genetic sequence has not been substantially altered in the past hundreds of years, it is thought that the rapid rise in obesity prevalence must be associated with a profound shift in our lifestyle, including unhealthy dietary patterns, overnutrition, and physical inactivity. Because the environmental risk factors are largely modifiable, the development of obesity would be essentially preventable. However, this endeavor will be helped by a better knowledge of the regulation of gene expression and metabolic pathways by nutrients and the understanding of clear gene–diet interactions.

Monogenic obesity is a rare and severe early-onset obesity associated with endocrine disorders. Monogenic obesity is mainly caused by mutations in genes that encode for proteins that have a key role in regulating and integrating signals that control food intake in the hypothalamus, including leptin, leptin receptor, melanocortin-4 receptor, prohormone convertase 1, proopiomelanocortin, and brain-derived neurotrophic factor, among others. Many of these forms of monogenic obesity are mediated by a sharp decrease or the inability to express the coded protein.

However, common obesity is a complex disease with multifactorial origins, which in most cases appears as a polygenic condition that is also affected by a myriad of environmental obesogenic and behavioral factors. Linkage and candidate gene association analyses, and especially genome-wide association studies, have allowed

the identification of a relevant number of novel loci and genetic variants associated with polygenic obesity/body mass index (BMI) changes in different populations by regulating a variety of biological processes, including central control of food intake, insulin signaling, adipocyte differentiation, thermogenesis, inflammation, and adipose tissue, liver, and muscle metabolism, and even gut microbiota composition. In contrast to genes involved in monogenic obesity, which largely affect phenotype, most genes associated with polygenic obesity have modest or small phenotype effects in themselves. However, they might have an additive effect by interacting with environmental and behavioral factors that fostering a relevant susceptibility to obesity or a different response to weight loss treatments.

It is conceivable that the effect of some genes on the control of body weight could be nutrient-sensitive. Some studies reported that the expression of multiple genes in different tissues is altered by reductions in energy intake or by diets rich in fat or carbohydrates. However, more studies are needed to demonstrate that the effects of diet on weight changes depend on DNA sequence variants regulating specific gene expression. For instance, integrating genome-wide linkage and/or association analyses with expression profiles from relevant tissues may be a good strategy to increase efficiency in the difficult task of identifying genes underlying obesity and body weight regulation. Microarrays and, more recently, whole-genome and RNA sequencing (RNA-seq) are facilitating this endeavor, although existing network databases have many missing interactions and contain many false-positive interactions that reduce the sensitivity of this type of network analysis. Anyway, only a combination of computational bioinformatics and experimental approaches will yield mechanistic insights into the process by which a genetic variant, or a combination of variants, affect a complex phenotype such as obesity by interacting with nutrients in regulating gene expression and metabolic response.

On the other hand, notable differences in gene expression levels and genetic regulation of transcripts are found when comparing different cell types and tissues, which adds a new layer of difficulty to understanding the effects of nutrients and diets on gene expression regulation.

ADIPOSE TISSUE IN OBESITY

Adipose tissues have crucial roles in the development of obesity. White adipose tissue (WAT) functions as an energy storage organ whereas brown adipose tissue is an energy-dissipating organ. However, WAT is also recognized as a highly active metabolic and endocrine organ. Adipocytes are important in buffering the daily

influx of dietary fat and WAT-derived factors exert autocrine, paracrine, and/or endocrine effects involved in the physiological regulation of fat storage, adipogenesis, energy metabolism, and food intake, and have a key role in developing obesity-associated metabolic disorders. Several studies aimed to identify differential gene expression patterns between lean and obese individuals using microarrays. The differential expression of genes involved in adipogenesis, mitochondrial oxidative metabolism, extracellular matrix remodeling, angiogenesis, inflammation and oxidative stress, apoptosis, signal transcription, and cell cycle-related genes has been described. However, in these studies comparing adipose tissue transcriptomic profiles between lean and obese individuals, it is difficult to distinguish whether the observed differential gene expression is a cause or consequence of obesity; more studies are needed to shed light on this area. Whole-genome blood transcriptomic analyses have been used to obtain insights into the protective molecular mechanisms acting in metabolically healthy obese subjects. These analyses revealed that the control of endoplasmic reticulum stress and/or ribosomal stress could be an effective strategy to prevent or delay the occurrence of metabolic disorders in obese individuals. Global transcriptomic analyses of adipose tissue or adipocytes have also allowed the identification of differential gene expression patterns between obese with low insulin resistance or high insulin resistance.

Concerning nutrigenomics, a number of adipocyte events may be affected by the interaction between genetic variants and dietary factors increasing the risk for developing obesity and insulin resistance. Some of the most important processes are adipogenesis (adipocyte differentiation), lipid turnover (synthesis and breakdown, including lipolysis), mitochondria respiratory function and fatty acid oxidation, endocrine and autocrine functions (including inflammation but also insulin and leptin signaling pathways), oxidative stress, and endoplasmic reticulum stress (which may induce activation of apoptosis and autophagy).

Finally, many complications linked to obesity, such as insulin resistance, mitochondrial dysfunction, lipotoxicity, and low-grade chronic inflammation, have been partially attributed to ectopic accumulation of fat in the liver, muscle, heart, pancreas, and even bone marrow. The cause of ectopic fat storage is largely unknown, although genetic, environmental, and behavioral factors are probably involved. Impaired Wnt signaling seems to be implicated in generating ectopic fat and insulin resistance. For the moment, diet and lifestyle intervention, particularly a combination of severe caloric restriction and exercise, are the only method to improve both ectopic fat deposition and the function of the organ in which ectopic fat is deposited.

Nevertheless, more research is needed about gene–diet interactions in this pathological condition. Characterization of the transcriptomic profile of adipose tissue by RNA-seq during a low-calorie diet has allowed the identification of predictors of weight outcomes in obese-nondiabetic responders and nonresponders. Short-term overfeeding also influences adipose tissue transcriptomics and epigenomics. Moreover, a strong body of evidence revealed that expansion of adipose tissue during obesity is also associated with altered expression of noncoding RNAs, including microRNA (miRNA) and long noncoding RNA (lncRNA). The main catabolic (lipolysis, fatty acid β -oxidation, and thermogenesis) and anabolic (adipogenesis and lipogenesis) processes are mainly controlled by adipose tissue, skeletal muscle, and liver and are regulated by nutrients, hormones, and environmental factors. Noncoding RNAs seem to have a key role in regulating these metabolic processes through interactions with proteins, messenger RNAs (mRNAs), or DNA.

In past years, some studies suggested that miRNA and lncRNA released from other cell types and tissues (immune cells, liver, muscle, etc.) could affect the expression of genes in adipose tissue and mediate changes in the metabolism of adipocytes. These noncoding RNAs are present within exosomes (alongside transcription factors and other transcriptional regulators) and may act locally or travel through the circulation, reaching distant sites. Exosomes mediate cell–cell communication through different mechanisms, including activation of intracellular signaling by ligand–receptor interaction at the plasma membrane, releasing their content into the cytoplasm of recipient cells after fusing with the plasma membrane, or being endocytosed and fusing with an endocytic compartment. There have been descriptions of exosomes and other extracellular vesicles released by the adipose tissue that can have therapeutic potential in vascular and neurodegenerative diseases or promote angiogenesis and cancer cell migration.

NUTRIENT REGULATION OF GENE EXPRESSION IN OBESITY AND WEIGHT LOSS

Nutrients and dietary factors are able to elicit health effects through specific interactions on a molecular level, such as by regulating gene expression by modulating the activity of transcription factors or even the epigenetic landscape. The science that studies the effects of nutrients on the expression of an individual's genetic makeup is called nutrigenomics. In a broader sense, this concept refers also to the impact of dietary compounds on the epigenome, proteome, and metabolome.

Translation of mRNAs represents the last step of genetic flow and primarily defines the proteome. Hence, transcriptional and translational regulation are critical for gene expression, especially under nutrient excess or deficiency. Studies focused on gene expression levels across tissues and how they are affected by body weight and adiposity status and dietary interventions to induce obesity, promote weight loss, or alter dietary nutrient profile have been investigated. However, studies documented a lack of correlation between genomic and proteomic profiling of a number of genes, which suggests that it is also important to analyze posttranscriptional regulation events, including noncoding RNAs, mRNA degradation, translational control, and protein turnover. Different dietary factors that affect gene expression and regulate molecular mechanisms that control body weight and insulin sensitivity are illustrated in [Fig. 58.1](#).

Many studies have analyzed the effects of different nutritional interventions on gene expression profiles. Some of them are shown in [Table 58.1](#).

Lipids

Dietary fat is an important macronutrient that has key functions in human physiology and well-being. Apart from being a source of metabolic energy, lipids are crucial for maintaining the structural integrity of cells. They have important roles in metabolism, regulating cell function, and synthesizing hormones. For example, some fatty acids can act as signaling molecules, regulating cell function. A growing body of evidence supports that the type and quality of fat have an important role in the development of obesity and metabolic-associated disorders. Indeed, many saturated fats have been related to the onset of obesity and insulin resistance, whereas diets high in monounsaturated and some polyunsaturated fatty acids (PUFA), particularly n-3 but also some n-6, could prevent or attenuate the development of these diseases. It has been suggested that an increase in the n-6/n-3 PUFA ratio raises the risk for obesity, possible through arachidonic acid metabolites and hyperactivation of the cannabinoid system.

Adipose tissue has an important role in handling fatty acids. Also, the function and metabolism of adipose tissue can be affected by the type of nutrients. Indeed, differential effects have been described on white fat adipogenesis, depending on the type of fatty acids (oleic acid, arachidonic acid, conjugated linoleic acid, or n-3 PUFA), based on their differential actions on key adipogenic transcription factors such as peroxisome proliferator activated receptor γ (PPAR γ) and CCAAT-enhancer-binding protein α . Interestingly, studies reported that the n-3 PUFA eicosapentaenoic acid could promote brown and white adipocyte differentiation.

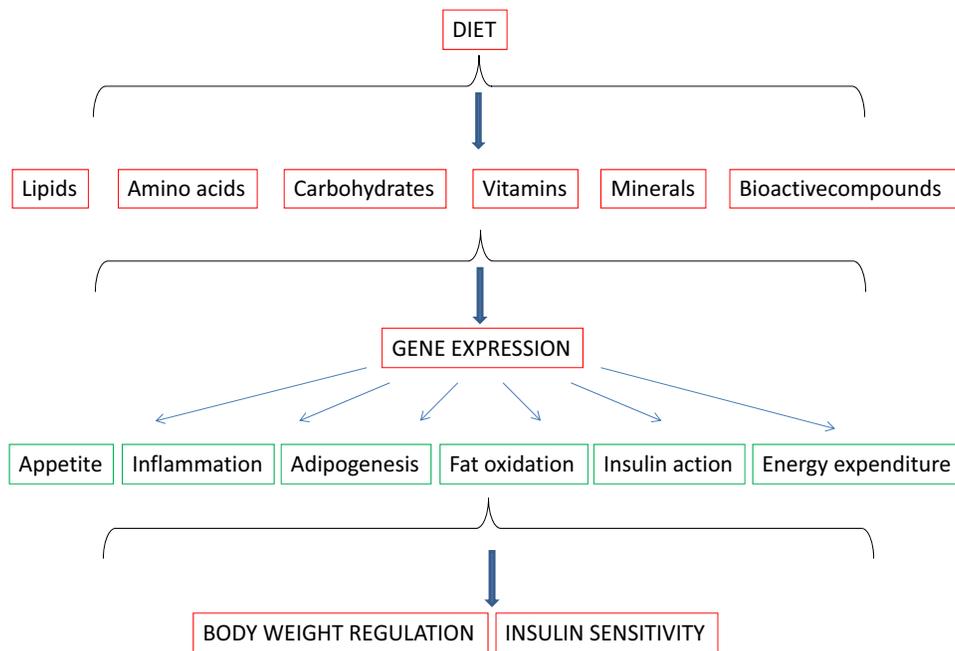


FIGURE 58.1 Dietary factors that affect gene expression and regulate some of the molecular mechanisms involved in controlling body weight and insulin sensitivity.

Dietary fat can also differentially affect the regulation of mitochondrial genes and mitochondrial function, which can be relevant because mitochondrial dysfunction can lead to the development of several metabolic diseases such as obesity and insulin resistance. It has been reported that diets high in saturated fat impair mitochondrial biogenesis by regulating master genes or transcription factors involved in controlling this process, such as peroxisome proliferator-activated receptor gamma coactivator 1 α . In contrast, n-3 PUFA have been shown to upregulate mitochondrial genes and biogenesis in adipose tissue under obesogenic conditions.

Dietary fatty acids can also influence obesity and the development of associated metabolic disorders by regulating the production of bioactive adipokines that have a critical role in controlling food intake and body composition, as well as in lipid and glucose homeostasis and insulin sensitivity. Some of these adipokines, such as leptin, adiponectin, or apelin, have been shown to be differentially regulated, depending on the type of fatty acid. For example, n-3 PUFA, particularly eicosapentaenoic acid and docosahexaenoic acid, ameliorate the low-grade inflammation of adipose tissue associated with obesity and induce changes in the pattern of secreted adipokines, resulting in improved systemic insulin sensitivity. n-3 PUFA affect adipocytes by different mechanisms, including the transcription factors PPAR α , PPAR γ , and PPAR δ , but some of the beneficial effects of n-3 PUFA depend on their active metabolites, especially eicosanoids and pro-resolving lipid mediators such as resolvins, protectins and maresins. On the other hand,

n-3 PUFA have been shown to reduce the transcriptional activation of several genes by inhibiting nuclear factor- κ B system activation, which in many cases occurs as a result of the decreased generation of intracellular reactive oxygen species.

Carbohydrates

A major function of dietary carbohydrates is as a metabolic fuel. Although they provide energy and may promote weight gain, the relation between carbohydrate intake and body mass index remains controversial. Carbohydrates can differentially regulate gene expression in obesity in different cell types and tissues. For example, high dietary fructose has been related to a higher risk for developing obesity and contributing to insulin resistance and metabolic syndrome. Excessive amounts of fructose lead to fatty liver, which is a key step in developing insulin resistance, by increasing de novo lipogenesis.

The metabolism of fructose is different from glucose; both monosaccharides display some differential metabolic actions. For example, fructose reduces the secretion of adiponectin and leptin from adipose tissue, whereas glucose metabolism stimulates leptin production. Fructose triggers AMP-activated protein kinase (AMPK)/malonyl-coenzyme A (CoA) signaling in hypothalamus, increasing food intake and the risk for obesity. Moreover, fructose alters gene expression patterns (such as PPAR γ coactivator-1 α/β in the liver), increases oxidative stress,

TABLE 58.1 Examples of Nutrigenomic Studies Assessing Gene Expression Profiles Associated With Nutritional Interventions, and the Potential Role of Nutrients and Bioactive Compounds.

Nutritional Interventions	Target Genes	Expression Changes	Potential Health Effects
Mediterranean diet	<i>NFKB1, IKKBK, MMP9, IL1B, MAPK8, XBP1</i>	–	Anti-inflammatory, antiatherogenic
Mediterranean diet plus olive oil	<i>NFKB1, MMP9, TNFA</i>	–	Anti-inflammatory, antiatherogenic
Mediterranean diet	<i>NFE2L2, SOD1, SOD2, TXNRD1</i>	–	Anti-inflammatory, antioxidant
High MUFA	<i>APOBR</i>	–	Antilipidemic, antiatherogenic
Energy-restricted diet plus EPA	<i>IL10</i>	+	Anti-inflammatory
High PUFA	<i>POMC, GALP</i>	+	Antiobesity
High PUFA	<i>HCRT, MCH</i>	–	Antiobesity
Energy-restricted diet plus EPA and α -lipoic	Lipid catabolism genes	+	Antilipidemic
Energy-restricted diet plus EPA and α -lipoic	Lipid storage genes	–	Anti-lipidemic
High protein	<i>PPARGC1A, PCK1, GSTA, CPT1A</i>	+	Antisteatotic
High protein	<i>FGF21, SCD1</i>	–	Antisteatotic
Curcumin	<i>MMP9, MMP13, EMMPRIN</i>	–	Antiatherogenic, anticancer
Resveratrol	<i>EMMPRIN</i>	–	Antiatherogenic
Apple polyphenols	<i>LEP, SREBF1, PLIN</i>	–	Antiobesity
Apple polyphenols	<i>PPARGC1A, AQP7, AEBP1</i>	–	Antiobesity
Flavonoid-fish oil supplement	Phagocytosis-related inflammatory genes	–	Anti-inflammatory
High n-3/n-6 PUFA ratio	<i>TLR4, TNFA, IL6, CRP</i>	–	Anti-inflammatory, antidiabetic
EGCG	<i>MMP9, MMP2</i>	–	Antitumorigenic
Theaflavin	<i>MMP2</i>	–	Antitumorigenic
Resveratrol	<i>FASN</i>	–	Antisteatotic
Sulforaphane	<i>EGR1</i>	+	Anticancer
Genistein	<i>P21, P16</i>	+	Anticancer
Genistein	<i>BMI1, c-MYC</i>	–	Anticancer

AEBP1, adipocyte enhancer binding protein 1; *APOBR*, apolipoprotein B receptor; *AQP7*, aquaporin 7; *CPT1A*, carnitine palmitoyltransferase 1A; *CRP*, C-reactive protein; *EGR1*, early growth response 1; *EMMPRIN*, extracellular matrix metalloproteinase inducer; *EPA*, eicosapentaenoic acid; *FASN*, fatty acid synthase; *FGF21*, fibroblast growth factor 21; *GALP*, galanin-like peptide; *GSTA*, glutathione S-transferase cluster; *HCRT*, hypocretin neuropeptide precursor; *IL10*, interleukin 10; *IL1B*, interleukin 1 β ; *IL6*, interleukin 6; *IKKBK*, inhibitor of κ light polypeptide gene enhancer in B cells, kinase β ; *LEP*, leptin; *MAPK8* (JNK1), mitogen-activated protein kinase 8; *MCH*, melanin concentrating hormone; *MMP13*, matrix metalloproteinase 13; *MMP2*, matrix metalloproteinase 2; *MMP9*, matrix metalloproteinase 9; *MUFA*, monounsaturated fatty acids; *NFE2L2*, nuclear factor erythroid 2-like 2; *NFKB1*, nuclear factor κ B subunit 1; *PCK1*, phosphoenolpyruvate carboxykinase 1; *PLIN*, perilipin; *POMC*, proopiomelanocortin; *PPARGC1A*, peroxisome proliferative activated receptor γ coactivator 1 α ; *PUFA*, polyunsaturated fatty acids; *SCD1*, stearoyl-coenzyme A desaturase 1; *SOD1*, superoxide dismutase 1; *SOD2*, superoxide dismutase 2; *SREBF1*, sterol regulatory element binding transcription factor 1; *TLR4*, Toll-like receptor 4; *TNFA*, tumor necrosis factor α ; *TXNRD1*, thioredoxin reductase 1; *XBP1*, X-box binding protein 1.

Based on Ramos-López, O., Milagro, F.I., Allayee, H., Chmurzynska A., Choi M.S., Curi R., Caterina R.De, Ferguson L.R., Goni, L., Kang, J.X., Kohlmeier, M., Marti, A., Moreno, L.A., Pérusse, L., Prasad, C., Qi, L., Reifem, R., Riezu-Boj, J.I., San-Cristobal, R., Santos, J.L., Martínez, J.A., 2017. Guide for current nutrigenetic, nutrigenomic, and nutriepigenetic approaches for precision nutrition involving the prevention and management of chronic diseases associated with obesity. *J Nutrigenetics Nutrigenomics* 10, 43–62.

inflammation, and portal endotoxin concentrations via Toll-like receptors, and induces leptin resistance. In contrast, high-glucose has been reported to inhibit AMPK activation in hypothalamus, leading to reduced food intake.

Proteins

The primary functions of proteins are to provide structure, repair damaged tissues, preserve immune function, synthesize hormones, and, in certain circumstances, provide energy. Diets moderately high in proteins are successfully used to treat obesity. These diets may increase energy use owing to enhanced protein catabolic processes, such as transamination, the tricarboxylic acid cycle, and oxidative phosphorylation. Two major pathways, target of rapamycin and amino acid response, sense both the sufficiency and deficiency of amino acids. They help the cell to detect nutrient availability and adjust cellular processes to meet bioenergetic needs.

Among amino acids, leucine is a well-known activator of the mammalian target of rapamycin. Although central injection of leucine reduces food intake, the efficacy of oral leucine supplementation to inhibit appetite significantly and reduce obesity remains unclear. On the other hand, leucine deprivation in mice stimulates fat loss by reducing lipogenesis in the liver and increasing fat mobilization in WAT, a result similar to that found in rodents fed diets deficient in other branched-chain amino acids, such as valine and isoleucine.

Dietary methionine restriction has been shown to reduce adiposity and liver steatosis and increase overall insulin sensitivity in mice, in part by reducing lipogenic gene expression in liver, inducing WAT browning, and enhancing the lipolytic and oxidative capacity of the remodeled WAT. Dietary methionine restriction has emerged as an effective mimetic of calorie restriction that reduces adiposity through a compensatory increase in energy expenditure and a decrease in the expression of inflammatory markers in adipose tissue and liver. Some, but not all of these effects could be mediated by the role of this amino acid as a precursor for the formation of *S*-adenosyl methionine and the methylation of macromolecules, including DNA. In any case, the role of specific amino acids on transcriptomic processes needs to be elucidated.

Vitamins

Obesity is often associated with low levels of B-group vitamins in serum, particularly folate. For example, low maternal folate supply in rats resulted in offspring that

were more susceptible to detrimental metabolic effects of a high-fat diet, which suggests that folate deficiency among pregnant women may exacerbate the risk for metabolic disease in offspring. Epigenetic mechanisms may be behind some of these effects, because folic acid is an important molecule involved in one-carbon metabolism and many cellular functions, including DNA, RNA, and protein methylation as well as DNA synthesis and repair. On the contrary, excess folic acid has been reported to increase lipid storage, weight gain, and adipose tissue inflammation in high-fat diet-fed rats. In 3T3-L1 adipocytes, excess folic acid induced PPAR γ expression and triglyceride accumulation. These data indicate that renewed efforts must to be focused on the effects of methyl donor (notably folate) supply, particularly during conception and pregnancy, owing to its potential long-term effects on offspring metabolic health.

As folate, vitamin B₁₂ is a necessary cofactor in the conversion of homocysteine to methionine. In observational studies, high concentrations of vitamin B₁₂ have been linked to reduced risk for obesity, whereas low plasma vitamin B₁₂ in pregnancy has been associated with increased adiposity and gestational diabetes. Also, vitamin B₁₂ is required as a coenzyme in the conversion of methylmalonyl-CoA to succinyl-CoA in the mitochondria; in its absence, accumulation of the former compound inhibits fatty acid oxidation, promoting lipogenesis. Low vitamin B₁₂ levels seem to be linked to adipocyte dysfunction and obesity-related complications by modulating cellular inflammation and lipid metabolism and causing hypomethylation of cholesterol biosynthesis pathways.

Biotin acts as a coenzyme for carboxylases regulating the synthesis of valine, isoleucine, and fatty acids, and in gluconeogenesis. It has been reported that biotin increases acetyl CoA carboxylase 2 gene expression in the hypothalamus, leading to the suppression of food intake in mice. Similarly, biotin seems to prevent the development of diabetes by increasing the functions of pancreatic β cells and improving insulin sensitivity in the periphery. It has also been reported that biotin restriction decreases lipid accumulation and respiration and alters mitochondrial morphology in adipocytes. The biotin-dependent functions are modified by adiposity and correlate with inflammation and hypertriglyceridemia.

Low intake levels of the main antioxidant vitamins (vitamins A, C, and E) have been associated with increased BMI and waist perimeter. For example, vitamin A has been inversely related to adiposity: the higher the consumption of vitamin A, the lower the risk for obesity. Chronic dietary vitamin A supplementation regulates obesity in obese-phenotype rats, possibly by upregulating the uncoupling protein 1 gene and causing adipose tissue loss. Binding of retinol-bound

retinol-binding protein to a membrane-binding protein is known to suppress insulin signaling. In this sense, all-trans retinoic acid, a derivative of vitamin A, is able to reverse these effects, resulting in increased insulin sensitivity, downregulation of the phosphoenolpyruvate carboxykinase gene, and upregulation of glucokinase expression.

In some studies, vitamin C levels were inversely related to BMI, body fat percentage, and waist circumference, modulating the expression of genes involved in adipogenesis. In rats supplemented with vitamin C, genes whose expression more strongly correlated with improvements in body fat and homeostasis model assessment-estimated insulin resistance (insulin resistance index) were those involved in adipocyte differentiation, lipid and glucocorticoid metabolism, cell cycle regulation, and insulin-induced processes; however, these results need to be confirmed in humans. Concerning vitamin E, it was reported that genetic variations in oxidative stress-related genes in humans (*TXN* and *COMT*) could modulate waist circumference in relation to dietary vitamin E intake. However, more intervention studies are needed to clarify the role of this vitamin in obesity development and treatment.

Vitamin D deficiency has been also linked with obesity. Furthermore, early pregnancy vitamin D status, particularly 25[OH]D₃, has been inversely associated with gestational diabetes risk. Many studies demonstrated the significant effect of calcitriol on adipocytes. However, in addition to its calcitropic effect, vitamin D is a regulator of gene expression as well as cell proliferation and differentiation. Vitamin D might exert its effect on obesity through different cell-signaling mechanisms, including prostaglandins, matrix metalloproteinases, mitogen-activated protein kinase pathways, reactive oxygen species, and nitric oxide synthase. The interactions of dietary vitamins with gene expression in the context of obesity need further investigations.

Minerals

Although the intake levels of different minerals have been associated with beneficial or deleterious effects on the control of body weight and adiposity, much is unknown about their molecular mechanisms and regulatory effects on gene expression. For example, chromium picolinate has been suggested to be involved in regulating glucose metabolism and weight loss. However, randomized controlled trials have not found evidence about the efficacy and safety of chromium picolinate supplementation in overweight or obese subjects.

High dietary selenium intake has been associated with a beneficial body composition profile, whereas high dietary magnesium intake has been associated

with low insulin resistance. Phosphorus status has been inversely correlated with body weight, and phosphorus supplementation might decrease body weight, BMI, waist circumference, and subjective appetite scores. Although some data suggest that dietary calcium intake may have a role in human body weight regulation, calcium dietary supplementation has not usually elicited statistically or clinically significant effects on weight loss.

The connection between iron and excessive adiposity has also received much research interest because iron deficiency has been usually found in obese individuals. However, low plasma zinc level has been associated with obesity, and a systematic review concluded that zinc supplementation could improve insulin resistance in obese individuals of both sexes. Finally, iodine is known to have an important role in treating obesity, because it might contribute to increase energy expenditure by stimulating the thyroid gland. Investigation of the role of the different dietary minerals in obesity through transcriptomic mechanisms may contribute to interpreting excessive fat deposition better.

Bioactive Compounds

Nutrigenomics also helps in understanding how bioactive dietary components influence human health by altering gene expression profile in cells and tissues. Several bioactive food components have been described to regulate the expression of key metabolic gene and proteins; they are potential strategies for obesity prevention and treatment, and include lipoic acid and resveratrol. In fact, both have been shown to regulate genes and proteins involved in food intake regulation, adipogenesis, lipogenesis, lipolysis, fatty acid oxidation, mitochondrial biogenesis, insulin sensitivity, and glucose homeostasis in important metabolic organs such as liver, adipose tissue, and muscle. Like other phenolic compounds, resveratrol has received widespread interest because of its ability to mimic calorie restriction. Resveratrol has demonstrated antiobesity effects in mice owing to its ability to inhibit adipogenesis, suppress lipogenesis, stimulate lipolysis, promote apoptosis, increase fatty acid oxidation and thermogenesis, and induce WAT browning. Some of these effects probably depend on the ability to activate AMPK, a critical enzyme regulating cellular energy metabolism. However, few human studies have found antiobesity effects for resveratrol, in part because of its low bioavailability. One study reported that resveratrol supplementation for 30 days significantly decreased adipocyte size in obese men. Microarray analysis revealed downregulation of Wnt and Notch signaling pathways and upregulation of pathways involved in cell cycle regulation, which

suggests an increase in adipogenesis. Moreover, the lysosomal–phagosomal pathway and transcription factor EB were upregulated, reflecting an alternative pathway of lipid breakdown by autophagy. However, more studies in humans are necessary to confirm the potential benefits of resveratrol supplementation (or other related phytochemicals) in human obesity.

Higher dietary choline and betaine intakes have been associated with better body composition and low insulin resistance. Both compounds may act as methyl donors and, like other lipotropic compounds, have been shown to attenuate nonalcoholic fatty liver disease and steatosis, although the molecular mechanisms must be elucidated.

Some alkaloids, such as capsaicin, caffeine, ephedrine, and nicotine, have also been related to weight loss, mainly through thermogenic effects.

Finally, several plant extracts rich in saponins, alkaloids, and phenolic compounds are being studied as potential weight loss aids. These extracts are sources of individual compounds whose effects are being more thoroughly analyzed. For instance, (–)-hydroxycitric acid from *Garcinia cambogia* seems to act as a potential appetite suppressant that could also inhibit adipocyte differentiation and reduce fatty acid synthesis.

Probiotics and Gut Microbiota Composition

Perturbations of microbial communities may increase the predisposition to different disease phenotypes. Numerous factors have been shown to influence gut microbiota composition, including diet, disease, age, genetics, and epigenetics. Among them, gut microbiome is extremely dependent on diet composition and diversity. Some food components may be converted into metabolites by intestinal microbes; they serve as biologically active molecules affecting diverse functions in the host, including the regulation of gene expression. It was reported that probiotics may restore the composition of the gut microbiome and induce beneficial functions to gut microbial communities, resulting in amelioration or prevention of gut inflammation and other intestinal or systemic disease phenotypes, but also metabolic diseases such as obesity-related complications. Intestinal microbes can affect the expression of many genes in the mammalian gut mucosa, including genes involved in immunity, nutrient absorption, energy metabolism, and intestinal barrier function. Some of these genes could be related to the chronic low-grade inflammatory state that leads to insulin resistance in obesity. Some of the effects attributed to the microbiota are mediated by postbiotics produced by microorganisms, such as short chain fatty acids (butyrate, acetate, and propionate), indole, or γ -aminobutyric acid. It was also reported

that gut microbiota modulates host miRNA expression profile, because mice colonized with specific pathogen free microbiota showed differential expression of nine miRNAs in their ileum and colon compared with germ-free controls. The gut microbiota–miRNA interactions are composed of two processes: (1) host-secreted miRNAs regulate gut microbiota composition; and (2) the gut microbiota affects the host by inducing special miRNAs expression (i.e., by production of specific metabolites).

Technological advances in RNA-seq are facilitating the possibility of gaining insights into genes that are actively expressed in complex microbial communities, such as the large intestine, enabling the elucidation of functional changes that regulate the microbiome functions in relation to different dietary patterns, its interactions with the host, and functional alterations associated with the disease-driving configuration of the microbiome.

Finally, apart from microbiota-produced miRNAs, food-derived miRNA (i.e., from rice and milk) have been described as being absorbed in the human intestine, which demonstrates a new possible role of food beyond minerals, vitamins, macronutrients, and bioactive compounds. Although food-derived miRNAs still represent a contradictory theme, they are a potential candidate as possible epigenetic communicators between the diet and host metabolism.

NUTRIGENOMICS IN PRECISION NUTRITION IN OBESITY

The integration of nutrigenomic, nutrigenetic, epigenetic, metagenomic, proteomic, and metabolomic data may help to design more personalized dietary treatments to prevent chronic diseases and optimize the individual's response to dietary interventions (Fig. 58.2).

This revolutionary approach has been proposed to improve the efficacy of interventions by better adapting to the nutritional requirements and metabolism of each subject. Gene expression data in adipose cells before the consumption of a low-fat diet have been reported to predict the success of weight loss treatment early. Also, adipose tissue transcriptome has been able to differentiate subjects with continued weight loss and subjects regaining weight 6 months after caloric restriction independent of energy intake, predicting subjects who experience successful short-term weight maintenance compared with those who experience weight regain. However, the use of adipose biopsies is not a good option to perform biomarker analyses, and research has focused on RNA obtained from less invasive samples. For example, gene expression profiles at baseline in peripheral blood mononuclear cells

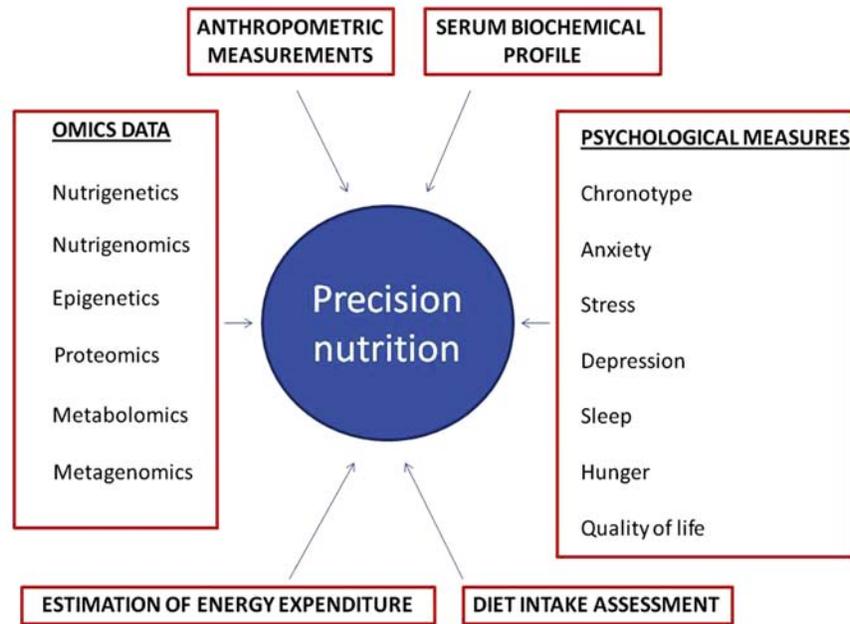


FIGURE 58.2 Integration of different “omics” technologies that, in addition to more conventional biomarkers, will be useful in implementing precision nutrition.

(PBMC) have been able to differentiate obese boys who are high and low responders to a moderate energy-restricted diet. In a similar way, basal expression of different miRNAs (such as mir-935 and mir-4772) measured by high-throughput sequencing in PBMC has been reported to forecast the response to a hypocaloric diet. Much effort is also being made to identify exosome miRNAs and proteins that could also predict the response to different diets and therapeutic strategies against metabolic diseases.

Studies have demonstrated the existence of several epigenetic markers that may modify gene expression and could be involved in the process of weight gain, the response to different diets in relation to other lifestyle factors, and the outcome of weight loss interventions. Studies found that BMI is associated with widespread changes in DNA methylation, but alterations in DNA methylation are predominantly the consequence of adiposity rather than the cause. On the other hand, disturbances in DNA methylation may predict future development of type 2 diabetes and other metabolic complications.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

In the next years, more studies will deepen the relations among genetic background, epigenetic processes, and gene expression regulation in response to different

nutritional states and interventions. They will undoubtedly boost the new field of precision nutrition to fight obesity and its related diseases better.

High-throughput “omics” and system biology studies hold the potential to improve our knowledge dramatically about nutrigenomics. They are helping to develop early diagnostic methods and effective therapeutic strategies for diet-related diseases, including obesity and its comorbidities, in the form of precision medicine.

Suggested Reading

- Abete, I., Navas-Carretero, S., Marti, A., Martinez, J.A., 2012. Nutrigenetics and nutrigenomics of caloric restriction. *Prog Mol Biol Transl Sci* 108, 323–346.
- Armenise, C., Lefebvre, G., Carayol, J., Bonnel, S., Bolton, J., Di Cara, A., Gheldof, N., Descombes, P., Langin, D., Saris, W.H., Astrup, A., Hager, J., Viguier, N., Valsesia, A., 2017. Transcriptome profiling from adipose tissue during a low-calorie diet reveals predictors of weight and glycemic outcomes in obese, nondiabetic subjects. *Am J Clin Nutr* 106, 736–746.
- Gao, X., Salomon, C., Freeman, D.J., 2017. Extracellular vesicles from adipose tissue—A potential role in obesity and type 2 diabetes? *Front Endocrinol* 8, 202.
- Gaye, A., Doumatey, A.P., Davis, S.K., Rotimi, C.N., Gibbons, G.H., 2018. Whole-genome transcriptomic insights into protective molecular mechanisms in metabolically healthy obese African Americans. *NPJ Genom Med* 3, 4.
- Gillberg, L., Perflyev, A., Brøns, C., Thomasen, M., Grunnet, L.G., Volkov, P., Rosqvist, F., Iggman, D., Dahlman, I., Risérus, U., Rönn, T., Nilsson, E., Vaag, A., Ling, C., 2016. Adipose tissue transcriptomics and epigenomics in low birthweight men and controls: role of high-fat overfeeding. *Diabetologia* 59, 799–812.

- Goni, L., Cuervo, M., Milagro, F.I., Martínez, J.A., 2016. Future perspectives of personalized weight loss interventions based on nutrigenetic, epigenetic, and metagenomic data. *J Nutr* 146, 905–912S.
- González-Plaza, J.J., Santiago-Fernández, C., Gutiérrez-Repiso, C., García-Serrano, S., Rodríguez-Pacheco, F., Ho-Plagaro, A., Garrido-Sánchez, L., Moreno-Ruiz, F.J., Rodríguez-Cañete, A., García-Fuentes, E., 2018. The changes in the transcriptomic profiling of subcutaneous adipose tissue after bariatric surgery depend on the insulin resistance state. *Surg Obes Relat Dis* 14, 1182–1191.
- Kulyté, A., Ehrlund, A., Arner, P., Dahlman, I., 2017. Global transcriptome profiling identifies KLF15 and SLC25A10 as modifiers of adipocytes insulin sensitivity in obese women. *PLoS One* 12, e0178485.
- Liu, B., Qian, S.B., 2011. Translational regulation in nutrigenomics. *Adv Nutr* 2, 511–519.
- Lorente-Cebrián, S., González-Muniesa, P., Milagro, F.I., Martínez, J.A., 2019. MicroRNAs and other non-coding RNAs in adipose tissue and obesity: emerging roles as biomarkers and therapeutic targets. *Clin Sci* 133, 23–40.
- Martínez-Fernández, L., Laiglesia, L.M., Huerta, A.E., Martínez, J.A., Moreno-Aliaga, M.J., 2015. Omega-3 fatty acids and adipose tissue function in obesity and metabolic syndrome. *Prostag Other Lipid Mediat* 121, 24–41.
- Massaro, M., Scoditti, E., Carluccio, M.A., Campana, M.C., De Caterina, R., 2010. Omega-3 fatty acids, inflammation and angiogenesis: basic mechanisms behind the cardioprotective effects of fish and fish oils. *Cell Mol Biol* 56, 59–82.
- Mutch, D.M., Temanni, M.R., Henegar, C., Combes, F., Pelloux, V., Holst, C., Sørensen, T.I., Astrup, A., Martínez, J.A., Saris, W.H., Viguerie, N., Langin, D., Zucker, J.D., Clément, K., 2007. Adipose gene expression prior to weight loss can differentiate and weakly predict dietary responders. *PLoS One* 2, e1344.
- Ramos-López, O., Milagro, F.I., Allayee, H., Chmurzynska, A., Choi, M.S., Curi, R., Caterina, R.De, Ferguson, L.R., Goni, L., Kang, J.X., Kohlmeier, M., Marti, A., Moreno, L.A., Pérusse, L., Prasad, C., Qi, L., Reifen, R., Riezu-Boj, J.I., San-Cristobal, R., Santos, J.L., Martínez, J.A., 2017. Guide for current nutrigenetic, nutrigenomic, and nutriepigenetic approaches for precision nutrition involving the prevention and management of chronic diseases associated with obesity. *J Nutrigenetics Nutrigenomics* 10, 43–62.
- Reynés, B., Palou, M., Palou, A., 2017. Gene expression modulation of lipid and central energetic metabolism related genes by high-fat diet intake in the main homeostatic tissues. *Food Funct* 8, 629–650.
- Wahl, S., Drong, A., Lehne, B., Loh, M., Scott, W.R., et al., 2017. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 541, 81–86.

Nutrients and Gene Expression in Type 2 Diabetes

Dominique Langin^{1,2,3,4}, *Nathalie Viguerie*^{1,2,3}

¹Institut National de la Santé et de la Recherche Médicale (Inserm), UMR1048, Institute of Metabolic and Cardiovascular Diseases, Toulouse, France; ²University of Toulouse, UMR1048, Institute of Metabolic and Cardiovascular Diseases, Paul Sabatier University, Toulouse, France; ³Franco-Czech Laboratory for Clinical Research on Obesity, Third Faculty of Medicine, Prague and Paul Sabatier University, Toulouse, France; ⁴Toulouse University Hospitals, Laboratory of Clinical Biochemistry, Toulouse, France

Glossary

BCAA branched-chain amino acid
ChREBP carbohydrate-response element-binding protein
FFA free fatty acid receptor
LXR liver X receptor
PPAR peroxisome proliferator-activated receptor
RXR retinoid X receptor
SREBP sterol-responsive element binding protein
TLR Toll-like receptor

The chapter sums up current knowledge on the influence of nutrients on gene expression in type 2 diabetes, with an emphasis on human studies.

NUTRIGENOMICS AND TYPE 2 DIABETES

Nutrition has a profound impact on human health, especially in the setting of metabolic diseases such as obesity and type 2 diabetes. The emergence of “omics” technologies has pushed the field forward and profoundly improved the understanding of gene–nutrient interactions. In this chapter, we will focus on gene expression (i.e., transcriptomics) studies in humans.

For chronic diseases such as type 2 diabetes, there is a complex interaction between environmental and genetic factors. The global surge in the prevalence of type 2 diabetes is generally considered to be caused by the rapid and widespread adoption of unhealthy lifestyle habits, including poor nutrition and lack of physical activity. Diets with highly refined sugar and saturated fat content have favored an increase in fat mass, insulin resistance,

and type 2 diabetes. However, the genetic susceptibility to developing the disease is high. This is exemplified by studies of different populations and ethnic groups (e.g., Greenlanders and Pima Indians), characterized by a high prevalence of diabetes. The susceptibility to developing diabetes in a given environment is highly variable. Diabetic types and progression toward complications are heterogeneous. There is also considerable variation in the response to lifestyle therapies, as well as to drugs. Personalized medicine may help in that context to improve the assessment of susceptibility to adverse lifestyle exposures, complications, and treatment response. Nutrigenomics may specifically contribute to personalized care in type 2 diabetes through the study of how food components modulate the expression of genetic information in an individual, and how an individual’s genetic makeup affects the response to nutrients and other bioactive components present in food.

Type 2 diabetes features a strong heritability; there is an approximately 70% lifetime risk of developing disease when both parents are affected. Genome-wide association studies have identified common variants associated with type 2 diabetes, which in aggregate contribute to a fraction of the disease heritability. Rare variants not detected by genome-wide association studies have been hypothesized to explain the missing part. Whereas such low-frequency alleles have been identified, especially in specific populations, they do not have a major role in the predisposition to type 2 diabetes (Fuchsberger et al., 2016). Unlike monogenic

forms, genetic testing of individuals for type 2 diabetes has little clinical relevance.

The identification of gene–environment interactions on complex traits in humans has proven to be difficult. However, some success has been achieved in exploring the regulation of gene expression through analyses of expression quantitative trait loci. Studies on the deCODE cohort revealed that unlike blood, adipose tissue gene expression profiles were strongly associated with obesity- and diabetes-related traits. Genome-wide linkage and association mapping revealed a highly significant genetic component to adipose gene expression traits (Emilsson et al., 2008). This tissue specificity was confirmed in a multitissue transcriptomic study that identified *cis* and *trans* gene–body mass index interactions only in adipose tissue (Glastonbury et al., 2016).

MACRONUTRIENTS AND TYPE 2 DIABETES

Several mechanisms, such as lipotoxicity, glucotoxicity, oxidative stress, and inflammation, have been involved in the pathogenesis of type 2 diabetes. Most of these mechanisms are directly controlled by nutrients at the transcriptional level. Nutrients or their metabolites may bind transcription factors, notably nuclear receptors, and influence the expression of genes involved in metabolic and inflammatory processes.

The peroxisome proliferator-activated receptors (PPARs) superfamily is one of the most investigated groups of nutrient sensors. The three members of the family act as heterodimers with the retinoid X receptor (RXR). PPARs control both metabolism and inflammation in a variety of tissues (Gross et al., 2017). PPAR α , the first PPAR to be identified, is predominantly expressed in the liver, heart, and brown adipose tissue. In the liver, it regulates various pathways of lipid and lipoprotein metabolism. There is ample evidence that PPAR α is the master regulator of lipid metabolism during fasting. It is also involved in the control of cholesterol, glucose, and bile acid metabolism. In the heart, PPAR α is essential for optimal substrate oxidation and lipid handling. In human brown-like adipocytes, PPAR α induces fatty acid anabolic and catabolic pathways, while repressing glucose oxidation. Of the three PPARs, PPAR β/δ shows the widest tissue distribution and is the predominantly expressed isotype in the skeletal muscle, integrating signals from physical exercise and fasting. It has a role in skeletal muscle remodeling, with a switch from glycolytic to oxidative fiber types, an increase in the capillary-to-fiber ratio, and the formation of new myofibers. PPAR γ , notably the adipose tissue-restricted PPAR γ 2 isoform, has a major role in fat cells. It is essential for adipogenesis and controls metabolism

in both energy-storing white and energy-dissipating brown adipocytes. One of the PPAR transcriptional cofactors, PPAR γ coactivator 1 α , initially described as a metabolic regulator of adaptive thermogenesis in brown adipose tissue, has been shown to act as a nutrient sensor that is essential for controlling mitochondriogenesis, fatty acid oxidation, and hepatic gluconeogenesis. In both the liver and skeletal muscle, it has a central role in adaptation to fasting. Ligands of the three PPAR isotypes exhibit antiinflammatory properties. This nonmetabolic functions can be ascribed to the transrepression of genes of proinflammatory mediators. PPARs can also interact with other transcription factors in a manner independent of DNA-binding and modulate inflammatory pathways. With respect to nutrient control of PPARs, it is puzzling that despite years of intense work, the precise nature of endogenous ligands remains elusive. As a unifying feature of most nuclear receptors, PPARs are activated by small hydrophobic ligands. Natural ligands of PPARs include fatty acids, eicosanoids, endocannabinoids, and phospholipids derived endogenously from cellular metabolism or exogenously from dietary lipids. Most show low affinity for PPARs. The distribution of these ligands in the body and the combination in which they occur in a given cell are highly variable, as is the combination of the three PPARs in a cell nucleus.

Liver X receptors (LXRs) α and β act as heterodimers with RXR. Whereas LXR β distribution is ubiquitous, LXR α is highly expressed in the liver, but it is also detected in the intestine, adipose tissue, kidneys, adrenals, and macrophages. LXRs function as intracellular cholesterol sensors and bind derivatives of cholesterol such as oxysterols. In response to nutrients, a number of genes, such as those involved in cholesterol conversion to bile acids, fatty acid synthesis, and secretion of very low–density lipoproteins are induced. LXRs function as a critical signaling node linking lipid metabolism, inflammation, and immune cell function.

Sterol-responsive element binding proteins (SREBPs) are another family of transcription factors controlling cholesterol and lipid metabolism (Wang et al., 2015). They activate a cascade of enzymes required for endogenous cholesterol, fatty acid, triglyceride, and phospholipid synthesis. Three members of the SREBP family have been described in several mammalian species: SREBP1a and 1c, produced from a single gene, and SREBP2, from a separate gene. SREBP2 mainly controls the expression of genes involved in cholesterol biosynthesis, whereas SREBP1a and SREBP1c induce transcription of genes involved in fatty acid synthesis. Cholesterol and derivatives modulate the processing of SREBP in the endoplasmic reticulum and its nuclear translocation that allows activation of target genes. Insulin controls SREBP1c expression and thereby de novo

lipogenesis (i.e., the conversion of glucose carbons into fatty acid). A complex interplay between LXRs and SREBPs exists for the control of hepatic gene expression.

Together with SREBP1c, carbohydrate-response element-binding protein (ChREBP) controls de novo lipogenesis in the liver and the adipose tissue. It acts as a glucose-responsive transcription factor (Filhoulaud et al., 2013). Two isoforms of ChREBP are generated by alternative promoter use in the *MLXIPL* gene. Glucose posttranslationally activates ChREBP α , which induces ChREBP β , a more potent transcriptional activator. They activate transcription of target genes as heterodimers with Max-like protein X. Whereas the induction of de novo lipogenesis contributing to the development of fatty liver is often viewed as a key event in the pathogenesis of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis, ChREBP induction of this pathway in the adipose tissue is associated with improved insulin sensitivity and may therefore prevent the development of insulin resistance and type 2 diabetes.

Amino acid regulation of gene expression occurs at distinct levels: transcriptional control, messenger RNA (mRNA) stabilization under starved conditions, and regulation of the translation rate by amino acid availability. In type 2 diabetes, branched-chain amino acids (BCAAs) deserve special attention (Bifari and Nisoli, 2017). These essential amino acids are critical nutrient signals that affect metabolism, either directly or indirectly. BCAA-rich diets have been reported to improve metabolic health, including the regulation of body weight, muscle protein synthesis, and glucose homeostasis. Paradoxically, plasma levels of BCAAs positively correlate with increasing risk of insulin resistance and type 2 diabetes in humans. The latter observation may be related to impaired BCAA catabolism. The effect of BCAAs drastically changes when they act under catabolic or anabolic conditions. In catabolic states, BCAAs can behave as energy substrates directly oxidized in the muscle or converted to gluconeogenic-ketogenic substrates. In contrast, under anabolic conditions, BCAAs stimulate protein synthesis and cell growth. Therefore, to predict their effects accurately, the overall catabolic or anabolic status of patients should be known.

Besides nuclear receptors, receptors present at the cell membrane recognize nutrients and transduce intracellular signals modulating insulin signaling. Extensive research has been conducted on fatty acids. Saturated fatty acids enhance inflammatory pathways through Toll-like receptor (TLR)-dependent and independent mechanisms, thereby promoting insulin resistance (Glass and Olefsky, 2012). Bacterial lipids activate TLRs. The main TLR4 ligand is lipopolysaccharide, a lipid component of the walls of gram-negative bacteria, whereas TLR2 ligands include the lipoteichoic acid, a component of gram-positive bacteria. Hence, diet-

induced changes in gut microbiota may modulate immune response through bacterial lipids. Several G protein-coupled receptors, previously considered to be orphan receptors, are activated by endogenous and dietary fatty acids. As an example, the free fatty acid receptors (FFA1–4), which are expressed in various cell types including pancreatic β cells, have provided exciting new data in nutrient sensing (Miyamoto et al., 2016). FFA1-4 bind different fatty acid species. FFA2 and FFA3 are activated by short-chain fatty acids such as acetate, propionate, and butyrate, which are the primary metabolic byproducts of anaerobic fermentation by intestinal microflora. FFA1 and FFA4 are both activated by medium- and long-chain fatty acids, including the ω 3 polyunsaturated fatty acids, primarily α -linolenic acid, eicosapentaenoic acid, and docosahexaenoic acid. ω 3 fatty acids produce antiinflammatory effects by stimulating FFA4, also known as GPR120, and the production of resolvins and protectins. Preclinical models show the ability of FFA4 agonists to improve glucose disposal and enhance insulin sensitivity.

MICRONUTRIENTS AND TYPE 2 DIABETES

Some vitamins with direct relevance to type 2 diabetes exert effects through nuclear receptors. Vitamin A derivatives, mainly retinoic acid, regulate gene expression by activating retinoic acid receptors and RXRs (Zhang et al., 2015). Because PPAR γ /RXR α heterodimers can also be activated by RXR agonists, they constitute attractive candidates for the treatment of type 2 diabetes. Retinoids have been shown to modulate gene expression in several insulin-sensitive tissues, notably the liver and adipose tissue, where they have a crucial role in adipogenesis. However, there is still no clear indication for the routine recommendation of vitamin A supplements in the management of type 2 diabetes.

Low vitamin D intake has been associated with a higher incidence of type 2 diabetes (Boucher, 2011). Vitamin D acts through vitamin D receptor–RXR heterodimers. It may enhance insulin sensitivity by modulating gene expression in the skeletal muscle and adipose tissue, and it affects insulin secretion by controlling calcium concentration and flux in pancreatic β cells. Metaanalyses of observational studies show an association between vitamin D status and the prevalence of glucose intolerance or type 2 diabetes. Randomized controlled trials investigating the effect of vitamin D supplementation on glucose homeostasis gave mixed results, from lack of effect to a positive impact.

Disturbed iron homeostasis is associated with hyperglycemia and diabetes. Elevated body iron stores and serum ferritin are risk factors for type 2 diabetes (Fernandez-Real et al., 2015). The liver is the major reservoir of iron in the body. Excess iron storage in the liver interferes with glucose metabolism, causing hyperinsulinemia, with both decreased insulin extraction and impaired insulin signaling. A hyperinsulinemic status, on the other hand, favors the intrahepatic deposition of iron. Iron deficiency is associated with enhanced hepatic glucose production and increased SREBP1c-controlled de novo lipogenesis. Iron also controls adipose tissue function both at the fat cell level, by modulating adipogenesis, and at the macrophage level, by modulating polarization and inflammatory response. Whereas key aspects of systemic iron metabolism are regulated transcriptionally (e.g., hepatic expression of hepcidin, a hormone-regulating iron fluxes and levels) and posttranslationally (e.g., function of the iron exporter ferroportin by hepcidin), cellular iron homeostasis is regulated posttranscriptionally by iron regulatory proteins. These RNA-binding proteins interact with conserved cis-regulatory hairpin structures present in the 5'- or 3'-untranslated regions of target mRNAs.

CALORIE RESTRICTION AND INSULIN RESISTANCE

Calorie restriction is a standard lifestyle therapy in obesity management and prevention of type 2 diabetes. Calorie restriction-induced weight loss improves the metabolic profile of most obese individuals. In terms of transcriptomics, hypocaloric diets have been the most comprehensively studied dietary interventions; adipose tissue gene expression profiles were thoroughly investigated. Several conclusions may be drawn from those studies. First, significant changes in adipose tissue gene expression profiles were observed with very low to low-calorie diets (i.e., diets with energy intake of 3.3 MJ [800 kcal] per day to diets with an energy deficit of 2.5 MJ [600 kcal] per day less than the individually estimated daily energy requirement). Expression of metabolism genes in fat cells is decreased, whereas expression of immune genes in macrophages is increased or unchanged (Capel et al., 2009). Second, energy restriction has a much more pronounced impact on variations in human adipose tissue gene expression than macronutrient composition (Capel et al., 2008; Viguerie et al., 2012). Differences in nutrient composition are less pronounced in human studies than in most rodent studies. Third, when very low-calorie diets are followed by weight maintenance diets, either ad libitum or isocaloric, an opposite regulation of gene expression is observed between the two phases (Capel et al., 2009;

Viguerie et al., 2012). The associations between adipose tissue gene expression and insulin sensitivity are different in the various phases of dietary interventions.

CONCLUDING REMARKS

Several specificities of human nutrigenomic studies need to be considered in the field of type 2 diabetes. Unlike studies of diets in model organisms, whose compositions are often extreme, differences in human trials of nutrient composition between diets usually remain within dietary recommendations and nutritional requirements defined by the Food and Agriculture Organization of the United Nations/World Health Organization or regional and national scientific societies. In interventional studies, the number of subjects is often too low to capture moderate effects. This is also a limitation for investigating genotype–nutrient interactions and the genetic control of gene expression. Because study design and study populations are never the same, true replication is rarely achieved. Unlike drug trials, there is no placebo arm. The studies are often not blinded. Moreover, the interdependence of nutrients needs to be considered (e.g., the proportions of fat and carbohydrate, changing in parallel during hypocaloric diets). Another limitation pertains to the availability of tissues. Blood is not the most suitable tissue for investigating metabolic diseases. Tissues involved in the pathogenesis of type 2 diabetes, such as the liver, skeletal muscle, and pancreas, are rarely available, and notably so in dietary interventions, when biopsies should be performed at multiple time points. The major exception is adipose tissue. Microbiopsy of subcutaneous fat allows fast and painless adipose tissue sampling, providing enough high-quality total RNA for high-throughput quantitative polymerase chain reaction, RNA sequencing, or DNA microarray analyses (Viguerie et al., 2012). Studies on this tissue have proven highly informative in investigating the effects of calorie restriction.

Most of these caveats are now being addressed. The number of participants in carefully controlled dietary interventions is increasing. In some of these trials (e.g., CALERIE and DiOGenes), consecutive collections of fat and skeletal muscle samples were adopted. Less invasive methods are becoming common for performing microbiopsies of adipose tissue, skeletal muscle, and liver. Integration of “omics” data is also a challenge that shows considerable development. Cross-talk between metabolic organs and gut microbiota is emerging as an important determinant of dietary responses and metabolic alterations (Zeevi et al., 2015). The same is occurring in epigenetic studies bridging nutrigenetics and nutrigenomics. Thus, there is promise that such developments will shed much more relevant light in this new area of science.

References

- Bifari, F., Nisoli, E., 2017. Branched-chain amino acids differently modulate catabolic and anabolic states in mammals: a pharmacological point of view. *Br J Pharmacol* 174, 1366–1377.
- Boucher, B.J., 2011. Vitamin D insufficiency and diabetes risks. *Curr Drug Targets* 12, 61–87.
- Capel, F., Klimcakova, E., Viguerie, N., Roussel, B., Vitkova, M., Kovacicova, M., Polak, J., Kovacova, Z., Galitzky, J., Maoret, J.J., Hanacek, J., Pers, T.H., Bouloumie, A., Stich, V., Langin, D., 2009. Macrophages and adipocytes in human obesity: adipose tissue gene expression and insulin sensitivity during calorie restriction and weight stabilization. *Diabetes* 58, 1558–1567.
- Capel, F., Viguerie, N., Vega, N., Dejean, S., Arner, P., Klimcakova, E., Martinez, J.A., Saris, W.H., Holst, C., Taylor, M., Oppert, J.M., Sorensen, T.I., Clement, K., Vidal, H., Langin, D., 2008. Contribution of energy restriction and macronutrient composition to changes in adipose tissue gene expression during dietary weight-loss programs in obese women. *J Clin Endocrinol Metab* 93, 4315–4322.
- Emilsson, V., Thorleifsson, G., Zhang, B., Leonardson, A.S., Zink, F., Zhu, J., Carlson, S., Helgason, A., Walters, G.B., Gunnarsdottir, S., Mouy, M., Steinthorsdottir, V., Eiriksdottir, G.H., Bjornsdottir, G., Reynisdottir, I., Gudbjartsson, D., Helgadóttir, A., Jonasdottir, A., Jonasdottir, A., Styrkarsdottir, U., Gretarsdottir, S., Magnusson, K.P., Stefansson, H., Fossdal, R., Kristjansson, K., Gislason, H.G., Stefansson, T., Leifsson, B.G., Thorsteinsdottir, U., Lamb, J.R., Gulcher, J.R., Reitman, M.L., Kong, A., Schadt, E.E., Stefansson, K., 2008. Genetics of gene expression and its effect on disease. *Nature* 452, 423–428.
- Fernandez-Real, J.M., McClain, D., Manco, M., 2015. Mechanisms linking glucose homeostasis and iron metabolism toward the onset and progression of type 2 diabetes. *Diabetes Care* 38, 2169–2176.
- Filhoulaud, G., Guilmeau, S., Dentin, R., Girard, J., Postic, C., 2013. Novel insights into ChREBP regulation and function. *Trends Endocrinol Metabol* 24, 257–268.
- Fuchsberger, C., Flannick, J., Teslovich, T.M., Mahajan, A., Agarwala, V., Gaulton, K.J., Ma, C., Fontanillas, P., Moutsianas, L., McCarthy, D.J., Rivas, M.A., Perry, J.R., Sim, X., Blackwell, T.W., Robertson, N.R., Rayner, N.W., Cingolani, P., Locke, A.E., Fernandez Tajos, J., Highland, H.M., Dupuis, J., Chines, P.S., Lindgren, C.M., Hartl, C., Jackson, A.U., Chen, H., Huyghe, J.R., van de Bunt, M., Pearson, R.D., Kumar, A., Muller-Nurasyid, M., Grarup, N., Stringham, H.M., Manning, A., Ng, M.C., Palmer, N.D., Balkau, B., Stancakova, A., Abboud, H.E., Boeing, H., Giedraitis, V., Prabhakaran, D., Gottesman, O., Scott, J., Carey, J., Kwan, P., Grant, G., Smith, J.D., Neale, B.M., Purcell, S., Butterworth, A.S., Howson, J.M., Lee, H.M., Lu, Y., Kwak, S.H., Zhao, W., Danesh, J., Lam, V.K., Park, K.S., Saleheen, D., So, W.Y., Tam, C.H., Afzal, U., Aguilar, D., Arya, R., Aung, T., Chan, E., Navarro, C., Cheng, C.Y., Palli, D., Correa, A., Curran, J.E., Rybin, D., Farook, V.S., Fowler, S.P., Freedman, B.I., Griswold, M., Hale, D.E., Hicks, P.J., Khor, C.C., Kumar, S., Lehne, B., Thuillier, D., Lim, W.Y., Liu, J., van der Schouw, Y.T., Loh, M., Musani, S.K., Puppala, S., Scott, W.R., Yengo, L., Tan, S.T., Taylor Jr, H.A., Thameem, F., Wilson Sr, G., Wong, T.Y., Njolstad, P.R., Levy, J.C., Mangino, M., Bonnycastle, L.L., Schwarzmayr, T., Fadista, J., Surdulescu, G.L., Herder, C., Groves, C.J., Wieland, T., Bork-Jensen, J., Brandslund, I., Christensen, C., Koistinen, H.A., Doney, A.S., Kinnunen, L., Esko, T., Farmer, A.J., Hakaste, L., Hodgkiss, D., Kravic, J., Lyssenko, V., Hollensted, M., Jorgensen, M.E., Jorgensen, T., Ladenvall, C., Justesen, J.M., Karajamaki, A., Kriebel, J., Rathmann, W., Lannfelt, L., Lauritzen, T., Narisu, N., Linneberg, A., Melander, O., Milani, L., Neville, M., Orholm-Melander, M., Qi, L., Qi, Q., Roden, M., Rolandsson, O., Swift, A., Rosengren, A.H., Stirrups, K., Wood, A.R., Mihailov, E., Blancher, C., Carneiro, M.O., Maguire, J., Poplin, R., Shakir, K., Fennell, T., DePristo, M., Hrabe de Angelis, M., Deloukas, P., Gjesing, A.P., Jun, G., Nilsson, P., Murphy, J., Onofrio, R., Thorand, B., Hansen, T., Meisinger, C., Hu, F.B., Isomaa, B., Karpe, F., Liang, L., Peters, A., Huth, C., O'Rahilly, S.P., Palmer, C.N., Pedersen, O., Rauramaa, R., Tuomilehto, J., Salomaa, V., Watanabe, R.M., Syvanen, A.C., Bergman, R.N., Bharadwaj, D., Bottinger, E.P., Cho, Y.S., Chandak, G.R., Chan, J.C., Chia, K.S., Daly, M.J., Ebrahim, S.B., Langenberg, C., Elliott, P., Jablonski, K.A., Lehman, D.M., Jia, W., Ma, R.C., Pollin, T.I., Sandhu, M., Tandon, N., Froguel, P., Barroso, I., Teo, Y.Y., Zeggini, E., Loos, R.J., Small, K.S., Ried, J.S., DeFronzo, R.A., Grallert, H., Glaser, B., Metspalu, A., Wareham, N.J., Walker, M., Banks, E., Gieger, C., Ingelsson, E., Im, H.K., Illig, T., Franks, P.W., Buck, G., Trakalo, J., Buck, D., Prokopenko, I., Magi, R., Lind, L., Farjoun, Y., Owen, K.R., Gloyn, A.L., Strauch, K., Tuomi, T., Kooner, J.S., Lee, J.Y., Park, T., Donnelly, P., Morris, A.D., Hattersley, A.T., Bowden, D.W., Collins, F.S., Atzmon, G., Chambers, J.C., Spector, T.D., Laakso, M., Strom, T.M., Bell, G.I., Blangero, J., Duggirala, R., Tai, E.S., McVean, G., Hanis, C.L., Wilson, J.G., Seielstad, M., Frayling, T.M., Meigs, J.B., Cox, N.J., Sladek, R., Lander, E.S., Gabriel, S., Burt, N.P., Mohlke, K.L., Meitinger, T., Groop, L., Abecasis, G., Florez, J.C., Scott, L.J., Morris, A.P., Kang, H.M., Boehnke, M., Altshuler, D., McCarthy, M.I., 2016. The genetic architecture of type 2 diabetes. *Nature* 536, 41–47.
- Glass, C.K., Olefsky, J.M., 2012. Inflammation and lipid signaling in the etiology of insulin resistance. *Cell Metab* 15, 635–645.
- Glastonbury, C.A., Vinuela, A., Buil, A., Halldorsson, G.H., Thorleifsson, G., Helgason, H., Thorsteinsdottir, U., Stefansson, K., Dermitzakis, E.T., Spector, T.D., Small, K.S., 2016. Adiposity-dependent regulatory effects on multi-tissue transcriptomes. *Am J Hum Genet* 99, 567–579.
- Gross, B., Pawlak, M., Lefebvre, P., Staels, B., 2017. PPARs in obesity-induced T2DM, dyslipidaemia and NAFLD. *Nat Rev Endocrinol* 13, 36–49.
- Miyamoto, J., Hasegawa, S., Kasubuchi, M., Ichimura, A., Nakajima, A., Kimura, I., 2016. Nutritional signaling via free fatty acid receptors. *Int J Mol Sci* 17, 450.
- Viguerie, N., Montastier, E., Maoret, J.J., Roussel, B., Combes, M., Valle, C., Villa-Vialaneix, N., Iacovoni, J.S., Martinez, J.A., Holst, C., Astrup, A., Vidal, H., Clement, K., Hager, J., Saris, W.H., Langin, D., 2012. Determinants of human adipose tissue gene expression: impact of diet, sex, metabolic status, and cis genetic regulation. *PLoS Genet* 8, e1002959.
- Wang, Y., Viscarra, J., Kim, S.J., Sul, H.S., 2015. Transcriptional regulation of hepatic lipogenesis. *Nat Rev Mol Cell Biol* 16, 678–689.
- Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M., Suez, J., Mahdi, J.A., Matot, E., Malka, G., Kosower, N., Rein, M., Zilberman-Schapira, G., Dohnalova, L., Pevsner-Fischer, M., Bikovsky, R., Halpern, Z., Elinav, E., Segal, E., 2015. Personalized nutrition by prediction of glycemic responses. *Cell* 163, 1079–1094.
- Zhang, R., Wang, Y., Li, R., Chen, G., 2015. Transcriptional factors mediating retinoic acid signals in the control of energy metabolism. *Int J Mol Sci* 16, 14210–14244.

This page intentionally left blank

Gene Expression in Dyslipidemias

Cátia Lira do Amaral¹, Rui Curi²

¹Universidade Estadual de Goiás, Campus de Ciências Exatas e Tecnológicas, Anápolis, Goiás, Brazil; ²Institute of Physical Activity and Sport Sciences, Cruzeiro do Sul University, Sao Paulo, Brazil

Glossary Terms

Apolipoproteins Proteins that are part of the lipoprotein structure (for example, apoAI, apoB100, and apoCII).

Lipids Organic compounds insoluble in water (for example, triglycerides, cholesterol, phospholipids, and fatty acids).

