Epigenetic Epidemiology

Karin B. Michels
Editor

Epigenetic Epidemiology

Editor
Karin B. Michels, ScD, PhD
Obstetrics and Gynecology Epidemiology Center
Department of Obstetrics, Gynecology and Reproductive Biology
Brigham and Women's Hospital
Harvard Medical School
221 Longwood Avenue
Boston, MA 02115
USA
kmichels@rics.bwh.harvard.edu

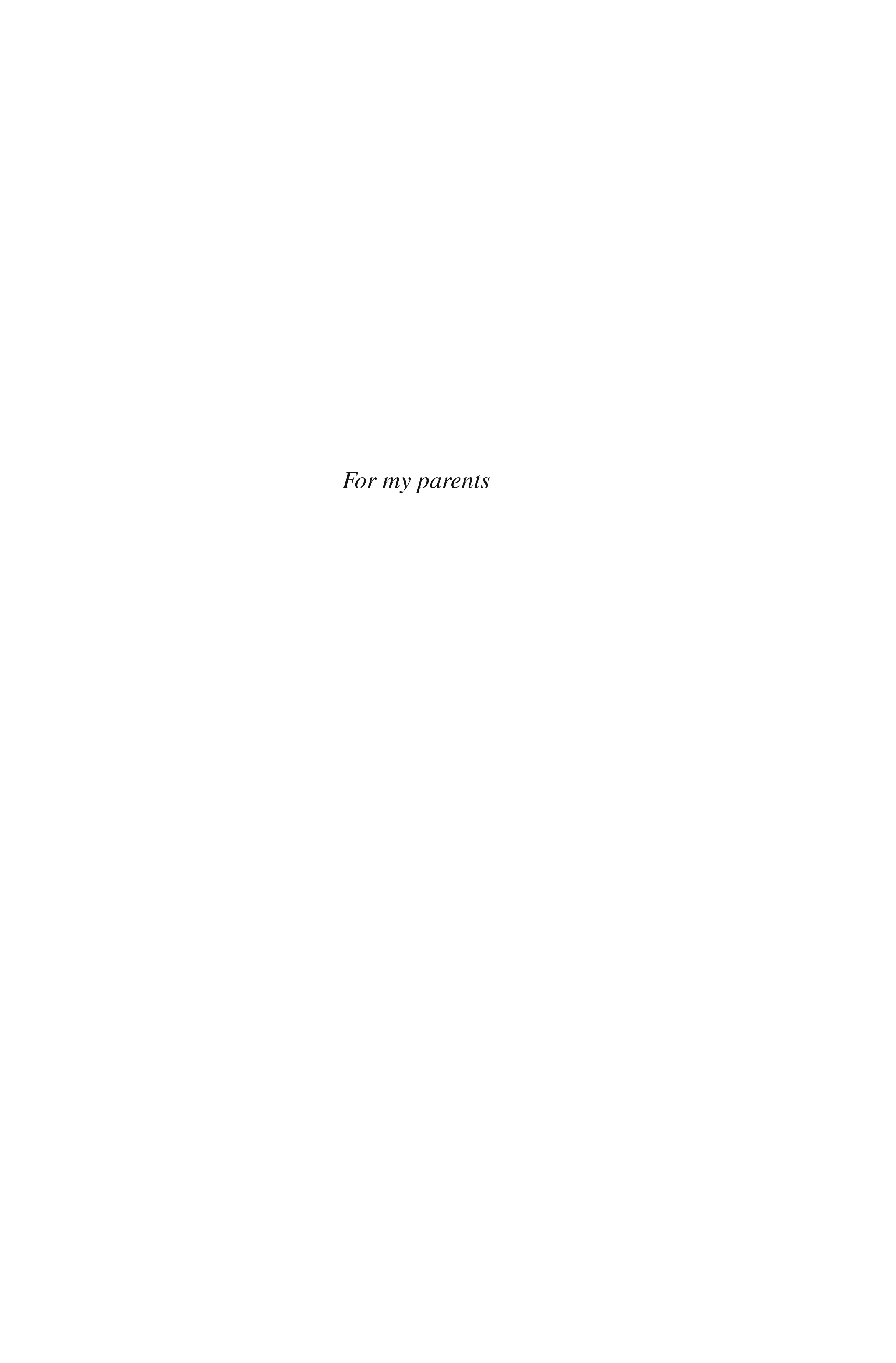
ISBN 978-94-007-2494-5 e-ISBN 978-94-007-2495-2
DOI 10.1007/978-94-007-2495-2
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2011943620

© Springer Science+Business Media B.V. 2012
No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)



For my parents

Contents

1 **Introduction** 1
Karin B. Michels

2 **The Human Epigenome** 5
Peter A. Jones and Gangning Liang

3 **Considerations in the Design, Conduct, and Interpretation of Studies in Epigenetic Epidemiology** 21
Karin B. Michels

4 **Laboratory Methods in Epigenetic Epidemiology** 37
Ludovic Barault and Rebecca C. Rancourt

5 **Biostatistical Methods in Epigenetic Epidemiology** 57
E. Andrés Houseman

6 **Epigenome Changes During Development** 77
Gavin Kelsey

7 **The Role of Epigenetics in the Developmental Origins of Health and Disease** 105
Karin B. Michels and Robert A. Waterland

8 **Epigenetics and Assisted Reproductive Technology** 117
Nicole Banks and James H. Segars

9 **Imprinting Disorders of Early Childhood** 137
I. Karen Temple, Jill Clayton-Smith, and Deborah J.G. Mackay

10 **The Utility of Twins for Epigenetic Analysis** 161
Richard Saffery, Ruth Morley, and Debra L. Foley

11 **Age-Related Variation in DNA Methylation** 185
Jean-Pierre Issa

12 **Influence of Environmental Factors on the Epigenome** 197
Brock C. Christensen, Carmen J. Marsit, and Karl T. Kelsey

13 **Epigenetic Epidemiology of Cancer** 225
Angela Risch, Narek Sarkisyan, David Scherf, Hanna Jacobsson, Wolfgang Hagmann, and Christoph Plass

14 **Epigenetic Epidemiology of Infectious Diseases**.................................. 269
Toshikazu Ushijima and Hideyuki Takeshima

15 **Epigenetic Epidemiology of Inflammation and Rheumatoid Arthritis**.. 289
Caroline Ospelt and Steffen Gay

16 **Asthma Epigenetics: Emergence of a New Paradigm?**........................ 307
Rachel L. Miller

17 **Epigenetic Epidemiology of Autism and Other Neurodevelopmental Disorders** 321
Rima Woods and Janine M. LaSalle

18 **Epigenetic Epidemiology of Psychiatric Disorders**.............................. 343
Bart P.F. Rutten and Jim van Os

19 **Epigenetic Epidemiology of Type 1 Diabetes** 377
Amanda J. MacFarlane

20 **Epigenetic Epidemiology of Obesity, Type 2 Diabetes, and Metabolic Disorders** 401
Reinhard Stöger

21 **Epigenetic Epidemiology of Atherosclerosis**... 423
Einari Aavik, Mikko P. Turunen, and Seppo Ylä-Herttuala

Index... 441

Contributors

Einari Aavik Department of Biotechnology and Molecular Medicine, A. I. Virtanen Institute, University of Eastern Finland, Neulaniementie 2, Kuopio, Finland, einari.aavik@uef.fi

Nicole Banks Program in Reproductive and Adult Endocrinology, NICHD, NIH, Building 10, CRC, Rm 1E-3140, 10 Center Drive, Bethesda, MD 20892, USA, banksnk@mail.nih.gov

Ludovic Barault Obstetrics and Gynecology Epidemiology Center, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA, Ludovic.barault@free.fr

Brock C. Christensen Department of Community and Family Medicine, Section of Biostatistics and Epidemiology, Department of Pharmacology and Toxicology, Dartmouth Medical School Hanover, Hanover, NH, USA, Brock.C.Christensen@Dartmouth.edu

Jill Clayton-Smith Genetic Medicine, Manchester Academic Health Sciences Centre, University of Manchester, Jill.clayton-smith@cmft.nhs.uk

Debra L. Foley Applied Genetics and Biostatistics, Orygen Youth Health Research Centre, Parkville, VIC, Australia

Centre for Youth Mental Health, The University of Melbourne, Parkville, VIC, Australia, dfoley@unimelb.edu.au

Steffen Gay Center of Experimental Rheumatology and Center of Human Integrative Physiology, University Hospital Zürich, Zürich, Switzerland, steffen.gay@usz.ch

Wolfgang Hagmann Division C010, Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, w.hagmann@dkfz.de

E. Andrés Houseman College Of Public Health and Human Sciences, Oregon State University, 309 Waldo Hall, Corvallis, OR 97331-6406, andres.houseman@oregonstate.edu

Jean-Pierre Issa Fels Institute for Cancer and Molecular Biology, Temple University, 3307 North Broad Street, Rm 154 – PAHB, Philadelphia, PA 19140, jpissa@temple.edu

Hanna Jacobsson Division C010, Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, h.Jacobsson@dkfz.de

Peter A. Jones Departments of Urology, Biochemistry and Molecular Biology, USC Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, 90089 Los Angeles, CA, USA, pjones@med.usc.edu

Gavin Kelsey Epigenetics Programme, The Babraham Institute, Babraham Research Campus, Cambridge, CB22 3AT, UK

Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EG, UK, gavin.kelsey@babraham.ac.uk

Karl T. Kelsey Department of Community Health, Center for Environmental Health and Technology, Providence, RI, USA

Pathology and Laboratory Medicine, The Warren Alpert Medical School of Brown University, Providence, RI, USA, Karl_Kelsey@brown.edu

Janine M. LaSalle Medical Microbiology and Immunology, Genome Center, and Medical Institute of Neurodevelopmental Disorders, University of California, Davis, CA 95616, USA, jmlasalle@ucdavis.edu

Gangning Liang Departments of Urology, Biochemistry and Molecular Biology, USC Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, 90089 Los Angeles, CA, USA, gliang@usc.edu

Amanda J. MacFarlane Department of Biology, Faculty of Science, Carleton University, Ottawa, ON, Canada

Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

Nutrition Research Division, Food Directorate, Health Canada, 251 Sir Frederick Banting Drive Ottawa, ON, K1A 0K9, Canada, Amanda.macfarlane@hc-sc.gc.ca

Deborah J.G. Mackay Division of Human Genetics, Faculty of Medicine, University of Southampton, Southampton, UK, D.J.G.Mackay@soton.ac.uk

Carmen J. Marsit Department of Pharmacology and Toxicology, Department of Community and Family Medicine, Section of Biostatistics and Epidemiology, Dartmouth Medical School, Hanover, NH, USA, carmen.marsit@dartmouth.edu

Karin B. Michels Obstetrics and Gynecology Epidemiology Center, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Avenue, Boston, MA 02115, USA

Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA, kmichels@rics.bwh.harvard.edu

Rachel L. Miller Division of Pulmonary, Allergy and Critical Care Medicine, Columbia University College of Physicians and Surgeons, New York, NY, USA, rlm14@columbia.edu

Ruth Morley Cancer, Disease and Developmental Epigenetics, Murdoch Childrens Research Institute (MCRI), Australia and Department of Paediatrics University of Melbourne, Parkville, VIC 3052, Australia, ruth.morley@mcri.edu.au

Jim van Os Department of Psychiatry and Psychology, School for Mental Health and Neuroscience, European Graduate School of Neuroscience (EURON), South Limburg Mental Health Research and Teaching Network (SEARCH), Maastricht University Medical Centre, PO BOX 616 (VIJV)6200, Maastricht, MD, The Netherlands

Department of Psychosis Studies, Institute of Psychiatry, King's College London, King's Health Partners, London, UK, j.vanos@maastrichtuniversity.nl

Caroline Ospelt Center of Experimental Rheumatology and Center of Human Integrative Physiology, University Hospital Zürich, Zürich, Switzerland, caroline.ospelt@usz.ch

Christoph Plass Division C010, Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, c.plass@dkfz.de

Rebecca C. Rancourt Obstetrics and Gynecology Epidemiology Center, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA, RRANCOUR@hsph.harvard.edu

Angela Risch Division C010, Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, a.risch@dkfz.de

Bart P.F. Rutten Department of Psychiatry and Psychology, School for Mental Health and Neuroscience, European Graduate School of Neuroscience (EURON), South Limburg Mental Health Research and Teaching Network (SEARCH), Maastricht University Medical Centre, PO BOX 616 (VIJV) 6200, Maastricht, MD, The Netherlands, b.rutten@maastrichtuniversity.nl

Richard Saffery Cancer, Disease and Developmental Epigenetics, Murdoch Childrens Research Institute (MCRI), Australia and Department of Paediatrics University of Melbourne, Parkville, VIC 3052, Australia, richard.saffery@mcri.edu.au

Narek Sarkisyan Division C010, Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, n.sarkisyan@dkfz.de

David Scherf Division C010, Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, d.scherf@dkfz.de

James H. Segars Program in Reproductive and Adult Endocrinology, NICHD, NIH, Building 10, CRC, Rm 1E-3140, 10 Center Drive, Bethesda, MD 20892, USA, segarsj@mail.nih.gov

Reinhard Stöger University of Nottingham, School of Biosciences, Nottingham, UK, reinhard.stoger@nottingham.ac.uk

Hideyuki Takeshima Division of Epigenomics, National Cancer Center Research Institute, 1-1 Tsukiji 5-chome, Chuo-ku, Tokyo, 104-0045, Japan, hitakesh@ncc.go.jp

I. Karen Temple Division of Human Genetics, Faculty of Medicine, University of Southampton, Southampton, UK, i.k.temple@soton.ac.uk

Mikko P. Turunen Ark Therapeutics LTD, Microkatu 1S, Kuopio, Finland, mikko.turunen@uef.fi

Toshikazu Ushijima Division of Epigenomics, National Cancer Center Research Institute, 1-1 Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan, tushijim@ncc.go.jp

Robert A. Waterland Department of Pediatrics, USDA/ARS Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX 77030, USA

Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA, waterland@bcm.edu

Rima Woods Medical Microbiology and Immunology, Genome Center, and Medical Institute of Neurodevelopmental Disorders, University of California, Davis, CA 95616, USA, rwoods@ucdavis.edu

Seppo Ylä-Herttuala Department of Biotechnology and Molecular Medicine, A. I. Virtanen Institute, University of Eastern Finland, Neulaniementie 2, Kuopio, Finland

Gene Therapy Unit, Kuopio University Hospital, Kuopio, Finland, Seppo. Ylaherttuala@uef.fi

Chapter 1
Introduction

Karin B. Michels

This book joins two fields, epidemiology and epigenetics, to take advantage of their respective strengths in creating the science of epigenetic epidemiology. Epidemiology is the study of the frequency, distribution, and determinants of health and disease in humans. As a science fundamental to the study of public health, epidemiology is concerned with the prevention and effective control of disease. Epidemiology has early roots with the Greek physician Hippocrates, was essential in resolving classic infectious disease epidemics such as cholera, and takes center stage in unveiling the causes of the chronic disease epidemics of our times including cardiovascular disease, diabetes, and cancer [1].

Epigenetics is the science of non-genetic mitotically heritable variation in the gene expression potential [2]. Gene expression represents the cell-specific response to intracellular and extracellular signals; epigenetic mechanisms such as DNA methylation and histone modification govern the ability to respond appropriately to these signals. While the field of epigenetics has evolved over the past three decades, interest is now increasing exponentially. Most work has been conducted in plants and animal models. While intriguing and important observations have emerged, many fundamental questions in epigenetic mechanisms in humans remain unanswered, providing abundant opportunities for discovery in the context of epidemiologic studies.

These two "epi" sciences (epidemiology: upon the people; epigenetics: above genetics) meet at the intersection of interindividual epigenetic variation and the distribution of disease [3]. Epigenetic epidemiology is defined as the study of the associations between epigenetic variation and the risk of disease in humans [2]. Marrying a bench science and a population science creates both challenges and opportunities. The amalgamation of the two fields creates a new science that supports

K.B. Michels (✉)
Obstetrics and Gynecology Epidemiology Center, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Avenue, Boston, MA 02115, USA

Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA
e-mail: kmichels@rics.bwh.harvard.edu

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_1,
© Springer Science+Business Media B.V. 2012

the study of the role of epigenetic modifications in human disease etiology, the appreciation of epigenetics as a possible mechanistic link between environmental exposures and disease outcomes, and the discovery of new disease biomarkers [3]. Since epigenetic states are amenable to changes induced by environmental stressors, identifying factors that create or correct disease-specific patterns provides new possibilities for prevention and treatment.

Epidemiology and epigenetics share the elements of time and variability. The risk of disease among individuals varies with age and is influenced by environmental factors. The epigenetic code – unlike the genetic code – is modifiable [4] and, while fairly robust [5] changes with age [6–8] and as a result of environmental influences [3, 9]. Epigenetic epidemiology relies on associations between those epigenetic marks with considerable interindividual variability and the incidence of disease.

This book is intended to be a resource for epidemiologists and epigeneticists alike. It provides insights into the mechanisms and methods in both fields to enable scientists to learn from each other, collaborate, and conduct superior studies. Epidemiologists wishing to incorporate an epigenetic component into their epidemiologic study will find useful tools here such as guidance on the appropriate epigenetic methods and specifics about strengths and weaknesses of various laboratory assays. Epigeneticists will find relevant information on how to embed their research ideas into a population-based study, how to choose their study population and design, what pitfalls to watch out for, and the appropriate statistical analyses of their research findings. The book also provides information on the formation of the epigenome during development, the role of genomic imprinting, the potential effect of assisted reproductive technology on the epigenome, the influence of age and environmental factors on the epigenetic profile, and the role of epigenetics as a mechanistic underpinning of the developmental origins of health and disease. The contributors to this book have collected and summarized the state of the art in epigenetic epidemiologic research on a number of important diseases including cancer, infectious diseases, inflammation and rheumatoid arthritis, asthma, autism and other neurodevelopmental disorders, psychiatric disorders, diabetes, obesity and metabolic disorders, and atherosclerosis.

I hope this book conveys a sense of the excitement pervading the nascer field of epigenetic epidemiology

References

1. Gordis L (2009) Epidemiology, 4th edn. Saunders, Elsevier, Philadelphia
2. Waterland RA, Michels KB (2007) Epigenetic epidemiology of the developmental origins hypothesis. Annu Rev Nutr 27:363–388
3. Michels KB (2010) The promises and challenges of epigenetic epidemiology. Exp Gerontol 45:297–301
4. Jaenisch R, Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 33(Suppl):245–254

5. Cruz-Correa M, Zhao R, Oviedo M, Bernabe RD, Lacourt M, Cardona A et al (2009) Temporal stability and age-related prevalence of loss of imprinting of the insulin-like growth factor-2 gene. Epigenetics 4:114–118
6. Wilson VL, Jones PA (1983) DNA methylation decreases in aging but not in immortal cells. Science 220:1055–1057
7. Issa JP, Ottaviano YL, Celano P, Hamilton SR, Davidson NE, Baylin SB (1994) Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. Nat Genet 7:536–540
8. Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H et al (2008) Intra-individual change over time in DNA methylation with familial clustering. JAMA 299:2877–2883
9. Christensen BC, Marsit CJ, Kelsey KT (2012) Chapter 12: Influence of environmental factors on the epigenome. In: Michels KB (ed). Epigenetic Epidemiology, Springer 197–224

Chapter 2
The Human Epigenome

Peter A. Jones and Gangning Liang

Contents

2.1 Introduction 6
2.2 Four Interacting Systems of Epigenetic Control 8
2.3 The Basics of DNA Methylation 10
2.4 Shaping of the Genome by DNA Methylation 12
2.5 Effects of the Environment on DNA Methylation 15
2.6 The Role of Aging 16
2.7 The Use of Surrogate Tissues 16
2.8 Appropriate Controls 17
2.9 Summary 18
References 18

Abstract The output of the genome is controlled by the interaction of transcription factors with the epigenome. Epigenetic processes such as DNA methylation, histone modification, histone variants, non-coding RNAs and nucleosomal remodeling machines interact with each other to ensure stable states of gene expression. These processes can become dysregulated during aging, exposure to environmental stressors and the development of cancer and other diseases. DNA methylation patterns can be relatively easily read by high throughput techniques and provide information reflecting the influence of the environment and aging on the functionality of the epigenome. Analysis of DNA methylation patterns therefore provides an exciting new route to understanding how the environment interacts with the epigenome to cause disease. Despite the promise of DNA methylation patterns for epidemiologic studies, caution in interpreting data from surrogate tissues is necessary and cellular heterogeneity can also complicate interpretation of the data. In addition,

P.A. Jones (✉) • G. Liang
Departments of Urology, Biochemistry and Molecular Biology, USC Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, 90089
Los Angeles, CA, USA
e-mail: pjones@med.usc.edu; gliang@usc.edu

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_2,
© Springer Science+Business Media B.V. 2012

DNA methylation within the body of genes can influence the response of the genome to the environment. Hypomethylation of repetitive elements can lead to genomic instability and ectopic gene expression. Methylation of coding regions can directly increase the rate of spontaneous hydrolytic mutations and increase the mutational frequency induced by carcinogens and radiation. Epigenetic processes can therefore contribute in multiple ways to the development of human diseases particularly cancer.

List of Abbreviations

ChIP	chromatin immunoprecipitation
DNMT1	DNA methyltransferase I
DNMT3A	DNA methyltransferase 3A
DNMT3B	DNA methyltransferase 3B
ncRNA	non-coding RNA
SAM	S-adenosine methionine
TDG	thymine DNA glycosylases
UDG	DNA glycosylase enzymes

2.1 Introduction

The genetic information encoded in the DNA of living organisms has to be read and interpreted in cells in such a way that its expression is highly controlled in response to developmental and environmental cues. Eukaryotic organisms, unlike prokaryotes, package their DNA into chromatin in which the fundamental building block is the nucleosome consisting of ~146 bp of DNA wrapped around an octamer of histones (Fig. 2.1). This packaging is essential to fit the DNA into the confines of the mammalian nucleus and also to provide functionality in different cell types. The combination of DNA and histones within the nucleosome is inherently refractory to transcription and nucleosomes have to be moved around or even evicted from particular places to allow gene expression to occur. This chromatin substrate, which is read by transcription factors in differentiated cell types, is what constitutes the epigenome. The accessibility or lack thereof of the genetic code is governed by chemical modifications which are applied to both the DNA and the protein components of chromatin and recent advances in high-throughput technologies now allow us to read these epigenetic modifications in their entireties in differentiated cell types.

The roles of the different modifications during development and stability of differentiated states are now beginning to become apparent, as are the switching mechanisms which occur during the development of a human from a fertilized egg. Epigenetic information is heritable in somatic cells and can be copied after DNA synthesis and mitosis to ensure stability of cellular states. However, the epigenome

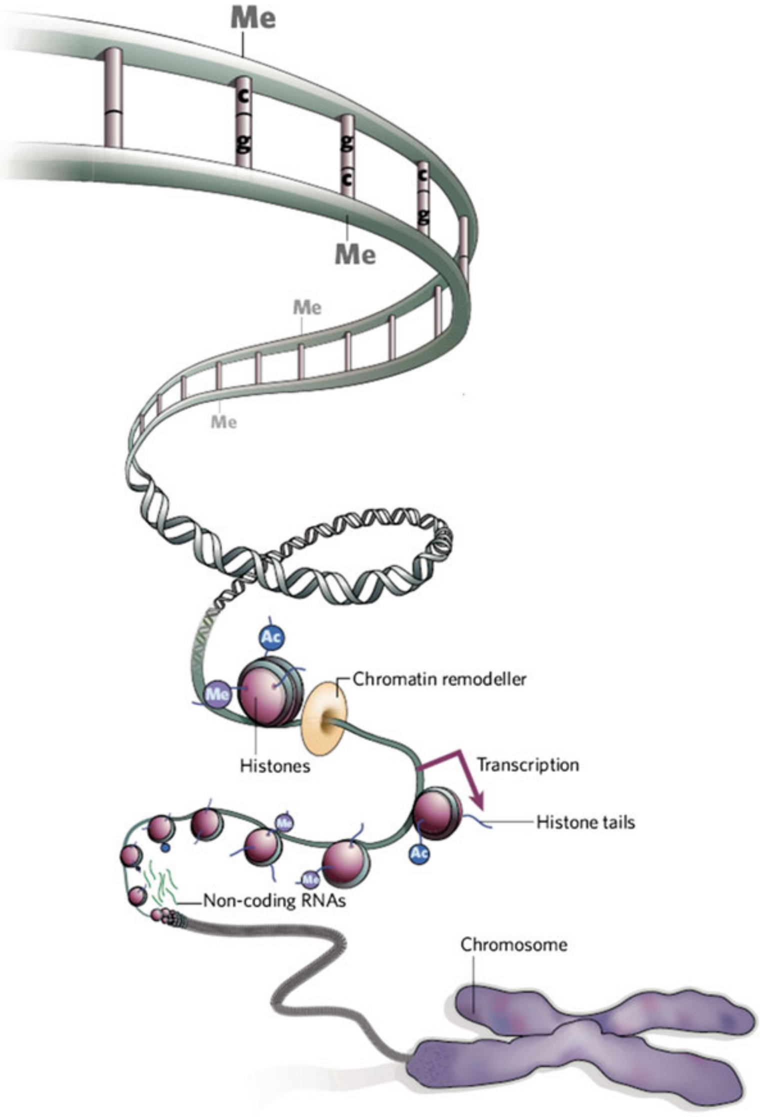


Fig. 2.1 Epigenetic processes. The DNA in living cells is complexed with proteins and RNA to fit into the structural confines of the mammalian nucleus. Most DNA is found in nucleosomes which contain about 146 base pairs of DNA wound around a histone octamer. The DNA can be modified by the application of methyl groups to cytosine residues in the simple palindromic sequence CpG. The tails of the histones and some internal amino acid residues are subject to posttranslational modifications which have significance in terms of dictating transcriptional competence. Nucleosome remodelers are necessary to expose regions of DNA so that it is accessible to the transcriptional and regulatory machinery. Non-coding RNAs also participate in the organization and functionality of chromatin. The various covalent marks communicate with each other and with remodeling machines to define the structure of different epigenomes and different cell types (Reprinted with permission from *Nature* [1])

is susceptible to alterations induced by the environment, nutrition, and other factors, so that potential changes in the packaging of genetic information can subsequently be copied in differentiated cells giving rise to both normal and abnormal cell states. Missteps in epigenetic processes can give rise to cancer and possibly to several other human diseases. We are just beginning to understand the multiple effects of the environment on epigenetic modifications and since these are potentially reversible, there is the possibility that several different diseases which have an epigenetic basis may be subject to pharmacological rectification [1].

2.2 Four Interacting Systems of Epigenetic Control

DNA in the nucleus of the cell is wrapped around an octamer containing 8 histone molecules in the fundamental structure of the nucleosome. The nucleosome contains about 146 bp of DNA and packages the DNA into the confines of the nucleus and also controls the output of the genome (Fig. 2.1). The nucleosome compactness is quite refractory to the initiation of transcription and nucleosomes need to be physically moved by nanomolecular machines to open up the DNA and allow transcription factors to initiate transcription.

The packaging and output is controlled by interaction between the various molecules which constitute chromatin and these systems interact with each other as depicted in Fig. 2.2. A variety of covalent marks and the presence of distinct histone variants, together with the involvement of non-coding RNA, are essential to the proper control of gene activity. DNA may become modified by the application of methyl groups to the 5 position of the cytosine ring and patterns of DNA methylation which are established during early development and differentiation can be copied giving rise to somatically heritable states of gene expression which can be passed from one daughter cell to the next.

The nucleosome contains two copies each of the basic histone proteins, H2A, H2B, H3 and H4, and the tails of these histones are subject to a large number of covalent modifications which convey information regarding the stability of the nucleosome and its accessibility to transcription (Fig. 2.2). Some of these covalent modifications, such as acetylation and methylation of particular lysine residues are associated with active gene transcription. On the other hand, modifications such as methylation of other lysine residues can result in transcriptional silencing. In addition, certain histone variants such as histone H3.3 or histone H2AZ are inserted into nucleosomes within specific places in the transcriptional unit and play important roles in gene activation or repression.

The covalent modification of DNA and of histones communicate with each other at a biochemical level in ways which are just now being unraveled. Together these processes can collaborate to ensure stable states of transcriptional competency and we should soon have a much better understanding of how alterations in the nucleosomal constituents collaborate to cause these relatively stable competency states.

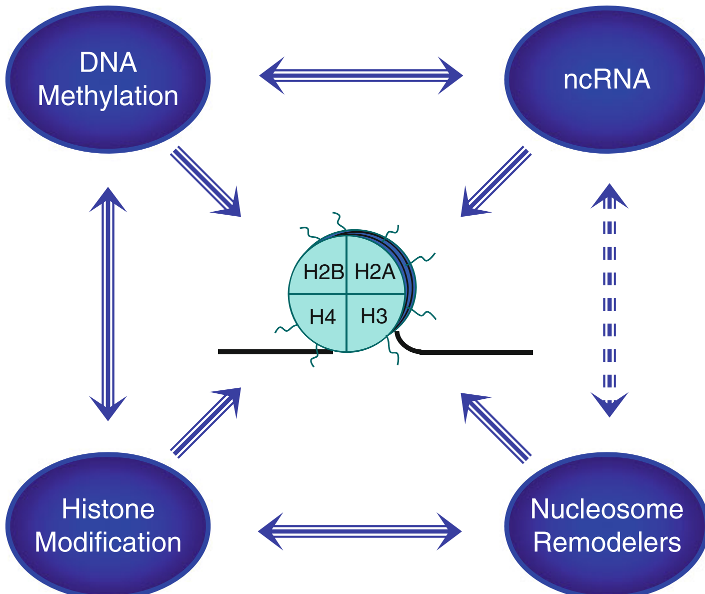


Fig. 2.2 Four interacting systems ensure epigenetic control. The four epigenetic processes outlined (among others) communicate to ensure somatically heritable states of gene expression in the context of the nucleosome. For example, DNA methylation and histone modification interact with each other and chromatin remodeling machines to position nucleosomes into active or repressive states. DNA methylation can also control the expression of non-coding RNAs which in turn can alter DNA methylation states in plants and possibly in human cells. These systems are mutually self reinforcing and therefore can both initiate silencing and maintain previously silenced states

Figure 2.2 also depicts the role of nucleosomal remodeling in epigenetic processes. As mentioned earlier, the presence of nucleosomes at transcriptional start sites is inherently refractory to transcriptional initiation and a whole series of multi protein complexes uses the energy of ATP to move nucleosomes around and expose different parts of the DNA thus allowing transcription to occur. Unfortunately, little is now known as to how the covalent marks on the DNA and histone proteins communicate with these remodeling machines.

Recent work has also suggested the potentially important role of non-coding RNA (ncRNA) species in the establishment and stability of epigenetic states. The role of RNA is quite well understood in organisms such as yeast and also in plants where it has been shown directly that ncRNAs can lead to DNA methylation and histone modifications which are important for keeping particular chromosomal regions silent by the formation of more densely packed configurations. Also certain microRNAs can downregulate chromatin modifiers such as DNA methyltransferases [2] or histone methyltransferases [3, 4]. The exact role of RNA in human epigenetic states is not clear at the present but it seems very likely that it will play a major role in their establishment and stability.

Recent developments in chromatin immunoprecipitation (ChIP) procedures and the coupling of these approaches with high throughput sequencing now makes it relatively simple to map the distribution of the different epigenetic marks on a genome-wide basis. It is likely that the availability of epigenomic maps will make it easier for us to understand how the marks communicate with each other and their potential roles in human development and disease.

Although it is clear that all of the processes outlined in Fig. 2.2 are contributing to epigenetic behavior and that all of them might be potentially altered by different environmental and nutritional influences. It is very likely that nutrition and the environment cause immediate and potentially reversible alterations to histone modifications which could be the subject of detailed epidemiologic studies. This review will, however, focus on the promise and potential of DNA methylation analysis for epidemiologic studies because the 5-methylcytosine mark is inherently more stable than the chromatin structure and the code of DNA methylation can be more easily and quantitatively read so that its role in disease states can be better understood.

2.3 The Basics of DNA Methylation

About 1% of the cytosine residues in human DNA become methylated after the DNA is synthesized by the application of a methyl group from S-adenosine methionine (SAM) to the 5 position of the cytosine ring (Fig. 2.3a). The modification occurs very shortly after DNA has been synthesized although there clearly is some methylation which occurs hours after the DNA has left the replication fork [5]. There are at least three enzymes which are responsible for setting up and maintaining DNA methylation patterns. DNA methyltransferase 3A (DNMT3A) and DNMT3B are thought to act early in development and can apply methyl groups to unmethylated and hemimethylated DNA (in which one strand has methylation and the other not). DNMT1 is thought to act primarily as a "maintenance enzyme" [6, 7] in that it has a preference for hemimethylated DNA and is the most active DNA methyltransferase in somatic cells [8]. All three of these enzymes have been shown, in gene knockout experiments in mice to be essential for mouse development demonstrating that DNA methylation is required for mammalian development.

These enzymes were thought to act largely in an autonomous manner with the DNMT3A and 3B "*de novo* methylases" required for early establishment of methylation patterns and DNMT1 then taking over and faithfully copy these patterns once they had been established. However, we have recently argued that this is unlikely to be the case and have proposed a cooperativity between the enzymes in order to appropriately maintain DNA methylation patterns in normal and transformed cells [9]. The methylation activities of these enzymes are also regulated by accessory proteins which alter the methylating capability of the enzymes. For example, the catalytically inactive form of DNMT3 (DNMT3L) is strongly expressed in embryonic stem cells and is highly stimulatory to the DNA methylating activity of both DNMT3A and 3B [8].

Fig. 2.3 Covalent modifications in DNA. (**a**) Almost all of the cytosine methylation in human DNA occurs in the simple palindrome CpG in which either both cytosines are methylated or neither are methylated as shown. The recent demonstration of 5-hydroxymethylcytosine in DNA suggests that a certain number of these sites might be further modified by the TET proteins as indicated above. Recently non-CpG methylation in the sequence CHG (where *H* represents any base other than *G*) have been observed in human embryonic stem cells. These very recent discoveries complicate the further dissection of the role of 5-methylcytosine in gene control and need to be considered in epidemiologic studies. (**b**) Cytosine residues in DNA are modified by the application of a methyl group from S-adenosyl methionine to the 5 position of the cytosine ring. Recently the TET proteins have been shown to be capable of further modification of the 5-methylcytosine to 5-hydroxymethylcytosine. This might represent a pathway to demethylation or alternatively have consequences for gene control

The existence of tissue-specific patterns of DNA methylation has been known for a long time and these patterns are known to be strongly associated with gene expression. Methylation of gene promoters is commonly linked to silencing whereas methylation in the gene body does not block transcription or silence alterative transcripts in a tissue specific manner [10]. The complexity of maintaining DNA methylation patterns is not completely understood, however, errors in the process can occur under normal conditions such as aging and in abnormal situations such as those which occur as a result of exposure to environmental insults. For this reason, the study of DNA methylation patterns in normal and diseased states has become of great importance. However, many of the observed alterations may have no direct role in the aging process or disease and it is still very difficult to distinguish between causative changes and alterations which have no functional consequence.

Almost all of the methylation of cytosine residues in somatic cell DNA occurs in the simple palindromic sequence, CpG (Fig. 2.3a). Most studies relating to DNA modification have focused on this covalent addition of the methyl group, however, it has recently become clear that other sequences and other modifications are also present in human DNA. For example, genome-wide studies in human embryonic stem cells have shown a high proportion of non-CpG methylation in particular regions of human DNA in this early developmental state [11]. The physiologic significance of this is not understood,

however, it may be due to the very high level of the DNMT3A and 3B enzymes in embryonic stem cells causing methylation at sites which subsequently lose their methylation at later stages of development possibly because the substrate generated after DNA synthesis would not be recognized and inherited by DNMT1.

Until very recently it was believed that 5-methylcytosine was the only modified base in human DNA but tremendous excitement has also been generated by the detection of 5-hydroxymethylcytosine in brain cell DNA and also in leukemia cells [12, 13]. The TET (Ten-Eleven Translocation) proteins are capable of oxidizing 5-methylcytosine to 5-hydroxymethylcytosine possibly on a pathway to active demethylation or to fulfill some currently unknown physiologic function (Fig. 2.3b). The hydroxymethylation state is not distinguished from 5-methylcytosine by standard bisulfite sequencing technology which is used to map DNA methylation in human cells. Thus, there is the potential for this modification to have been missed until now and it will likely attract a great deal of interest over the next few years.

A variety of methodologies have been used to determine DNA methylation levels and patterns, including restriction enzyme degradation, high pressure liquid chromatography and bisulfite sequencing among others. The fact that the mark can be read in DNA extracted from formalin-fixed material and seems to be stable in specimens which have been kept for a long time has encouraged the use of DNA methylation as a marker for environmental exposures with the goal of determining the influence of these exposures on epigenetic processes.

2.4 Shaping of the Genome by DNA Methylation

DNA cytosine methylation has had a profound effect on the structure of the genome because the application of the methyl group to the 5 position on the cytosine ring creates mutational hotspots in DNA [14]. The methylation of DNA in germ cells has resulted in the depletion of the methylation acceptor site CpG during the course of evolution [15]. Cytosine residues paired with guanines are known to undergo spontaneous hydrolytic deamination reactions of the order of 100 deaminations per genome per day (Fig. 2.4). The product of this deamination is a uracil residue which is not normally found in DNA and which can be rapidly and accurately repaired by ubiquitous and highly expressed uracil DNA glycosylase enzymes (UDG) which remove the uracil base and this results in the reinsertion of the cytosine residue so that no mutational events occur.

The presence of a methyl group on the 5 position of the DNA increases the rate of spontaneous deamination by about 2.5-fold [16], but more importantly, results in the generation of a thymine residue as a deamination product rather than a uracil. Thymine, being a normal constituent of DNA, is more difficult to repair. Although thymine DNA glycosylases (TDG) or MBD4 are capable of repairing most deamination sites in the correct direction there is an increased possibility of the C to T transition mutation occurring following DNA methylation and deamination. As mentioned above, this process has led to the depletion of CpG sites in the bulk of human DNA because most of the CpG sites in germ cells are methylated and

Fig. 2.4 5-Methylcytosine as a mutational hotspot. CpG sites in DNA are hotspots for C to T transition mutations in human DNA. Both cytosine and 5-methylcytosine can undergo spontaneous hydrolytic deamination to form uracil and thymine respectively. Uracil is not a DNA base and is rapidly and accurately repaired by uracil DNA glycosylase (UDG). Thymine, being a natural component of DNA, is more difficult to accurately repair by thymine DNA glycosylase (TDG) or methyl binding protein domain 4 (MBD4). This has led to the loss of CpG sites which previously were methylated in the germline during evolution and has led to the generation of CpG islands which somehow have escaped methylation in the germline. CpG methylation in the germline contributes to about 30% of all new disease causing familial mutations in humans [15]. In somatic cells, it can lead to mutations in tumor suppressor genes thus causing cancer

therefore subject to this increased mutability. Regions of DNA which are not methylated in germ cells have escaped this depletion of CpG sites and this has resulted in the presence of so-called "CpG islands" which are small regions of DNA about 1 kb in length which occur in the promoters of a substantial portion of human genes [17]. These CpG islands usually remain unmethylated in all normal tissues and this is associated with transcriptional competency. On the other hand, the abnormal methylation of these CpG sites can cause gene silencing resulting in cancer and other diseases.