Lipoproteins Assembly of proteins and lipids that transport lipids bound to apolipoproteins in blood (for example, very low-density lipoprotein, low-density lipoprotein, high-density lipoprotein, and chylomicron).

density lipoprotein (VLDL), low-density lipoprotein (LDL), and high-density lipoprotein (HDL). The apolipoproteins are on the lipoprotein surface and the lipids form an amphiphilic layer. The lipoproteins are synthesized in distinct tissues and are used for the transport of lipids, as detailed in the legend of Fig. 60.1. The process involves several proteins; their changes in expression are associated with the development of dyslipidemias.

INTRODUCTION

Dyslipidemias are abnormal circulating concentrations of lipids that mainly manifest by increasing or decreasing cholesterol or triglycerides (TG). **Primary** dyslipidemias are associated with gene expression alterations. Other diseases, drug interventions, and environmental factors cause **secondary** dyslipidemias. Environmental factors also cause dyslipidemias and associated phenotypes. Some lipids are directly related to dyslipidemia, such as free fatty acids (FAs), TG (known as neutral lipids), free cholesterol or esterified to one FA (cholesteryl esters), and phospholipids (PL). Other lipid classes are eicosanoids, steroid hormones (estradiol and testosterone), sphingolipids (sphingosine, ceramide, cerebroside, and sphingomyelin), and glycerophospholipids (glycerol-3-phosphate, phosphatidic acid, and phosphatidylcholine) with important regulatory cell functions.

Lipids can be obtained through the diet or synthesized in the body from carbohydrates and/or proteins. The lipids are transported in the blood bound to proteins to be soluble in water. The FAs are transported bound to albumin. TG and cholesterol are transported in lipoproteins. The lipoproteins are categorized into four main types according to the composition of apolipoproteins and lipids: chylomicron (CM), very low-

GENES EXPRESSION CHANGES ASSOCIATED WITH THE DEVELOPMENT OF DYSLIPIDEMIA

Changes in the amount and proportion of lipoproteins are sometimes associated with changes in the expression of genes involved in lipid metabolism.

Genes Involved in Uptake of Dietary Lipids

Lipids ingested in the diet are hydrolyzed by lipases present in the gastrointestinal tract (Fig. 60.2). The **pancreatic lipase (PNLIP)** hydrolyzes long-chain TG to form FA, diglycerides and monoglycerides. At the congenital deficiency of PNLIP (OMIM 614338), the FA absorption occurs partly because of the presence of other intestinal lipases. After hydrolysis, the FAs are absorbed by passive diffusion. Carriers present in the luminal membrane of the enterocytes favor the absorption of lipids as the scavenger receptor CD-36 for FA and **Niemann-Pick C1-like intracellular cholesterol transporter 1 (NPC1L1)** for cholesterol (Fig. 60.2). This receptor is the target of the drug ezetimibe, which inhibits its activity. The expression of *NPC1L1* messenger RNA (mRNA) occurs primarily in enterocytes of the small intestine and hepatocytes. Patients with nonalcoholic fatty liver disease (NAFLD) feature hypermethylation at the

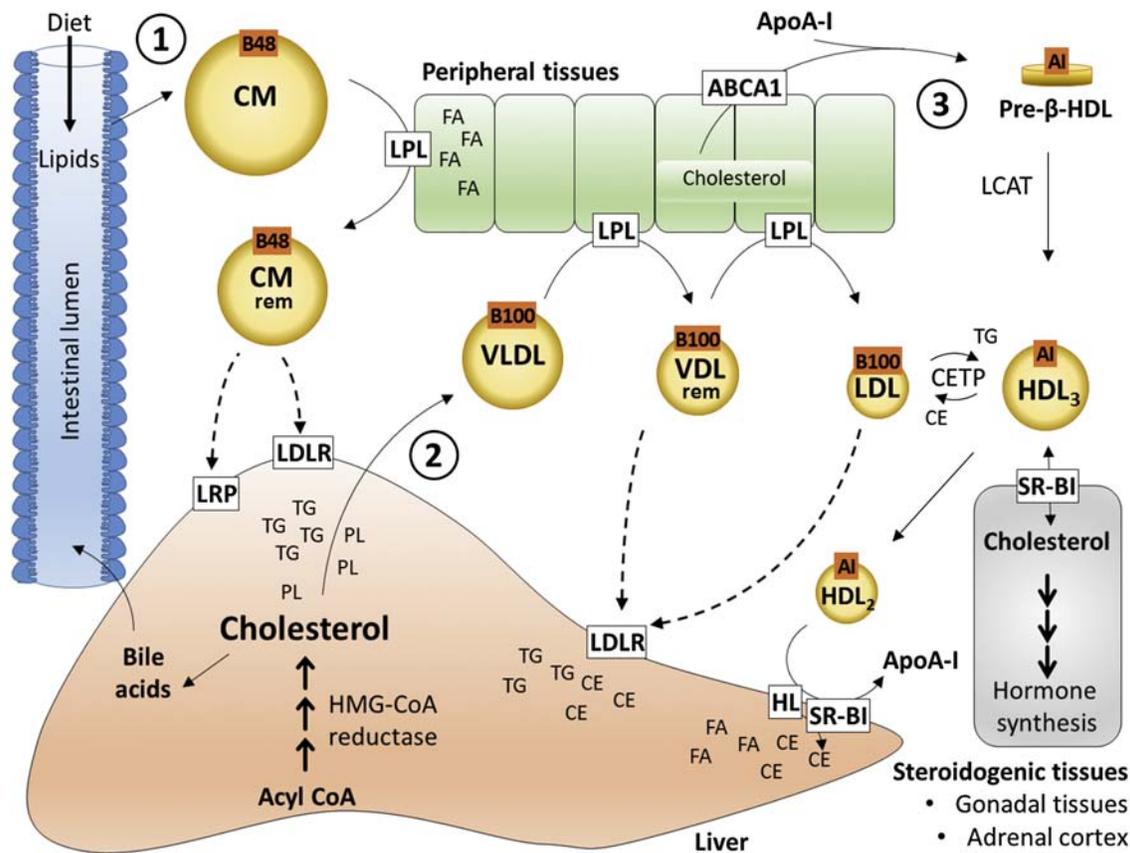


FIGURE 60.1 Summary of lipoprotein metabolism. (1) Lipids are absorbed by enterocytes and secreted with lipoprotein chylomicron (CM) that contains apolipoprotein B (apoB48). Lipoprotein lipase (LPL) located in peripheral tissues releases fatty acids (FA) from the CM to the tissues to form the CM remnant (CMrem). This is taken up by the liver through carriers such as LDL (low-density lipoprotein) receptor-related (LDLR) protein 1 (LRP1) and LDLR. (2) The liver synthesizes cholesterol that is exported with triglycerides (TG) and phospholipids (PL) through very low-density lipoprotein (VLDL). In peripheral tissues, LPL converts VLDL into VLDL remnant (VLDLrem). VLDLrem is then converted into LDL. The VLDL, VLDLrem, and LDL contain apoB100. This is recognized by hepatic receptors LDLR. (3) Reverse cholesterol transport is performed from peripheral tissues. The adenosine triphosphate binding cassette A1 (ABCA1) receptor secretes cholesterol that is removed by circulating apoA-I and forms pre-β-high-density lipoprotein (HDL). The enzyme lecithin: cholesterol acyltransferase (LCAT) promotes the esterification of cholesterol in cholesteryl esters (CE), forming a more oval particle, the HDL₃. The cholesteryl ester transfer protein (CETP) enzyme interacts with apoB-containing lipoproteins and performs the transfer of CE to LDL and TG to HDL. This transfer reduces the size of HDL. Successive actions of LCAT and CETP originate the HDL₂. The HDL is recognized by the scavenger receptor class B member 1 (SR-BI) receptor. This is expressed in steroidogenic tissues and the liver; CE is used in the synthesis of hormones and bile salts, respectively. apoB-containing lipoproteins also perform reverse cholesterol transport to the liver because they receive HDL cholesterol. Hepatic lipase contributes to hydrolyze TG present in HDL.

promoter region inversely correlated with the transcription. Epigenetic factors modulate the expression of the receptor. In the luminal membrane, there are also efflux carriers of **adenosine triphosphate binding cassette (ABC)** type that limit the absorption of steroids originating from animals or plants (Fig. 60.2). ABCG5 and ABCG8 function as dimers. Sitosterolemia (OMIM 210250) is a disease with autosomal recessive inheritance and mutation in one of the ABC receptors. Accumulations of steroids, especially those derived from plants such as campesterol, β-sitosterol, and stigmasterol, are reported. Because there is no cholesterol efflux in the intestine, patients have hypercholesterolemia, which can be mitigated with a lower intake of foods rich in steroids. In enterocytes of patients with type 2

diabetes, the mRNA expression of *ABCG5/ABCG8* is reduced, whereas that of *NPC1L1* is increased. This metabolic feature favors hypercholesterolemia normally present in patients with metabolic syndrome.

Genes Involved in Synthesis of Lipoproteins

Lipoproteins are synthesized mainly in the small intestine and liver. **Apolipoprotein B (apoB)** is incorporated into CM (apoB48) (Fig. 60.2) and in the VLDL (apoB100) (Fig. 60.3). The initial lipoprotein derivatives also feature apoB48 in CM remnants (CMrem) and apoB100 in VLDL remnants (VLDLrem) and LDL. The expression is controlled at the transcription and post-translation levels and modulates the content of apoB

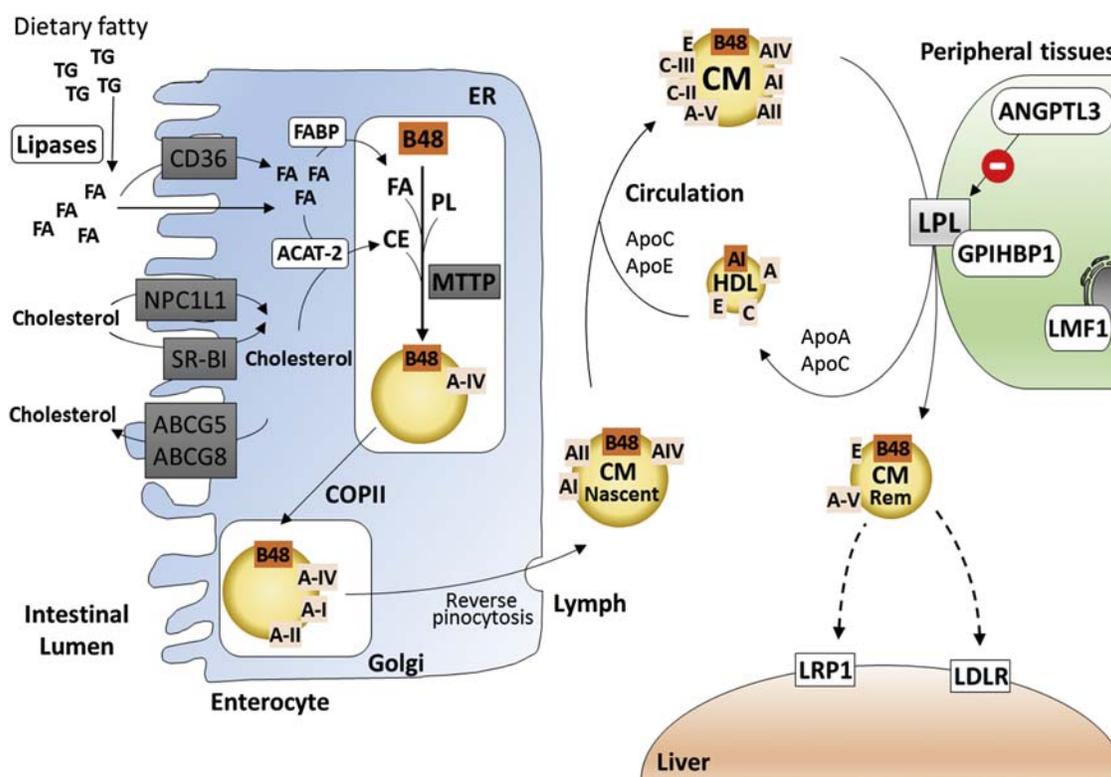


FIGURE 60.2 Synthesis and metabolism of chylomicrons. The diet provides mainly triglycerides (TG), phospholipids (PL), and cholesterol. The TGs are partially hydrolyzed to fatty acids (FA) by gastric, intestinal, and pancreatic lipases. The FAs are absorbed by simple diffusion or via carriers such as CD-36 located in the luminal membrane of the enterocytes. At this membrane, cholesterol is absorbed via carriers such as Niemann-Pick C1-like intracellular cholesterol transporter 1 (NPC1L1) and scavenger receptor class B member 1 (SR-BI), whereas adenosine triphosphate binding cassette (ABCG5) and ABCG8 carriers promote cholesterol active efflux, reducing absorption. The FAs are transported to the endoplasmic reticulum (ER) by fatty acid binding protein and reesterified to TG. ACAT-2 esterifies cholesterol with FAs and forms cholesterol ester (CE). In the ER, newly synthesized apoB48 incorporates TG, CE, phospholipids, and apoAIV through the action of microsomal triglyceride transfer protein (MTP) forming primary chylomicron (CM). These lipids are transported in vesicles to the Golgi system by the action of proteins of the COPII complex. At the Golgi system, the CMs incorporate apoAI and apoAII to form nascent CM. These are exported in vesicles by reverse pinocytosis at the lymphatic system. In the circulation, the nascent CM acquires apolipoproteins from high-density lipoprotein (HDL) to form mature CM. The CM in the circulation can undergo the action of lipoprotein lipase (LPL) located in the peripheral tissues. Proteins present in lipoproteins (apoCII, ApoCIII, and ApoAV) and tissues (glycosylphosphatidylinositol anchored high-density lipoprotein binding protein 1 [GPIHBP1], lipase maturation factor 1 [LMF1], and angiopoietin-like [ANGPTL3]) regulate lipoprotein lipase (LPL) activity. Cleavage of TG releases FA into the cells and diacylglycerol into the medium. Owing to the action of the LPL, the quantity of TG is reduced in the CMs. apoA and apoC are transferred to HDL. The remainder is then identified as CM remnant (CMrem), which contains apoB48 and apoE besides apoAV. These apolipoproteins are recognized by hepatic receptors such as low-density lipoprotein receptor (LDLR) and LDL receptor-related protein 1 (LRP1), which are internalized.

present in the liver or intestine. In patients with NAFLD, the promoter region of the *APOB* gene was associated with changes in DNA methylation and correlated with transcription activity.

Mutations in apoB are associated with dyslipidemia. In familial hypobetalipoproteinemia (FHBL1) (OMIM 615558), the mutations generate a protein that is unable to form lipoproteins. There is an absence of apoB in plasma, which results in hypochylomicronemia in addition to a reduction of other apoB100-containing lipoproteins. The inheritance is autosomal codominant. The heterozygotes have half the normal concentration of apoB and an accumulation of lipids in the enterocytes and hepatocytes owing to a reduction in the TG export.

In more severe cases, plasma LDL cholesterol levels are below 20 mg/dL. Mutation with a deletion of two amino acids (Thr26_Tyr27del) at the amino terminal portion of the apoB48 is associated with a deficiency in apoB48 secretion. In contrast to the alteration of apoB synthesis, type B hypercholesterolemia (OMIM 144010) results in a mutation with dominant inheritance that impairs the association of apoB to LDL receptor (LDLR) (Fig. 60.3). The lower affinity for the LDLR contributes to the accumulation of apoB100 in the circulation and increases plasma cholesterol levels.

Microsomal triglyceride transfer protein (MTP) has a key role in incorporating neutral lipids in apoB (apoB lipidation) (Figs. 60.2 and 60.3). When this protein

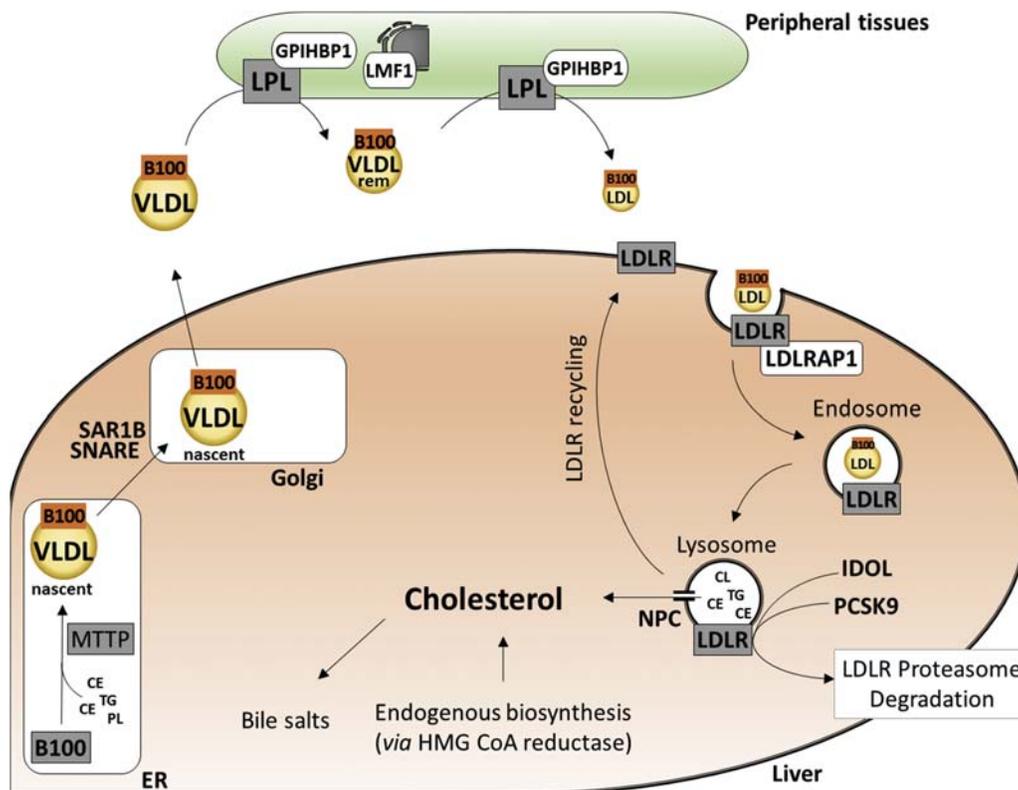


FIGURE 60.3 Synthesis and metabolism of very low-density lipoprotein (VLDL). It is performed in the liver and involves the same proteins as chylomicron (CM) synthesis (see Fig. 60.2). However, the apolipoprotein is apoB100. In the peripheral tissues, lipoprotein lipase (LPL) hydrolyzes triglyceride (TG) of VLDL, converting it into very low-density lipoprotein (LDL) remnant (VLDLrem). Successive actions of the LPL convert VLDLrem into LDL. This lipoprotein is recognized by LDL receptor (LDLR) receptors, which along with LDLR adaptor protein 1 (LDLRAP1), internalize the lipoprotein forming the endosome. After hydrolysis in the lysosome, the cholesterol is exported to the cytoplasm by means of the Niemann-Pick C (NPC) carrier. When there is an increase in the expression of inducible LDLR degrading factor (IDOL) and pro-protein convertase subtilisin/kexin type 9 (PCSK)9 proteins, the LDLR receptor is directed for degradation instead of recycling in the plasma membrane.

is inactive, the intestine and liver do not release lipoproteins. Similar to FHBL1, there is an absence of apoB in the circulation. Mutations in the *MTTP* gene cause abetalipoproteinemia (OMIM 200100). The heterozygotes have normal values of lipids because of autosomal recessive inheritance. Enterocytes of diabetic patients have a lower expression of *MTTP*. In epigenetic studies of the gene *MTTP*, the importance of the expression of this gene in the lipid metabolism was proved. Rodents fed a high-fat diet have hypermethylation of the promoter region of *MtTp*. An inverse correlation was reported in relation to gene expression. In humans, the expression of *MTTP* is not correlated with the methylation of the promoter region.

apoBs containing lipids need to be transported from the endoplasmic reticulum to the Golgi complex. This transport is mediated by the coat protein (COPII) complex containing the subunit secretion-associated **Ras-related GTPase 1B**, which directs the formation of vesicle transports (Fig. 60.2). Mutations in this gene cause CM retention disease (CMRD) (OMIM 246700),

also known as Anderson disease, which is a hypocholesterolemia with autosomal recessive inheritance. Patients do not have CM in plasma and accumulate lipids in the enterocytes. Because the protein SAR1B also acts in the secretion of hepatic VLDL (Fig. 60.3), patients with CMRD have steatosis and low production of apoB100 and apoA1.

The biogenesis of HDL depends on **apoAI** (Fig. 60.4). This gene is in an *APOA1C3A4A5* cluster located on chromosome 11. The liver and intestine synthesize and secrete apoAI as a lipid-poor protein form that interacts with lipids through the ABCA1 receptor to form the nascent HDL (Fig. 60.4). The lipid-poor form may be metabolized in the kidneys. ApoAI has a central region that is important for activating the enzyme lecithin:cholesterol acyltransferase (LCAT) to perform cholesterol esterification. Deficiency in the synthesis of apoAI leads to a reduction in the HDL concentration in plasma. A mutation in the *APOA1* gene results in primary hypoalphalipoproteinemia (OMIM 604091). The changed apoAI has no ability to capture cholesterol or activate

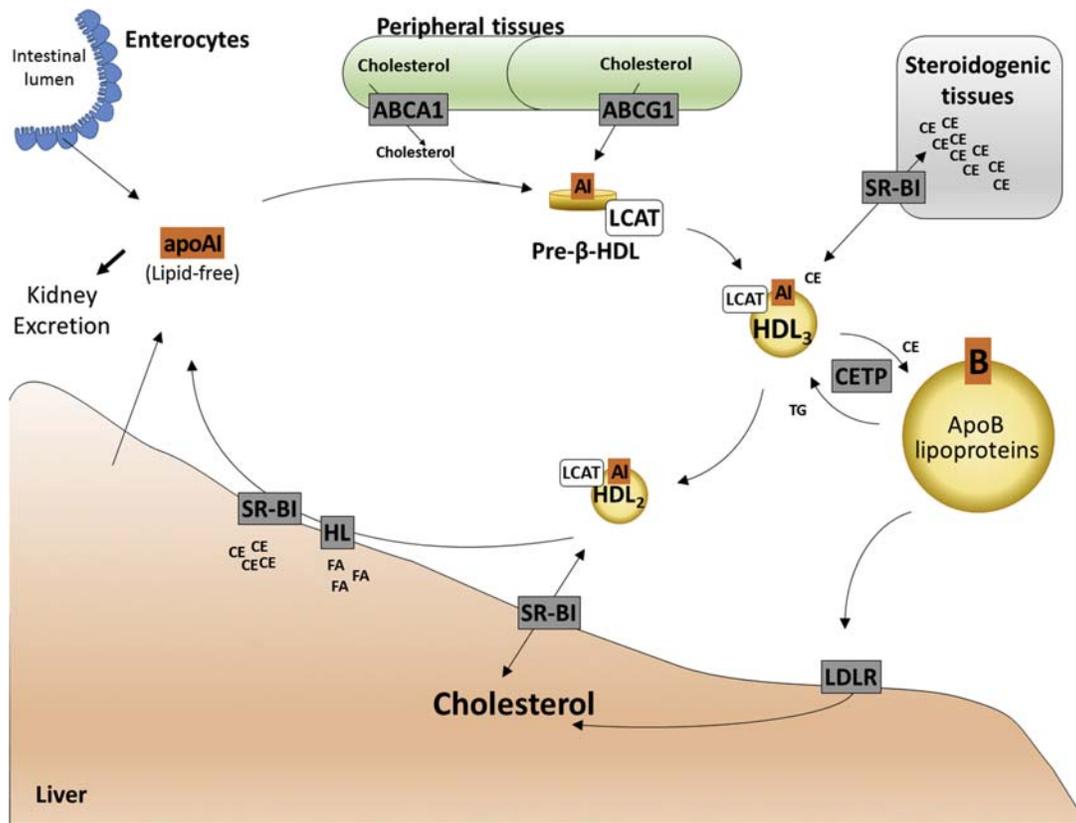


FIGURE 60.4 Synthesis and metabolism of high-density lipoprotein (HDL). Apolipoprotein (apo)AI is synthesized in the liver and enterocytes in a lipid-free manner. This apoAI incorporates lipids through the action of the adenosine triphosphate binding cassette (ABC)A1 transporter that transfers cholesterol from peripheral tissues and forms the discoid lipoprotein pre- β -high-density lipoprotein (HDL). The cholesterol is esterified by the action of the enzyme lecithin: cholesterol acyltransferase (LCAT) and transformation into HDL₃ with an oval shape. The cholesterol ester (CE) is distributed to apoB-containing lipoproteins or to the tissues. The enzyme cholesteryl ester transfer protein (CETP) mediates the transfer of CE for apoB-containing lipoproteins in an exchange of triglycerides (TG) to HDL. This change reduces the size of the lipoprotein and forms HDL₂. The CE distribution to the tissues is mediated by scavenger receptor class B member 1 (SR-BI) receptors located in steroidogenic or hepatic tissues. The transfer is bidirectional and depends on the concentration of CE in the tissue and in the HDL. The HL performs hydrolysis of TG of HDL and contributed to releasing lipid-free apoAI, which is excreted in the kidneys. The cholesterol transfer process from the tissues to the liver is known as reverse cholesterol transport and it is mediated by HDL via SR-BI receptors and by low-density lipoprotein (LDL) via low-density lipoprotein receptors (LDLR).

the LCAT. The demethylation of *APOA1* in intestinal cells by exogenous demethylating agents results in increased gene expression. Long noncoding RNAs also act as negative regulators of *APOA1* transcription.

Genes Involved in Lipoprotein Metabolism

Lipoprotein lipase (LPL) hydrolyses TG either from the diet or endogenously (Figs. 60.2 and 60.3). Mutations in LPL may abolish enzyme synthesis, generate unstable transcripts, prevent the cells from secreting the enzyme, interfere with the stability of the homodimer, lose the ability to bind to glycosylphosphatidylinositol anchored high-density lipoprotein binding protein 1 (GPIHBP1), or raise cleavage susceptibility. LPL mutations cause type I hyperlipoproteinemia (OMIM 238600). As LPL catalyzes hydrolysis and TG uptake in peripheral tissues (Figs. 60.2 and 60.3), the absence of functional LPL leads

to an accumulation of particles rich in TG as CM and VLDL. Fasting plasma TG levels are above 1000 mg/dL. In healthy subjects, the expression of LPL depends on the tissue metabolic demand. Hormones such as insulin, glucocorticoids, and adrenaline are the main regulators of LPL expression. LPL activity is increased in heart muscle and reduced in adipose tissue under fasting conditions. Proteins located in peripheral tissues such as **lipase maturation factor 1 (LMF1)**, GPIHBP1, and angiopoietin-like (ANGPTL) and apolipoproteins present in lipoproteins, such as apoCIII, apoCII, and apoAV, also regulate LPL activity.

LMF1 is a chaperone required for the folding and expression of LPL on the surface of endothelial cells (Figs. 60.2 and 60.3). Changes in cell function are associated with adjustments to the catabolism of triglyceride-rich lipoproteins owing to the reduction or increase in LPL activity. Patients with combined lipase deficiency

(OMIM 246650), a disease with autosomal recessive inheritance, have severe hypertriglyceridemia caused by the absence of LMF1. The protein **GPIHBP1** directs the transendothelial transport of LPL and anchors the CMs and VLDLs close to the endothelial surface (Figs. 60.2 and 60.3). The absence of GPIHBP1, which reduces the activity of LPL, increases plasma CM levels; it is characterized as type ID hyperlipoproteinemia (OMIM 615947). **ANGPTL** are proteins that inhibit LPL activity (Fig. 60.2). In familial hypobetalipoproteinemia 2 disease (FHBL2) (OMIM 605019), patients have low blood levels of apoB. Different from FHBL that occurs as the result of a mutation in the *APOB* gene, changing the affinity for LPL, the absence of activity of the inhibitor **ANGPTL3** causes an increase in LPL activity. Low concentrations of apoB, total cholesterol, LDL cholesterol, and HDL cholesterol are reported. **ANGPTL4** also inhibits LPL activity. Hypermethylation of **ANGPTL3** is reported in patients with NAFLD.

In the bloodstream, the lipoproteins rich in TG receive apoCII, apoCIII, and apoAV from HDL. These apolipoproteins regulate LPL activity (Fig. 60.2). ApoCII is a cofactor activator of LPL. It is present in HDL, CM, and VLDL. The hepatic expression of *APOC2* is upregulated by a high-fat diet and by agonists of peroxisome proliferator-activated receptor α (PPAR- α). Apolipoprotein CII deficiency (OMIM 207750) is associated with increased plasma lipoprotein levels. **ApoCIII** is also present in HDL, CM, and VLDL; it inhibits LPL activity. The mutation with loss of function causes apoCIII deficiency (OMIM 614028). Low levels of the inhibitor apoCIII increase the containing apoB lipoprotein catabolism. Carriers of the mutation have low plasma concentrations of TG and reduction in the risk for cardiovascular diseases. Patients with hypertriglyceridemia exhibit increased apoCIII levels. Both *APOC2* and *APOC3* are hypermethylated in the liver tissue of the patients with NAFLD. However, the effect of these epigenetic changes on the expression of these genes is not yet clear. **ApoAV** stabilizes the enzyme–lipoprotein complex and increases lipolysis. In type V hyperlipoproteinemia (OMIM 144650), an absence or deficiency in apoAV is reported. Under these conditions, decreased hydrolysis of TG by LPL and, consequently, accumulation of particles rich in TG are described. Transcription factor PPAR- α causes upregulation of the *APOA5* gene. Genetic polymorphisms as well as epigenetic modifications at the *APOA5* gene are associated with increased plasma TG levels.

ABCA1 is a carrier that promotes the transfer of cholesterol and PL for apoAI (Fig. 60.4). Tangier disease (OMIM 205400) is an autosomal recessive deficiency of ABCA1. In this condition, a reduction in plasma HDL levels caused by the inability to form pre- β -HDL has been reported. Patients have impaired cholesterol

exported from peripheral tissues and an accumulation of cholesteryl esters in adipose tissue. Another carrier, the **ABCG1** is also involved in the transport of cholesterol and PL to HDL (Fig. 60.4). Although *ABCG1* does not have a primary association with dyslipidemia, prospective studies in Europeans and Indians have identified that methylation of the *ABCG1* gene in the peripheral blood is associated with an increased risk for developing type 2 diabetes. Furthermore, African American adults with metabolic syndrome have increased *ABCG1* DNA methylation. In macrophages, PPAR- α and PPAR γ induce the expression of *ABCA1* and promote an accumulation of cholesterol. A third receptor that recognizes HDL is the **scavenger receptor class B member 1 (SR-BI)**, also known as **SCARB1** (Fig. 60.4). Expressed predominantly in the liver, adrenal glands, and ovaries, SR-BI transports cholesteryl esters derived from HDL to the liver and steroidogenic tissues. Mutations in SR-BI cause HDL cholesterol level quantitative trait locus 6 (OMIM 610762). Carriers exhibit an increase in HDL cholesterol levels.

ApoAI is the main activator of **LCAT**, which is found in HDL and LDL. Absence of the activity of LCAT in HDL and LDL causes LCAT deficiency (OMIM 245900). With autosomal recessive inheritance, patients have a reduction in plasma HDL cholesterol levels. The quantity of apoAI is low owing to the fast renal catabolism of poor lipid apoAI. The absence of activity only in HDL causes fish-eye disease (OMIM 136120), which also has low plasma HDL cholesterol levels. The enzyme **cholesteryl ester transfer protein (CETP)** transfers cholesteryl esters between containing apoB lipoproteins and HDL (Fig. 60.4). Mutations with loss of function cause hyperalphalipoproteinemia 1 (OMIM 143470). An increase in HDL resulting from the absence of transfer of cholesteryl esters to the containing apoB lipoproteins, is described. **Hepatic lipase (HL)** catalyzes TG hydrolysis into FAs from several lipoproteins. HL is important in the metabolism of HDL because it converts HDL rich in TG into HDL poor in TG (Fig. 60.4). Mutation in this gene causes hepatic lipase deficiency (OMIM 614025). Carriers have both HDL and LDL with a high quantity of TG.

Genes Involved in Tissue Clearance of Lipoproteins

The clearance of plasma lipoproteins depends on proteins related to lipoprotein activation and degradation. The increase in plasma cholesterol levels may result from a deficit in LDL uptake and processing. Mutations can affect **LDLR** activity, preventing internalization, or even affect receptor transcription. Some patients with increased plasma cholesterol levels may have the

polygenetic form. Several genes are affected by mutations that alter LDLR function. The control of LDLR expression depends on the hepatic cholesterol concentration. When the intracellular concentration of cholesterol rises, there is activation of the liver X receptors, a transcription factor that induces the expression of the inducible LDLR degrading factor (IDOL). IDOL is a posttranslational regulator of LDLR by adding ubiquitin to the LDLR and directing it for degradation in the lysosome (Fig. 60.3). When IDOL is enabled, it reduces the uptake of exogenous cholesterol by LDLR. Familial hypercholesterolemia disease (OMIM 143890) is caused by a mutation in the *LDLR* gene with autosomal dominant inheritance. An increase in plasma LDL cholesterol levels is reported with this condition.

Various proteins are involved in LDL uptake by the liver (Fig. 60.3). **LDLR adaptor protein 1 (LDLRAP1)** is required to internalize LDLR (Fig. 60.3). A null recessive mutation in the *LDLRAP1* causes autosomal recessive hypercholesterolemia (OMIM 603813). Because the endocytosis of the receptor LDLR is impaired, an increase in plasma cholesterol levels occurs. Another protein expressed in the liver, intestines, and kidneys and also found in plasma is **proprotein convertase subtilisin/kexin type 9 (PCSK9)**. This protein affects recycling of the LDLR. When there is an increase in intracellular cholesterol levels, PCSK9 promotes degradation of the LDLR. The transcription is controlled by transcription factors such as SREBPs, hepatocyte nuclear factor 1, and PPAR- γ . Mutations with a gain of function in PCSK9 lead to an increase of plasma cholesterol levels caused by the reduction in the number of receptors on the plasma membrane surface. Patients bearing this gene mutation exhibit autosomal dominant hypercholesterolemia 3 disease (OMIM 603776). In contrast, the loss of function in PCSK9 contributes to increase the quantity of LDLR receptors and thus reduce plasma

LDL levels. Pharmacological inhibitors of PCSK9 are being tested for the treatment of hypercholesterolemia. **Niemann-Pick C1 (NPC1)** is a membrane protein of the lysosome (Fig. 60.3). NPC1 promotes the export of cholesterol derived from endocytosis mediated by the LDLR. Mutations in this gene cause Niemann-Pick disease type C1 (OMIM 257220). Under this genetic condition, accumulation of cholesteryl esters and other lipids in the liver and cells of the central nervous system is reported. Uncontrolled lipid accumulation causes cell death and irreversible damage to the tissues.

Another receptor involved in the clearance of lipoproteins is the VLDL receptor (VLDLR). VLDLR is expressed in adipose tissue, skeletal muscle, endothelial cells of capillaries, and arterioles, but it is not expressed in hepatic tissue. In liver, the LDLR is widely distributed. Tissue VLDLR provides fuel substrates for cells. Although DNA hypermethylation suppresses VLDLR gene expression in gastric cancer cell lines, there are no studies in dyslipidemic patients. ApoE has one region that recognizes LDLR and interacts with LDL receptor-related protein 1 (LRP1) and heparan sulfate proteoglycan 2. LDLR is involved in the clearance of CMrem. In type III hyperlipoproteinemia, there is an increase in apoE triglyceride-rich lipoproteins levels. As a consequence, an increase in plasma TG and plasma total cholesterol levels resulting from CMrem and VLDLrem (or IDL) accumulation is found.

ROLE OF NUTRITIONAL GENOMIC IN DYSLIPIDEMIA

Table 60.1 summarizes dyslipidemias associated with function modulation through changes in gene expression. As observed, most dyslipidemias are associated with specific gene mutations (Table 60.1). However,

TABLE 60.1 Dyslipidemias Associated with Changes in Gene Expression Owing to Gene Mutation.

Disease	Gene	Manifestation	Comments
LIPID ABSORPTION			
Pancreatic lipase deficiency (OMIM 614338)	<i>PNLIP</i>	Oily/greasy stools	Lipids partial absorption owing to other intestinal lipases
Sitosterolemia (OMIM 210250)	<i>ABCG5/G8</i>	Plants steroids accumulation	No cholesterol efflux in intestine Genes are downregulated in type II diabetes
LIPOPROTEIN SYNTHESIS			
Familial hypobetalipoproteinemia (OMIM 615558)	<i>APOB</i>	Absence of plasma apolipoprotein B (apoB) Hypocholesterolemia Low-density lipoprotein (LDL) < 20 mg/dL	apoB is unable to form lipoproteins Malabsorption of lipid-soluble vitamins

Continued

TABLE 60.1 Dyslipidemias Associated with Changes in Gene Expression Owing to Gene Mutation.—cont'd

Disease	Gene	Manifestation	Comments
Abetalipoproteinemia (OMIM 200100)	<i>MTTP</i>	Absence of plasma apoB	No neutral lipid incorporation in apoB
Chylomicron retention disease (OMIM 246700) or Anderson disease	<i>SAR1B</i>	Absence of plasma chylomicron (CM)	No formation of vesicles transport in Golgi complex Enterocytes lipid accumulation Steatosis Low production: apoB100 and apoA1
Primary hypoalphalipoproteinemia (OMIM 604091)	<i>APOAI</i>	Reduction in cellular cholesterol efflux	No apoAI function (capture cholesterol or activate lecithin: cholesterol acyltransferase)
LIPOPROTEIN METABOLISM			
Type I hyperlipoproteinemia (OMIM 238600)	<i>LPL</i>	↑ Triglycerides (TG) (fasting > 1000 mg/dL)	Several mutations Lipoprotein lipase (LPL) function absent or altered accumulates CM and very low-density lipoprotein (VLDL) LPL expression and activity are controlled by tissue metabolic demand
Combined lipase deficiency (OMIM 246650)	<i>LMF1</i>	↑ TG	Absence of lipase maturation factor 1 (LMF1) LMF1 is responsible for folding and expression of LPL on surface of endothelial cells
Type ID hyperlipoproteinemia, (OMIM 615947)	<i>GPIHBP1</i>	↑ TG (↑ CM)	GPIHBP1 absence reduces LPL activity.
Familial hypobetalipoproteinemia 2 disease (FHBL2, OMIM 605019)	<i>ANGPTL3</i>	Low plasma apoB ↓ Total cholesterol ↓ LDL ↓ High-density lipoprotein (HDL)	Absence increases LPL activity (angiopoietin-like3 is inhibitor of LPL)
Apolipoprotein CII deficiency (OMIM 207750)	<i>APOCII</i>	↑ Plasma lipoprotein	High fat diet and peroxisome proliferator-activated receptor α agonists upregulate gene expression
Apolipoprotein CIII deficiency (OMIM 614028)	<i>APOCIII</i>	↓ TG	Loss of function increases LPL activity (apoCIII is an inhibitor of LPL)
Type V hyperlipoproteinemia (OMIM 144650)	<i>APOAV</i>	↑ TG	Absence decreases LPL activity (apoAV stabilizes enzyme–lipoprotein complex).
Tangier disease (OMIM 205400)	<i>ABCA1</i>	↓ HDL	Deficiency causes inability to form pre- β -HDL. Cholesteryl ester accumulation in no adipose tissue
High-density lipoprotein cholesterol level quantitative trait locus 6 (OMIM 610762).	<i>SR-BI</i>	↑ HDL	Mutation
Lecithin: cholesterol acyltransferase 'deficiency' (OMIM 245900)	<i>LCAT</i>	↓ HDL	Absence of activity in HDL and LDL Fast renal catabolism of poor lipid apoAI.
Fish eye disease (OMIM 136120)	<i>LCAT</i>	↓ HDL	Absence of activity only in HDL
Hyperalphalipoproteinemia 1 (OMIM 143470)	<i>CETP</i>	↑ HDL	Loss of function with no transfer cholesteryl ester to apoB lipoproteins
Hepatic lipase deficiency (OMIM 614025)	<i>HL</i>	↑ TG in HDL and LDL	
LIPOPROTEIN CLEARANCE			
Familial hypercholesterolemia disease (OMIM 143890)	<i>LPLR</i>	↑ Cholesterol levels ↑ LDL	Prevents internalization Affects receptor transcription LDL receptor (LDLR) expression depends on hepatic cholesterol levels

TABLE 60.1 Dyslipidemias Associated with Changes in Gene Expression Owing to Gene Mutation.—cont'd

Disease	Gene	Manifestation	Comments
Autosomal recessive hypercholesterolemia (OMIM 603813)	<i>LDLRAP1</i>	↑ Cholesterol levels	Endocytosis of LDLR is impaired
Autosomal dominant hypercholesterolemia 3 disease (OMIM 603776)	<i>PCSK9</i>	↑ Cholesterol levels	Gain of function reduces number of receptors on plasma membrane (PCSK9 promotes LDLR degradation)
		↓ Cholesterol levels	Loss of function increases LDLRs
Niemann-Pick disease type C1 (OMIM 257220)	<i>NPC1</i>	Accumulation of cholesteryl esters in the liver and CNS	Export cholesterol derived from endocytosis mediated by LDLR
Type B hypercholesterolemia (OMIM 144010)	<i>APOB</i>	apoB100 accumulation ↑ Cholesterol	ApoB has low affinity for LDLR

OMIM, Online Mendelian Inheritance in Man (<https://www.omim.org/>).

several dyslipidemias are not explained by single mutations. These dyslipidemias are probably associated with epigenetic mechanisms that may be influenced by environmental factors such as diet, lifestyle, contaminants, and drugs. To date, only few genes have been investigated concerning epigenetic mechanisms, as summarized in Table 60.2.

The lack of studies on epigenetics of dyslipidemia is illustrated in Fig. 60.5. A database search identified only 107 publications on dyslipidemias and epigenetic mechanisms (search terms dyslipidemia* and epigen*) in the Web of Science Core Collection for all publications (<https://www.webofknowledge.com>). Nutritional

longitudinal studies can increase the knowledge because epigenetic marks are prone to alterations during life. New gene and protein expressions and activities analysis and systems biology approaches such as genomics, transcriptomics, methylomics, proteomics, metabolomics, and phenomes should be measured concomitantly to understand polygenic dyslipidemia. Other applications of the epigenetics are for identifying biomarkers. Global DNA methylation is not a good biomarker for predicting plasma lipoproteins levels. Specific individual gene methylations such as *ABCG1*, *CPT1A*, *TNNT1*, *MIR33B*, *SREBF1*, and *TNIP* have shown lipid plasma level associations.

TABLE 60.2 Genes Related to Dyslipidemias That Have Epigenetic Control.

Lipid metabolism function	Gene	Epigenetic alteration
Lipid absorption	<i>NPC1L1</i>	<i>NPC1L1</i> promoter hypermethylation is inversely correlated with transcription.
Lipoprotein synthesis	<i>APOB</i>	<i>APOB</i> DNA methylation changes is related to transcription in NAFLD.
	<i>MTTP</i>	<i>MTTP</i> expression is not correlated to promoter DNA methylation.
	<i>APOA1</i>	<i>APOA1</i> demethylation increases gene expression.
Lipoprotein metabolism	<i>ANGPTL3</i>	Nonalcoholic fatty liver disease (NAFLD) patients have promoter hypermethylation.
	<i>APOCII</i>	Liver <i>APOC2</i> is hypermethylated in NAFLD.
	<i>APOCIII</i>	Liver <i>APOC3</i> is hypermethylated in NAFLD.
	<i>APOAV</i>	Epigenetic modifications at the <i>APOA5</i> gene are associated with increased plasma TG.
	<i>ABCG1</i>	Methylation of the <i>ABCG1</i> gene in peripheral blood is associated with increased risk for developing type 2 diabetes.
Lipoprotein clearance	<i>VLDLR</i>	DNA hypermethylation suppress gene expression in gastric cancer cell lines. No studies in patients with dyslipidemia patients.

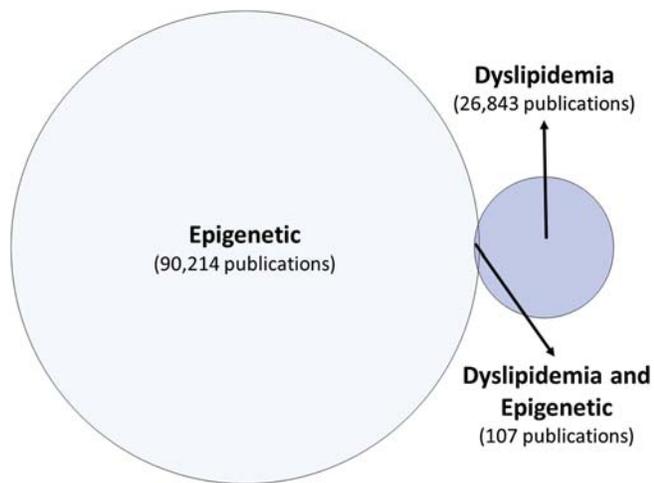


FIGURE 60.5 Studies that investigate the association of epigenetic with dyslipidemias. A database search of the Web of Science identified only 107 publications with the topic dyslipidemia and epigenetic (search terms dyslipidemia* and epigen*) (<https://apps.webofknowledge.com/>).

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

The absorption, synthesis, metabolism, and clearance of lipids are finely regulated by the activities and expressions of specific genes. Genetic alterations associated with primary dyslipidemia allow us to understand the implications of specific genes in interactions of

lipoproteins with other lipoproteins and tissues. Several secondary dyslipidemias that may have polygenic causes remain to be described. Epigenetic mechanisms are involved in controlling the expression of these genes; as a result, new therapeutic approaches for dyslipidemias are proposed. Genomic and epigenomic studies need to be performed and analyzed together to define the mechanisms involved in secondary and polygenic dyslipidemias. Clarification of genetic and epigenetic mechanisms involved in the development of dyslipidemias is required to establish the impact of nutrients on the expression of the genes as well as their application in preventing and controlling dyslipidemias.

Suggested Reading

- Braun, K.V., Voortman, T., Dhana, K., Troup, J., Bramer, W.M., Troup, J., Chowdhury, R., Dehghan, A., Muka, T., Franco, O.H., 2016. The role of DNA methylation in dyslipidaemia: a systematic review. *Prog Lipid Res* 64, 178–191.
- Mwinyi, J., Boström, A.E., Pisanu, C., Murphy, S.K., Erhart, W., Schafmayer, C., Hampe, J., Moylan, C., Schiöth, H.B., 2017. NAFLD is associated with methylation shifts with relevance for the expression of genes involved in lipoprotein particle composition. *Biochim Biophys Acta* 1862, 314–323.
- Rodríguez, A., Pajukanta, P., 2018. Genomics and systems biology approaches in the study of lipid disorders. *Rev Invest Clin* 70, 217–223.
- Trinh, M.N., Lu, F., Li, X., Das, A., Liang, Q., De Brabander, J.K., Brown, M.S., Goldstein, J.L., 2017. Triazoles inhibit cholesterol export from lysosomes by binding to NPC1. *Proc Natl Acad Sci U S A* 114, 89–94.

Nutrients and Gene Expression in Inflammation

Charles N. Serhan

Center for Experimental Therapeutics and Reperfusion Injury, Hale Building for Transformative Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States

Glossary Terms and Keywords

LC-MS-MS liquid chromatography tandem mass spectrometry

LM lipid mediator

LOX lipoxygenase

MaR1 maresin 1 (7*R*,14*S*-dihydroxy-docosa-4*Z*,8*E*,10*E*,12*Z*,16*Z*,19*Z*-hexaenoic acid)

PD1 protectin D1 (10*R*,17*S*-dihydroxy-docosa-4*Z*,7*Z*,11*E*,13*E*,15*Z*,19*Z*-hexaenoic acid)

RvD1 Resolvin D1 (7*S*,8*R*,17*S*-trihydroxy-docosa-4*Z*,9*E*,11*E*,13*Z*,15*E*,19*Z*-hexaenoic acid)

RvD2 Resolvin D2 (7*S*,16*R*,17*S*-trihydroxy-docosa-4*Z*,8*E*,10*Z*,12*E*,14*E*,19*Z*-hexaenoic acid)

RvD3 Resolvin D3 (4*S*,11*R*,17*S*-trihydroxy-docosa-5*Z*,7*E*,9*E*,13*Z*,15*E*,19*Z*-hexaenoic acid)

RvD4 Resolvin D4 (4*S*,5*R*,17*S*-trihydroxydocosa-6*E*,8*E*,10*Z*,13*Z*,15*E*,19*Z*-hexaenoic acid)

SPM specialized pro-resolving mediator

INFLAMMATORY RESPONSES: THE SEQUENCE, INITIATION, AND RESOLUTION

Acute inflammation is a defensive physiological response occurring in vascularized tissues to protect the host against microbial invaders and injuries. It is divided into two phases: initiation and resolution (or termination). This ally manifests itself in the early phase after a microbial infection, to fight against invading microbes before the adaptive or antibody-dependent response is engaged. The cardinal signs of inflammation, described by the Roman physician Celsus in the first century as *rubor* (redness), *tumor* (swelling), *calor* (heat), and *dolor* (pain), are the visible manifestation of changes occurring at the molecular and cellular levels within inflamed tissues. Edema is one of the earliest events in an inflammatory response that arises from increased vascular permeability

(Fig. 61.1). Then, leukocytes are recruited to sites of inflammation via crossing postcapillary venules. Polymorphonuclear neutrophils (PMN) are among the first leukocyte responders that accumulate in inflamed sites. These cells kill pathogens by engulfing them (*phagocytosis*), releasing microbicidal agents stored in their granules and producing reactive oxygen species in phagolysosomal vacuoles to kill invaders. Next, mononuclear cells enter from blood that can differentiate into macrophages (MΦs) (Fig. 61.2). Together, these cells clear microbes, cellular debris, and apoptotic PMN via phagocytosis in a non-phlogistic process termed *efferocytosis*, which is pivotal to the resolution of inflammation.

Clearance and efflux of phagocytes (*egress*) allows for resolution, permitting the return to homeostasis. To maintain a healthy status, both the initiation of acute inflammation and its resolution must be efficient. Uncontrolled or unresolved inflammation is recognized as a major driver of human pathologies, including arthritis, asthma, neurodegenerative diseases, cancers, and cardiovascular diseases. Also, in humans and other species, aging and diet are tightly linked to the magnitude of inflammation (Fig. 61.1). Because of the high occurrence of these diseases, understanding how acute inflammation resolves is of wide interest, because the language of cell movements or cellular traffic is governed by chemical mediators that serve as signals in these vital processes.

This chapter focuses on the newly elucidated specialized pro-resolving mediators (SPMs) biosynthesized from essential polyunsaturated fatty acid (PUFA) precursors (arachidonic acid [AA], eicosapentaenoic acid [EPA], and docosahexaenoic acid [DHA]) as well as the chemical mediator families lipoxins (LX), resolvins (Rv), protectins (PD), and maresins (MaR) (Fig. 61.2),

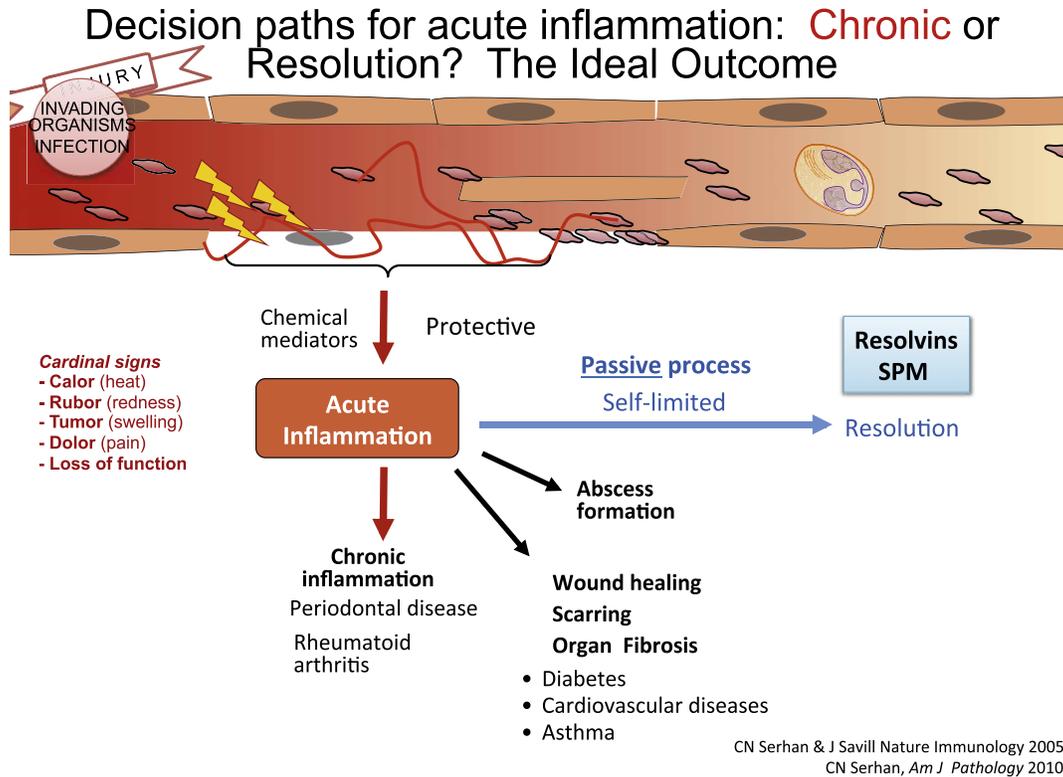


FIGURE 61.1 Decision paths for acute inflammation: chronic or resolution? The ideal outcome. SPM, specialized pro-resolving mediator.

and reviews SPM biosynthetic pathways, receptors, and microRNAs (miRNAs) that control self-limited inflammation to promote timely resolution.

RESOLUTION IS AN ACTIVE PROCESS: CHEMICAL MEDIATORS ARE THE LANGUAGE

At the level of tissue histology, resolution was described by pathologists more than 100 years ago as the time when neutrophils that infiltrated inflamed tissue sites leave or are lost from the site. Resolution was thought to be a passive process, simply resulting from attenuation or dissipation of chemotactic and proinflammatory initiating signals. Results from the author's laboratory demonstrated that resolution is instead an *active process* orchestrated by special novel chemical mediators that *turn on* biosynthetic and cellular pathways to enable return to homeostasis. Lipid mediators (LM) from PUFAs have essential roles in each phase of acute inflammation, with prostaglandins (PGs) and cysteinyl leukotrienes (cysLTs) promoting an early increase in vascular permeability, and leukotriene (LT) B₄, a potent and key leukocyte chemoattractant that summons PMN to the inflammatory site (Fig. 61.2). Prostaglandins

also contribute to fever and pain. Chronic inflammation was widely viewed as an excess of proinflammatory mediators (Fig. 61.1). The resolution phase is characterized by *active biosynthesis* of specific LMs that operate as *resolution agonists* to (1) keep inflammation within physiological boundaries and (2) expedite return to homeostasis (Fig. 61.1). Identification of this new array of SPM and their relation to proinflammatory LM was achieved using self-limited or naturally resolving acute inflammation models in vivo and a systems approach. This new array of chemical signals is the language recognized as a genus of SPMs that have two broad functions: (1) antiinflammation and (2) pro-resolving, by stimulating multilevel clearance actions. SPM exert potent antiinflammatory, pro-resolving, and regenerative activities in the low nanogram to microgram concentration range in experimental disease models (infections and inflammation). Accumulating evidence indicates that failure or disruption of the body's pro-resolution pathways can be detrimental and underlie some mechanisms of chronic inflammatory diseases.

Additional chemical mediators are operative in inflamed tissues to switch off leukocyte infiltration and restore physiological functions. Among these are several cytokines (e.g., transforming growth factor β and interleukin (IL)-10), which accumulate in resolving exudates;

Ideal Outcome of Inflammation : Complete Resolution

Systems Approach to Mapping Resolution

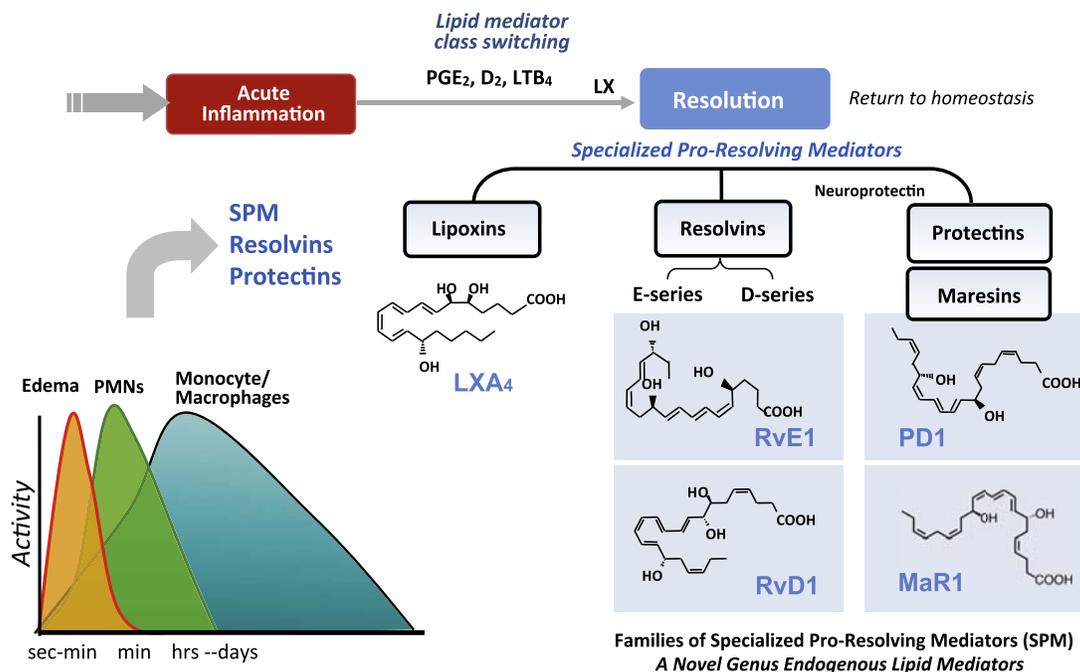


FIGURE 61.2 **Ideal outcome of inflammation: complete resolution.** Systems approach to mapping resolution. LTB_4 , leukotriene B_4 ; LX , lipoxin; $MaR1$, maresin 1; $PD1$, protectin D1; PGE_2 , prostaglandin E_2 ; $PMNs$, polymorphonuclear neutrophils; $RvE1$, $RvD1$, E-series and D-series resolvins, D4; SPM , specialized pro-resolving mediator.

glucocorticoids and the glucocorticoid-induced annexin-1 protein, which tune the inflammatory response; and the transcription factor nuclear factor (NF)- κB , which also has some antiinflammatory properties. Inducing PMN apoptosis as well as lymphoid cells while stimulating their prompt removal by M Φ s also can promote resolution. In addition to SPMs themselves, small inhibitors of cyclin-dependent kinases fulfill this goal, as do annexin 1 peptides, which demonstrates that the resolution process can be pharmacologically targeted as a new means to control excessive inflammation.

Importantly, resolution is not synonymous with antiinflammation. This is because, to be considered pro-resolving, in addition to serving as a stop signal for neutrophil trafficking and other cardinal signs of inflammation (e.g., swelling, pain) (Figs. 61.1 and 61.2), a mediator must also stimulate PMN efferocytosis by M Φ , favor antibacterial activities, and promote tissue repair and regeneration. Active resolution includes gene expression regulation of several soluble chemical mediators (e.g., cytokines, chemokines), receptors (e.g., Toll-like receptors), and transcription factors. Many genes are under tight control of miRNAs (short noncoding RNA molecules) that act as translational repressors of messenger RNA (mRNA) transcripts. They are involved

in many physiological and pathological processes, including cell development, cancer, and inflammation. Specific miRNAs in self-limited inflammation and in the resolution phase are regulated by $RvD1$ -G-protein coupled receptor (GPCR)-dependent gene networks in the resolution of acute inflammation as part of the circuitry that controls active resolution (see subsequent discussion). Interested readers are directed to the suggested reading list for further detail.

STRUCTURAL ELUCIDATION OF RESOLUTION PHASE MEDIATORS: RESOLVINS

An unbiased systems approach was taken to identify endogenous SPMs and decode their mechanisms of action. For this, mouse dorsal air pouches and self-limited acute inflammation proved ideal because they permitted isolation of contained inflammatory exudates (namely, *pus*) and enabled direct untargeted and targeted LM-lipidomics of bioactive products, as well as their precursors and further metabolites, proteomics, miRNAs, and analyses of cellular composition of the resolving exudate: hence, the natural means by which

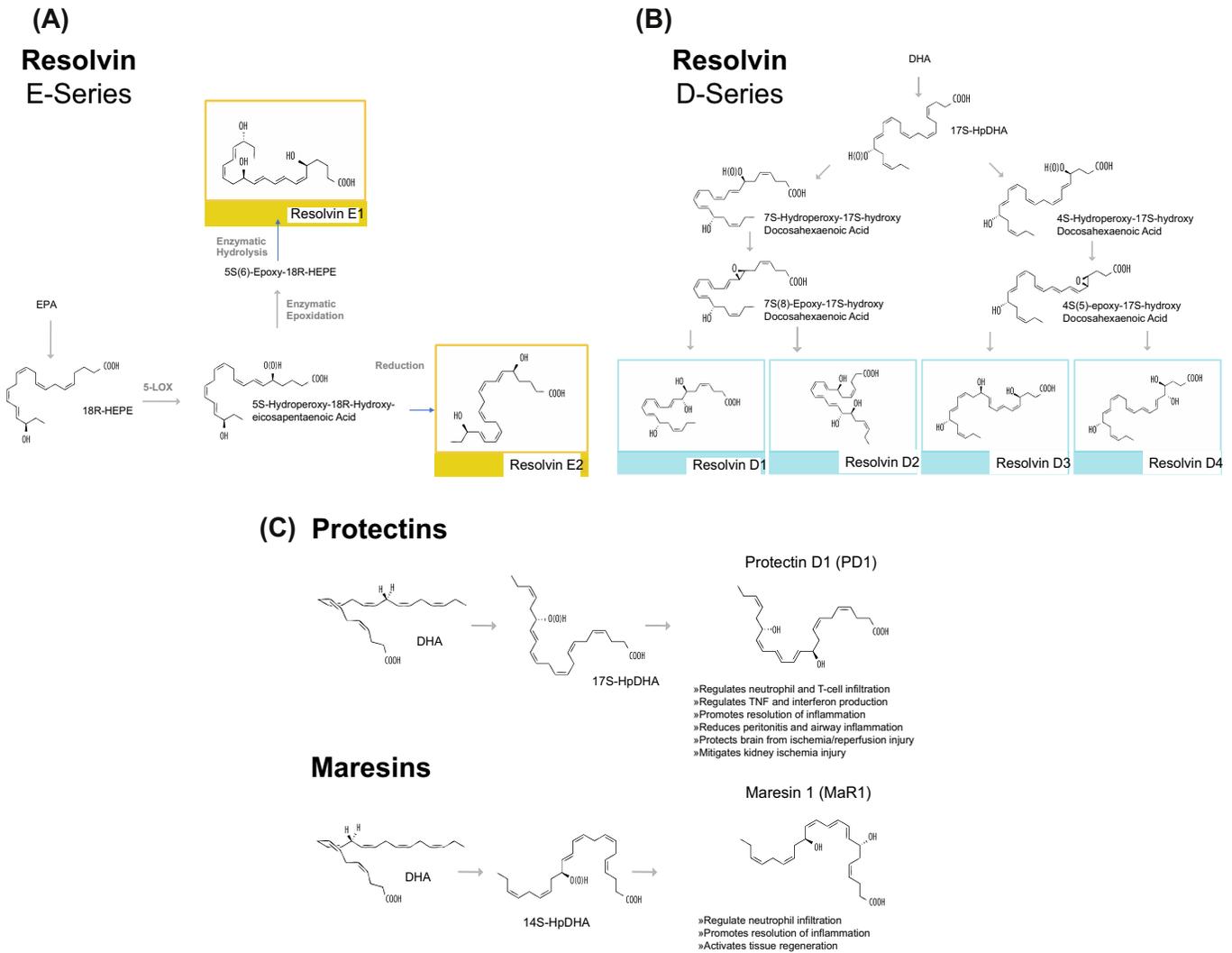


FIGURE 61.3 **Specialized pro-resolving mediators.** (A) E-series resolvins; (B) D-series resolvins; (C) protectins and maresins. *DHA*, docosahexaenoic acid; *EPA*, eicosapentaenoic acid; *HEPE*, 18R-hydro(peroxy)-EPA.

inflammation returns to homeostasis via the newly uncovered resolution cascade. With this systems approach, it was also possible to establish the local and temporal dissociation of LM-SPM biosynthesis. For example, upon initiation of inflammation with sterile tumor necrosis factor (TNF)- α or live microbes, a typical acute-phase response characterized by rapid PMN infiltration is preceded by local generation of both PGs and LTs. Unexpectedly, eicosanoids undergo what we earlier termed a class switch, and the profiles of LM made within this milieu switch with time (Fig. 61.2). Potent chemoattractants such as LTB_4 , typically go signals, are inactivated, and 15-lipoxygenase required for LX, Rv, and most SPM production is transcriptionally activated. The omega-3 fatty acid SPM precursors DHA and EPA (Fig. 61.3A–C) are rapidly carried into the exudates via edema (Fig. 61.1) and are then made available for conversion by congregated neutrophils within exudates

or pus. This, which we coined *LM-class switching*, is driven in part by cyclooxygenase-derived PGE_2 and D_2 via transcriptional regulation of enzymes in lipoxin biosynthesis. Hence, the concept that alpha signals omega, (the beginning signals the end in inflammation), denotes that at time zero, mediators are biosynthesized that accumulate and then signal to limit and stop PMN influx, terminating the contained inflammatory response.

To qualify as a chemical mediator in the language of inflammation-resolution, a biosynthetic product must be stereoselective in its actions and be produced at the site in amounts commensurate with its potency. Lipoxins, resolvins, and protectins each are present in human serum (picomolar to nanomolar) (Table 61.1) concentrations (e.g., LXA_4 , ~ 1.4 nM; RvD1, ~ 50 pM; RvE1, ~ 0.5 nM) (see suggested readings for RvE1, RvE2, RvD1, and RvD2 reported values in human peripheral

TABLE 61.1 Human Specialized Pro-resolving Mediator via Mass Spectrometry: Liquid Chromatography–Tandem Mass Spectrometry–Based Identification.

Tissue/Organ	Specialized Pro-resolving Mediator
Milk	RvE1, RvD1
Serum	RvE1, RvE2, RvE3 RvD1, RvD2, RvD3, RvD5, RvD6, AT-RvD1, AT-RvD3 PD1, AT-PD1 MaR1
Lymph nodes	RvE2, RvE3 RvD1, RvD3, RvD5, RvD6, AT-RvD1, AT- RvD3 AT-PD1
Spleen	RvE1, RvE2, RvE3 RvD5 PD1 MaR1
Brain	PD1
Urine	RvE2 RvD1, AT-RvD1
Plasma	RvE1, RvE2, RvE3 RvD1, RvD2, RvD3, RvD5, RvD6, AT-RvD1, AT-RvD3 PD1, AT-PD1
Adipose tissue	RvD1, RvD2
Cerebrospinal fluid	RvD1, PD1 RvD1
Breath condensate	PD1
Placenta	RvD1, RvD2, AT-RvD1 PD1
Sputum (cystic fibrosis)	RvE1
Synovial fluid	RvD5 MaR1

AT, aspirin-triggered; MaR1, maresin 1; PD1, protectin D1; Rv, resolvin.

Table References:

Clària, J., Dalli, J., Yacoubian, S., Gao, F., Serhan, C.N., 2012. Resolvin D1 and resolvin D2 govern local inflammatory tone in obese fat. *J Immunol* 189, 2597–2605.

Colas, R.A., Shinohara, M., Dalli, J., Chiang, N., Serhan, C.N., 2014. Identification and signature profiles for pro-resolving and inflammatory lipid mediators in human tissue. *Am J Physiol Cell Physiol* 307, C39–54.