Figure 2.5 outlines how some cytosine methylations can be involved in normal gene control and can directly interact with the environment. About 50% of human genes contain unmethylated CpG islands in their promoters and first exons, whereas repetitive elements including Alus and LINES tend to be methylated, as do the coding portions of genes within the exons as indicated. Abnormal methylation of the CpG islands, which can be the result of copying errors associated with cell division, aging, diet or exposure to carcinogens or other environmental stressors can result in the silencing of genes as indicated in Fig. 2.5a. This process has been well studied in cancer where between 1% and 10% of the CpG islands within genes have acquired abnormal methylation patterns during transformation [18]. The Fig. 2.5b also shows that demethylation of repetitive elements such as those within Alus and LINES can frequently occur [19] and this alteration is often related to disease outcomes in epidemiologic studies since

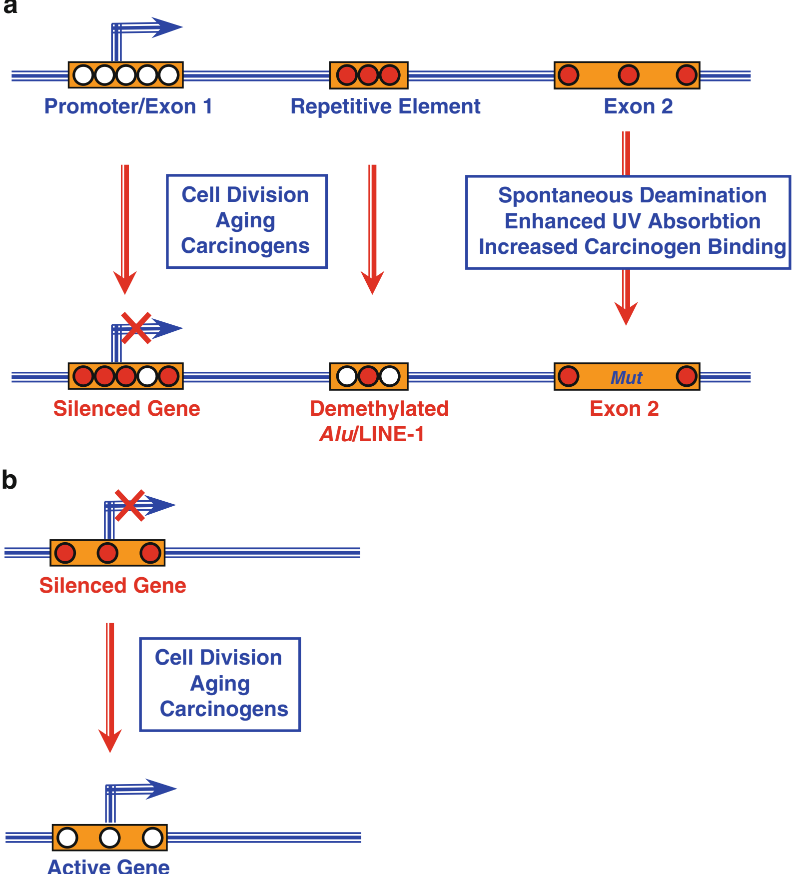


Fig. 2.5 The methylation status and effects of methylation on the carcinogenic process. (**a**) A promoter CpG island containing gene in which the gene is actively expressed due to a lack of methylation (shown as open circles) at the transcriptional start site. Oncogene and/or repetitive elements such as Alus or LINES located in the gene body are generally extensively methylated (closed circles) as are the CpG sites which are found within the coding regions of the gene. Promoter CpG islands can undergo inappropriate silencing and methylation of the CpG sites in response to aging, cell division, nutrition or exposure to environmental carcinogens. Demethylation of repetitive elements has often been observed during carcinogenesis and can be easily measured because of the high copy number of these elements in human DNA. The methylation of CpG sites within the exon can increase the rate of somatic mutations directly by increasing the frequency of C to T transition mutations. It also can alter the interaction of the DNA with the environment. For example, it can increase UV absorption and increase the binding of carcinogens to DNA. (**b**) Often overlooked, is the demethylation of non-CpG island promoters which can result in inappropriate gene inactivation as opposed to silencing during carcinogenesis

these elements are abundant in DNA and their methylation status can be measured relatively easily using quantitative techniques such as pyrosequencing [20].

Figure 2.5a also shows that the gene body methylation which occurs in exons can have profound effects on carcinogenesis. For example, the spontaneous deamination

of these methylated sites can give rise to mutations in tumor suppressor genes [21]. The presence of 5-methylcytosine in the coding regions of genes increases the rate at which mutations are introduced by ultraviolet light during the development of skin cancers [22]. This is because 5-methylcytosine absorbs UV light at a wavelength more prevalent in sunlight than cytosine, thus increasing the chances of mutations. Pfeifer et al. [22] have also shown that methylated CpG dinucleotides are the preferred targets for G to T transversions which are the most common mutations induced in mammalian cells by benzo(a)pyrene derived from tobacco smoke. Analysis of the mutational spectrum in human cancers can therefore provide evidence of value to the epidemiologist because the pattern of mutations can provide evidence of the most likely environmental cause [23]. For example, the high prevalence of C to T transition mutations at CpG dinucleotides in the p53 gene in colorectal cancer argues that these are most likely induced by endogenous processes potentially involving increased cell division [24] rather than being caused directly by exposure to carcinogens in the gut.

Figure 2.5b also shows that many tissue-specific genes which do not have CpG islands in their promoters can be ectopically activated as a function of aging and cell division. The potential role of non-CpG island methylation in gene control has been largely neglected in the field even though there is strong evidence that methylation of such regions can preclude gene expression [25]. Several studies have recently pointed to widespread hypomethylation of such regions in tumors and apparently normal cells adjacent to the tumor [19]. Since chemical carcinogens can inhibit DNA methylation, these processes can potentially result in the ectopic activation of genes which could play a significant role in the tumorigenic process.

2.5 Effects of the Environment on DNA Methylation

Soon after the discovery of the presence of 5-methylcytosine in DNA, work began to determine whether the levels of the modified base were altered in cancer and to determine whether chemical carcinogens could directly influence the methylation process. Lapeyre and Becker [26] showed that primary hepatocarcinomas and transplantable mouse liver tumors contained decreased levels of 5-methylcytosine relative to normal liver. Subsequently human leukemias and other uncultured tumors were found to have altered levels of DNA methylation [27]. Many such studies showed alterations of the overall levels of DNA methylation in cancer cells, however, Feinberg and Vogelstein [28] were the first to show that the methylation of specific sites within individual gene bodies were decreased in uncultured tumors. These early studies summarized by Riggs and Jones [27] established clearly that DNA methylation was fundamentally altered in cell lines and cancers.

Given the emerging interest in the potential role of 5-methylcytosine in controlling gene expression [6, 7] there was increasing research activity in determining the potential role of DNA methylation in cancer. The thrust of this work was on the potential for carcinogens to heritably alter the regulation of genes rather than on their abilities to cause mutations as discussed earlier. Early pioneers such as

Drahovsky and Morris [29] began work to determine whether chemical carcinogens could influence DNA methylation reactions in the test tube. These studies, also summarized in Riggs and Jones [27], pointed strongly to the possibility that chemicals in the environment including benzo(a)pyrene might be able to influence the DNA methylation machinery and that this could participate in the oncogenic process. Evidence that this was indeed the case came from studies of Wilson and Jones [30] using cultured cells and Wilson et al. [31] using freshly explanted normal human bronchial epithelial cells. These studies provide a logical hypothesis that analysis of human DNA extracted from cancers might provide evidence for the nature of the chemical insult which caused the methylation change.

2.6 The Role of Aging

A potential effect of aging on DNA methylation was suggested by Wilson and Jones [30] who showed that the lifespan of cells in culture was linked to the rate of overall loss of 5-methylcytosine levels. Subsequent experiments by Mays-Hoopes et al. [32] showed that alterations in DNA cytosine methylation also occurred in the intercisternal A particles (IAP genes) within mice. These studies which demonstrated hypermethylation as a function of age again emphasized that DNA methylation levels, while somatically heritable, were not completely invariant and could be used as markers for aging and exposure to chemicals.

The pioneering work of Ahuja and Issa and colleagues [33] in showing that hypermethylation of CpG islands in the colonic epithelium of people without cancer could be directly linked to the age, was also of great value in showing that the epigenome reacts to the increased cell division which accompanies aging. More recent work in mice has shown widespread and tissue specific DNA methylation changes showing that epigenetic regulation is a common feature of aging in mammals [34]. Since aging is a major risk factor for cancer, these alterations might provide a biochemical basis for the subsequent development of tumors.

Observations that DNA methylation patterns can be profoundly altered in aging in people without cancer show the plasticity of the epigenome. They also underline the importance of using age-matched controls in epidemiologic studies to investigate their alterations in this process and the relevance to development of cancer and other diseases. As mentioned previously, many DNA methylation changes may have no known significance in terms of genome function making it important for the important causative alterations to be determined in the future.

2.7 The Use of Surrogate Tissues

Epigenetic landscapes are tissue-specific and contribute to the phenotype of the cell. Unlike genetic studies, in which all differentiated cells in a subject have essentially the same markers such as SNPs etc., one cannot assume that surrogate tissues will

necessarily have the same value in assessing the effect of the environment on a given marker such as DNA methylation. It is therefore not always feasible to take an easily available tissue such as peripheral blood and use this to examine DNA methylation changes which might be occurring in a different target tissue. Epigenomic epidemiologic studies are therefore more difficult to perform and evaluate than genetic epidemiologic studies. Another complicating factor, which limits the use of blood cells, is that they are a heterogeneous mixture whose composition can change dramatically in response to other cues such as infections etc. Because each specialized type of cell in the peripheral circulation would be expected to have a different epigenomic profile, a measured change might reflect a change in cellular composition rather than a change in the pattern in a given cell type. Since epidemiologic studies often demand a large number of subjects and sometimes repeat sampling, other cells to consider are buccal cells, urine sediments, sputum and epidermal cells which can be relatively easily obtained.

Despite these reservations, useful information can be obtained from peripheral blood DNA methylation studies which might have value in determining the influence of nutrition or age for example on particular epigenomic marker. For example, the DNA methylation patterns of imprinted genes which are possibly methylated to similar extents in different tissues might be suitable as a surrogate although this remains to be shown. Another would be the methylation status of repetitive DNAs such as Alus and LINEs which do not show a great deal of inter tissue variations and which have been successfully used to measure changes in response to benzene exposure [35]. It is also important to consider the potential biological significance of relatively small changes which might be uncovered by these studies if there is an attempt being made to link the changes to a particular disease state. For example, it is not known whether small changes in the methylation status of a given promoter necessarily translates an alteration in gene expression.

2.8 Appropriate Controls

As mentioned earlier, epigenetic analyses differ from genetic analysis because the epigenome is cell type specific and is altered by environmental factors. This makes the appropriate selection of normal controls of great importance. For example, comparison of the DNA methylation patterns in a set of tumors should be compared to age-matched controls because the epigenome is known to change with aging. Another complicating issue is the fact that epigenetic changes can often be observed in the cells surrounding the tumor. For example, we found that the entire urothelium of the bladders of patients with bladder cancer is altered with respect to DNA methylation patterns [19]. Thus, the surrounding normal tissue may already harbor DNA methylation changes which are either selected for in the tumor or more probably allow the tumor to grow by altering the integrity of the epithelium. Therefore, comparisons of normal appearing surrounding tissue to similar tissues from age-matched controls who do not have the particular disease of interest, is often necessary to fully appreciate the changes which occur during the process of transformation.

Despite this complication, the existence of DNA methylation changes in normal tissues surrounding a tumor may be of great value in understanding the mechanism of carcinogenesis. It is still not known whether these changes precede the formation of the tumor or are a response of the epithelium to the presence of a tumor in a bladder. It might be possible in the future to conduct these analyses on high risk populations without cancer and predict cancer susceptibility. However, this will be limited to easily biopsied tissue and may not be applicable to all cancers, like brain cancer.

2.9 Summary

Interacting epigenetic processes ensure the somatic heritability of differentiated cell states and are set up early in development. These processes reinforce each other and can be influenced by environmental factors to alter gene expression in heritable ways which can cause disease. DNA methylation is a particularly attractive epigenetic process for epidemiologic studies since DNA methylation patterns can be quantitatively measured, are known to influence gene expression when located in controlling regions of genes and are subject to alterations associated with aging and exposure to environmental toxins. High-throughput approaches allow for the concomitant analysis of thousands of DNA methylation sites in large numbers of samples thus opening the door to future studies to link the influence of the environment to the epigenome. Caution however needs to be used when interpreting DNA methylation data, particularly because many changes may have little functional significance and there is a need to examine the cell type of origin in order to gain useful information.

Although much neglected in the field, the methylation of cytosine residues within the coding regions of genes can directly contribute to carcinogenesis by increasing the frequency of both spontaneous and induced mutations. Analysis of the mutational spectrum in different disease states can give an indication of likely exogenous or endogenous causes. The large number of new epidemiologic studies including epigenetic analyses, such as those discussed in this book suggests that we are entering an age of epigenetic epidemiology and that much will be learned about the interaction of the epigenome and the environment and its relation to disease.

References

1. American Association for Cancer Research Human Epigenome Task Force and the European Union, Network of Excellence, Scientific Advisory Board (2008) Moving AHEAD with an international human epigenome project. Nature 454:711–715
2. Iorio MV, Piovan C, Croce CM (2010) Interplay between microRNAs and the epigenetic machinery: an intricate network. Biochim Biophys Acta (BBA) – Gene Regul Mech 1799:694–701
3. Friedman JM, Liang G, Liu CC, Wolff EM, Tsai YC, Ye W et al (2009) The putative tumor suppressor microRNA-101 modulates the cancer epigenome by repressing the polycomb group protein EZH2. Cancer Res 69:2623–2629

4. Varambally S, Cao Q, Mani RS, Shankar S, Wang X, Ateeq B et al (2008) Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. Science 322:1695–1699
5. Liang G, Chan MF, Tomigahara Y, Tsai YC, Gonzales FA, Li E et al (2002) Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. Mol Cell Biol 22:480–491
6. Riggs AD (1975) X inactivation, differentiation, and DNA methylation. Cytogenet Cell Genet 14:9–25
7. Holliday R, Pugh JE (1975) DNA modification mechanisms and gene activity during development. Science 187:226–232
8. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z et al (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. Nature 448:714–717
9. Jones PA, Liang G (2009) Rethinking how DNA methylation patterns are maintained. Nat Rev Genet 10:805–811
10. Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD et al (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 466:253–257
11. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J et al (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462:315–322
12. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324:930–935
13. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324:929–930
14. Coulondre C, Miller JH, Farabaugh PJ, Gilbert W (1978) Molecular basis of base substitution hotspots in *Escherichia* coli. Nature 274:775–780
15. Jones PA, Rideout WM 3rd, Shen JC, Spruck CH, Tsai YC (1992) Methylation, mutation and cancer. Bioessays 14:33–36
16. Shen JC, Rideout WM 3rd, Jones PA (1994) The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. Nucleic Acids Res 22:972–976
17. Takai D, Jones PA (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci USA 99:3740–3745
18. Jones PA, Baylin SB (2007) The epigenomics of cancer. Cell 128:683–692
19. Wolff EM, Byun HM, Han HF, Sharma S, Nichols PW, Siegmund KD et al (2010) Hypomethylation of a LINE-1 promoter activates an alternate transcript of the MET oncogene in bladders with cancer. PLoS Genet 6:e1000917
20. Yang AS, Estecio MR, Doshi K, Kondo Y, Tajara EH, Issa JP (2004) A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. Nucleic Acids Res 32:e38
21. Rideout WM III, Coetzee GA, Olumi AF, Jones PA (1990) 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. Science 249:1288–1290
22. Pfeifer GP, Tang M, Denissenko MF (2000) Mutation hotspots and DNA methylation. Curr Top Microbiol Immunol 249:1–19
23. Jones PA, Buckley JD, Henderson BE, Ross RK, Pike MC (1991) From gene to carcinogen: a rapidly evolving field in molecular epidemiology. Cancer Res 51:3617–3620
24. Preston-Martin S, Pike MC, Ross RK, Jones PA, Henderson BE (1990) Increased cell division as a cause of human cancer. Cancer Res 50:7415–7421
25. Simonsson S, Gurdon J (2004) DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. Nat Cell Biol 6:984–990
26. Lapeyre JN, Becker FF (1979) 5-Methylcytosine content of nuclear DNA during chemical hepatocarcinogenesis and in carcinomas which result. Biochem Biophys Res Commun 87:698–705
27. Riggs AD, Jones PA (1983) 5-Methylcytosine, gene regulation, and cancer. Adv Cancer Res 40:1–30

28. Feinberg AP, Vogelstein B (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature 301:89–92
29. Drahovsky D, Morris NR (1972) The mechanism of action of rat liver DNA methylase. 3. Nucleotide requirements for binding and methylation. Biochim Biophys Acta 277:245–250
30. Wilson VL, Jones PA (1983) Inhibition of DNA methylation by chemical carcinogens in vitro. Cell 32:239–246
31. Wilson VL, Smith RA, Longoria J, Liotta MA, Harper CM, Harris CC (1987) Chemical carcinogen-induced decreases in genomic 5-methyldeoxycytidine content of normal human bronchial epithelial cells. Proc Natl Acad Sci USA 84:3298–3301
32. Mays-Hoopes LL, Brown A, Huang RC (1983) Methylation and rearrangement of mouse intracisternal a particle genes in development, aging, and myeloma. Mol Cell Biol 3:1371–1380
33. Ahuja N, Li Q, Mohan AL, Baylin SB, Issa JP (1998) Aging and DNA methylation in colorectal mucosa and cancer. Cancer Res 58:5489–5494
34. Maegawa S, Hinkal G, Kim HS, Shen L, Zhang L, Zhang J et al (2010) Widespread and tissue specific age-related DNA methylation changes in mice. Genome Res 20:332–340
35. Bollati V, Baccarelli A, Hou L, Bonzini M, Fustinoni S, Cavallo D et al (2007) Changes in DNA methylation patterns in subjects exposed to low-dose benzene. Cancer Res 67:876–880

Chapter 3
Considerations in the Design, Conduct, and Interpretation of Studies in Epigenetic Epidemiology

Karin B. Michels

Contents

3.1 The Objectives of Studies in Epigenetic Epidemiology 22
3.2 The Tissue Question 23
3.3 Selecting the Epigenetic Mark to Study 24
3.4 Study Designs in Epigenetic Epidemiology 26
3.4.1 Cross-Sectional Studies 26
3.4.2 Retrospective Case-Control Studies 27
3.4.3 Cohort Studies 27
3.4.4 Nested Case-Control Studies 27
3.4.5 Intervention Studies 28
3.4.6 Family-Based Studies 28
3.4.7 Birth Cohorts 28
3.4.8 Summary 28
3.5 Other Methodologic Considerations in Epigenetic Epidemiology 29
3.5.1 Choice of the Study Population 29
3.5.2 Sample Size 29
3.5.3 Effect Modification 30
3.5.4 Confounding 30
3.5.5 Misclassification 31
3.6 Evaluating and Interpreting Results of Epigenetic Epidemiology Studies 31
References 33

Abstract Studies in epigenetic epidemiology may identify epigenetic aberrations associated with disease, link environmental and lifestyle factors to the epigenetic profile, or unveil epigenetic mechanisms underlying statistical associations between risk factors and disease outcomes. Epidemiologic studies provide the framework for

K.B. Michels (✉)
Obstetrics and Gynecology Epidemiology Center, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Avenue, Boston, MA 02115, USA

Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA
e-mail: kmichels@rics.bwh.harvard.edu

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_3,

© Springer Science+Business Media B.V. 2012

identifying epigenetic biomarkers for disease risk or early detection of disease. Appropriate design considerations for studies in epigenetic epidemiology are imperative for their success. The tissue specificity of epigenetic marks represents a challenge in epigenetic epidemiology, and identification of disease markers in easily accessible surrogate tissues are essential for large-scale population-based studies. Nested case-control studies using biospecimens collected prior to onset of disease provide appropriate data to identify epigenetic changes preceding disease. Selecting a representative study population with sufficiently large sample size and appropriate comparison group is crucial for the validity and reproducibility of the results. Challenges in epigenetic epidemiology studies include confounding and effect modification, and identifying epigenetic marks with sufficient systematic interindividual variation.

3.1 The Objectives of Studies in Epigenetic Epidemiology

Epidemiology is primarily concerned with the frequency, distribution, and determinants of health and disease in humans [1]. Epidemiologic studies connect risk factors with disease outcomes based on distributions, often producing observations that inform basic research to identify mechanisms. A familiar example is the link between cigarette smoking and lung cancer [2], which guided basic research studies to classify nicotine as a carcinogen. Integrating epigenetics into an epidemiologic study recognizes the mechanistic link between a risk factor and disease risk [3] (Fig. 3.1). Epidemiologic studies support an association between folate deficiency and neural tube defects (NTD) [4]. In a case-control study including 48 induced abortions with NTD and 49 elective induced abortions without NTDs, the risk of NTDs increased with decreasing levels of LINE-1 methylation in brain tissue [5]. Maternal serum folate levels were lower in NTD cases than in controls and positively correlated with DNA methylation in fetal brain tissue, with a stronger correlation among cases [6].

The implications of epigenetic modifications in the risk of human disease can also be explored using the framework of an epidemiologic study (Fig. 3.1). Loss of imprinting (LOI) of insulin-like growth factor (*IGF*) 2, defined as biallelic expression of the normally monoallellically expressed gene in a parent-of-origin fashion,

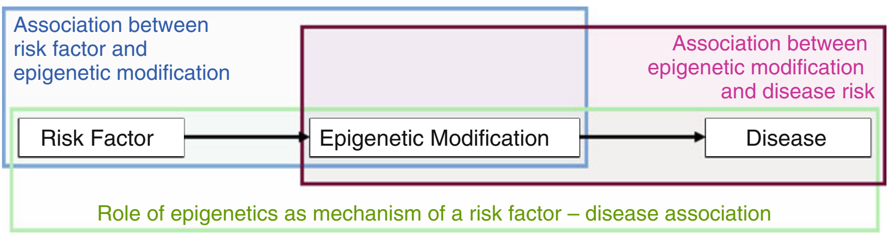


Fig. 3.1 Objectives of studies in epigenetic epidemiology

is an important etiologic factor in Beckwith-Wiedeman Syndrome [7] and Wilms' tumor [8] and was discovered by comparing the frequency of LOI of *IGF2* in children with and without these syndromes. We are only beginning to understand the role of epigenetic modifications in disease etiology; epigenetic epidemiology will provide the relevant methodologic underpinning to explore important correlations in human populations.

Whether aberrant DNA methylation, chromatin marks, or non-coding RNAs may be candidates for biomarkers of disease risk or early detection of disease can also be explored in epidemiologic studies. To date, few epigenetic biomarkers with acceptable sensitivity and specificity for disease have been identified [9–11]. Tissue-specificity of epigenetic patterns, requirements for sophisticated lab methods (e.g., capturing cancer DNA from serum) and equipment, and heterogeneity in sample collection and laboratory routines are just some of the obstacles that have hindered progress in identifying such biomarkers and moving them towards clinical application [12].

Observational studies also allow investigation of the influence of environmental and lifestyle factors on the epigenome (Fig. 3.1). Several lifestyle factors including alcohol consumption and smoking affect DNA methylation, as discussed in more detail in Chap. 12. Insights into environmental influences on the epigenome may provide targets for disease prevention.

Understanding the role of epigenetic changes in the disease process will stimulate the development of targeted interventions to prevent and treat disease. Indeed, the first epigenetic drugs have already been used to target hematologic cancers. DNA methyltransferase inhibitors (which induce hypomethylation) have been approved in the U.S. and elsewhere on the basis of randomized clinical trials to treat patients with myelodysplastic syndrome [13], and histone deacetylase inhibitors are available for the treatment of cutaneous T-cell lymphoma [14]. Moreover, DNA methylation may also predict response to therapy and foster individualized medicine. Promoter hypermethylation of the DNA repair protein MGMT is associated with a poor prognosis in various cancers because of the accumulation of mutations. However, these patients may respond well to alkylating drugs [15].

3.2 The Tissue Question

Due to its significant role in cell differentiation, one of the most profound characteristics of the epigenomic signature is its tissue specificity. Thus, the choice of the appropriate tissue is central to the success of an epigenetic epidemiology study. Changes in DNA methylation or chromatin structure are often restricted to the target organ affected by disease. Thus, when studying the contribution of epigenetics to cancer, it is important to microdissect tumor tissue to ensure a uniform cancer cell population. Moreover, tissue specificity can make epidemiologic studies of aberrant epigenetic marks very difficult. Not only is it challenging to collect samples of difficult to obtain target tissue from a sufficient number of patients with illnesses,

e.g., brain tissue from patients with Alzheimer's disease or pancreatic tissue from patients with pancreatic cancer, but it is nearly impossible to collect control tissues from healthy individuals to study the "normal" tissue-specific epigenome.

These barriers (as well as the problem of interindividual confounding discussed below) have seduced cancer researchers into using tumor-adjacent tissue that has been histologically rated tumor-free (from the same patient) for comparison. Of course, using control tissue samples from the same person avoids confounding by age or other interindividual differences. However, morphologically normal tissues adjacent to tumors harbor a number of genetic abnormalities [16–18], and epigenetic changes identified in cancer tissue have also been found in adjacent cancer-free tissue as far as 4 cm away from primary tumors [19–21]. This field effect renders the use of adjacent tissues in the quest for epimutations obsolete, and potential DNA methylation markers of tumorigenesis may be missed due to the use of inappropriate control tissue [22].

However, some tissue types are more accessible than others. For example, tissue harvested from invasive tumors in the intestinal wall of the colon displays characteristic epimutations when compared to colon mucosa from cancer-free patients [23]. Human skin biopsies show distinct methylation patterns according to sun exposure [24].

When obtaining tissues from a person with a particular disease, it is important to keep in mind that epigenetic changes may have either preceded or may be a result of the disease. Thus, target tissues should be obtained prior to disease onset to preclude reverse causation. This approach is unfortunately prohibitive for most tissues, given the number of individuals from whom target tissues would have to be harvested while they are phenotypically healthy. Prospective sampling may be possible only for specimens that can be reasonably easily collected such as blood, saliva, buccal cells, skin cells, urine, and feces.

Aberrant DNA methylation or histone modifications in such tissues may serve as surrogate markers of disease risk for many illnesses, however, the epigenetic pattern will not reflect the pattern in the target tissue. Such surrogate markers have been identified for some solid tumors [19, 25], and other chronic diseases including heart disease, asthma, and depression [26–28]. The use of blood to identify changes in DNA methylation associated with disease is hindered by the cell type-specificity of DNA methylation; any differences may be due to shifts in the distributions of the different cell types resulting from infection, inflammation, or disease. Alternatively, in patients with cancer it is possible to capture tumor DNA in the serum, however, identifying sufficient quantities of tumor DNA in the serum may prove impractical due to technical and quantitative challenges. Conversely, bladder cancer DNA has been retrieved from urine sediments [29].

3.3 Selecting the Epigenetic Mark to Study

The research question of interest generally determines the epigenetic mark to study. Generally, DNA methylation is a suitable marker for epidemiologic studies, because methyl groups are covalently bound to CpG dinucleotides and remain intact during routine DNA extraction, and cytosine methylation is fairly stable over the long term

if samples are properly processed and stored [30]. Very few epidemiologic studies have incorporated histone modifications, since the chromatin structure requires an additional immunoprecipitation step in sample preparation, which rules out the use of most existing biorepositories.

In the absence of reference epigenomes, the intraindividual variation of an epigenetic characteristic over time or the stochastic interindividual variation are not well understood. Epigenetic traits with large interindividual variation are the best candidates for epidemiologic studies, assuming that the majority of variation is systematic (and separates individuals with different characteristics such as differences in lifestyle or environmental factors or individuals with and without disease) rather than stochastic. Furthermore, for an epidemiologic study to detect interindividual variation in an epigenetic mark and link it to disease occurrence, interindividual variation has to exceed intraindividual variation. Some DNA methylation and genomic imprinting marks exhibit considerable interindividual variation and reasonable intraindividual stability over time [31], which makes them preferred markers for epidemiologic studies. However, a more detailed understanding of the DNA methylome will be necessary to identify prime candidates.

In epidemiologic studies the influence of environmental, nutritional, and lifestyle factors on the epigenome may be of interest. The focus will be on *de novo* methylation of unmethylated CpG islands in promoter regions. The stability of the methyl-cytosine bond makes active demethylation difficult to induce. While there is evidence of active demethylation in somatic cells [32], the mechanisms facilitating this process remain to be elucidated. Histone modifications are considered to be more volatile than CpG dinucleotide methylation and environmental factors may affect transcription also via this pathway. This makes chromatin an attractive target for studies in epigenetic epidemiology. The world of histone modifications, however, is complex and includes methylation, acetylation, phosphorylation, and ubiquitination making the choice of an appropriate candidate to study difficult.

Transcription is likely governed by the interplay of DNA methylation, histone modifications, and miRNAs [33]. Moreover, gene expression is not governed by epigenetic features alone, and changes in DNA methylation or chromatic structure may not affect transcription levels. Unmethylated gene promoters (or enhancers) do not necessarily facilitate expression. The functional relevance of methylation of individual CpGs within or outside of a CpG island is not always evident and may have pleiotrophic effects. Similarly, whether a difference in methylation of a few percent has functional consequences is likely gene-specific.

The candidate-gene approach targets DNA methylation or histone modification at certain candidate genes suspected or known to play a role in a specific disease process. To identify epigenetic changes associated with a disease or resulting from a particular exposure, e.g., folic acid supplementation or exposure to a chemical like bisphenol A, a genome-wide approach such as DNA methylation microarrays or ChIP-on-chip may be preferable to the candidate-gene approach in order to allow new discoveries (discussed in more detail in Chap. 4). However, genome-wide approaches are generally less sensitive than the assays used for candidate-gene approaches and rely heavily on sophisticated bioinformatics methods (see Chap. 5 for further details), which complicates interpretation of results. Various genome-wide

approaches are currently employed, and microarray-based methods, in particular the Illumina Infinium arrays for DNA methylation are widely used [34]. With the rapid advance in sequencing technologies, seq-based methods for both bisulfite converted DNA and chromatin immunoprecipitation are gaining popularity. Integrating genome-wide epigenetic scans into epidemiologic studies has become more feasible. For epigenome-wide association studies (EWAS), samples from prospective cohorts are required to ensure the temporal sequence of epimutations and disease incidence.

3.4 Study Designs in Epigenetic Epidemiology

A number of study designs can be employed in epigenetic epidemiology; the choice of the appropriate design depends on the research question (Table 3.1).

3.4.1 Cross-Sectional Studies

In a cross-sectional study, all factors of interest are assessed at one time point. For example, the proportion of individuals with a methylated CpG in a specific gene promoter in a population defined by a special characteristic, such as female smokers age 20–25 years, or the prevalence of LOI of a particular gene, say *IGF2*, in newborns, can be studied most easily with a cross-sectional design. Similarly, comparing the degree of global DNA methylation in two groups, e.g., among Caucasians age 60–65 years, males vs. females, can be accomplished effectively with a cross-sectional design.

Table 3.1 Study designs used in epigenetic epidemiology and their applications

Study design	Application
Cross-sectional study	Prevalence of an epigenetic mark in a well-defined population
Retrospective case-control study	Permanent epigenetic marks among individuals with and without disease
Cohort study	Epigenetic mechanisms underlying a risk factor-disease association
	Basis for nested case-control study
Nested (prospective) case-control study	Epigenetic marks predisposing to disease
	Biomarkers for early disease detection
	Biomarkers for disease risk
Intervention study	Effect of intervention on epigenetic pattern
Crossover study	Effect of epigenetic therapies on disease
Randomized controlled trial	
Family-based study	Transgenerational inheritance of epigenetic traits
Birth cohort	Influence of preconceptional and prenatal factors on establishment of the epigenome

3.4.2 Retrospective Case-Control Studies

In a case-control study, individuals with a disease and appropriately selected individuals free of the disease are sampled from the same source population. To study epigenetic variation, relevant biospecimens would be obtained and DNA methylation or histone modification assessed among cases and controls. In the context of epigenetics, this type of design bears similarities to a cross-sectional study, except for the particular control selection, which must be independent of the epigenetic state. The purpose of the controls is to estimate the epigenetic state of cases had they not contracted the disease under study; if the controls are not properly chosen, selection bias results. In neither the cross-sectional study nor the retrospective case-control study can it be determined whether the particular epigenetic signature assessed at the time of case and control selection preceded (and possibly caused) the disease among cases or whether it may be a consequence of (and possibly have been caused by) the disease. Therefore, epigenetic differences in retrospective case-control studies have to be interpreted with appropriate caution. Biomarkers for disease risk or early disease detection cannot be identified on the basis of retrospective case-control studies alone.

3.4.3 Cohort Studies

In a cohort, healthy individuals are recruited to participate in a longitudinal study over a certain period of follow-up time (weeks, months, years). At baseline, easy-to-obtain biospecimens are collected from all participants and stored. During follow-up, additional biospecimens may be collected. Such a biorepository provides the opportunity to study changes in DNA methylation or other epigenetic marks over time. During follow-up, disease outcomes are recorded. Cohort studies are usually large and may include many thousands of participants. Because it is not cost-effective to analyze the samples from all participants obtained at baseline and during follow-up, a nested case-control study is usually embedded in the cohort to study epigenetic marks predisposing to disease. In addition, cohort studies provide the opportunity to study epigenetic mechanisms underlying a risk factor-disease association. The major limitation of prospective studies is that for most target organs biospecimens cannot be obtained in healthy individuals.

3.4.4 Nested Case-Control Studies

A nested or prospective case-control study is embedded in a cohort study. All individuals who develop the disease of interest at any time during follow-up are selected, and appropriate controls (often 2 controls per case) are selected from those who remained free of the disease throughout follow-up. Since biospecimens from both cases and controls were obtained prior to the associating disease (and have been stored since

then), this study design permits associating the prediagnostic epigenetic status with disease outcome, i.e., it is clear that the epigenetic state reflected in the biospecimens preceded the diagnosis of disease and is not influenced by (phenotypically manifest) disease. Though the nested case-control study is a cost-effective study design, of course, prospective cohorts are generally expensive to maintain.

3.4.5 Intervention Studies

The effect of interventions such as folate and other supplements that affect the one-carbon metabolism on the epigenome or the effect of demethylating agents on disease progression can be studied in intervention studies. In an intervention study, the effect of supplements on the epigenetic pattern can be explored by comparing the profile before and after supplement use in the same person. In a randomized controlled trial, a demethylating agent is randomly assigned to treat half of the patients in the study, while the other half receives a different dose, a different drug, or even a placebo. Clinical outcomes are recorded in both groups.

3.4.6 Family-Based Studies

Transgenerational inheritance of epigenetic traits have to be studied using triads of mother, father, and child. For an epigenetic mark to be inherited transgenerationally, the putatively inherited epigenetic change must be present in both the contributed gametes and the offspring soma. Studying an additional generation would permit examining the persistent inheritance of the epigenetic mark. To date, transgenerational inheritance of epigenetic marks in humans has not been established.

3.4.7 Birth Cohorts

In a birth cohort, preconceptional and prenatal exposures can be assessed and related to DNA methylation, imprinting profiles, and chromatin states of the offspring at birth in tissues that can be easily obtained such as cord blood, cord, placenta, and saliva. Additional follow-up of the birth cohort permits tracking of developmental stages, anthropometric variables, disease outcomes, and changes in epigenetic marks over time.

3.4.8 Summary

Thus, if the goal is to identify epigenetic changes that may predispose to disease, a nested case-control study would be the most appropriate design. Similarly, epigenetic

biomarkers of susceptibility would be best studied using a nested case-control design. A retrospective case-control study may provide some useful information about an epigenetic marker of disease, but epimutations may actually be a consequence of the disease. The effect of a lifestyle factor such as nutrition or alcohol consumption on the epigenetic profile can be studied in either a cohort study or an intervention study. Epigenetics as a causal link connecting a risk factor and a disease requires samples from a cohort study, collected after the risk factor is manifest and before disease occurrence.

3.5 Other Methodologic Considerations in Epigenetic Epidemiology

Besides identifying the appropriate tissue, settling on the epigenetic marks to study, and choosing a suitable study design, other considerations are essential to ensure the success of the study (Table 3.2). Among them are the study population and its characteristics.

3.5.1 Choice of the Study Population

As only a sample of the population of interest can be studied, the characteristics of this sample determine the extent to which the results can be generalized to the population of interest. A convenience sample for a birth cohort may focus on women with uncomplicated deliveries because of the logistical ease of obtaining the biospecimens. However, the epigenetic profile among the newborns may differ from those of infants born prematurely or those whose mothers suffered from preeclampsia. Similarly, if individuals with certain characteristics are over- or underrepresented in the sample drawn, the result, e.g. the prevalence of hypermethylation of a particular tumor suppressor gene promoter, may not be representative of the prevalence in the underlying population of interest from which the sample was drawn.

3.5.2 Sample Size

One of the most important – and common – limitations of basic science research is small sample size. A sufficiently large sample is a fundamental requirement of a

Table 3.2 Some important considerations in planning and interpreting a study in epigenetic epidemiology

- Tissue-specificity: target tissue or surrogate tissue
- Restriction to well-defined representative population
- Sufficient sample size
- Effect modification
- Adjustment for confounding

high-quality study in epigenetic epidemiology: it increases the likelihood of a valid study result and is necessary to achieve adequate precision of the findings. Moderate to modest differences in epigenetic patterns between two groups may be important but cannot be detected in a small sample. Some factors may exert only modest changes on DNA methylation, e.g., diet or dietary supplements. A sufficiently large sample is necessary to detect such differences. Power calculations conducted during the planning phase of a study determine the number of participants necessary to detect a certain difference that may be of interest. Even if the expected difference is large, the number of individuals studied has to be sufficient to ensure precision of the results reflected in a narrow confidence interval around the measure of interest.

3.5.3 Effect Modification

Epigenetic marks are known not only to be tissue-specific but also to vary by race and ethnicity, sex, and age [35, 36]. If the prevalence of a particular epigenetic characteristic is of interest, e.g., the prevalence of LOI of *IGF2* among healthy individuals, such frequency is best described in the context of a well-defined substratum of the population, e.g., Caucasian females age 30–40 years. Inferences about the prevalence in other population subgroups, e.g., Hispanic males age 60–70 years, cannot be made unless the prevalence is assessed directly.

When the goal is to study the relevance of an environmental factor for the epigenetic profile (e.g., the impact of smoking on DNA methylation) or the association between an epigenetic trait and a disease endpoint (e.g., a certain histone modification and the incidence of asthma), it is important to consider whether these associations may differ by sex, age, race or ethnicity, or other factors. Is there reason to suspect smoking may induce different methylation changes in Hispanics than in African-Americans? If the answer is yes or if there is sufficient uncertainty, the heterogeneity of this association ("effect modification") can be studied across ethnicity using a statistical test for interaction. If this test yields insignificant differences, the association between smoking and DNA methylation does not differ substantially between the different ethnic subgroups considered and can be assessed in a population with mixed ethnicity. Beware, however, that the strata of different ethnicities is not too small when applying the test for interaction. Similarly, whether smoking affects DNA methylation differently in men and women can be tested. If significant effect modification by sex is identified, the association between smoking and DNA methylation should be reported separately for males and females. Each of the population subsets created by stratification needs to be sufficiently large to generate valid and reliable stratum-specific estimates.