Giera, M., Ioan-Facsinay, A., Toes, R., Gao, F., Dalli, J., Deelder, A.M., Serhan, C.N., Mayboroda, O.A., 2012. Lipid and lipid mediator profiling of human synovial fluid in rheumatoid arthritis patients by means of LC-MS/MS. *Biochim Biophys Acta* 1821, 1415–1424.

Keelan, J.A., Mas, E., D'Vaz, N., Dunstan, J.A., Li, S., Barden, A.E., Mark, P.J., Waddell, B.J., Prescott, S.L., Mori, T.A., 2015. Effects of maternal n-3 fatty acid supplementation on placental cytokines, pro-resolving lipid mediators and their precursors. *Reproduction* 149, 171–178.

Levy, B.D., Kohli, P., Gotlinger, K., Haworth, O., Hong, S., Kazani, S., Israel, E., Haley, K.J., Serhan, C.N., 2007. Protectin D1 is generated in asthma and

dampens airway inflammation and hyper-responsiveness. *J Immunol* 178, 496–502.

Lukiw, W.J., Cui, J.G., Marcheselli, V.L., Bodker, M., Botkjaer, A., Gotlinger, K., Serhan, C.N., Bazan, N.G., 2005. A role for docosahexaenoic acid-derived neuroprotectin D1 in neural cell survival and Alzheimer disease. *J Clin Invest* 115, 2774–2783.

Mas, E., Croft, K.D., Zahra, P., Barden, A., Mori, T.A., 2012. Resolvins D1, D2, and other mediators of self-limited resolution of inflammation in human blood following n-3 fatty acid supplementation. *Clin Chem* 58, 1476–1484.

Pruss, H., Rosche, B., Sullivan, A.B., Brommer, B., Wengert, O., Gronert, K., Schwab, J.M., 2013. Proresolution lipid mediators in multiple sclerosis - differential, disease severity-dependent synthesis - a clinical pilot trial. *PLoS One* 8, e55859.

Psychogios, N., Hau, D.D., Peng, J., Guo, A.C., Mandal, R., Bouatra, S., Sinelnikov, I., Krishnamurthy, R., Eisner, R., Gautam, B., Young, N., Xia, J., Knox, C., Dong, E., Huang, P., Hollander, Z., Pedersen, T.L., Smith, S.R., Bamforth, F., Greiner, R., McManus, B., Neuman, J.W., Goodfriend, T., Wishart, D.S., 2011. The human serum metabolome. *PLoS One* 6, e16957.

Sasaki, A., Fukuda, H., Shiida, N., Tanaka, N., Furugen, A., Ogura, J., Shuto, S., Mano, N., Yamaguchi, H., 2015. Determination of omega-6 and omega-3 PUFA metabolites in human urine samples using UPLC/MS/MS. *Anal Bioanal Chem* 407, 1625–1639.

Wang, X., Zhu, M., Hjorth, E., Cortés-Toro, V., Eyrjofsdottir, H., Graff, C., Nennesmo, I., Palmblad, J., Eriksdotter, M., Sambamurti, K., Fitzgerald, J.M., Serhan, C.N., Granholm, A.C., Schultzberg, M., 2015. Resolution of inflammation is altered in Alzheimer's disease. *Alzheimers Dement* 11, 40–50.e42.

Weiss, G.A., Troxler, H., Klinke, G., Rogler, D., Braegger, C., Hersberger, M., 2013. High levels of anti-inflammatory and pro-resolving lipid mediators lipoxins and resolvins and declining docosahexaenoic acid levels in human milk during the first month of lactation. *Lipids Health Dis* 12, 89.

Yang, J., Eiserich, J.P., Cross, C.E., Morrissey, B.M., Hammock, B.D., 2012. Metabolomic profiling of regulatory lipid mediators in sputum from adult cystic fibrosis patients. *Free Radic Biol Med* 53, 160–171.

blood). LM-lipidomics using liquid chromatography–tandem mass spectrometry (LC-MS/MS) coupled with informatics permits profiling of closely related compounds and identification of new molecules. Retrograde synthesis, both biogenic and total organic, allows the complete elucidation of chemical structure, stereochemistry, and physical properties, along with the recapitulation of their biosynthetic pathway (Fig. 61.3A–C). The roles of omega-3 PUFA in health were evident in 1929, and ω -3 fatty acids, also known as n-3 PUFAs, have beneficial effects on human diseases, including potential antithrombotic, immunoregulatory, and antiinflammatory properties. To address the molecular basis for anti-inflammatory properties of ω -3 fatty acids, the unbiased LC-MS/MS-based informatics approach outlined earlier was used to identify novel SPMs biosynthesized from ω -3 precursors during acute inflammation. Using this approach, EPA and DHA were enzymatically converted into novel potent bioactive products coined **resolvins**, the acronym for *resolution phase interaction products*, because they (1) are produced during cell–cell interactions occurring in the resolution phase of acute inflammatory response, (2) stop further neutrophil entry to sites of inflammation, and (3) reduce exudates. Resolvins represented an entirely new family of mediators

produced from the ω -3 fatty acids. They appeared during the resolution phase via active biosynthetic processes. The biosynthesis of resolvins results in stereospecific local mediators that have potent actions and activate specific receptors (Fig. 61.3).

RESOLVINS: E-SERIES

EPA-derived E-series Rv (Fig. 61.3A) are biosynthesized in resolving murine exudates and by isolated human cells (e.g., endothelial cell–leukocyte interaction) and in whole blood (vide infra). The complete stereochemistry of the first bioactive member of this family, resolvin E1 (RvE1), is 5*S*,12*R*,18*R*-trihydroxy-6*Z*,8*E*,10*E*,14*Z*,16*E*-EPA. Within vascular endothelial cells, aspirin-acetylated cyclooxygenase (COX)-2 converts EPA into 18*R*-hydro(peroxy)-EPA (HEPE), which is rapidly taken up by activated leukocytes (e.g., neutrophils) and further converted into RvE1. Notably, quantitative chiral high-performance liquid chromatography analysis indicated that the 18*R*-HEPE isomer was dominant compared with its epimer 18*S*-HEPE in human plasma from healthy subjects taking EPA. Human subjects who were administered aspirin before EPA had more 18*S*- than 18*R*-HEPE, which indicated that aspirin might promote 18*S*-HEPE production as well as 18*R*-HEPE. The 18*S*-HEPE can also be converted to epimeric RvE1 and RvE2 by human recombinant 5-LOX and LTA₄ hydrolase, known as proinflammatory LTB₄-synthesizing enzymes. RvE1 is also produced from an aspirin-independent pathway via cytochrome P450-driven oxygenation of EPA. RvE2 (5*S*,18-dihydroxy-EPE) is biosynthesized in resolving exudates and in human whole blood via reduction of 5*S*-hydroperoxy,18-hydroxy-EPE, an intermediate. EPA is also precursor to RvE3, which also limits PMN infiltration (Fig. 61.3A).

RESOLVINS: D-SERIES

Investigations using LC-MS/MS lipidomics of resolving exudates from mice given DHA and aspirin provided the first evidence of novel endogenous routes that lead to formation of 17-hydroxy-containing mediators. Gaining information on how human tissue and cells produce D-series resolvins (Fig. 61.3B) involved recapitulation of biosynthetic pathways with isolated human cells and recombinant enzymes. For example, hypoxic human endothelial cell COX-2 converts DHA to 13-hydroxy-DHA, which switched with aspirin to 17*R*-hydroxydocosahexaenoic acid (HDHA). Human neutrophils transformed 17*R*-hydroxy-DHA into two series of di-

and trihydroxy bioactive products, one initiated via oxygenation at carbon-7 and the other at carbon-4 (Fig. 61.3B). These were coined the *aspirin-triggered* (AT)-D-series resolvins. D-series resolvins carrying a 17*S*-hydroxy group were also identified in murine exudates and isolated human cells without aspirin. The enzymatic pathway leading to formation of RvD1 is shown in Fig. 61.3B, and 17*R*-D-series are similar save for the chirality at carbon 17 as *R* when aspirin is given. After the complete organic synthesis, stereochemistry of 17*S*-, 17*R*-RvD1, and RvD2 were each established: 7*S*,8*R*,17*S*-trihydroxy-4*Z*,9*E*,11*E*,13*Z*,15*E*,19*Z*-DHA (RvD1), 7*S*,8*R*,17*R*-trihydroxy-4*Z*,9*E*,11*E*,13*Z*,15*E*,19*Z*-DHA (17*R*-RvD1 or AT-RvD1), and 7*S*, 16*R*,17*S*-trihydroxy-4*Z*,8*E*,10*Z*,12*E*,14*E*,19*Z*-DHA (RvD2) (Fig. 61.3B). Additional members of this family were identified (RvD3–RvD4), with confirmed bioactions (interested readers are directed to references within the suggested readings).

PROTECTINS

DHA is also the precursor of another family of LM characterized by a conjugated-triene double bond system called PD. The name PD accounts for their protective actions observed in neural tissues and within the immune system, whereas the prefix “neuro” gives the tissue localization and site of action. The structure of the founding member of this family, PD1, was first disclosed in a report on the isolation and elucidation of resolvins, and its complete stereochemistry is 10*R*,17*S*-dihydroxy-docosa-4*Z*,7*Z*,11*E*,13*E*,15*Z*,19*Z*-hexaenoic acid (Fig. 61.3C). In addition to PD1, several stereo- and positional isomers that were also identified in human and mouse tissues had lower bioactivity than PD1. These include 10*S*,17*S*-diHDHA (also known as PDX), 4*S*,17*S*-diHDHA, and 22-hydroxy-10,17*S*-docosatriene (a putative inactivation product of PD1). The geometry of the double bonds in PD1, their positions during biosynthesis, and chirality of C10 indicate that PD1 biosynthesis proceeds via a confirmed C16(17)-epoxide intermediate that requires specific enzymatic steps to biosynthesize the potent bioactive PD1, confirmed by the identification of alcohol-trapping products as well as two vicinal diol 16,17-docosatrienes as minor products of PD1 biosynthesis and total organic synthesis. The aspirin-triggered COX 2-driven pathway was uncovered that biosynthesizes the 17*R*-epimeric form of PD1, namely AT-PD1 (Fig. 61.3C). Both PD1 and AT-PD1 reduced leukocyte infiltration in murine peritonitis, reduced PMN transmigration with endothelial cells, and enhanced efferocytosis of apoptotic PMN by human M Φ , hallmark actions of an SPM (Fig. 61.3).

MARESINS

MΦs have critical roles orchestrating the return to homeostasis and biosynthesizing SPM. For example, MΦ ingesting apoptotic cells initiate the biosynthesis of LXA₄, RvE1, and PD1, but not LTB₄. In addition, a family of SPM biosynthesized by MΦs was identified. Unbiased LM LC-MS/MS-based metabolite-lipidomics during self-limited peritonitis led to the identification of a novel pathway that converted DHA into 14-HDHA (Fig. 61.3C). Experiments with 12/15-LO^{-/-} mice or with inhibitor confirmed that 14-HDHA was via a DHA carbon 14 lipoxygenation pathway. 14-H(p)DHA was rapidly converted by isolated human and mouse MΦ into a new set of bioactive products (Fig. 61.3C) whose molecular structure was established and confirmed by total organic synthesis. A major product of this new pathway, first in this family, proved to be 7R,14S-dihydroxydocosa-4Z,8E,10E,12Z,16Z,19Z-hexaenoic acid, denoted macrophage mediator in resolving inflammation 1 (maresin 1) (MaR1). Similar to other potent SPMs, MaR1 biosynthesis proceeds via an epoxide intermediate, specifically maresin-13,14-epoxide, which opens during enzymatic conversion to set the double-bond geometry and obtain chirality of carbon-7 alcohol (Fig. 61.3C). MaR1 and RvE1 are potent stimulators of organ regeneration uncovered using planaria regeneration. Hence, SPMs are primordial molecules that signal from the inflammatory site to generate the aftermath of inflammation and tissue injury and are produced across the animal kingdom (i.e., planaria, fish, minipigs, baboons, and human tissues).

SPECIALIZED PRO-RESOLVING MEDIATOR RECEPTORS

Resolvin E1 Receptors: We identified orphan GPCR ChemR23 for RvE1 (Fig. 61.4) that binds ³H-RvE1 and stereoselectively transduces signals to monocytes and dendritic cells. RvE1 also directly interacts with BLT1, a leukotriene B₄ receptor. Therefore, RvE1 gives cell-type specific actions, serving as an agonist for ChemR23 on mononuclear and dendritic cells as well as an antagonist for BLT1 signals on PMN. RvE1–ChemR23 interactions also stimulate macrophage phagocytosis via phosphorylation signaling pathways including ribosomal protein S6, a downstream target of PI3K/Akt signaling and the Raf/ERK pathways. 18S-RvE1 also binds to ChemR23 with increased affinity and potency compared with the R-epimer but is rapidly inactivated. RvE2 is a partial agonist for ChemR23.

Resolvin E1 and LTB₄ receptor (BLT1) Axis: BLT1 also directly interacts with RvE1, which inhibits calcium mobilization, NF-κB activation, and PMN infiltration

in vivo. Both 18S-RvE1 and RvE2 also bind to BLT1, demonstrating ligand specificity and related structural features of resolvins.

Resolvin D1, Lipoxin A₄ Receptor, and G-protein coupled receptor 32 (GPR32) Receptors: The pro-resolving actions of RvD1 are mediated via both lipoxin A₄ receptor (ALX) and human GPR32. The two separate GPCRs for RvD1 are on human phagocytes (ALX and GPR32). RvD1 displays specific binding, reduces actin polymerization and CD11b on PMN, and stimulates macrophage phagocytosis in an ALX- and GPR32-dependent manner. The aspirin-triggered epimer 17R-RvD1 and stable analog mimetic 17-R/S-methyl-RvD1 each dose-dependently activate ALX/formyl peptide receptor 2 (FPR2) and GPR32 in GPCR-overexpressing β-arrestin systems and electric cell-substrate impedance sensing. RvD5 also activates human GPR32, and macrophage phagocytosis of *Escherichia coli* is enhanced by GPR32. RvD3 and AT-RvD3 each activate this GPCR, contributing to their pro-resolving actions. Thus, D-series Rv ligands for GPR32 include RvD5, RvD3, and AT-RvD3 (Fig. 61.4A). PD1 displays specific binding with human retinal pigment epithelial cells and neutrophils. However, neither RvE1 nor LXA₄ competes for [³H]-NPD1/PD1 specific binding with human neutrophils, which suggests specific receptors for PD1.

Transgenic Mice and Knockout Mice: The pro-resolving actions of RvD1 were further enhanced in ALX-transgenic (TG) mice overexpressing human ALX. Transgenic mice overexpressing human ChemR23, the RvE1 receptor, on myeloid cells were also prepared. In these TG mice, RvE1 is 10-fold more potent in limiting PMN infiltration in zymosan-initiated peritonitis, and ligature-induced alveolar bone loss was diminished in ChemR23tg mice. Hence, RvE1 modulates osteoclast differentiation and bone remodeling by direct actions on bone, in addition to antiinflammation and pro-resolution. In ALX/fpr2-deficient mice (mouse ortholog of human ALX), antiinflammatory actions of RvD1 are lost (Fig. 61.4A). RvD1 regulates acute inflammation in part via human ALX and GPR-32 receptors via miRNA and elucidates the target organs for SPM biosynthesis and action (Fig. 61.4) (see subsequent discussion).

SPECIALIZED PRO-RESOLVING MEDIATORS OF RESOLUTION IN REGENERATION

Complete resolution requires regeneration of destroyed tissues without affecting their functionality, as in the case of fibrosis or scarring. Pro-resolving MΦ have key functions in tissue remodeling in both homeostatic (e.g., postparturition) and pathological (e.g., removal of microbes from infected tissues) settings. In

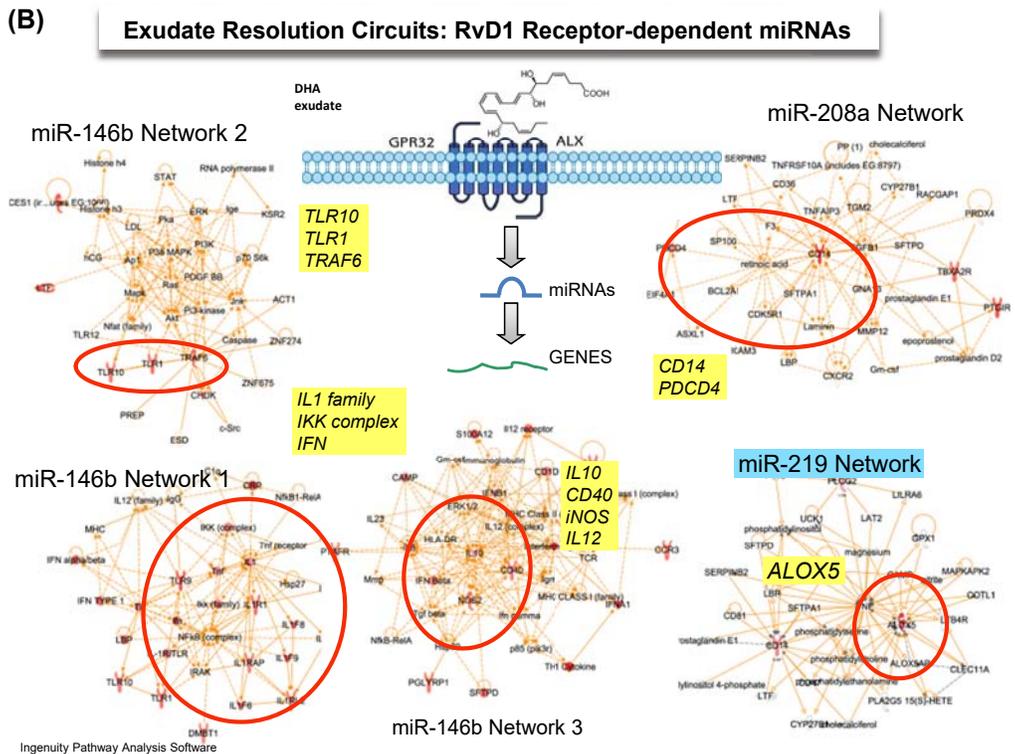
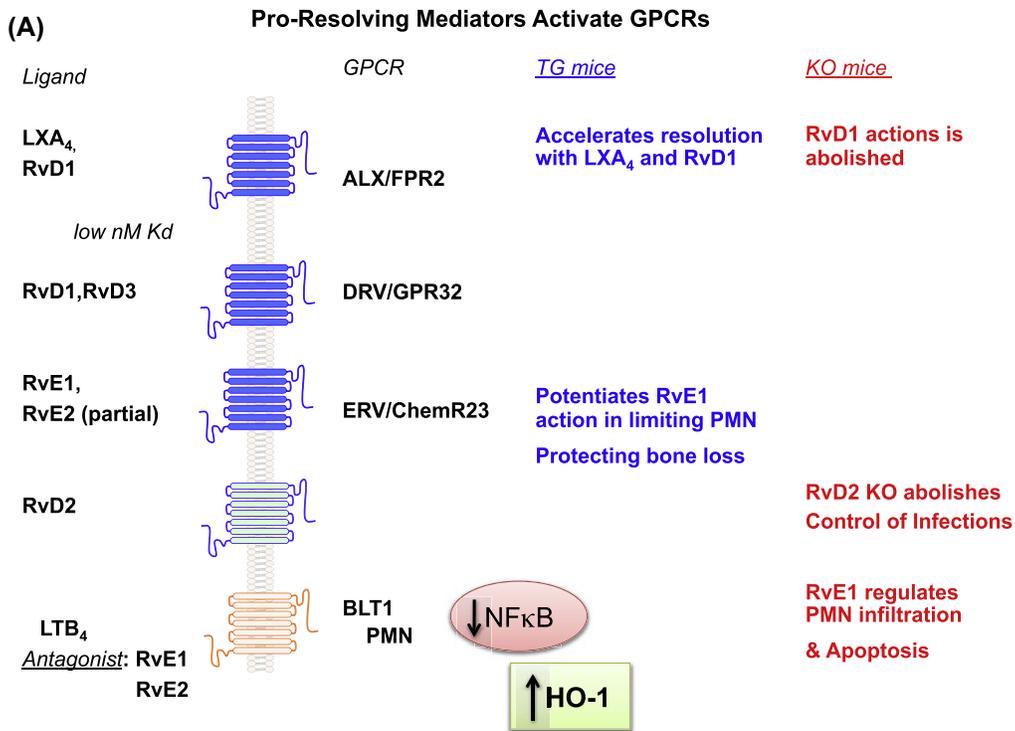


FIGURE 61.4 Resolvin (Rv)D1-G-protein coupled receptor (GPCR) signaling network. (A) Pro-resolving mediators activate GPCRs; (B) exudate resolution circuits: RvD1-GPCR-dependent micro RNAs (miRNAs). *ALX/FPR2*, lipoxin A₄ receptor/formyl peptide receptor 2; *ERV/ChemR23*, *RvE1* receptor; *HO-1*, heme oxygenase-1; *KO*, knockout; *LTB₄*, leukotriene B₄; *LXA₄*, lipoxin A₄; *NFκB*, nuclear factor κB; *PMN*, polymorphonuclear neutrophil; *RvD1*, resolvin D1; *TG*, transgenic; *ALOX5*, 5-lipoxygenase; *ALX*, lipoxin A₄ receptor; *CD40*, cluster of differentiation 40; *IFN*, interferon; *IKK*, IκB kinase; *IL1*, interleukin 1; *iNOS*, inducible nitric oxide synthase; *miRNA*, microRNA; *PDCD*, pyruvate dehydrogenase complex deficiency; *RvD1-GPCR*, resolvin D1–G-protein coupled receptor; *TLR*, Toll-like receptor; *TRAF*, tumor necrosis factor receptor associated factor.

this regard, SPMs are considerably interesting in view of their roles in regulating M Φ activities. LX, Rv, and PD stimulate non-phlogistic efferocytosis by M Φ . In addition, RvD1 regulates M Φ accumulation in diabetic obese mice and reduces arthritic pain. Failure of the M Φ -driven pro-resolution program can support persistent inflammation associated with many human diseases, such as periodontitis. In addition to enhancing M Φ phagocytosis, MaR1 biosynthesized in vivo during tissue injury repair accelerated tissue regeneration in planaria (*Dugesia tigrina*) after surgical head removal.

GENES AND MICRORNAS IN RESOLUTION CIRCUITS

Because miRNAs emerged as the fine-tuners of many cellular processes, including immune responses and cancer, it was interesting to investigate whether they also had roles in resolution and to detect SPM-regulated specific miRNA as part of their mechanisms of action. To identify an miRNA signature of resolution in self-resolving exudates, a strategy with resolving exudates from murine peritonitis was used (Fig. 61.4B). Zymosan A particles, a Toll-like receptor 2 and 4 ligand from the yeast *Saccharomyces cerevisiae*, were injected intraperitoneally and peritoneal exudates were collected to monitor temporal changes in both leukocyte numbers and composition. Rapidly after injection, leukocytes infiltrate the peritoneal cavity during the onset phase of acute inflammation (4 h), reaching a maximum ($\sim 22.0 \times 10^6$ cells) at about 12 h and declining at 24–48 h. Because it is important to define resolution in unbiased terms, resolution indices, introduced by the author's laboratory, were calculated (i.e., $\Psi_{\max} \sim 15.0 \times 10^6$; $T_{\max} \sim 12$ h; $\Psi_{50} \sim 7.5 \times 10^6$; $T_{50} \sim 30$ h; $R_i \sim 18$ h) to characterize the resolution phase and determine temporal changes in miRNA expression during this time interval. Hierarchical clustering grouped the approximately 300 miRNAs examined into distinct clusters based on their relative abundance at the different time intervals, indicating that specific miRNAs are temporally regulated during acute inflammation and its natural self-limited resolution. In these, miR-21, miR-146b, miR-208a, and miR-219 were significantly regulated at 12 and 24 h compared with 4 h (Fig. 61.4B), which suggested a role in resolution. For example, miR-21 proved critical for the production of antiinflammatory IL-10 in experimental peritonitis. Quantitative real-time polymerase chain reaction (PCR) analyses carried out with exudates revealed that RvD1 temporally controls the specific sets of pro-resolving miRNAs miR-21, miR-146b, miR-208a, and miR-219, coined "resolution miRs" in exudates in vivo. For translation to humans, actions of RvD1 on these miRNAs were assessed in ALX/FPR2- and human RvD1 receptor denoted DRV1/

GPR32-overexpressing isolated human cells, namely M Φ , the master regulator cells of resolution. Hence, RvD1 regulates miR-21, -146b, -208a, and -219 in a GPCR-dependent manner.

miRNAs regulate mRNAs and protein levels of hundreds of genes involved in biological processes; hence, miR-target genes were clustered using the Ingenuity Pathway Analysis knowledge database based on their physical and/or functional interactions, creating the first identified RvD1-GPCR-regulated gene networks involved in inflammation and resolution (Fig. 61.4B). For example, miR-146 networks 1 and 3 included genes of the NF- κ B activation pathway (e.g., inhibitor of κ B [I κ B] kinase and TNF-associated factor 6) and innate response to pathogens (e.g., Toll-like receptors, S100 protein, C-reactive protein, peptidoglycan recognition protein) (Fig. 61.4B). Several cytokines and chemokines (IL-8, 10, and 12 and interferon- α and β) belonged to the miR-146b network 2 (Fig. 61.4B). NF- κ B is a critical transcription factor involved regulating cell functions in inflammation and resolution. RvD1 in human monocytes reduces the nuclear translocation of NF- κ B and TNF α -induced phosphorylation of I κ B, counteracts NF- κ B activation in ALX/FPR2 and DRV1/GPR32 recombinant cells, dampens acute inflammation in murine dorsal air pouches evoked by local administration of TNF- α , and downregulates I κ B kinase levels in murine peritonitis. Therefore, regulation of miRNAs and the TNF- α /NF- κ B axis appears to be a key component in the RvD1-GPCR signaling network (Fig. 61.4).

The microRNA miR-208a downregulates CD14, CD40 ligand, PGI $_2$ receptor, thromboxane A $_2$ receptor, and programmed cell death (Fig. 61.4B), and a tumor suppressor molecule and translational repressor of IL-10, consistent with the existence of a seed region for miR-208a in the 3'UTR of PDCD4. In self-limited peritonitis, RvD1 reduced PDCD4 and increased IL-10 production, providing in vivo correlation for RvD1-miR-dependent gene regulation. The miR-219 network includes CD14 and 5-lipoxygenase (LOX) (Fig. 61.4B), a key enzyme for biosynthesis of LTs, and SPMs. A significant reduction in 5-LOX protein and LTB $_4$ production is found in human M Φ overexpressing RvD1-regulated miR-219, translating findings in mouse peritonitis to human M Φ (Fig. 61.4B). Endogenous regulatory mechanisms of 5-LOX are of wide interest because of the important roles of LTs, LXs, and Rvs in inflammation and resolution. In addition to transcriptional regulation by cytokines and growth factors, miR-219 can provide a rapid mean to balance the abundance of 5-LOX protein in cells under dynamic conditions such as during inflammation. RvD1 actions were also tested in genetically engineered mice to obtain further evidence of ALX/FPR2-dependent action. Transgenic mouse colonies were created in which the human ALX/FPR2

Steps Toward Human Translation

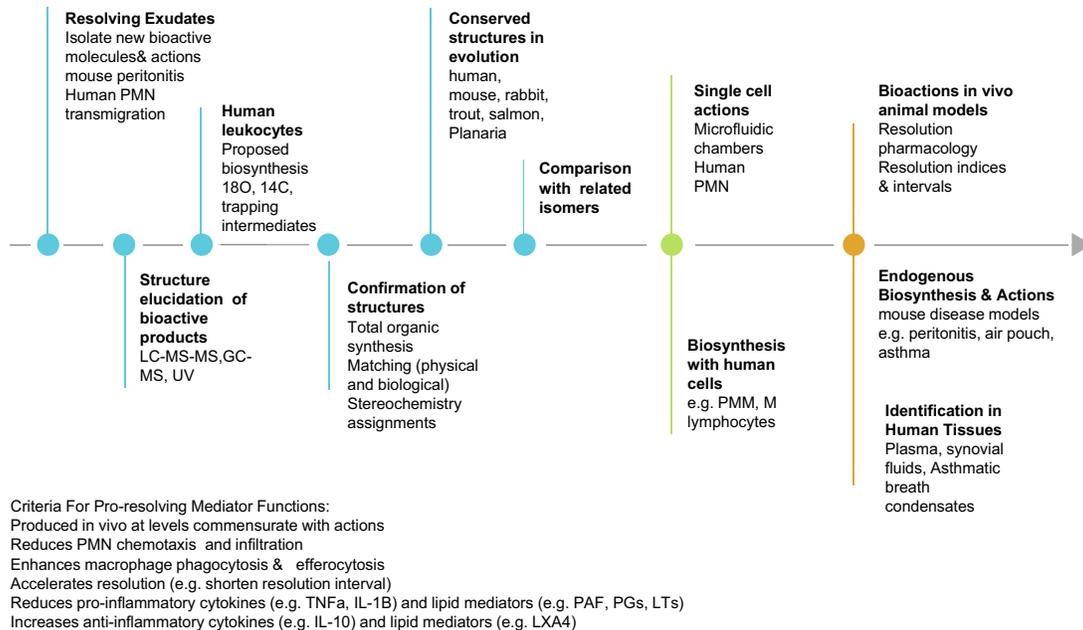


FIGURE 61.5 Pro-resolving mediators: toward human translation. Focus on structure function.

gene was placed under control of a CD11b promoter that drives transgene expression in mature murine myeloid cells. Total peritoneal leukocytes at 24 h peritonitis are significantly lower in ALX/FPR2 transgenic mice treated with zymosan alone compared with their non-transgenic littermates, and RvD1 administration results in a further significant decrease (~53% reduction) in peritoneal exudate leukocyte numbers in ALX/FPR2 transgenic mice compared with non-transgenic littermates. Quantitative PCR analysis of miRNAs isolated from peritoneal lavages collected 24 h after injection inflammatory challenge showed that RvD1 significantly upregulated identified resolution miRNAs miR-208a and miR-219 in vivo (Fig. 61.4B). In ALX/FPR2^{-/-} mice, RvD1 does not regulate PMN infiltration, nor does it significantly alter miR-208a or miR-219 expression at 24 h (Fig. 61.4B). Hence, RvD1 controls leukocyte infiltration and specific resolution-phase miRNAs in acute inflammation in an ALX/FPR2-dependent manner in mice and DRV1/GPR32-dependent interactions in human cells. Fig. 61.5 illustrates the main steps taken to validate and establish the functions of SPMs in humans toward translation.

CONCLUSION AND SUMMARY KEY POINTS

- The first response in acute inflammation is ubiquitous and mounts throughout vascularized organs in the body (Figs. 61.1 and 61.2).
- The SPM families and pathways are agonists of resolution; they stimulate resolution and counter proinflammatory signals rather than block inflammation-protective pathways. As autacoids (local-acting mediators), when and where count for LM and SPM production in vivo.
- Each SPM (Fig. 61.3) stimulates cessation of neutrophil influx (*efferocytosis*) and enhances phagocytosis for microbial containment. These steps are defining SPM functions for the structural elucidation of SPM and are the key signs of resolution of inflammation.
- Each SPM pathway has additional, nonredundant functions on target cell types.
- At the cellular and molecular levels, SPMs counterregulate proinflammatory mediators (eicosanoids, chemokines, cytokines, and

adipokines), regulate specific miRNAs and cell traffic, and enhance microbial killing by receptor-mediated mechanisms in animal models *in vivo* and with human neutrophils and macrophages.

Results from the first human clinical phase I/phase II trial with an Rv mimetic analog appear to be encouraging and can open new opportunities for resolution pharmacology based on endogenous mediators to terminate inflammation and treat inflammation-related diseases. It is hoped that more human trials will be launched to test the notion that stimulating resolution mechanisms can improve disease and health status with both SPMs and their precursors in nutritional studies relating their production to dietary supplementation and clinical outcomes.

Acknowledgments

This work was supported in part by the US National Institutes of Health (Grant R01GM38765).

Suggested Readings

- Serhan, C.N., Chiang, N., Dalli, J., 2015. The resolution code of acute inflammation: novel pro-resolving lipid mediators in resolution. *Semin Immunol* 27, 200–215.
- Serhan, C.N., 2014. Pro-resolving lipid mediators are leads for resolution physiology. *Nature* 510, 92–101.
- Serhan, C.N., Ward, P.A., Gilroy, D.W., 2010. *Fundamentals of Inflammation*. Cambridge University Press, New York.
- De Caterina, R., 2011. n-3 fatty acids in cardiovascular disease. *N Engl J Med* 364, 2439–2450.
- Samuelsson, B., 1982. From Studies of Biochemical Mechanisms to Novel Biological Mediators: Prostaglandin Endoperoxides, Thromboxanes and Leukotrienes, *Les Prix Nobel: Nobel Prizes, Presentations, Biographies and Lectures*. Almqvist & Wiksell, Stockholm, pp. 153–174.
- Samuelsson, B., Dahlen, S.E., Lindgren, J.A., Rouzer, C.A., Serhan, C.N., 1987. Leukotrienes and lipoxins: structures, biosynthesis, and biological effects. *Science* 237, 1171–1176.
- Perretti, M., Leroy, X., Bland, E.J., Montero-Melendez, T., 2015. Resolution pharmacology: opportunities for therapeutic innovation in inflammation. *Trends Pharmacol Sci* 36, 737–755.
- Barden, A.E., Mas, E., Mori, T.A., 2016. n-3 Fatty acid supplementation and proresolving mediators of inflammation. *Curr Opin Lipidol* 27, 26–32.

This page intentionally left blank

Nutrients and Gene Expression in Cardiovascular Disease

Marika Massaro¹, Egeria Scoditti¹, Nadia Calabriso¹,
Maria Annunziata Carluccio¹, Paul Hugenholtz², Raffaele De Caterina³

¹National Research Council (CNR), Institute of Clinical Physiology, Lecce, Italy ²Erasmus Medical Centre, Rotterdam, The Netherlands ³Cattedra di Malattie dell'Apparato Cardiovascolare, Università degli Studi di Pisa, Italy

Glossary

Cardiovascular disease Cardiovascular disease (CVD) is a general term used to indicate pathological conditions affecting the heart or blood vessels. The largest proportion of CVD is atherosclerotic vascular disease, often used as a synonym for CVD. This affects the intima of medium to large arteries and is usually associated with a buildup of inflammatory fatty deposits, with an increased risk of thrombosis, affecting the brain, heart, kidneys, and other organs. CVD is the main cause of death and disability worldwide, but it can be largely prevented by adopting healthy lifestyles.

Nutrigenomics Nutrigenomics is the area of medicine that uses complementary molecular tools, including biochemistry, physiology, proteomics, metabolomics, transcriptomics, and epigenomics, to determine and elucidate how the components of a particular diet (bioactive compounds) may affect the expression of our genes. It focuses on understanding the interaction between nutrients with the genome at the molecular level, to explain how specific nutrients or dietary regimes may affect human health.

Pathway analysis Pathway analysis is related to the bioinformatic evaluation performed to interpret “omics” data in a structured fashion to form integrated flow diagrams. It enables distinct cellular processes, diseases, or signaling pathways to be found that are statistically associated with a selection of differentially expressed genes between two samples. Often, but erroneously, pathway analysis is used as synonym for network analysis, which conversely aims to highlight functional interaction between biological molecules.

Systems medicine Systems medicine is a novel discipline in biomedical research closely related to systems biology. It aims to study complex interactions within the human body in light of patients' genomics, behavior, and environment, with the major focus on developing computational models that describe disease progression and the effect of therapeutic interventions.

INTRODUCTION

Cardiovascular diseases (CVDs) have atherosclerosis, the inflammatory-mediated degenerative disease of the vascular wall, at their core. The best recognized therapeutic approach to CVD is to acquire and implement adequate lifestyle habits (Sacks et al., 2017). Because many recognized dietary risk factors contribute to disease progression in connection with genetic settings, an understanding of the impact of nutrients on the complex networking of the entire genome has been envisaged as a highly advisable medical research goal. Although for years this has turned out to be a daunting enterprise, improvements in high-throughput technologies and their wider application to nutritional studies in animals and humans are constantly disclosing potential pathogenic players and targets in atherosclerosis and its related thrombotic complications, paving the way to adopting comprehensive and effective dietary management of cardiovascular risk (Merched and Chan, 2013).

The Basis of Cardiovascular Disease: A Brief Summary

Inflammation is widely established as the driving force for CVD development. Early vascular lesions develop at specific sites of the arterial tree, such as curvatures and bifurcations, characterized by nonlaminar and turbulent flow, and reduced shear stress. In these

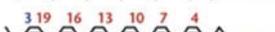
regions, risk mediators, in particular high levels of low-density lipoprotein (LDL) cholesterol, but also persistent high levels of glucose, insulin, homocysteine, and/or products of cigarette smoking, synergize with nonlaminar flow in promoting endothelial endoplasmic reticulum stress and an increase in the permeability to lipoproteins. The ensuing endothelial activation promotes the recruitment of circulating monocytes through the de novo expression and/or upregulation of a number of endothelial-leukocyte adhesion molecules (ELAMs) and chemokines that sustain the attraction and transmigration of circulating monocytes within the subendothelial space. Here monocytes are prompted to take-up lipoproteins, turning them into macrophage-derived foam cells, pathognomonic signatures of the early atherosclerotic lesions in the arterial intima. Concordantly, LDL-lowering agents, such as the statins, are still the most utilized therapeutic agents in current clinical practice. Despite statins, patients appear to have a residual risk of CVD-related events, only in part related to the small proportion of patients unable to achieve target LDL cholesterol levels. Therefore, the search for new therapeutics that can either be taken alone or combined with statins is warranted and should be encouraged. One complementary research is the search for vasculoprotective dietary micro- and macronutrient components (Merched and Chan, 2013). A deeper understanding of the gene expression modulatory activity exerted by selected dietary factors promises to characterize unsuspected biological pathways and identify gene expression regulations and modulated pathways as the biological basis for new therapeutic strategies. In subsequent discussion, we will describe how typical dietary vasculoprotective components include unsaturated fatty acids (FAs) and antioxidant polyphenols as ways to modulate expressions of genes involved in the pathogenesis of atherosclerosis and contributing to atherosclerotic cardiovascular risk.

OMEGA-3 POLYUNSATURATED FATTY ACID GENE REGULATORY ACTIVITIES IN INFLAMMATORY MODELS OF ATHEROSCLEROSIS

As reviewed in nutritional advisories by the American Heart Association (Sacks et al., 2017), in the past half-century, large epidemiological and experimental evidence has accumulated regarding the role of nutritional saturated fatty acids (SFAs) as CVD risk factors, and the benefits associated with the omega-3 polyunsaturated fatty acids (omega-3 or n-3 PUFAs) in managing CVD (Siscovick et al., 2017) (Table 62.1). Because omega-3 PUFAs have been recognized to exert antiinflammatory and antithrombotic properties, interest in

evaluating their properties of regulating gene expression in vascular cells, including vascular endothelial cells and monocytes, has received unprecedented attention (De Caterina, 2011). By using a candidate-gene approach, with an a priori selection of molecules and genes of interest based on the knowledge of pathogenetic pathways, our own research disclosed potential mechanistic explanations for the preventive and therapeutic use of omega-3 PUFAs. We used human adult saphenous vein endothelial cells and human umbilical vein endothelial cells (HUVEC) activated by proinflammatory cytokines as in vitro models of the early steps in atherogenesis. Here, we assessed the effects of the most abundant omega-3 fatty acid, docosahexaenoic acid (DHA) on several determinants of endothelial inflammation. We observed that, when added to cultured endothelial cells early enough to allow a significant incorporation in cell membrane phospholipids, DHA inhibited the expression of genes involved in endothelial activation, including the athero-ELAMs vascular cell adhesion molecule (VCAM)-1, E-selectin, and intercellular adhesion molecule (ICAM)-1, as well as the proinflammatory enzyme cyclooxygenase (COX)-2. The dampening by DHA of the endothelial proinflammatory pathways elicited by inflammatory cytokines occurred in a range of DHA concentrations compatible with the nutritional intake of DHA when added to a normal Western diet by increased fish intake or pharmacological supplementation and was strictly related to the extent of incorporation into total cell lipids. This effect appeared to include also other cytokine-activated products, such as the soluble proteins interleukin (IL)-6, IL-8, monocyte chemoattractant protein (MCP)-1 and macrophage colony-stimulating factor (M-CSF), all involved in amplifying the inflammatory response. From a mechanistic point of view, we observed that incorporation of DHA in cell membranes decreased the generation of intracellular reactive oxygen species (ROS) and the related activation of a series of signaling switches, converging into the inhibition of the redox-sensitive transcriptional factor nuclear factor (NF)- κ B, involved in the expression of most proinflammatory genes in vascular endothelial cells and monocytes (De Caterina, 2011). Other known signaling systems affected by omega-3 PUFAs include the Toll-like receptors (TLRs) and the sarcoma family kinases. Part of the effects of omega-3 PUFAs may occur through more or less specific receptors, such as the recently discovered G-protein coupled receptor (GRP)120. In macrophages, its direct activation by omega-3 PUFAs, was shown to inhibit the expression of tumor necrosis factor (TNF)- α , IL-6, and MCP-1. Furthermore, by regulating the expressions and the transcriptional activities of sterol regulatory element-binding proteins and peroxisome proliferator-activated receptor (PPAR) α , omega-3 PUFAs may

TABLE 62.1 Chemical Structure and Dietary Sources of the Most Common Nutritional Fatty Acids.

Chemical Structure	Full Name (Abbreviation)	Carbon Number (ω Position)	Type of Fatty Acids	Essential (Yes/No)	Common Dietary Sources
	Palmitic acid (PA)	16	Saturated	No	Meats, dairy, palm and coconut oil
	Stearic acid (SA)	18	Saturated	No	Meats, dairy, palm and coconut oil
	Oleic acid (OA)	18 (ω -9)	Monounsaturated	No	Olive oil, canola oil, high oleic safflower oil
	Linoleic acid (LA)	18 (ω -6)	Polyunsaturated	Yes	Safflower, sunflower, sesame, soybean and canola oil, nuts and peanuts
	α -Linolenic acid (ALA)	18 (ω -3)	Polyunsaturated	Yes	Flaxseeds, and flaxseed oil, canola oil, soybeans and soybean oil, pumpkin seeds and pumpkin seed oil, perilla seed oil, almonds
	Arachidonic acid (AA)	20 (ω -6)	Polyunsaturated	Yes	Meat, poultry, and eggs
	Eicosapentaenoic acid (EPA)	20 (ω -3)	Polyunsaturated	Yes	Fish and fish oil
	Docosahexaenoic acid (DHA)	22 (ω -3)	Polyunsaturated	Yes	Fish and fish oil

The ω double bond position (from methyl end) is indicated in blue; the double bond positions from the carboxyl end are indicated in red.

downregulate the expression of several lipogenic genes, including the fatty acid synthase, stearoyl-coenzyme A desaturase 1, and PCSK9 (De Caterina, 2011). The list of the gene regulatory activities of omega-3 PUFAs has grown exceptionally thanks to the applications of high-throughput genomic tools that allow the simultaneous comparison of thousands of genes and gene products in a fashion not biased or restricted by a priori hypotheses. In particular, we have evaluated the transcriptomic profile of HUVEC activated by the proinflammatory cytokine IL-1 β after exposure to DHA. We observed that DHA deeply modifies the transcriptomic profile of endothelial cells in both basal and stimulated conditions. DHA alone altered the expression of 188 genes, decreasing 92 and increasing 96 of them. IL-1 β changed the expression of 2031 genes, decreasing 997 and increasing 1034 of them. Treatment with DHA after endothelial cell inflammatory stimulation significantly affected the expression of 116 IL-1 β -deregulated genes. Functional and network analyses identified immunological, inflammatory, and metabolic pathways as most affected by DHA exposure. Newly identified DHA-regulated genes are involved in stemness, cellular growth, cardiovascular system function, and cancer and included, among others, previously totally unsuspected genes of potential cardiovascular interest, such as junctional adhesion molecule-A, transcriptional regulator caspase recruitment domain 11, trombospondin-1 receptor (also known as Cluster of Differentiation [CD] 47), and the vasoconstrictor enzyme phosphodiesterase 5 α (Massaro et al., 2015).

Transcriptomic Evaluation of Omega-3 Polyunsaturated Fatty Acids in Human Adipose Tissue and Peripheral Blood Cells

There is a growing understanding of the limitations of using animal models and cell lines to explore human pathogenic processes. The degree of match between gene activation in inflammatory disease in humans and mice has been found to be low, and research in human cellular models is more suitable than in animal models for a thorough understanding of tissue-specific mechanisms of nutrient action (Seok et al., 2013). This consideration highlights the need to perform experiments to evaluate the activity of nutrients directly in humans. Because blood sample and subcutaneous adipose tissue collection is minimally invasive for use in clinical settings, the evaluation of gene expression signatures in blood cells and the adipose tissue may be particularly advantageous to combine epidemiological and molecular information in nutrigenomics. In particular, as part of the transport system, blood cells interact with all tissues and are exposed to nutrients,

metabolites, and waste products. Furthermore, blood mononuclear cells together with endothelial cells represent key players in atherosclerosis. For these reasons, blood cells have been suggested as the best surrogate tissues to study responses to dietary interventions in the context of CVD (de Mello et al., 2012). With this background, intervention studies aiming to evaluate the gene regulatory effects of omega-3 PUFAs in human blood and adipose tissue cells have increased considerably. In the attempt to review the nutrigenomic effects of nutritional FAs in humans, we have selected nine studies answering the following criteria: to be human intervention studies testing unsaturated FAs and addressing the transcriptomic profile of blood cells or adipose tissue. Results from this selection are compiled and compared in Table 62.2. The first two studies listed here evaluated how monounsaturated fatty acids (MUFAs) and omega-3 PUFAs affect the transcriptomic profile of adipose tissue in overweight subjects. Both studies concluded that the administration of unsaturated FAs dampens the activation of proinflammatory genes in adipose tissue. The next study, by Bouwens et al., was the first to report the effect of omega-3 PUFA supplementation on the whole-genome expression profile in human peripheral blood mononuclear cells (PBMCs). As expected, triglyceride plasma levels were significantly reduced in the omega-3 PUFA (fish oil) group. Although changes in C-reactive protein plasma levels were not evident, the genomic effects of omega-3 PUFAs were prominent, because 1040 gene transcripts turned out to be altered after fish oil intervention. Pathway analysis of regulated genes showed that fish oil significantly decreased the expression of genes involved in the inflammatory and hypoxia signaling pathways. Of relevance, the expression of genes most tightly related to atherosclerosis, including cell adhesion molecules and scavenger receptors, as well as of genes involved in activating proinflammatory and metabolic transcriptional factors NF- κ B, TLRs, PPARs, and the liver X receptor-retinoid X receptor (LXR-RXR) also decreased. Consistent with these results, Rudkowska et al. evaluated the effect of fish oils on the PBMC transcriptomic profile in overweight insulin resistant patients and overweight normolipidemic subjects. Here, gene pathway analysis showed that genes connected with NF (erythroid-derived 2)-like (NFE2L)2, PPAR α , hypoxia-inducible factors, and NF- κ B pathways were affected by fish oil supplementation in both male and female subjects, and this was accompanied by a downregulation of COX-2, IL-1 β , and TNF- α . Correspondently, a reduced expression of cytokine-related genes was also observed by Gorjao and by Myhrstad in lymphocytes and PBMCs of healthy adults exposed to omega-3 PUFAs for 2 months.

TABLE 62.2 Whole-blood Gene Expression Profiling Studies Examining the Effects of Nutritional Fatty Acids in Humans

First author, DOI or PMID	Design	Health Status	Supplement	Placebo	Analyzed Tissue	Enriched Pathways or Annotated Functions of Cardiovascular Interest	Selected or Validated Differentially Expressed Genes	Main Results
van Dijk, 10.3945/ajcn.2009.27792	Intervention, two-arm, parallel	Overweight and obese subjects	Saturated fatty acids (19%,2% of daily energy); monounsaturated fatty acids (20%,3% of daily energy); 8 weeks	No placebo	Adipose	T- and B-cell receptor signaling; leukocyte extravasation signaling; natural killer cell signaling; TLR signaling; Il-6-NF-κB, IL-1, and p38 signaling	Saturated fatty acids: ↑CCL5/RANTES; ↑ CD14; CD163; ↑ CTSS; ↑ C1QB; ↑ ITGB2; ↑ RAC2; ↓ ADIPOQ; ↓ PPARγ. Monounsaturated fatty acids: ↑CCL5/RANTES; ↓CD14; ↓CD163; ↓CTSS; ↓C1QB; ↓TGB2; ↓RAC2; ↑ADIPOQ; ↓PFARg.	SFA-rich diet induced a pro-inflammatory gene expression profile in adipose tissue; consumption of MUFA-rich diet resulted in activation of antiinflammatory gene profile.
Kabir, 18065585	RCT, two-arm, parallel	Overweight type 2 diabetic women	EPA/DHA, 2.8 g; 8 weeks	Paraffin oil	Adipose	Not evaluated	*↓ADN; ↓ATP5F1; ↓ATP6V1E1; ↓CEBP; ↓CES2; ↓COX-1; ↓JKTBP; ↓LIPL; ↓MMP9; ↓RPLP1; ↓PRDX3; ↑CAV1; ↑GCK; ↑E2IG1; ↑LRPAP1; ↑SLC25A1; ↑MAP2K2; ↑POM121; ↑PCCB; ↑PKM2; ↑RAE1; ↑STAT6; ↑STX4A; ↑TM4SF2	n-3 PUFAs supplementation reduced adiposity, atherogenic markers, and adipose tissue inflammation.
Bouwens, 10.3945/ajcn.2009.27680	RCT, double blinded, three-arm, parallel	Overweight elderly subjects	EPA/DHA, 1.8 g; EPA/DHA, 0.4 g; 26 weeks	High-oleic sunflower oil	PBMC	Decrease in inflammatory signaling pathways including eicosanoid metabolism, IL-6, MAPK signaling, NF-κB and TLR signaling; oxidative stress, cell adhesion, PPAR signaling, LXR/RXR, and hypoxia initiation	↓HBEGF; ↓PDK4; ↓ZEB2; ↓CXCL16; ↓BACH1; ↓MITF; ↓MYADM; ↓CD38; ↓MARVELD; ↓TNFRSF8; ↓DPYSL2; ↓HIF1A; ↓VEGFA;	n-3 PUFA supplementation resulted in antiinflammatory and antiatherogenic changes in PBM C gene expression profiles of elderly subjects.

Continued

TABLE 62.2 Whole-blood Gene Expression Profiling Studies Examining the Effects of Nutritional Fatty Acids in Humans—cont'd

First author, DOI or PMID	Design	Health Status	Supplement	Placebo	Analyzed Tissue	Enriched Pathways or Annotated Functions of Cardiovascular Interest	Selected or Validated Differentially Expressed Genes	Main Results
Rudkowska, 10.1159/000330226	RCT, two-arm, repeated measure	Overweight insulin resistant subjects	EPA/DHA, 1.8 g; EPA/DHA, 1.8 g plus fish gelatin (25% daily intake), 8 weeks	No placebo	PBMC	factor 1 (HIF) ^a signaling. Oxidative stress response mediated by NF (erythroid-derived 2)-like 2 (Nrf2); mechanism of gene regulation by PPAR α , HIF signaling, NF- κ B signaling pathway, and oxidative stress	The kind of regulation (up or down) is not clearly specified.	Both n-3 PUFAs and fish gelatin improved inflammatory and metabolic profile of overweight and insulin-resistant subjects.
Gorjao, 10.1016/j.clnu.2006.03.004	Pre-post intervention study, one-arm, repeated measure	Healthy subjects	EPA/DHA, 2.4 g; 8 weeks	No placebo	Lymphocytes	Signaling pathway including oncogenes and tumor suppressor, nuclear receptor, intercellular transducers, transcription activator, and suppressor	↓ FGR; ↓ CSF1R; ↓ v-src; ↓ Ret; ↓ SHB; ↓ SMAD; ↓ MYBL2; ↓ CDKN1C; ↓ DST; ↓ CDC25B; ↓ CDC20; ↓ NOP2; ↓ EFNA4; ↓ MARCKSL1; ↓ PRKACA; ↓ PRKCD; ↓ MAPK8; ↓ MAP2K1; ↓ CCR2; ↓ CSF2RA	n-3 PUFAs improved gene expression profiles connected to cellular defense and suppressed expression of proapoptotic genes in healthy men.
Myhrstad, 10.1111/joim.12217	RCT, double-blinded, two-arm, parallel	Healthy subjects	DHA/EPA, 1.6 g; 7 weeks	High-oleic sunflower oil	PBMC	Cell cycle, endoplasmic reticulum stress and apoptosis pathways	↑ CDK2; ↑ CCNE1; ↑ MLL5; ↑ TUBB; ↑ BUB1; ↑ MAPRE2; ↑ H2AFX; ↑ GFI1; ↑ HELLS; ↑ UBA52; ↑ H2AFV; ↑ NAP1L3; ↑ NCAPH2; ↑ ASF1B; ↑ IRF4; ↑ PRMT7; ↑ TERF2IP; ↑ JAG2; ↑ MCM3; ↑ INO80; ↑ KIFC1; MND1; UBE2C; DDIT3;	n-3 PUFAs improved gene expressions involved in cellular homeostasis in healthy subjects.
Schmidt, 10.1186/1476-511X-11-105	RCT, double-blinded, four-arm, parallel	Normolipidemic and dyslipidemic subjects	EPA/DHA, 2.7 g; 12 weeks	LA, 3.05 g	Whole blood	Inflammation, immune response, and cardiovascular disease	↓ NOS2; ↓ CCR3; ↓ TNFRSF18; ↓ CD40; ↓ IL8; ↓ MMP-2; ↓ MMP-3; ↑ PTGDR; ↑ ALOX5AP; ↓ IL-31RA; ↓ IL2RB; ↓ IL3RA; ↑ CCL5; ↓ SOCS2; ↓ ADAM 10; ↓ ADAMTS10; ↓ NOS3; ↓ PTGER3; ↑ CD99L2; ↓ CRLF1	n-3 PUFAs downregulated proinflammatory gene expressions in dyslipidemic subjects.

Schmidt, 10.1186/1476- 511X-11-172	RCT, double- blinded, four-arm, parallel	Normolipidemic and dyslipidemic subjects	EPA/DHA, 2.7 g; 12 weeks	LA, 3.05 g	Whole blood	Not evaluated	↑RXRG; ↓HNF6; ↓HNF1; ↓MOGAT3; ↑PLTP; ↑ABCG5; ↓SOAT1	n-3 PUFAs regulated lipid-associated transcriptional factors and genes in dyslipidemic subjects.
Tsunoda, 10.1016/ j.atherosclerosis. 2015.05.015	RCT, double- blinded, three- arm, parallel	Subjects with mild elevation in Lp-PLA ₂	EPA (1.8 g) + olive oil (3 g); DHA (1.8 g); olive oil (6 g); 6 weeks	Not placebo controlled	PBMC	DHA group: immune response and metabolism; EPA + olive oil group: cellular immune response, intracellular and second messenger signaling; metabolic pathways	DHA group: ↓ HIF1A; ↓ NT5M; ↑ DNAJA1; ↑ PSMC6; ↓ PTK2; ↑ TICAM1; ↑ SLAMF7; ↓ OAZ2; EPA + olive oil group: ↑ OAS1; ↑ MX1; ↑ IFIT3; ↑ IFIT1; ↑ OAS3; ↑ IRF7; ↑ OAS2; ↓ CREB1; ↓ HIF1A; ↓ MTPN; ↓ HMGB1	EPA and DHA improve common inflammatory and metabolic pathways. However, they regulate expression of distinct genes.

ABCG1, ATP-binding cassette, subfamily G (WHITE), member 1; *ABCG5*, ATP-binding cassette subfamily G member 5; *ACOT13*, acyl-coenzyme A thioesterase 13 (THEM2); *ADAM10*, disintegrin and metalloproteinase domain-containing protein 10 precursor; *ADAMTS10*, disintegrin and metalloproteinase with thrombospondin motifs 10 precursor; *ADCY*, adenylate cyclase; *ADIPOQ*, adiponectin; *ALOX5AP*, arachidonate 5-lipoxygenase-activating protein (FLAP); *ASCC2*, activating signal cointegrator 1 complex subunit 2 (p100); *ATP5F1*, H+ transporting mitochondrial F1; *ATP6V1E1*, H+ transporting lysosomal; *C12orf43*, chromosome 12 open reading frame; *CIQB*, complement component 1q subcomponent B chain; *CACNB1*, calcium channel, voltage-dependent, β 1 subunit; *CASP4*, apoptosis-related cysteine peptidase SUPT4H1, suppressor of Ty 4 homolog 1; *CAV1*, caveolin 1; *CCDC78*, coiled-coil domain containing; *CCL5*, C-C motif chemokine 5 precursor; *CCL5/RANTES*, regulated on activation normal T cell-expressed and secreted/chemokine CC motif ligand 5; *CCR2*, C-C, motif chemokine receptor 2; *CCR3*, C-C chemokine receptor type 3; *CD14*, cluster of differentiation 14; *CD40*, tumor necrosis factor receptor superfamily member 5 precursor; *CD99L2*, CD99 antigen-like protein 2 precursor; *CDC20*, cell division cycle 20 homolog 2.1; *CDC25B*, cell division cycle 25B; *CDKN1C*, cyclin-dependent kinase inhibitor 1C (p57); *CEBPB*, enhancer binding protein; *CES2*, carboxyl esterase 2; *COX*, cyclooxygenase 1; *CREB1*, cAMP-responsive element binding protein; *CRLF1*, cytokine receptor-like factor 1 precursor; *CSF1R*, colony stimulating factor 1 receptor; *CSF2RA*, colony stimulating factor 2 receptor α subunit; *CTSS*, cathepsin S; *DHX34 DEAH*, (Asp-Glu-Ala-His) box polypeptide; *DNAJA1*, DnaJ (Hsp40) homolog, subfamily A, member 1; *DRG1*, developmentally regulated GTP binding protein 1; *DST*, dystonin; *E2IG1*, heat shock protein 8; *EFNA4*, ephrin A4; *ERPINB9*, serpin peptidase inhibitor, clade B (ovalbumin), member 9 (P19); *FGR*, Gardner-Rasheed feline sarcoma viral oncogene homolog; *GCK*, glucokinase; *GGT*, γ -glutamyltransferase; *AND*, adipisin; *HGS*, hepatocyte growth factor-regulated tyrosine kinase substrate; *HIF*, hypoxia inducible factor; *HIF1A*, hypoxia inducible factor 1, α subunit; *HK2*, hexokinase 2; *HMGB1*, high mobility group box 1; *HNF6*, hepatocyte nuclear factor 6; *HSD17B11*, hydroxysteroid (17- β) dehydrogenase 11; *IFIT*, interferon-induced protein with tetratricopeptide repeats; *IL2RB*, interleukin-2 receptor subunit β precursor; *IL31RA*, interleukin-31 receptor A precursor; *IL-6*, interleukin-6; *IRF*, interferon regulatory factor; *IRF1*, interferon regulatory factor 1; *ISLR*, immunoglobulin superfamily containing leucine-rich repeat; *ITGB2*, integrin β 2; *RAC2*, ras-related C3 botulinum toxin substrate 2; *JKTB*, heterogeneous nuclear ribonucleoprotein D-like; *LIPL*, lipoprotein lipase; *Lp-PLA₂*, lipoprotein-associated phospholipase A2; *LRPAP1*, LDL receptor-related protein associated protein 1; *LXR/RXR*, liver X receptor/retinoid X receptor; *MAP2K1*, mitogen-activated protein kinase kinase 1; *MAP2K2*, mitogen-activated protein kinase 2; *MAPK8*, mitogen-activated protein kinase 8; *MARCKSL1*, macrophage myristoylated alanine-rich C kinase; *MMP9*, matrix metalloproteinase 9/gelatinase B; *MOGAT3*, 2-acylglycerol O-acyltransferase 3; *MRPL55*, mitochondrial ribosomal protein L55; *MS4A3*, membrane-spanning-4 domains subfamily A; *MSMO1*, methylsterol monooxygenase 1 (SC4MOL); *MTPN*, myotrophin; *MX1*, MX dynamin-like GTPase 1; *MYBL2*, MYB proto-oncogene-like 2; *NAIP*, NLR family apoptosis inhibitory protein; *NELF*, nasal embryonic luteinizing hormone-releasing hormone factor; *NOP2*, NOP2 nucleolar protein; *NOS2*, inducible nitric oxide synthase; *NT5M*, 5',3'-nucleotidase, mitochondrial; *OAS1*, oligoadenylate synthetase 1; *OAZ2*, ornithine decarboxylase antienzyme 2; *OSBP2*, oxysterol-binding protein 2; *PBMC*, peripheral blood mononuclear cell; *PCCB*, propionyl CoA carboxylase; *PKM2*, pyruvate kinase; *PLA2G6*, phospholipase A2, group VI (cytosolic, calcium-independent); *PLTP*, phospholipid transfer protein; *POM121*, POM121 membrane glycoprotein; *PRAC*, prostate cancer susceptibility candidate (C17orf92); *PRDX3*, peroxiredoxin 3; *PRKACA*, protein kinase cAMP-activated catalytic subunit α ; *PRKCD*, protein kinase C δ ; *PSMC6*, proteasome (prosome, macropain) 26S subunit, ATPase, 6; *PTGDR*, prostaglandin D2 receptor; *PTGER3*, prostaglandin E2 receptor EP3 subtype; *PTGS2*, prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase); *PTK2*, protein tyrosine kinase 2; *RAB27A*, member RAS oncogen family 1; *RAE1*, RNA export 1 homolog; *Ret*, proto-oncogene (multiple endocrine neoplasia and medullary thyroid carcinoma 1, Hirschsprung disease); *RPLP*, ribosomal phosphoprotein; *RXRG*, retinoic acid receptor RXR- γ ; *SCARF1*, scavenger receptor class F member 1; *SHB*, SH2 domain containing adaptor protein B; *SLAMF7*, SLAM family member 7; *SLC25A1*, oxoglutarate/malate carrier; *SLC30A5*, solute carrier family 30 (zinc transporter), member 5; *SMAD*, SMAD family member 1; *SNAP23*, synaptosomal-associated protein, 23 kD; *SOAT1*, sterol O-acyltransferase 1; *Socs2*, suppressor of cytokine signaling 2; *STAT6*, signal transducer and activator of transcription 6; *STX4A*, syntaxin 4A; *TICAM1*, Toll-like receptor adaptor molecule 1; *TLR*, Toll-like receptor; *TM4SF2*, Transmembrane 4 superfamily member t3; *TNFRSF18*, Tumor necrosis factor receptor superfamily member 18 precursor; *UBE2V1*, ubiquitin-conjugating enzyme E2 variant 1; *USP5*, ubiquitin specific protease 5 (isopeptidase T); *v-src*, sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog.

Up to a maximum of 20 genes were reported. For the other genes, see original references.

*Results represent the change between samples after 2 months of treatment and those at baseline in the same fish oil group.

Nutrigenomic studies in dyslipidemic patients were also performed. Schmidt et al. performed a four-arm, randomized, controlled study to evaluate the nutrigenomic effects of DHA plus eicosapentaenoic acid (EPA) versus linoleic acid (an omega-6 PUFA) in normolipidemic and dyslipidemic subjects. In this setting, fish oil supplementation downregulated the expression of several proinflammatory genes that were activated in dyslipidemic subjects, including IL-8, metalloproteinase (MMP)-2 and MMP-3, and CD40. Also, the expression of PPAR α and RXR and the expression of genes related to triglyceride and cholesterol metabolism, including 2-acylglycerol O-acyltransferase (MOGAT)2, MOGAT3, diacylglycerol O-acyltransferase 1, and sterol O-acyltransferase 1 were downregulated, especially in men. Similarly, in a 3-arm randomized design, Tsunoda et al. compared the effects of DHA, EPA, and olive oil on the transcriptomic profile of PBMCs in mildly dyslipidemic patients. The results demonstrated that besides affecting the expression of several unsuspected genes, including proteasome 26S subunit ATPase 6, protein tyrosine kinase 2, and Toll-like receptor adaptor molecule 1, both EPA and DHA quenched inflammatory and metabolic pathways, in agreement with all the earlier reported results.

Although these studies are heterogeneous as to subjects' health status, treatment duration, and therapeutic doses, they collectively show that the health effects of fish oils invariably associate with a favorable modulation of classical and new biological pathways involved

in regulating inflammation, lipid metabolism, and antioxidant enzyme systems. However, they are also associated with previously unsuspected genes and transcriptional factors. Overall, they provide the first qualitative and quantitative evaluations of how networks of known and emerging inflammatory and metabolic pathways are favorably affected in a remarkably complex fashion by unsaturated FAs from available foods, providing protection from atherosclerosis and its consequences.

POLYPHENOLS AND GENE EXPRESSION IN CARDIOVASCULAR DISEASE

The health benefits of vegetable-based food in CVD have been attributed not only to their low energy content and low dietary glycemic load associated with their intake but also, and perhaps predominantly, to the high micronutrient content, particularly of polyphenols. Polyphenols are ubiquitously present in vegetables as secondary metabolites and represent the most relevant source of phytochemicals supplied by the diet. To date, no less than 8000 polyphenols have been identified and classified into different subclasses according to their structure and broadly categorized into flavonoids and nonflavonoids (Fig. 62.1). Beside fruits and vegetables, the most important nutritional sources of polyphenols include red wine, extravirgin olive oil, green and black tea, coffee, cocoa, herbs, spices, and nuts. Although the

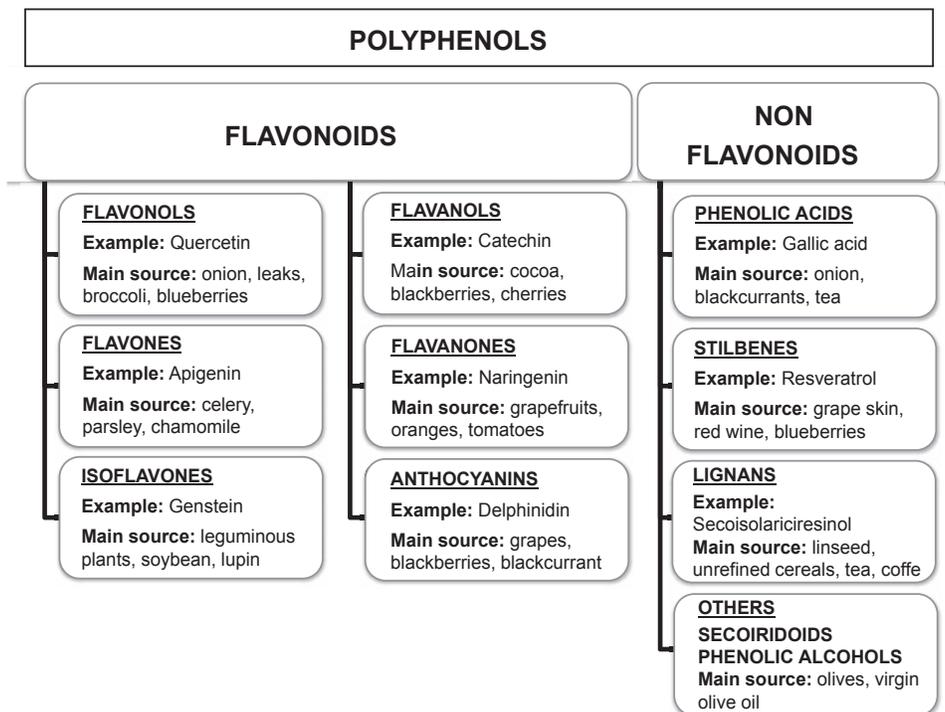


FIGURE 62.1 Classification of polyphenols.

daily polyphenol assumption is estimated to be high (1500 mg gallic acid equivalents per person), the bioavailability of polyphenols is generally low; polyphenols are rapidly absorbed, metabolized, and excreted. This is because the forms present in circulating blood are the product of extensive metabolic transformation catalyzed by endogenous enzymes and enzymes present in gut microbiota, so that the native compounds are converted into conjugated compounds, such as glucuronidated, sulfated, and methylated derivatives. Despite their low bioavailability, evidence is rapidly accumulating that beneficial activities of polyphenols occur in humans. This apparent paradox may be explained by the potential synergistic action of the different polyphenol metabolites produced endogenously and/or by the combination of polyphenol activities with those of other (micro)nutrients. Historically, biologic actions of polyphenols have been attributed to their antioxidant properties, because they are free radical scavengers, but evidence suggests that antiinflammatory, immunomodulatory, and vasodilatory properties of polyphenols may also contribute to CVD risk reduction. From a mechanistic point of view, the healthful properties of polyphenols occur mainly through their nutrigenomic effects, because they affect the expression of numerous genes involved in chronic inflammation and in the development and progression of CVD. Such nutrigenomic effects, as demonstrated by applying targeted and nontargeted gene expression profiling in cell models relevant to CVDs and in humans, are briefly summarized next.

Flavonols

Flavonols (Figs. 62.1 and 62.2) are the flavonoid subclass most commonly represented in our diet. The main food sources are onions, spinach, asparagus, and some types of berries. Quercetin and its conjugates are the most widely consumed type of flavonols (in amounts of about 22 mg/day) in Europe. Several epidemiological studies linked flavonol consumption with a low prevalence of cardiovascular risk factors, such as hypertension, and a low incidence of CVD. Correspondingly, human intervention studies showed that quercetin supplementation lowers blood pressure in hypertensive and type 2 diabetic patients, as well as in smokers and in overweight or obese subjects. Furthermore, by reducing plasma levels of oxidized LDL and of the inflammatory cytokines IL-6 and TNF- α , quercetin also improves blood biomarkers of vascular health.

Mechanisms accounting for these protective effects are incompletely understood; multiple modulations of cell signaling and gene expression are the most widely documented and probable. Growing evidence attributes healthful effects of quercetin to the improvement of

endothelial function, which is mediated by its ability to modulate the expression of a series of vascular genes. Such effects include a reduction in endothelin (ET)-1 gene expression and an increase in endothelial nitric oxide synthase (eNOS) involved in the production of the vasodilator, antiaggregating, and antiinflammatory nitric oxide, thus improving the overall balance between vasoconstrictive and vasodilating factors. In addition, *in vitro* studies showed that flavonols inhibit the early development of atherosclerosis by decreasing the adhesion of monocytes to endothelial monolayers through a significant reduction of the expression of several ELAMs, including E-selectin, VCAM-1, and ICAM-1, as well as the expression and release of MCP-1 and M-CSF from activated endothelial cells (EFSA Panel on Dietetic Products NaAN, 2011). Such effects appear to be mediated by a reduced activation of NF- κ B and activator protein (AP)-1, two pivotal regulators of the expression for atherogenic and inflammatory genes. Furthermore, in human vascular endothelial cells, quercetin reduced inflammatory angiogenesis, a key pathogenic process contributing to atherosclerotic lesion progression and vulnerability, by inhibiting MMP-9 and COX-2 gene expression and activity (Scoditti et al., 2012). Against the background of this large mass of interesting results obtained applying candidate gene approaches, looking for genes or gene product candidates for exerting a modulating function in atherosclerosis, only one study exploited nontargeted approaches to unveil unsuspected molecular targets of flavonols. Here, the *in vivo* effect of quercetin on the transcriptomic profile of PBMCs from healthy subjects supplemented with 150 mg/day quercetin for 2 weeks was evaluated by applying a microarray technique (Boomgaarden et al., 2010). Data analysis showed that quercetin supplementation profoundly modified the transcriptomic profile of PBMCs, altering the expression of 788 genes, 503 of which were upregulated and 285 downregulated. With a significant modulation of core one galactosyltransferase 1, GM2 ganglioside activator, heparin binding growth factor, and Serpin family B member 9 genes, quercetin affected glycolipid catabolism, cell proliferation, and apoptotic processes. Overall, despite incomplete knowledge of the mechanisms of action, a substantial body of evidence suggests that quercetin and related flavonols may effectively contribute to prevent or reduce CVD risk by modulating the expression of CVD-related genes.

Stilbenes

Stilbenes belong to the class of nonflavonoid polyphenols. Resveratrol (3,5,4'-trihydroxy-trans-stilbene), typically present in grapes and red wines, is the primary representative of this class of compounds (Figs. 62.1

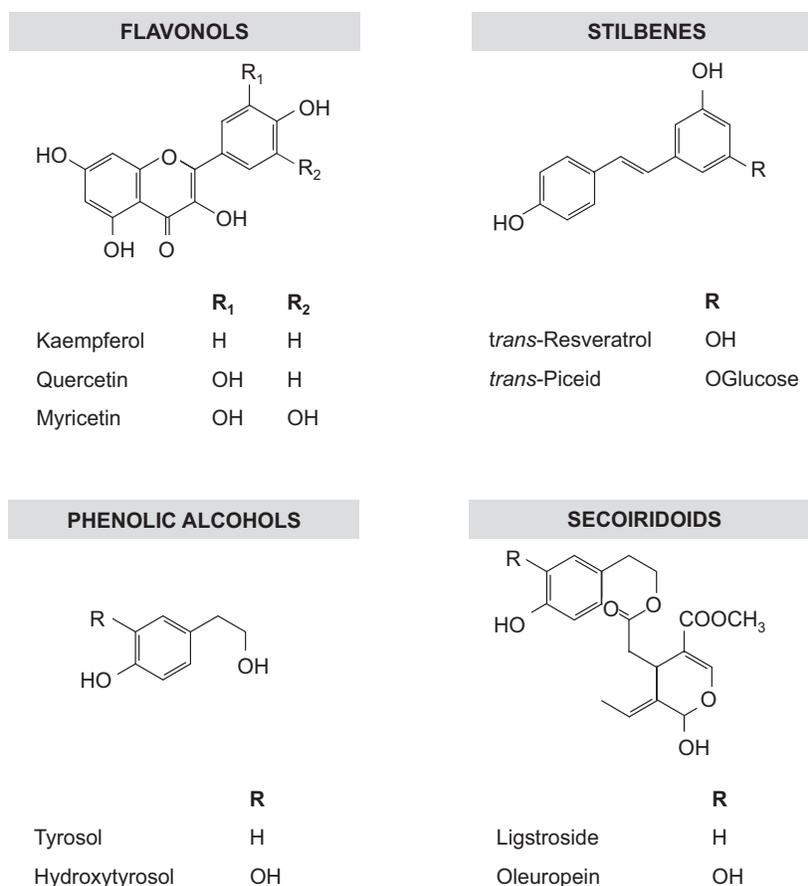


FIGURE 62.2 Chemical structures of main nutritional polyphenols.

and 62.2). The presence of resveratrol, along with other polyphenolic compounds, including quercetin, in red wine has garnered much attention in France, especially with regard to the potential role of these compounds in protecting against CVD. Indeed, red wine is thought to be mainly responsible for the French Paradox, a term used to describe the low incidence of CVD among the French despite their high intake of saturated fats. Many preclinical studies have identified a number of mechanisms and targets by which resveratrol could exert benefits against CVD. By applying candidate gene approach techniques, the ability of resveratrol to modulate the expression of several inflammation-related genes involved in the atherosclerotic process has been shown. In vitro studies showed that resveratrol decreased monocyte cell adhesion to human endothelial cells by reducing VCAM-1, ICAM-1, and E-selectin, as well as the proinflammatory mediators MCP-1, M-CSF, and IL-6 in both endothelial cells and monocytes. In human endothelial cells, monocytes, and smooth muscle cells (SMCs), resveratrol strongly inhibited the expression of MMPs responsible for the degradation of extracellular matrix, a critical event in plaque rupture. Moreover, in endothelial and mononuclear cells,

resveratrol decreased the stimulated expression of tissue factor, a key regulator in the extrinsic pathway of blood coagulation (Pendurthi et al., 1999). These antiinflammatory and antithrombotic effects of resveratrol were at least partly mediated by lowered levels of intracellular ROS and reduced activation of NF- κ B and AP-1. Furthermore, resveratrol was reported to decrease the expression of NFE2L2, IL-8, ET1, eNOS, and sirtuin 1 genes in endothelial cells. The corresponding gene products are involved in the regulation of cell proliferation, blood pressure, oxidative stress, and vascular tone. Resveratrol in a blend of polyphenols from grape extracts exhibited greater inhibitory effects on the expression of inflammatory markers in vascular cells than resveratrol alone, which confirmed the occurrence of a cooperative action or synergism among polyphenols (Calabriso et al., 2016). Although the potential utility of resveratrol in preventive medicine has been demonstrated using in vitro and animal models, only few clinical trials have evaluated the effects of resveratrol on the expression of genes clinically relevant to CVD. In a trial evaluating the effects of 1-month resveratrol supplementation on endothelial physiology and inflammatory plasma biomarkers in healthy subjects, although

resveratrol did not affect circulating biomarkers, the exposure of cultured endothelial cells to plasma drawn after the resveratrol supplementation resulted in significantly lower mRNA expression of VCAM-1, ICAM-1, and IL-8 than plasma drawn from the same subjects at baseline. These data indicate that resveratrol may have protective effects against atherosclerosis in healthy individuals and suggests a potential role in primary prevention. In a secondary prevention trial, a 1-year dietary intervention with resveratrol-rich grape extracts (GE-RES) affected the inflammatory and fibrinolytic status of patients with stable coronary artery disease (CAD) compared with a resveratrol-free grape extracts. Furthermore, GE-RES increased adiponectin levels while decreasing plasminogen activator inhibitor levels. In PBMCs from these patients, GE-RES modulated the transcriptional profile of inflammatory genes and the activity of several inflammatory transcription factors, including Kruppel-like factor 2, AP-1, C-Jun, activating transcription factor-2, and cAMP response element-binding protein. Functional analysis showed a significant inhibition of signaling pathways related to atherothrombotic signals, including inflammation, cell migration, and T-cell interaction. Finally, in a subgroup of type 2 diabetic and hypertensive patients with CAD, 1-year supplementation with GE-RES modulated several inflammatory-related microRNAs (miR), including miR-21, miR-181b, miR-663, miR-30c2, miR-155, and miR-34a (Tome-Carneiro et al., 2013). These results provide evidence that resveratrol may exert cardioprotective actions improving vascular function by affecting biological pathways involved in regulating atherothrombotic signals.

Phenolic Alcohol and Secoiridoids

Phenolic alcohols and secoiridoids are nonflavonoid polyphenols typical of Mediterranean foods; they are abundant in olives and extravirgin olive oil. These foods are particularly rich in the phenolic alcohols hydroxytyrosol (HT) (3,4-dihydroxyphenylethanol) and tyrosol (4-hydroxyphenylethanol), as well as in the secoiridoids oleuropein and ligstroside (the ester of elenolic acid with HT or with tyrosol, respectively) (Figs. 62.1 and 62.2). A great number of studies have been devoted to characterizing the effects of olive oil polyphenols on atherothrombotic risk factors or mechanistic pathways, including plasma lipid profile, oxidative stress, endothelial dysfunction, hypertension, platelet aggregation, diabetes, and inflammation. After the publication of the *Study on the Effect of Olive Oil on Oxidative Damage in European Populations* (EUROLIVE) results, the European Food Safety Authority released a claim supporting the benefits derived from the daily ingestion of 5 mg olive

oil polyphenols (EFSA Panel on Dietetic Products NaAN, 2011). Current research indicates that olive oil polyphenols reduce markers of oxidative stress, inflammation, and atherosclerotic risk (including oxidized LDL, diabetes, and hypertension). In a randomized crossover trial, a postprandial gene expression microarray analysis on PBMCs from patients with metabolic syndrome revealed that a breakfast based on virgin olive oil, which is high in polyphenols, repressed the expression of proinflammatory genes linked to atherosclerosis, obesity, dyslipidemia, and type 2 diabetes mellitus compared with a common olive oil-based breakfast low in polyphenols. Among these, several genes pathways involved were related to cytokines and other proteins controlled by NF- κ B and AP-1, including several cytokines, mitogen-activated protein kinases (MAPKs), as well as eicosanoids. In addition, exploratory microarray reports indicated that acute and chronic intake of virgin olive oil induced changes in the expression of genes related to insulin resistance, oxidative stress, and inflammation in healthy volunteers. In a transcriptomic study on PBMCs of healthy volunteers, sustained (3-month) consumption of olive oil rich in polyphenols induced favorable changes in the expression profile of genes involved in atherosclerosis, inflammation, and oxidative stress. In a subgroup of the EUROLIVE study, after sustained consumption of olive oil with high versus low polyphenol content, PBMC transcriptomic profiling showed a significant decrease in the expression of proatherogenic genes related to the CD40/CD40L cascade, in association with a reduction in LDL oxidation and an increase in urine antioxidant polyphenols. In a transcriptomic substudy of the *Prevencion Con Dieta Mediterranea* randomized controlled trial in PBMCs of individuals at high cardiovascular risk, a Mediterranean diet supplemented with extravirgin olive oil, compared with a low-fat diet, modified several disease pathways in a cardiovascular protective direction. These included pathways involved in atherosclerosis, the renin-angiotensin system, nitric oxide, angiopoietin, hypoxia, and eNOS signaling. Downregulated genes included IL-1 β , IL-1 receptor antagonist, TNF- α , ICAM-1, vascular endothelial growth factor, NF- κ B, and c-jun. These results highlight the in vivo nutrigenomic protective effects of olive oil polyphenols in humans within the frame of the Mediterranean diet.

In agreement with results from human intervention trials, in vitro studies provide mechanisms for the nutrigenomic effects of olive oil polyphenols in vascular protection. Olive oil polyphenols HT and oleuropein, at nutritionally relevant concentrations, downregulated the gene expression of adhesion molecules, chemoattractants, MMPs, and proinflammatory enzymes, through mechanisms including the downregulation of the NF- κ B subunit p65, the upregulation of the NF- κ B

inhibitor $\text{I}\kappa\text{B-}\alpha$, as well as the reduction of intracellular oxidative stress involved in triggering the MAPK–NF- κB signaling cascade. Other studies have shown that HT upregulates the gene expression of antioxidant/detoxifying enzymes by inducing the expression and nuclear translocation of the transcription factor NFE2L2. Furthermore, olive oil polyphenols have been shown to influence the expression of genes related to obesity. Indeed, HT and oleuropein inhibited the early stages of adipocyte differentiation by suppressing the expression of adipogenesis-related genes. HT also prevented the deleterious downregulation of gene and protein expression of adiponectin in human adipocytes in inflammatory conditions (Scoditti et al., 2015).

CONCLUSIONS

The pathogenesis of atherosclerosis features a complex interplay of genetic and environmental factors, among which are specific nutrient intakes and nutrient combinations in diets. A large number of in vitro and in vivo studies have consistently found that specific unsaturated FAs and polyphenols affect the expression levels of genes related to the cardiovascular systems in a protective fashion, thus contributing to explaining benefits associated with diets and foods rich in these components. The possibility of reducing major cardiovascular events by administering a monoclonal antibody against IL-1 β in a trial (CANTOS) of secondary CVD prevention (Ridker et al., 2017) shows that vascular inflammation is a true viable target to reduce the burden of CVD. The greatest challenge is now to interfere with vascular inflammatory pathways physiologically, exploiting natural strategies without appreciable side effects. These have the potential to tackle the roots of vascular disease before the occurrence of its clinical manifestations (i.e., in primary cardiovascular prevention). The road ahead is indicated by the scientific evidence reviewed earlier, linking the intake of natural compounds, through mechanisms being unraveled in molecular detail, with favorable outcomes.

The use of global transcriptional profiling is a powerful tool in current nutrigenomic studies. This allows dynamic changes in gene expression to be measured before and after adding a nutrient or a nutrient metabolite in cultured endothelial cells, monocytes and macrophages, and SMCs, all of which are relevant for atherosclerosis. This highlights novel nutrient-responsive genes and detailing associated plausible nodes of signaling pathways. To date, most in vivo human nutrigenomic studies have used PBMCs as a feasible cell type for gene expression studies. However, nutrigenomic effects can be tissue-dependent and gene expression measurements in human tissues other than

PBMCs, such as adipose tissue cells, have been increasingly considered.

An important challenge for cardiovascular nutrigenomic studies in humans is to correlate tissue-specific transcriptomic signatures with proteomic, metabolomic, and microbiome data and with systemic effects. Integration of different “omics” techniques and the classical biomarkers with nutritional information using a systems biology approach is instrumental to achieving a comprehensive view of the biological response to nutrients to be exploited for preventing and treating CVD. Furthermore, more attention is warranted to examine the variability in cell or tissue response to such nutrients, examining gene expression profiles in responders versus nonresponders, because vascular healthful effects of specific nutrients and nutrient combinations may depend on inherited genetic variants affecting their uptake, metabolism, and responses. Probing this complexity will lead to a more global understanding of the influence of environmental and genetic factors conditioning the specific response to nutrients.

Acknowledgments

We would like to thank Mr. Massimo Scoditti for valuable support in the audio registration of this section.

References

- Boomgaarden, I., Egert, S., Rimbach, G., Wolfram, S., Muller, M.J., Doring, F., 2010. Quercetin supplementation and its effect on human monocyte gene expression profiles in vivo. *Br J Nutr* 104, 336–345.
- Calabriso, N., Massaro, M., Scoditti, E., et al., 2016. Red grape skin polyphenols blunt matrix metalloproteinase-2 and-9 activity and expression in cell models of vascular inflammation: protective role in degenerative and inflammatory diseases. *Molecules* 21.
- De Caterina, R., 2011. n-3 fatty acids in cardiovascular disease. *N Engl J Med* 364, 2439–2450.
- de Mello, V.D.F., Kolehmanien, M., Schwab, U., Pulkkinen, L., Uusitupa, M., 2012. Gene expression of peripheral blood mononuclear cells as a tool in dietary intervention studies: what do we know so far? *Mol Nutr Food Res* 56, 1160–1172.
- EFSA Panel on Dietetic Products NaAN, 2011. Polyphenols in olive related health claims. *EFSA Journal* 9, 2033–2048.
- Massaro, M., Martinelli, R., Gatta, V., et al., 2015. Transcriptome-based identification of new anti-inflammatory and vasodilating properties of the n-3 fatty acid docosahexaenoic acid in vascular endothelial cell under proinflammatory conditions. *PLoS One* 10, e0129652, 2016;11.
- Merched, A.J., Chan, L., 2013. Nutrigenetics and nutrigenomics of atherosclerosis. *Curr Atheroscler Rep* 15.
- Pendurthi, U.R., Williams, J.T., Rao, L.V.M., 1999. Resveratrol, a polyphenolic compound found in wine, inhibits tissue factor expression in vascular cells - a possible mechanism for the cardiovascular benefits associated with moderate consumption of wine. *Arterioscler Thromb Vasc Biol* 19, 419–426.
- Ridker, P.M., Everett, B.M., Thuren, T., et al., 2017. Antiinflammatory therapy with canakinumab for atherosclerotic disease. *N Engl J Med* 377, 1119–1131.

- Sacks, F.M., Lichtenstein, A.H., Wu, J.H.Y., et al., 2017. Dietary fats and cardiovascular disease: a presidential advisory from the American heart association. *Circulation* 136, e1–e23.
- Scoditti, E., Calabriso, N., Massaro, M., et al., 2012. Mediterranean diet polyphenols reduce inflammatory angiogenesis through MMP-9 and COX-2 inhibition in human vascular endothelial cells: a potentially protective mechanism in atherosclerotic vascular disease and cancer. *Arch Biochem Biophys* 527, 81–89.
- Scoditti, E., Massaro, M., Carluccio, M.A., et al., 2015. Additive regulation of adiponectin expression by the Mediterranean diet olive oil components oleic acid and hydroxytyrosol in human adipocytes. *PLoS One* 10.
- Seok, J., Warren, H.S., Cuenca, A.G., et al., 2013. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc Natl Acad Sci U S A* 110, 3507–3512.
- Siscovick, D.S., Barringer, T.A., Fretts, A.M., et al., 2017. Omega-3 polyunsaturated fatty acid (fish oil) supplementation and the prevention of clinical cardiovascular disease: a science advisory from the American heart association. *Circulation* 135, e867–e884.
- Tome-Carneiro, J., Larrosa, M., Gonzalez-Sarrias, A., Tomas-Barberan, F.A., Garcia-Conesa, M.T., Espin, J.C., 2013. Resveratrol and clinical trials: the crossroad from in vitro studies to human evidence. *Curr Pharmaceut Des* 19, 6064–6093.

This page intentionally left blank

Nutrients and Gene Expression in Cancer

Ana B. Crujeiras^{1,2}, David de Lorenzo^{3,4}

¹Epigenomics in Endocrinology and Nutrition Group, Instituto de Investigacion Sanitaria (IDIS), Complejo Hospitalario Universitario de Santiago (CHUS/SERGAS) and Santiago de Compostela University (USC), Santiago de Compostela, Spain; ²CIBER Fisiopatología de la Obesidad y Nutricion (CIBERObn), Madrid, Spain; ³Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, CEXS-UPF-PRBB, Barcelona, Catalonia, Spain; ⁴Centro de Estudios en Genómica y Nutrición-CESGEN, Parc Científic i Tecnològic Agroalimentari de Lleida-PCiTAL, Lleida, Catalonia, Spain

INTRODUCTION

There is a need for effective strategies for preventing cancer because the incidence of several cancers is rapidly increasing owing to aging populations in many countries and harmful dietary patterns.

Cancer is caused by a series of genetic changes in tumor suppressor genes and oncogenes. A small proportion of tumors can be promoted by acquired genetic modifications; however, most cancers result from changes that accumulate throughout life because of exposure to various factors such as nutrients, infections, physical inactivity, harmful social behaviors, and other environmental factors. In fact, it is currently estimated that 30%–40% of cancer types are preventable over time with appropriate food and nutrition and regular physical activity, and by avoiding obesity (Fig. 63.1).

The effect of these environmental factors on the silencing or expression of genes is mediated by epigenetic regulation. Epigenetics is defined as reversible heritable changes in gene expression that occur without alterations in DNA sequence. These changes need to be sufficiently powerful to regulate the dynamics of gene expression. Epigenetic mechanisms include DNA methylation, histone modifications, and gene regulation by noncoding RNAs. All of these mechanisms work synergistically and cooperatively to determine the expression of genes. The clinical relevance of these mechanisms is the reversibility of epigenetic alterations (i.e., epigenetically modified genes can be restored via drugs or dietary interventions). Therefore, targeting the epigenetic regulation of the expression of genes involved in cancer development and progression offers

an opportunity to design optimal chemopreventive and therapeutic anticancer strategies. Taking this fact into account, the potential to prevent disease and promote overall health has been explored and identified in several bioactive dietary components. Many of these compounds feature anticancer properties and may have a role in cancer prevention. They modulate gene expression by means of epigenetic mechanisms and have an impact on the initiation and progression of oncogenesis. This chapter provides an overview of the impact of dietary factors on the expression of genes involved in carcinogenesis pathways, with a special focus on the epigenetic regulation of these procarcinogenic genes mediated by dietary factors.

NUTRIGENOMICS AND NUTRIEPIGENETICS

Although the capacity of a dietary component to overcome a defective pathway involved in cancer susceptibility or progression has been clearly demonstrated, it is difficult to provide general dietary advice for all individuals because good dietary advice for one individual with susceptibility to the disease may not be most appropriate for others (Kohlmeier et al., 2016). For these reasons, a high-throughput screening of the regulatory process involved in the disease modified by a certain dietary component or pattern is important. This screening may provide new therapeutic avenues for treatment and/or prevention of cancer. In this sense, nutrigenomics and nutriepigenetics are two sciences that have the potential to provide information about

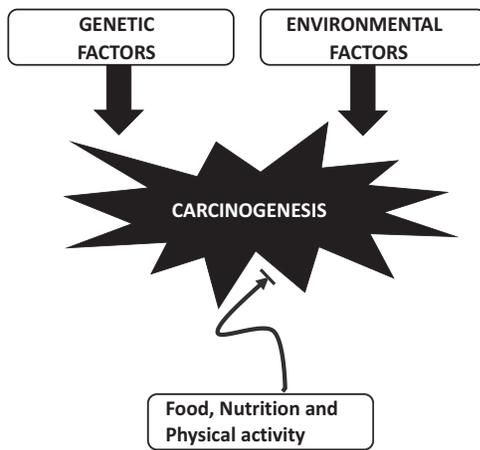


FIGURE 63.1 Factors related to the promotion or prevention of carcinogenesis. Cancer is caused by a series of genetic changes in tumor suppressor genes and oncogenes; however, a proportion of tumors can be promoted by environmental factors, such as nutrients, infections, toxics, physical inactivity, and excess body weight. Thus, it appears that 30%–40% of cancers are preventable with appropriate food and nutrition and regular physical activity, and by avoiding obesity.

mechanisms involved in the effect of nutrients, as well as on biomarkers for precision nutrition in preventing and managing cancer (Fig. 63.2).

Nutrigenomics is the science that studies the effect of nutrition on gene expression and, consequently, on the proteome and the metabolome; it also studies differences in response to dietary factors based on the individual

genetic makeup. Nutriepigenetics explores the chromatin structure and DNA modifications with nutrients that do not alter the underlying DNA sequence, but affect gene expression (Kohlmeier et al., 2016). Unlike genetic mutations, changes in gene expression owing to epigenetic regulation during carcinogenesis can be reversed or prevented by chemicals. Therefore, the pharmacological targeting of epigenetic events has emerged as a promising approach to treating or preventing various types of cancers (Kohlmeier et al., 2016).

The epigenetic machinery has several levels of regulation: DNA methylation, posttranslational histone modifications, nucleosome positioning, and noncoding RNAs including microRNAs (miRNAs) and long-noncoding RNAs (Fig. 63.2). Among these mechanisms, DNA methylation is the best known. It can affect not only intragenic but also genome intergenic regions, affecting the transcription of genes by preventing the binding of key transcription factors (Crujeiras and Casanueva, 2015). DNA methylation usually occurs in certain areas of the genome with a high density of cytosine-phosphate-guanine (CpG) dinucleotides, thus named CpG islands (CGIs), a process catalyzed by DNA methyltransferases (DNMTs). Moreover, DNA methylation occurs in other genomic regions to maintain the conformation and integrity of chromosomes, as well as to avoid potential damage by mobile genetic elements. Array-based genome-wide methylation analyses have shown that DNA methylation can also affect areas of the genome with low CpG density increasingly far

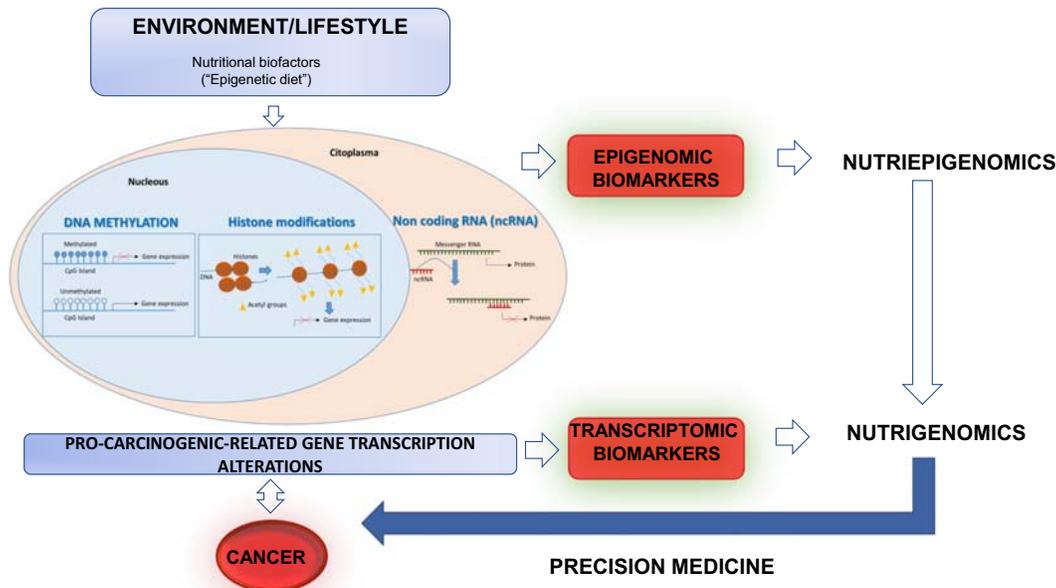


FIGURE 63.2 Epigenomic mechanisms involved in the promotion and prevention of carcinogenesis. Nutrients and biofactors may contribute to the prevention of carcinogenesis by regulating the epigenetic machinery. This regulation can modify the activity of procarcinogenic-related gene transcription. Therefore, the analysis of epigenomic or transcriptomic biomarkers provides insights for precision medicine in counteracting cancer onset and progression. CpG, cytosine-phosphate-guanine; lncRNA, long noncoding RNA; miRNA, microRNA.

from CGIs: CGI shores (up to 2 kb from CGIs), shelves (2–4 kb from CGIs), and open sea (>4 kb from CGIs). These low CpG-density regions may be especially important for gene regulation, as reflected by the conservation of small CpG clusters, despite the high mutation rates observed in CpG sites. Methylation is known to modify the function of DNA. When it occurs in the promoter region of the gene, hypermethylation is known to be associated with transcriptional repression and hypomethylation with transcriptional activation. However, the correlation of gene body methylation and expression was observed to be not only positive or negative, but also dependent on the cell type. Global hypomethylation is associated with nearly all human cancers, and altered DNMT expression and activity is associated with numerous diseases, including cancer (Crujeiras and Casanueva, 2015).

Epigenetic modifications are relevant in cancer because they are involved in cell proliferation, differentiation, and survival, as well as in tumor development and progression. Moreover, at least half of all tumor suppressor genes are inactivated through epigenetic mechanisms in tumorigenesis (Crujeiras and Casanueva, 2015).

As mentioned, several bioactive food components consumed by the ingestion of natural products, including fruits and vegetables, can influence the human genome directly or indirectly and thereby influence the expression of genes and gene products (Link et al., 2010). As a consequence, dietary patterns and/or specific dietary components may modify multiple cellular processes that can determinate the onset, incidence, progression, and/or severity of multiple diseases, including cancer. Thus, the use of high-throughput omics technologies holds the promise for developing more efficient nutritional interventions that might improve the management of cancer onset and progression. In this context, it was demonstrated in vitro and in vivo that several nutrients and bioactive compounds may exert chemopreventive effects in part by modulating various components of the epigenetic machinery in humans, such as DNA methylation, histone modifications, and regulation of expression of noncoding miRNAs (Link et al., 2010). Therefore, it is established that consumption of dietary agents can alter the epigenetic status and, as consequence, regulate the expression of genes involved in the hallmarks of cancer.

MOLECULAR PATHWAYS INVOLVED IN CARCINOGENESIS

Carcinogenesis consists of different stages, from initiation to promotion and progression of tumors. In this process, hallmarks such as cell proliferation, apoptosis, invasion and metastasis, angiogenesis, immortalization,

inflammation, immunity, genome instability and mutation, and cell energetics and metabolism are involved (Hanahan and Weinberg, 2011).

The most fundamental trait of cancer cells is their capacity to sustain chronic proliferation; this property is involved in the initiation, promotion, and progression of the tumor. Normal tissues carefully control proliferation to ensure a normal tissue architecture and function. By contrast, cancer cells develop deregulation of the proliferative control signals, such as JAK/Stat, mTOR, or PI3K/Akt. To acquire the capacity to sustain proliferative signaling, cancer cells activate the expression of growth factor ligands and receptors that activate survival pathways in an autocrine fashion. Moreover, growth factor ligands are secreted and poised to interact with normal cells within the tumor-associated stroma, in turn to induce the expression of growth factors that potentiate the proliferative capacity of surrounding cancer cells. This capacity is acquired when an oncogenic agent induces DNA damage in genes involved in controlling proliferation (Hanahan and Weinberg, 2011).

Cancer cells also promote the capacity to sustain proliferation and growth by negatively regulating the expression and activity of tumor suppression genes, such as *TP53*, retinoblastoma, and phosphatase and tensin homolog, or positively regulating the expression of oncogenes, such as *Myc* and *Ras* (Hanahan and Weinberg, 2011).

Another relevant hallmark of cancer is the inhibition of apoptosis or programmed cell death. This process is triggered in response to various physiologic stresses, such as DNA damage experienced by cells. The apoptotic machinery is composed of circuits that receive and process extracellular death-inducing signals, whereas other circuits integrate intracellular signals and both culminate in the activation of proteases (caspases 8 and 9), which in turn prompts a cascade of proteolysis to destroy cells and remove cell remnants by phagocytic cells. Cancer cells are able to develop strategies to evade the apoptotic machinery; this has become a target for anticancer therapies (Hanahan and Weinberg, 2011).

To generate macroscopic tumors, cancer cells require unlimited replicative potential, which the most normal cell lineages in the body do not have. This capacity is termed immortalization; it is centrally conditioned by the maintenance of the length of telomeric DNA achieved by upregulating the expression of telomerase. Once tumor mass is built, it requires nutrients and oxygen and the ability to evacuate metabolic wastes and carbon dioxide. These needs are addressed by the tumor-associated neovasculature, generated by the process of angiogenesis. This process is governed by angiogenic regulators, such as vascular endothelial growth factor-A, thrombospondin-1, and members of the fibroblast growth factor family. The upregulation of

expression of the angiogenic factors is induced by oncogenes such as *Myc* and *Ras*, or indirectly by immune-inflammatory cells. Finally, carcinomas progress to higher pathological grades of malignancy by gaining the capability of invasion and disseminating metastases. This process is induced by changes in the cancer cell shape and in their attachment to other cells and the extracellular matrix. The best-characterized changes here involve the loss of E-cadherin by carcinoma cells. The mechanisms by which cancer cells acquire the ability to invade are mediated by the regulatory epithelial-mesenchymal transition program, orchestrated by a set of pleiotropic transcription factors such as *Snail*, *Slug*, *Twist*, and *Zeb1/2* (Hanahan and Weinberg, 2011).

Regarding the mediators of carcinogenesis, oxidative stress and inflammatory agents have a relevant role because these agents induce DNA damage, inhibit pathways of DNA repair and detoxification, and induce immune evasion. Thus, cancer is caused by a series of transcriptional changes in tumor suppressor genes and oncogenes. These changes can be mediated by epigenetic alteration, including the DNA methylation profile, histone modifications, and the expression of noncoding RNAs. In fact, epigenetic regulation is an emerging area of interest in the study of cancer hallmarks, because functionally significant epigenetic changes are likely factors not only in cancer cells but also in altered cells of the tumor-associated stroma (Hanahan and Weinberg, 2011).

Moreover, certain nutrients and bioactive compounds feature antitumor activity because they are able to interact with specific molecular targets involved in several hallmarks of carcinogenesis (Gonzalez-Vallinas et al., 2013).

EFFECT OF NATURAL PRODUCTS AND DIETARY CONSTITUENTS ON THE REGULATION OF PROCARCINOGENIC-RELATED GENES

Several natural products and dietary constituents have been identified with the potential to prevent cancer and inhibit cancer progression. These products, especially phytochemicals (i.e., nonnutritive compounds with disease-preventive properties found in plants), contribute to cancer prevention by acting at different stages of carcinogenesis, from tumor initiation through all of the hallmarks of cancer (Gonzalez-Vallinas et al., 2013). Accumulating evidence suggests that some dietary phytochemicals may exert cancer chemopreventive and anticancer effects by regulating gene expression. This regulation mainly occurs through epigenetic modifications (Hardy and Tollefsbol, 2011) (Fig. 63.3).

A relevant example is the case of extravirgin olive oil (EVOO) and its phenolic compounds. It represents the typical lipid source of the Mediterranean diet, an eating

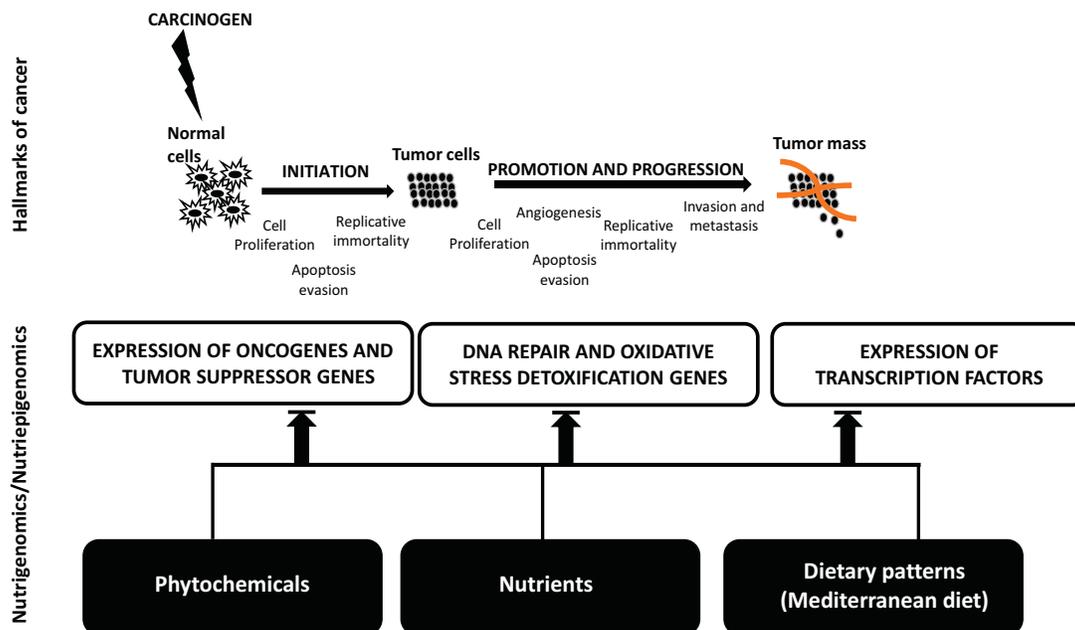


FIGURE 63.3 Interaction between nutrients and the molecular regulation of cancer hallmarks. When a carcinogen induces damage in a normal cell of the organism, it activates cellular processes, such as cell proliferation, immortalization, apoptosis evasion, angiogenesis, invasion, and metastasis (all hallmarks of cancer). These processes are mediated by regulation in the expression of genes involved in carcinogenesis and tumor suppression, DNA repair and oxidative stress detoxification, as well as transcription factors. Phytochemicals, nutrients, and specific dietary patterns such as the Mediterranean diet(s) may prevent and/or inhibit carcinogenesis-related processes modulating the expression of genes (nutrigenomics) by epigenetic mechanisms (nutria-epigenomics).

habit pattern associated with a significant reduction in cancer risk. EVOO and other dietary lipids may exert modulatory effects on cancer by influencing oxidative stress, a modulation of cell signaling transduction pathways, regulation of gene expression, and influences on the immune system. Treatment of cancer cells with EVOO or its administration in animal models of cancer induces a decrease in cell proliferation that goes hand in hand with the regulation of genes involved in the pathogenesis of colorectal cancer, such as those encoding for type 1 cannabinoid receptor and certain miRNAs (miR23a and miR-301a). This regulation is mediated by epigenetic mechanisms (Di Francesco et al., 2015). Curcumin is another relevant example of bioactive compounds with anticancer properties. It is a natural plant-derived compound that can slow proliferation and induces apoptosis in cancer cell lines by potently affecting gene expression profiles. In most studies, differentially expressed genes treated with this compound appeared to be involved in cell signaling, apoptosis, and the control of cell cycle. Curcumin can also induce specific methylation changes; it is a powerful regulator of the expression of miRNAs controlling oncogenesis (Hardy and Tollefsbol, 2011). A low intake of the dietary antioxidant micronutrient selenium (Se) is associated with increased risk for colorectal adenoma and with colorectal cancer mortality. In an integrative transcriptomic analysis, in rectal biopsy specimens from healthy participants, differences in Se status within the physiologic range were shown to affect global expression patterns of genes implicated in inflammatory signaling, immune function, and cytoskeleton remodeling. These findings provide insights into mechanisms through which Se may influence cancer risk and may help in the development of early biomarkers of disease risk (Hardy and Tollefsbol, 2011). Genistein, a soy-derived bioactive isoflavone, activates tumor suppressor genes and affects cancer cell survival by modulating chromatin configuration and DNA methylation (Hardy and Tollefsbol, 2011). Resveratrol, which is present in red wine and grapes, possesses a natural antiproliferative activity because it functions similarly to phytoalexin, a plant antibiotic. Pharmacological doses of this polyphenol can affect proliferation, apoptosis, cell cycle progression, inflammation, angiogenesis, invasion, and metastasis in multiple cancer cells by controlling signaling pathways related to Akt, PI3K/Akt, and nuclear factor- κ B (Gonzalez-Vallinas et al., 2013; Hardy and Tollefsbol, 2011). At the level of epigenetic regulation, resveratrol has limited demethylating activity, but its cancer-preventive effects have been attributed to its activation of sirtuin protein (SIRT1), for which increased expression and activity are related to proapoptotic effects, repression of transcription of oncogenes, or repression of the activity of oncoproteins. Moreover, SIRT1

regulates energy homeostasis, which is crucial for tumor development.

These health-promoting effects of dietary factors have been conventionally revealed by epidemiological studies, which initially found a consistent relation between abundant consumption of fruits and vegetables and a reduction in the risk for developing several diseases, including cancer. No mechanistic preclinical or interventional studies have been performed to date, although promising results have been found in vitro in several cell systems (Gonzalez-Vallinas et al., 2013). Therefore, it is imperative to carry out interventional studies to evaluate targeted pathways that are altered in patients with specific types of cancer and to evaluate which patients would benefit from these properties. Whole-genome transcriptomic analyses in samples collected during human dietary intervention studies provide insight into underlying mechanisms of chemoprevention by identifying biomarkers of exposure and the effects of nutrients and phytochemicals (van Breda et al., 2015). Thus, genomic techniques provide valuable tools for establishing mechanisms involved in preventing disease induced by dietary factors, as well as disclosing potential therapeutic targets to manage cancer progression (van Breda et al., 2015).

CONCLUSIONS AND FUTURE PERSPECTIVES

Several natural products and dietary constituents, such as phytochemicals, have been identified as able to contribute to cancer prevention by acting at different stages of carcinogenesis, from tumor initiation through all of the hallmarks of cancer. Some of these dietary components may exert cancer chemopreventive and anticancer effects by regulating gene expression. This regulation appears to occur mainly through epigenetic modifications. Diet-based interventions that target the regulation of gene expression pathways, including epigenetics, are promising for cancer prevention and the inhibition of cancer progression. However, translating these scientific findings into clinical or public health practices is not immediate. The amount of dietary components and the optimal dietary patterns needed to induce biological effects, as well as the timing of exposure and other variables that can influence cellular responses are still open questions. Moreover, most studies were conducted using a particular dietary factor, and it is probable that most such dietary factors may be consumed in the context of a complete food or dietary pattern, such as those rich in fruits and vegetables or the Mediterranean diet. Thus, additional clinical work is required to improve our understanding of molecular mechanisms underlying the chemopreventive effect of dietary factors. The

“omics” (e.g., genomics, epigenomics, transcriptomics, proteomics, metabolomics) approach may provide useful clues to cancer prevention, early disease, or nutritional status, and identify potential molecular targets in cancer processes that are modulated by dietary constituents and/or dietary patterns. Further studies are needed to translate preclinical studies to clinical settings and to evaluate the practical applications of such knowledge in conjunction with other cancer prevention and chemotherapeutic approaches.

References

- Crujeiras, A.B., Casanueva, F.F., 2015. Obesity and the reproductive system disorders: epigenetics as a potential bridge. *Hum Reprod Update* 21, 249–261.
- Di Francesco, A., Falconi, A., Di Germanio, C., et al., 2015. Extravirgin olive oil up-regulates CB(1) tumor suppressor gene in human colon cancer cells and in rat colon via epigenetic mechanisms. *J Nutr Biochem* 26, 250–258.
- Gonzalez-Vallinas, M., Gonzalez-Castejon, M., Rodriguez-Casado, A., Ramirez de Molina, A., 2013. Dietary phytochemicals in cancer prevention and therapy: a complementary approach with promising perspectives. *Nutr Rev* 71, 585–599.
- Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: the next generation. *Cell* 144, 646–674.
- Hardy, T.M., Tollefsbol, T.O., 2011. Epigenetic diet: impact on the epigenome and cancer. *Epigenomics* 3, 503–518.
- Kohlmeier, M., De Caterina, R., Ferguson, L.R., et al., 2016. Guide and position of the International Society of Nutrigenetics/Nutrigenomics on personalized nutrition: part 2 - ethics, challenges and endeavors of precision nutrition. *J Nutrigenet Nutrigenomics* 9, 28–46.
- Link, A., Balaguer, F., Goel, A., 2010. Cancer chemoprevention by dietary polyphenols: promising role for epigenetics. *Biochem Pharmacol* 80, 1771–1792.
- van Breda, S.G., Wilms, L.C., Gaj, S., et al., 2015. The exposome concept in a human nutrigenomics study: evaluating the impact of exposure to a complex mixture of phytochemicals using transcriptomics signatures. *Mutagenesis* 30, 723–731.

Nutrients and Gene Expression Affecting Bone Metabolism

José Luis Pérez-Castrillón¹, José Antonio Riancho del Moral²

¹Center of Investigation of Endocrinology and Nutrition, Medical School and Department of Internal Medicine, Hospital Universitario Río Hortega, University of Valladolid, Valladolid, Spain; ²Department of Medicine, School of Medicine, Department of Internal Medicine, Hospital Universitario Marques de Valdecilla, University of Cantabria, Santander, Spain

Glossary

AP-1 Activator protein 1
AFT-4 Activating factor transcription 4
BMD Bone mineral density
BMP Bone morphogenetic protein
BMU Bone multicellular unit
c-Fos Proto-oncogene c-Fos
c-myc Proto-oncogene c-myc
CYP24A1 Cytochrome P450 family 24 subfamily A member 1
CYP27B1 Cytochrome P450 family 27 subfamily B member 1
CYP3A4 Cytochrome P450 family 3 subfamily A member 4
DMP Dentin matrix acidic phosphoprotein
FA Fatty acids
FGF Fibroblast growth factor
GLI1 gene Glioma associated oncogene
GWAS Genome-wide association studies
IGF-1 Insulin-like growth factor 1
IL-6 Interleukin 6
LPR5/LPR6 Low-density lipoprotein receptor-related
MAPK Mitogen-activated protein kinase
MC3TB-E1 Cell of osteoblast cell line mouse
NFAT1c Nuclear factor of activated T cells cytoplasmic 1
OSCAR Osteoclast-associated immunoglobulin-like receptor
PTH Parathyroid hormone
Pt11 gene Pistil transmitting tissue 1
PUFAS Polyunsaturated fatty acids
RANK/RANKL/OPG Receptor activator of nuclear factor κ B/nuclear factor κ B ligand/osteoprotegerin
RUNX2 Runt-related transcription factor 2
RXR Retinoid X receptor
SMAD Small drosophila
SOST sclerostin gene
SOX9 Sex determining region Y-box 9
SRC Proto-oncogene tyrosine-protein kinase
SXR Xenobiotic steroid receptor
TERM2 Transcription factor term 2
TGF- β Transforming growth factor β
TNFSF11 Tumor necrosis factor ligand superfamily 11
TNF α Tumor necrosis factor α
TRAP Triiodothyronine receptor auxiliary protein

VDR Vitamin D receptor
VEGF Vascular endothelial growth factor
Wnt wingless

One sentence: Bone mineral density is genetically determined and may be modified by exogenous factors such as nutrients, which can modify the expression of genes encoding proteins involved in bone remodeling.

INTRODUCTION

Bone is one of the most metabolically active human tissues. It is constantly changing and is responsible for growth and maintenance of the body through two mechanisms: modeling and remodeling. Modeling is the process by which bone tissue is formed and constantly renewed before the end of the growth process. It takes place through membranous and endochondral ossification. Bone remodeling is a physiological process, the functions of which are permanent renewal of the skeleton for a biomechanically correct bone and the provision of minerals (calcium, phosphorus, and magnesium) required for the body's needs. Both modeling and remodeling occur in the bone multicellular units (BMU), formed by osteoclasts, osteoblasts, and osteocytes (Zaidi et al., 2018). During modeling, there is a predominance of activity of the osteoblasts, the bone-forming cells that determine skeletal growth. The remodeling cycle lasts 120 days and consists of these phases:

1. phase 1: initiation and activation of remodeling in a specific location;
2. phase 2: bone resorption and recruitment of mesenchymal stem cells and osteoblast progenitors;

3. phase 3: differentiation and activation of osteoblasts with osteoid synthesis; and
4. phase 4: mineralization of the osteoid.

Reports have described BMUs that are directly in contact with vessels, a highly vascularized region, and are lined by a canopy formed by lining cells. This process involves three types of cells.

Osteocytes

Osteocytes are cells derived from osteoblasts that are embedded in the bone matrix. Between 5% and 20% of osteoblasts are transformed into osteocytes. Osteocytes express some genes expressed by osteoblasts and others, such as sclerostin and genes related to bone mineralization and phosphorus metabolism, such as dentin matrix acidic phosphoprotein (DMP) and fibroblast growth factor (FGF)23, which are osteocyte-specific. They are located in lacunae and feature cytoplasmic dendritic processes in the canaliculi within the bone matrix that permit intercellular communication. In addition, because they are related to bone cells, they communicate with the blood vessels. These cells behave like mechanostats, responding to mechanical forces regulating bone formation and resorption. They are the main producers of sclerostin, which is encoded by the sclerostin (SOST) gene. Sclerostin antagonizes various bone morphogenetic proteins (BMPs) and binds to low-density lipoprotein receptor-related 5/6 (LPR5/LPR6), blocking the canonical wingless (Wnt) pathway. In addition to inhibiting bone formation, osteocytes may also initiate resorption, either by segregating the receptor activator for nuclear factor- κ B ligand (RANKL) or indirectly after apoptosis, by stimulating the production of RANKL by osteoblasts or cells of the osteoblast lineage. These cells produce osteoprotegerin that competes with RANKL and blocks resorption (Bonewald and Johnson, 2008).

Osteoclasts

Osteoclasts are cells derived from the hematopoietic system; they have two main functions: bone resorption and the stimulation of osteoblasts to initiate formation. They are directed to the BMU either from the bone marrow or from precursors in the bloodstream. Their surface contains many receptors that determine their proliferation, differentiation, and survival, the most important of which is the receptor activator for nuclear factor- κ B (RANK), the RANKL binding site. Osteoclasts release small extracellular vesicles of RANK. This vesicles trigger mineral deposition, indicative of bone formation (Zaidi and Cardozo, 2018). The binding of RANK to RANKL activates different kinases, whose final pathway is the calcineurin-dependent cytoplasmic

nuclear factor of activated T cells cytoplasmic 1 (NFAT1c) factor (NAFATc1), which is translocated to the nucleus and activates genes involved in the proliferation and maturation of osteoclasts. Other important receptors are the tumor necrosis factor- α (TNF α) receptor, the proto-oncogene tyrosine-protein kinase, and transcription factor term 2, which cooperate in activating NFATc1 (Zaidi and Cardozo, 2018; Henriksen et al., 2011).

To participate in resorption, osteoclasts require activation. First, they adhere to the bone surface through podosomes, with the participation of molecules called integrins. Their adhesion, by means of a sealing area, creates the resorption lacuna. Hydrogen ions are released in the lacunae and acidify the medium, allowing the activation of enzymes such as cathepsin, which degrades the bone matrix. Subsequently, the loose material (collagen and calcium) is reabsorbed and transported to the interior of the osteoclast, and subsequently released to the exterior. Once this process is completed, apoptosis, or programmed cell death, occurs.

However, before this occurs, osteoblast activation is necessary (Long, 2012). This is carried out in three ways:

1. the release of growth factors embedded in the bone matrix called matrikines, the most important of which are transforming growth factor β (TGF β), insulin-like growth factor (IGF)-1, and BMP2;
2. cell–cell contact (osteoclast–osteoblast) through transmembrane proteins, such as the ephrin system; and
3. the release of proteins that may be stimulatory, such as sphingosine-1-phosphate, BMP6, Wnt10b, and ephrin B₂; or inhibitory, such as semaphorin 4D.

Osteoblasts

The osteoblastic line includes cells derived from mesenchymal progenitors that can be differentiated into various populations: osteoblast precursors, which are attracted to the resorption lacunae by substances released from the bone matrix; surface osteoblast cells, which retract after receiving information from osteocytes, allowing the osteoclasts to act; mature osteoblasts and osteocytes; previously described osteoblasts embedded in the bone matrix (Yamaguchi et al., 2000).

Osteoblasts have a basophilic cytoplasm, abundant mitochondria, and a well-developed Golgi apparatus. This structure explains their capacity to synthesize proteins, such as alkaline phosphatase, osteocalcin, and large amounts of collagen type I, which forms the osteoid that is then mineralized by calcium hydroxyapatite crystals. Various transcription factors intervene in their differentiation and maturation, including sex

determining region Y-box 9 (SOX9), runt-related transcription factor 2 (RUNX2), activating factor transcription 4 (AFT-4), and activator protein 1 (AP-1). SOX9 is not expressed in mature osteoblasts but it favors differentiation from preosteoblasts. RUNX2 is a factor involved in mature osteoblast differentiation and function. AFT-4 regulates the expression of osteocalcin, an osteoblast protein involved in regulating glucose, and RANKL, which promotes osteoblast differentiation and function. All of these transcription factors are activated by exogenous factors (TGF β , IGF-1, and fibroblast growth factor [FGF]) that activate different metabolic pathways. TGF β is released from the bone matrix during resorption, facilitating the proliferation of osteoblast precursors and their transport to the resorption lacunae. TGF β also stimulates osteoblast proliferation and function, increasing the matrix synthesis and inhibiting late-maturing osteoblasts and osteocyte apoptosis, as well as regulating the osteoblast–osteoclast ratio. IGF-1 has an anabolic function, increasing osteoblast maturation and favoring the chemotaxis of precursors. FGF is composed of a long family of 22 proteins, some of which, such as FGF2 and FGF10, have anabolic actions, whereas others (FGF23) regulate phosphorus metabolism. They act on receptors located on the surface of preosteoblasts and osteoblasts, promoting differentiation and maturation (Yamaguchi et al., 2000).

All of these proteins act on receptors located on the osteoblast surface, activating various metabolic pathways that stimulate the previously described

transcription factors, including the Notch metabolic pathway, the Hedgehog protein pathway, and the Wnt and BMP pathways (Neve et al., 2011). BMP, Notch, and Wnt pathways are also regulated epigenetically by microRNA (Zaidi et al., 2018).

The Wnt metabolic pathway has attracted the highest pharmacological interest. The Wnt family consists of proteins involved in the bone-forming system of the osteoblasts. It has a β catenin–dependent signaling system and an independent system. Initially, Wnt binds to Frizzled receptors and their coreceptors (LRP4, LPR5, and LPR6), stabilizing β -catenin, which is then translocated to the nucleus and stimulates genes responsible for osteoblast maturation. This association is regulated by inhibitors, such as sclerostin and Dickkopf-related protein 1. There is a second signaling pathway, independent of β -catenin, which induces the activity of phosphatidylinositol coupled with protein G, thus activating protein kinase C, which in turn promotes osteoblast activity in a RUNX2-modulated process. Moreover, elements of the Wnt pathway, such as Wnt10A and Wnt5A, promote osteoblast differentiation, inhibiting adipogenic differentiation (Neve et al., 2011). Fig. 64.1 shows the genes involved in bone remodeling.

Initially, and until the third decade of life, there is a balance between bone formation and resorption, with the same amount of bone being formed by osteoblasts as is destroyed by osteoclasts. After this point, resorption, mediated by osteoclasts, predominates and increases in the perimenopausal and early menopausal

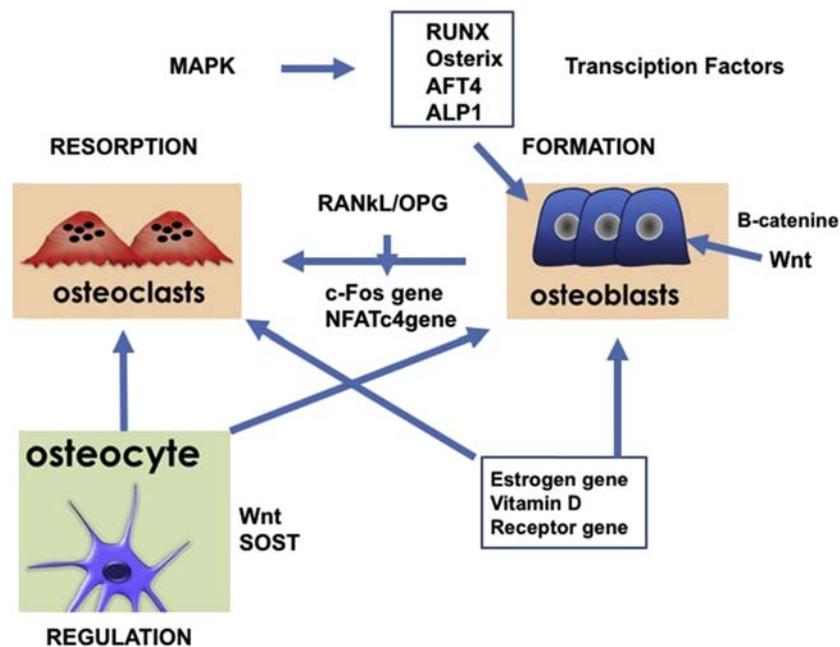


FIGURE 64.1 Genes and bone remodeling. *AFT4*, activating factor transcription 4; *ALP1*, actin-like protein 1; *MAPK*, mitogen-activated protein kinase; *NFATc*, nuclear factor of activated T cells cytoplasmic 1; *RANKL/OPG*, nuclear factor κ B ligand/osteoprotegerin; *RUNX*, runt-related transcription factor; *SOST*, sclerostin; *Wnt*, wingless.

periods, coinciding with the fall in estrogen. Correct bone health is determined by a balance between bone formation and resorption. In the early stages of development, formation predominates over resorption and is responsible for bone growth and the achievement of a peak bone mass. Bone mass is 60%–90% genetically determined and is the main determinant of fracture risk. Osteoporosis is the most common metabolic disease; it features fractures as the most serious complication. There are other genetically regulated risk factors for fractures, such as bone microarchitecture and bone geometry. Genetically, osteoporosis is a complex disease; many genes have small effects that contribute to the appearance of the characteristic phenotype. The genetic study of osteoporosis may be approached by studying polymorphisms involved in monogenic bone diseases or polymorphisms that modify the action of factors clearly involved in bone metabolism, such as estrogen or vitamin D. Another approach is through genome-wide association studies (GWAS), which have identified metabolic loci related to metabolic pathways such as Wnt/ β -catenin, RANK/RANKL/OPG, the endochondral ossification and transcription factors RUNX2, and osterix (Clark and Duncam, 2015).

Nutrients, and especially calcium and vitamin D, have an important role in the development and maintenance of bone mass. Other nutrients that influence bone mass and the risk for fractures in experimental or clinical studies are vitamins (A, B, C, and K), phytonutrients, and fatty acids. The effects of these nutrients on osteoblastic and osteoclastic genes are based on cell studies and experimental animal studies.

CALCIUM AND VITAMIN D

Calcium is an important nutritional element that acts in conjunction with vitamin D. In osteoclasts, it induces apoptosis and reduces the RANKL–OPG ratio. The reduction in RANKL inhibits expression of the *c-Fos* and *NAFATc1* genes, producing a reduction in osteoclast maturation and proliferation. In osteoblasts, calcium stimulates the mitogen-activated protein kinase (MAPK) pathway, acting on various transcription factors (RUNX2 and osterix) that activate osteoblastic genes (alkaline phosphatase, osteocalcin, and collagen I) responsible for forming the bone matrix. In addition, calcium stimulates the *CYP27B1* gene, which encodes the enzyme that converts calcidiol into calcitriol, the active metabolite of vitamin D (Shu et al., 2011).

Vitamin D facilitates the intestinal absorption of calcium and bone mineralization. It is synthesized in the skin from ultraviolet radiation, but in the winter months, and especially in the elderly, it depends on the exogenous supply, food, or supplements. It acts through an

active metabolite, calcitriol, which exerts its effect by binding to a receptor that is widely distributed in the cells and is a nuclear receptor member of the family of transcription factors. Its key feature is its nuclear localization, even in the absence of the ligand. It is associated with corepressor complexes that silence transcription. Ligand binding to the receptor requires the retinoid X receptor, a heterodimer that binds to positive or negative response elements of vitamin D. The heterodimer forms a coactivator complex, which permits transactivation of the target gene by direct and indirect mechanisms. The vitamin D receptor (VDR) regulates the expression of genes involved in different biological functions, including organ development, cell cycle control, calcium and phosphorus metabolism, detoxification, and the control of innate and adaptive immunity. In addition to facilitating intestinal calcium absorption, it activates osteoblastic genes, favoring the expression of the osteoblast phenotype, cell proliferation, and mineralization, and increases the expression of the *RUNX2* gene, osteocalcin, *c-myc*, and vascular endothelial growth factor (Bouillon et al., 2018).

The VDR gene is located on the long arm of chromosome 12 and is composed of a region of approximately 100 kb of DNA, although only 4.6 kb encodes the protein. More than 900 allelic variants have been described at this locus, some of which have been associated with chronic disease. Initially, various restriction enzymes were used to identify polymorphisms in this gene. The most studied are *Apal* (rs7975232), *BSML* (rs1544410), *TaqI* (rs731236), and *FokI* (rs10735810) (92). These polymorphisms seem to influence the stability of the receptor messenger RNA. The only exception is *FokI*, whose restriction site is located on exon 2 of the 5' coding region. This polymorphism results in a shorter protein, with three fewer amino acids, which increases the stability of the receptor and has greater activity (Uitterlinden et al., 2004).

Vitamin D is an element that regulates receptor expression. During the winter months, vitamin D levels depend on the exogenous supply, and a reduction below the threshold value stimulates the release of parathyroid hormone (PTH). In osteoblasts, PTH activates the *TNFSF11* gene, which is responsible for RANKL synthesis, and inhibits the gene responsible for osteoprotegerin, a RANKL antagonist. The activation of the *c-Fos* gene and *NFATc1* increases osteoclast maturation and activation, resulting in a reduction in bone mass. In addition, PTH stimulates *CYP27B1*, a gene encoding α -1 hydroxylase, which increases levels of the active metabolite of vitamin D, calcitriol. This increases the expression of the VDR in osteoblasts and the expression of the *LPR5* gene, a coreceptor of the Wnt pathway, all elements that increase bone formation. In turn, in a negative feedback, it inhibits *CYP27B1* and stimulates *CYP24A1*, decreasing the

synthesis of calcitriol (Van de Peppel and Van Leeuwen, 2014; Saccone et al., 2015). In various studies and meta-analyses, the administration of calcium and vitamin D has been shown to lead to a significant 15–18% reduction in the risk for fracture (Bolland et al., 2014). Reyes García et al. (2018) showed that daily intake of milk enriched with calcium and vitamin D in postmenopausal healthy women induced an improvement in vitamin D levels and bone mineral density (BMD) femoral neck. Another metaanalysis suggested that vitamin D does not prevent fractures and has no clinically significant effects on BMD. However, that study had serious methodological flaws (Bolland et al., 2018).

OTHER VITAMINS

Vitamin A

Studies showed that a high intake of vitamin A has a detrimental effect on bone health by increasing the risk for fractures, especially hip fractures. Retinoids affect both osteoclasts and osteoblasts. Chronic exposure to retinoids inhibits the expression of osteoblastic genes, reducing osteoblast formation. This effect is facilitated by the transformation of these cells into osteocytes, increasing the expression of SOST (the sclerostin gene, which inhibits the Wnt pathway) and DMP1. In addition, it activates the RANKL gene and increases bone resorption (Green et al., 2016).

Vitamin B

Plasma levels of homocysteine are the main metabolic biomarker related to vitamin B. The main B vitamins (folic acid, vitamin B6, and vitamin B12) reduce homocysteine levels; high homocysteine levels has been associated with a reduction in bone mass and an increased risk for fracture. Vitamin B deficiency and changes in the metabolism of tetrahydrofolate reductase are the leading causes of elevated homocysteine levels. Folic acid, the nutrient with the greatest effect on this, is suggested to have the capacity to inhibit the CYP24A1 gene and thereby increase calcitriol levels. In addition to lowering homocysteine levels, through this genetic mechanism, folic acid has a beneficial effect on bone health (Bailey and Van Wijngaarden, 2015). However, a prospective study found no evidence that daily supplementation with B vitamins reduced fracture risk in middle-aged and older women at high risk for cardiovascular disease (Stone et al., 2017).

Vitamin C

Epidemiological studies showed that vitamin C deficiency was associated with a decrease in bone mass and

an increased risk for fracture. In vitro studies showed that osteoblastic cells cultured with ascorbic acid induce the appearance of specific genes associated with osteoblasts. Vitamin C also upregulates Hedgehog proteins, promoting expression of the Gli1 and Ptc1 genes. The effect on osteoclasts is variable, depending on their maturity. In the early stages, vitamin C increases RANKL synthesis, stimulating proliferation and maturation; in later stages, it increases osteoclast death (Aghajanian et al., 2015).

Vitamin K

Vitamin K1 (phylloquinone) is a nutrient widely found in plants, whereas vitamin K2 (menaquinone) is of animal origin or derived from bacterial synthesis. These vitamins act by binding to a xenobiotic steroid receptor, which modulates gene transcription through a kinase-dependent mechanism. In the osteoclasts, it acts as an antagonist between vitamin K1 and vitamin K2. K1 stimulates genes (c-Fos, NFATc1, osteoclast-associated immunoglobulin-like receptor, and triiodothyronine receptor auxiliary protein) that increase osteoclast maturation, whereas vitamin K2 inhibits them. Moreover, vitamin K2 activates osteoblastic genes by means of CYP3A4. Both vitamins behave as cofactors for the carboxylation of osteocalcin, a protein necessary for bone formation (Azuma et al., 2014).

Phytochemicals

Phytochemicals (phytonutrients) are compounds derived from dietary fruit and vegetables. They can be classified into various groups of polyphenols, according to the number of phenol rings they contain and the associated structural elements. These include flavonoids, phenolic acid, tannins, stilbenes, coumarins and lignans, carotenoids, phytosterols, alkaloids, terpenes, and sulfur-containing compounds. Studies in humans are scarce, except for those in phytoestrogens and isoflavones.

Isoflavones are a subclass of flavonoids contained in soy, of which genistein and daidzein are the most common. Reports showed the benefits of administering isoflavones on bone health. Isoflavones are estrogen receptor modulators and may act as agonists, partial agonists, or antagonists. The effect mediated through their binding with the receptor, predominantly the β receptor, is genomic. A complex is formed with the receptor that interacts with the nuclear DNA, modifying protein synthesis. It acts on the MAPK pathway, stimulating transcription factors RUNX2, osterix, AP-1, and ATF-4, which are responsible for the increased activity of various osteoblastic genes. In in vitro studies using

MC3T3-E1 cells of osteoblastic origin, daidzein increased osteoblast activity, manifested by an increase in alkaline phosphatase activity, DNA synthesis, and a greater concentration of alkaline phosphatase in the supernatant (Horcajada and Offord, 2012). Moreover, genistein increases the production of osteoprotegerin (OPG) by osteoblasts and inhibits the activity of NF- κ B, reducing the synthesis of IL-6, a cytokine that increases bone resorption. Through this mechanism, isoflavones can modulate the RANK/RANKL/OPG metabolic pathway.