3.5.4 Confounding

A distinctly different consideration from the above described effect modification is whether a factor may influence the strength or even direction of an association.

When evaluating whether LINE-1 hypomethylation in mammary tissue is more common among women with or without breast cancer, the distribution of age may distort the findings. Women with breast cancer may be older than women free of breast cancer, since risk increases with age. However, global methylation of repetitive elements decreases with age. Unequal distribution of age may create a spurious or exaggerated association between LINE-1 hypomethylation and breast cancer that vanishes or at least diminishes after appropriate statistical adjustment for age differences. Confounding can be detected when proper statistical adjustment changes the association of interest, i.e., the effect estimates with and without statistical adjustment for the confounder differ. A factor is a confounder if it is associated with both factors of interest, the exposure or risk factor and the disease of interest but is not in the casual pathway between risk factor and disease outcome; in the above example age is associated with both LINE-1 methylation and breast cancer. Confounding is one of the most important threats to the validity of an epidemiologic study.

3.5.5 Misclassification

As the field of epigenetics matures, so do its methods. Genome-wide scans now allow us to assess the methylation status of an increasing number of individual CpGs within and outside of CpG islands. The assessment of loss of imprinting with allele-specific expression assays using pyrosequencing has become more quantitative; previous methods relied on gel electrophoresis and radioactive labeling. Each of these improvements in technology decreases random misclassification of the epigenetic state and improves precision. Misclassification may be further reduced by the next generation of genome-wide CpG dinucleotide methylation assessment methods. Correspondingly, increasingly sophisticated bioinformatics tools allow us to distinguish between true signals and noise.

3.6 Evaluating and Interpreting Results of Epigenetic Epidemiology Studies

The establishment of a benchmark of a "normal" or "disease-free" state is fundamental in genetics and epigenetics. Any departure from this state may be linked to diseases or other phenotypes of interest. If an epigenetic change is already manifest prior to diagnosis of a disease, it may be a candidate for an early detection biomarker [3]. If it is present prior to disease development, it may be useful as a biomarker of disease risk. To date, human reference epigenomes have not been established, although the first genome-wide, single-base-resolution of methylated cytosines was recently presented for human embryonic stem cells and fetal fibroblasts [37]. In the absence of reference epigenomes, results of epigenetic epidemiology studies are often based on comparisons of epigenetic marks between two or more groups.

In epidemiology, comparing two groups (i.e., individuals with and without a certain characteristic of interest) is a substitute for the ideal (but impossible) set of information: data on the same individual measured at the same time twice, once with the characteristic of interest and once without. For example, two sets of data on the same woman, once with regular consumption of one glass of wine per day and once without alcohol consumption, would allow inference of whether her subsequent breast cancer had also arisen if no alcohol was consumed, since all other factors are identical and no confounding was possible. Such counterfactual information would permit causal inference about the role of alcohol consumption in breast cancer etiology. Comparing the fate of two individuals, one with the characteristic and one without, even in the context of a randomized controlled trial, cannot exclude differences between these individuals besides the trait of interest.

Similarly, in epigenetic epidemiology information on the same person is not available twice: with and without a characteristic to see its bias-free effect on the epigenetic mark of interest, or with and without the epigenetic trait to evaluate its effect on disease risk. In order to reduce the risk of confounding, an intraindividual comparison can be conducted to explore the effect of an intervention, such as supplementation with methyl group donors but the different time points of assesment may still introduce bias.

While DNA methylation changes with age (discussed in more detail in Chap. 11), age-related changes occur slowly, and the degree of change differs for individual genes and repetitive elements. Hence, intraindividual variation in cytosine methylation is limited. Moreover, genome-wide methylation studies have generally revealed statistically significant interindividual differences in only a fraction of genes studied, suggesting that DNA methylation is reasonably tightly controlled [38, 39]. More profound differences are generally restricted to the comparison of cancerous and normal tissues [25, 40].

In the literature, most methylation differences between two groups of individuals with and without a certain characteristic are small and amount to less than 5%. This raises the question of the functional implication of differences in LINE-1 methylation of, let us say, 80% vs. 82% or in the DMR of an imprinted gene of 42% vs. 46%. While it is advisable to examine the expression profile of the gene studied, methylation and expression status of a gene have recently been found to be less closely correlated than previously assumed [41–43]. Nevertheless, downstream effects of methylation are possible but may be difficult to determine. Moreover, differences in methylation are often reported for individual CpGs, and the functional relevance of such differences may be difficult to assess. Neighboring CpGs seem to influence each other's DNA methylation states [44] and at least in cancer tissues, entire CpG islands rather than individual CpGs are aberrantly methylated [45].

The lack of precise definitions and standardization of epigenetic phenomena further complicates the interpretation of results of studies in epigenetic epidemiology. While hyper- and hypomethlation are terms commonly used in evaluating the results of global or locus-specific methylation experiments in populations, cut-off values applied in the literature vary considerably: above and below the median [46, 47], above and below the 25th and 75th percentile [5], statistically significantly higher or

lower than the controls [48], and percent methylated reference with a differently chosen cut-off for each gene studied [49]. Such lack of standardization makes results from different studies difficult to compare. Similarly, there is no distinct definition of loss of imprinting, even in individuals. Often the normally silenced allele is expressed but at a considerably lower level than the normally expressed allele. Yet there is no set threshold level that defines loss of imprinting, leaving it to the individual investigator to make the call. A more standardized approach may aid comparability across studies.

There is increasing evidence that DNA methylation is correlated with DNA sequence and single nucleotide polymorphisms, suggesting that DNA composition predisposes CpG islands to DNA methylation [39, 50]. Furthermore, DNA methylation, the chromatin state, and non-coding RNAs interrelate. For a comprehensive understanding of the structural and functional implications of epimutations, it may be necessary to examine genetics, methylation status, histone modification, non-coding RNAs, and expression in the loci studied.

References

1. Gordis L (2009) Epidemiology, 4th edn. Saunders, Elsevier, Philadelphia
2. Doll R, Hill AB (1950) Smoking and carcinoma of the lung; preliminary report. Br Med J 2:739–748
3. Michels KB (2010) The promises and challenges of epigenetic epidemiology. Exp Gerontol 45:297–301
4. Daly LE, Kirke PN, Molloy A, Weir DG, Scott JM (1995) Folate levels and neural tube defects. Implications for prevention. JAMA 274:1698–1702
5. Wang L, Wang F, Guan J, Le J, Wu L, Zou J et al (2010) Relation between hypomethylation of long interspersed nucleotide elements and risk of neural tube defects. Am J Clin Nutr 91:1359–1367
6. Chang H, Zhang T, Zhang Z, Bao R, Fu C, Wang Z et al (2011) Tissue-specific distribution of aberrant DNA methylation associated with maternal low-folate status in human neural tube defects. J Nutr Biochem. Feb 16 [epub]
7. Ohlsson R, Nystrom A, Pfeifer-Ohlsson S, Tohonen V, Hedborg F, Schofield P et al (1993) IGF2 is parentally imprinted during human embryogenesis and in the Beckwith-Wiedemann syndrome. Nat Genet 4:94–97
8. Ogawa O, Eccles MR, Szeto J, McNoe LA, Yun K, Maw MA et al (1993) Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. Nature 362:749–751
9. Lofton-Day C, Model F, Devos T, Tetzner R, Distler J, Schuster M et al (2008) DNA methylation biomarkers for blood-based colorectal cancer screening. Clin Chem 54:414–423
10. Tanzer M, Balluff B, Distler J, Hale K, Leodolter A, Rocken C et al (2010) Performance of epigenetic markers SEPT9 and ALX4 in plasma for detection of colorectal precancerous lesions. PLoS One 5:e9061
11. Chung W, Bondaruk J, Jelinek J, Lotan Y, Liang S, Czerniak B et al (2011) Detection of Bladder Cancer Using Novel DNA Methylation Biomarkers in Urine Sediments. Cancer Epidemiol Biomarkers Prev 20:1483–1491
12. Kagan J, Srivastava S, Barker PE, Belinsky SA, Cairns P (2007) Towards Clinical Application of Methylated DNA Sequences as Cancer Biomarkers: A Joint NCI's EDRN and NIST Workshop on Standards, Methods, Assays, Reagents and Tools. Cancer Res 67:4545–4549

13. Silverman LR, Demakos EP, Peterson BL, Kornblith AB, Holland JC, Odchimar-Reissig R et al (2002) Randomized controlled trial of azacitidine in patients with the myelodysplastic syndrome: a study of the cancer and leukemia group B. J Clin Oncol 20:2429–2440
14. Piekarz RL, Frye R, Turner M, Wright JJ, Allen SL, Kirschbaum MH et al (2009) Phase II multi-institutional trial of the histone deacetylase inhibitor romidepsin as monotherapy for patients with cutaneous T-cell lymphoma. J Clin Oncol 27:5410–5417
15. Jacinto FV, Esteller M (2007) MGMT hypermethylation: a prognostic foe, a predictive friend. DNA Repair (Amst) 6:1155–1160
16. Deng G, Lu Y, Zlotnikov G, Thor AD, Smith HS (1996) Loss of heterozygosity in normal tissue adjacent to breast carcinomas. Science 274:2057–2059
17. Ellsworth DL, Ellsworth RE, Love B, Deyarmin B, Lubert SM, Mittal V et al (2004) Genomic patterns of allelic imbalance in disease free tissue adjacent to primary breast carcinomas. Breast Cancer Res Treat 88:131–139
18. Kurose K, Hoshaw-Woodard S, Adeyinka A, Lemeshow S, Watson PH, Eng C (2001) Genetic model of multi-step breast carcinogenesis involving the epithelium and stroma: clues to tumour-microenvironment interactions. Hum Mol Genet 10:1907–1913
19. Cui H, Cruz-Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S et al (2003) Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. Science 299:1753–1755
20. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J et al (2005) MGMT promoter methylation and field defect in sporadic colorectal cancer. J Natl Cancer Inst 97:1330–1338
21. Yan PS, Venkataramu C, Ibrahim A, Liu JC, Shen RZ, Diaz NM et al (2006) Mapping geographic zones of cancer risk with epigenetic biomarkers in normal breast tissue. Clin Cancer Res 12:6626–6636
22. Riazalhosseini Y, Hoheisel JD (2008) Do we use the appropriate controls for the identification of informative methylation markers for early cancer detection? Genome Biol 9:405
23. Ahuja N, Li Q, Mohan AL, Baylin SB, Issa JP (1998) Aging and DNA methylation in colorectal mucosa and cancer. Cancer Res 58:5489–5494
24. Gronniger E, Weber B, Heil O, Peters N, Stab F, Wenck H et al (2010) Aging and chronic sun exposure cause distinct epigenetic changes in human skin. PLoS Genet 6:e1000971
25. Marsit CJ, Koestler DC, Christensen BC, Karagas MR, Houseman EA, Kelsey KT (2011) DNA methylation array analysis identifies profiles of blood-derived DNA methylation associated with bladder cancer. J Clin Oncol 29:1133–1139
26. Kim M, Long TI, Arakawa K, Wang R, Yu MC, Laird PW (2010) DNA methylation as a biomarker for cardiovascular disease risk. PLoS One 5:e9692
27. Nadeau K, McDonald-Hyman C, Noth EM, Pratt B, Hammond SK, Balmes J et al (2010) Ambient air pollution impairs regulatory T-cell function in asthma. J Allergy Clin Immunol 126:845–852, e810
28. Uddin M, Koenen KC, Aiello AE, Wildman DE, De los Santos R, Galea S (2011) Epigenetic and inflammatory marker profiles associated with depression in a community-based epidemiologic sample. Psychol Med 41:997–1007
29. Friedrich MG, Weisenberger DJ, Cheng JC, Chandrasoma S, Siegmund KD, Gonzalgo ML et al (2004) Detection of methylated apoptosis-associated genes in urine sediments of bladder cancer patients. Clin Cancer Res 10:7457–7465
30. Waterland RA, Michels KB (2007) Epigenetic epidemiology of the developmental origins hypothesis. Annu Rev Nutr 27:363–388
31. Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G et al (2010) Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. FASEB J 24:3135–3144
32. Ma DK, Jang MH, Guo JU, Kitabatake Y, Chang ML, Pow-Anpongkul N et al (2009) Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. Science 323:1074–1077
33. Bernstein BE, Meissner A, Lander ES (2007) The mammalian epigenome. Cell 128:669–681

34. Laird PW (2010) Principles and challenges of genomewide DNA methylation analysis. Nat Rev Genet 11:191–203
35. Issa JP (2003) Age-related epigenetic changes and the immune system. Clin Immunol 109:103–108
36. Zhang FF, Cardarelli R, Carroll J, Fulda KG, Kaur M, Gonzalez K et al (2011) Significant differences in global genomic DNA methylation by gender and race/ethnicity in peripheral blood. Epigenetics 6:623–629
37. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J et al (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462:315–322
38. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL et al (2009) Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet 5:e1000602
39. Langevin SM, Houseman EA, Christensen BC, Wiencke JK, Nelson HH, Karagas MR et al (2011) The influence of aging, environmental exposures and local sequence features on the variation of DNA methylation in blood. Epigenetics 6:630–637
40. Marsit CJ, Christensen BC, Houseman EA, Karagas MR, Wrensch MR, Yeh RF et al (2009) Epigenetic profiling reveals etiologically distinct patterns of DNA methylation in head and neck squamous cell carcinoma. Carcinogenesis 30:416–422
41. Gort EH, Suijkerbuijk KP, Roothaan SM, Raman V, Vooijs M, van der Wall E et al (2008) Methylation of the TWIST1 promoter, TWIST1 mRNA levels, and immunohistochemical expression of TWIST1 in breast cancer. Cancer Epidemiol Biomarkers Prev 17:3325–3330
42. Diplas AI, Lambertini L, Lee MJ, Sperling R, Lee YL, Wetmur J et al (2009) Differential expression of imprinted genes in normal and IUGR human placentas. Epigenetics 4:235–240
43. Fourkala EO, Hauser-Kronberger C, Apostolidou S, Burnell M, Jones A, Grall J et al (2010) DNA methylation of polycomb group target genes in cores taken from breast cancer centre and periphery. Breast Cancer Res Treat 120:345–355
44. Bock C, Walter J, Paulsen M, Lengauer T (2008) Inter-individual variation of DNA methylation and its implications for large-scale epigenome mapping. Nucleic Acids Res 36:e55
45. Jones PA, Baylin SB (2002) The fundamental role of epigenetic events in cancer. Nat Rev Genet 3:415–428
46. Zhang FF, Morabia A, Carroll J, Gonzalez K, Fulda K, Kaur M et al (2011) Dietary patterns are associated with levels of global genomic DNA methylation in a cancer-free population. J Nutr 141:1165–1171
47. Zhu ZZ, Sparrow D, Hou L, Tarantini L, Bollati V, Litonjua AA et al (2011) Repetitive element hypomethylation in blood leukocyte DNA and cancer incidence, prevalence, and mortality in elderly individuals: the Normative Aging Study. Cancer Causes Control 22:437–447
48. Yoshida T, Yamashita S, Takamura-Enya T, Niwa T, Ando T, Enomoto S et al (2011) Alu and Satalpha hypomethylation in Helicobacter pylori-infected gastric mucosae. Int J Cancer 128:33–39
49. Saito K, Kawakami K, Matsumoto I, Oda M, Watanabe G, Minamoto T (2010) Long interspersed nuclear element 1 hypomethylation is a marker of poor prognosis in stage IA non-small cell lung cancer. Clin Cancer Res 16:2418–2426
50. Bock C, Paulsen M, Tierling S, Mikeska T, Lengauer T, Walter J (2006) CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure. PLoS Genet 2:e26

Chapter 4
Laboratory Methods in Epigenetic Epidemiology

Ludovic Barault and Rebecca C. Rancourt

Contents

4.1 Introduction 38
4.2 Choice of the Study Design 38
  4.2.1 Biospecimens and Materials 38
  4.2.2 Experimental Backgrounds for Epidemiologic Epigenetic Research Studies 39
4.3 Epigenetic Methods Targeting Genomic DNA: DNA Methylation and Hydroxymethylation 40
  4.3.1 DNA Methylation 40
  4.3.2 DNA Hydroxymethylation 44
  4.3.3 Genome-Wide Methods 45
  4.3.4 Tips and Tricks to Develop a Methylation Assay 46
4.4 Epigenetic Studies Involving RNA: Gene Expression 47
  4.4.1 Types of RNA 47
  4.4.2 Methods for Gene Expression 48
  4.4.3 Tips and Tricks to Consider for Expression Analyses 51
4.5 Epigenetic Studies Targeting Chromatin 52
  4.5.1 Nucleosome Occupancy 53
  4.5.2 Histones Modifications 53
  4.5.3 Chromatin Conformation 53
4.6 Recommendation for the Method Section in Communications and Publications 54
4.7 Conclusion and Perspective 55
References 55

L. Barault (✉)
Obstetrics and Gynecology Epidemiology Center, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
e-mail: Ludovic.barault@free.fr

R.C. Rancourt
Obstetrics and Gynecology Epidemiology Center, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
e-mail: RRANCOUR@hsph.harvard.edu

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_4,

© Springer Science+Business Media B.V. 2012

Abstract Epidemiology seeks to identify risk factors for disease and to quantify the association between environmental and lifestyle factors and disease frequency in human populations. Epigenetics may provide mechanistic clues on a cellular and molecular level behind epidemiologic links. Combining these two fields along with collecting relevant biospecimens will aid in understanding the benefits and the limitations of genome-wide and gene-specific approaches in decoding the epigenome and its functional relevance. This chapter addresses the importance of the experimental design, optimal preparation of biological samples in population-based studies, provide an overview of the available techniques to analyze **DNA methylation**, **gene expression** and **histone modification**, and highlight possible pitfalls of these approaches.

4.1 Introduction

Most scientific questions rely on good laboratory practices, and epigenetics is no different. At the start, scientists who want to perform epigenetic experiments must consider the choice of biological samples and study population, the study design, and the appropriate data analysis. This chapter provides a brief introduction to the most commonly used methods for epigenetic research.

4.2 Choice of the Study Design

4.2.1 Biospecimens and Materials

Once a hypothesis is established, it is important to first decide which samples are most suited to address the research aim. An individual has one genome but many different epigenomes, and an observation in one specific tissue type may not be the same in other tissues.

As described in Chap. 2, epigenetic alterations can affect different elements of the cells (DNA methylation, transcription levels and histone modifications) that all require specific preparation. The type and preparation of biospecimen determines which level of the epigenome can be assessed (Fig. 4.1).