Crisafulli et al., (2004) found that genistein inhibits RANKL production and increases OPG production, which is related to changes in bone mass and in deoxy-pyridinoline, a marker of bone resorption, in healthy postmenopausal women. The effect was different from that produced by estrogens, which decreased the levels of both RANKL and OPG. Other beneficial effects of isoflavones are increased expression of the VDR, which facilitates the intestinal absorption of calcium and mineralization, a reduction in angiogenesis, an antioxidant effect, and inhibition of the activity of the enzyme aromatase.

Experimentally, other polyphenols have shown beneficial effects on osteoblastic genes mediated by the estrogenic receptors or the phosphorylation of various proteins (small drosophila, p38, and MAPK). In vitro studies using phenolic acids derived from olive oil found a 12–16% increase in osteoblast proliferation. These components can act through direct and indirect mechanisms. The indirect mechanisms are mediated by the estrogen receptors, which modify the expression of osteoblastic genes, whereas the direct mechanisms work through transcription factors (RUNX2 and osterix), BMPs, the osteocalcin gene, and genes of the proteins of the Wnt pathway (García-Marquez et al., 2016).

Fatty Acids

PUFAs are classified into two groups: omega-3 fatty acids (n3 FA) and omega-6 fatty acids (n6 FA). Studies of the effects of PUFAs on bone health and fracture risk are inconsistent. The effect seems beneficial for n3 FA, whereas the effect of n6 FA is not clearly established. n3 FA favor the differentiation of preosteoblasts to osteoblasts through an increase in the RUNX2 and osterix transcription factors. In mature osteoblasts, n3 FA inhibit the production of RANKL and OPG, thereby reducing the activity of c-Fos and NF- κ B in osteoclasts, suppressing their activation and proliferation. Furthermore, they reduce the activity of the cathepsin gene, a key enzyme in the destruction of the bone matrix. The effect is beneficial, because it consists of the stimulation of bone formation and inhibition of resorption. n6 FA

might have a detrimental effect, mediated through prostaglandin E₂. This stimulates the production of RANKL and inhibits osteoclast apoptosis, increasing bone destruction (Boeyens et al., 2014). The intake of long-chain omega-3 PUFA seems to be positively associated with BMD at the level of both the hip and the lumbar spine in normal and osteopenic women (Lavado-García et al., 2018).

Linoleic acid is an n-6 PUFA found in milk and meat; it modifies osteoblastic genes. In vitro studies showed that it blocks the expression of the peroxisome proliferator activator receptor γ gene, which inhibits osteoblast differentiation by mesenchymal cells and blocks the expression of genes involved in signaling pathways and transcription factors that promote bone formation. These studies have not been confirmed in humans, in which a neutral effect was found on bone mass despite the loss of fat and body weight (Ing and Belury, 2011).

The dietary intake of olive oil, which is rich in oleic acid, was positively associated with a better volumetric BMD in Spanish women. Moreover, the oil olive is a source of phenolic compounds that have a role in preventing osteoporosis (Roncero-Martin et al., 2018).

CONCLUSIONS

Cells (osteoblasts, osteoclasts, and osteocytes) involved in bone remodeling are genetically regulated by proteins that are transcription factors or part of metabolic pathways. In vitro studies of nutrients (calcium, vitamin D, vitamins B, C, and A, phytonutrients, and PUFAs) modify the expression of genes of the osteoblastic line encoding proteins involved in bone formation and resorption. These nutrients have a beneficial effect, except for vitamin A at high doses and n6 FA. However, these benefits do not have a clear impact on bone health in humans. Only the administration of calcium and vitamin D has shown a significant benefit in reducing osteoporotic fractures.

References

- Aghajanian, P., Hall, S., Wongworawat, M.D., Mohan, S., 2015. The roles and mechanisms of action of vitamin C in bone: new developments. *J Bone Miner Res* 30, 1945–1955.
- Azuma, K., Ouchi, Y., Inoue, S., 2014. Vitamin K: novel molecular mechanisms of action and its roles in osteoporosis. *Geriatr Gerontol Int* 14, 1–7.
- Bailey, R.L., Van Wijngaarden, J.P., 2015. The role of B-vitamins in bone health and disease in older adults. *Curr Osteoporos Rep* 13, 256–261.
- Boeyens, J., Deepak, V., Chua, W.H., Kriger, M.C., Joubert, A.M., Coetzee, M., 2014. Effects of ω - and ω 6-polyunsaturated fatty acids on RANKL-induced osteoclast differentiation of RAW264.7 cells: a comparative study. *Nutrients* 6, 2584–2601.

- Bolland, M.J., Grey, A., Gamble, G.D., Reid, I.R., 2014. The effect of vitamin D supplementation on skeletal, vascular or cancer outcome: a trial sequential meta-analysis. *Lancet Diab Endocrinol* 2, 307–320.
- Bolland, M.J., Grey, A., Avenell, A., 2018. Effects of vitamin D supplementation on musculoskeletal health: a systematic review, meta-analysis, and trial sequential analysis. *Lancet Diabetes Endocrinol* 6, 847–858.
- Bonewald, L.F., Johnson, M.L., 2008. Osteocytes, mechanosensing and Wnt signaling. *Bone* 42, 606–615.
- Bouillon, R., Marcocci, C., Carmeliet, G., et al., 2018. Skeletal and extra-skeletal actions of vitamin D: current evidence and outstanding questions. *Endocr Rev.* <https://doi.org/10.1210/er.2018-00126>.
- Clark, G.R., Duncan, E.L., 2015. The genetics of osteoporosis. *Br Med Bull* 1–9.
- Crisafulli, A., Altaville, D., Squadrito, G., Romeo, A., Adamo, E.B., Marini, R., Inferrera, M.A., Marini, H., Bitto, A., D'Anna, R., Corrado, F., Bartolone, S., Frisina, N., Squadrito, F., 2004. Effects of the phytoestrogen genistein on the circulating soluble receptor activator of nuclear factor κ B ligand-osteoprotegerin in early postmenopausal women. *J Clin Endocrinol Metab* 89, 188–192.
- García-Marquez, O., De Luna-Bertos, E., Ramos-Torrecilla, J., Ruiz, C., Milia, E., Lorenzo, M.L., Jimenez, B., Sanchez-Ortiz, A., Rivas, A., 2016. Phenolic compounds in extra virgin olive oil stimulate human osteoblastic cell proliferation. *PLoS One* 1–16.
- Green, A.C., Martin, T.J., Purton, L.E., 2016. The role of vitamin A and retinoic acid receptor signaling in post-natal maintenance of bone. *J Steroid Biochem Mol Biol* 155, 135–146.
- Henriksen, K., Bollerslev, J., Everts, V., Karsdal, M.A., 2011. Osteoclast activity and subtypes as a function of physiology and pathology. Implications for future treatments of osteoporosis. *Endocr Rev* 32, 31–63.
- Horcajada, M.N., Offord, E., 2012. Naturally plant-derived compounds: role in bone anabolism. *Curr Mol Pharmacol* 5, 205–218.
- Ing, S.W., Belury, M.A., 2011. Impact of conjugated linoleic acid on bone physiology: proposed mechanism involving inhibition of adipogenesis. *Nutr Rev* 69, 123–131.
- Lavado-García, J., Roncero-Martin, R., Moran, J.M., et al., 2018. Long-chain omega-3 polyunsaturated fatty acid dietary intake is positively associated with bone mineral density in normal and osteopenic Spanish women. *PLoS One*.
- Long, F., 2012. Building strong bones: molecular regulation of the osteoblast lineage. *Mol Cell Biol* 13, 27–38.
- Neve, A., Corrado, A., Cantatore, F.P., 2011. Osteoblast physiology in normal and pathological conditions. *Cell Tissue Res* 343, 298–302.
- Reyes García, R., Mendoza, N., Palacios, S., et al., 2018. Effects of daily calcium and vitamin D-Enriched milk in healthy postmenopausal women: a randomized controlled, double-blind nutritional study. *J Women's Health* 27, 561–568.
- Roncero-Martin, R., Aliaga Vera, I., Moreno-Corral, L.J., et al., 2018. Olive oil consumption and bone microarchitecture in Spanish women. *Nutrients* 10, 968.
- Saccone, D., Asani, F., Bornman, L., 2015. Regulation of the vitamin D receptor gene by environment, genetics and epigenetics. *Gene* 561, 171–180.
- Shu, L., Ji, J., Zhu, Q., Cao, G., Karaplis, A., Pollak, M.R., Brown, E., Goltzman, D., Miao, D., 2011. The calcium-sensing receptor mediates bone turnover induced by dietary calcium and parathyroid hormone in neonates. *J Bone Miner Res* 26, 1057–1071.
- Stone, K.L., Lui, L.Y., Christen, W.G., et al., 2017. Effect of combination folic acid, vitamin B₆, and vitamin B₁₂ supplementation on fracture risk in women: a randomized, controlled trial. *J Bone Miner Res* 32, 2331–2338.
- Uitterlinden, A.G., Fang, Y., Van Meurs, J.B.J., Pols, H., Van Leeuwen, J., 2004. Genetics and biology of vitamin D receptor polymorphisms. *Gene* 338, 143–156.
- Van de Peppel, J., Van Leeuwen, J., 2014. Vitamin D and gene networks in human osteoblasts. *Front Physiol* 5.
- Yamaguchi, A., Komari, T., Suda, T., 2000. Regulation of osteoblast differentiation mediated by bone morphogenic proteins, hedgehogs and Cbfa1. *Endocr Rev* 21, 393–411.
- Zaidi, M., Cardozo, C.P., 2018. Receptor became a ligand to control bone remodeling. *Nature* 56, 180–189.
- Zaidi, M., Yuen, T., Sun, L., Rosen, C.J., 2018. Regulation of skeletal homeostasis. *Endocr Rev* 39, 701–718.

This page intentionally left blank

S E C T I O N I V

Translational Nutrigenetics
and Nutrigenomics

This page intentionally left blank

Precision Nutrition Interventions Based on Personalized Genetic Advice

Rodrigo San-Cristobal¹, Santiago Navas-Carretero²,
Martin Kohlmeier³, J. Alfredo Martínez⁴

¹Cardiometabolic Nutrition Group, Precision Nutrition and Cardiometabolic Health, IMDEA Food Institute, Madrid, Spain; ²Centre for Nutrition Research, School of Pharmacy and Nutrition, University of Navarra, Pamplona, Spain; ³Human Research Core and Nutrigenetics Laboratory, UNC Nutrition Research Institute; and Schools of Medicine and Public Health, Department of Nutrition, Kannapolis, NC, United States; ⁴Ciencias de la Alimentación y Fisiología, Universidad de Navarra, Navarra, Spain

Glossary

DoHAD Developmental Origins of Health and Diseases. The concept that maternal nutrition during organ development could lead to modification of the risk for disease in adult life.

Gene–Environment-Wide Interaction studies (GEWIS) Observational study of a wide set or entire genome for the determination of interactions between gene variants and environmental exposure in relation to specific traits.

Genome-Wide Association Studies (GWAS) Observational study of a wide set of genetic variants in an extensive sample of population to analyze associations with specific traits. This analysis could be also applied on epigenetic marks (EWAS).

Nutrigenetic Field involved in personalized nutrition or precision nutrition that studies the effect of genetic background on individual response to different dietary interventions.

Omics From the Greek, the suffix *-ome* is used in molecular biology to refer to a totality of some sort. Refers to a field of study in biology involved in analyzing a wide set of data of genes (genomics), epigenetic marks (epigenomics), proteins (proteomics), or metabolites (metabolomics).

Orexigenic The propensity of increasing feeding and intake behavior mediated by specific neuronal circuitry and hormonal effects.

Precision Nutrition Field of nutrition focused on the design of nutritional recommendations to treat or prevent metabolic disorders based on parameters such as eating habits, food behavior, physical activity, as well as genetics, epigenetics, microbiota, or metabolome, among other factors.

Reference Intakes Recommended amounts of nutrient for daily intake to the public ensuring nutritional adequacy to meet nutrient requirements and prevent nutritional imbalances (under- or overconsumption). Usually, these recommendations are clustered by sex and age.

Single Nucleotide Polymorphism (SNP) Common variation in DNA sequence on one base or nucleotide (DNA unit) from the more common form; the acronym should not be used to indicate other types of polymorphisms, such as deletions, insertions, or repeats.

PRECISION NUTRITION

Public health nutrition guidelines involving reference intakes and recommended dietary allowances are estimated nutrient requirements of people without a known disease. These norms are intended to meet the nutritional needs of most members of a community and to prevent common nutrition-related harm such as obesity, diabetes, and cardiovascular diseases. It is important to recognize that reference intakes are not intended to manage patients with metabolic disorders or other diseases. Actually, it would be harmful for carriers of numerous inherited dispositions, such as phenylketonuria, urea cycle defects, hemochromatosis, and celiac disease, to follow all recommendations for healthy populations. Although individually each of these conditions is rare, in the aggregate they affect more than 1% of some populations. People with additional common genetic variants are also not well-served by the recommendations because some are known to cause increased vulnerability to some foods or nutrients and others require higher intake to prevent disease optimally. For example, two-thirds of the world population do not well-tolerate high consumption of lactose-rich foods such as milk. Another example is the many women (more than one in five of Mexican Americans and Southern Italians) who, owing to their genetic disposition, need much higher folate intake than most others if they are to avoid an increased risk for having children with birth defects.

The success of precision nutrition is likely to depend on the ability to integrate multiple factors, such as information about the individual's genome, epigenome, and microbiome, and other "omics" data sources. It is also important to consider prenatal and postnatal nutrition, heritable epigenetics, allergies, intolerances, religious, cultural and ethnic aspects, personal food preferences and lifestyle habits, diseases, and medication use.

Nutrigenetics is currently devoted to consider the impact of the genetic background on the body's metabolic response to foods or specific nutrients. New genetic approaches and nutritional scores enable researchers to screen deeply and assess the influence of genetic variations in complex metabolic responses to specific individual nutrients, dietary prescriptions, or particular foods. Newer nutrigenetic technologies involving microarrays and sequencing strategies are able to explore inherited DNA variations, as was already done for certain newborn test screens, to confirm syndromic diagnoses, for carrier screening, and to assess disease risk prediction and prognosis of asymptomatic individuals. Therefore, nutrigenetics has great potential as a tool for personalizing disease prevention, diagnosis and prognosis, and treatment in wide fields of nutrition, obesity, and other metabolic diseases by facilitating the customization of individualized healthy lifestyle and nutrition counseling.

The current flood of discoveries has raised expectations about screening tests for disease susceptibility variants that can identify individuals at risk for developing metabolic disorders. A major concern arises from the rising amount of information that is not based on sound research practices and has not undergone peer review, often spread by Internet sites. To counter unreliable information about nutrigenetic tests and support consumers considering direct-to-consumer (DTC) genetic test services (see Chapter 77), various authorities have started initiatives that register and evaluate such offerings. The National Center for Biotechnology Information website is a good example (<https://www.ncbi.nlm.nih.gov/gtr/>). An additional initiative funded by the European Union was the Food4me project, which aims to investigate the awareness of European populations, prepare to set boundaries for personalized nutrition, and provide advice for audiences at all levels (from general recommendations to genetic advice). This project is also intended to explore the economic potential for when dietary advice is provided and monitored by credible experts with demonstrated nutrigenetic expertise.

Genetic disposition can vary greatly among individuals with different ethnic backgrounds. Knowledge of such diversity is critically important for the development and formation of health professions that can adapt existing programs and services to a wide range of individuals and communities in a personalized manner.

Furthermore, advances in epigenetics and other knowledge in gene function and expression not directly linked to the DNA sequence (changes that are not necessarily fixed for the entire lifetime) give new insights and promise the development of specific dietary treatments for slowing age-related impairments and illnesses. For instance, age at onset as well as the severity of lactose intolerance vary, which makes it difficult to apply population-based and personalized recommendations. However, for polygenic diseases such as obesity, diabetes, dyslipidemia, and cancer, current knowledge is often insufficient to prescribe detailed personalized protocols aiming at primary prevention for all high-risk individuals. There is some promise for the development of better and clinically useful nutrigenetic algorithms and bioinformatics approaches.

Emerging technologies, such as *omics* platforms, are convenient integral approaches in which current precision nutrition reference intakes may be refined and the huge possibilities of individualized nutrition implemented. Therefore, it seems likely that in the near future, it will be possible to offer personalized recommendations to more subgroups. There is a great need for continued research that includes more diverse populations to understand interactions between genetic polymorphisms and nutritional imbalances.

Nutrigenetic tests usually do not target rare mutations or variants causing serious illness with high penetrance, but instead focus on common variants that increase the susceptibility to deficiencies or increase the risk for chronic diseases. This risk is often initially assessed by genome-wide association studies (GWAS) or by the investigating the predictive power of common gene variants. Understanding the complex interactions between genes (or genes and environmental factors) may lead to more precise and focused guidelines for complex diseases. Genetic tests are useless without evidence-based individualized interpretation and professional support. The health care professional must be competent to translate and recommend the best course of action based on the analyzed genetic variants, costs, quality, clinical validity, and other relevant information. It is important first to consider best clinical practice based on environmental factors and family history, and to use genetic information in that context. The formulation and application of genetic scores computing the number of effect alleles is gaining acceptance and there is an increasing number of examples of their use in practice. A well-established example is pharmacogenomics, in which gene–drug interactions are being employed to select and medications and their dosages. This textbook contains a chapter that goes to the next level and explores the use of pharmacogenetic considerations (see Chapter 68).

NUTRITIONAL INTERVENTIONS BASED ON GENOTYPE

There is a wide number of reported single nucleotide polymorphisms (SNPs) concerning weight and metabolic control in relation to obesity and the onset of other related metabolic diseases, characterized by metabolic disequilibria such as diabetes or dyslipidemia. Some of these SNPs have been repeatedly associated with a higher risk for having a determined disease in GWAS studies or through candidate gene approaches (Fig. 65.1). Thus, common variants in the *FTO* gene (rs9939609) have been related to obesity and metabolic syndrome traits in a number of GWAS in populations of different ethnic origins, but also to several variants near the melanocortin 4 receptor (*MC4R*) gene (rs17782313 and rs12970134). Regarding cardiovascular risk, some SNPs have been identified in GWAS studies, such as several variants in the *APOA5* gene (rs651821 and rs662799). Other variants and SNPs are being investigated to have a deeper regulatory genetic understanding in relation to fat consumption, folate requirements, vitamin D functions, sodium intake outcomes, and the role of healthy dietary regimens and Mediterranean food-based patterns with an emphasis on precise nutritional interventions for obesity and diabetes management with different energy-restricted diets and diverse macronutrient distribution. The implementation of nutrigenetic strategies in distinct populations and ethnic groups is also a major current challenge. Some relevant examples concerning gene variants, nutrigenetics tools, and potentially suitable nutritional approaches are reported in different nutritional conditions.

In general, reported SNPs in this section were studied in Caucasian populations, because they are more widely

studied than other populations. Information on genotype frequencies in different ethnic groups is readily available (<http://www.ncbi.nlm.nih.gov/SNP>).

Technical and scientific advances involving large-scale studies have allowed the identification of many genetic variants, with a significant impact on risks associated with having obesity, diabetes, and cardiovascular diseases. Thus, a GWAS together with a metaanalysis performed in more than 339,000 individuals identified 97 body mass index (BMI)-associated loci related to obesity, 56 of which were novel. A number of these SNPs achieved significance not only in relation to obesity but also in respect to other metabolic phenotypes related to diabetes, hypertension, dyslipidemia, and inflammation. Reports based on GWAS suggested that common variation accounts for more than 20% of variation in BMI, much more than the less than 5% variation in BMI explained by the 97 loci mentioned earlier. Subsequent pathway analyses from such investigations provide support for a function of the central nervous system in obesity susceptibility and point to specific genes involved in cell signaling, insulin secretion and action, energy and lipid metabolism, thermogenesis, and adipogenesis. Furthermore, body fat distribution is heritable and a predictor of adverse cardiometabolic traits, as ascertained in another genome-wide association metaanalysis of outcomes related to abdominal obesity estimated through waist and hip circumferences in up to 224,459 individuals, which was able to recognize 49 loci (33 new) associated. In total, 19 of the 49 loci showed significant sexual dimorphism with a more sizable effect in women, for whom at least 20 of 49 involved transcriptional regulation, angiogenesis, adipogenesis, and insulin resistance as featured processes participating in fat distribution and accompanying putative adverse pathophysiological mechanisms. Another epigenome-wide association study (EWAS) carried out in 10,261 samples pointed to 187 CpGs with a causal role or otherwise related to obesity. Such information will pave the way for personalized medicine and precision nutrition.

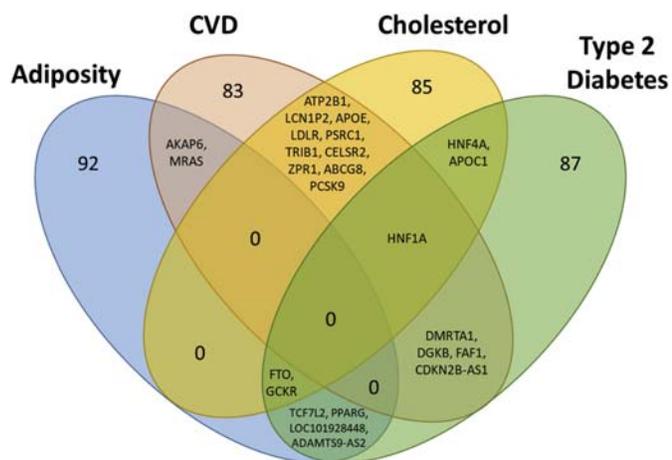


FIGURE 65.1 Shared genes among associated traits (adiposity, type 2 diabetes, cardiovascular diseases (CVD), and total cholesterol levels) from the top 100 genes in genome-wide association studies analyses.

Precision Nutrition Interventions Relating to Specific Nutrients

A number of instances reveal nutrigenetic interactions concerning dietary practices or specific nutrients (Table 65.1). Thus, candidate gene and genome linkage–association studies have involved pro-opiomelanocortin loci in features related to appetite, metabolic regulation, and adiposity. Nutrient choice was influenced by the rs1042571 genotype, which may affect an individual's energy intake. Also, a comparison of a healthy dietary pattern rich in vegetables, fruits, legumes, low-fat dairy, whole grains, vegetable oils, and fruits in relation to a

TABLE 65.1 Personalized nutrition interventions and some of the most important studies performed.

<i>ADBR2</i>	Gln27Glu	Carbohydrate intake	Consumption of carbohydrates (<49% energy) increases risk for obesity depending on <i>ADBR2</i> genotype	Caucasian population (2003)
<i>MTTP, FABP, others</i>	23 SNPs	Mediterranean-type diets	Several interactions between particular SNPs and metabolic responses to diet are reported	French population (2009)
<i>GCKR</i>	rs1260326	High-fat challenge	<i>GCKR</i> gene polymorphism affects postprandial lipidemia responses	Amish cohort (2009)
<i>CETP</i>	4502 C>T, TaqIB	TaqI Mediterranean diet	<i>CETP</i> diplotype TT/B2B2 was associated with higher HDL cholesterol levels compared with the diplotype CC/B1B1, but were not modified by mediterranean diet	Mediterranean cohort (2010)
<i>POMC</i>	rs1042571	Energy intake and food choice	rs1042571 polymorphism participates in appetite and food intake	Dutch population (2011)
<i>ADIPOQ</i>	rs182052	Monounsaturated fatty acids (MUFA)	10066 G/G homozygotes. Enriched MUFA diet increases adiponectin concentration	Caucasian Europeans (2011)
<i>PAR</i>	Pro12Ala, Leu162Val	MUFA	Interaction between <i>PPAR</i> genotypes and fatty acids influences cholesterol metabolism	British individuals (2011)
<i>PPAR</i>	rs1800206, rs1801282, rs2016520	Fat content and quality	Different <i>PPAR</i> polymorphisms interact with dietary fats on cardiovascular disease risk	Canadian subjects (2011)
<i>ADBR2</i>	rs1042713	Energy restriction	<i>ADBR</i> genetic variants influence body weight loss and composition changes	Spanish (2011)
<i>IRS1</i>	G972R	Dietary fat	Insulin response is determined by interaction between G972R polymorphism and dietary fat	Mediterranean population (2011)
<i>VDR</i>	rs3135506	Vitamin D consumption	Vitamin D intake influences plasma HDL cholesterol levels	Utah Females (2012)
<i>FTO</i>	rs9939609	Saturated fatty acids (SFA) and polyunsaturated fatty acids (PUFA)	High intake of SFA exacerbated fat accumulation	Mediterranean population (2012)
<i>IRS1</i>	rs7578326, rs2943641	SFA/carbohydrate ratio and MUFA	Carbohydrate and fatty acids intake influences insulin resistance depending on genotype	American mixed populations (2013)
<i>SREBF1</i>	rs12953299, rs4925118, rs4925115	Fish oil	<i>SREBF1</i> gene variants modulate insulin sensitivity in response to fish oil supplementation	Canadian subjects (2014)
<i>MTHFR</i>	TT	Folate supplementation	TT genotype is associated with lower response to folate intake	Australian people (2015)
<i>APOA2 APOA5 LEPR</i>	rs3813627/rs5082 rs662799, rs3135506, s8179183, rs1137101	Fat quantity and quality	Dietary fat interacts with <i>APOA2</i> , <i>APOA5</i> , and <i>LEPR</i> , affecting lipidemia	Mexican individuals (2015)
<i>MC4R</i>	rs12970134	Healthy versus western diet	Interaction between rs12970134 and fat and vegetable intake on risk for metabolic syndrome	Iranian volunteers (2016)

<i>NPPA</i> <i>NPPC</i> others	rs5063 rs2077386	Dietary sodium	Differential nutrigenetic responses depending on genes related to natriuretic peptide system to dietary sodium	Chinese population (2016)
<i>IL-6</i> <i>PPARG</i>	174 G>C Pro12Ala	10-week energy restriction	Both polymorphisms interact to protect against weight regain	Spanish subjects (2006)
<i>PGC-1α</i>	rs8192673	Low-calorie diet and 1-year follow-up	Enhanced insulin sensitivity in carriers of Gly482Ser variant after weight loss	Spanish overweight and obese subjects (2008)
<i>CD36</i>	rs2151916	Weight management	<i>CD36</i> gene is associated with improved lipid profiles during weight loss and subsequent maintenance	Twins United Kingdom cohort and obese Spanish subjects (2008)
<i>LEPR</i>	Lys109Arg	Hypocaloric diet	AA genotype group had higher body mass loss and cholesterolemia reduction compared with minor allele carriers	Caucasian subjects (2009)
<i>ADIPOQ</i>	rs17300539	Energy restriction and weight maintenance followed by 60 weeks	Carrying A-allele protects against weight regain	Spanish cohort (2009)
<i>TEAP2B</i>	rs987237	Macronutrient distribution and weight control	High-fat hypocaloric diet produced lower fat loss in T-allele carriers	European overweight/obese (2010)
<i>NPY</i>	rs16147	Energy restriction with different macronutrient distribution	<i>NPY</i> genotypes influence on adiposity phenotypes and is modified by dietary fat within hypocaloric diet	American obese subjects (2015)
<i>DHCR7</i>	rs12785878	Low-calorie high-protein diets	T-allele carriers improve insulin resistance more effectively when consuming energy-restricted diets enriched in protein	American overweight and obese individuals (2015)
<i>CETP</i>	rs3764261	Two independent trials with different weight loss diets varying in fat content	Genetic variants on <i>CETP</i> modify effect of weight-loss diets depending on fat content	Two different American cohorts (2015)
<i>FTO</i>	rs1558902	Hypocaloric diets with different fat and protein contents	Carriers of risk alleles (A) benefit more under low-calorie diets with high-fat content	American cohort (2015)
<i>FABP2'</i>	Ala54Thr	Moderately high-protein diet versus standard low-calorie diet	Total and LDL cholesterol decreased only in wild-type genotype	Spanish cohort (2015)
<i>CNR1</i>	rs1049353	100-kcal/day diet with different macronutrient compared with conventional hypocaloric diet	Similar weight reduction with both diets, but better glucose response in low-carbohydrate diet	Caucasian females (2015)
<i>TNF-α</i>	rs1800629	High-protein/low-carbohydrate diet versus control energy-restricted diet	Better metabolic response in -308 GA obese with high-protein hypocaloric diet.	Spanish population (2016)
<i>FGF21</i>	rs838147	Low calorie diet and different macronutrient distribution	High-carbohydrate/low-fat diet was associated with higher fat loss in C-allele carriers.	American overweight and obese (2016)
<i>Genetic Risk Score</i>	14 SNPs	Weight loss and macronutrient content	Subjects with higher genetic risk score improve glucose homeostasis with low-fat diets	Americans with excessive body weight (2016)

CETP, cholesteryl ester transfer protein; *HDL*, high-density lipoprotein; *LDL*, low-density lipoprotein; *PPAR*, peroxisome proliferator-activated receptor; *SNP*, single nucleotide polymorphism.

Westernized dietary pattern consisting of a high intake of soft drinks, fast foods, refined grains and sweets, solid fats, red meats and poultry, salty snacks, high-fat dairy, and eggs revealed that saturated fatty acid intake could modulate the association of A-allele carriers of MC4R (rs12970134) with abdominal obesity and associated metabolic syndrome features. Other intervention concerning Mediterranean-type diets, which are commonly related to a reduction in cardiovascular events, showed that the genetic background may be important to explain differential dietary outcomes. Thus, interactions from SNP in genes coding for microsomal transfer protein or intestinal fatty acid binding protein with Mediterranean-type diets produced different metabolic responses associated with the genotype concerning triglyceride, low-density lipoprotein (LDL) cholesterol, or Framingham score lowering. However, the association of the cholesteryl ester transfer protein (CETP) SNPs and high-density lipoprotein (HDL) cholesterol levels was not statistically modified by the adherence to a Mediterranean pattern or by other environmental factors.

Research on the relations of consuming a high-fat diet in carriers of specific genetic polymorphisms on the APOA2 (rs3813627 and rs5082), APOA5 (rs662799 and rs3135506), and LEPR (rs8179183 and rs1137101) genes demonstrated that dietary fat intake and quality modify the effects of APOA5 and LEPR genetic variants on serum triglycerides and cholesterol concentrations in young subjects. In this field, the apolipoprotein E (APOE) risk allele ($\epsilon 4$) has been related to an amplified response to saturated fatty acid (SFA) reduction, but this finding is still inconclusive. Nevertheless, a gene-based personalized nutrition targeted toward APOE was more effective in reducing SFA intake than standard dietary advice, but there were no differences between APOE "risk" and "non-risk" groups concerning the response to fat intake. A gene transcript that modulates postprandial triglyceridemia is expressed by the glucokinase regulatory protein (GCKR) gene, which influences lipid response after an acute high-fat challenge. Data from a dietary intervention study showed that the rs1260326 T-allele of GCKR is associated with fasting hypertriglyceridemia, which may explain a higher atherogenic risk depending on both genetic disposition and nutritional exposure.

The peroxisome proliferator-activated receptors (PPARs) are transcription factors regulated by fatty acid consumption. Thus, an intervention study reported an interaction between PPARG Pro12Ala and PPARA Leu162Val genotypes on circulating LDL cholesterol levels after a high monounsaturated fatty acid diet. Furthermore, adiponectin messenger RNA levels are modulated by PPAR γ . Another trial with Canadian subjects evidenced gene–diet interactions for the PPAR δ

-87T \rightarrow C polymorphism with the polyunsaturated/saturated fat ratio and for the PPAR γ P12A genetic variant with saturated fat intake, which may determine the effect of diets with different lipids on cardiovascular disease events.

Genetic variants in the fat mass and obesity-associated (FTO) gene have been associated with excessive fat deposition, reduced insulin sensitivity, and other unhealthy metabolic traits; its control may be related to macronutrient consumption. The FTO rs9939609 genotype is a relevant obesity genetic determinant of adiposity involving dietary interactions, because higher dietary SFA consumption ($\geq 15.5\%$ energy) and a lower dietary polyunsaturated fatty acids to SFA intake ratio (< 0.38) promoted abdominal obesity and increased cardiometabolic risk. Furthermore, the FTO risk allele is associated with a lower reported total energy intake and altered patterns of macronutrient consumption. These authors concluded that although significant, these differences are small. Further research is required to ascertain whether the relations are independent of dietary reporting.

Furthermore, a case–control study (normal weight controls versus obese people) devised to evaluate a possible dietary effect modification on obesity risk of the Gln27Glu polymorphism for the $\beta(2)$ -adrenoceptor gene (ADRB2) showed that a high intake of carbohydrates (more than 49 energy %) was associated with higher insulin levels among women carrying such a genetic variant. This finding concerning a gene–nutrient interaction emphasized the importance of examining the outcome of some obesity-related mutations depending on dietary sugars and also the gender.

Moreover, a metaanalysis involving observational and intervention studies confirmed that the MTHFR TT genotype is associated with higher homocysteinemia and lower serum folate as well as by smaller responses to short-term folate supplementation, which may lead to prescribe individualized intake gene-based recommendations for this nutrient. Another micronutrient with putative nutrigenetic interactions is vitamin D, whose deficiency is a potential risk factor for cardiovascular disease, because this micronutrient may network at established lipid metabolism-regulating gene loci, which explains the variability associated with dietary lipid intake. Thus, rs3135506 minor allele carriers were associated with smaller HDL-C concentrations in subjects with reduced dietary 25-hydroxyvitamin D (25OHD). This research also suggested that the 25OHD receptor site modulating an APOA5 promoter polymorphism is related to lower HDL-C in vitamin D–deficient persons. Moreover, another nutrient relation with genetic variants was found in the natriuretic peptide (NP) system with blood pressure response to dietary sodium intervention. A 7-day low-sodium intake followed

by 7-day high-sodium ingestion evidenced significant associations with blood pressure response to low-sodium dietary intervention under a recessive genetic model concerning SNP rs5063 in the NPPA gene and SNP rs2077386 in the NPPC gene.

Precision Nutrition Interventions for Managing Obesity and Cardiometabolic Traits

A number of nutritional interventions and studies with nutrigenetic foundations have repeatedly focused on inducing weight loss and managing metabolic impairments. A systematic review and metaanalysis from randomized controlled trials to assess the effect of the FTO genotype on weight lowering after a number of nutrigenetic-based interventions showed that the genetic predisposition to excessive weight associated with the FTO minor allele can be at least partly overcome through precision nutritional prescriptions, although the conclusion remained elusive. Another research revealed that carriers of the risk alleles of rs1558902 in FTO differentially reduced insulin resistance by consuming high-fat weight loss diets instead of low-fat diets. Another nutritional intervention explored several appetite-controlling polymorphisms: rs10830963 (MTNR1B), rs9939609 (FTO), and rs17782313 (MC4R), which may affect weight loss in response to an energy-restricted diet. Among women, higher total and animal protein intakes were associated with lower weight loss in G-allele carriers of the MTNR1B variant; this association was conjointly influenced by FTO and MC4R loci.

The rs1049353 polymorphism of the cannabinoid receptor (CNR1) gene is involved in appetite and energy homeostasis regulation; it also has a role in cardiovascular risk factors and weight loss. However, an interventional study comparing a high-protein/low-carbohydrate versus a standard hypocaloric diet (1000 kcal/day) over 9 months did not find a relation between the rs1049353 CNR-1 polymorphism and body weight response, but improved glucose metabolism in subjects independent of the genotype. Within the neuronal circuitry involving orexigenic pathways, neuropeptide Y is a key peptide regulating food intake; its genetic variants may interact with dietary-induced changes in adiposity markers. The rs16147 genotype role in abdominal obesity and central fat distribution phenotypes was modified by dietary fat.

Adrenoceptors are involved in lipid metabolism and body weight control. The response to a 12-week energy-restricted diet in women was affected by ADRB2 rs1042714 (Gln27Glu) and rs1042713 (Arg16Gly) polymorphisms. The experimental outcomes suggested that the ADRB2 Gln27Glu polymorphism had an effect

on hypocaloric-induced evolution on body weight and composition mediated by lipolytic mechanisms. Another lipid metabolism regulatory protein involving a polymorphism of the CD36 fatty acid transporter gene (rs2151916) was examined in obese subjects during an 8-week low-calorie diet and 6-month weight maintenance intervention. The CD36 promoter SNP allele 22674C was associated with lower serum LDL cholesterol concentrations in female twins and also with an improvement in circulating lipid levels during weight loss and ponderal maintenance in obese persons. Also, CETP genetic variants are involved in lipid turnover regulation and may modify the effects of low-calorie diets with varying macronutrient content on plasma lipid levels. Replicated analyses from two independent American cohorts showed that individuals with the CETP rs3764261 CC genotype might have greater healthy benefits with decreased triglycerides and increased HDL cholesterol concentrations with a low-carbohydrate/high-fat weight-lowering diet compared with a low-fat diet. In this area of investigation, it was also reported that the expression of the fatty acid-binding protein 2 (FABP2) gene depends on dietary intake; the Ala54Thr polymorphism of this protein is associated with insulin functions. The analysis of weight loss changes subsequent to a high-protein/low-carbohydrate versus a standard hypocaloric diet for 9 months revealed that only in the FABP2 wild genotype, total and LDL cholesterol levels decreased with both energy-restricted diets.

The association of variants of the insulin receptor substrate 1 (IRS1) gene with insulin resistance and accompanying phenotypes was tested for interactions with diet. Thus, in the Genetics of Lipid lowering Drugs and Diet Network (GOLDN) cohort, rs7578326 G-allele carriers and rs2943641 T-allele carriers had a significantly lower risk for insulin insensitivity than noncarriers only in subjects in whom the dietary saturated fatty acid to carbohydrate ratio was low (≤ 0.24), whereas in the GOLDN and Boston Puerto Rican Health Study populations, individuals with the rs7578326 G-allele had a lower risk for metabolic syndrome than noncarriers only when the intakes of monounsaturated fatty acids were inferior than the median consumption of participants in each cohort. Moreover, an association between insulin sensitivity changes associated with fish oil supplementation in carriers of different polymorphisms (SNPs) involving the sterol regulatory element binding transcription factor 1 gene was reported, which may contribute to a more precise dietary prescription in diabetic subjects. Moreover, the G972R polymorphism at IRS1 was reported to interact with dietary fat to modulate insulin resistance. Thus, insulin sensitivity increased in GR carriers of the G972R polymorphism at the IRS1 locus after intake of a carbohydrate-rich

diet. Transcription factor AP-2 β (TFAP2B) rs987237 was related to fat deposition showing interactions with the dietary fat to carbohydrate ratio and subsequent effects on weight loss. The TFAP2B rs987237 polymorphism interplays with dietary protein–carbohydrate modifying weight maintenance outcomes. Those authors suggested that TFAP2B dietary protein–glycemic interactions depend on the nutritional state, although the dietary carbohydrate distribution and quality of the interaction were different from the previously reported rs987237 fat to carbohydrate ratio involvement in weight loss. In this area of research, the transcription factor 7-like 2 rs7903146, which operates via glucagon-like peptide 1 secretion, is upregulated more by fat than by carbohydrate intake in such a way that in a high-fat diet group, each additional T-allele was associated with a reduced loss of fat mass. Also, the Gly482Ser missense mutation of the transcriptional coactivator, peroxisome proliferator-activated receptor- γ coactivator-1 α , participates in insulin functions derangements. A trial investigating the relations of carrying this polymorphism with insulin resistance during an 8-week-long weight loss program and the subsequent weight follow-up found an increased risk of insulin resistance in obese subjects carrying the rs8192673 genotype, which was markedly reduced in such carriers by an energy-restricted diet and sustained at least 1 year after the dietary intervention.

Obesity and diabetic outcomes are mediated by vitamin D–related genetic markers, which were analyzed for interactions after subjects followed weight loss diets varying in macronutrient content. In such an experimental intervention, subjects with the T-allele of DHCR7 rs12785878 has an increase in insulin sensitivity compared with noncarriers when consuming high-protein, weight-lowering diets. The implementation of a genetic risk score tool based on 14 glucose-regulating SNPs indicated that subjects at a higher diabetic risk may have some metabolic advantages by eating a low-fat diet to improve glucose metabolism.

The tumor necrosis factor α (TNF- α) gene is a key regulator of inflammatory processes as well as the control of specific metabolic responses, depending on energy intake. Mutation screening featured a G-> A transition in the promoter region of the TNF- α gene (rs1800629), which was associated with the benefits of circulating triglycerides in obese subjects without the A-allele, whereas a diet enriched in protein exhibited reductions in insulin levels and homeostasis model assessment for insulin resistance only in wild genotype. Adiponectin is another adipose tissue-specific adipokine related to inflammation that often declines in patients with excessive body weight. Carrying the A-allele grants protection from weight regain; this effect was particularly evident 32–60 weeks after the

nutritional intervention, when improvement in GG subjects had disappeared. Another gene expressed in adipose cells is the leptin receptor gene (LEPR), which has been associated with appetite and lipid metabolism control and has an impact on inflammatory and stress states. The AA genotype of the Lys109Arg LEPR gene polymorphism accompanied a higher fat mass loss and total cholesterol decrease compared with the minor allele carriers after they followed a hypocaloric diet. Further research on polymorphic interactions involving two genes related to adipocyte inflammation and adipogenesis, such as interleukin-6 174G>C and PPAR- γ 2 Pro12Ala variants in persons following a body weight management protocol, revealed a synergetic effect of both variants on weight maintenance when they were prescribed a diet to lose weight. In addition, the fibroblast growth factor 21 (FGF21) gene participates in regulating energy and adipose tissue metabolism and also in macronutrient intake choices, which was examined in a 2-year randomized diet intervention trial including four low-calorie diets with different macronutrient distributions. A low-calorie, high-carbohydrate/low-fat diet was favorable in overweight or obese subjects, who were carriers of the carbohydrate intake-reducing allele of the FGF21 variant by improving abdominal adiposity.

CONCLUSIONS

Nutrigenetics is foreseen to be a potent instrument for complementing dietary advice in health care and prevention. It is a useful tool for decreasing public health costs based on a more personalized management of disease by precise dietary advice. Currently analyzed SNPs for nutrigenetic purposes require consensus documents and more investigations with larger populations and different ethnic backgrounds. The cost of the genetic analysis and the price of conjoint personalized advice considering the medical history, anthropometric measurements, biochemical determinations, and dietary and lifestyle factors are affordable. Thus, it is envisaged that new protocols will be implemented to extend the use of nutrigenetics in public health nutrition and clinical dietary practices, although the massive availability of individuals' genetic data requires sophisticated statistical analyses to interpret nutrient–gene interactions, and environmental factors (nutrition and physical activity) need to be included in such analyses. Qualified professionals, who have the responsibility for translating all of the genetic information into nutritional advice and deciding what the best method of assessment, are critical. This scenario of developments would increase the confidence of health professionals regarding the genetic diagnosis. The objective is to transmit personalized advice to consumers and achieve an increase in reliable

demand for these tests to sustain optimal health status. Companies offering genetic tests have an important role. They should transmit as much information as is available to health professionals to help to select the most appropriate data for to each patient with scientific criteria.

Thus, future research considering sequencing and microarray approaches should be oriented on finding new variants related to regulating cell signaling, oxidative stress, and inflammatory and other metabolic pathways, emphasizing the health interplay with environmental factors to achieve precision nutrition. Genetic variants can be complemented by epigenetics, because recent discoveries demonstrated an epigenetic modulation of different metabolic disorders such as obesity, cardiovascular diseases, and type 2 diabetes as well as cancer and some specific processes involved in the Developmental Origins of Health and Diseases (DoHAD).

The search for new genetic polymorphisms in populations with different genetic and epigenetic makeups (i.e., GWAS and EWAS strategies), in addition to advances in analytical methods and bioinformatic techniques, allow an interpretation of specific interactions and relations among genes and dietary factors necessary for characterizing genetic predisposition in complex metabolic diseases and providing precision nutrigenetic recommendations in marketing functional foods specifically devised for personalized nutrition. Precision nutrition should take advantage of available genetic information, but also of as much data as possible, including health and associated profiles as well as consumer and patient interest and ethical or genetic privacy concerns in receiving individualized guidelines designed to reduce disease risk. In the Food4Me study, it is remarkable that the genetic diagnosis was found to be useful as a motivational tool. However, there is a suspected inaccuracy of some DTC genetic tests, due to either insufficient clinical value, or insufficient scientific validity, as well as an unbalanced or biased selection of genetic markers with applicable and validated nutritional interactions, as has been already warned by the Nutrigenomics Organization and the International Society of Nutrigenetics and Nutrigenomics. This apparent lack of evidence regarding some of these products, or even for some of the analyzed SNPs, makes the existence of tools identifying inconsistencies of the products, such as the genetic trial registry, more necessary had been predicted.

Thus, based on the higher or lower predisposition of patients to develop obesity or nutrition-related diseases, nutrigenetics interventions are proposed to be a potential instrument that may complement dietary advice in primary care and prevention. They are a useful tool for reducing public health costs based on the ability to provide personalized management of disease by dietary

advice. Because the SNPs may have differing importance depending on the ethnic background, more research is required in mixed populations. Finally, interpretation of the genetic analyses and the price of personalized advice (always considering the medical history, anthropometric, biochemical, and nutritional and lifestyle factors) must be improved to strengthen the utility of nutrigenetic approaches in public health and individualized clinical interventions. Chapter 73 provides further details on practical approaches for implementing genotype-informed nutrition guidance.

References

- Abete, I., Navas-Carretero, S., Marti, A., Martinez, J.A., 2012. Nutrigenetics and nutrigenomics of caloric restriction. *Prog Mol Biol Transl Sci* 108, 323–346.
- Bouchard, C., Ordovas, J.M., 2012. Fundamentals of nutrigenetics and nutrigenomics. *Prog Mol Biol Transl Sci* 108, 1–15.
- de Toro-Martín, J., Arsenault, B.J., Després, J.P., Vohl, M.C., 2017. Precision nutrition: a review of personalized nutritional approaches for the prevention and management of metabolic syndrome. *Nutrients* 9 (8). <https://doi.org/10.3390/nu9080913> pii: E913.
- Drabsch, T., Gatzemeier, J., Pfadenhauer, L., Hauner, H., Holzapfel, C., 2018. Associations between single nucleotide polymorphisms and total energy, carbohydrate, and fat intakes: a systematic review. *Adv Nutr* 9 (4), 425–453.
- Ferguson, L.R., De Caterina, R., Görman, U., Allayee, H., Kohlmeier, M., et al., 2016. Guide and position of the international society of nutrigenetics/nutrigenomics on personalised nutrition: Part 1 - fields of precision nutrition. *J Nutrigenetics Nutrigenomics* 9 (1), 12–27.
- Fitó, M., Melander, O., Martínez, J.A., Toledo, E., Carpéné, C., Corella, D., 2016. Advances in integrating traditional and omic biomarkers when analyzing the effects of the mediterranean diet intervention in cardiovascular prevention. *Int J Mol Sci* 2 (9), 17.
- Goni, L., Cuervo, M., Milagro, F.I., Martínez, J.A., 2016. Future perspectives of personalized weight loss interventions based on nutrigenetic, epigenetic, and metagenomic data. *J Nutr* 9 pii: jn218354.
- González-Muniesa, P., Martínez-González, M.A., Hu, F.B., Després, J.P., Matsuzawa, Y., Loos, R.J.F., Moreno, L.A., Bray, G.A., Martinez, J.A., 2017. Obesity. *Nat Rev Dis Primers* 3, 17034.
- Grimaldi, K.A., van Ommen, B., Ordovas, J.M., Parnell, L.D., Mathers, J.C., et al., 2017. Proposed guidelines to evaluate scientific validity and evidence for genotype-based dietary advice. *Genes Nutr* 12, 35.
- Kohlmeier, M., De Caterina, R., Ferguson, L.R., Görman, U., Allayee, H., Prasad, C., Kang, J.X., Nicoletti, C.F., Martinez, J.A., 2016. Guide and position of the international society of nutrigenetics/nutrigenomics on personalized nutrition: Part 2 - ethics, challenges and endeavors of precision nutrition. *J Nutrigenetics Nutrigenomics* 9 (1), 28–46.
- Livingstone, K.M., Celis-Morales, C., Papandonatos, G.D., Erar, B., Florez, J.C., et al., 2016. FTO genotype and weight loss: systematic review and meta-analysis of 9563 individual participant data from eight randomised controlled trials. *Br Med J* 20, 354.
- Martinez, J.A., Navas-Carretero, S., Saris, W.H., Astrup, A., 2014. Personalized weight loss strategies-the role of macronutrient distribution. *Nat Rev Endocrinol* 10 (12), 749–760.
- Melhorn, S.J., Askren, M.K., Chung, W.K., Kratz, M., Bosch, T.A., et al., 2018. FTO genotype impacts food intake and corticolimbic activation. *Am J Clin Nutr* 107 (2), 145–154.

- Ramos-Lopez, O., Milagro, F.I., Allayee, H., Chmurzynska, A., Choi, M.S., et al., 2017. Guide for current nutrigenetic, nutrigenomic, and nutriepigenetic approaches for precision nutrition involving the prevention and management of chronic diseases associated with obesity. *J Nutrigenetics Nutrigenomics* 10 (1–2), 43–62.
- San-Cristobal, R., Milagro, F.I., Martínez, J.A., 2013. Future challenges and present ethical considerations in the use of personalized nutrition based on genetic advice. *J Acad Nutr Diet* 113 (11), 1447–1454.

Relevant Websites

dbSNP: <http://www.ncbi.nlm.nih.gov/SNP>.

GTR: <https://www.ncbi.nlm.nih.gov/gtr/>.

GWAS catalog: <https://www.ebi.ac.uk/gwas/>. This page intentionally left blank

Genetically Informed Nutrition Guidance

Martin Kohlmeier

Human Research Core and Nutrigenetics Laboratory, UNC Nutrition Research Institute; and Schools of Medicine and Public Health, Department of Nutrition, Kannapolis, NC, United States

Glossary

- DRI** Dietary Reference Intakes are recommendations for intakes of macronutrients and essential nutrients, defining amounts in terms of both requirements and excess
- EAR** Estimate of average requirement used to calculate typical dietary needs
- PONG** Precision Online Nutrition Guidance tool collates individual information, generates a personalized nutrition report, and suggests meal plans
- RDA** Recommended Dietary Allowance intended to cover 97%–98% of dietary needs of a healthy reference population

PRACTICE FRAMEWORKS FOR NUTRITION COUNSELING

Current nutrition counseling is governed by a patchwork of laws, regulations, and voluntary guidelines that differ among countries and often even within the same country.

In the United States, the strongest element of regulating practice is professional licensing. In most states, nutrition counseling requires a valid license, usually as a Licensed Dietary Nutritionist. Weight loss counseling is governed by a partially overlapping regulatory framework requiring a valid license. Licensed physicians, osteopaths, and some other health professionals are also entitled to provide nutritional counseling. The assumption here is that trained and licensed health professionals will know what needs to be done. Dietitians have to complete continuing professional education as part of their self-designed continuing study program. Physicians and other health professionals only have to keep up with advances in general medicine and their board specialties, which in most cases do little to address nutrition counseling skills.

In the United Kingdom and some other countries, the actual practice of Licensed Dietitians and Nutritionists is

supervised slightly more; it is mainly enforced through reprimands, ultimately removing the license to practice.

The knowledge base on which nutrition counseling practice rests in the United States consists of best practice information, some professional guidelines, and government recommendations. The latter include the Dietary Guidelines for Americans, newly issued every 5 years, and the Dietary Reference Intakes (DRI) issued by the Food and Nutrition Board of the Institute of Medicine. Although originally intended, the DRI are not updated with regularity owing to persistent lack of funding. A triage approach is used to identify the nutrients most in need for updates ([Brannon et al., 2016](#)), but even this is hampered by fiscal limitations.

The DRIs are limited by both their insufficient specificity and their excessive specificity. The insufficient specificity is largely because the intake targets are intended to minimize the risk that individuals get either too little or too much ([DRI: Applications in Dietary Planning, 2003](#)). Although the recommendations distinguish among 22 population categories, they are still set up for entire groups without recognizing the great diversity among individuals, for instance by weight, body composition, or lifestyle. On the other hand, they prescribe with much greater specificity Recommended Dietary Allowance (RDA) or Adequate Intake levels than cited evidence supports.

Individual differences are explicitly ignored by using inappropriate models to estimate RDA levels ([Kohlmeier, 2013](#)). The current approach assumes that within the 22 categories bounded by gender and age, actual individual intake requirements follow a normal distribution that can be approximated by an estimate of average requirement (EAR) and an estimate of the standard deviation (SD). The current version of the DRI for folate describes in great detail that the EAR was set by relying mainly (with some other evidence considered)

on a study of five women who were fed 200, 320, and 500 μg dietary folate equivalents (DFE)/day and had homocysteine (Hcys) and folate concentrations monitored (O'Keefe et al., 1995). Three of the five women had elevated Hcys with the 320- μg diet and all had normal Hcys with the 500- μg diet. Genotypic variation was not known or was known. This convinced the committee that 320 μg accurately estimated the EAR, not 280 or 360 μg , which would also have been compatible with available data. When adding 20% of the EAR (because the SD was unknown, it was assumed to be 10%), they arrived at an RDA of 400 μg (not 384 μg , as the prescribed calculation would suggest). The added irony is that the most common genetic variant with an impact on dietary requirements (gender) was elided by this evaluation based largely on data in women. For that matter, folate RDA came out the same for all age groups over 14 years. This addresses the complaint about excessive specificity that suggests a level of precision not supported by the cited evidence.

The standard DRI model does not allow for genotype-specific differences even when they are well-understood. For example, individuals with the common methylenetetrahydrofolate reductase (MTHFR) rs1801133 (MTHFR 677C>T) genotypes CC, CT, and TT differ greatly in terms of the dietary intakes needed to achieve the Hcys concentration in blood considered to reflect adequate folate nutrition (Ashfield-Watt et al., 2002). These differences in Hcys are to be expected because the allozymes encoded by the two alleles have different specific activities (Martin et al., 2006). Based on currently available data, it is obvious that carriers of the TT genotype need more than 800 μg DFE, whereas those with the CT and CC genotypes are well-served with 400 μg . In other words, current DRI guidelines do not meet their self-defined goal and health professionals who rely on them will fail the 10%–20% of their patients with the MTHFR 677 TT genotype. This must not lead to the conclusion that everybody should take a high-dose folate supplement because the resulting folate excess has a significant potential for increased cancer risk (Ericson et al., 2009) and other harm.

INDIVIDUALLY TAILORED NUTRITION GUIDANCE

Ideally, we should use as much information as practically possible to shape what an individual should eat and drink (Fig. 66.1). To calculate a reasonable estimate of energy requirement, at a minimum, we need to consider gender, age, height, weight, and physical activity. In practice, we tend to evaluate body size and note that individuals with excess weight consume too many

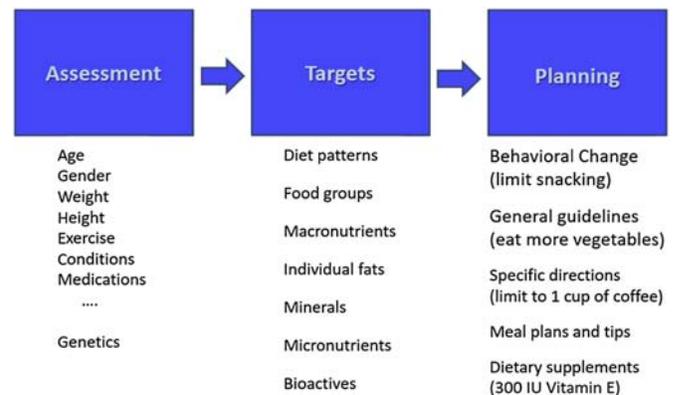


FIGURE 66.1 Stepwise progression from assessment of current circumstances of a client to estimating appropriate nutrition target, and ultimately providing dietary guidance.

calories, whatever that number may be, and likewise with evident underweight. This is not so easy for minerals, micronutrients, and bioactives. We cannot look at clients and conclude that their folate intake is too low. Of course, we will ask about their consumption of major folate sources (dark green vegetables, oranges, and legumes) or even conduct a comprehensive evaluation. However, we will still not know how much of those folate-rich foods they need. There is the possibility of measuring a marker of folate status, such as Hcys or folate concentration in blood and using repeated measure to learn with what kind of meal plan they get enough. This approach is already difficult to do for investigational purposes, and usually it is not an option in practice. Besides, many different nutrients would have to be evaluated.

The availability of genotyping can overcome this difficulty, at least in part (Kohlmeier et al., 2016). One attractive feature of genetic information is that it never changes and a single analysis lasts for a lifetime (of course, this also means that this potentially sensitive information has to be carefully guarded for an equally long time).

The development of dietary targets should be more broadly conceived. The nutrition professional needs to decide what fits the individual's dietary needs best: what kind of dietary pattern (Mediterranean, Mind, vegan, etc.), what food groups to include or exclude (more dark green vegetables, gluten-free foods, pork avoidance, etc.), macronutrient balance (low-carbohydrate, hypocaloric, etc.), individual nutrient amounts, an array of bioactives (docosahexaenoic acid, eicosapentaenoic acid, lycopene, flavonoids, etc.), and also of dietary supplements as needed. Genetic information can help with setting targets in all of these groups.

The final challenge is then the actual counseling. This should include general advice on behaviors (limit snacking, have breakfast, avoid late-night meals) and general

guidelines (eat more vegetables, avoid salty foods, limit saturated fat intake), all the way to very specific suggestions, detailed meal plans, and what dietary supplements do or do not fit that individual's profile.

Composing a pleasing and well-tolerated concert of food combinations that meets all of an individual's requirements, preferences, and intolerances is difficult. Combining a moderate number of targets usually can be done by experienced professionals. Getting the pre-specified number of calories with enough protein and vegetables and not too much saturated fat or sodium is one of the easier exercises. But adding to the mix the need to get that extra amount of folate or to avoid too much iron because of an inborn tendency to accumulate iron stretches our skills. The main problem is that it takes a lot of time to put together sufficiently specific suggestions and meal plans for people who are all different. At that moment, we wish for some computer support.

COMPUTER-BASED MEAL PLANNING

Nutrition guidance programs are intended to help users consume the amounts of foods and critical nutrients that are likely optimal for them. Such a program was developed for use with computers, portable devices, and smartphones. Individual targets are estimated based on gender, age, and calculated energy requirement. The program then uses parabolic distance functions with variable-specific weights and adjustments to calculate the goodness of fit for a large number of food combinations. The user can review all daily menus

that are a good fit (Fig. 66.2). The selections present specific foods, products, dishes, and beverages as needed. Upon request, the program displays suitable alternatives for individual items and allows users to incorporate favored foods to the greatest extent possible. Disliked foods can also be deleted and additional foods added. The program always indicates instantly which foods fit the user's profile best. The user can also explore for some dishes external links to popular cooking sites and obtain detailed ingredient lists and cooking instructions. Favored daily menu plans can be printed out to support meal planning and purchasing of ingredients.

This type of easy-to-use nutrition guidance and meal planning program is essential for unlocking the potential of nutrigenomics in practice. Both health professionals and the public can benefit from the decision support that offers food selection choices most people can understand. Various additional functions can increase the utility of the tool, such as optimizing food cost, reduce carbon dioxide footprint, and avoiding unwanted allergens and ingredients.

References

- Ashfield-Watt, P.A., Pullin, C.H., Whiting, J.M., Clark, Z.E., Moat, S.J., Newcombe, R.G., Burr, M.L., Lewis, M.J., Powers, H.J., McDowell, I.F., 2002. Methylenetetrahydrofolate reductase 677C>T genotype modulates homocysteine responses to a folate-rich diet or a low-dose folic acid supplement: a randomized controlled trial. *Am J Clin Nutr* 76 (1), 180–186.
- Brannon, P.M., Weaver, C.M., Anderson, C.A., Donovan, S.M., Murphy, S.P., Yaktine, A.L., 2016. Scanning for new evidence to prioritize updates to the Dietary Reference Intakes: case studies for thiamin and phosphorus. *Am J Clin Nutr* 104 (5), 1366–1377.
- DRI: Applications in Dietary Planning, 2003. Chapter 2, Using Dietary Reference Intakes in Planning Diets for Individuals. National Academies Press, DC.
- Ericson, U.C., Ivarsson, M.I., Sonestedt, E., Gullberg, B., Carlson, J., Olsson, H., Wirfält, E., 2009. Increased breast cancer risk at high plasma folate concentrations among women with the MTHFR 677T allele. *Am J Clin Nutr* 90 (5), 1380–1389.
- Kohlmeier, M., 2013. Nutrigenetics. Applying the Science of Personal Nutrition. Academic Press, San Diego, ISBN 9-7801-2385-9-006, 386 pp.
- Kohlmeier, M., De Caterina, R., Ferguson, L.R., Görman, U., Allayee, H., Prasad, C., Kang, J.X., Nicoletti, C.F., Martinez, J.A., 2016. Guide and position of the International Society of Nutrigenetics/nutrigenomics on personalized nutrition: Part 2: ethics, challenges and endeavors of precision nutrition. *J Nutrigenetics Nutrigenomics* 9, 28–46.
- Martin, Y.N., Salavaggione, O.E., Eckloff, B.W., Wieben, E.D., Schaid, D.J., Weinshilboum, R.M., 2006. Human methylenetetrahydrofolate reductase pharmacogenomics: gene resequencing and functional genomics. *Pharmacogenetics Genom* 16, 265–277.
- O'Keefe, C.A., L.B., Bailey, Thomas, E.A., Hofler, S.A., Davis, B.A., J.J., Cerda, Gregory, J.F. 1995. 3rd. Controlled dietary folate affects folate status in nonpregnant women. *J Nutr* 125 (10), 2717–2725.

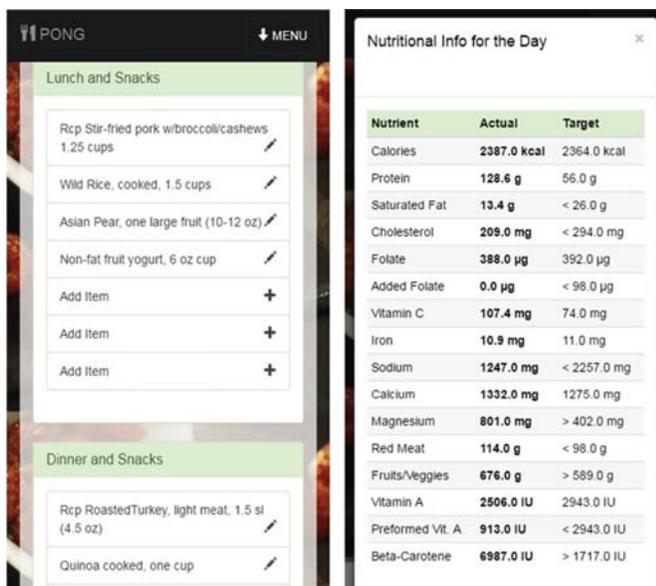


FIGURE 66.2 Precision Online Nutrition Guidance program offers individually tailored meal plans that users can modify to meet their individual needs and preferences.

This page intentionally left blank

Nutrigenomics of Food Pesticides

Laura Bordoni, Rosita Gabbianelli

Unit of Molecular Biology, School of Pharmacy, University of Camerino, Camerino, Italy

Glossary

3-PBA 3-Phenoxybenzoic acid is the biomarkers of pyrethroid pesticide exposure discoverable in the people's urine.

ADI Acceptable Daily Intake is "defined as the amount of specific substance that can be ingested on a daily basis without an appreciable health risk" (WHO, 1987).

CYP Cytochrome P450 describes a group of enzymes whose main function is the oxidative catalysis of several exogenous and endogenous substances.

GST Glutathione S-transferase catalyzes the reaction of glutathione with electrophiles of both endogenous and xenobiotic origins contributing to detoxification and protection against oxidative stress.

HI The Hazard Index is used to assess the cumulative chronic health risk of exposure to a group of pesticides; it defines the consumer risk in food.

LEARn Latent Early life Associated Regulation.

MRLs Maximum residue levels are the upper levels of pesticide residues that are legally permissible in food or animal feed, based on good agricultural practice and the lowest consumer exposure necessary to protect vulnerable consumers.

NOAEL No Observed Adverse Effect Level is the dose that gives no toxic effect. It is extrapolated by chronic neurotoxicity studies done performed with animals.

PON Paraoxonase is a group of proteins present in three forms (PON1, PON2, and PON3) encoded by genes *PON1*, *PON2*, and *PON3*.

PUFAs Polyunsaturated fatty acids.

workers exposed to pesticides revealed that these compounds can be responsible for adverse effects. The acute toxicity of pesticides is well-known and their use is regulated on the basis of maximum residue levels (MRLs), which are established by considering, among others, the acceptable daily intake (ADI) for any particular pesticide in humans. Nevertheless, uncertainties remain concerning the chronic and long-term effects of exposure to pesticides.

Delayed health effects attributed to pesticide exposure range from general malaise to chronic and long-term severe effects on the nervous system, including cognitive and psychomotor dysfunctions, mild cognitive dysfunctions (e.g., mood changes, neurobehavioral changes), depression, minor psychiatric morbidity, neurodegenerative (e.g., Parkinson and Alzheimer diseases), and neurodevelopmental effects. Reproductive functions can also be affected, with impaired fecundity, infertility, birth defects, and altered growth.

Exposure to deceptively subtoxic doses of pesticides during a developmental period such as early life is a critical time window of particular concern. Numerous studies showed that developing fetuses, newborns, infants, and developed children are highly susceptible to late functional toxicity, which becomes manifest in adult life as the result of early-life exposure. In particular, developmental functional toxicity may affect the central nervous system, but effects have also been observed on the reproductive, endocrine, and immune systems. These observations suggest that professionals should know and consider these aspects to preserve health through good nutritional practices.

NONNUTRITIVE CHEMICALS IN FOODS

Organizations in charge of food safety across the world constantly analyze levels of chemical residues (i.e., pesticides, insecticides, fungicides, herbicides) in

INTRODUCTION

Pesticides are chemicals used to protect crops against noxious or unwanted living species such as insects, weeds, fungi, and other pests; they have a significant role in food production. Pesticides protect or increase yields and the number of times per year a crop can be grown on the same land, which is particularly important in countries facing food shortages. Nevertheless, pesticides have caused increasing concerns for human health, and their potentially harmful effects have been extensively studied over past decades. Both animal models of pesticide toxicity and observational studies on

foods. The United States, Peoples' Republic of China, and European Union have different laws and screening programs aimed at controlling and defining acceptable levels of chemicals that should have no toxic effects. A report on pesticide residues in 5989 samples of foods published by the US Food and Drug Administration stated that over 98% of domestic and 90% of imported foods were compliant with federal standards. No pesticide chemical residues were identified in 49.8% of domestic and 56.8% of imported human food samples analyzed (<https://www.fda.gov/downloads/Food/FoodborneIllnessContaminants/Pesticides/UCM582721.pdf>). However, the ToxCast high-throughput screening program discovered that foods in the United States contained several chemical compounds that can be classified into four groups: chemicals added to foods for functional purposes (556 direct food additives); chemicals that may migrate to food from packaging, processing, or clinical chemicals (339 indirect additives); pesticide residues (406 compounds); and nonfood (319 compounds) (Karmaus et al., 2017). A list of these compounds is available online at <http://ilsina.org/curation-of-food-relevant-chemicals-in-toxcast>. Use of these compounds is allowed by the Food and Drug Administration even though about 70% of these additives have no experimental data to clearly demonstrate their safety for human health. Chemicals included in the ToxCast classification were analyzed for their capacity to induce various effects, such as interactions with receptors, enzyme inhibition, induction of stress, and cytotoxicity. Within the pesticide group, all pesticide residues registered for food use in the United States, as well as pesticides not included in the accepted list in the United States but used elsewhere, were included. Of particular note is that among the nonfood group, 219 compounds were detected and recognized as pesticides not actually been demonstrated as being toxic, or that have nonfood crop uses. Among the nonfood compounds, sunscreen ingredients, insect repellents, anticoagulants, and anti-neoplastic and antipsychotic drugs have been identified.