- *Body fluids, fresh tissues or relevant cell lines*: if the samples are freshly collected, nucleic acid isolation (DNA and RNA), and/or histone modification analyses (if the number of cells in the sample is large enough) can be performed. Once frozen, cells burst and can no longer be sorted or isolated. In the context of epigenetic epidemiology, cell lines can be useful for development and optimization of experimental protocols, and to test the effect of certain drugs on loci of interest.
- *Formalin fixed Paraffin embedded tissues* (FFPE): up to now FFPE tissues have been used for DNA but not for RNA analysis, as the latter is often highly degraded. Although DNA is often fragmented, the ability to study large genomic

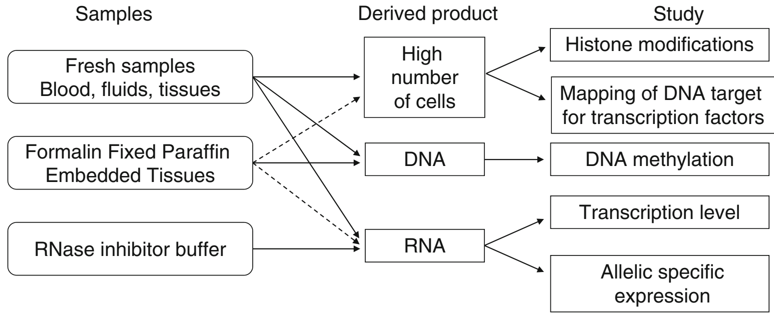


Fig. 4.1 Epigenetic studies possible according to sample availability

fragments (more than 500 base pair [bp]) may be impossible. New RNA kits have been recently developed enabling a better recovery of RNA from FFPE tissue, but the yield and quality do not allow all types of experiments (*e.g*. microarrays). Also FFPE are usually not suited for chromosomal immuno-precipitation (ChIP) analyses as the chromatin structure can be disrupted by fragmentation and the absence of good extraction methods [1].

- *Genomic DNA*: DNA can be isolated to assess cytosine methylation and hydroxy-cytosine-methylation analyses, immunoprecipitation targeting DNA (methyl cytosine) or DNA bound with supplemented transcription factors. However, it cannot be utilized for histone modification analyses since nucleosomes are disrupted during the extraction,
- *RNA*: RNA can be isolated for gene expression analyses and optimal conditions must be ensured to obtain high quality RNA. RNA can be extracted from tissues and bodily fluids. It is possible to isolate RNA from fresh frozen tissue and from frozen blood and buccal cells; however, addition of a buffer that inhibits RNAse at the time of collection (*e.g.* RNA later) is essential to aid in preventing RNA degradation. Unfortunately some experiments will become impossible to perform due to the initial cells lyses (*e.g.* micro-dissection, cell sorting, and phase separation for blood, and possibly expression microarrays).

4.2.2 *Experimental Backgrounds for Epidemiologic Epigenetic Research Studies*

Epigenetic studies require various expertise which may lead to the requirement for a multi-disciplinary team or collaborations. For each step of the process key aspects must be considered: (Fig. 4.2):

- *Sample collection:* Collection of samples may need to be performed by a pathologist or someone skilled in histology in order to select the desired cell-type since each differentiated tissue has a specific epigenome. Micro-dissection of tissue may be necessary to ensure a homogenous cell type.

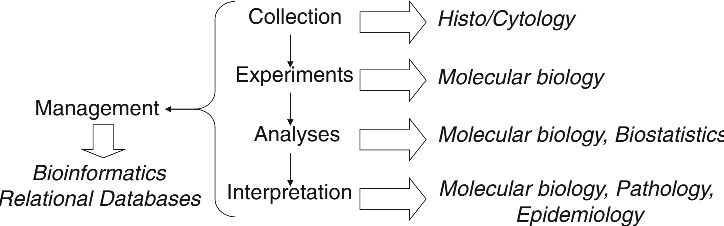


Fig. 4.2 Schematic summary of multiple-disciplines essential for good practice of epigenetic epidemiology studies

- *Experiments:* While commercially available kits have made lab experiments more robust, epigenetic experiments require a strong background in molecular biology to allow interpretation of the results at the cellular level and to avoid flaws in study design, conduct, and interpretation.
- *Analyses:* Statisticians and bioinformatics specialists can play a key role in experimental analyses especially since most of the new and high-throughput techniques require complex analyses (see Chap. 5). Moreover, it is important to involve statisticians at the study design level to ensure that the hypothesis can be properly addressed.
- *Interpretation:* Results should optimally be reviewed and interpreted by a multidisciplinary team including molecular biologists, pathologists, epidemiologist, statisticians, and other.
- *Sample and data management:* Epidemiologic studies typically involve a large number of samples, leading to a vast amount of samples and data to manage (for example, a study examining the methylation of 7 CpG sites in a gene, performed in 200 samples with replicates will produce 2,800 data points). Sample labeling and tracking could be made more efficient with the use of bar-coded tubes which are read by a scanner generating a unique sample barcode number. Spreadsheets are not an efficient solution for data management of this magnitude therefore; knowledge in bioinformatics or database management systems (SQL, Oracle, etc.) will be beneficial. Acquisition of a LIMS (Laboratory Information Management Systems) can become essential to deal with a growing through-put of the laboratory analyses.

4.3 Epigenetic Methods Targeting Genomic DNA: DNA Methylation and Hydroxymethylation

4.3.1 DNA Methylation

Discovered in the early twentieth century, 5-Methyl Cytosine (5mC; referred to as the fifth base in DNA) has been increasingly studied since the 1980s. Two different approaches are available: loci-specific and global estimation (Fig. 4.3).

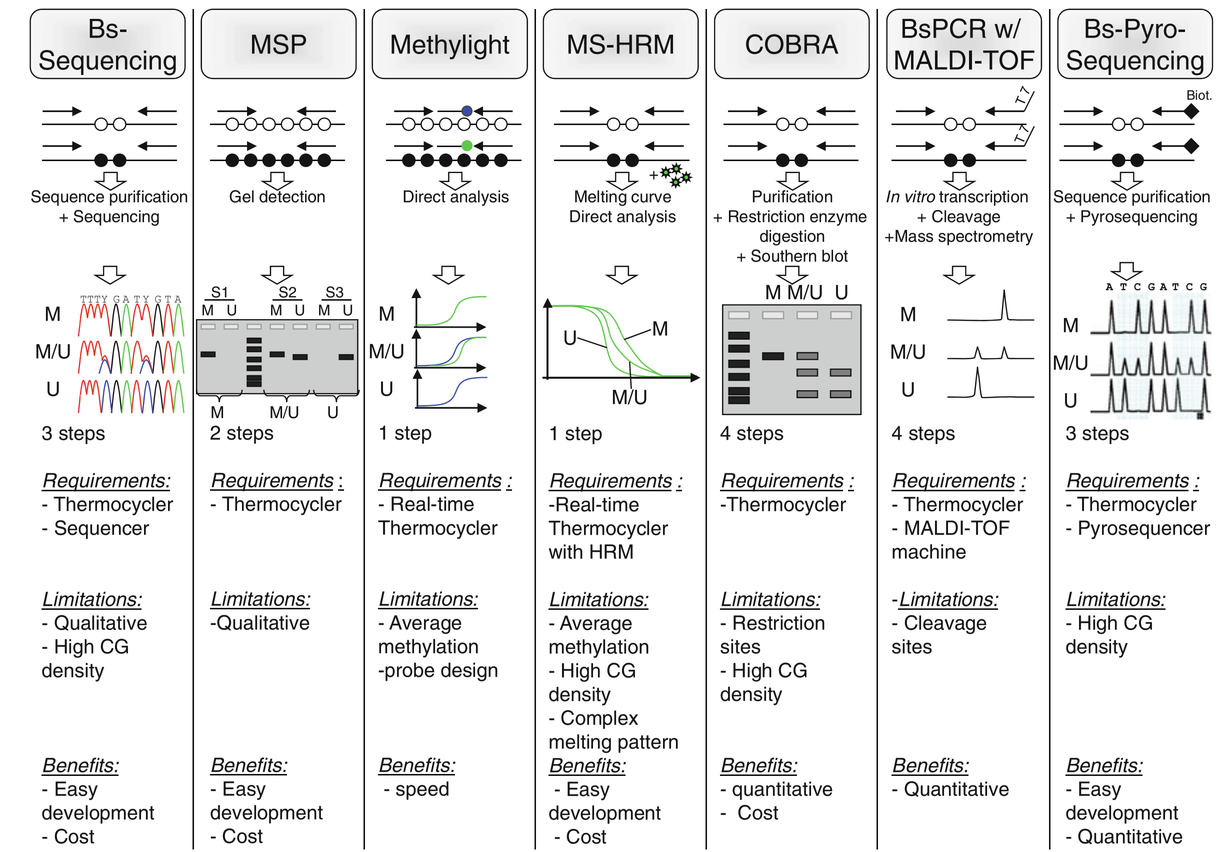


Fig. 4.3 DNA methylation methods. M: methylated; M/U: partially methylated; U: unmethylated; S1, S2, S3: sample 1, 2, 3

4.3.1.1 Loci Specific Analyses

Bisulfite treatment is a key requirement for many of the current techniques assessing DNA methylation. Various bisulfite treatment kits are now commercially available, enabling high recovery rate with a short incubation time, lower starting amounts of DNA and high conversion efficiency. Bisulfite treated DNA (BsDNA) results in single-stranded DNA of low complexity, since conversion of all unmethylated cytosines into uracil will drastically decrease the percentage of GC in the sequence. Different techniques can be used to evaluate the difference between the unmethylated and methylated DNA (see below). Single-stranded DNA, is very fragile and must be handled and stored properly (*e.g.* –20° C for short-term use, –80° C for long term storage and repeated freeze-thawing should be avoided).

- **Sensitive methods**
  Bisulfite sequencing (Bs-Sequencing) was one of the first techniques developed to assess DNA methylation [2]. Amplification is first performed on BsDNA followed by sequencing. Setting aside its cost, this technique has the benefit of providing information at each CpG sites in the sequence of interest, and can be useful as an exploratory step. However, the methylation levels cannot be quantified limiting the ability to draw associations with any transcriptional consequences.

  MSP (Methyl specific PCR): developed by Herman and colleagues in 1996 [3], this technique consists of an amplification specific to the methylated or unmethylated sequence. The targeted amplification is performed using primers with CG sites (or UG after treatment for the unmethylated DNA) in the 3′ sequence stabilizing the annealing specificity. This technique is sensitive, and requires the amplification of unmethylated sequence in order to confirm that the absence of an amplicon for the methylated sequence is really due to the lack of methylation and not a problem with the amplification protocol. MS-PCR can be performed using fluorescent primers and detection on capillary electrophoresis, as well as classical gel electrophoresis analyses.
- **Semi-quantitative methods**
  Methylight is a real-time MSP [4]. Primers and probes specific to the methylated or unmethylated sequences are used and the methylation ratio is calculated by the fluorescence ratio of the probes. The quantification of one amplicon is relative to the other making this technique semi-quantitative as it is not compared to a standard. Efficiencies of the two amplifications (methylated and unmethylated) have to be equivalent. This technique is dependent on the quality and the careful positioning of the probes. Moreover, the methylation ratio resulting from the experiment is an average of the CpG sites on the probe and the primers. It is important to note that a mutation at one of the sites will completely alter the percentage of methylation and should be avoided by carefully confirming the sequence of interest.

  MS-HRM (Methylation Specific High Resolution Melting curve analysis) is a recently developed method using new fluorescent dye chemistry [5]. Contrary to

Methylight, a PCR is performed using primers (independent of methylation) in the presence of a high saturating intercalating agent. After amplification, a melting curve is performed and is compared to a methylation scale obtaining a relative quantification of the methylation. As in Methylight, the percentage of methylation is only an average of all the CpG sites present within the amplicon. This technique requires the new generation of a real-time PCR machine, but is labor- and cost-effective compared to most other protocols.

- **Quantitative methods**

  COBRA (COmBined Restriction enzyme Analyses) was the first quantitative technique developed in 1997 [6]. This technique is labor intensive, cost-effective and does not require any specific equipment. In this protocol, methylation detection is performed using restriction enzymes. Bisulfite conversion will create a new restriction site or allow retention of a pre-existing site which will be assessed by enzyme digestion after PCR purification. Quantification of methylation is next evaluated using a Southern blot protocol (gel-based detection). The region of analysis is limited to the presence/location of a specific restriction site and the methylation is assessed for only one CpG site.

  Bisulfite PCR followed by MALDI-TOF (Bs-PCR w/MALDI-TOF) [7]. With this approach, a targeted region is amplified using primers containing a T7 promoter. Next, the amplicon is transcribed *in vitro* to generate single-stranded RNA which is then cleaved at multiple sites using RNAse specific for a nucleotide (*e.g.* RNAse A) and separated using MALDI-TOF according to their mass (base content). A methylation ratio can be calculated by the different fragment masses and intensities that are generated from the methylated and unmethylated sequence. The development of such assays can be labor intensive and dependent upon the sequence context (which will be defined by the presence of a cleavage site). However, once an assay is optimized, this technique allows high-throughput analyses

  Bisulfite Pyrosequencing (Bs-pyrosequencing) is currently the most commonly used technique to assess cytosine methylation [8]. PCR amplification is performed on Bs-DNA using primers specific to the region of interest. One of the primers is biotinylated, allowing for the binding of avidin-coated sepharose beads. The amplicon is denatured, and the non-biotinylated strand is washed off. The remaining single strand is then bound to the bottom of the reaction plate, and buffers, sequencing primer and DNA synthase enzyme are added to the reaction. Each nucleotide is then sequentially added to the reaction following a specific dispensation order. If the dispensed base is complementary to the upcoming base in the DNA strand, this nucleotide will be added in the sequence by the DNA synthase, which releases pyrophosphates that are converted into light. The emitted light is proportional to the amount of base incorporated into the DNA sequence. Therefore, the ratio of the signal following the dispensation of a C or a T at the CpG site will represent the level of methylation at this site. This technique requires a pyrosequencer, but is highly quantitative, very reproducible, and easy to develop. Bs-Pyrosequencing can also be used for allele specific methylation, by the use a sequencing primer with a single nucleotide polymorphism (SNP) in its 3′ region [9].

4.3.1.2 Global Methylation of the Genome

CpG methylase assay [10] involves a technique where the genomic DNA is treated with a CpG methylase in presence of S-adenosyl-methionine (SAM) labeled with tritium, which will transfer a methyl group on each unmethylated CpG sites. DNA is next transferred to a membrane and the radioactivity is quantified. The level of radioactivity indicates the degree of global methylation, with higher radioactivity indicative of lower methylation. Although this assay has been used for a long time, due to the use of radio-isotope this technique tends to be avoided if possible.

ELISA-type methods: in these assays, genomic DNA is coated on a plate in the presence of 5-Methyl-Cytosine antibodies. The incorporation of the antibodies is quantified (like any standard ELISA). The development of recent assay kits have decreased the starting amount of DNA, and processing time. The disadvantage of this technique is under estimation of the global methylation since 5-Methyl-Cytosine antibodies preferentially bind to areas of high CG density.

Cytosine extension assay [11] is based on single base pair extension quantification after enzyme digestion of genomic DNA. Two separate reactions are performed, one with a methylation-sensitive enzyme and the other with its isoschizomer (a methylation-insensitive enzyme). Both enzymes produce sticky ends and cut before the cytosine. The incorporation of biotinylated cytosine in the two different reactions is then quantified. This protocol was optimized for a pyrosequencer which is called LUMA [12]. The detection of the methylation is limited to the presence of the restriction site and isoschizomer.

Analysis of repetitive elements is considered a surrogate of the global methylation assays previously presented [13]. This technique consists of amplification (post bisulfite conversion) of repetitive elements like LINE1 or ALU, which are normally highly methylated, analyzing a potential hypomethylation of the genome. This protocol has been developed using different methods previously presented [14–16], though it is not as labor intensive, and requires less starting DNA than other methods. It is important to consider that repetitive elements methylation cannot be directly compared to other global methylation assays, and that there may be different mechanisms which regulate the methylation of repetitive elements and for the whole genome (global) that can be also dependent on pathology and/or the tissue/cell type. Repetitive elements may not provide an accurate representation for specific genomic regions of interest.

4.3.2 DNA Hydroxymethylation

5-hydroxymethylcytosine (5hmC) which is referred to as DNA's sixth base was first discovered in bacteriophage DNA in the 1950s. 5hmC has been a fascinating discovery as it was found present in mammalian embryonic stem cells and in cerebellar Purkinje cells [17, 18]. 5hmC is known to be derived from 5mC by the action of TET family enzymes [18].

The role of 5hmC is still not clear, but is thought have the following effects:

- to affect the annealing of methyl-binding protein to the DNA;
- to induce demethylation through the DNA base pair repair mechanism;
- initiating chromatin modification via the recognition of protein specific interactions.

An important point regarding 5hmC is the absence of conversion into uracil, during bisulfite treatment, creating the possibility of some regions of the genome being mislabeled as methylated when in fact they may be hydroxymethylated. Currently there are only a few techniques to discriminate 5hmC from 5mC:

Detection by restriction enzyme: 5hmC and 5mC are recognized by the same enzymes, but 5hmC can be protected against the enzymatic digestion by adding a protective group [*e.g.* glucosyl, or glucose]. This difference could be assessed by PCR amplification, with cleaved DNA not amplified. However, this method is limited to the recognition site of the enzyme.

Detection by antibodies: similar to the ELISA methods for global methylation, antibodies for 5hmC are available and can be used to determine the global content of 5hmC in the DNA. Up to now, no chromatin immunoprecipitation has been performed using these antibodies, but it could shed light on identifying regions affected by the 5 hydroxymethylation. One issue of using antibodies is the analysis is limited to area with dense representation of the 5hmC alteration.

GLIB (GLucosylation, perIodate oxidation, Biotinylation): A new method was recently described [19]; each 5hmC are conjugated with a molecule of glucose, which enables the subsequent addition of biotin molecules following this, streptavidin is used to pull down the DNA which can be analyzed by microarray or sequencing.

4.3.3 Genome-Wide Methods

Increasingly, more and more epigenetic studies are performed at a genome-wide scale [20]. To date these projects have mainly focused upon DNA methylation and histone modification.

- *Methylation microarray after bisulfite treatment* is currently the most used genome-wide application. In this protocol, BsDNA is amplified and put in contact with probes recognizing one CpG site each. A single base extension with labeled nucleotides will define the ratio of methylation of each CpG sites. However, this method has several drawbacks: firstly, the microarray's targets are limited to the available probes defined by the company, and are not always related to any functional annotations; secondly, the use of bisulfite treatment (as previously described) does not allow the discrimination of 5hmC. The most popular arrays have been the Illumina Golden Gate, the Infinium 27K and the Infinium 450K.
- Several genome-wide assays use restriction enzymes, and one the first methods was *MCA* (Methylated CpG Island amplification) [21]. Enrichment for

methylated DNA is performed by digestion with methylation sensitive and methylation-insensitive restriction enzymes, followed by ligation of adaptors and PCR amplification. The resulting amplicons, which are representative of the methylated fraction, are labeled and co-hybridized in a microarray platform or sequenced.

- With *RRBS* (*Reduced Representation Bisulfite Sequencing*) [22], a restriction enzyme digest is followed by purification of the fragments (within a range of specific lengths), thus enabling the enrichment of sequences containing CpGs. DNA fragments are next bisulfite converted, amplified, and finally sequenced. This labor intensive technique allows the widespread study of methylation across the genome (not exclusively in the promoter region) and works with a very low amount of DNA (less than 100 ng).
- *ChIP-Seq and ChIP-on-chip* are two methods using a chromatin immunoprecipitation as the first step. In this protocol, DNA is precipitated using antibodies specific to 5mC (MeDIP), methyl-binding proteins or histone modifications. The precipitated DNA is next used for a whole genome sequencing analysis (ChIP-Seq) or a microarray (ChIP-on-chip). Currently this is among the most expensive methods, yet these protocols offer an in-depth screening assay. The disadvantages with antibodies are the need for a high density of the target, and the possibility of relatively low specificity creating potential false positives. Moreover, these methods do not allow for the direct measurement of a methylation ratio.