The European Food Safety Authority evaluated the presence of pesticides in food within MRLs allowed in European Union legislation. The European Food Safety Authority report revealed that more than 97% of foods analyzed in Europe contain the legal level of pesticide residues, 96.5% of foods for infants and young children were free of residues, and 99.3% of organic foods were residue-free or within the permitted limits (<https://www.efsa.europa.eu>). Currently, 385 substances are authorized in the European Union, 26 of which are also accepted in organic agriculture.

The risk assessment for food pesticides should consider not only the presence of chemicals above the authorized limit but also the presence of multiple kinds of pesticides, each within the MRLs permitted by the

legislation. In this case, the amplified effect of a mixture of pesticides should be considered, because the final biological effect is likely to reflect the sum of all chemical residues contained in food. Of particular relevance is the observation that many pesticide metabolites are as active as their parental compound and that the same class of pesticides results in the same kind of metabolites.

Another key point for the controlled assessment of pesticide residues in food is related to the analytical method used in their extraction and detection, which can lead to significant differences in the evaluation of MRLs. The analytical recovery and final quantification can vary by 80% of the total, depending on the solvent used (i.e., ethyl acetate, dichloromethane, cyclohexane).

In Asia, the wide presence of pesticides in food is of concern, as demonstrated by the significant increase in published articles on this topic in past couple of decades. The governing bodies are working to inform farmers about the need to regulate the quantity and quality of chemicals required to control crop production. As reported in samples collected in other geographic areas, chemicals can also be identified in organic products: 15 different pesticides were identified in samples of cabbage, celery, lettuce, broccoli, mustard, spinach, and cauliflower from organic farms in the Cameron Highlands, Malaysia (Farina et al., 2017). Among these vegetables, samples without detectable residues ranged from 6.6% to 30%, samples containing residues less than or at MRLs ranged between 60% and 90%, whereas 4%–13.3% contained residues at MRLs or higher. Only one sample among the 109 vegetables analyzed contained no residue over the MRLs.

The Hazard Index (HI) has been used to assess the cumulative chronic health risk related to the exposure to a group of pesticides contained in foods for humans. This index is useful in screening the content of chemical residues in organic, conventional, and imported foods, and can provide a quantification of residues in different diets even when the pesticide level is below the no observed adverse effect level (NOAEL). The HI has been used to predict the cumulative effects of pesticides with different target organs. Although HI is useful for monitoring food quality, screening of biological markers of exposure is a major goal of health investigations in this area, because the metabolite residues of chemicals in urine, plasma, and blood adequately describe individual metabolic responses to pesticides according to one's own genome.

BIOMARKERS OF PESTICIDE EXPOSURE IN HUMANS

Biomarkers of pesticide exposure are mainly discoverable in the urine (McKelvey et al., 2013). For example,

the wide use of pyrethroids is demonstrated by the presence of 3-phenoxybenzoic acid (3-PBA) in people's urine in China, the European Union, and the United States. The concern is linked to the observation that this metabolite is higher in children than in adults. Pyrethroids and their common metabolite, 3-PBA, can cross the blood–brain barrier and, owing to their lipophilicity, can be stored in the long term in the brain. Together with pyrethroids, organophosphates are the most common pesticides used in Europe. Their presence can be monitored in people's urine owing to the presence of the dialkyl phosphate metabolites.

Of particular concern are data obtained from the Center for the Health Assessment of Mothers and Children of Salinas cohort in the United States regarding the association between increased levels of organophosphate exposure biomarkers in the urine of farm worker mothers and abnormal mental development in their children. Dialkyl phosphates and 3-PBA were increased in children's urine affected by attention-deficit hyperactive disorder; the risk for developing attention-deficit hyperactive disorder increased by 55% with a 10-fold increase in urine dialkyl phosphates, and the risk was twice as high in children with detectable concentrations of 3-PBA. A metaanalysis of articles published between 1979 and 2016 supported an association between residential exposure to pesticides and the development of childhood brain cancer. The Endocrine Disruptors : Longitudinal Study on Pregnancy Anomalies, Infertility, and Childhood cohort in France associated urine 3-PBA levels with defects in verbal and memory functions in 6-year-old children. The social cost of this decrease in intelligence quotient levels in children exposed to organophosphates has been estimated at about €125 billion.

Constructive criticism underlines the relevance of statistical heterogeneity in these results, suggesting that further strategies are mandatory to improve knowledge regarding the link between markers of exposure and health effects. Furthermore, because of the various types and toxicity of pesticides, the development of assays aimed at identifying the cumulative risk for dietary chemical exposure in biological fluids (i.e., urine, plasma) should be a priority to establish the link between biomarkers of chemical exposure and diseases.

INDIVIDUALS' RESPONSES TO PESTICIDES

Exposure to the same environment does not always have the same effect on different subjects. On the contrary, important interindividual variability has emerged in responses to chemicals. Genetic differences in food tolerance, taste preferences, nutrient absorption, and metabolism all potentially affect the influence of diet

on disease risk. Dietary constituents of foods, including both micronutrients and non-nutrients, act as inducing agents of metabolizing enzyme systems through several molecular mechanisms. In addition to direct dietary inducers, several nutritional factors (i.e., protein intake, obesity, fasting) have been shown to influence the activity of oxidation enzymes. Because diet is a combination of mutagens and protective agents that are all metabolized by detoxification enzymes, genetic variations in their gene structure, considered in the context of differential environmental exposures, may account for individual variation in disease risk. The identification of single nucleotide polymorphisms in human populations exposed to different environmental hazards has an essential role in detecting the genetic risks for inducing several important human diseases. For this purpose, numerous studies identified an association between single nucleotide polymorphisms in family of genes, such as those of paraoxonase (PON), cytochrome P450 (CYP), or glutathione S-transferases (GSTs), and pathologic conditions related to long-term exposure to pesticides (Lampe, 2007).

For example, genotypes and consequent enzyme activities of human PON1 are associated with increased susceptibility to pesticide-related damage and oxidative stress-related health conditions (e.g., neurodegenerative disorders, cardiovascular diseases, diabetes, and obesity). In particular, several studies suggested the existence of a relation between PON1 polymorphisms and low PON1 protein levels as well as adverse developmental and cognitive behavior in children. Moreover, it has been shown that progeny of mothers with susceptible genotypes are more prone to toxicity related to prenatal exposure to organophosphates.

CYP and *N*-acetyltransferase-2 are other xenobiotic-metabolizing enzyme systems, the genetics of which has been studied for susceptibility to DNA damage resulting from pesticide exposure. DNA damage in workers exposed to organophosphate pesticides was higher in people with one particular allele compared with those with a different allele. Studies have shown associations between CYP1A1-PON1 polymorphisms and several diseases in individuals occupationally exposed to polychlorinated biphenyls.

GSTs are other enzymes involved in detoxification processes. Prenatal exposure to pesticides in children of mothers with GSTM1 and GSTT1 alleles, which result in a lack of GSTM1 and GSTT1 proteins, respectively, exposes them to a greater risk for fetal growth restriction. Furthermore, polymorphisms in the GSTM1 and GSTT1 genes were associated with a higher risk for cancer, which, by contrast, is strongly reduced with the intake of cruciferous vegetables. Thus, it is clear that an interaction exists between susceptibility to xenobiotics and diet, and that further knowledge of

susceptibility based on genetic factors could target dietary interventions for individuals most likely to benefit.

MODULATION BY PESTICIDES OF GENE EXPRESSION AND EPIGENETICS

Food pesticides contribute to the human exposome, which is one of the most promising research areas for the therapy and prevention of disease. The exposome is defined as “the cumulative measure of environmental influences and associated biological responses throughout the life span, including exposures to the environment, diet, behavior, and endogenous processes” (Dennis and Jones, 2016). Thus, exposure to contaminants is emerging as related to the contaminant content in foods and as having a strong impact on health: the presence of certain chemicals per se has a critical role in health, but also how those molecules alter our biology and how resilient the organism is in maintaining dynamic homeostasis.

Environmental exposure can be responsible for several potential effects, such as DNA mutation, adducts, epigenetic changes or alterations in enzyme functions, or damage induced through reactive oxygen or nitrogen species. Furthermore, environmental exposure can cause structural changes as well as changes in gene expression profile in exposed tissues. Numerous studies showed that lipophilic pesticides can accumulate in adipose tissue, exerting an obesogenic effect through subtle changes in gene expression and tissue organization. Transcriptomic analysis also revealed changes in gene expression in animal brains exposed to pesticides, and linked changes in the expression patterns of messenger RNAs with neurodegeneration induced by the exposure.

Moreover, besides adductomic and metabolomic approaches, increasing importance is being attributed to evidence in the literature that environmental factors may cause diseases through changes in gene expression, in turn mediated by epigenetic mechanisms such as DNA methylation, histone modifications, and chromatin remodeling, all of which are potentially implicated in disease triggering induced by numerous environmental factors (Collotta et al., 2013). A few examples are exposure in early life to subtle concentration of permethrin, which can lead to global DNA methylation changes; levels of persistent organic pollutants in the blood, which were inversely associated with global DNA methylation; exposure to dichlorodiphenyltrichloroethane, which was linked to altered DNA methylation patterns in the rat hypothalamus; paraquat, which can induce histone H3 acetylation and decrease histone deacetylase activity; and dieldrin-induced proteasomal dysfunction, which can result in the accumulation of

histone acetyltransferases. Whenever exposure to dieldrin is prolonged, it induces histone hyperacetylation in the corpus striatum and *substantia nigra* in mice.

Furthermore, among many other factors, both genetic and epigenetic determinants contribute to the final biological response. PON1 expression is an example of how both of these aspects regulate responses to xenobiotics. Even when PON1 expression is strongly regulated by genetics (see the previous discussion), other factors such as epigenetic changes are involved in controlling PON1 enzyme variability. This occurs because the *PON1-108* promoter polymorphism is associated with methylation of particular CpG sites, which suggests that DNA methylation can be mediate PON1 genetics and its expression. These findings are an example of how the integration of genetic, epigenetic, and expression data could clarify functional mechanisms involved in susceptibility to xenobiotics, and that the totality of genetic background and environmental exposures has to be considered in evaluating disease prevention strategies at the individual and population levels (Declerck et al., 2017). Advances in transcriptomics, genomics, and epigenomics thus provide insights into susceptibility to diseases, increasing our understanding of the molecular mechanisms behind their onset. Furthermore, these advances help to identify new kinds of biomarkers, particularly referring to epigenetic changes, and provide information about exposures, but also biological effects, thus offering an important set of tools in the area of personalized medicine and prevention.

PESTICIDES AS EARLY DETERMINANTS OF LATE-ONSET DISEASES

Because of all of the previously listed complex molecular mechanisms through which pesticides can interact (including those contained in foods), numerous diseases have been related to exposure to environmental toxicants. Particularly when taking place in the specific window of plasticity, they may contribute to the occurrence of adverse birth outcomes, neurodevelopmental deficits, increased risk for cancer, and other multifactorial diseases such as asthma and diabetes. Furthermore, not only the generation directly exposed but also its offspring can be affected through mechanisms of epigenetic memory. These revolutionary discoveries that have emerged from both epidemiological and in vivo studies emphasize the importance of considering the subtle effect of pesticides on health. For example, the epigenetic effect of bisphenol A was originally shown in viable yellow mice by decreasing CpG methylation upstream of the *Agouti* gene, and maternal diet supplementation with methyl donors such as folic acid or the phytoestrogen genistein, was shown to prevent the

hypomethylating effect of the toxicant. In addition, epigenetic effects of bisphenol A and phthalates were also demonstrated and linked to complex diseases such as cancer and diabetes in humans. Moreover, pesticides such as dieldrin, paraquat, maneb, and permethrin were shown to lead to Parkinson-like disease through dysfunction of the nigrostriatal dopaminergic system and/or abnormalities in motor response in animal studies, which proves that early-life exposure to these chemicals can exert adverse effects later in life. Furthermore, Skinner and colleagues demonstrated that exposure of gestating rats to methoxychlor increases the incidence of kidney disease, ovary disease, and obesity in offspring, spanning three generations. Thus, nutrition can influence epigenetic homeostasis by providing substrates, cofactors for epigenetic reactions, and epigenetically active molecules, but also by being a source of exposure to exogenous molecules able to interfere with this equilibrium (Skinner, 2016).

Several different models have been hypothesized to explain the role of epigenetics in disease onset. Barker hypothesized that adult diseases are consequences of fetal adverse conditions, because the fetus adapts to a certain environment determined by environmental stimuli in early life. Adaptive responses may be in the form of metabolic changes, hormonal release, or sensitivity of the target organs to hormones, all of which then affect the development of target organs with no immediate consequences to the newborn, but with later disturbances in physiologic and metabolic functions. Gluckman and Hanson suggested that when the fetus is exposed to stress or adverse conditions, immediate reversible changes occur. However, when stress conditions are prolonged, changes become irreversible and able to persist throughout life, influencing the individual persistently in adulthood. They defined this phenomenon using the term "predictive adaptive response." Effects of these irreversible changes on the fetus in the long run may be favorable or unfavorable. The Developmental Origin of Health and Disease model postulated that not only embryonic stages but also the period of development during infancy is responsible for the late-life risk for disease. A further evolution of these models aiming to explain the role of environmental factors in disease etiology, with a particular focus on neurodegeneration is represented by the Latent Early life Associated Regulation (LEARn) model. This theory states that environmental agents, including nutrition, metal exposure, head traumas, and lifestyle, are hits related to the cause and progression of sporadic neurodegenerative disease onset. LEARN differs from other models in that it is neither an acute nor a chronic model, but is based on the idea that latent epigenetic changes induced in early life, which result in no disease symptoms, create a perturbation in the epigenome, causing

manifest consequences later in life after the occurrence of a second triggering agent. The period between the epigenetic perturbation and the second triggering event is called the latency period, and genes that respond late in relation to early-life responses are called LEARNed genes. The process of responding to the early-life environmental triggers after the long latency period is termed LEARNing.

NUTRIGENOMICS AS A STRATEGY FOR BIOREMEDIATION

The sensitivity of populations to the consequences of food pesticides is increasing, as demonstrated by the growth of the organic retail market by 107% between 2006 and 2015 in the European Union; it reached a total of €27 billion. Generally, people consuming organic food have a healthy lifestyle, and their food is rich in fresh fruits and vegetables useful for the presence of bioactive compounds able to modulate gene expression beneficially. A metaanalysis of 343 peer-reviewed publications underlined the presence of higher levels of antioxidants and lower concentrations of cadmium in organically grown crops compared with conventionally grown ones (Barański et al., 2014). A metaanalysis of 170 articles related to organic milk composition revealed that polyunsaturated fatty acids (PUFAs), specifically n-3 PUFAs, were significantly higher, by 7% and 56%, respectively, compared with conventional milk, whereas no change in saturated fatty acids and monounsaturated fatty acids was observed. Furthermore, the concentrations of conjugated linoleic acid, α -tocopherol, and iron increased. Another metaanalysis of 67 articles highlighted an increase in PUFA and n-3 PUFA, by 23% and 47%, respectively, in organic meat compared with conventional meat, which underlines that these parameters strongly depend on differences among animal species providing the source of meat.

The consumption of organic food is associated with a lower occurrence of allergy and atopic diseases in children. In particular, a reduction by 36% in the risk for eczema in 2-year-old infants was demonstrated when their mothers (n = 2700) consumed organic food during pregnancy and breastfeeding. A study of 62,000 participants consuming organic food demonstrated a 31% decrease in body mass index compared with those who consumed conventional foods. Another study among 630,080 middle-aged United Kingdom women reported a reduction in the risk for non-Hodgkin lymphoma in subjects consuming more organic foods compared with the control group. The use of HI in Sweden was proposed to quantify the exposure to toxicants owing to the consumption of 500 g fruits and vegetables divided into three groups: imported conventional,

domestic conventional, or organically produced foods. The HI indicated that exposure to toxicants was about 70 times lower when organic food was consumed.

According to this evidence, organic food may be a strategy for increasing the level of bioactive compounds able to modulate gene expression favorably and also reduce the exposure to pesticides. An alternative strategy might be to consider using selected bioactive food or supplements; attention should be given to the molecular mechanisms of pesticide damage. To partially counterbalance adverse effects, dietary agents with antioxidant and antiinflammatory properties, as well as bioactive components able to mediate epigenetic modifications, might be useful. As reported in experimental animal models, coadministration of low-dose pesticides and antioxidants significantly decreased parameters associated with oxidative stress caused by pesticide exposure. In particular, the moderate use of vitamin C, glutathione, and vitamin E can favorably modulate the expression or activity of transcription factors such as nuclear factor- κ B, Nrf2, and cytokines.

Green tea polyphenols, black raspberries, soy isoflavones, curcumin, apple/coffee polyphenols, broccoli isothiocyanates, vitamin A, selenium, lycopene, folate, and vitamins B12 and B6 work as preventative agents modulating DNA methylation and preventing aberrant promoter hypermethylation or genome-wide hypomethylation for cancer and chronic diseases. Sulforaphane from broccoli, diallyl disulfide from garlic, and butyrate from fiber by gut microbiota are histone deacetylase inhibitors useful for controlling inflammation. Studies show that consumption of 105 mg sulforaphane contained in 68 g broccoli sprouts or 570 g mature broccoli can significantly inhibit histone deacetylase activity after 3 h. Sulforaphane has also been demonstrated to be useful in promoting detoxifying enzymes required to metabolize pesticides.

Despite these outcomes, additional epidemiological studies are required on the efficacy of organic food or supplements to contrast pesticide-induced damage, together with an evaluation of pesticide residue metabolites in biological fluids that provide data on human exposure and the individual genetic capacity to metabolize these chemicals.

References

- Barański, M., Srednicka-Tober, D., Volakakis, N., Seal, C., Sanderson, R., Stewart, G.B., Benbrook, C., Biavati, B., Markellou, E., Giotis, C., Gromadzka-Ostrowska, J., Rembiałkowska, E., Skwarło-Sońta, K., Tahvonon, R., Janovská, D., Niggli, U., Nicot, P., Leifert, C., 2014. Higher antioxidant and lower cadmium concentrations and lower incidence of pesticide residues in organically grown crops: a systematic literature review and meta-analyses. *Br J Nutr* 112 (5), 794–811.
- Collotta, M., Bertazzi, P.A., Bollati, V., 2013. Epigenetics and pesticides. *Toxicology* 307, 35–41.
- Declerck, K., Remy, S., Wohlfahrt-Veje, C., Main, K.M., Van Camp, G., Schoeters, G., Vanden Berghe, W., Andersen, H.R., 2017. Interaction between prenatal pesticide exposure and a common polymorphism in the PON1 gene on DNA methylation in genes associated with cardio-metabolic disease risk—an exploratory study. *Clin Epigenetics* 9, 35.
- Dennis, K.K., Jones, D.P., 2016. The exposome: a new frontier for education. *Am Biol Teach* 78, 542–548.
- Farina, Y., Abdullah, M.P., Bibi, N., Khalik, W.M., 2017. Determination of pesticide residues in leafy vegetables at parts per billion levels by a chemometric study using GC-ECD in Cameron Highlands, Malaysia. *Food Chem* 224, 55–61.
- Karmaus, A.L., Trautman, T.D., Krishan, M., Filer, D.L., Fix, L.A., 2017. Curation of food-relevant chemicals in ToxCast. *Food Chem Toxicol* 103, 174–182.
- Lampe, J.W., 2007. Diet, genetic polymorphisms, detoxification, and health risks. *Altern Ther Health Med* 13 (2), S108–S111.
- McKelvey, W., Jacobson, J.B., Kass, D., Barr, D.B., Davis, M., Calafat, A.M., Aldous, K.M., 2013. Population-based biomonitoring of exposure to organophosphate and pyrethroid pesticides in New York city. *Environ Health Perspect* 121 (11–12), 1349–1356.
- Skinner, M.K., 2016. Endocrine disruptors in 2015: epigenetic transgenerational inheritance. *Nat Rev Endocrinol* 12 (2), 68–70.

Pharmaconutrigenetics: The Impact of Genetics on Nutrient–Drug Interactions

*Olivia Dong*¹, *Tim Wiltshire*²

¹Duke Center for Applied Genomics and Precision Medicine, Duke University School of Medicine; ²UNC Eshelman School of Pharmacy and UNC Center for Pharmacogenomics and Individualized Therapy, Kannapolis, NC, United States

INTRODUCTION

Whereas the focus of this textbook is on the ways in which genetics affects the body's response to nutrients, this is one of many factors that explain variations in health status observed among individuals. The medications to which our bodies are exposed is another factor that can affect our health status, and genetic variations can alter the way these medications are processed in the body. The emerging field of pharmaconutrigenetics lies at the intersection where genetic variation shapes individual responses to the combination of nutrients and drugs. In the future, a feature of precision medicine will likely incorporate individualized therapies that tailor both nutritional and pharmacological interventions in closely coordinated fashion to achieve better outcomes in the prevention and treatment of diseases. Some clinical sites have started to use genetic testing as a valued component of patient care to guide nutrition and drug interventions, although widespread adoption has not yet occurred. For many genes, there is enough evidence for its incorporation into patient care because the impact it has on health is well-defined, but barriers to moving genetic information into patient care are associated with implementation challenges. Factors such as a lack of training among health care providers or a lack of supporting infrastructure within the health care system to integrate genetic information efficiently within electronic health care records are some examples of major implementation barriers that are keeping genetic information from being integrated into patient care. In addition, the nutrition and pharmacy fields do not typically collaborate optimally to ensure that the interventions with which they are providing patients work synergistically. Shifting toward stronger collaboration between the fields could

improve outcomes in the prevention and treatment of many diseases. Background information on the main genes implicated in drug biotransformation, a major type of pharmacogenes and several examples of the genetic impact on nutrient–drug interactions will be offered to demonstrate some ways in which the nutrition and pharmacy fields can work collaboratively to achieve better patient outcomes.

GENETIC VARIATION IN DRUG BIOTRANSFORMATION

The goal of prescribing drugs is to elicit an optimal response in the prevention or treatment of a condition or disease. Prescribing too much drug can lead to toxicity; prescribing too little drug can lead to subtherapeutic effects. The optimal drug concentration range that allows drugs to be efficacious while avoiding unbearable side effects is the ideal scenario; this therapeutic range can be narrow for some pharmacological agents, which makes precise dosing extremely important. Once a pharmacological agent enters the body, the agent undergoes metabolism, or biotransformation, to facilitate excretion. This excretion process involves making the drug more water-soluble or charged to avoid absorption, which also supports the elimination of the compound's pharmacological effects and lessens its potential toxicity within the body. Some drugs are prodrugs; they must go under biotransformation to produce an active metabolite to deliver effects in the body. However, most are given as active drugs, in which they undergo biotransformation to become inactivated. Genetic variation in the activity levels of these enzymes that are responsible for biotransformation can vary drug

exposure in the body. This means that genetic variation in these enzymes responsible for biotransformation is an important factor to consider in the drug prescribing process. These genetic variants can vary the ideal pharmacological agent and dose for patients from standard dosing. For instance, genetic variations that result in a nonfunctional enzyme can result in overexposure to active drugs and, on the other hand, lead to reduced exposure to an active drug if it is a prodrug. Drugs are dosed with the assumption that the patient possesses genes that all code for fully functional enzymes, but this is not the case for a subset of the population.

The CYP1, 2, and 3 families of the cytochrome P450 (CYP450) enzymes are responsible for the biotransformation of 70–80% of drugs used clinically (Zanger and Schwab, 2013). Approximately 86% of the drugs that are metabolized by a CYP enzyme are done so by a CYP gene that has both common and rare polymorphisms (Fig. 68.1), and the functional activity of these genes are often described as ultrarapid, normal/extensive, intermediate, and poor metabolizers phenotypes. For example, poor metabolizer often means the gene results in a loss-of-functioning protein, and carriers of this phenotype are unable to metabolize drugs and thus do not benefit from the pharmacological effects of that drug, if it is a prodrug. Under standard of care in the clinical setting, there may be a drug that is typically used as the first go-to agent, but genetic variations can change this established standard to achieve optimal drug response in a subset of patients. Therefore, the pharmacogenetic field aims to consider these varying levels of functioning enzymes at the genetic level dose drugs to more precisely and guide selection of the ideal pharmacological agent to optimize clinical outcomes. Genes other than the CYP450 enzymes have been established within the pharmacogenetic field as important in

the drug prescribing process. Examples include drug transporters along with enzymes beyond the cytochrome P450s that also metabolize drugs.

Currently, there are more than 160 unique drugs spanning various therapeutic areas with established genetic information in germline mutations in the drug prescribing process from the Food and Drug Administration (FDA) or the Clinical Pharmacogenetics Implementation Consortium (CPIC) (Food and Drug Administ, 2017; CPIC, 2017). More than 30 genes are implicated in the dosing of these drugs, many of which are CYP enzymes. The likelihood of carrying a variant genotype that could result in an altered response to standard medication is common. In fact, almost everyone will carry at least one high-risk gene variant, and one study found that 54% of adults in one study were prescribed at least one of the top 30 drugs with high pharmacogenetic risk over a 1-year period (Dunnenberger et al., 2015). Some of these drugs are commonly prescribed, such as codeine, antidepressants, and simvastatin. The chances of being prescribed a drug with established pharmacogenetic guidelines genetic information and carrying a genotype that results in a deviation from standard of care is a likely scenario.

It is imperative to remember that the prevalence of genotypes that result in a change in drug prescribing from standard of care can vary depending on the ethnic group being considered. For instance, the proportion of poor metabolizers for one group may be very rare, but it may be more common in other ethnic groups. It would be ideal to obtain genetic information in patients so that clinical decisions can be made from actual genotype results instead of using ethnicity as a predictor for how individuals will respond to drugs.

Now that the role of genetic variation within the pharmacogenetic field has been discussed, the next step is to combine the consequences of these genetic variations along with variants that are important in nutrigenetics. The following examples demonstrate the impact genetics can have on nutrition and pharmacological interventions.

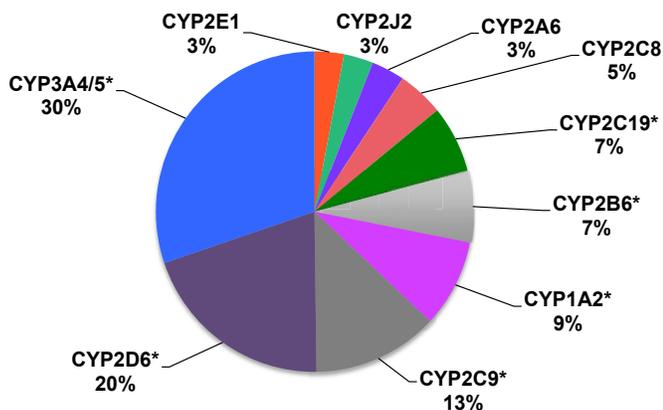


FIGURE 68.1 Breakdown of drugs ($n = 248$) that are metabolized by cytochrome P (CYP)450 isoforms (Adapted from Zanger and Schwab, 2013). *Polymorphic genes with established genetic dosing guidelines from the Food and Drug Administration or Clinical Pharmacogenetics Implementation Consortium (Food and Drug Administ, 2017; CPIC, 2017).

HYPERTENSION AND SALT SENSITIVITY STATUS

The salt sensitivity phenotype that was discussed in Chapter 41 (Electrolytes) is an example of pharmacogenetics. The salt sensitivity phenotype, which has genetic underpinnings, categorizes individuals into defined groupings, including individuals who are salt-resistant and those who are salt-sensitive. Determining the salt sensitivity phenotype is important because it is an independent risk factor for mortality even when blood pressure is considered. Individuals who are hypertensive and salt-sensitive have the worst mortality risk when

considering blood pressure and salt sensitivity status. Evaluating individuals for salt sensitivity could help clinicians choose how to intervene and reduce blood pressure for their patients. For those who are hypertensive and salt-sensitive, aggressive dietary sodium restriction (1500 mg dietary sodium/day) would be effective, but for those who are hypertensive and salt-resistant, severe sodium restriction may not be effective. Instead, other dietary changes such as reducing excessive weight need to be prioritized and antihypertensive medications may have to be considered. Several drug classes are available to help control blood pressure, and the process to find the specific antihypertensive that is effective for patients is still a trial-and-error approach. Pharmacogenetic research is ongoing in this area to find key genetic variants that can inform the drug selection process (Arnett and Claas, 2009).

Knowing the most effective and optimal way to prevent or treat disease is the cornerstone of precision medicine. In this example, a salt-restrictive diet (1500 mg sodium/day) is difficult to implement successfully because sodium is found so readily in the food supply in high amounts. Sodium restriction is a burdensome type of dietary intervention, so adequate evidence is needed to impose such a dietary prescription. Being able to identify individuals who would benefit most from this type of intervention may also improve the adherence to such a dietary restriction. Patients who use multiple antihypertensive agents and who are salt-sensitive may be able to reduce the number and amounts of medications by curbing their sodium consumption. Patients often feel overwhelmed by too many dietary prescriptions; genetics can offer a way to prioritize changes and even indicate which will or will not be worth pursuing.

GRAPEFRUIT–DRUG INTERACTIONS AND CYP3A4

CYP3A4 is an enzyme that metabolizes a substantial number of drugs that are biotransformed by CYP3A4 (Lynch and Price, 2007). CYP3A4 is a highly conserved gene (wild type allele: 94.9%–99.7%) across Europeans, Africans, East Asians, South Asians, and admixed Americans (Zhou et al., 2017). Specific food items can inhibit or induce the CYP3A4 enzyme, and components in grapefruit have been identified as inhibitors of this enzyme (Fukuda et al., 1997). For active compounds that are metabolized by this gene, inhibiting it will lead to excessive drug exposure to the body and can lead to adverse drug events. For prodrugs that are metabolized by this gene, therapeutic failure owing to underexposure to the active metabolites can occur if this enzyme is inhibited. As little as a single glass of

grapefruit juice can interfere significantly with drug action (Bailey et al., 1998). It is important for registered dietitians and pharmacists to be aware of such interactions and work jointly with patients, informing them about grapefruit-containing foods and understand the importance of avoiding them while they use medications that are substrates of CYP3A4.

Although pharmacists may be trained to inform patients of the interaction between grapefruit juice and any medications that are metabolized by CYP3A4, there may be food sources of grapefruit that are missed. Registered dietitians are knowledgeable about nutritional interventions but they need to pay particular attention to such nutrient–drug interactions. At the least, basic knowledge of pharmacology concepts to understand how CYP enzymes can increase or decrease drug levels in the body are core competencies for effective dietary guidance of patients using relevant drugs.

WARFARIN, VITAMIN K, AND CYP2C9/ VKORC1/CYP4F2

Warfarin has been used extensively worldwide as an anticoagulant, or blood thinner, for many cardiovascular indications such as atrial fibrillation, deep venous thrombosis, and pulmonary embolism. Considerable interindividual variability has been observed among patients when trying to maintain the narrow therapeutic index that is required to get patients to achieve target anticoagulation. Suboptimal warfarin dosing is one of the most commonly reported adverse drug events and one of the most common reasons for visits to the emergency room (Johnson et al., 2017), which makes it critical to achieve optimal warfarin dosing. A pharmacogenetic algorithm has been developed for warfarin dosing and illustrates the integration of genetic and multiple clinical factors in determining a more tailored dose that is specific to patients (Gage et al., 2008). These genetic factors include a VKORC1 polymorphism (rs9923231), CYP2C9*3, CYP2C9*2; clinical factors include body surface area, age, target international normalized ratio, amiodarone use, smoker status, race, and current thrombosis (Gage et al., 2008). These clinical factors alone explain 17%–22% of the variation observed in response to warfarin, whereas 53%–54% of the variability can be explained when genetics and clinical factors are considered together (Gage et al., 2008). The CYP4F2 gene accounts for 11% of the variability observed in response to warfarin among Europeans and has been added to the official dosing consensus statement from the CPIC (Johnson et al., 2017). Approximately 38% of patients would require a daily dose that is lower than standard dosing because of the presence of either the CYP2C9*2 or *3 or VKORC1-1639A variants (Johnson et al., 2017). It has been shown

that using pharmacogenetic information when dosing warfarin results in fewer numbers of dosing adjustments to achieve therapeutic range compared with standard dosing (Anderson et al., 2012). Fewer dosing adjustments would result in more efficient patient care and reduce the amount of time clinicians need to spend on each patient. These genetic variants in CYP2C9 and VKORC1 result in an increase in sensitivity to warfarin and a lower dose of warfarin is needed to achieve optimal anticoagulation. Excessive warfarin for those who are sensitive to warfarin increases the chances of a major or fatal bleed. In addition to these pharmacogenetic genes, it is also important to consider dietary intake of vitamin K from food sources and multivitamins. Patients receiving warfarin are advised to keep vitamin K intake consistent because dietary vitamin K is a compound required in the clotting cascade and promotes coagulation. Inadequate warfarin dosing may occur if dietary vitamin K suddenly increases; excessive warfarin dosing may result if dietary vitamin K suddenly decreases. To achieve optimal management of warfarin, collaborative efforts are needed from pharmacists and nutritionists to reduce unwanted side effects that have routinely been observed in patients prescribed this drug.

Warfarin dosing provides a prime example of the multifactorial approach in achieving more precise dosing. The algorithm integrates numerous clinical factors to account for the variability in warfarin that is required to achieve therapeutic range. Genetics plays a key role in explaining this variation but additional factors also affect this variability and algorithms such as that for warfarin may be the future of precision medicine. Combining many clinical factors together in a more nuanced manner to predict precise drug dosing may be a more accurate model. There are many known biomarkers that influence variability in drug and nutrient response but identifying the most impactful ones will be critical. For now, consistent vitamin K intake is still the way to manage patients receiving warfarin, but perhaps future algorithms can allow for greater flexibility of vitamin K intake.

THE ROLE OF CYP1A2 AND CYP2C19 IN REDUCING ISCHEMIC HEART ATTACKS

Lowering the risk of a myocardial infarction can be achieved through the interaction between optimal caffeine consumption using CYP1A2 genetic information and the optimal selection of an antiplatelet therapy using CYP2C19 genetic information for acute coronary syndrome patients receiving a stent placement.

CYP1A2 is responsible for metabolizing more than 95% of consumed caffeine, and the metabolic rate strongly depends on common genetic variants. Fast metabolizers usually carry two copies of the CYP1A2*1A

haplotype whereas slow metabolizers typically have the CYP1A2*1F haplotype. In the case for carriers of CYP1A2*1F, metabolizing caffeine more slowly would increase caffeine exposure within the body compared with the rapid metabolizer phenotype. This difference in caffeine exposure and excretion leads to different cardiovascular risk. There is evidence supporting the guidance that slow caffeine metabolizers should limit caffeine intake to 200 mg to avoid increased risk for a myocardial infarction (Cornelis et al., 2006).

Caffeine intake and its associated risk for myocardial infarction may seem benign, but under the right circumstances the added effort of paying close attention to caffeine intake may be necessary, as in the case for patients who have acute coronary syndrome undergoing a percutaneous intervention with stenting. For these patients, an antiplatelet therapy for at least 12 months is indicated and clopidogrel, a prodrug, is a highly prescribed antiplatelet medication and often the standard of care for this indication. As discussed earlier, prodrugs need to be biotransformed into the active drug for the desired protective pharmacological effects on the body. Adverse cardiovascular events result for CYP2C19 intermediate and poor metabolizers, who are carriers of the loss-of-function enzymes (*2–*8 haplotypes) because they cannot metabolize clopidogrel into the active drug and therefore do not receive the appropriate antiplatelet effects of clopidogrel as intended. FDA boxed warnings on the drug label and CPIC guidelines are available to inform clinicians about how to adjust clopidogrel prescribing based on CYP2C19 genotypes. This involves switching intermediate and poor metabolizers to an alternate antiplatelet therapy to reduce adverse cardiovascular events, including myocardial infarctions. It is estimated that 2%–15% of patients are poor metabolizers and 18%–45% of patients are intermediate metabolizers (Scott et al., 2013). The *2 haplotype is the most common CYP2C19 loss-of-function haplotype; the prevalence varies by ethnicity (15% in Africans, 29% in East Asians, 15% in Europeans, and 35% in South/Central Asians) (Scott et al., 2013). Although pharmacogenetic guidance is available to optimize drug dosing for antiplatelet therapies, this testing is only slowly coming to the clinical setting and those carrying a loss-of-function CYP2C19 genotype may receive subtherapeutic levels of an antiplatelet therapy when prescribed clopidogrel.

Clopidogrel users with loss-of-function CYP2C19 alleles and with the CYP1A2*1F allele are at increased risk for heart attacks when they consume more than 200 mg of caffeine/day. In this case, a clear intervention can be offered to patients to reduce their risk for myocardial infarctions. Genetic testing for these genes and assessing caffeine intake would be important for these patients, especially if the patients are receiving

suboptimal antiplatelet therapy. Some may argue that it might be better just to cut caffeine intake across all individuals when taking clopidogrel, but we are far enough along in genomic research to be able to provide individualized guidance to allow patients to choose the most fitting foods that are best-suited to their genetic profile. This example also emphasizes the importance of the nutrition and pharmacy fields understanding how therapies within their respective fields affect management of therapies in the other field.

WEIGHT GAIN MANAGEMENT WITH PSYCHOTROPIC DRUGS

Genetic information can optimize psychotropic drug selection among patients with psychiatric illnesses. Therapeutic failure and substantial drug-associated side effects are common for psychiatric conditions; turning to genetic information to avoid further therapeutic failures is a strategy clinicians have often employed in this field. Having patients undergo genetic testing when these agents are initiated could potentially help achieve better achieving better patient outcomes sooner and could potentially shorten the time it takes to reach therapeutic success. Pharmacogenetic information can improve the efficacy of the selected pharmacological agent by guiding either the dose or agent selection for patients. Using pharmacogenetic information when choosing psychotropic drugs can cut down on the time that is commonly used in the trial-and-error approach in selecting an agent. Approximately a fifth of drugs with genetic guidance from the FDA information are used for psychiatric illnesses (i.e., antipsychotics, antidepressants) (Eap, 2016). According to CPIC guidelines, deviation from the standard recommended dose for many of these psychotropic drugs is recommended for CYP2D6 ultrarapid, intermediate, and poor

metabolizers (8%–23% of the population) and CYP2C19 ultrarapid and poor metabolizers (7%–45% of the population) (Hicks et al., 2015, 2017).

Individuals taking psychotropic drugs are a high-risk group for developing metabolic disturbances. A major reason for therapeutic failure among patients receiving psychotropic drugs is the amount of weight gain that results as a side effect, which also compromises adherence to these therapies. The risk for obesity, diabetes, metabolic syndrome, and cardiovascular events is higher in this population, shortening life by 10–25 years (Delacretaz et al., 2017).

Commonly dosed psychotropic drugs that are prescribed are second-generation antipsychotics to treat a range of disorders, including schizophrenia, bipolar disorders, schizoaffective disorders, major depressive disorders, posttraumatic disorders, substance abuse disorders, and anxiety disorders, but these agents are associated with substantial weight gain, or antipsychotic-induced weight gain (AIWG). The amount of weight gain can vary depending on the specific second-generation antipsychotic, as illustrated in Fig. 68.2.

As a result of this increase in premature mortality among these patients, genetic polymorphisms that are associated with psychotropic-induced weight gain have been studied to identify patients at risk for weight gain, to match them with therapeutic agents that do not induce as much weight gain. Many promising nutrigenetic markers have been identified that can predict the risk for weight gain with these psychotropic agents. For instance, the most consistently associated genetic markers with AIWG are rs3813929C > T allele in the HTR2C gene and the rs7799039G > A allele in the leptin gene (Lett et al., 2012). In addition, the rs3888190C > A allele in the SH2B1 gene and the rs1000940A > G allele in the RABEP1 gene have been associated with metabolic decline during treatment with psychotropic agents (Delacretaz et al., 2017). Other genes have been implicated in

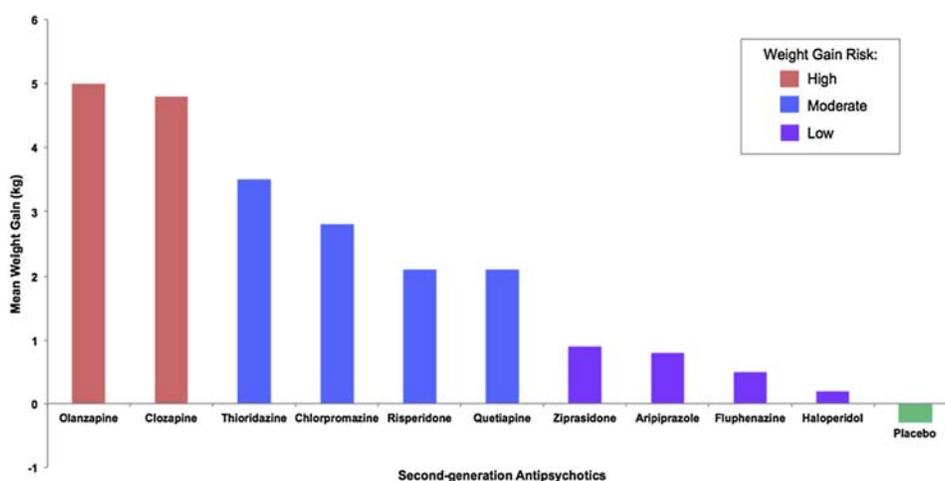


FIGURE 68.2 Mean weight gain with second-generation antipsychotics categorized by weight gain risk (Adapted from Lett et al., 2012).

weight gain associated with these medications, and the research in this area continues to grow because of the adverse effect of these drugs and the high rate of treatment failure experienced by these patients. Losing weight among this patient population is a priority because of the metabolic disturbances that occur with these psychotropic drugs. If weight loss medication is the strategy that patients end up using because of failed lifestyle attempts, there is also evidence to suggest genetic individualization for weight loss agents. For example, the weight loss medication sibutramine has been associated with greater weight loss among carriers of G/G and A/A alleles at the –148 position compared with carriers of the G/A allele in the phenylethanolamine *N*-methyltransferase gene (Peters et al., 2003). Using genetic information when selecting medications can also provide patients with some confidence that they will experience success with these agents. The psychiatry field is a therapeutic area that has seen a wealth of genetic research helping clinicians pinpoint better therapeutic strategies because of the failure rate of these pharmacological agents and the associated nutrition-related side effects. This is a worthwhile effort because of the health impacts observed in this patient population.

CONCLUSION

Some genetic testing already occurs for both nutrigenetics and pharmacogenetics in the clinical setting but increasing use of this information in the future is highly promising. More pharmaconutrigenetic examples that would lead to better outcomes for patients are being developed and will be ready for use in practice. These examples highlight the importance of genetic information for pharmacological and nutritional interventions to ensure optimal treatment and management of patients. Better integration of science and practice in pharmacy and nutrition will achieve new synergies when coordinating interventions that are now prescribed separately in these fields.

References

- Anderson, J.L., Horne, B.D., Stevens, S.M., et al., 2012. A randomized and clinical effectiveness trial comparing two pharmacogenetic algorithms and standard care for individualizing warfarin dosing (CoumaGen-II). *Circulation* 125 (16), 1997–2005.
- Arnett, D.K., Claas, S.A., 2009. Pharmacogenetics of antihypertensive treatment: detailing disciplinary dissonance. *Pharmacogenomics* 10 (8), 1295–1307.
- Bailey, D.G., Malcolm, J., Arnold, O., Spence, J.D., 1998. Grapefruit juice-drug interactions. *Br J Clin Pharmacol* 46 (2), 101–110.
- Cornelis, M.C., El-Sohehy, A., Kabagambe, E.K., Campos, H., 2006. Coffee, CYP1A2 genotype, and risk of myocardial infarction. *JAMA* 295 (10), 1135–1141.
- CPIC: Clinical Pharmacogenetics Implementation Consortium 2017.
- Delacretaz, A., Zdravovic, A., Vandenberghe, F., et al., 2017. Association of variants in SH2B1 and RABEP1 with worsening of low-density lipoprotein and glucose parameters in patients treated with psychotropic drugs. *Gene* 628, 8–15.
- Dunnenberger, H.M., Crews, K.R., Hoffman, J.M., et al., 2015. Preemptive clinical pharmacogenetics implementation: current programs in five US medical centers. *Annu Rev Pharmacol Toxicol* 55, 89–106.
- Eap, C.B., 2016. Personalized prescribing: a new medical model for clinical implementation of psychotropic drugs. *Dialogues Clin Neurosci* 18 (3), 313–322.
- US Food and Drug Administration Table of Pharmacogenomic Biomarkers in Drug Labeling, 2017.
- Fukuda, K., Ohta, T., Oshima, Y., Ohashi, N., Yoshikawa, M., Yamazoe, Y., 1997. Specific CYP3A4 inhibitors in grapefruit juice: furocoumarin dimers as components of drug interaction. *Pharmacogenetics* 7 (5), 391–396.
- Gage, B.F., Eby, C., Johnson, J.A., et al., 2008. Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. *Clin Pharmacol Ther* 84 (3), 326–331.
- Hicks, J.K., Bishop, J.R., Sangkuhl, K., et al., 2015. Clinical pharmacogenetics implementation consortium (CPIC) guideline for CYP2D6 and CYP2C19 genotypes and dosing of selective serotonin reuptake inhibitors. *Clin Pharmacol Ther* 98 (2), 127–134.
- Hicks, J.K., Sangkuhl, K., Swen, J.J., et al., 2017. Clinical pharmacogenetics implementation consortium guideline (CPIC) for CYP2D6 and CYP2C19 genotypes and dosing of tricyclic antidepressants: 2016 update. *Clin Pharmacol Ther* 102 (1), 37–44. <https://doi.org/10.1002/cpt.597>. PMID: 27997040.
- Johnson, J.A., Caudle, K.E., Gong, L., et al., 2017. Clinical pharmacogenetics implementation consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 update. *Clin Pharmacol Ther* 102 (3), 397–404.
- Lett, T.A., Wallace, T.J., Chowdhury, N.I., Tiwari, A.K., Kennedy, J.L., Muller, D.J., 2012. Pharmacogenetics of antipsychotic-induced weight gain: review and clinical implications. *Mol Psychiatry* 17 (3), 242–266.
- Lynch, T., Price, A., 2007. The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. *Am Fam Physician* 76 (3), 391–396.
- Peters, W.R., MacMurry, J.P., Walker, J., Giese Jr., R.J., Comings, D.E., 2003. Phenylethanolamine *N*-methyltransferase G-148A genetic variant and weight loss in obese women. *Obes Res* 11 (3), 415–419.
- Scott, S.A., Sangkuhl, K., Stein, C.M., et al., 2013. Clinical Pharmacogenetics Implementation Consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update. *Clin Pharmacol Ther* 94 (3), 317–323.
- Zanger, U.M., Schwab, M., 2013. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther* 138 (1), 103–141.
- Zhou, Y., Ingelman-Sundberg, M., Lauschke, V.M., 2017. Worldwide distribution of cytochrome P450 alleles: a meta-analysis of population-scale sequencing projects. *Clin Pharmacol Ther* 102 (4), 688–700.

Setting Genome-Directed Guidelines and Algorithms

Martin Kohlmeier

Human Research Core and Nutrigenetics Laboratory, UNC Nutrition Research Institute; and Schools of Medicine and Public Health, Department of Nutrition, Kannapolis, NC, United States

Glossary

Clinical utility The balance of health improvement with a test algorithm and the burden from costs and harms

GWAS Genome-wide association study is often used as a first screen in the search for genetic predictors

Haplotype Combination of genetic variants on the same DNA strand often extending over hundreds of thousands of bases

PUTTING GENETIC INFORMATION TO WORK

Raw, uninterpreted genetic information is largely an industrial commodity. It takes just drooling into a spit kit, a modest amount of money, and a few weeks' waiting time to get more genetic information than we can fit on a typical thumb drive. The sequencing of a whole human genome costs barely US\$1000, with prices continuing to fall. The expensive part is the mining of these raw data and using them to spin meaning and guidance.

At the simplest level, the derived information can be dichotomous, the presence or absence of a disposition. Are we likely to continue expressing lactase in our small intestines, which deals with lactose in unprocessed milk? The answer tends to be yes or no. Sometimes several variants affect the same gene in a way that is best predicted by looking at the different haplotypes. In other cases, a more complex relation existing among variants in several genes adds up to one cumulative outcome.

So, how can we find out what the genes are telling us?

A CASE OF DICHOTOMY: LACTASE PERSISTENCE

It was a major achievement when Ennatah and colleagues found the locus that determines whether lactase continues to be expressed in the small intestines after weaning ([Ennatah et al., 2002](#)). Within a short time, it became clear that some herders and dairy farmers around the world have other variants associated with lactase persistence. Because we now know the responsible regions far upstream of the lactase transcription start site, we can systematically screen those regions by targeting the known variants individually or by sequencing. An individual is likely to be lactase persistent with at least one copy of the few predictive variants (−13907G, −13910T, −13915G, −14009G, or −14010C), regardless of ethnicity. Two allele copies reinforce lactase production tends even more ([Dzialanski et al., 2016](#)).

The genotype is not the only factor determining our ability to consume large amounts of milk, but it is an important one. In addition, there is an age effect. Many notice as they approach their forties and fifties that they become more sensitive to large amounts of milk and eventual develop abdominal discomfort with even modest amounts. This increasing lactose intolerance is understood to be mediated by age-related epigenetic silencing of the enhancer region. The effect occurs much earlier in nonpersistent individuals, but much later particularly in heterozygous carriers of the persistence variant. By the second half of life, the digestive capacity of the small intestines declines at a modest rate, but these effects tend to add up. This means that both the specific

genotype (such as –13910 CT versus –13910 TT) and age have to become part of the algorithm, in which advancing age predicts less lactase production, particularly in people with the –13910 CT genotype.

Finally, there is increasing recognition that the microbiome in some populations mitigates the discomforting consequences of undigested lactose in the large intestines, possibly owing to adaptation of the bacterial species to the presence of this disaccharide. How this might be incorporated into a further refined predictive algorithm will have to be seen.

SINGLE-VARIANT ASSOCIATION STUDIES

The simplest genetic risk predictors use just a single variant and therefore search for the one with the closest association. This is how most genome-wide association studies (GWAS) are performed (Fig. 69.1), but it is also widely used across a single gene or with a number of selected candidate variants in genes of interest.

A special case is the scanning of the variant near and within a particular target gene. The expectation is that this scan identifies a good predictor of the function of interest. Many important predictive polymorphisms have been identified this way.

ALLELE COUNTING

More sophisticated approaches count alleles across several variants, even across genes, and distinguish between individuals with high scores and those with low scores. This approach is relatively common in cancer research to predict disease risk.

Another possibility is that it is not the sum of risk alleles that predict the outcome but the presence of a

number of alternative risk variants, often in different genes that functionally may have similar end effects. An example for this approach can be seen with the choline requirement that was predicted to be high by alleles in either of two genes: SLC44A1 (rs7873937 G, rs2771040 G, rs6479313 G, rs16924529 A, and rs3199966 C) and CHKB (rs1557502 A) (da Costa et al., 2014).

HAPLOTYPES

There are hundreds of common variants on an average-sized gene and it is difficult to know their functional relevance. Many are said to be silent because they do not change the encoded amino acid in the protein sequence. However, there may still be functional consequences owing to differences in splicing and RNA editing. Nevertheless, the entire DNA sequence works together and the individual variants do not function in isolation. Often, variants across different segments of a gene are linked in common haplotypes, potentially affecting (or not) promoter regions, exons, and 3'-untranslated regions with their effect on messenger RNA stability and interactions with regulatory factors such as microRNAs.

The many common variants in the GC gene (encoding vitamin D-binding protein) may serve as an example because they strongly modulate 25-hydroxyvitamin D (25OH-D) concentration in blood. Upon closer examination, we find that the numerous polymorphic variants across the GC gene are grouped into a limited number of haplotypes. The three haplotypes defined by the variants rs7041 and rs4588 comprise more than 90% of Caucasians in Central Europe. The haplotype names are based on the banding in isoelectric focusing electrophoresis. 1s (rs4588C–rs7041G) is associated with a much higher 25OH-D concentration than haplotype 2 (rs4588A–rs7041T). The third haplotype 1f (rs4588C–rs7041T) is associated with a somewhat lower 25OH-D concentration than 1s (Abbas et al., 2008). Looking at the two variants individually, we would not have known about the existence of the metabolically distinct third haplotype 1f. These three haplotypes differ in many other variants, but rs4588 and rs7041 are sufficient to distinguish them. This means that it is not necessarily those two that are responsible for the functional difference between them, but it could be any of the other variants associated with their respective haplotype or some combination. As a first approximation of a predictive algorithm for predicting 25OH-D concentration in Germany, rs4588 and rs7041 will work even better combined with a seasonal adjustment because concentrations are generally known to be much lower during the darker months than during summer.

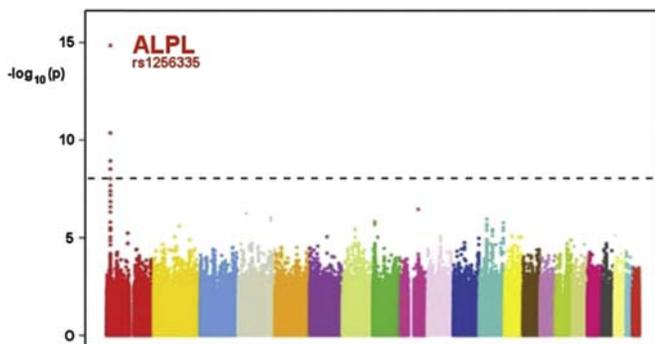


FIGURE 69.1 Manhattan plot of all associations of the concentration of pyridoxal phosphate (vitamin B6) with variants across the entire human genome. From Kohlmeier, M., 2013. *NutriGenetics: Applying the Science of Personal Nutrition*. Academic Press, pp. 288.

Moreover, vitamin D-binding protein has an effect on circulating 25OH-D concentrations, but also on postmenopausal breast cancer risk (Abbas et al., 2008). This is contrary to what one might expect based on cancer risk reduction with higher 25OH-D concentration. When trying to build predictive algorithms, their target has to be clear: in this case, 25OH-D concentration or breast cancer risk.

CRITERIA FOR SELECTING GENETIC TESTS

Before genetic variants can reasonably be used for clinical testing, the expected effect must have been replicated in at least one other study by other investigators (Kohlmeier et al., 2016). The actionable information coming from the test has to be analytically reliable. The algorithm should be well-tested and documented. In addition, the mechanism of the intervention effect should be reasonably understood to avoid unexpected pleiotropic harm (at another site or system).

The most important question is thus whether testing provides adequate clinical utility. This means that the outcome after using the test to inform a clinical decision is clearly better than the outcome without using the test. The outcome is the balance of benefits and harms. It is easy to forget that this is not just about the direct effects, but also about the harms associated with the test itself. Take, for example, a test, methylenetetrahydrofolate reductase (MTHFR) 677 C/T (rs1801133), which can predict which individuals will have lower blood pressure with supplemental riboflavin in Ireland (Wilson et al., 2013). The algorithm would identify among adults with elevated blood pressure all of those with the MTHFR 677 TT genotype and add a modest amount of riboflavin (1.6 mg/day) to their usual medication. The expected benefit would be on average 5–6 mm Hg lower systolic blood pressure. Let us assume that the beneficial effect does not wear off over a longer time (which cannot be assumed in the absence of actual long-term trial data). This may then seem an easy call because the benefit is so clear. The blood pressure, for which the individuals require treatment, is being lowered. However, before we endorse the test too readily, we should consider the other side, the potential harm, which may not be as obvious. First, there are expenses associated with the test, maybe additional office visits, and the additional treatment. Then there is the potential

for psychological and social burdens. That will depend on the specifics of the test, whether it is a dedicated test that targets only this one variant or one with broader coverage. The more of the genome that is analyzed, the more can be discovered, including unpleasant or hurtful information (such as about paternity). Moreover, there is always the possibility that understanding of the health significance of the test changes. The APOE4 allele was originally seen as just a minor contributing factor to blood cholesterol concentration and cardiovascular risk. Only later was it understood that this allele is an important predictor of late-onset Alzheimer disease. Finally, harm may come from insurance and employment risks. After the result of the test is known, it may become more difficult to obtain life insurance or other insurance, or the person who is tested may have difficulty in finding or keeping employment. On balance, the clinical utility will depend on the handling of the risks, burdens, and cost. Much of that can be mitigated with informed planning.

References

- Abbas, S., Linseisen, J., Slanger, T., Kropp, S., Mutschelknauss, E.J., Flesch-Janys, D., Chang-Claude, J., 2008. The Gc2 allele of the vitamin D binding protein is associated with a decreased postmenopausal breast cancer risk, independent of the vitamin D status. *Cancer Epidemiol Biomark Prev* 17 (6), 1339–1343.
- da Costa, K.A., Corbin, K.D., Niculescu, M.D., Galanko, J.A., Zeisel, S.H., 2014. Identification of new genetic polymorphisms that alter the dietary requirement for choline and vary in their distribution across ethnic and racial groups. *FASEB J* 28 (7), 2970–2978.
- Dzialanski, Z., Barany, M., Engfeldt, P., Magnuson, A., Olsson, L.A., Nilsson, T.K., 2016. Lactase persistence versus lactose intolerance: is there an intermediate phenotype? *Clin Biochem* 49 (3), 248–252.
- Enattah, N.S., Sahi, T., Savilahti, E., Terwilliger, J.D., Peltonen, L., Järvelä, I., 2002. Identification of a variant associated with adult-type hypolactasia. *Nat Genet* 30 (2), 233–237.
- Kohlmeier, M., 2013. *Nutrigenetics. Applying the Science of Personal Nutrition*. Academic Press, San Diego, ISBN 9-7801-2385-9-006, p. 386.
- Kohlmeier, M., De Caterina, R., Ferguson, L.R., Görman, U., Allayee, H., Prasad, C., Kang, J.X., Nicoletti, C.F., Martinez, J.A., 2016. Guide and position of the International Society of Nutrigenetics/Nutrigenomics on personalized nutrition: Part 2: ethics, challenges and endeavors of precision nutrition. *J Nutrigenet Nutrigenom* 9, 28–46.
- Wilson, C.P., McNulty, H., Ward, M., Strain, J.J., Trouton, T.G., Hoelt, B.A., Weber, P., Roos, F.F., Horigan, G., McAnena, L., Scott, J.M., 2013. Blood pressure in treated hypertensive individuals with the MTHFR 677TT genotype is responsive to intervention with riboflavin: findings of a targeted randomized trial. *Hypertension* 61 (6), 1302–1308.

This page intentionally left blank

Direct-to-Consumer Testing

Lynnette R. Ferguson¹, Virginia R. Parslow²

¹Discipline of Nutrition and Dietetics, Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand ²Department of Optometry and Vision, Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand

Glossary

Utility Assessment of a test by asking about the number of cases for which the outcome is better with the information versus without it, considering the impact and usefulness of the test results to the individual, the family, and society, as well as psychological, social, and economic consequences of testing and implications for health outcomes.

BACKGROUND

The rapid rise in -omics knowledge and technologies comes with the promise to support a shift from a disease-focused approach to science, medicine, and health to one that focuses on systems biology (Bland, 2015). The resulting personalized medical, pharmaceutical, and nutritional recommendations reflect further shifts in medicine from a diagnostic to a more prognostic approach (Paul et al., 2014) and from a paternalistic model to one in which individuals share in health-related decision-making (Middleton, 2012). However, even health professionals are commonly out of their depth when dealing with increasingly complex -omics and other individual information. It is impractical to confront lay users with data of such breadth and depth and expect them to make sense of them. Personalized technologies purport to tailor anything from exercise to nutritional intake, sleep patterns, and beyond (Gibney et al., 2015). With increased access to information and services, many people expect the right to know and make decisions based on their personal genetic makeup (Ahlgren et al., 2013; Su et al., 2013) (Fig. 70.1).

Personalized nutrigenetic testing is often marketed and sold directly to consumers. Although it is often not credited as such, the first company to offer direct-

to-consumer (DTC) nutrigenetic testing was Sciona, which offered its first product in 2001. A short year later, Sciona switched to a direct-to-practitioner model owing to fierce resistance from opponents (Finogold et al., 2005). Interleukin Genetics, another early starter, began life as Medical Sciences Systems, founded in San Antonio, Texas, and became Interleukin Genetics in 2000, when it shifted to Waltham, Massachusetts. Following from a strong interest in pharmacogenomics, the company became interested in nutrigenomics, which resulted in its partnering with Alticor, the parent company of Amway, to develop personalized nutrition products. In 2003, Alticor bought the controlling interest in Interleukin and it became a separate research and development subsidiary of Alticor (Finogold et al., 2005). Interleukin's Inherent Health Weight Management Genetic Test forms part of Amway's personalized nutrition offerings; the website for Inherent Health also offers DTC, the Weight Management Test, plus a Nutritional Needs Genetic Test and a Heart Health Genetic Test (<http://www.xcode.in/nutrition-and-health-genomics>).

In 2007–10, there was a boom in DTC testing; many companies hoped to capitalize on direct market possibilities. The variety and quality of testing varied considerably, as did the particular genes and mutations studied. To a large extent, this continues to be the case (Grimaldi, 2014).

Things slowed and changed from 2010 onward, with the Food and Drug Administration (FDA) sending 23 letters to DTC companies in 2010 and 2011 stating that in the FDA's view, the companies were offering medical devices, and that as such, they needed regulatory approval (Vorhaus et al., 2011). As a result of these letters, numerous companies ceased trading or scaled back their operations considerably, whereas others shifted to marketing tests only through physicians

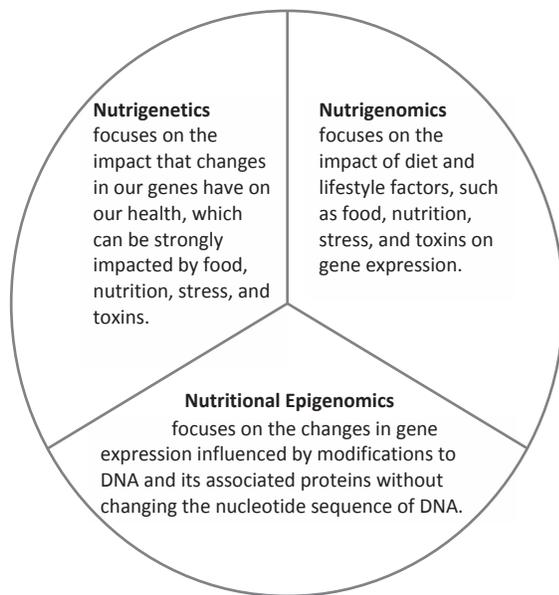


FIGURE 70.1 Clarification of terms. Nutritional genomics studies the impact of genetic makeup on the risk for disease and dysfunction and how that can be mitigated by nutritional intervention, as well as the impact that food, nutrition, stress, and toxins have on the expression of genes. It is an umbrella term that includes nutrigenetics, nutrigenomics, and nutritional epigenomics.

(Allison, 2012). Reasons for this (aside from regulatory pressure) were that consumer demand for DTC tests was not as hoped, and the clinical testing market was considered more lucrative than the DTC market (Allison, 2012). In 2012, Navigenics and deCODEme were both bought by firms that ceased offering tests via the DTC model. During those years (from 2010 onward), prices decreased dramatically. The cost of an entire genome scan dropped from approximately \$250,000 to just under \$1000. Companies also began searching for more nuanced business models with which to advance personal genomic testing (Allison, 2012).

23andMe, which is seen as the flagship DTC genetic test (GT) company by many, ceased offering its Personal Genome Service (PGS) in the United States in 2013 after a cease and desist letter from the FDA. In Feb. 2015, 23andMe announced that it had received FDA approval to sell its Bloom test DTC, the first such regulatory approval given in the United States. The FDA stated that part of the reason for this approval was because the test is a carrier test rather than a diagnostic one (Lab Tests Online, 2015). The FDA also converted 23andMe's application for approval from the 510K status under which it was originally submitted to a de novo request (Health and Medicine Week, 2015). In Oct. 2015, 23andMe relaunched a pared-down PGS at US\$199 (Hernandez, 2015a). The only wellness tests currently sold by 23andMe in the United States are for a predisposition to caffeine-liking, lactose intolerance, face-flushing response to drinking alcohol, and muscle

composition (Park, 2015). Since 2014, 23andMe has targeted its fuller range of tests to customers in Canada, the United Kingdom (where they can be purchased through the nationwide chain of Superdrug pharmacies (23andMe, 2015), Denmark, Finland, Ireland, Sweden, and the Netherlands (Kalokairinou et al., 2015); however 85% of 23andMe's customer base is in the United States (Hernandez, 2015a).

THE REGULATORY ENVIRONMENT

Regulatory Situation in the United States

In the United States, three government bodies have jurisdiction over GT:

1. The FDA has authority over whether GT are considered medical devices and whether regulatory approval is therefore required. To date, commercial test kits (groups of reagents used in gene testing and sold to laboratories, hospitals, and doctors' offices) have been the main focus of regulation, whereas laboratory-developed tests (LDTs) (i.e., tests developed and used by a single laboratory) have largely been exempt. The FDA announced in 2010 that it planned to regulate all GT, soon after Pathway Genomics announced plans to market DNA test kits through US retail pharmacy Walgreens. These plans were abandoned when Walgreens changed its mind about selling the tests (Pollack, 2010), subsequent to Pathway Genomics receiving an FDA letter stating that the tests in question need FDA approval. Despite FDA's notice to Congress in Jul. 2014 that it would be announcing draft guidelines on the regulation of LDTs in the next 60 days (National Human Genome Research Institute), no further action was taken with regard to the timing (Evans to Parslow, 2015, personal communication).
2. The Centers for Medicare and Medicaid Services regulates clinical laboratories by ensuring their compliance with the Clinical Laboratory Improvement Amendments of 1988 in terms of staff qualifications, quality control, and proficiency testing processes, but they do not examine whether the tests performed are clinically valid or clinically useful (Kohlmeier et al., 2016).
3. The Federal Trade Commission regulates how tests are advertised.

Direct-to-Consumer Regulations in Europe

In Europe, DTC GT currently come under the In-Vitro Diagnostic Medical Devices Directive, but no universal

law regarding DTC GT is applied throughout every European Union (EU) Member State, resulting in some countries allowing DTC GT (e.g., the United Kingdom [Saukko, 2013] and Greece [ter Meulen et al., 2012]) whereas other countries require a doctor's involvement in GT (e.g., France and Germany [Saukko, 2013]). In fact, because of its newness and uniqueness, personalized nutrition itself, falls under numerous European legal domains rather than one clear one (Kohlmeier et al., 2016). Discussion is currently under way among the European Parliament, the Council of the EU, and the European Commission to replace the current directive with a Regulation on In-Vitro Diagnostic Medical Devices. The impact of the new regulation is potentially far-reaching, including the possibility that nutrigenetic tests will be regulated identically to disease-risk tests and that residents of EU Member States will not be permitted to use tests bought outside the EU if those tests do not meet the new EU regulation's requirements (Kalokairinou et al., 2015).

Direct-to-Consumer Regulations in Countries Other Than Europe and North America

Just as the situation in Europe is far from homogeneous, a similar situation exists throughout other parts of the world. The regulatory situation in Australia has been ambiguous since a regimen for regulating in vitro diagnostic devices was introduced in 2014 (Nicol and Hagger, 2013), whereas in New Zealand, DTC testing is available but people are required to disclose resulting information if they purchase new health or life insurance policies (Lintott and Dixon, 2014). In Latin America, most countries do not have specific rules for GT (Kohlmeier et al., 2016), which has resulted in the growth of DTC in parts of that area (Ashton-Prolla et al., 2015). There are scattered pockets of DTC nutrigenetic testing throughout Asia; India has a proliferation of such companies.