4.3.4 Tips and Tricks to Develop a Methylation Assay

- *Bisulfite conversion efficiency* is an important aspect in 5mC assessment. Incomplete conversion may lead to false estimation of the methylation (particularly true with homemade bisulfite treatment and methylation specific amplification methods such as MSP, and Methylight).
  - One assessment can involve the ratio of converted BsDNA by evaluating an amplification product bearing a C outside CpG sites, and using primers independent of the conversion (without any cytosine in their sequence) [4]. The conversion of cytosine in the amplicon can be checked by probes, melting curve analysis or even sequencing.
  - Another assessment is by investigating the complete conversion of an individual DNA strand for methods which look at individual CpG sites (Bs-sequencing, Bs-PCR w/ MALDI-TOF, and Bs-Pyrosequencing). Failure of complete conversion will result in the presence of cytosine outside of CpG sites, visible in Bs-sequencing and Bs-PCR w/MALDI-TOF. This is available in Bs-Pyrosequencing by adding a C in the dispensation order (at the position of potential unconverted C).
- *Primer design* is one of the most important features of methylation assay optimization [23]. There are several rules to follow, according to the experiment to perform.

  - If the assay is using a methylation-specific amplification (MSP, Methylight), primers should contain one or more CpG site in their 3′ region to ensure a specific amplification of the (un)methylated sequence. Freeware are available for primer design (*e.g.* Methprimer).
  - For an assay using a methylation independent amplification, primers should avoid having CG in their sequence or these dinucleotides must be limited to one or two in the 5′ region. When CpG sites in the primers cannot be avoided, the Cytosine of the dinucleotide (or its complementary depending on the primer direction) should be replaced by a mismatched base independent of the methylation status (A instead of C/T or T instead of G/A).
  - The salt adjusted melting temperature (Tm) should be higher than 60°C, to enable a PCR Tm at 60°C.
  - The 3′ region containing a T derived from the conversion of cytosines not included in a CpG site will ensure strict amplification of the BsDNA.

- *Linearity of detection:* once the bisulfite conversion and the primer design are optimal, the ability to assess the different degree of methylation needs to be checked. A methylation scale should be created by pooling (in a dilution scale) unmethylated DNA with completely methylated DNA. Whole genome amplification can get rid of all methyl groups and complete methylation can be obtained with methylase treatment. The methylation scale provides a useful tool to confirm the accuracy of the primers and overall assay.
- *Reproducibility* must be confirmed by performing at least duplicates of the bisulfite treatment (not only the PCR!). Correlation between replications must to be evaluated to ensure the best assay quality and accuracy of results (for example, a difference between the bisulfite replicates less than two times the standard deviation).

4.4 Epigenetic Studies Involving RNA: Gene Expression

One important way to examine any phenotypic differences across individuals is to investigate biological function of specific genes by measuring gene expression. Isolating RNA can provide valuable samples and important factors must be considered when planning expression experiments. Selection of the tissue and bodily fluid samples should be considered due to the nature of cell-type specific expression of certain genes. Additional transcriptional features such as: the presence of multiple splice variants and isoforms, or other transcripts that may span the region of interest (*e.g.* antisense RNAs) must be thought of prior to experiments.

4.4.1 Types of RNA

Translational RNA: Messenger RNA, transfer RNA, and ribosomal RNA work together in the process of protein synthesis by translating the genetic code into

protein. Messenger RNA (mRNA) carries the protein sequence information to the ribosomes, while Transfer RNA (tRNA) is a small RNA chain (~80 nucleotides) that transfers specific amino acids for protein synthesis during translation. Ribosomal RNA (rRNA) combines with proteins in the cytoplasm to form ribosomes (the nucleoprotein). 18S, 5.8S, 28S and 5S are four the rRNA molecules found in eukaryotic ribosomes.

Non-coding RNA: Not all RNAs encode for protein structures. Many do not and these are referred to as non-coding RNAs (ncRNAs). Non-coding RNAs can have regulatory functions, and can act to down-regulate gene expression. This regulation can be direct or through the targeting of other repressive modifications which will affect gene expression (*e.g*. targeting methylation). Some types of NcRNAs are long non-coding RNAs (lncRNAs) which are classed as greater than 200 nucleotides, as well as various short ncRNAs including small interfering RNA (siRNA, ~21–22 nucleotides), microRNA (miRNA, ~22 nucleotides), small nucleolar RNAs (snoRNAs, ~80–300 nucleotides), and PIWI-interacting RNAs (piRNAs, ~26–30 nucleotides). Taft et al. have written a comprehensive review of ncRNAs and their regulatory role in diseases [24].

Antisense RNAs are transcribed in the opposite direction of one or more endogenous mRNA transcripts which are in the sense orientation. Antisense RNA can act to either activate or down-regulate gene expression by binding to the endogenous sense mRNA [25]. By definition the antisense transcript shares the same sequence of the sense counterpart and raises the need for strand-specific consideration for this common region for expression analysis.

4.4.2 Methods for Gene Expression

Expression analysis is most commonly used to examine the function of a gene(s) at the transcriptional level in a specific tissue or cell type, and for certain instances it is important to investigate the expression patterns in a strand or parent of origin specific manner.

4.4.2.1 Overall Gene Expression

Classical methods for gene expression include Northern blotting and ribonuclease protection assay (RPA). **Northern blots** generate a membrane from gel electrophoresis of RNA samples. Radio-labeled probes specific to the transcript are used to visualize RNA expression. **RPA** is a technique for quantifying and detecting RNAs with the use of radio-labeled complementary *in vitro* transcribed strand-specific probes. The probes and target RNA are hybridized and then treated with a ribonuclease mix which degrades all remaining single-stranded RNA. RNA which is hybridized to the probe will be protected from digestion and is visualized on an

Table 4.1 Methods to study gene expression

Technology	Northern blot and RPA	Expression arrays	Quantitative real-time PCR
Type of RNA assessment	Semi-quantitative and quantitative	Quantitative	Quantitative
Typical starting amount of total RNA	As low as 1 μg, (ideally 5–10 μg)	As low as 100–200 ng	as low as 50 ng (ideally 100–200 ng)
Sample size	Smaller scale	High-throughput	Moderate throughput
Benefits	Cost effective, not difficult to develop, capable of strand-specific analyses	Thousand of genes examined, little sample used	Cost effective, easy to develop, capable of strand-specific analyses
Limitations	Labor intensive, amount of RNA used	High cost	Labor intensive
Key requirements	Radioactivity	Array core facility, bioinformatics analysis	Real-time Thermocycler

acrylamide gel. Some important considerations for these well established classical methods are (Table 4.1): these assays are not ideal for large sample sizes, require a large amount of starting RNA material, and the laboratories must be capable of working with radioactive isotopes, although recent modifications to both protocols allow for non-radioactive labeling using (Digoxigenin).

Generation of cDNA. For some expression methods (see below) mRNA must be first converted to cDNA by reverse transcription. Some essential points for cDNA synthesis are outlined below in the Sects. 4.3.4 and 4.4.3.

Gene expression arrays [26] allow for a high-throughput gene expression profile which measures the expression of thousands of genes in one sample. Microarrays are glass slides coated with transcript specific probes. The RNA sample is fluorescently labeled and hybridized to the microarray slide which is processed and laser scanned. The normalization of the RNA sample for quality control is done using reference RNA (fluorescently labeled differently than the RNA sample). There are various commercially available platforms of expression arrays and some companies offer custom made arrays, for the researcher to choose their favorite genes of interest. Using low amounts of starting RNA, this technology generates vast amounts of data which requires bioinformatics analysis.

Quantitative PCR (qPCR) quantifies gene expression on a smaller scale than the arrays while requiring a small amount of starting material (this is ideal for human samples that are often precious). Unlike classical PCR, which provides an end point measurement, this reaction is done in real-time providing analysis as the reaction progresses. qPCR is monitored with the use of a fluorescent reporter molecule, which accumulates with the product amplification at each cycle. There are two common detection methods when performing real-time qPCR: non-specific detection using DNA intercalating dyes and specific detection with probes designed to a

specific target region (gene of interest). Real-time probes are labeled with a quencher dye and a reporter dye, and are available with different chemistries (hydrolysis or hybridization). Probe-based real-time PCR can be useful for qPCR of mRNA, allele specific discrimination (SNP assays), and microarray validation experiments. qPCR provides an amplification plot containing the point of amplification threshold, and the cycle number this threshold is crossed (Ct). There are different algorithms used to quantify the levels of expression [27, 28].

4.4.2.2 Allele-Specific Gene Expression

Definition of Imprinting in Relation to Expression

The aspects of genomic imprinting are described in more depth in Chaps. 6 and 9. Imprinted genes exhibit monoallelic expression which is specific to the parent-of-origin. A disruption in imprinting (loss of imprinting) could result in either a gain in expression on the normally silent allele causing biallelic gene expression or a loss of expression on the normally expressed allele which would result in an absence of expression.
The study of imprinted genes represents a challenge for different reasons:

- The nature of genomic imprinting requires the discrimination of parental origin which can be addressed by using a SNP located in the coding region (exons).
- Non-coding RNA transcripts are found in most imprinted clusters; the majority of these are antisense to reciprocally imprinted protein coding genes. The ability of strand-specific detection is often essential to address overlapping transcripts.
- Some imprinted genes have multiple transcriptional variants/ isoforms and not all variants are necessary imprinted or expressed in the same tissues (*e.g. GRB10* [29] and *IGF2* [30]).

Selection for Informative Individuals

In the example of imprinted genes, whose function is specific to the parent of origin (inheritance) it is important to discriminate between the maternal and the paternal chromosomes. When using human samples, allelic discrimination is limited to the availability of informative single nucleotide polymorphisms (SNPs). Genotyping of the samples is performed on the DNA to assess the heterozygosity of specific SNPs (defining informative individuals) for a gene. Information on the location of SNPs can be found in genome browser databases (e.g. NCBI and UCSC). If possible, multiple SNPs should be assayed for a specific gene to increase the number of informative samples from the sample pool. Although many SNPs have been well characterized in the literature, it is ideal to validate the presence of a SNP by sequencing. Both genome-wide SNP arrays and gene-specific SNP assay kits are also commercially available.

Methods for Detection of Allele-Specific (Parent of Origin) Expression

All of these methods require the region containing the SNP to be amplified via PCR [31]. These methods can be performed on both DNA for genotyping and RNA for measuring transcriptional levels.

Some SNPs generate a restriction site which can be recognized by an enzyme. The amplicon can be digested with a restriction enzyme, creating a difference in fragment sizes which distinguishes the parental alleles (digested or undigested).

If the SNP does not generate a restriction enzyme site or a more quantitative method is desired, sequence based methods are ideal (*e.g.* pyrosequencing across the SNP).

SNuPE assays (single nucleotide primer extension assay) identify and quantify allelic variants with the use of a labeled (fluorescent or radioactive) primers specific to one or each of the SNP variants in a primer extension step [32]. SNuPE can be modified to measure allele specific transcriptions using qPCR analysis on cDNA [33, 34].

An additional quantitative technique for allelic-specific detections is **Hot-Stop PCR** which involves the addition of a radio-labeled PCR primer in the final amplification cycle [35].

4.4.3 Tips and Tricks to Consider for Expression Analyses

- Knowing the type RNA of interest is important when planning experiments as certain RNA kits do not allow for the isolation of small RNAs.
- RNA quality can be assessed on a gel by measuring the ratio of 18S/28S or by using a more accurate technique like a bioanalyzer (which also provides important quantity values and requires less RNA). Stringent RNA quality assessment must be performed prior to experiments. Possible degradation and contamination will impair the quality and validity of the experiments [36].
- Since transcriptional variants partially share the same sequence and may have a different expression pattern it is important to find a specific SNP to distinguish the individual variants.
- Choosing a reverse transcriptase primer for cDNA synthesis: cDNA synthesis kits provide OligoDT, random hexamers, or a mixture of both for priming oligos.
  - OligoDT offers a more specific priming to poly(A) mRNA tail (as well as polyA rich regions) giving greater assurance that the mRNA is intact.
  - Random hexamer is recommended for total RNA especially if the transcription of interest does not have a polyA tail or is unknown.
- A gene-specific primer is needed to identify the orientation when characterizing a new transcript or when measuring expression of a gene that has an additional

transcript spanning the region in the opposite orientation (antisense). In order to perform RT-PCR in a strand-specific manner (*e.g.* orientation specific), the use of a primer specific to the gene of interest is used during the initial cDNA synthesis instead of random hexamer and oligoDT primers. Therefore, two separate cDNAs are generated: one with a forward gene specific primer and one with a reverse gene specific primer.

- *Controls:* All experiments are only as good as the controls they have. Without the proper controls, the analysis is next to impossible to interpret.
  - Whenever possible the primers should be exon spanning to allow specific amplification of the RNA instead of any possible contaminating gDNA.
  - During the cDNA synthesis, up to half of the reaction must be kept as the negative control (RT minus) with water added in place of the reverse transcriptase enzyme to detect possible gDNA contamination.
  - A housekeeping gene provides the needed normalization step to accurately measure the levels of cDNA samples.
    - Commonly used housekeeping genes include: GAPDH, β-actin, α-tubulin, HPRT, and 18S.
      [Note: OligoDT cannot prime 18S rRNA eliminating the possibility of 18S being used as a housekeeping gene.]
    - It is important to confirm that the housekeeping gene has ubiquitous expression in the tissue type studied.
- **Replicates** (biological and technical) are essential due to possible well to well variation, pipetting error, or poor efficiency of cDNA synthesis.
- **Criteria for primer design and target amplicon**
  - Primers should be around 15–30 base-pairs in length and the melting temperature (Tm) of both primers should be similar or equal to each other. Ideal melting temperatures are around 58–62°C (low temperature can result in non-specific annealing).
  - Confirm that the primers are specific to the targeted region by running the sequence around the primers and probe through a genome database software. Primers with low specificity can amplify other regions of the genome. Avoid repetitive sequences and runs of identical nucleotides.
  - The ideal size for the amplicon length is 50–150 bp, and should not exceed 400 bp. Avoid heavily repetitive sequences by checking the target sequence in repeat masker software.

4.5 Epigenetic Studies Targeting Chromatin

Epigenetic control of gene expression can involve the chromatin structure, through proteins called histones that together with the DNA form the repeating nucleosomal units of chromatin. Modifications to the histone N-terminal tails are known to be somatically heritable, have a role in regulating gene expression and occur

post-translationally through methylation, acetylation, phosphorylation, ubiquitination and by others methods. Various histone modifications have been identified in many regions throughout the genome. Ongoing genome-wide screens are being performed in order to understand the possible link of histone marks with gene expression and chromatin conformation [37, 38].

4.5.1 Nucleosome Occupancy

Nucleosome occupancy reflects the condensation of the chromatin structure, which can affect gene transcription. Nucleosome positioning in a specific locus can be assessed by treating cells or tissue samples with a DNA methylase (*e.g.* CpG Methyl transferase M.SssI), which will add a methyl group on cytosines that are not at the level of the histones core [39]. It results in a difference of methylation pattern depending on the presence of histones and this difference can be assessed using DNA methylation methods. Due to the nature of this protocol, DNA derived from these experiments can be used only for this type of analysis.

4.5.2 Histones Modifications

Most histone modification analyses are through immunoprecipitation [1] where cells are incubated with antibodies targeting a specific modification after crosslinking the histones to the DNA (chemical fixation of the nucleosome). The chromatin immunoprecipitation (ChIP) protocol allows for the enrichment of DNA associated with histone modifications which can then be analyzed by sequencing (ChIP-Seq), or microarray (ChIP-on-chip). The results of previously performed ChIP assays are available in different web databases (*e.g.* UCSC) for various tissue types.

4.5.3 Chromatin Conformation

The chromosome conformation capture assay is a technique used to identify and quantify possible physical interactions that are occurring between any two genomic loci enabling the study of *in vivo* genomic organization and interacting chromatin segments over vast regions of up to several hundreds of kilobases in size [40]. The captured interactions could be specific interactions (*e.g.* between genes and regulatory elements), or random collisions between loci [41]. In the chromatin conformation protocol, cells are first crosslinked which leads to stabilization or capture of the DNA-DNA, as well as DNA-protein interactions. The DNA is then digested with a restriction enzyme and the fragments are ligated, creating a mixture interacting fragments. Sequence analysis of the area of ligation allows the identification

of the interacting chromatin. A variant of this technique was described by Lieberman-Aiden and colleagues, which enables the genome-wide analysis of the chromatin conformation [42].

4.6 Recommendation for the Method Section in Communications and Publications

Providing detailed information such as experimental controls, replicates, chromosomal location and primer details (including the original sequence for DNA methylation assays) are essential for the experimental reproducibility by the scientific community (Table 4.2).

Table 4.2 Key requirements for information to be provided in publications on epigenetic epidemiology studies

Experimental analysis	Key requirements for publication	Examples
Cohort samples	Collection	Cohort details
	Isolation protocol	Kits used, time until sample processing
	Samples	Sample/cell type
	Storage	Storage temperature and duration
Epigenetic study focused on DNA	Target sequence	Accession number, probes filtering for arrays
	Primers sequence	Unconverted and converted sequence
	Reproducibility	Coefficient of variation
	Replication	Number and type of replicates
	Controls	Number and type of controls
Epigenetic study focused on RNA	Target sequence	Accession number
	Primers and probe sequence	Sequence or reference number
	Variant assessment	SNP reference and location
	Replication	Number and type of replicates
	Controls	Housekeeping gene, RT minus
Epigenetic study focused on chromatin	Chromatin preparation	Tissue/cell type, isolation details, cross-linking
	Primers, probes and target sequences	Accession number, and assay conditions
	Antibody information	Reference, dilution
	Replication	Number and type of replicates
	Controls	Type of controls (*e.g.* negative control for IP, antibody specificity control, chromatin conformation ligation template control)

4.7 Conclusion and Perspective

An in-depth knowledge of the relevant methods is essential for defining the experimental hypotheses and aims, as well as making the most efficient use of biological samples and cohorts.