Dialogue Within the Scientific Community Regarding Best Approach to Regulating Nutrigenetic Tests

Even within the international scientific community, disparate views are held regarding whether nutrigenetic testing should be categorized with testing for disease risk. Gorman et al. (Gorman et al., 2013), quoting Paynter et al. (Paynter et al., 2010), point out that unlike disease-risk tests that are based on association studies, nutrigenetic tests use specific information based on specific diet–gene interactions, and that the two should not be confused. They therefore argue that because nutrigenetics tests simply study genetic influences on biological

processes, it is not necessary to offer anything more than metabolic information as evidence of clinical utility. Similarly, Grimaldi argues that because nutrigenetics is not intended for specific disease prevention but rather for the long-term promotion of health based on best available evidence, evidence required for its use in personalized nutrition should be the same as that required for existing nutritional guidelines (Grimaldi, 2014). Such evidence is mainly epidemiological and interventional in nature (Gorman et al., 2013). Further, Grimaldi suggests that nutrigenetic tests should be judged on their health utility rather than their clinical utility, to fit with their differentiated status compared with and to medical tests (Grimaldi, 2014).

El-Sohehy and Nielson argue that nutrigenetic tests promise to be much more reliable than those that predict disease risk, because they are based on dietary responsiveness of particular phenotypes rather than just gene–disease association studies. They further state that a nutrigenetic approach to gene testing is likely to increase utility and validity of disease–risk associations discovered and to provide a more suitably focused approach to genetic testing feedback (Nielsen and El-Sohehy, 2012). This overcomes the fact that gene–disease associations are increasingly being discovered and that each time it happens, it changes the degree of disease susceptibility that existing genes are thought to have on disease risk.

Others see nutrigenetic advice as straddling a border between medicine and nutrition/lifestyle (Ahlgren et al., 2013), or between medicine and consumer culture (Saukko et al., 2010). Those who hold these views quote the fact that knowledge gained about links between genetic variants and disease has informed current understanding of gene–diet effects on homeostasis and disease (Fenech et al., 2011) and that such knowledge guides personalized nutrition in support of an individual's optimal health by developing personalized nutritional care (Ferguson et al., 2016). Furthermore, the International Society of Nutrigenomics and Nutrigenetics states that “nutrigenetics considers the influence of individual genetic variation on differences in response to dietary components, nutrient requirements and predisposition to disease” (Kohlmeier et al., 2016).

The idea that nutrigenetic testing is inextricably linked with disease risk is further supported by the fact that, just as drugs can modulate molecular mechanisms, so, too, can nutrients (Ferguson et al., 2016), and hence, personalized nutrition by its very nature focuses on using nutrients and their interaction with an individual's genome to affect disease risk (Kang, 2014). This is supported by the fact that targeted dietary interventions to targeted genotypes can be a more refined and effective tool in prevention (San-Cristobal et al., 2013) and or amelioration of chronic disease (Nielsen

and El-Sohemy, 2012) New understandings of the ways in which macronutrients and micronutrients interact with our genetic makeup and an improved understanding of the mechanisms that underlie chronic inflammation (a central factor in many common diseases) enhance the prospect of nutrigenetic testing assisting with the prevention and treatment of disease and disease risk as well as a role in promoting general health and maintaining homeostasis.

One solution to differing views regarding whether nutrigenetic tests should have to prove clinical utility in the same way as other GT is to ensure they are marketed as lifestyle rather than as medical, and to make available only tests that have been clinically proven to be useful (Grimaldi, 2014; Gorman et al., 2013; De Caterina and El-Sohemy, 2016). As Howard and Borry point out, it is impractical and likely impossible to ban genetic testing, given the reach of the Internet, and therefore, it is much more useful to offer relevant guidance on how best to use GT (Howard and Borry, 2012).

CURRENT PERCEIVED ADVANTAGES OF DIRECT-TO-CONSUMER NUTRIGENETIC TESTING

- Improved health care based on personalized dietary recommendations (Ferguson et al., 2016)
- Ease of accessibility (Kohlmeier et al., 2016)
- Sense of personal empowerment and control in learning information about oneself and acting on the information learned (Consumers on the Internet)
- Ability to have tests done from a distance (i.e., no need to attend the clinic in person)
- Potential for preventive measures by individuals in choosing dietary supplements, nutraceuticals, and functional foods relevant to an individual's genotype
- Assistance with chronic issues such as obesity through gene-targeted recommendations (Kohlmeier et al., 2016)

CURRENT CONCERNS ABOUT DIRECT-TO-CONSUMER NUTRIGENETIC TESTING

Scientific Concerns

- Uncertainty about:
 - Scientific validity: the accuracy with which a test predicts an outcome (Grimaldi, 2014)
 - Analytical validity: whether a test accurately identifies the gene variant it is supposed to identify (Saukko, 2013)
- Clinical validity: the strength of association between a gene variant and a disease (Saukko, 2013)
- Health utility (Grimaldi, 2014)/clinical utility (Saukko, 2013): how likely it is that an intervention will lead to a beneficial outcome (Grimaldi, 2014).
- When considering nutrigenetic testing under the umbrella of (or on the boundary with) testing for disease risk, the following must be considered: more gene–disease associations are being discovered with time, and whenever this happens, it changes the degree of disease susceptibility that existing genes are thought to have for disease risk. This concern would be somewhat mitigated if our previously suggested approach is taken with nutrigenetic tests marketed as lifestyle rather than as medical, and if only tests that clinically proven to be useful are made available (Grimaldi, 2014; Gorman et al., 2013; De Caterina and El-Sohemy, 2016).
- A systems approach is needed when offering and interpreting testing, which is not always available from or through DTC companies. An example of complexities that need considering is that although genome-wide association studies (GWAS) linked at least 150 genetic variants to body mass index, waist circumference, and obesity risk, it appears that the microbiome and epigenome also affect weight loss, gain, and maintenance (Bray et al., 2016).
- GWAS data are further limited in that they usually use correlation rather than gene–environment factors and hence provide only a fragment of the overall picture, particularly in which heritability is concerned (Ferguson et al., 2016).
- Similarly, the accuracy of any risk assessment will depend on many factors, including which single-nucleotide polymorphism (SNP) panels are used and whether or which environmental factors are incorporated into assessments (Camp and Trujillo, 2014).
- There is a lack of accompanying tools to support gene testing, such as metabolomics, which would potentially increase both the accuracy of results (clinical validity), given impacts such as environmental factors, and more useful dietary guidance (clinical utility).
- There is the belief that the benefit of nutrigenomics for disease prevention has not yet been proven and that consequently, there is a need for more research.
- There is potential medicalization of food (Görman, 2006).
- Although certain SNPs may have an important role in nutritional needs, nutrient metabolism, and disease risk (Ferguson et al., 2016), using single genes to explore complex disease risk and on which to base genomic feedback is simplistic (Madden et al., 2011).

When assessing risk for chronic diseases such as cancer, diabetes, and heart disease, SNPs are only predictive of an altered disease risk, and only in a limited way (Camp and Trujillo, 2014). Gene–gene interactions must be considered, multiple variants and their impact on particular genes must be assessed, and an awareness should be held or applied that (ultimately) numerous genes can affect disease risk, development, and progression (Fenech et al., 2011). DTC GT for complex disorders needs a full understanding of complexities involved when selecting tests to use and skilled interpretation and genetic counseling, which are unlikely to be consistently available via DTC purchases (Federal Trade Commission).

- Complexities in how nutrients interact with one another (the food matrix) are often not well-understood by consumers who might take high doses of particular compounds that could ultimately cause problems.
- Misunderstanding and inconsistencies of attitude exist regarding nutrigenetic testing displayed by so-called experts in the medical literature, which lead to both public and professional confusion.
- More education is needed for professionals (Roberts and Ostergren, 2013).
- Even within the field of direct-to-provider testing, clinicians rely mainly on commercial laboratories offering genetic risk assessment services to evaluate knowledge and assess risks (McGowan et al., 2014).
- The likely importance and relevance of factors such as gender and ethnicity on effects of gene-based, personalized nutrition have not yet been widely understood or applied (San-Cristobal et al., 2013).

Consumer Safety and Care

- There is a tendency for those who are marketing tests also to market supplements that may be of unproven value to the recipient; in some cases, they may actually be harmful (Webborn et al., 2015).
- There is a lack of accurate interpretation to consumers and of qualified guidance from a qualified professional regarding how to apply results
- Even when a professional is involved in the process, that professional is often part of the company selling the tests rather than an independent professional such as a doctor looking after the patient or a dietician looking after the client (Paul et al., 2014).
- Regulations controlling gene testing are potentially inadequate, particularly but not only, regarding minors and others unable to give informed consent.

- People are being given more information and/or unwanted results than they asked for to the point where it creates anxiety (Kohlmeier et al., 2016).
- There is disparity in the ability to afford both tests and products recommended as a result of those tests.
- More education is needed for consumers (Roberts and Ostergren, 2013; Skirton et al., 2012).
- Seemingly different results come from different laboratories for the same tests on the same person (which does not consider that tests performed in different laboratories and using different methods often produce different results in science (Perbal, 2014a).

Regulatory Issues

- There is a need for better regulation (Middleton, 2012).

Privacy and Security Issues

- Potential discrimination by insurance companies, employers, and others exists based on genetic aspects of individuals (Kohlmeier et al., 2016).
- Privacy and safety of genetic information are necessary (Stewart-Knox et al., 2015; Perbal, 2014b), particularly in light of:
 - the growing trend of DTC companies to collect genetic data in the name of voluntary research to sell and/or share it. Some 23andMe customers feel or felt duped that personal data gifted to 23andMe for the greater good were used to establish, among other things, a patent on polymorphisms associated with Parkinson disease (Perbal, 2014a). In Jan. 2015, 23andMe signed a USD\$10 million partnership with Genentech to access part of the genomic information on those at increased risk for Parkinson disease, and in Mar. 2015, 23andMe created a therapeutic group whose goal was to translate genomic information into drugs (Park, 2015).
 - There are uncertainties regarding how robustly health insurance companies adhere to legal requirements for privacy of information.
 - The trend of employee wellness schemes includes forms of genetic testing (Silverman, 2015; Shaywitz, 2015) and encourages the purchase of nutritional supplements by the same company that is performing the testing (Genetic Alliance et al., 2008)

RECENT AND CURRENT TRENDS IN DIRECT-TO-CONSUMER NUTRIGENETIC TESTING

Success or Otherwise of Direct-to-Consumer Marketing Model

- To date, the DTC market for nutrigenetic testing has neither become a huge industry nor disappeared. Currently, genetic ancestry still dominates the field, with health-related testing forming a smaller niche (Turrini and Prainsack, 2016). However, ancestry companies themselves are now looking toward gene testing for health (Hernandez, 2015b; Phillips, 2016).

Breadth of Use of Direct-to-Consumer Marketing Model

- The growing trend is for DTC genetic testing to be offered by alternative and complementary practitioners such as naturopaths, homeopaths, and chiropractors (Caulfield et al., 2015).
- The tendency for companies to offer personalized nutrition advice, meal plans, and food supplements has extended to some offering actual meals (<http://www.ingredientsnetwork.com/habit-reveals-personalised-nutrition-plans-news043497.html>).

Changes Away From Direct-to-Consumer Marketing Model

- There is a shift from the DTC to the direct-to-provider model (McGowan et al., 2014).
- Former DTC companies are moving, ceasing, or lessening their test offerings and joining either collaborative projects in research or those belonging to bigger biotech companies (Perbal, 2014a).
- Sometimes including a professional, which is often required, is suggested; the quality and training of professionals varies. However, it seems that consumers are tending to want health care professionals involved with genetic testing (Goldsmith et al., 2013).
- Practitioners in relevant professions are focusing on how to adopt genomics technologies and knowledge in more structured approaches. For example, the Academy of Nutrition and Dietetics in the United States has a specialty practice group of registered dietitians that focuses on the nutritional systems biology approach and its clinical application. This group of integrative practitioners, named Dietitians in Integrative and Functional Medicine (DIFM), rose to 3000 members in 2012. The stated mission of DIFM is “To empower members to be leaders in

personalized genomics, holistic care, and functional nutrition therapies” (Swift, 2012).

Range and Types of Tests Available

- There is a diversity of tests available (e.g., microbiome test DTC from companies such as Biome).
- Whole-genome and exome sequencing are overtaking SNP chip or single-gene tests (McGowan et al., 2014; Perbal, 2014a).
- Relevant technologies (Gibney et al., 2015; Kohlmeier et al., 2016; Perbal, 2014a) are being developed, such as Illumina’s app for exploring a real human genome.

Accessibility of Tests

- The cost of testing, including full-genome sequencing is dropping and will continue to fall (Wadhwa, 2015; Davies, 2010).
- An increasing number of genetic services are being covered by insurers (NIH US National Library of Medicine).
- The growing trend is for companies to offer GT to employees as part of wellness program and their like, including gene testing, to personalize diet and exercise recommendations (Silverman, 2015; Shaywitz, 2015; Steinberg et al., 2015).

LOOKING FORWARD

Likely Trends

- Giants such as Ancestry.com and IBM are likely to increase their presence in the health data marketplace (Hernandez, 2015a).
- The cost and time required for genotyping will continue to decrease (Grimaldi, 2014).
- It is likely that some information relating to genotype will be incorporated into public health guidelines and into recommended daily intakes of nutrients (Grimaldi, 2014).
- Dietitians can and will probably be primary professionals in the mix (Nielsen and El-Sohemy, 2012).
- Technology will continue to develop in scope, sophistication, and form regarding the way in which personalized nutrition information is communicated.

Where to Go From Here?

- A way forward that is both ethically sound and scientifically well-founded is to offer nutrigenetic testing for personalized dietary advice only when there is robust scientific evidence for health effects,

and to follow up tests with careful monitoring of behavioral and psychological effects (Gorman et al., 2013). This approach addresses the needs for clinical and health utility as well as professional interaction with consumers.

- It is essential for personalized advice to avoid paternalism and to support the autonomous choice of consumers (Gorman et al., 2013).
- It is important for genetic evidence to be assessed at the same level as phenotypic evidence (Gorman et al., 2013).
- Robust and unbiased information needs to be made available to health professionals and consumers so that they can make sound decisions regarding genetic testing (Grimaldi, 2014).
- There is a need to ensure that the difference between nutrigenetic testing and disease-risk testing is understood, while honoring the role that dietary choices can have in disease risk and development.
- Agreement is required regarding what professionals need to understand about testing and subsequent coaching. This will result in something like a registered society, so that consumers will know practitioners are qualified to be involved. Registration with such a society should include the requirement for continued uptake and application of newer nutrigenetic knowledge as it is manifest.
- Pieces that must be in place include:
 - adequate education and buy-in of the medical and professional community (Roberts and Ostergren, 2013); this needs to occur over a broad range of professions, including (but not limited to) family doctors, relevant specialist practitioners, genetic counselors, nutritionists, and dieticians
 - awareness of and willingness to incorporate numerous environmental factors
 - awareness of and willingness to incorporate an individual's microbiome
 - consideration of the complexities of how food compounds work, and the need to consider the food matrix rather than single nutrients: e.g., altering levels of food components can be harmful to health and/or negate benefits of other food components
 - awareness that rather than SNPs alone, the link between nutrients and metabolic pathways can affect disease outcome
 - education and support of the general public with regard to what nutrigenetics and nutrigenomics are, and how best to select a test (including choosing a process that encompasses appropriate professional support)
 - consistent regulations that apply as universally as possible
 - bioinformatic and mathematical tools that need to continue to develop to handle data and support

nutritional strategies arising from the use of nutrigenetic tools (Kohlmeier et al., 2016)

CONCLUSION

The horse has bolted so to speak, and it is not all bad. We believe that little will be gained in trying to stop consumers from accessing their individual data and endeavoring to use it to improve health. Rather, what is needed is an immediate investment of resources into cross-discipline consultation and regulation, plus concerted efforts to educate both the public and professionals to pave a surer and safer way to the future.

References

- 23andMe, March 30, 2015. 23andMe's Personal Genome Service® Is Now Available for Purchase at Superdrug. <http://mediacenter.23andme.com/en-eu> [press release].
- Ahlgren, J., Nordgren, A., Perrudin, M., Ronteltap, A., Savigny, J., van Trijp, H., et al., 2013. Consumers on the Internet: ethical and legal aspects of commercialization of personalized nutrition. *Genes Nutr* 8 (4), 349–355.
- Allison, M., 2012. Direct-to-consumer genomics reinvents itself. *Nat Biotechnol* 30 (11), 1027–1029.
- Ashton-Prolla, P., Goldim, J.R., e Vairo, F.P., da Silveira Matte, U., Sequeiros, J., 2015. Genomic analysis in the clinic: benefits and challenges for health care professionals and patients in Brazil. *J Commun Genet* 6 (3), 275–283.
- Bland, J., 2015. Functional medicine: an operating system for integrative medicine. *Integr Med* 14 (5), 18–20.
- Bray, M.S., Loos, R.J.F., McCaffery, J.M., Ling, C., Franks, P.W., Weinstock, G.M., et al., 2016. NIH working group report—using genomic information to guide weight management: from universal to precision treatment. *Obesity* 24 (1), 14–22.
- Camp, K.M., Trujillo, E., 2014. Position of the academy of nutrition and dietetics: nutritional genomics. *J Acad Nutr Diet* 114 (2), 299–312.
- Caulfield, T., Borry, P., Toews, M., Elger, B.S., Greely, H.T., McGuire, A., 2015. Marginally scientific? Genetic testing of children and adolescents for lifestyle and health promotion. *J Law Biosci* lsv038.
- Davies, K., July 29, 2010. Consumer Genetics Puts on a Show: Big Genomics News from Illumina. GnuBio. Available from: <http://www.bioitworld.com/2010/issues/julaug/consumer-genetics.html>.
- De Caterina, R., El-Soheily, A., 2016. Moving towards specific nutrigenetic recommendation algorithms: caffeine, genetic variation and cardiovascular risk. *J Nutrigenetics Nutrigenomics* 9 (2–4), 106–115.
- Evans, J.P., to Parslow, V.R., 2015. Query Regarding FDA Regulation of Genetic Testing. Personal Communication.
- Federal Trade Commission: Consumer Information: Direct-to-Consumer Genetic Tests. Available from: <https://www.consumer.ftc.gov/articles/0166-direct-consumer-genetic-tests>.
- Fenech, M., El-Soheily, A., Cahill, L., Ferguson, L.R., French, T.A., Tai, E.S., et al., 2011. Nutrigenetics and nutrigenomics: viewpoints on the current status and applications in nutrition research and practice. *J Nutrigenetics Nutrigenomics* 4 (2), 69–89.
- Ferguson, L.R., De Caterina, R., Görman, U., Allayee, H., Kohlmeier, M., Prasad, C., et al., 2016. Guide and position of the international society of nutrigenetics/nutrigenomics on personalised nutrition: Part 1 - fields of precision nutrition. *J Nutrigenetics Nutrigenomics* 9 (1), 12–27.

- Finegold, D.L., Bensimon, C.M., Daar, A.S., Eaton, M.L., Godard, B., Knoppers, B.M., et al., 2005. *BioIndustry Ethics*. Academic Press.
- Genetic Alliance, Terry, S.F., President & CEO Genetic Alliance, et al., January 28th, 2016. Letter to U.S. Equal Employment Opportunity Commission: Comments on Proposed Rule, Amendments to Regulations under the Genetic Information Nondiscrimination Act of 2008, RIN 3046--AB02 – Attachment B, GINA NPRM Sign on [Internet].
- Gibney, M., Walsh, M., Goosens, J., 2015. Personalized Nutrition: Paving the Way to Better Population Health. White paper from the Food4Me Project No: 265494. <https://repositorio-abertoupt/bitstream/10216/82229/2/130854pdf>.
- Goldsmith, L., Jackson, L., O'Connor, A., Skirton, H., 2013. Direct-to-consumer genomic testing from the perspective of the health professional: a systematic review of the literature. *J Commun Genet* 4 (2), 169–180.
- Görman, U., 2006. Ethical issues raised by personalized nutrition based on genetic information. *Genes Nutr* 1 (1), 13–22.
- Gorman, U., Mathers, J.C., Grimaldi, K.A., Ahlgren, J., Nordstrom, K., 2013. Do we know enough? A scientific and ethical analysis of the basis for genetic-based personalized nutrition. *Genes Nutr* 8 (4), 373–381.
- Grimaldi, K.A., 2014. Nutrigenetics and personalized nutrition: are we ready for DNA-based dietary advice? *Pers Med* 11 (3), 297–307.
- Health & Medicine Week, March 13, 2015. 23andMe authorized by FDA to market first direct-to-consumer genetic test, p. 436. Available from: <http://www.newsrx.com/newsletters/Health-and-Medicine-Week.html>.
- Hernandez, D., October 22, 2015. Thanks to the Feds, Competitors Dethroned 23andMe as the King of Consumer Genetics. Available from: <http://fusion.net/story/218329/23andme-dna-test-revival>.
- Hernandez, D., April 3, 2015. Ancestry.com Is Quietly Transforming Itself into a Medical Research Juggernaut. Available from: fusion.com.
- Howard, H.C., Borry, P., 2012. To ban or not to ban? *EMBO Rep* 13 (9), 791–794.
- Kalokairinou, L., Howard, H.C., Borry, P., 2015. Current developments in the regulation of direct-to-consumer genetic testing in Europe. *Med Law Int*, 0968533215618441.
- Kang, J.X., 2014. Future directions in nutrition research. *J Nutrigenetics Nutrigenomics* 6 (4–5), I–III.
- Kohlmeier, M., De Caterina, R., Ferguson, L.R., Görman, U., Allayee, H., Prasad, C., et al., 2016. Guide and position of the international society of nutrigenetics/nutrigenomics on personalized nutrition: Part 2—ethics, challenges and endeavors of precision nutrition. *J Nutrigenetics Nutrigenomics* 9 (1), 28–46.
- Lab Tests Online. FDA Approves First Direct to Consumer Genetic Test to Determine Carrier Status, March 16, 2015. Available from: <https://labtestsonline.org>.
- Lintott, C., Dixon, J., 2014. Genetic Health Services New Zealand: What You Need to Know.
- Madden, J., Williams, C.M., Calder, P.C., Lietz, G., Miles, E.A., Cordell, H., et al., 2011. The impact of common gene variants on the response of biomarkers of cardiovascular disease (CVD) risk to increased fish oil fatty acids intakes. *Annu Rev Nutr* 31, 203–234.
- McGowan, M.L., Fishman, J.R., Settersten Jr., R.A., Lambrix, M.A., Juengst, E.T., 2014. Gatekeepers or intermediaries? The role of clinicians in commercial genomic testing. *PLoS One* 9 (9), e108484.
- Middleton, A., 2012. Communication about DTC testing: commentary on a 'family experience of personal genomics'. *J Genet Couns* 21 (3), 392–398.
- National Human Genome Research Institute. Regulation of Genetic Tests: FDA Plans to Regulate LDTs. Available from: <http://www.genome.gov/10002335>.
- Nicol, D., Hagger, M., 2013. Direct-to-consumer genetic testing—a regulatory nightmare? Will the current framework protect consumers effectively? *Med J Aust* 198 (9), 501–502.
- Nielsen, D.E., El-Sohehy, A., 2012. Applying genomics to nutrition and lifestyle modification. *Pers Med* 9 (7), 739–749.
- NIH US National Library of Medicine. Will health insurance cover the costs of genetic testing? Available from: <https://ghr.nlm.nih.gov/primer/testing/insurancecoverage>.
- Park, A., October 21, 2015. Genetic testing company 23andMe returns to market. Time.
- Paul, N.W., Banerjee, M., Michl, S., 2014. Captious certainties: makings, meanings and misreadings of consumer-oriented genetic testing. *J Commun Genet* 5 (1), 81–87.
- Paynter, N.P., Chasman, D.I., Paré, G., Buring, J.E., Cook, N.R., Miletich, J.P., et al., 2010. Association between a literature-based genetic risk score and cardiovascular events in women. *JAMA* 303 (7), 631–637.
- Perbal, B., 2014. Communication is the key - Part 2. *J Cell Commun Signal* 8 (4), 275–287.
- Perbal, B., 2014. Communication is the key. *J Cell Commun Signal* 8 (4), 275–287.
- Phillips, A.M., 2016. Only a click away—DTC genetics for ancestry, health, love... and more: a view of the business and regulatory landscape. *Appl Transl Genom* 8, 16–22.
- Pollack, A., June 12, 2010. FDA faults companies on unapproved genetic tests. *N Y Times*.
- Roberts, J.S., Ostergren, J., 2013. Direct-to-consumer genetic testing and personal genomics services: a review of recent empirical studies. *Curr Genet Med Rep* 1 (3), 182–200.
- San-Cristobal, R., Milagro, F.I., Martinez, J.A., 2013. Future challenges and present ethical considerations in the use of personalized nutrition based on genetic advice. *J Acad Nutr Diet* 113 (11), 1447–1454.
- Saukko, P., 2013. State of play in direct-to-consumer genetic testing for lifestyle-related diseases: market, marketing content, user experiences and regulation. *Proc Nutr Soc* 72 (01), 53–60.
- Saukko, P.M., Reed, M., Britten, N., Hogarth, S., 2010. Negotiating the boundary between medicine and consumer culture: Online marketing of nutrigenetic tests. *Soc Sci Med* 70 (5), 744–753.
- Shaywitz, D., December 18, 2015. The science – or lack of it – behind genetic tests offered in the workplace. *Forbes*.
- Silverman, R., December 16, 2015. Genetic testing may be coming to your office; a handful of firms are offering employees free or subsidized tests for markers linked to metabolism, weight gain and overeating. *Wall Str J*.
- Skirton, H., Goldsmith, L., Jackson, L., O'Connor, A., 2012. Direct to consumer genetic testing: a systematic review of position statements, policies and recommendations. *Clin Genet* 82 (3), 210–218.
- Steinberg, G., Scott, A., Honcz, J., Spettell, C., Pradhan, S., 2015. Reducing metabolic syndrome risk using a personalized wellness program. *J Occup Environ Med* 57 (12), 1269–1274.
- Stewart-Knox, B., Rankin, A., Kuznesof, S., Póinhos, R., de Almeida, M.D.V., Fischer, A., et al., 2015. Promoting healthy dietary behaviour through personalised nutrition: technology push or technology pull? *Proc Nutr Soc* 74 (02), 171–176.
- Su, Y., Otte, I., Borry, P., Howard, H., September 2013. "It's our DNA, we deserve the right to test!" A qualitative analysis of a petition for the right to access direct-to-consumer genetic testing without the intermediate of a health care professional. *Per Med* 10 (7), 729–739.
- Swift, K.M., 2012. The changing landscape of nutrition and dietetics: a specialty group for integrative and functional medicine. *Integr Med* 11 (2), 19.
- ter Meulen, V., Boccia, S., Cornel, M., Delpech, M., De Paepe, A., Estivill, X., et al., 2012. Direct-to-consumer genetic testing for health-related purposes in the European Union. *EASAC Policy Rep* 18.
- Turrini, M., Prainsack, B., 2016. Beyond clinical utility: the multiple values of DTC genetics. *Appl Transl Genom* 8, 4–8.

- Vorhaus, D., MacArthur, D., Jostins, L., June 16, 2011. DTC Genetic Testing and the FDA: Is There an End in Sight to the Regulatory Uncertainty? Available from: <http://genomesunzipped.org/?s=DTC+Genetic+Testing+and+the+FDA%3A+is+there+an+end+in+sight+to+the+regulatory+uncertainty%3F>.
- Wadhwa, V., March 3, 2015. As FDA loosens its grip on consumer DNA tests, a win for innovation. Wash Post.
- Webborn, N., Williams, A., McNamee, M., Bouchard, C., Pitsiladis, Y., Ahmetov, I., et al., 2015. Direct-to-consumer genetic testing for predicting sports performance and talent identification: consensus statement. *Br J Sports Med* 49 (23), 1486–1491.

This page intentionally left blank

Privacy Risks and Protective Measures

Francesco Visioli^{1,2}

¹Department of Molecular Medicine, University of Padova, Padova, Italy; ²IMDEA-Food, CEI UAM+CSIC, Madrid, Spain

Glossary

IT, information technology The use of computers to store, manage, and calculate large amounts of data.

Data sharing The practice of making one's own data available to other investigators to advance knowledge further. Current advocates of the correct use of data sharing include the International Committee of Medical Journal Editors: for example, the *New England Journal of Medicine* (see Longo and Drazen, *New Engl J Med*, 2016).

Big data Colloquial way to define a set of data so large they require appropriate software and handling for visualization and real-time analytics. Mostly used in predictive analysis.

Cloud computing The ensemble of computer system resources that can be operated on demand, usually via the Internet.

INTRODUCTION

The decision by the International Committee of Medical Journal Editors to make data sharing mandatory before clinical trials are published in major medical journals has important consequences in terms of trial planning and financing, as well as for future data storage and handling. On the one hand, this initiative will certainly increase the amount of information that can be extracted from a trial. On the other hand, we should be aware that deidentified patient data will eventually circulate among various research groups (which is exactly the goal of this program), but also that this process might spin out of control once subanalyses come into play.

In this chapter, I review some critical issues associated with privacy and data protection. Because the field of data storage and data mining is moving quickly and legislation often trails behind, many views are purely personal and therefore should be framed within the context of obtaining the most valuable clinical data with the minimum financial and labor efforts.

DATA IMMORTALIZATION

Progress in informatics allows for the storage, processing, and interchange of an ever-increasing amount of data. Because of the liquid form of current knowledge, in a way, such data become immortalized.

As put forward by other authors, data collected during clinical trials and/or genetic screenings will survive our life spans and become immortal. Of course, there is no such thing as nonmutation because, as the Buddha said in his final words, "All composite things (Saṅkhāra) are perishable"; it is just a matter of time. However, we should not devalue the problem of having personal data taking a life of their own and becoming available to more and more people, with consequences that are currently difficult to understand fully. Jorge Luis Borges wrote that "What is divine, terrible, and incomprehensible is *to know oneself immortal*." We now know that some important pieces of information (e.g., our own genetic code) might become available to many for years to come and might be used in various ways, only one of which is in the context of ecological studies of disease prevalence, for which they were originally collected. The long-term consequences of having genetic codes and biochemical parameters available to governmental bodies are, of course, unknown.

Finally, we should not rule out the possibility of data being hacked. It already happened and will certainly happen again (*vide infra*).

PRIVACY AND MARKETING

The (in)famous anecdote of the father who learned of his daughter's pregnancy when he received a coupon from a chain store indicates that data leaking occurs

and that any storage of large amounts of biochemical and genetic data is accompanied by some risk for (voluntary or involuntary) dissemination. Patients have the right to control their personal data (opt-in decision), which is also recognized by several legislators (one example is that of Italy, where Act 221/2012 recognizes the patient's control rights over his or her own electronic health records).

Medical data (genetic, but also biochemical) are stored for long periods of time, hence increasing their vulnerability. It is tempting to rely on cryptography as a surefire tool, but sooner or later it is likely that data will become available to unauthorized people. From a practical viewpoint, sharing encrypted data with other researchers precludes others' ability to perform meaningful analyses. One of the most logical solutions that has been advocated is that of data being handled by professional third-party agencies. Indeed, having a scientifically neutral yet information technology (IT)-savvy ombudsman or data keeper would certainly satisfy many of the privacy requirements. Yet, we should also foresee increased costs (for data storing and handing) and delays in data analyses once this third party has been charged with encryption, decryption, and subsequent distribution of data to authorized research bodies.

PROS AND CONS OF DATA AVAILABILITY

One of the most obvious consequences of data sharing is that research costs will decline if we can use data generated by a clinical trial to test hypotheses without the need to embark on an entirely new project. In this respect (and pushing data sharing to the limit), citizen scientists or amateur researchers or skilled statisticians might be able to obtain the maximum possible amount of information from a given clinical trial or observational study. The costs associated with subanalyses will drop, and the amount of results extracted from the accumulated data will climb, in turn making large trials cost-effective. The most obvious downside of this approach is the publication of fishing expeditions (i.e., trial results that are not hypothesis-driven and at risk for type 1 errors). In addition, we should anticipate, with such a freely accessible data availability, the publication of redundant papers. We already witness a steady increase in the publication of randomized controlled trials, which are currently in excess of 30,000/year.

All of this will likely cause publication pollution, and it will become increasingly difficult to discern publications based on original hypothesis from those that

try to take advantage of existing and publicly available data.

CLOUD COMPUTING AND PRIVACY

As mentioned, data availability in an electronic cloud will foster collaboration across health care institutions and greatly advance our clinical knowledge, although redundant data will certainly be generated and published. Indeed, the National Institutes of Health requires researchers to share patients' data with other researchers via ad hoc platforms. The amount of available data linked to an individual is increasing steadily and currently includes financial information, tax exemptions, diet, lifestyle, and social behavior in addition to health and genetic profiles. This will facilitate the development of precision medicine and targeted diagnosis and therapy, but it will concomitantly expose individuals to the risk for being identified and profiled for marketing purposes (see earlier discussion). The current cybersecurity mantra is "It is not whether you will be attacked, but when." Health care is one of the preferred targets for hackers of all kind, and medical data theft reportedly increased by over 21% in 2014.

What are the weak points of such data sharing, from a security viewpoint? In cloud computing, several nodes can become exposed to hacking. This includes researcher-to-patient, statistician-to-researcher, and researcher-to-researcher nodes, among others. One way to decrease the risk for data being stolen or hacked is to use digital signatures (which are different from electronic signatures in that they are cryptography-processed) when providing informed consent. This precaution will also limit the inclusion of false patients (i.e., sock puppetry), which can be generated by computer scripting in large trials. Patients' recruitment is progressively performed online via dedicated websites or social networks, which increases the risk for private data manipulation. Last but not least, we should be aware that a large proportion of medical doctors (either general practitioners or specialists) communicate among themselves, as well as with patients, by email, which further increases the risk for data leaking; this calls for additional cautionary security measures.

One final consideration concerns the adequate training of health care personnel. Not everyone is proficient in IT technologies and the present-day curricula do not address such issues in depth. Therefore, there is an urgent need to introduce IT security in university curricula and to train all staff members carefully who are prospectively involved in clinical trials or simple health data handling.

CONCLUSIONS

The debate on privacy and associated issues is as old as civilization. It probably started as a mean of protecting one's house and commerce, then evolved into more variegated expressions, which now involve self-determination (i.e., how much of our selves is communicated to others). This includes personal beliefs, political or religious, sexual orientation, as well as biometrical and clinical data. It all boils down to what American lawyers Samuel Warren and Louis Brandeis, in *The Right to Privacy* (1891), stated: "Privacy is the right to be let alone." Yet, in an ever-connected world, it becomes easy for information to circulate without much control over it. Biochemical data repositories have already been hacked and chances are that skilled hackers will reiterate that in the future. Personal health data can be exploited by a variety of subjects and organizations, including, but not limited to, insurance companies, employers, and pharmaceutical and food companies.

There is no easy way out of this. Increased data sharing, with all the caveats exemplified earlier, will foster clinical research and in turn benefit patients. On the other hand, we might have to surrender part of our privacy in a way similar to what happened after the 9/11 terrorist attacks. In a nutshell, the issue of privacy risk and data handling appears to be fluid and evolving

with time, along with technology and awareness. Long-term predictions are impossible, and it might be better to rely on scientific integrity, political wisdom, and a mutual sense of progress to harmonize one's right to confidentiality with population well-being.

Suggested Readings

- Costa, F.F., 2014. Big data in biomedicine. *Drug Discov Today* 19, 433–440.
- Filkins, B.L., Kim, J.Y., Roberts, B., Armstrong, W., Miller, M.A., Hultner, M.L., Castillo, A.P., Ducom, J.C., Topol, E.J., Steinhubl, S.R., 2016. Privacy and security in the era of digital health: what should translational researchers know and do about it? *Am J Transl Res* 8, 1560–1580.
- Garrison, N.A., Sathe, N.A., Antommaria, A.H., Holm, I.A., Sanderson, S.C., Smith, M.E., McPheeters, M.L., Clayton, E.W., 2015. A systematic literature review of individuals' perspectives on broad consent and data sharing in the United States. *Genet Med* 18 (7), 663–671.
- Hoffman, S., 2016. The promise and perils of open medical data. *Hastings Cent Rep* 46, 6–7.
- Tucker, K., Branson, J., Dilleen, M., Hollis, S., Loughlin, P., Nixon, M.J., Williams, Z., 2016. Protecting patient privacy when sharing patient-level data from clinical trials. *BMC Med Res Methodol* 16 (Suppl. 1), 77.
- Vayena, E., Gasser, U., 2016. Between openness and privacy in genomics. *PLoS Med* 13, e1001937.
- Zarate, O.A., Brody, J.G., Brown, P., Ramirez-Andreotta, M.D., Perovich, L., Matz, J., 2016. Balancing benefits and risks of immortal data: participants' views of open consent in the personal genome project. *Hastings Cent Rep* 46, 36–45.

This page intentionally left blank

Ethical Considerations in Nutrigenetics and Nutrigenomics

Ulf Görman¹, Jennie Ahlgren², Karin Nordström³

¹Centre for Theology and Religious Studies, Ethics division, Lund University, Lund, Sweden ²School of Education and Communication, Jönköping University, Jönköping, Sweden ³Theologische Fakultät, Universität Luzern, Luzern, Switzerland

Glossary of Terms

EGAPP Working Group The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative, established by the US Office of Public Health Genomics at the Centers for Disease Control and Prevention, supports the development and implementation of a rigorous, evidence-based process for evaluating genetic tests and other genomic applications for clinical and public health practice in the United States.

Food4Me Food4Me was a 4-year project that received funding from the European Commission. The main objectives of the project were to explore the scientific, business, and consumer aspects of personalized nutrition and to determine whether dietary advice, including knowledge of a person's genes, could deliver consumer benefits.

Oviedo Convention The Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine (ETS No 164) was opened for signature among the member states of the Council of Europe on Apr. 4, 1997 in Oviedo, Spain.

PKU Phenylketonuria is an inborn error of metabolism that results in decreased metabolism of the amino acid phenylalanine.

SNP A single-nucleotide polymorphism is a variation in a single nucleotide that occurs at a specific position in the genome.

INTEGRITY RISKS

Sensitive Data and Protection

Personalized nutrition is not a new concept, nor is it limited to nutrigenetics and nutrigenomics. Personalization of food advice related to lifestyle and health circumstances is common. Is genetic information inherently unique, so that it should be treated differently from other forms of personal or medical information, or is it just another new type of information to which we

need to become accustomed? This question has been asked since personal genetic data became available. The answer may not be found in a simple classification of the character of such knowledge. It is evident that genetic information is generally perceived to be particularly complex and personal. This understanding may be clarified as groups of people become more accustomed to take in and use it. What remains, however, is that genetic information is particularly personal and sensitive, because it uniquely identifies a person, and some of this information can have far-reaching implications for personal life as well as for close relatives.

Genetic data are sensitive information and must be treated with utmost care. Procedures must be set up to protect the privacy of such personal data. Within health services, attention is generally paid to this need, even if it is often weakened by the need to share such information within the profession. As the collection and further use of genetic information become more common, there is a risk that such confidentiality is not adequately upheld. Nevertheless, genetic information may be interesting to outsiders with questionable reasons for access, such as insurance companies and current or future employers. Revelation of such information can have far-reaching implications for those concerned.

Personalized nutrition services require personal health data to be processed. Such data, in particular genetic information, is of a sensitive character and must be treated with adequate protection. Regardless of whether they are genetic, phenotype, or lifestyle data, these data have to be collected and used in an ethically responsible way. All personalized nutrition services should include a transparent strategy of how the health data provided by the customer are stored and for what purpose they

are used. The integrity of the person providing personal health data has to be adequately protected.

Most countries have strict regulations for personal data protection, in addition to collection, storage, disclosure, and dissemination. All persons involved in processing personal health data must be fully aware of their sensitive character and take necessary steps to maintain a satisfactory level of security to protect their confidentiality. National legislation, international conventions, and other relevant agreements must be followed.

Relatives

Genetic test results can reveal predispositions to a variety of diseases and health problems for the tested individual; however, there is also a probability that his or her relatives share the same hereditary characteristics. Consequently, it is important whether or how genetic information should be shared with close relatives. Feelings of care and goodwill can easily come into conflict with respect for the autonomy of others. Relatives' autonomy and the possible wish to know or not know about future risks should be respected.

In specific circumstances, the involvement of relatives may be necessary. For instance, in the case of a new diagnosis of colorectal cancer, the Evaluation of Genomic Applications in Practice and Prevention Working Group recommended a genetic test for Lynch syndrome, a hereditary type of colorectal cancer, to reduce morbidity and mortality in relatives, because the etiology is a mutation that can reliably be identified, and relatives inheriting the mutation have a high risk for developing colorectal cancer. In cases of genetics-based nutrition advice, however, the evidence base is still weak and much more evidence is required regarding the causal relations among genetic polymorphism, food intake, and health to take such actions, which may unnecessarily alarm relatives. It cannot be considered ethically responsible to transfer personalized nutrition recommendations to relatives, or to recommend them to take such tests, without further validation and analysis by expert groups. If relevant, such disclosure to relatives should be mediated by medical professionals with adequate genetic education, and after informed consent.

Results from our study indicated that parents may be interested in testing their children as a basis for an early start to personalized nutrition. Such an action would counteract the child's ability to make an autonomous and informed decision at a later time. From an ethical point of view, it is reasonable to argue in a similar way as in a clinical setting, that the child's autonomy should be respected and given priority, and exceptions should be made only in cases in which an early start to an intervention may reduce serious health consequences.

Incidental Findings

Several studies indicated that the possibility of tailoring one's food habits to one's genes is attractive to persons who are informed about this idea and its prospects. A frequent first reaction seems to be optimistic and related to the prospect of gaining empowerment over one's health. However, some gene–diet–health connections offer a more complex scenario. The *APOE* gene provides instructions for making apolipoprotein E, which regulates the forming of lipoproteins. In the practice of gene-based personal nutrition, it is commonly noted that carriers of the variant *APOE* e4 have a higher risk than the average population for gaining high levels of cholesterol. Such carriers are then advised to take precautions by decreasing saturated fat intake and being physically active. However, *APOE* e4 is also related to a higher risk for Alzheimer disease and subsequent dementia, which is currently not a curable disease, although some means are available for palliative treatment. Is it then in line with ethical responsibility to offer information to a consumer that is connected to a higher risk for a serious incurable disease? Even if the dietician or physician, or the commercial company, who submits the information avoids mentioning this aspect, the person concerned may still find this information easily, such as on the Internet when Googling for more information.

According to the Oviedo Convention on Human Rights and Biomedicine, Article 10.2, “[E]veryone is entitled to know any information collected about his or her health. However, the wishes of individuals not to be so informed shall be observed.” This article has been difficult to implement, and there is an extensive discussion about how to develop a good practice in this respect. It is also ethically questionable to give a person unsolicited information of this kind, and the right to know or not know should be respected. A general solution may be to bring up this matter in connection to initial information and consent. In personalized nutrition, the complexity of such side effects must receive attention, and whenever relevant, this question should be handled at an early stage. Information should be given that a genetic analysis may reveal uncomfortable information about personal properties and health risks, and the question may be asked whether the recipient wants to receive such information.

SOCIAL RISKS

Social Aspects of Eating

Individual attitudes toward eating and drinking are divergent, but the social dimension is also evident. Food evokes enjoyment as well as cultural and personal identity. Food is an important aspect of human

happiness and well-being, not only an instrument for health, and meals are often social events, important manifestations of the relationship with others. Individualization of diet based on genetic information adds a complication that may negatively affect the social aspects of eating. However, the individual choice of food in a social setting has become much more common and accepted. Such an individual choice is often based on personal taste references as well as individual understandings of the health effects of food, environmental considerations, and moral convictions. In many cases, genetics-based personalization of food may result in minor modifications of population-based recommendations, and more seldom, in far-reaching alterations of food composition. Factors such as these may mitigate the social complexity of adding one more basis for the individual choice of food in a social setting.

Individual Responsibility for Health

The availability of personalized, adequate information about a healthy diet and lifestyle may be perceived as empowerment for the individual. Once a person has relevant and precise information about his or her personal health risks, as well as knowledge about dietary means to mitigate such risks, he or she is potentially able to make a difference in the health status. Personalized nutrition can thus be seen as a means of providing information concerning people's diet by which they can better influence and control their health.

However, this improved information and control creates expectations and potential intolerance toward people when information is not noticed and advice is not followed. Thus, by individualizing health, personalized nutrition is also part of the dilemma of individualization. On the one hand, information empowers the individual by improving the means to have influence and take control. On the other hand, it may limit personal freedom owing to expectations from society to follow advice and adapt proper dietary behavior. Reinforcing individual health control (and thus potentially individual autonomy) may provoke a clash between individual and societal interests with regard to responsibility for health. Discerning retrospective and prospective responsibility is ethically essential within the context of personalized nutrition. It means that when offered personalized nutrition services, people should be prospectively encouraged to improve diet and lifestyle, but not be blamed retrospectively.

Furthermore, personalized nutrition should offer truly targeted information, considering individual differences with regard to food traditions, as well as respecting variations of preferences and ambitions concerning well-being and health. A promising approach may be to focus personalized information on menu

choice and menu composition instead of single nutrients. Such communication tools may be automated and based on individual data input including health status, lifestyle, and specific genetic factors.

On a policy level, the legitimacy and implications of incentives, for instance, from health insurance companies or employers, have to be objectives of a continuous public debate.

Justice and the Risk for Discrimination

Investigations indicate that personalized nutrition has the potential to support a healthy diet and lifestyle. In the first place, this may involve a group of health seekers; it may also find support in a growing part of the population. However, such efforts may not involve everyone, which reinforces the problem of justice.

Many noncommunicable diseases (NCDs) are strongly influenced by environmental factors other than nutrition and lifestyle choices. This is particularly relevant for NCDs such as malnutrition, cardiovascular diseases, type 2 diabetes, and obesity. The prevalence of NCDs has a strong connection to socioeconomic factors. Information on such diseases and how to avoid them is easily available in society, but individual interest and the capacity to take in such information and integrate it into one's life, and subsequent compliance with advice based on such knowledge is divergent. This justice problem needs to be managed adequately within any responsible society.

Another problem is the complexity of information received through genetic investigations. Especially when based on genetics, personalized nutrition implies intricate information. Socioeconomic factors influence responsiveness to such advice. Depending on the context, there is a risk that personalized nutrition contributes to social inequalities. To promote social justice in health insurance and health systems, societies have to consider such differences and counterbalance them to promote solidarity and social justice. From an ethical point of view, this should influence the construction of health insurance and the distribution of resources when health systems set principles of priority.

The frequencies of specific genetic variants differ among ethnic groups, which also raises ethical issues. The discovery of a higher frequency of a specific genetic variant in an ethnic group can be greatly valuable for those concerned. But it may also be difficult to foresee whether findings of genetic differences among population groups can lead to discrimination. Care should be taken to avoid such consequences. An ethically responsible attitude is to be observant of this risk and to try to counteract misuse, in particular unjustified simplifications that can be used as a tool for discrimination. International and national recommendations and regulations suggest or require

precautions against discrimination and other aspects to protect persons whose personal data are processed. These concerns must be considered and national and professional restrictions must be followed.

QUALITY RISKS

Several studies show that most consumers initially have a positive attitude toward the concept of gene-based personalized nutrition. The possibility of having personalized nutritional advice based on genetic tests is often perceived as empowering, to make everyday food choices easier and contribute to good health and a sense of higher quality of life. Although consumers identify some risks with undergoing genetic tests, such as concerns related to confidentiality of the genetic information, and higher demands on the individual to take responsibility for his or her health, potential benefits are perceived as more important. However, it is not yet clear whether genetics-based personalized nutrition can live up to these expectations.

Evidence

Is the current scientific evidence on the interaction among genes, diet, and health sufficient to make an ethically responsible decision to offer personalized nutritional advice? There are strong arguments to be careful in this respect. Many studies on the interaction between nutrition and the genome have provided interesting results, both on how individual genotypes moderate the metabolism of different nutrients and on how food components influence gene expression. However, even if knowledge of the relations and interplay among genes, food, metabolism, and health is growing, this knowledge is still fragmentary. Only a limited number of genes have been investigated, and causality and multifactorial aspects such as the interplay between different genes remains poorly understood. For some monogenetic health problems, clear evidence for personalized adjustment of nourishment can be found, such as phenylketonuria and lactose intolerance, but knowledge is weaker when it comes to many multifactorial health problems.

Motivation

It is not well-known how personal food advice based on genetic information can influence motivation of the recipient of this information. A few early studies indicated that nutritional advice based on genetic information offers stronger support for motivation than ordinary population-based advice. The proof-of-principle study in the European Union FP7 project Food4Me indicated that participants who received personalized

nutrition advice kept healthier diets than did the control group that received only general population-based advice. However, no significant difference could be discerned, depending on whether the personal advice was based on diet and lifestyle alone and whether it was based on phenotypic measurements or genotype information. This indicates that there may be no added value in general with regard to motivation or compliance from using genomic information to personalize advice on food and lifestyle habits.

In this situation, genetic analysis may not provide an added incentive and it seems not to be a cost-effective way to motivate people to follow standard population-based nutritional recommendations, because lifestyle-based advice is cheaper, more easily understood, and has fewer ethical complexities. If so, there may be reason to focus the interest of genetics-based personalized nutrition on such cases where an intervention can clearly contribute with added value compared with other ways of personalization.

In some cases, evidence is available showing that individuals with specific genetic properties will benefit from deviating from customary dietary recommendations. One important example is the case of folate. Carriers of the 677T polymorphism of the *MTHFR* gene have reduced folate availability, resulting in a higher level of homocysteine, which is a risk factor for cardiovascular and other diseases. Such persons, in particular homozygote TT carriers, may benefit from a higher folate intake than standard population-based recommendations to keep homocysteine below risk levels.

Commercialization

Despite the common view among researchers that more research and knowledge are needed, several companies have been offering genetic testing to consumers for more than 10 years, often direct-to-consumer services based on a saliva sample sent by mail. After criticism, including from US public authorities claiming that the services were of inferior quality or lacked required authorization, most of these companies closed their businesses or reduced the services offered, but a limited number still offer genetic analyses online. Results are often delivered over email, sometimes with a personal contact, but in many cases without mediation by a health care provider. Their analyses include a wide range of results, such as risks for disease, kinship, and lifestyle advice, but also personalized nutrition advice.

Crucial quality problems connected to direct-to-consumer services remain. In addition to general problems mentioned elsewhere in this chapter, this is particularly relevant to privacy concerns, because the regulation of such services and their confidentiality of data is unclear

or inadequately perceived. Another important issue is the lack of adequate and transparent information about the benefits and limitations of the services.

MITIGATION

Precautionary Principle

The precautionary principle may be used as a tool to find a way through ethical concerns such as those mentioned earlier. This principle has often been promoted as a means for handling situations in which we only have limited knowledge of possible consequences that may follow from different actions taken. However, it has been understood in many ways, and some of these may be difficult to use or may have far-reaching consequences. For practical use, not least in cases such as these, we suggest a simple interpretation that looks back at the origin of the concept as a tool for prudent housekeeping. With this understanding, it is appropriate to find a balance between incomes and expenses by being careful not to overestimate incomes and not underestimate expenses. Translated to a risk–benefit scenario, benefits must not be overestimated, whereas risks must not be underestimated, and in a situation of insecurity, a margin must be kept for possible misjudgments. The balance between benefits and risks can also be influenced by mitigation measures (i.e., to improve the benefits and limit the risks). A number of such precautionary measures may be used, as discussed next.

Informed Consent

For a long time, informed consent has been regarded as an important tool to protect persons involved in biomedical activities, not only risky or burdensome interventions, but also in the case of retrieval and further processing of sensitive personal data. In many countries, informed consent is required to retrieve genetic information. For instance, in Europe, the treaty on genetic testing for health purposes declares that genetic probes may be carried out only under quality control and individual supervision, and after informed consent.

Because of the complexity and sensitivity of genetic information, there is reason to establish a practice in which recipients of genetics-based personalized nutrition advice always receive detailed information describing the character of the information to be received. This will need to include a description of the fragmentary knowledge of the causality, as well as the fact that the genetic analysis may entail uncomfortable information. This is a required practice in several professions, but it will need to be adopted even for others who may be involved in distributing such information.

A Wider Understanding of Personalization

A wider understanding of personalization, which includes all kinds of individual information, is relevant. All kinds of individually based and individually targeted advice needs to be handled with care and based on solid knowledge. This is a precondition for necessary trustworthiness. Because knowledge of the gene–diet–health connection is fragmentary, genetics-based dietary advice needs to be handled on a step-by-step basis, in which only connections that are well-understood are used for counseling.

Education of Professionals

Dietitians and other health professionals will need to follow developments in nutrigenetics and nutrigenomics. This may be supported by university courses and quality literature. For instance, the availability on the Internet of personalized nutrition services may result in a number of clients arriving at dietitians, asking for assistance in understanding the genetic information they have received after sending in their specimens. This probably means that dietitians need to follow developments in this area irrespective of whether they think it is a mature field.

Need for Revision of Legislation

Commercial products, adjusted to serve genetics-based personalized nutrition, are expected to appear. This already occurring, with examples such as vitamin supplements specifically designed to address single-nucleotide polymorphism–specific needs. The marketing of such products may be connected to more or less explicit promises of health benefits, as well as contrasting disclaimers indicating that the products are not intended to diagnose, treat, cure, or prevent any disease. Because consumer protection is crucial from an ethical point of view, such a situation brings to the fore the need for clear legal regulation. Policy makers should carefully follow developments, including new products and new routes of contact, marketing, and distribution, and update legislation accordingly.

Another development is the growing number of health and fitness applications, such as apps for mobile phones, personal digital assistants, and other wireless devices. Such health and well-being applications require updates of legal regulation to ensure their quality and transparency and thus achieve the necessary consumer protection.

The Food4Me project concluded that personalized nutrition services and products currently fall between the regulation of health care services and products and conventional business contracts that are direct-to-consumer. As a result, legislation applicable to personalized nutrition services and products has turned out to be

fragmented and to result in legal uncertainty with regard to protecting the consumer and the resulting obligations of personalized nutrition providers.

Step-by-Step Approach

The ethically responsible use of genetics-based dietetic advice must be based on a validated understanding of its effects. Such evidence is currently available in only a small number of cases. More evidence of the actual effect of dietary changes based on genetic information must be obtained before nutritional genomics is used more widely in dietetics practice. From an ethical point of view, it is advisable to introduce personal genome-based dietary advice on a step-by-step basis, in which such advice is offered only when clearly validated knowledge of causality and correlations becomes available. It is also important to observe how such information is received, and to be ready to assist in case of unforeseen behavioral and psychological effects.

Suggested Reading

- Ahlgren, J., Nordgren, A., Perrudin, M., Ronteltap, A., Savigny, J., van Trijp, H., Nordström, K., Görman, U., 2013. Consumers on the Internet: ethical and legal aspects of commercialization of personalized nutrition. *Genes Nutr* 8 (4), 349–355.
- Caulfield, T., McGuire, A.L., 2012. Direct-to-Consumer genetic testing: perceptions, problems, and policy responses. *Annu Rev Med* 63, 23–33.
- Camp, K.M., Trujillo, E., 2014. Position of the academy of nutrition and dietetics: nutritional genomics. *J Acad Nutr Diet* 114 (2), 299–312.
- Celis-Morales, C., et al., 2016. Effect of personalized nutrition on health-related behaviour change: evidence from the Food4Me European randomized controlled trial. *Int J Epidemiol* 1–11.
- Görman, U., Mathers, J.C., Grimaldi, K.A., Ahlgren, J., Nordstrom, K., 2013. Do we know enough? A scientific and ethical analysis of the basis for genetic-based personalized nutrition. *Genes Nutr* 8 (4), 373–381.
- Nordstrom, K., Coff, C., Jönsson, H., Nordenfelt, L., Görman, U., 2013a. Food and health: individual, cultural, or scientific matters? *Genes Nutr* 8 (4), 357–363.
- Nordstrom, K., Juth, N., Kjellström, S., Meijboom, F.L.B., Görman, U., 2013b. Values at stake: autonomy, responsibility, and trustworthiness in relation to genetic testing and personalized nutrition advice. *Genes Nutr* 8 (4), 365–372.
- Nordstrom, K., Goossens, J., 2016. Personalized nutrition and social justice: ethical considerations within four future scenarios applying the perspective of Nussbaum's capabilities approach. *J Agric Environ Ethics* 29, 5–22.

Impact of Nutrigenetics and Nutrigenomics on Society

Artemis P. Simopoulos

The Center for Genetics, Nutrition and Health, Washington, DC, United States

List of Acronyms

ApoE4	Apolipoprotein E4
DHA	Docosahexaenoic acid
DNA	Deoxyribonucleic acid
IL-1 β	Interleukin 1 β
LDL	Low-density lipoprotein
mRNA	Messenger RNA
PKU	Phenylketonuria
PUFA	Polyunsaturated fatty acids
SNP	Single nucleotide polymorphisms
VDR	Vitamin D receptor

Glossary

Metabolome	The total number of metabolites present within an organism, cell, or tissue.
Proteome	The entire complement of proteins that is or can be expressed by a cell, tissue, or organism.
Transcriptome	The sum total of all messenger RNA molecules expressed from the genes of an organism.

INTRODUCTION

The interaction of genetics and environment, nature and nurture, is the foundation for all health and disease (Simopoulos and Childs, 1990). Nutrition is an environmental factor of major importance. Major advances have occurred over the past 15 years in the fields of both genetics and nutrition. Methodological advances in molecular biology and genetics have facilitated the study of inherited disease at the DNA level, and the study of nutrients at the molecular level. This has led to (1) the development of concepts and research on genetic variation and dietary response (nutrigenetics, or individuals responding differently to the same diet by having different levels of, for example, serum cholesterol and blood pressure because of genetic variation); and (2)

studies on the evolutionary aspects of diet and the role of nutrients in gene expression (for example, polyunsaturated fatty acids [PUFA] suppress fatty acid synthase (messenger RNA [mRNA]) gene expression [nutrigenomics]). In addition to nutrients, nonnutritive dietary phytochemicals, such as phenolic compounds, are being studied for their effects on various aspects of human metabolism. Nutrigenomics could provide a framework for developing novel foods that will be genotype-dependent for promoting health and preventing and managing chronic diseases. In the United States and other countries, general dietary guidelines have been issued for preventing chronic diseases. In the development of dietary recommendations, the effects of genetic variation on dietary response have not been considered, despite such evidence (Simopoulos, 2010).

GENETIC VARIATION AND DIETARY RESPONSE

Genetics deals with variation. A fundamental aspect of the genetics approach to disease is an appreciation of human variation: its nature and extent, origin and maintenance, distribution in families and populations, interaction with environment, and consequences for normal development and homeostasis.

The extent of the genetic variation depends on how it is measured. At the level of DNA, genetic variation is considerable because in every 1000–2000 nucleotides there is a substitution leading to single nucleotide polymorphisms (SNP). SNPs refer to alterations of single bases (adenine, guanine, cytosine, or thymine) in the 1.83-m-long string of bases that make up human DNA. At the level of protein diversity (that is, variation in the sequence of amino acids), genetic variation is much less. In humans, approximately

30% of loci have polymorphic variants, defined as two or more alleles with frequencies of at least 1% or more in the population.

Advances in human biochemical genetics have produced data that suggest considerable biochemical variability within and between human populations (Simopoulos and Childs, 1990). Therefore, the relevance of this genetic information for human nutrition is considerable. Variation in nutritional requirements and the interaction of certain nutrients with genetically determined biochemical and metabolic factors suggest different requirements for individuals. This variation (like sex differences) is inborn and needs to be differentiated from variations caused by the life cycle (growth, pregnancy, and old age). Using the tools of molecular biology and genetics, research is defining the mechanisms by which genes influence nutrient absorption, metabolism and excretion, taste perception, and the degree of satiation, as well as the mechanisms by which nutrients influence gene expression. Furthermore, advances in molecular and recombinant DNA technology have led to exquisite studies in the field of genetics, as well as the recognition in a much more specific way, through DNA sequencing, of how unique each one of us is, and the extent to which genetic variation occurs in humans. The importance of the effects of genetic variation has been extensively studied and applied by pharmacologists in drug development and the evaluation of drug metabolism and adverse reactions to drugs. In the past 2 decades, physicians, geneticists, and nutritionists began to study the effects of genetic variation, and gene–nutrient interactions in managing chronic diseases (Goldbourt et al., 1994).

Data from around the world indicate that the incidence and prevalence of chronic diseases vary among individuals, families, and nations. Genetic predisposition, diet, physical activity, other environmental factors, and quality of care all contribute to these variations. Advances in genetics and molecular biology indicate that to a great extent, susceptibility to chronic diseases, such as coronary artery disease, hypertension, diabetes, obesity, osteoporosis, alcoholism, and cancer, is genetically determined. Because of genetic variation, not everybody is susceptible to chronic diseases to the same degree.

ETHNIC DIFFERENCES AND GENE VARIANTS

Although humans have evolved to be able to feed on a variety of foods and adapt to them, certain genetic adaptations and limitations have occurred in relation to diet. Because there are genetic variations among individuals, however, changes in dietary patterns affect a genetically heterogeneous population, although populations with

similar evolutionary background have more similar genotypes. Therefore, to be successful, dietary interventions must be based on knowing the frequency of genes the effects of which we are attempting to control or modify.

In northwest Europe, the frequency of phenylketonuria (PKU) is approximately one in 10,000 live births, but much lower in Africans and American Indians (Roberts et al., 1985). Some rare disorders in most populations are not as rare in some isolated populations. For example, hereditary tyrosinemia occurs in approximately 1 in 1000 live births in a French Canadian group, and hypertyrosinemia is more common in those of Finnish descent (Roberts et al., 1985). Hereditary fructose intolerance in Switzerland occurs in 1 in 20,000, and essential pentosuria in people of Jewish origin occurs at 1 in 2500 in some groups (Roberts et al., 1985).

Populations whose diet did not include wheat, barley, rye, or oats manifest gluten sensitivity upon the introduction of those foods in their diet. Celiac disease, which is triggered by the presence of gluten in the diet, occurs at 1 in 3000 live births in the United States, but 1 in 200 in Ireland. Removing gluten from the diet improves the general condition (Roberts et al., 1985).

Compared with Caucasians (15%), extremely high frequencies of the apolipoprotein E4Apo (E4) allele have been found in African and Asian populations, such as New Guineans (35%) and Nigerians (30%). Among European populations, northern countries have higher frequencies (22.7% in Finland and 20.3% in Sweden) than southern countries (9.4% in Italy), which suggests that Apo E4 may partly account for differences in cardiovascular disease prevalence in the two European regions. The relation between low-density lipoprotein (LDL) cholesterol levels and Apo E genetic variation depends on environmental and ethnic factors. The association of the Apo E4 isoform with elevated serum cholesterol levels is greater in populations consuming diets rich in saturated fat and cholesterol than in other populations. Data indicate that the higher LDL cholesterol levels observed in subjects carrying the Apo E4 isoform are manifested primarily in the presence of an atherogenic diet characteristic of certain societies, and that the response to saturated fat and cholesterol differs among individuals with different Apo E phenotypes (Simopoulos, 2010).

Racial differences have been noted in osteoporotic fracture risk. Fracture rates in Africans and Asians are considerably lower than in white populations despite low dietary intakes of calcium. Alleles of the vitamin D receptor (VDR) gene have been related to bone mineral density, bone turnover, and osteoporotic fracture risk. Age-related changes in bone mass and the influence of calcium intake on bone status have also been related to VDR alleles. Studies on the distribution of the VDR alleles show a higher frequency of the b allele in The Gambia

and China, where osteoporotic fractures are rare, and a much lower distribution in Cambridge, England. The bb distribution was 37.2% in England (similar to other reports of Northern European ancestry), 76.1% in The Gambia, and 85.3% in China, which suggests that the high frequency of the bb genotype in China and The Gambia may be associated with the low fracture incidence in these countries. The bb distribution in African Americans in Boston was 43.1%, which is much lower than in The Gambia but higher than in England.

THE ROLE OF NUTRIENTS IN GENE EXPRESSION

There has been an increase in the use of concepts evolved from molecular biology to the study of food components and essential nutrients as factors in the control of gene expression. In terms of chronic diseases, the effects of dietary cholesterol and fatty acids on gene expression are particularly relevant. Dietary cholesterol exerts a profound inhibitory effect on the transcription of the gene for β -hydroxy- β -methyl-glutaryl-coenzyme A reductase. Dietary PUFAs suppress the hepatic mRNA production of fatty acid synthase for lipoproteinemia in adult and weanling rats. This ability to suppress the abundance of mRNA for lipogenic proteins depends on the degree of fatty acid unsaturation. Eicosapentaenoic acid and docosahexaenoic acid (DHA) in the form of fish oils are thus more effective than arachidonic acid. Dietary omega-3 fatty acids reduce levels of mRNA for platelet-derived growth factor and for interleukin (IL)-1 β , indicating regulation at the transcriptional level. This area of research is expanding rapidly (Table 73.1) (Simopoulos, 1996).

THE FUTURE OF GENETIC NUTRITION: NUTRIGENETICS/NUTRIGENOMICS AND THE POTENTIAL FOR NOVEL FOODS

The main use of the human SNP map will be in dissecting contributions of individual genes to diseases that have a complex, multigene basis. Knowledge of genetic variation already affects patient care to some degree. For example, gene variants lead to tissue and organ incompatibility, affecting the success of transplants. Although the mainstay of medical genetics has been the study of the rare gene variants that lie behind inherited diseases such as PKU and cystic fibrosis, attention has focused on multigenic, multifactorial diseases such as cardiovascular disease, diabetes, cancer, obesity, and osteoporosis and their response to diet (Simopoulos, 2010). PKU, of course, was the first genetic disease for which screening programs were developed that was

controlled by diet (a low-phenylalanine diet) (Simopoulos et al., 1975). However, variations in genome sequences underlie differences in our susceptibility to (or protection from) all kinds of diseases, in the age at onset and sensitivity of illness, and in the ways our bodies respond to treatment with nutrients or drugs. By comparing patterns and frequencies of SNPs in patients and controls, research can identify which SNPs are associated with which diseases. Such research will enhance our understanding of both genetic medicine and genetic nutrition and will expand pharmacogenetics and nutrigenomics/nutrigenetics. This knowledge of our uniqueness will alter all aspects of medicine, nutrition, and pharmacotherapy. Because genes and genomes do not act in a vacuum, environmental factors are equally important in human biology. Genotyping will discover those who require special diets because of genetic variation; hence, diet will have to be modified to suit the individual or subgroup.

It is estimated that in the next few years, the pace of discovery of SNPs and developments in high-throughput genotyping may lead to the identification of many susceptibility genes for complex disorders. To take a hypothetical example, a 45-year-old man with hypertension would have a buccal scrape sample sent for genetic testing to enable molecular classification of the disease as type I, II, or III. He would then be prescribed a dietary regimen known to be effective in the relevant disease subtype (salt sensitive or not) with or without a specific drug. In a preventive setting, a 5-year-old boy whose older sibling has developed asthma might be tested (with parental consent and appropriate counseling) for asthma susceptibility genes. If the results were positive, he might be prescribed prophylactic drugs to prevent asthma. At the same time, he would receive dietary advice to increase omega-3 fatty acid intake in his diet (e.g., eat more fish) while decreasing omega-6 fatty acids by avoiding vegetable oils such as corn oil and soybean oil. It has been shown that decreasing the omega-6-omega-3 ratio by changing the background diets of persons with autoimmune disorders such as asthma, rheumatoid arthritis, and ulcerative colitis leads to decreases in IL-1 β and leukotriene B₄; improvements in clinical symptoms, and decreases in the dose of the nonsteroidal antiinflammatory agents. Genotyping will become part of routine management of an expanding range of human diseases over the decade, and nutrigenomics will supplement pharmacogenetics. Knowing who is at risk would be useful if it meant that one could avoid environmental triggers that convert susceptibility into disease.

How will genetics and nutrigenetics look in the future? Genetics will not remain the exclusive prerogative of regional genetic centers. Instead, every physician will need to use genetic knowledge and combine it with

TABLE 73.1 Effects of Polyunsaturated Fatty Acids on Several Genes Encoding Enzyme Proteins Involved in Inflammation, Adhesion Molecules, Cell Growth, Early Gene Expression, β -Oxidation, and Growth Factors.

Fatty acid	Cell growth and early gene expression c-fos, early growth response protein 1	Adhesion molecules VCAM-1 mRNA ^b	Inflammation interleukin-1B mRNA	β -Oxidation acyl-coenzyme A oxidase mRNA ^b	Growth factor platelet-derived growth factor mRNA
Linoleic acid				↑	
Linolenic acid				↑	
Arachidonic acid	↑	^c	↑	↑	↑
Eicosapentaenoic acid	↓	^a	↓	↑	↓
DHA	↓	↓	↓	↑↑	↓

↓ Suppress or decrease ↑ induce or increase; DHA, deoxyribonucleic acid; mRNA, messenger RNA; VCAM-1, vascular cell adhesion molecule 1.

^aEicosapentaenoic acid has no effect by itself but enhances the effect of DHA.

^bMonounsaturates also suppress VCAM-1 mRNA, but to a lesser degree than DHA, and induce acyl-coenzyme A oxidase mRNA.

^cArachidonic acid suppresses VCAM-1 mRNA, but to a lesser degree than DHA.

From Simopoulos AP. *Annals of Nutrition and Metabolism* (1996); 40:303–311.

appropriate dietary regimen, type, and amount of physical activity, and, if needed, drugs. For common adult-onset conditions such as diabetes and circulatory disorders, the interaction between genes and the environment is starting to be understood, and there is great interest in the potential for DNA diagnostics. Products are available that are enriched with various nutrients and nonnutrients: functional foods to prevent or treat disease. A major area reviewed is the enrichment of products with omega-3 fatty acids.

Pharmacogenetics is defined as the inherited basis for individual differences in drug response. It is currently dominating pharmaceutical research and development. Nutrigenetics is defined as the inherited basis for individual differences in nutrient (diet) response; it is only beginning to claim its potential. One can visualize the development of beverages and foods either as preventive agents or to treat individuals, families, or subgroups predisposed to a particular disease. There is a precedent for this in pediatrics, where ketogenic diets are used to treat patients with intractable epilepsy. Background diets balanced in essential fatty acids are paramount for patients with chronic inflammatory diseases, such as arthritis, asthma, ulcerative colitis, and lupus, as well as in patients with coronary artery disease and hypertension. Specific foods and diets are used for patients with celiac disease, PKU, and other single-gene diseases. From the marketing standpoint, Fig. 73.1 shows the flow of information from the identification of the individual based on genetic screening to marketing of the product. Table 73.2 shows expected achievements in genetics for the first quarter of the 21st century.

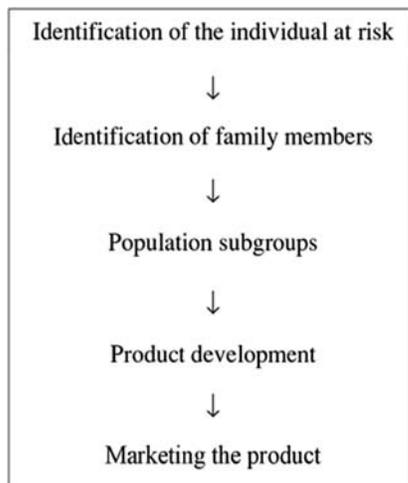


FIGURE 73.1 Flow of information from genetic screening to marketing of the product.

TABLE 73.2 Genetics: Expectations for the 21st Century.

First quarter

- Identify disease–gene associations for many illnesses and function of genes.
- Expand nutrigenomics and pharmacogenomics (treatment and prevention).
- Responsiveness to interventions (nutrients and drugs) will be predicted, because variation in the responses is often attributable to the genetic profile of the individual.
- Individualize prescriptions, diets, and lifestyle modifications and/or drug treatment.
- Develop gene-based designer diets for coronary artery disease, hypertension, diabetes, arthritis, asthma, mental health, etc.

SOCIETAL ISSUES

In the relation between nutrition and health as a result of genetic variants, it is necessary to develop a new concept of biomarkers, as well as to focus on the evolutionary aspects of diet. Major changes have taken place in our food supply to such an extent that current Western diets are not consistent with the diets humans evolved and to which our genes were programmed to respond. As an example, major changes have occurred in the production of oils from vegetables high in omega-6 fatty acids, as a result of the technology that made it possible to produce large amounts of oils from oil seeds such as corn, sunflower, safflower, soybean, cottonseed, etc. In addition to changes in the oils, there was a major change in animal feeds from eating grass (grazing) to grains high in omega-6 fatty acids. In the meantime, people ate less fish and more meat. Fish is high in omega-3 fatty acids, whereas meat from grain-fed animals is high in omega-6 fatty acids. These changes led to a ratio of omega-6 to omega-3 of 20:1, whereas during evolution the ratio was 1:1. The imbalanced omega-6 to omega-3 ratio led to a proinflammatory diet. A high omega-6 to omega-3 ratio and a proinflammatory diet are at the base of all chronic diseases, also known as diseases of civilization.

Functional genomics provides the tools to generate new hypothesis on the mechanisms of action of nutrients. A goal for applying genetics to nutrition science is to prevent diet-related diseases. Nutrients are dietary signals detected by cellular sensor systems that influence gene and protein expression, and subsequent metabolic products. Thus, new terms are being used, such as “transcriptome,” “proteome,” and “metabolome.” The National Institutes of Health’s (NIH, Bethesda, MD) NIH Road Map states that “to better understand the proteome, innovative tools must be developed that will enable researchers to determine in real time the

amounts, locations, and interactions of large numbers of individual proteins within a single cell.” We are just about there now. Researchers are eager for technologies that will enable them to measure local concentrations of carbohydrates, lipids, amino acids, and other metabolites within a single cell or even a specific part of a single cell. Molecular diagnostic tests eventually will have the potential to affect every area of health care, ranging from predicting who is at risk for developing disease, early diagnosis of disease, determining optimal treatment regimens based on nutrients, or a combination of drugs and nutrients, thus decreasing the dose of the drug and avoiding its side effects. This approach has both health and economic benefits to individuals, families, and society.

Academia and industry are carrying out studies to define candidate genes because the technology is available to generate data. What is lacking are models to interpret the data. Models are needed based on biological principles, not only mathematical and statistical ones. Diseases are indeed based on the interaction of genetic and many environmental factors. The frequency of genetic variants and the environmental factors vary in different populations. There are many association studies relating genes to risk factors for disease that cannot be reproduced and unconfirmed in other populations. Some scientists believe that increasing the number of subjects will take care of the problem. Others carry out metaanalyses. But if genetics deals with individuals, how can studies carried out in overall populations tell an individual’s risk? The “average individual,” such a popular term with epidemiologists, really does not exist in genetics. Nutrigenetics/nutrigenomics is about genetic individuality and nutrition.

The term “nutrigenetics” was first defined by R.O. Brennan in 1975 in his book *Nutrigenetics: New Concepts for Relieving Hypoglycemia*. It is essential that nutrigenomics consider the importance of ethical, legal, economic, scientific, and behavioral aspects of genetic screening, as recommended by the National Academy of Sciences National Research Council in its 1975 report, *Genetic Screening: Programs, Principles, and Research* (Simopoulos et al., 1975). Although a number of other reports have been published since 1975, this book remains a classic, as shown by the quotation:

This early classic in the field of genetic screening provides an ongoing framework to study the prospects, history, and development of principles, legislation, and program guidelines applicable to genetic testing aims, methodology and education. Ethical aspects are presented from the view of a ‘perfect’ screener, who would have all relevant facts to provide both error-free testing and effective counseling; possess a strong sense of the thoughts and emotions of those screened; be as free as possible from self-interest and inappropriate emotionalism; and apply principles consistently.

Scientific institutions and the World Health Organization have recognized the importance of the advances in human genetics to health. They have proposed the development of programs to enhance technical capacity in genetic science in the developing regions of the world, to reduce disparities in health status between developed and developing countries by using genetic science. In doing so, special consideration should be given to privacy issues and avoiding discrimination by health insurance companies and employers. Direct-to-consumer genetic testing has expanded; as a result, the need for evidence in genomic, nutrition, and genome medicine is more important than ever (Khoury, 2017; Simopoulos and Milner, 2010).

In the future, the focus will shift to prevention. It will be easier, indeed routine, to identify genetic predispositions in family members at risk, and probably those at risk in the population at large. As a consequence, it will be necessary to educate people at risk to the advantages of maintaining healthy lifestyles, avoiding risk, and seeking preventive therapies. Furthermore, knowledge of the genetic susceptibility to disease will identify those at higher risk, as well as their response to diet. As a result, there will be a need to develop novel foods targeted to individuals, families, and subgroups within populations.

As we advance our knowledge of gene nutrient interactions, society will need to create or use appropriate social, ethical, legal, educational, and economic frameworks to gain the benefits of such knowledge. Public health and regulatory processes will need to be established to define when genomic discoveries, such as gene–nutrient disease associations, are ready to be evaluated as potential tools to improve health screening and recommend dietary values.

CONCLUSIONS AND RECOMMENDATIONS

The new genetics has enormous implications for nutrition research both in preventing and managing chronic diseases. This new millennium is a remarkable time for medicine and nutrition, when the potential for improving human health is unequalled in history. With the unfolding genomic and technological revolution, continuing investments in research offer unprecedented opportunities to understand disease processes, prevent intrinsic and environmental risks to health, and develop new approaches to improve the quality of life worldwide. Although the emphasis of new genetics has been on pharmacogenetics, it is the responsibility of the nutrition scientists to expand the relation of genetics and nutrition, and in parallel to establish nutrigenomics and nutrigenetics as a major discipline in nutrition in

the 21st century. The International Society of Nutrigenetics/Nutrigenomics has a leadership role and is the premiere international organization in nutrigenetics and nutrigenomics.

There is no single universal approach for what we are calling the lifestyle approach to diseases with genetic predisposition. The approach will have to vary with national dietary patterns and national economy. Therefore, it will be necessary to promote lifestyle patterns that will be compatible with a healthier phenotypic expression of genotypes evolved under different conditions, which means individualized prescriptions and gene-based designer diets.

The prospect of targeting specific dietary treatment to those predicted to gain the most therapeutic benefit clearly has important clinical and economic consequences, particularly in diseases of high prevalence such as coronary artery disease, hypertension, diabetes, osteoporosis, and possibly cancer. Because the same genotype may not confer the same risk in all populations, populations should not copy each other's dietary recommendations for the prevention of coronary artery disease, and cancer, or any other disease, for that matter.

Universal dietary recommendations have been used by nutritionists who were concerned with undernutrition, but universal dietary recommendations are

inappropriate when the problem is one of overnutrition. Individual dietary recommendations considering genetic predisposition and energy expenditure are now easily foreseeable.

References

- Goldbourt, U., de Faire, U., Berg, K., 1994. Genetic Factors in Coronary Heart Disease. Kluwer, Dordrecht.
- Genetic variation and nutrition. In: Simopoulos, A.P., Childs, B. (Eds.), *World Rev Nutr Diet* 63, 1–300.
- Khoury, M.J., June 1, 2017. No shortcuts on the long Road to evidence-based genomic medicine. *J Am Med Assoc.* <https://doi.org/10.1001/jama.2017.6315>.
- Personalized nutrition. In: Simopoulos, A.P., Milner, J.A. (Eds.), 2010. *World Review of Nutrition and Dietetics*, vol. 101. Karger, Basel.
- Roberts, D.F., 1985. Genetics and nutritional adaptation. In: Blaxter, K., Waterlow, J.C. (Eds.), *Nutritional Adaptation in Man*. Libbey, London, pp. 45–59.
- Simopoulos, A.P., 1996. The role of fatty acids in gene expression: health implications. *Ann Nutr Metab* 40, 303–311.
- Simopoulos, A.P., 2010. Nutrigenetics/nutrigenomics. *Annu Rev Public Health* 31, 53–68.
- Simopoulos, A.P., Childs, B., 1975. Committee for the Study of Inborn Errors of Metabolism, Division of Medical Sciences, Assembly of Life Sciences, National Research Council. *Genetic Screening, Programs, Principles and Research*. National Academy of Sciences, Washington, DC.

This page intentionally left blank

Index

'Note: Page numbers followed by "f" indicate figures and "t" indicate table.'

- A**
ABCA1, 452
ABCG1, 452
N-Acetyltransferase-2, 515
Activated protein kinase (AMPK), 329
Acyl-CoA:retinol acyltransferase (ARAT), 241
Adaptive thermogenesis, in humans, 178
ADB1B gene, 13
5' adenosine monophosphate-activated protein kinase (AMPK), 398–399
ADH1B, 231–232
Adipogenesis, 290–291
ADIPOQ, 328–329
Adipose tissue, 373–375
 obesity, 432–433
 Omega-3 polyunsaturated fatty acids, 472–476
β2-Adrenergic receptor gene (*ADRB2*), 178
Adrenoceptors, 505
Adsorption methods, 43
Agilent 2100 Bioanalyzer, 44–45
Aging. *See* Calorie restriction (CR)
Alcohol consumption, 231
 ADH1B and ALDH2, 231–232
 alcohol detoxification pathway, 232f
 CYP2E1, 233–234
 MTHFR C677T, 235
 PNPLA3, 234–235
 pre-miRNA-27A, 232–233, 233t
Alcohol dehydrogenase(s) (ADHs), 242
Alcohol dehydrogenase 1B gene (ADH1B), 13
Alcohol detoxification pathway, 232f
Alcoholic liver disease (ALD), 231
Aldehyde dehydrogenases (ALDHs), 242
ALDH2, 231–232
Allele counting, 526
Alleles, 4
Allergic diseases, 404–406
Allozyme, 13–14
Allozyme *6 (429 Ala), 15, 15f
Allozyme *7 (519 His), 15, 15f
Allozyme *16 (429 Ala/519 His), 15, 15f
ALOX5, 14
Amino acids glutamate, 161
 γ-Aminobutyric acid (GABA), 159–160
AMP-activated protein kinase (AMPK), 23
Amplicon sequence variants (ASVs), 82–83
AMY1 gene, 16
Analogous technology, 82
Anderson disease, 450
Angiotensinogen (*AGT*), 189
ANGPTL, 451–452
Antipsychotic-induced weight gain (AIWG), 523
APOA5, 101–102
ApoAV, 452
ApoB48, 12
APOB-3500 variant, 12
ApoCIII, 452
Apolipoprotein A-II (APOA2), 219–220
Apolipoprotein B (ApoB), 448–449
Apolipoprotein E (APOE), 101–102, 189, 227–229, 266, 356, 381
Appetite traits, 162–163
Area under the receiver operator curve (AUC), 76–77
L-Ascorbic acid (ASC), 279
Aspartate, 161
Atherosclerosis, 470–476
ATP-binding cassette subfamily G member 1 (ABCG1), 256
ATP-binding cassette subfamily G member G5 (*ABCG5*) and G8 (*ABCG8*), 229
Autism, 78
Autism spectrum disorder (ASD), 78
Autosomal dominant inheritance, 5, 5f
Autosomal recessive inheritance, 5, 5f
Autosomes, 4
- B**
BCO1, 238
BDNF. *See* Brain-derived neurotrophic factor (BDNF)
Biallelic variant, 12–13
Bile acids, 412
Bioinformatics, 50, 57
Bioinformatics data analysis tools, 84
Biomarkers, of dietary intake, 92–93
Bioremediation, 517–518
Biotin, 272–274, 273f
 deficiency and supplementation, 274
 energy metabolism, 274
 genetic disorders, 272–273
 nutrigenomics, 273–274
 sources and requirements, 274
Bitterness, 160
Body composition
 adiposity traits, 168–170
 genetic variations, 167–168
 leanness phenotypes, 172
 lipodystrophy, 172t, 173
 monogenic and syndromic obesity, 170–172
 obesity, 167
Body mass index (BMI), 77, 181, 198f
Bone mass
 caffeine and vitamin D receptor genotype interaction, 380
 calcium, 492
 calcium and vitamin D interaction, 378–379
 fat intake-genotype visfatin interaction, 381
 fatty acids, 494
 modeling, 377
 osteoblasts, 490–492
 osteoclasts, 490
 osteocytes, 490
 phytochemicals, 493–494
 remodeling, 377
 remodeling cycle, 489–490
 soy and vitamin D receptor genotype interaction, 380
VDR gene, 492
vitamin B complex and methylenetetrahydrofolate reductase genotype interaction, 380–381
vitamin D and vitamin D receptor genotype interaction, 379–380, 492
vitamin K and ApoE genotype interaction, 381
vitamins A, 493
vitamins B, 493
vitamins C, 493
vitamins K, 493
Brain-derived neurotrophic factor (BDNF), 162, 176–178
Branched chain amino acids (BCAA), 77, 397, 443
Bread exposure, 93–94, 95f
- C**
Caffeine
 chemical structure, 336f
 health effects, 335–337
 interindividual variability in responses, 337–338
 mechanism of action, 335, 336f

- Caffeine (*Continued*)
 nutrigenetics, 338–339, 338f–339f
 source, 335
 and vitamin D receptor genotype interaction, 380
- Calcium
 bone metabolism, 492
 and vitamin D interaction, 378–379
- Calorie restriction (CR), 395–396
 aging, 398–399
 amino acid sensing, 397
 critical regulator, 396
 dietary amino acids mediate, 396–397
 nutrigenetics, 399–400
 rapamycin, 397–398
 type 2 diabetes (T2D), 444
- Cannabinoid receptor (CNR1) gene, 505
- Carbohydrate response element binding protein (ChREBP), 23, 206, 386–387, 443
- Carbohydrates
 absorption, 203, 204f
 concentration and transformation, 203
 different biological processes, 203
 intake and metabolism, 204–206
 macronutrients, 203
 nutrigenomics, 206–208
 obesity, 434–436
 relevance, 203–204
- Carcinogenesis
 environmental factors, 483
 molecular pathways, 485–486
 natural products and dietary constituents, 486–487, 486f
 nutriepigenetics, 483–485, 484f
 nutrigenomics, 483–485, 484f
 prevention, 483, 484f
- Cardiovascular disease (CVD), 211, 469
 apolipoprotein E (ApoE), 356
 epidemiologic studies, 355
 fatty acid desaturases, 356–357
 inflammation, 469–470
 5-lipoxygenase (5-LO), 357, 357f–358f
 manifestations, 355
 methylenetetrahydrofolate reductase (MTHFR), 357–359, 358f
 Omega-3 polyunsaturated fatty acids, 472–476
 peroxisome proliferator-activated receptors (PPARs), 359, 359f
 polyphenols, 476–480, 476f
 flavonols, 477
 phenolic alcohols and secoiridoids, 479–480
 stilbenes, 477–479
- L-Carnitine, 416
- Carnitine palmitoyltransferase (CPT), 328
- Carotenoids
 body transport and storage, 239–241
 enormous diversity, 237
 metabolism, 238
 retinoids
 absorption, 239–241
 conversion to, 238
 visual chromophore metabolism, 243–244
- Catechins, 329–331
- Catechol-O-methyltransferase (COMT), 329–331, 338
- Celiac disease, 418
- Cell-free mRNA, 43–44
- Cell-free nucleic acids, 43–44
- Cellular proteins, 18
- Ceruloplasmin gene (CP), 324
- ChIP-Chip, 54
- Chloride, 315
- Chloroform, 43
- Cholecystokinin (CCK), 161
- Cholesterol, 115
- Cholesterol 7 α -hydroxylase (Cyp7a1), 384
- Cholesteryl ester transfer protein (CETP), 229, 452
- Choline
 diets, 304–305
 functions, 304
 intake, 304
 next-generation nutrigenetics, 307
 nutrigenetics, 305–306
 nutrigenetics and metabolic variation, 303–304
 in pregnant women, 304
- ChREBP. *See* Carbohydrate response element binding protein (ChREBP)
- Chromatin accessibility, 38
- Chromatin immunoprecipitation and sequencing (ChIP-seq), 38
- Chromatin looping, 19
- Chromatin remodeling and remodelers, 20
- Chromosomes, 4
- Chronic kidney disease (CKD), 292
- Chronodisruption(CD)
 central clock genes, 143–144
 experimental animals mutations, 143
 genetics and weight loss, 144–146
 genetic variations, 143–144
 definition, 142
 epigenomics, 146–149
 nutrigenetics, 146, 147t–148t
- Chylomicron triglycerides, 256–257
- Circular mRNA, 14
- Circular RNAs (circRNAs), 41
- Circulating homocysteine (Hcy), 235
- Circulating nucleic acids. *See* Cell-free nucleic acids
- Classical eating disorders, 163
- Classical epidemiologic approaches, 100
- Clinical Pharmacogenetics Implementation Consortium (CPIC), 520
- Cloud computing, 540
- Cobalamin. *See* Vitamins B₁₂
- Coding variants, 8
- Codominant, 4
- Coenzyme A (CoA), 276
- Coffee, 337
- Cognition
 development, 350–351
 maternal *FADS* genotypes, 351–353, 351f
 research, 347–348
- Complementary DNA (cDNA), 13
- Computer-based meal planning, 511
- Congenital generalized lipodystrophy (CGL), 173
- Contigs, 83
- Continuous traits, 4–6
- Copper
 distribution, 323
 excessive accumulation, 323
 genetic variants, 323–324
 sources, 323
- Copy number variations (CNVs), 8, 15–16, 34–35, 54
- Coregulators, 19, 21
- CpG. *See* Cytosine-phosphate-guanine (CpG)
- CpG islands (CGIs), 131–132
- CR. *See* Calorie restriction (CR)
- C975T, 163
- CUBN, 250
- Curcumin, 331–332, 333f
- CVD. *See* Cardiovascular disease (CVD)
- Cyclic adenosine monophosphate (cAMP), 290
- CYP27B1, 250
- CYP2C9, 266
- CYP2E1, 233–234
- CYP2R1, 249
- Cystathionine β -synthase (CBS), 301
- Cytochrome 2C9, 521–522
- Cytochrome 4F2, 521–522
- Cytochrome P450 (CYP), 515
- Cytochrome P1A2, 522–523
- Cytochrome P3A4 (CYP3A4), 521
- Cytochrome P450 1A2 (CYP1A2), 338
- Cytochrome P2C19, 522–523
- Cytochrome P450 2C19 (CYP2C19), 416
- Cytochrome P450-4F2 (CYP4F2), 256–257
- Cytochrome P450 family 7 subfamily A member 1 (CYP7A1), 227–229
- Cytosine-phosphate-guanine (CpG), 484–485
- D**
- Data-driven analysis methods, 31
- Data immortalization, 539
- Dehydroascorbic acid (DHA), 279, 347–348
 influence of genetic variants, 349–350
- Deoxynucleic acid (DNA), 11
 extraction, 84
 sequencing, 84
- Developmental origins of health and disease (DOHaD), 122–123
- DHA. *See* Dehydroascorbic acid (DHA)
- DHCR7, 248–249
- DHFR rs7387829, 14
- Diabetes Prevention Program (DPP), 200
- Diacylglycerol (DAG), 77
- Dietary folate, 300–301

- Dietary intake
 biomarkers, 92–93
 quantitative assessment of, 108
- Dietary Reference Intakes (DRI), 509–510
- Dietary regulation
 chromatin regulators, 24
 gene expression, 21–22, 22f
 miRNAs modulators, 24
 transcription factors (TFs), 22–24
 abundance, 23
 activity, 22–23
 metabolic fluxes, 23
 nutrient/energy sensing signaling pathways, 23–24
- Dietary supplements, 154
- Diet-heart hypothesis, 211–212
- Diet-induced molecular processes, 373
- Differential methylation hybridization (DMH), 65
- Digital polymerase chain reaction (dPCR), 46
- Dihydrofolate reductase (DHFR), 300
- Direct-to-consumer (DTC) testing
 advantages, 532
 genetic test (GT), 530
 personalized nutrigenetic testing, 529
 regulatory environment
 Europe, 530–531
 North America, 531
 nutrigenetic tests, 531–532
 United States, 530
- Discrete/discontinuous traits, 4
- Discrete traits, 4
- DNAase I hypersensitivity, 7
- DNA damage prevention, 28–31
- DNA hydroxymethylation, 133
- DNA methylation, 56, 193–194, 424
 data, 38
 epigenetic modifications, 427
 epigenetics, 59–60, 130–133, 132f
 gene-silencing effects, 132
 global quantification, 61–63
 heritability, 132f
 locus-specific quantification, 63–64, 63f
 mass spectrometry methylation assays (MassARRAY), 64
 methylated DNA affinity enrichment assay, 64
 methylation-sensitive high-resolution melt analysis (MS-HRM), 63–64
 methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA), 64
 methylation-specific polymerase chain reaction, 63
 pyrosequencing, 64
- DNA methylome, protein restriction, 116–117, 117f
- DNA methyl-transferase (DNMT) enzymes, 132–133
- DNA repair, 28–31
- DNA replication, 28–31
- DNA sequencing, 38, 53
- Docosahexaenoic acid (DHA), 78
- Dominant, 4
- Dopamine D2 receptors (DRD2), 163
- Dopamine receptor D1 (D1R), 163
- DRB sensitivity-inducing factor (DSIF), 20
- Drug biotransformation, 519–520
- Dyslipidemia
 genes expression changes
 dietary lipids uptake, 447–448, 449f
 lipoprotein metabolism, 451–452
 lipoproteins synthesis, 448–451, 450f
 tissue clearance, of lipoproteins, 451–452
 primary, 447
 role, 453–455, 453t–455t
 secondary, 447
- E**
- Eating behavior, 159
 appetite traits, 162–163
 meal size and frequency, 161–162
 new genomic approaches, 164
 taste, 160–161
- Electrolytes
 chloride, 315
 definition, 309
 dietary recommendations, 309, 309t
 potassium, 314–315
 sodium
 mineralocorticoids, 311–314, 312f
 mineralocorticoids receptor cascade, 311–314, 312f
 monogenic disorders, 310
 renin-angiotensin system (RAS), 311, 312f
 salt sensitivity phenotype, 310–311, 310f
- Electrospray ionization (ESI), 55
- Embryonic development, 423–424
- Embryonic transcriptome, 115–116
- ENCODE Project, 7–8
- Endothelial dysfunction, 291
- Energy expenditure, 178–179
- Energy homeostasis, 175
 genetic influence, 175–176
- Epigallocatechin gallate (EGCG), 329
- Epigenetic markers, 193–195
- Epigenetic modifications, 125, 125f
- Epigenetic regulation, 126
- Epigenetics
 chromatin remodeling, 136–137
 components, 59–60, 60f
 definition, 59–60, 129
 in development, 424–425
 DNA hydroxymethylation, 133
 DNA methylation, 59–60, 130–133, 132f
 global quantification, 61–63
 locus-specific quantification, 63–64, 63f
 environmental influences, 137f
 food allergies
 allergic diseases, 404–406
 early life, 406–407
 GXE interaction, 404
 genome-wide DNA methylation analysis, 65–67
 microarray-based methylation analyses, 65
 next-generation sequencing, 65–67, 66t
 in health and disease, 137–138
 histone modifications, 133–136, 134f
 marks and DNA packaging, 131f
 next-generation sequencing (NGS), 60
 nucleotide methylation analysis, 60–61
 nutrients, 59
 pesticides, 516
 philosophical notions, 130
 regulation, 130
 RNA interference, 137
 Waddington's landscape, 129, 130f
- Epigenomics, 50, 56–57
- Epistasis, 6
- Epistemology
 collecting data, 108–109
 initial building blocks of knowledge, 107–108
 turning data into knowledge, 109
- Epithelial sodium channel (EnaC), 161
- Ethics
 integrity risks
 genetic test results, 544
 incidental findings, 544
 sensitive data and protection, 543–544
 mitigation
 informed consent, 547
 legislation, 547–548
 precautionary principle, 547
 professionals education, 547
 step-by-step approach, 548
 wider understanding of personalization, 547
 quality risks, 546–547
 commercialization, 546–547
 evidence, 546
 motivation, 546
 social risks
 individual responsibility for health, 545
 justice and the risk for discrimination, 545–546
 social aspects of eating, 544–545
- Eukaryotic DNA, 18
- Expression QTLs (eQTLs), 8–9
- Extranuclear genome, 4
- Extrinsic pathway, 361–362, 362f
- F**
- Fantom5, 7
- Fat intake-genotype visfatin interaction, 381
- Fat mass and obesity-associated (FTO), 159–160, 162, 168–169, 178, 182–185
- Fatty acids
 bone metabolism, 494
 in brain, 348–349
 desaturases, 356–357
 genetics, 349
- Fecal microbiota transplants (FMTs), 85

- Fetal development, 423–424
 Fetal transcriptome, 115–116
 Flavin adenine dinucleotide synthase, 285
 Flavonols, 477
 Folate, 28, 28f
 coenzymes, 295
 enzymes, 299–301
 genetic variants, 297–301
 Food allergies, 403
 development, 404
 epigenetics
 allergic diseases, 404–406
 early life, 406–407
 GXE interaction, 404
 and food intolerance, 403–404
 GXE interactions, 404
 Foodomics, 89–90, 91f
 Food pesticides, 513
 biomarkers, 514–515
 bioremediation, 517–518
 early determinants, 516–517
 individuals' responses, 515–516
 modulation, 516
 nonnutritive chemicals, 513–514
 Fragments per kilobase per million reads (FPKM), 38
 Frameshift, 14
 Fructose, 206
 Fructose-2,6-bisphosphate, 23
 FTO. *See* Fat mass and obesity-associated (FTO)
- G**
 Gametes, 5
 Gas6, 269–270
 Gene, 3–4
 Gene-activating switch/sucrose nonfermentable (SWI/SNF) complex, 136
 Gene expression, 18, 117f
 changes in, 18
 chromatin structure, 18–19
 dietary regulation, 21–22, 22f
 DNA interactions, 18–19
 genome-wide monitoring of, 54–55
 microarray analysis of, 116
 microarrays, 36, 55
 nutrients, 551
 nutrients and, 427–428
 pesticides, 516
 phytanic acid (PA), 343–344
 protein-coding genes, 18–20
 signals impact, 21, 21f
 transcription factor, 18–19
 Gene–gene interactions, 54
 Gene–nutrient interactions, 372–373
 Gene–protein interactions, 54
 General transcription factors (GTFs), 18
 Genetic disorders, 5
 mitochondrial flavin adenine dinucleotide transport deficiency, 285
 transporter deficiency, 284–285
 Genetic information, 525
 Genetic risk score (GRS), 169–170, 198–199, 219–220
 Genetics, 33
 Genetics of Lipid Lowering Drugs and Diet Network (GOLDN), 219–220
 Genetic variation
 body weight and obesity, 198
 complex patterns, 14–16
 copper, 323–324
 copy number variants, 15–16
 deletions, insertions and repeats, 14
 and dietary response, 549–550
 in drug biotransformation, 519–520
 EPA, DHA, and AA levels, 223–224
 ethnic differences, 550–551
 frequency and relevance, 12–13
 FTO gene, 182–185
 functional significance, 8–9
 genotype, diet interventions and weight loss, 199–200
 habitual dietary intakes, 198–199
 haplotypes, 14–15, 15f
 hemizygosity, 15
 obesity, 204–205. *See also* Obesity
 Single nucleotide polymorphism (SNP).
 See Single nucleotide polymorphism (SNP)
 sweetness perception, 205
 transporters, 205–206
 weight loss diets, 182
 zinc, 322–323
 Genome, 3–4
 1000 Genomes Project, 7–8
 Genome-wide association studies (GWAS), 8, 34–35, 178, 188, 255–256
 genomic epidemiology, 100–102
 Genome-wide DNA methylation analysis, 65–67
 microarray-based methylation analyses, 65
 next-generation sequencing, 65–67, 66t
 Genome-wide scans, 56
 Genome-wide screening studies, 191–193
 Genomic epidemiology
 association study, 100
 different study designs, 105f
 family-based association studies, 103
 gene–environment interactions, 103–104
 genome-wide association studies (GWAS), 100–102
 homozygosity mapping approach, 105
 linkage studies, 105–106
 mediation studies, 104
 Mendelian randomization studies, 104–105
 multifactorial traits and diseases, 99–100, 101f, 101t
 next-generation sequencing (NGS) methods, 105–106
 observational studies, 104f
 pathway-based association studies, 102
 segregation studies, 105
 Genomic methods, 49–50, 53–54
 copy number variations, 54
 DNA sequence, 53
 gene–gene interactions, 54
 gene–protein interactions, 54
 haplotype analysis, 53–54
 microarrays, 53
 nutrition assessment, 54
 Genomics
 approaches, 35
 definition, 33
 Genomic variation
 feature, 8
 types, 8
 Genotyping, 4, 13
 single nucleotide polymorphisms (SNP)-array, 35–36
 GGXX enzyme, 266
 Ghrelin, 161
 GHRL, 176
 GHSR, 176
 Gla-rich protein (GRP), 268
 Glucose, 206
 metabolism, 268–269, 290
 Glucose-6-phosphate, 23
 Glucose response conserved element (GRACE), 386
 Glutamate decarboxylase 2 (GAD2), 159–160
 Glutathione S-transferases (GSTs), 515–516
 Glutathione transferases, 280–281
 Gout, 389
 clinical significance, 392
 GPIHBP1, 451–452
 Green tea catechins (GTCs), 329, 331
 Group-specific component (GC), 251
 Guanidinium thiocyanate (GTC), 43
 Gut microbiome, 81
 Gut microbiota, 81
 and celiac disease, 418
 composition, 438
 design and interpretation, 411–416
 dietary effects, 412–414, 414f
 diurnal cycling, 414–415
 Helicobacter pylori vs. host genetic polymorphism, 418–419
 host genetic polymorphism, 417–418
 host genetics influence, 416–417
 inflammasome, 415
 L-carnitine, 416
 microbiome-based personalized nutrition, 415–416
 studies, 411
- H**
 Haplotype analysis, 53–54
 Haplotypes, 14–15, 15f, 526–527
 HapMap Project, 7, 100–101
 Haptoglobin, 281
 Health Professional Follow-Up Study (HPFS), 198–199
 Hemi-insufficiency, 15
 Hemizygosity, 15
 Hemochromatosis
 genetic risk factors, 319
 health consequences, 318–319
 interventions, 319

- pathophysiology, 318–319
- Hemostasis
- coagulation FVII, 363
 - coagulation process, 362–363
 - diet
 - extra virgin olive oil and monounsaturated fatty acids, 363
 - n-3 polyunsaturated fatty acids, 364–365
 - polyphenols and platelet aggregation, 365–366
 - postprandial factor VII, 364
 - triglycerides, 363–364
 - factor VII, 363
 - human pathology, 362–363
 - nutrigenetics, nutrigenomics and coagulation system
 - gene–environment interaction, 366–367
 - mediterranean diet, 367
 - polymorphisms of factor VII gene, 366
 - primary, 361
 - secondary hemostasis pathways, 361–363
- Hemotrophic nutrition, 425
- Hepatic lipase (HL), 452
- Heritability, 5–6
- HERITAGE Family Study, 191
- Heterozygous, 4, 13
- High-density lipoprotein (HDL), 290
- High-fat diet (HFD), 415, 429
- Histiocytic nutrition, 425
- Histone acetylation, 424–425
- Histone acetyl transferase (HAT) enzymes, 134, 136f
- Histone deacetylase (HDAC) enzymes, 134, 136f
- Histone methylation, 134–136, 424–425
- Histone modifications, 7, 133–136, 134f
- Histone posttranslational modifications, 56, 133–134
- Homologous, 4
- Homozygosity mapping approach, 105
- Homozygous, 4, 13
- Human Genome Project, 7–8
- Human genome structure, 11–12
- Human genomic variation, 7–9
- Humanized gnotobiotic mice, 85
- Human Microbiome Project, 85
- Hyperkalemia, 314–315
- Hyperphagia, 163, 168–169
- Hypertension, 520–521
- Hyperuricemia
 - asymptomatic treatment, 392–393
 - clinical significance, 391–392
 - endogenous purine metabolism, 390
 - genetics and environmental factors affecting, 393
 - intestinal urate excretion, 391
 - major urate transporters, 393–394
 - nutrients affecting, 390f
 - renal urate excretion, 390
- Hypocaloric diet (HP diet), 185
- Hypokalemia, 314
- I**
- IEM. *See* Inborn errors of metabolism (IEM)
- Illumina technology, 82
- Inborn errors of metabolism (IEM)
 - with dietary interventions, 156t–157t
 - medical foods, 154
 - newborn screening programs, 154
 - nutritional interventions, 154
 - principle, 153–154
 - screening, 153–154
 - significant number, 154
- Inflammation
 - D-series, 462
 - E-series, 462
 - genes and microRNAs, 465–466
 - ideal outcome, 459f
 - maresins, 463
 - niacin, 291–292
 - protectins, 462
 - regeneration, 463–465
 - resolution, 458–459
 - specialized pro-resolving mediator receptors, 463
 - structural elucidation, 459–462
- Insulin growth factor 2 (IGF-2), 124–125, 428
- Insulin receptor substrate 1 (IRS1) gene, 505–506
- Insulin signaling, 290
- Interindividual genetic variation, 8
- Intermediate dominance, 4
- International Human Genome Sequencing Consortium, 129
- Intracisternal A-particle (IAP), 427
- Intrauterine programming, 427
- Intrinsic pathway, 361, 362f
- Iodine, 324–325
- Ion S5, 82
- Ion Torrent Personal Genome Machine, 82
- Iron
 - absorption, 317, 318f
 - accumulation, 318–319
 - deficiency risk, 317–318
 - homeostasis, 317
- Iron-sulfur (Fe–S) cluster proteins, 28
- Ischemic heart attacks, 522–523
- Isoenergetic oligonucleotide microarray mapping, 47
- Isoflavones, 493–494
- Isopycnic gradient method, 43
- L**
- Lab-on-chip technologies, 44
- Lactase persistence, 525–526
- Lambda, 99–100
- Latent Early life Associated Regulation (LEARn) model, 517
- Late-onset diseases, 516–517
- LCAT, 452
- LDLR adaptor protein 1 (LDLRAP1), 453
- Lecithin retinol acyltransferase (LRAT), 240–241
- Lectin affinity chromatography, 74
- Leptin, 161, 170–171
- Leptin receptor gene (LEPR), 506
- Ligand-modulated TFs, 21
- Linkage, 4
- Linkage disequilibrium (LD), 35, 100–101
- Lipase maturation factor 1 (LMF1), 451–452
- Lipid mediators (LM), 458
- Lipid metabolism, 290–291
- Lipids, 433–434
- Lipodystrophy, 172t, 173
- Lipoproteins, 447
 - mechanism, 448f
 - metabolism, 451–452
 - synthesis, 448–451, 450f
 - tissue clearance of, 451–452
- 5-Lipoxygenase (5-LO), 357, 357f–358f
- Liver X receptor (LXR), 23
 - hepatic glucose signaling, 386–387
 - hepatic lipogenesis, 385–386
 - in lipogenesis, 383, 385f
 - nuclear receptor, 383–384
- Liver X receptors (LXRs), 442
- Locus, 4
- Logistic regression, 76–77
- Long interspersed nuclear elements (LINEs), 7–8
- Long noncoding RNAs (lncRNAs), 7, 123
- Long-term potentiation (LTP), 348
- Low density lipoprotein (LDL), 356
- Low-fat diet intervention (LFD), 93
- Low-protein (LP)
 - diet in pregnancy, 113
 - fetal development, 117
 - programming effects, 114
- LRP2, 250
- Luteolin, 332–333, 333f
- Lysine 16 of histone 4 (H4K16), 134
- Lysine 4 on histone 3 (H3K4me), 134–136
- Lysine 9 or 27 on histone 3 (H3K9 and H3K27), 134
- M**
- MALDI-TOF, 56
- Mass spectrometry (MS), 55, 93–94
- Mass spectrometry methylation assays (MassARRAY), 64
- Maternal Agouti mouse diet, 138
- Maternal perinatal nutrition
 - applications, 125–126
 - current challenges, 125–126
 - macronutrient and micronutrient exposure, 124–125
 - obesity, 121–122
 - offspring development, 122–123
 - during preconception period, 121–122
 - during pregnancy, 121–122, 122f
 - programming mechanisms, 123–124
- Matrix-assisted laser desorption/ionization (MALDI), 55, 71
- MC1R, 248–249
- Mechanistic target of rapamycin (mTOR) pathway, 24
- Mediterranean diet (MD), 93, 94f
- Melanocortin, 171

- Melanocortin 4 receptor (*MC4R*), 168–169, 171, 178
- Mendelian inheritance, 5
of quantitative traits, 5–6, 6f
types, 5, 5f
- Mendelian traits, 5
- Mendelian transmission, 5–6
- Menkes disease protein (*ATP7A*), 323–324
- Mesolimbic pathway, 163
- Messenger RNA (mRNA), 11, 36, 41
- Metabolic Diet App Suite, 154
- Metabolic stress response signaling, 288
- Metabolomics, 50, 55–56, 86, 90–92, 92f
dysfunctions gene variants, 77–78
invasive fluids, 75
mechanistic insights, 76
noninvasive biological fluids, 75
obesity, 77
organ tissue, 75
placental abruption (PA), 76–77
reveal biomarkers, 76
study designs, 75–76
targeted and untargeted analysis, 76
- Metagenomics, 38–39, 50
approaches, 82–83
current gut microbiota status, 85
DNA sequences, 85–86
history of methods used, 81–82
microorganisms, 85
Pacific Biosciences (PacBio), 82
principle, 82
single-molecule real-time (SMRT), 82
targeted methods, 82–83
untargeted methods, 83
workflow, 83–85, 84f
- Metaproteomics, 86
- Metatranscriptomics, 86
- Methionine synthase, 298–299
- Methionine synthase reductase, 298–299
- Methylated DNA affinity enrichment assay, 64
- Methylated DNA immunoprecipitation (MeDIP), 64
- Methylation cycle, 60, 61f
- Methylation QTLs (meQTLs), 9
- Methylation-sensitive high-resolution melt analysis (MS-HRM), 61–64
- Methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA), 64
- Methyl-binding domain (MBD), 132
- Methyl-CpG binding protein 2 (MeCP2), 132
- 5-Methylcytosine, 28
- Methyl DNA-binding proteins (MBD), 64
- Methyl donors, 28
- Methylenetetrahydrofolate reductase (MTHFR), 15, 15f, 108–109, 299–300, 329, 357–359, 358f, 510
- Methylnicotinamide (MNA), 292
- 5-Methyltetrahydrofolate, 28
- Metric traits. *See* Quantitative traits
- Microarrays, 47, 53
- Microbiome, 411
- Microbiota, 411
- Micronucleus (MN) frequency, 30–31
- Micronutrients, 28–31, 29t
deficiency, 28–30
- Micro RNA (miRNAs), 7, 56–57, 123, 137, 424–425
modulators, 24
- Microsatellites, 14
- Microsomal triglyceride transfer protein (MTTP), 449–450
- Minerals, 28, 437
- Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines, 46
- Minor allele frequency (MAF), 101–102
- Mitigation
informed consent, 547
legislation, 547–548
precautionary principle, 547
professionals education, 547
step-by-step approach, 548
wider understanding of
personalization, 547
- Mitochondrial DNA, 11
- Mitochondrial flavin adenine dinucleotide transport deficiency, 285
- Mitochondrial genome, in mammals, 4
- Mobile elements, 7–8
- Modern postgenomic datasets, 57
- Modern spectrophotometric methods, 44
- Monoallelic expression, 13
- Monogenic obesity, 167–168, 170–172, 431
- Monogenic traits, 4
- Monounsaturated fatty acids (MUFAs), 212
- Monozygotic (MZ) *vs.* dizygotic (DZ) twins, 100
- MTHFR. *See* Methylenetetrahydrofolate reductase (MTHFR)
- MTHFR C677T, 235
- Multimetabolite biomarker panel (MBP), 92–93
- N**
- NADSYN1, 248–249
- NanoDrop, 44
- National Institute on Aging (NIA), 400
- Nature, 6
- ND-3300, 44
- Negative elongation factor (NELF), 20
- Network biology analysis, 94–95
- Neural tube defects (NTD), 427–428
- Newborns, 154
- Newborn screening programs, 154
- Newborn Screening Translational Research Network (NBSTRN), 154
- Next-generation analytical tools, 57
- Next-generation nutrigenetics, 307
- Next-generation sequencing (NGS), 38–39, 47, 81–82
alignment and mapping, 37f
analyses, 35, 35f
epigenetics, 60
genome-wide DNA methylation analysis, 65–67, 66t
genomic epidemiology, 105–106
- Niacin, 287, 288f
endothelial dysfunction, 291
inflammation, 291–292
ischemia, 291
metabolism regulation and signaling, 287–290
metabolic stress response signaling, 288
oxidative stress response, 289–290
poly(adenosine diphosphate-ribose) polymerases, 288–289
sirtuins, 289
neuroprotective effects, 292
physiology
adipogenesis and lipid metabolism, 290–291
glucose metabolism and insulin signaling, 290
pleiotropic effect, 292
vascular homeostasis, 291
- Niemann-Pick C1 (NPC1), 453
- Niemann-Pick C1-like 1 (*NPC1L1*), 229
- N*-methyl-*D*-aspartate (NMDA), 348
- Nonalcoholic fatty liver disease (NAFLD), 387
- Noncoding variants, 8
- Nonprotein-coding genes, 7
- Nuclear genome, 4
- Nuclear magnetic resonance spectroscopy, 55–56
- Nuclear receptor (NR), 21–23, 383–384
- Nucleic acid modification, 62t
- Nucleosomal histones, 20
- Nurture, 6
- Nutrigenetics, 33, 554
- Nutrigenomics, 33
bioinformatics, 50, 57
epigenomics, 50, 56–57
genomic methods, 49–50, 53–54
copy number variations, 54
DNA sequence, 53
gene–gene interactions, 54
gene–protein interactions, 54
haplotype analysis, 53–54
microarrays, 53
nutrition assessment, 54
metabolomics, 50, 55–56
metagenomics, 50
methods, 51t–53t
proteomics, 50, 55
statistical methods, 50
study design, 57
transcriptomics, 50, 54–55
- Nutrimetabolomics, 90–92
- Nutriome, 31
- Nutrition
confounders, 347–348
and development, 425–429, 426f
DNA damage prevention, 28–31
DNA repair, 28–31
DNA replication, 28–31
- Nutritional genomics, 327
- Nutritional phenotype, 203
- Nutrition guidance

- computer-based meal planning, 511
individually tailored, 510–511
nutrition counseling practice,
509–510
- O**
- Obesity, 167, 175, 181, 269–270, 328
adipose tissue, 432–433
carbohydrates, 204–205
energy expenditure, 178–179
energy homeostasis, 175
genetic influence, 175–176
energy intake, 176–178
etiology of, 197
genetic polymorphisms, 328–329
genetic predisposition, 198–199
maternal perinatal nutrition, 121–122
nutrient regulation, 433–438, 434f
bioactive compounds, 437–438
carbohydrates, 434–436
gut microbiota composition, 438
lipids, 433–434
minerals, 437
probiotics, 438
proteins, 436
vitamins, 436–437
nutrition and genomics, 329
pathogenesis, 431
polyphenols, 329
catechins, 329–331
curcumin, 331–332, 333f
luteolin, 332–333, 333f
resveratrol, 331, 333f
precision nutrition, 438–439, 505–506
prevent, 150t, 151
susceptibility genes, 162–163
- Odorant receptors (ORs), 15
- Omega-3 fatty acids (n-3 FAs), 223
genetic influences, 224–225
- Omega-3 long chain polyunsaturated fatty acid supplementation (*n*-3 LCPUFA), 78
- Omics, 49, 50f, 500
food intake measurement, 92–94, 93f
integration, 91f
interindividual variability, 90–92
novel nutritional biomarkers development, 90–92
nutritional systems biology, 94–95, 96f
nutrition research, 89–90
postgenomic era, 89
technologies and molecules of interest, 90t
tools, 89–90
- One-carbon groups (OCG), 295–296
- One-carbon metabolism (OCM), 295–297, 296f
- Operational taxonomic units (OTUs), 82–83
- Optical density (OD) measurement, 44
- Osteoblasts, 490–492
- Osteoclasts, 490
- Osteocytes, 490
- Oxidative bisulfite sequencing (oxBS-seq), 66
- Oxidative stress response, 289–290
- Oxylipins, 223
- Oysters, 322
- P**
- PA. *See* Phytanic acid (PA)
- Pacific Biosciences (PacBio), 82
- Pancreatic lipase (PNLIP), 447–448
- Pantothenic acid, 276–277
coenzyme A synthesis and distribution, 276–277
deficiency and supplementation, 277
genetic disorders, 276
sources and requirements, 277
- Paraoxonase (PON), 515
- Parkinson disease (PD), 292
- PCSK9, 102
- Peptide mass fingerprinting (PMF), 71
- Peptide tyrosin-tyrosin (PYY), 162
- Peripheral blood cells, omega-3 polyunsaturated fatty acids, 472–476
- Peroxisome proliferator-activated receptor- α (PPAR- α), 343–344
- Peroxisome proliferator activated receptor γ (PPAR- γ), 159–160, 328
- Peroxisome proliferator-activated receptor gamma coactivator 1 α (PPARGC1A), 189
- Peroxisome-proliferator activated receptors (PPARs), 22, 359, 359f, 442, 504
- Personal Genome Service (PGS), 530
- PGC-1 α , 23
- Phenol:chloroform method, 43
- Phenolic alcohols, 479–480
- Phenotype, 4
- Phenotypically discordant twins, 100
- Phenotypic expression, 6–7
- Phenotypic traits, 3–7
- Phosphatidylethanolamine-*N*-methyltransferase (PEMT), 77
- Photoreceptor outer segments (POS), 243–244
- Physical fitness and performance
exercise training
genetic component, 187–188
individual differences, 187
trainability, 188–195
candidate gene studies, 188–191
epigenetic markers, 193–195
genome-wide screening studies, 191–193
- Phytanic acid (PA)
and cancer, 344
chemical preparation, 341
gene expression, 343–344
human plasma, 341–342
inborn errors, 342–343
metabolism, 342
origin, 341–342
pathophysiological mechanisms, 345
physiological effects, 344
- Placental abruption (PA), 76–77
- Placental transcriptome, 114–115
- Plant sterols, 227–229
- PNPLA3, 234–235
- Pol II core promoters, 19
- Polycomb group (PcG) proteins, 136–137
- Polygenic inheritance, 5–6
- Polymerase chain reaction (PCR), 53
methylation-sensitive high-resolution melt analysis (MS-HRM), 63–64
methylation-specific, 63
- Polymorphism, 13
- Polyphenols, 327–328
catechins, 329–331, 333f
classification, 476–480, 476f
curcumin, 331–332, 333f
flavonols, 477
luteolin, 332–333, 333f
phenolic alcohols and secoiridoids, 479–480
resveratrol, 331, 333f
stilbenes, 477–479
- Poly(adenosine diphosphate-ribose) polymerases, 288–289
- Polyunsaturated fatty acids (PUFA), 22–23, 349, 517
- POMC, 171
- Poor nutrition, 211
- Posttranscriptional regulatory mechanisms, 18
- Posttranslational modification (PTM) analysis
massive screening, 73
protein glycosylation, 74
protein lysine acetylation, 73
protein phosphorylation, 73
- Potassium, 314–315
- Prader-Willi syndrome (PWS), 171–172
- Precision nutrition, 49, 499–500
genotype, 501–506
obesity, 438–439
obesity and cardiometabolic traits, 505–506
specific nutrients, 501–505, 502t–503t
- Precision nutrition approach, 89
- Pregnancy, maternal perinatal nutrition during, 121–122, 122f
- Preimplantation development, 423–424
- Preinitiation complex (PIC), 18–20
- Preventing Overweight Using Novel Dietary Strategies (POUNDS Lost), 199
- Primary dyslipidemias, 447
- Privacy risk and data protection, 539
cloud computing, 540
data availability, 540
data immortalization, 539
marketing, 539–540
- Probiotics, 438
- Promoter-proximal pausing (PPP), 20
- Proopiomelanocortin (POMC) methylation, 123–124
- Proprotein convertase subtilisin/kexin type 9 (PCSK9), 453
- Prostaglandin-endoperoxide synthase-2 (PTGS-2), 77
- Proteasome, 19

- Protein-coding genes, 7
 Protein-coding RNAs, 41
 Protein glycosylation, 74
 Protein kinase positive transcription elongation factor b (P-TEFb), 20
 Protein lysine acetylation, 73
 Protein phosphorylation, 73
 Protein restriction
 DNA methylome, 116–117, 117f
 early life programming of disease, 113
 early life programming, of gatekeeper approach, 116f
 embryonic and fetal transcriptome, 115–116
 placental transcriptome, 114–115, 115f
 in pregnancy, 117–118
 programming mechanisms, 113–114
 Proteins, 436
 Proteomics, 50, 55, 69–70
 bottom-up methods
 gel-based strategies, 71
 shotgun analysis, 71–72
 definition, 69–70
 genome-wide sequencing projects, 69
 huge dimension, 70–71
 posttranslational modification (PTM)
 analysis
 massive screening, 73
 protein glycosylation, 74
 protein lysine acetylation, 73
 protein phosphorylation, 73
 steps, 70f
 successful experiment, 70–71
 targeted, 72–73
 selected reaction monitoring (SRM), 72–73
 SWATH mass spectrometry, 73
 top-down, 72
 typical experiment, 70–71
 Proton pump inhibitor (PPI), 416
 Psychotropic drugs, 523–524
 Pyridoxal phosphate (PLP), 297
 Pyridoxine. *See* Vitamins B₆
 Pyrosequencing, 64
- Q**
 Quantitative Insights Into Microbial Ecology (QIIME), 84
 Quantitative trait locus (QTL), 5–6, 8–9
 Quantitative traits, 4, 6, 35
 Quercetin, 477
- R**
 Reads per kilobase per million reads (RPKM), 38
 Real-time reverse transcription polymerase chain reaction (qRT-PCR), 45–46
 Recessive, 4
 Recommended Dietary Allowance (RDA), 509–510
 Red blood cell (RBC), 223–224
 Refsum disease (RD), 342–343
 Regulatory elements, 7
 Regulome, 7
 Renin-angiotensin system (RAS), 311, 312f
 Repeats, 7–8
 Resveratrol, 331, 333f
 Retinal pigment epithelium (RPE), 243–244
 Retinal pigment epithelium-specific protein 65 kDa (RPE65), 243–244
 Retinoic acid (RA), 241–243
 Retinoid-binding proteins, 239–240
 Retinoids
 absorption, 239–241
 conversion to, 238
 Retinoid X receptor (RXR), 343–344, 442
 Retinol dehydrogenases (RDHs), 242
 Reverse transcription dPCR (RT-dPCR), 46
 Reverse transcription polymerase chain reaction (RT-PCR), 45–46
 Reversible chromatin remodeling, 20
 Reversible protein phosphorylation, 73
 Rhizomelic chondrodysplasia punctata type 1 (RCDP type 1), 343
 Riboflavin, 283
 dietary deficiency, 283
 function, 283–284
 genetic disorders, 284–285
 flavin adenine dinucleotide synthase, 285
 mitochondrial flavin adenine dinucleotide transport deficiency, 285
 transporter deficiency, 284–285
 uptake, 283, 284f
 Ribonucleic acid (RNA), 41, 161
 analyses, 48f
 classes, 41, 42t
 digital polymerase chain reaction (dPCR), 46
 extraction methods
 adsorption methods, 43
 cell-free mRNA, 43–44
 isopycnic gradient method, 43
 phenol:chloroform method, 43
 small RNA molecules, 44
 special methods, 43
 isolation, 41–44
 isolation methods, 45f
 microarrays, 47
 prokaryotic and eukaryotic organisms, 41
 quantity and quality assessment, 44–45
 real-time reverse transcription polymerase chain reaction (qRT-PCR), 45–46
 reverse transcription, 45
 sequencing, 47
 Ribosomal Database (RBD) Project, 82–83
 Ribosomal RNA (rRNA), 36
 RM-493, 171
 RNA interference, 137
 RNA-Sequencing (RNA-Seq), 38, 47
 Roadmap Epigenomics Project, 7
 Runt-related transcription factor 2 (RUNX2), 490–491
 RXRA, 253
- S**
 S-adenosylhomocysteine (SAH), 133
 S-adenosylmethionine (SAM), 133
 Salt sensitivity, 310–311, 310f, 520–521
 Sanger sequencing, 33–34
 Scaffold, 83
 Scavenger receptor class B member 1 (SR-BI), 452
 Secoiridoids, 479–480
 Secondary dyslipidemias, 447
 Selected reaction monitoring (SRM), 72–73
 Selenium, 324
 Semidehydroascorbate (SDA), 279
 Seminal studies, 138
 Serine hydroxymethyltransferase (SHMT), 301
 Serum uric acid (SUA), 389
 Sex determining region Y-box 9 (SOX9), 490–491
 Sex-influenced inheritance, 7
 Sex-limited inheritance, 7
 Sexual chromosomes, 4
 Short-chain fatty acids (SCFAs), 412
 Short interspersed nuclear elements (SINEs), 7–8
 Short tandem repeats (STRs), 14
 Simple sequence repeats (SSRs), 14
 Single-molecule real-time (SMRT), 82
 Single nucleotide polymorphism (SNP), 13–14, 162
 Single Nucleotide Polymorphism Database (dbSNP), 13
 Single nucleotide polymorphisms (SNP), 8–9, 34–35, 102, 176, 179, 182, 305–306, 501, 549–550
 weight loss diets, 182
 Single nucleotide variants (SNVs), 168–169
 Single-variant association studies, 526
 Sirtuin 1 (SIRT1), 23, 486–487
 Sirtuins (SIRTs), 289
 Sitosterolemia, 447–448
 SLC23A1, 280
 SLC23A2, 280
 SLC40A1, 319
 Small interfering RNAs (siRNAs), 137
 Small intestinal bacterial overgrowth (SIBO), 268
 Societal issues, 553–554
 Sodium
 mineralocorticoids, 311–314, 312f
 mineralocorticoids receptor cascade, 311–314, 312f
 monogenic disorders, 310
 renin-angiotensin system (RAS), 311, 312f
 salt sensitivity phenotype, 310–311, 310f
 Sodium-dependent multivitamin transporter (SMVT), 276–277
 Specialized pro-resolving mediators (SPMs), 458
 Sphingosine 1-phosphate (S1P), 78
 SREBP1C, 385–386

- Statistical methods, 50
- Sterol regulatory element-binding protein-1 (SREBP-1), 23
- Sterol-responsive element binding proteins (SREBPs), 442–443
- Stilbenes, 477–479
- Structural variants (SV), 8
- Sucrase-isomaltase, 206–207
- Sugar-sweetened beverages (SSB), 198–199
- SWATH mass spectrometry, 73
- Sweetness perception, 205
- Syndromic obesity, 170–172
- T**
- Targeted metabolomics, 76
- TAS2R38, 160–161
- TCF7L2, 101–102
- T2D. *See* Type 2 diabetes (T2D)
- Ten-eleven transferase (TET) enzyme, 133, 133f
- Thiamine
- deficiency and supplementation, 276
 - functions, 275
 - genetic disorders, 275
 - nutrigenomics, 276
 - sources and requirements, 276
- Threonine 666 (Thr66), 386–387
- Thrombosis
- coagulation FVII, 363
 - coagulation process, 362–363
 - diet
 - extra virgin olive oil and monounsaturated fatty acids, 363
 - n-3 polyunsaturated fatty acids, 364–365
 - polyphenols and platelet aggregation, 365–366
 - postprandial factor VII, 364
 - triglycerides, 363–364
 - factor VII, 363
 - human pathology, 362–363
 - nutrigenetics, nutrigenomics and coagulation system
 - gene–environment interaction, 366–367
 - mediterranean diet, 367
 - polymorphisms of factor VII gene, 366
 - primary, 361
 - secondary hemostasis pathways, 361–363
- Thyroid peroxidase, 325
- Tissue factor (TF)-activated extrinsic pathway, 361–362, 362f
- Tissue remodeling, 114, 114f
- Trace elements
- copper
 - distribution, 323
 - excessive accumulation, 323
 - genetic variants, 323–324
 - sources, 323
 - iodine, 324–325
 - selenium, 324
 - zinc
 - enzymes, 321–322
 - genetic variation, 322–323
 - insufficient, 322
 - intestinal absorption and transfer, 322
 - meats and organ meats, 322
- Trainability, 188–195
- candidate gene studies, 188–191
 - epigenetic markers, 193–195
 - genome-wide screening studies, 191–193
- Traits, 4
- types, 4
- Transcobalamin II, 298
- Transcription
- chromatin remodeling and remodelers, 20, 21f
 - of protein-coding genes, 19–20
 - signals and mechanisms, 21
 - subsequent steps, 20
- Transcription factor 7-like 2 (TCF7L2), 212–219
- Transcription factors (TFs), 18–19, 21
- dietary regulation, 22–24
 - abundance, 23
 - activity, 22–23
 - metabolic fluxes, 23
 - nutrient/energy sensing signaling pathways, 23–24
- Transcription start site (TSS), 18–19
- Transcriptomics, 50, 54–55
- gene expression microarrays, 36
 - generation of libraries, 36
 - integration of results, 38
 - number of replicates, 36
 - quantification, 38
 - RNA-seq analysis, 36
 - RNA sequencing data analysis, 37
 - sequencing depth, 37
- Transcripts, 12
- Transcripts per million (TPM), 38
- Transgenerational epigenetic inheritance, 404
- Transmission/disequilibrium test (TDT), 103
- Transporter deficiency, 284–285
- biochemical abnormalities, 284
 - clinical symptoms, 284
 - diagnosis, 284
 - pathophysiology, 284
 - treatment and prognosis, 284–285
- Triallelic variant, 12–13
- Trimethylamine *N*-oxide (TMAO), 414
- Tübingen Lifestyle Intervention Program (TULIP), 200
- Tumor necrosis factor α (TNF- α) gene, 506
- Twin and adoption studies, 100
- Two-dimensional electrophoresis (2DE), 69–71
- Type 2 bitter taste receptor genes (TAS2R), 160
- Type 2 diabetes (T2D), 163
- adipose tissue, 373–375
 - calorie restriction (CR), 444
 - diet-induced molecular processes, 373
 - gene–nutrient interactions, 372–373
 - genes *vs.* environment, 371–372
 - insulin resistance, 444
 - macronutrients, 442–443
 - micronutrients, 443–444
 - nutrigenomics, 441–442
 - progressive phenotype, 371, 372f
- Type I and II HDACs inhibitors, 24
- U**
- Umami taste gene (TAS1R3), 161
- Uncoupling protein 1 (UCP-1), 344
- Uncoupling protein (UCP), 179
- Untargeted metabolomics, 76
- Upper phase, 43
- Uric acid, 389. *See also* Hyperuricemia
- concentration, 389
 - metabolism, 391f
- Uridine diphosphate N-acetylglucosamine (UDP-GlcNAc), 23
- V**
- Vascular disease, 267
- Vascular homeostasis, 291
- Very-low-density lipoprotein (VLDL), 257
- Viruses, 85
- Visual chromophore metabolism, 243–244
- Vitamin A, 237–238, 240f
- from holo-RBP4, 241
- Vitamin B3. *See* Niacin
- Vitamin B12, 28
- Vitamin B complex, 380–381
- Vitamin C, 493
- absorption, metabolism and disposition, 279–280, 280f
 - deficiency, 279
 - function, 279
 - genetic variation, 280–281
 - glutathione transferases, 280–281
 - haptoglobin, 281
 - health significance, 281
 - SLC23A1, 280
 - SLC23A2, 280
- Vitamin D
- bone metabolism, 492
 - calcium interaction, 378–379
 - deficiency and excessive intake, 247
 - function, 246
 - major proteins and genetic variation, 248–253
 - circulation, 250–251
 - kidney, 249–250
 - liver, 249
 - skin, 248–249
 - target cells, 251–253
 - metabolism, 247–248, 247f
 - metabolite, 246
 - natural source, 246
 - requirement, 246–247
 - and vitamin D receptor genotype interaction, 379–380
- Vitamin D receptor (VDR), 22, 178, 253, 492, 550–551
- Vitamin E (VE)
- absorption, 257–259
 - concentrations, 259–260

- Vitamin E (VE) (*Continued*)
 corresponding genes, 255–257
 future research, 260
 intake, 255
 role, 255
 tocopherols, 255
 tocotrienols, 255
- Vitamin K, 263, 521–522
 and ApoE genotype interaction, 381
 bone metabolism, 493
 calcification, 267–270
 catabolism, 266–270
 deficiency and excessive intake,
 263–264
 epoxide cycle, 265–266
 function, 263
 metabolism, 265, 265f
 pharmacogenetic aspects, 265–266
 requirements, 263
 significance, 270
 sources, 263
- Vitamin K-dependent proteins (VKDPs),
 263–265, 264t, 270
- Vitamins A, 493
- Vitamins B, 493
- Vitamins B₆
 cystathionine β -synthase (CBS), 301
 serine hydroxymethyltransferase
 (SHMT), 301
 structure, 297f
- Vitamins B₉. *See* Folate
- Vitamins B₁₂
 folate trap, 299
 methionine synthase, 298–299
 methionine synthase reductase,
 298–299
 structure, 297f
 transcobalamin II, 298
- VKORC1, 266
- W**
- Waist circumference (WC), 160–161
- Warfarin, 521–522
- Weight loss diets, 182
- White adipose tissue (WAT), 373, 374f, 432
- Whole-genome shotgun (WGS)
 sequencing, 83
- Wilms tumor, aniridia, genitourinary
 anomalies, and mental retardation
 (WAGR) syndrome, 162
- Wilson protein (ATP7B), 323
- Women Genome Health Study (WGHS),
 198–199
- X**
- X-linked dominant inheritance, 5, 5f
- X-linked recessive inheritance, 5, 5f
- Xylulose 5-phosphate, 23
- Y**
- Y-linked inheritance, 5, 5f
- Z**
- Zellweger spectrum disorders (ZSDs), 343
- Zinc
 enzymes, 321–322
 genetic variation, 322–323
 insufficient, 322
 intestinal absorption and transfer, 322
 meats and organ meats, 322

Principles of Nutrigenetics and Nutrigenomics

Fundamentals for Individualized Nutrition

Edited by

Raffaele De Caterina

University Cardiology Division, University of Pisa, Pisa, Italy

J. Alfredo Martinez

Department of Nutrition, University of Navarra, Pamplona, Spain

IMDEA food (CEI UAM + CSIC) and CIBERObn Madrid, Spain

Martin Kohlmeier

UNC Nutrition Research Institute (NRI), and Department of Nutrition, University of North Carolina School of Medicine and Gillings School of Global Public Health, Chapel Hill, NC, USA

Principles of Nutrigenetics and Nutrigenomics: Fundamentals for Individualized Nutrition is the most comprehensive foundational text on the complex topics of nutrigenetics and nutrigenomics. Edited by three leaders in the field, with contributions from the most well-cited researchers conducting groundbreaking research in the field, this book addresses how the genetic makeup influences the response to foods and nutrients as well as how nutrients affect gene expression. *Principles of Nutrigenetics and Nutrigenomics: Fundamentals for Individualized Nutrition* is broken into four parts, providing a valuable overview of genetics, nutrigenetics, and nutrigenomics, as well as a conclusion translates research into practice.

With an overview of the background, evidence, challenges, and opportunities in the field, readers will come away with a strong understanding of how this new science is the frontier of medical nutrition. *Principles of Nutrigenetics and Nutrigenomics: Fundamentals for Individualized Nutrition* is a valuable reference for students and researchers studying nutrition, genetics, medicine, and related fields.

Authoritative and foundational, this book clearly presents evidence and practical guidance from the ground up.

Key Features

- Offers a uniquely foundational, comprehensive, and systematic approach with full evidence-based coverage of established and emerging topics in nutrigenetics and nutrigenomics
- Includes a valuable guide to ethics for genetic testing for nutritional advice
- Contains definitions, methods, summaries, figures, and tables to help students and researchers grasp key concepts



ACADEMIC PRESS

An imprint of Elsevier

elsevier.com/books-and-journals