References

1. Collas P (2010) The current state of chromatin immunoprecipitation. Mol Biotechnol 45:87–100
2. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW et al (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci USA 89:1827–1831
3. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB (1996) Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci USA 93:9821–9826
4. Campan M, Weisenberger DJ, Trinh B, Laird PW (2009) MethyLight. Methods Mol Biol 507:325–337
5. Wojdacz TK, Dobrovic A, Hansen LL (2008) Methylation-sensitive high-resolution melting. Nat Protoc 3:1903–1908
6. Xiong Z, Laird PW (1997) COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res 25:2532–2534
7. Ehrich M, Zoll S, Sur S, van den Boom D (2007) A new method for accurate assessment of DNA quality after bisulfite treatment. Nucleic Acids Res 35:e29
8. Tost J, Dunker J, Gut IG (2003) Analysis and quantification of multiple methylation variable positions in CpG islands by Pyrosequencing. Biotechniques 35:152–156
9. Wong HL, Byun HM, Kwan JM, Campan M, Ingles SA, Laird PW et al (2006) Rapid and quantitative method of allele-specific DNA methylation analysis. Biotechniques 41:734–739
10. Balaghi M, Wagner C (1993) DNA methylation in folate deficiency: use of CpG methylase. Biochem Biophys Res Commun 193:1184–1190
11. Fujiwara H, Ito M (2002) Nonisotopic cytosine extension assay: a highly sensitive method to evaluate CpG island methylation in the whole genome. Anal Biochem 307:386–389
12. Karimi M, Johansson S, Stach D, Corcoran M, Grander D, Schalling M et al (2006) LUMA (LUminometric Methylation Assay) – a high throughput method to the analysis of genomic DNA methylation. Exp Cell Res 312:1989–1995
13. Yang AS, Estecio MR, Doshi K, Kondo Y, Tajara EH, Issa JP (2004) A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. Nucleic Acids Res 32:e38
14. Irahara N, Nosho K, Baba Y, Shima K, Lindeman NI, Hazra A et al (2010) Precision of pyrosequencing assay to measure LINE-1 methylation in colon cancer, normal colonic mucosa, and peripheral blood cells. J Mol Diagn 12:177–183
15. Stanzer S, Balic M, Strutz J, Heitzer E, Obermair F, Hauser-Kronberger C et al (2010) Rapid and reliable detection of LINE-1 hypomethylation using high-resolution melting analysis. Clin Biochem 43:1443–1448
16. Weisenberger DJ, Campan M, Long TI, Kim M, Woods C, Fiala E et al (2005) Analysis of repetitive element DNA methylation by MethyLight. Nucleic Acids Res 33:6823–6836
17. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324:929–930
18. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324:930–935

19. Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M et al (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature 473:394–397
20. Laird PW (2010) Principles and challenges of genome-wide DNA methylation analysis. Nat Rev Genet 11:191–203
21. Estecio MR, Yan PS, Ibrahim AE, Tellez CS, Shen L, Huang TH et al (2007) High-throughput methylation profiling by MCA coupled to CpG island microarray. Genome Res 17:1529–1536
22. Smith ZD, Gu H, Bock C, Gnirke A, Meissner A (2009) High-throughput bisulfite sequencing in mammalian genomes. Methods 48:226–232
23. Wojdacz TK, Hansen LL, Dobrovic A (2008) A new approach to primer design for the control of PCR bias in methylation studies. BMC Res Notes 1:54
24. Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS (2010) Non-coding RNAs: regulators of disease. J Pathol 220:126–139
25. Thakur N, Tiwari VK, Thomassin H, Pandey RR, Kanduri M, Gondor A et al (2004) An antisense RNA regulates the bidirectional silencing property of the Kcnq1 imprinting control region. Mol Cell Biol 24:7855–7862
26. Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R et al (2000) A concise guide to cDNA microarray analysis. Biotechniques 29:548–550, 552–544, 556 passim
27. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45
28. Stahlberg A, Hakansson J, Xian X, Semb H, Kubista M (2004) Properties of the reverse transcription reaction in mRNA quantification. Clin Chem 50:509–515
29. Monk D, Arnaud P, Frost J, Hills FA, Stanier P, Feil R et al (2009) Reciprocal imprinting of human GRB10 in placental trophoblast and brain: evolutionary conservation of reversed allelic expression. Hum Mol Genet 18:3066–3074
30. Monk D, Sanches R, Arnaud P, Apostolidou S, Hills FA, Abu-Amero S et al (2006) Imprinting of IGF2 P0 transcript and novel alternatively spliced INS-IGF2 isoforms show differences between mouse and human. Hum Mol Genet 15:1259–1269
31. Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. Genet Sel Evol 34:275–305
32. Kuppuswamy MN, Hoffmann JW, Kasper CK, Spitzer SG, Groce SL, Bajaj SP (1991) Single nucleotide primer extension to detect genetic diseases: experimental application to hemophilia B (factor IX) and cystic fibrosis genes. Proc Natl Acad Sci USA 88:1143–1147
33. Lombardo AJ, Brown GB (1996) A quantitative and specific method for measuring transcript levels of highly homologous genes. Nucleic Acids Res 24:4812–4816
34. Singer-Sam J, LeBon JM, Dai A, Riggs AD (1992) A sensitive, quantitative assay for measurement of allele-specific transcripts differing by a single nucleotide. PCR Methods Appl 1:160–163
35. Uejima H, Lee MP, Cui H, Feinberg AP (2000) Hot-stop PCR: a simple and general assay for linear quantitation of allele ratios. Nat Genet 25:375–376
36. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55:611–622
37. Kouzarides T (2007) Chromatin modifications and their function. Cell 128:693–705
38. McEwen KR, Ferguson-Smith AC (2010) Distinguishing epigenetic marks of developmental and imprinting regulation. Epigenet Chromatin 3:2
39. Fatemi M, Pao MM, Jeong S, Gal-Yam EN, Egger G, Weisenberger DJ 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:e176
40. Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. Science 295:1306–1311
41. Dekker J (2006) The three 'C' s of chromosome conformation capture: controls, controls, controls. Nat Methods 3:17–21
42. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A et al (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326:289–293

Chapter 5
Biostatistical Methods in Epigenetic Epidemiology

E. Andrés Houseman

Contents

5.1 Introduction 58
5.2 General Statistical Considerations in Analyzing Epigenetic Data 61
5.3 Univariate or Locus-by-Locus Analysis 62
5.4 Unsupervised Analysis 63
5.5 Supervised Analysis 66
5.6 Special Considerations for the Analysis of Microarray Data 68
5.7 Multiple Comparisons and Study Objectives 70
5.8 Study Design 72
5.9 Useful Software 73
References 73

Abstract Epidemiologic studies involving epigenetics have, to date, focused mostly on DNA methylation, though studies investigating other epigenetic measures such as loss-of-imprinting and micro-RNAs are becoming increasingly popular. Analysis of data arising from such studies can be complicated by issues such as multiple comparison or lack of clarity about the goals of analysis. In this chapter, we provide an overview of key considerations in the analysis of such data sets, as well as a review of modern statistical techniques that have been used in the analysis of epigenetic data. We distinguish between exploratory studies and confirmation studies, providing guidance on statistical analysis in each case. In particular, we indicate when the control of familywise error rate or of false discovery rate is to be preferred. We provide an overview of unsupervised and supervised multivariate analysis, guidance on the analysis of microarray data, as well as guidance on study design considerations.

E.A. Houseman (✉)
College Of Public Health and Human Sciences, Oregon State University,
309 Waldo Hall, Corvallis, OR 97331-6406
e-mail: andres.houseman@oregonstate.edu

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_5,
© Springer Science+Business Media B.V. 2012

List of Abbreviations

ANOVA	analysis of variance
CART	classification and regression tree
ChIP-on-chip	chromatin immunoprecipitation with combined with microarray technology
CIMP	CpG Island Methylator Phenotype
COBRA	combined bisulfite restriction analysis
FA	factor analysis
FDR	false discovery rate
FWER	familywise error rate
GWAS	genome-wide association studies
HOPACH	hierarchical ordered partitioning and collapsing hybrid
LDA	linear discriminant analysis
LINE	long interspersed nuclear elements
LOI	loss of imprinting
LOOCV	leave-one-out-cross-validation
MCAR	missing completely at random
MeDIP	methyl-DNA immunoprecipitation
miRNA	micro-RNA
mRNA	messenger-RNA
MSP	methylation-specific PCR
PCA	principal components analysis
PCR	Polymerase chain reaction
RPMM	recursively partitioned mixture model
SAM	significance analysis of microarrays
SNP	single-nucleotide polymorphism
SVM	support vector machine

5.1 Introduction

Epidemiologic studies involving epigenetics have, to date, focused mostly on DNA methylation, though emerging work is beginning to address the role of histone modifications, microRNA (miRNA) expression, and loss of imprinting (LOI) in exploring disease causation. The primary reason for the focus on DNA methylation is that DNA methylation has been easier to measure than other epigenetic phenomena. Measurement of each of the many possible histone modifications requires a specific immunoprecipitation assay followed by microarray analysis or sequencing [1], and is thus expensive, especially when the number of combinations of possible histone marks is considered. In contrast, the measurement of DNA methylation requires only bisulfite modification followed by microarray or (pyro)sequencing, though sequencing is complicated by the problem of aligning modified DNA. So-called *global* DNA methylation is actually a measure of methylation of a particular repetitive element sequence, LINE1, a long interspersed nuclear element. Micro-RNA

expression is an emerging area of epigenetic research, requiring more or less the same biotechnology as messenger RNA (mRNA) expression, though normalization of miRNA arrays may require special considerations [2, 3]; aside from these, tools for the analysis of mRNA expression can be used for the analysis of miRNA microarrays. Finally, as discussed below, LOI data typically have statistical properties similar to DNA methylation data, aside from missing data considerations. Specifically, both DNA methylation and LOI assay data arise as percents or proportions, and their interpretations are roughly the same, so that statistical analysis of LOI data proceeds in a manner similar to that used for DNA methylation.

In the following chapter, we will focus primarily on the analysis of DNA methylation, with some discussion on variations applicable to the analysis of miRNAs and LOI. In addition, we will focus on single-locus assays and microarrays, since the cost of next-generation sequencing prevents its routine usage in most epidemiologic investigations to date. The assays and microarray platforms available for measuring DNA methylation of other sequences are numerous [4]; popular examples include Methylation-specific PCR (MSP) [5], COBRA [6], PCR followed by pyrosequencing [7], MethyLight [8], Sequenom [9], Illumina methylation arrays [10–12], MeDIP [13], and CHARM [69] . A more comprehensive list is provided in a recent review by Peter Laird [4]. The first five of these are examples of labor-intensive approaches that are most appropriate for candidate gene studies or validation of microarray results, i.e. studies involving a small number of loci, while the latter two admit high-throughput analysis. MSP produces a dichotomous value for each interrogated locus (gene), unmethylated (coded as 0) and methylated (coded as 1). COBRA, pyrosequencing, MethyLight, Sequenom, and Illumina methylation arrays produce interval scaled values between 0 and 1 (or 0 and 100) for each interrogated locus. MeDIP, on the other hand, produces fluorescence intensity values measured at regularly spaced intervals, typically with very dense coverage of the genome, and is typically reduced to peak intensities associated with p-values. MeDIP arrays are similar to the ChIP-on-chip arrays also manufactured by Roche to assay chromatin modification, and represent similar data complexity and expense. Therefore, they are uncommon in epidemiologic studies, and consequently their analysis is beyond the scope of this chapter. Note that because ChIP-on-chip assays are one of the few popular methods for studying chromatin modification, and due to the complexity and expense of the assays, chromatin modification is still relatively impractical to interrogate in a population-based study.

Most modern genetic and epigenetic studies involve the interrogation of multiple genes, CpG sites, or miRNA sequences; this multiplicity can lead to a number of test statistics, e.g. one for each interrogated locus, thus requiring a method for addressing multiple comparisons. The proper type of adjustment depends on the primary objectives of the study: whether the study is meant to be *hypothesis testing* (a “confirmation study”) or *hypothesis generating* (an “exploratory study”). The former requires rigorous control of Type I error probability (the probability of making any false claim) while the latter admits less stringent methods that control the *false discovery rate* (FDR, the rate of purported discoveries that cannot subsequently be reproduced). These two quantities are quite different, as will be discussed in detail below. A study can be designed to have multiple phases, wherein earlier phases are

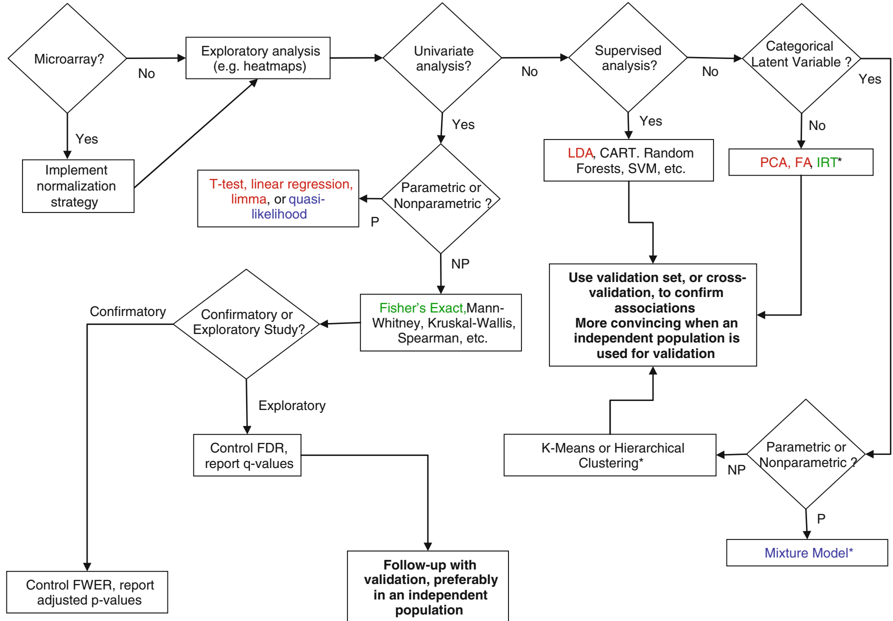


Fig. 5.1 Schematic diagram summarizing the principles and methods discussed

hypothesis generating and later phases are designed to test those hypotheses that arise from earlier phases. For the purposes of this chapter, different phases can be considered distinct studies. However, note that some studies are structured to confirm hypotheses in a population that is independent of the one used in the hypothesis-generating phase, also frequently employing a different assay, while other studies use the same population for both exploration and confirmation, using a different (usually more labor-intensive) assay for the latter. Use of independent populations is often more biologically convincing than the latter strategy, and has been adopted as the standard in genome-wide association studies (GWAS).

The remainder of this chapter is organized as follows. First, general statistical considerations in analyzing epigenetic data are presented. Next, an overview of statistical procedures is given, subdivided into three categories: *univariate* or *locus-by-locus* analysis, *unsupervised* multivariate analysis, and *supervised* multivariate analysis. These sections are followed by a discussion of special considerations for the analysis of microarray data. The subsequent section presents an overview of procedures for adjusting for multiple comparisons and their selection based on overall study objectives. The concluding sections briefly discuss study design considerations and software available for conducting analysis. Figure 5.1 presents a schematic diagram summarizing the principles and methods discussed in this chapter.

5.2 General Statistical Considerations in Analyzing Epigenetic Data

A variety of methods exist for analyzing genetic and epigenetic data, with the choice depending on the anticipated products of the analysis, as will be discussed below. However, a primary consideration is the distribution of the data to be analyzed. Crude versions of MSP produce dichotomized values (0 = unmethylated, 1 = methylated), while other platforms produce continuous values that can typically be scaled to lie between 0 and 1. In principle, some platforms, such as MethyLight, can produce values that are arbitrarily large, lying beyond 1 on any reasonable 0–1 scale, though such values can be considered problematic from the standpoint of interpretation. Thus, for MSP, care should be taken to use methods such as logistic regression that respect the dichotomous nature of the methylation measurement, while the others require techniques that respect the distribution properties of the methylation variable. An important point of consideration is the possible non-normality: because the methylation measurement can be influenced by a high fraction of methylated (or unmethylated) targets, their distributions can be bimodal with modes near 0 and 1, or else profoundly skewed. Consequently, methods with strong normality assumptions, such as t-tests, linear regression, and moderated versions of these techniques, should be avoided unless normality can first be confirmed with an appropriate distribution diagnostic. Nonparametric methods such as the Wilcoxon/Mann-Whitney and Kruskal-Wallis tests can often be used with great success, though they are difficult to extend to more complex regression models. When a need for a regression model arises, parametric approaches may be unavoidable, but several options are still available. One option is to model the methylation variable with a beta distribution rather than a Gaussian distribution, since the beta distribution can adequately capture the bimodality and skewness that is not uncommon with methylation data. This can be done explicitly with specialized maximum-likelihood procedures (e.g. [14]), or else employing quasi-likelihood (e.g. [15]). In either case, the following relation is enforced:

$$\mathrm{var}(Y) = E(Y)[1 - E(Y)]$$

for all methylation variables Y. This relation arises from the properties of a beta distribution, and is identical to the relation arising from a binomial distribution, with which the beta distribution shares many mathematical properties. Note that the beta distribution has been well-motivated for DNA methylation arrays on Illumina platforms [14, 16]. A somewhat simpler approach is to use an arcsine transformation $Y \rightarrow \arcsin(\sqrt{Y})$, which is the *variance-stabilizing transformation* for a beta distribution, i.e. the transformation that removes the dependence of the variance of any beta-distributed variable on its corresponding mean, thus making a Gaussian approximation more plausible; however, this approach is reasonable only when the methylation variables do not display marked bimodality. Another popular transform is the logit (e.g, [70]).

Micro-RNA expression data, generated using identical technology as that used for mRNA expression, will often present as lognormal variables, so that the log

expression values typically arising from these platforms can safely be analyzed as if they were Gaussian. However, it nevertheless is prudent to confirm distribution properties! LOI data typically present as bimodal interval-scaled data bounded between 0 and 1 [17], i.e. they are similar to the methylation assays discussed above. However state-of-the-art assays commonly used to measure LOI are limited by their inability to measure samples that are homozygous for the reporter SNP used to identify parent-of-origin [17]. Consequently, LOI data sets can have a substantial portion of missing data. Assuming that the genotype of the reporter SNP has no biological association with its corresponding (possibly unmeasured) LOI state, the missing data mechanism can be considered to be *missing completely at random* (MCAR), thus the missing values can be safely ignored in data analysis without biasing conclusions [18]. However, in the design of such studies, the anticipated proportion of data missing due to homozygosity must be considered in power calculations.

Choice of analysis is often driven by the objectives of the study. Baker and Kramer [19] prevent an overview of methods appropriate for each of various objectives in the context of microarray analysis, and Siegmund and Laird [20] provide an overview with a focus on DNA methylation. However, many of the same general principles apply to the analysis of smaller-scale studies, i.e. studies considering a small number of features (loci), such as candidate gene studies. These more focused studies will typically employ methods that are simpler, with fewer complications, than those used for the large-dimensional measurements typical of microarrays, but the overarching issues are similar. Broadly speaking, there are three possible approaches: *univariate* or *locus-by-locus* analysis, *unsupervised* multivariate techniques, and *supervised* multivariate techniques. In locus-by-locus techniques, analysis focuses on the univariate associations between methylation and phenotype ("differential expression" when the phenotype is dichotomous). These result in an association test statistic for each individual interrogated locus. Alternatively, using unsupervised multivariate methods, one can attempt to reduce higher dimensionality to lower dimensionality through the use of principal components analysis (PCA) or cluster analysis, obtaining estimates (or, more formally, *predictions*) of low-dimensional latent structure, subsequently assessing associations between phenotypes and these latent variables. Finally, using supervised multivariate techniques, one can directly model a phenotype on the entire multivariate realization of DNA methylation response. Many general variations of the latter two classes of analysis are presented in great depth in Hastie et al. [21]. In the sections below, we present an overview of each of these three classes of techniques.

5.3 Univariate or Locus-by-Locus Analysis

Univariate analysis is usually the most straightforward. In this type of analysis, each locus or feature is analyzed more-or-less separately from the others. For dichotomous methylation values Y such as those arising from MSP, Fisher's Exact test or its large-sample equivalent, the Pearson chi-square test, can be used to quantify an

association between methylation and a categorical phenotype Z (e.g. labels for two distinct groups), while a simple t-test or Mann-Whitney test can be used to quantify associations between Y and a continuous phenotype Z (e.g. age). Similarly, a Mann-Whitney or Kruskal-Wallis test can be used to quantify an association between continuous Y (e.g. % methylation) and categorical Z, while the Spearman correlation coefficient can be used to quantify an association between continuous Y and continuous Z. Note that these nonparametric tests are robust to violations of distributional and other modeling assumptions. In principle, t-tests, analysis of variance (ANOVA), and Pearson correlation could be used instead of their nonparametric analogs, but care should be taken to verify any distributional assumptions of both phenotype and methylation; the arcsine transformation can sometimes be used to transform methylation measures to variables that are more plausibly Gaussian.

When more than one locus is interrogated, it is essential to adjust for multiple comparisons. Broadly speaking, there are two classes of multiple comparison adjustments, those that control *familywise error rate*, or FWER, and those that control the *false discovery rate* [22, 23, 71], or FDR. Familywise error rate is akin to Type I error and represents the probability of any test resulting in significance under the assumption that all hypotheses are truly null; applicable multiple comparison procedures include many older, familiar methods such as Bonferroni-adjustment. On the other hand, FDR, which represents the fraction of significant tests for which the hypothesis is truly null, is increasingly being used to address multiple comparisons in the analysis of microarrays. We will discuss these options in more detail below.

5.4 Unsupervised Analysis

When analyzing data from microarrays, or even more focused candidate-gene studies where a moderate number of loci are being interrogated, it is sometimes advantageous to characterize the correlations among the assayed loci, both for pattern recognition and for dimensionality reduction (i.e. separating signal from noise with the objective of characterizing a large number of loci using a much smaller number of derived measures). The objective of unsupervised analysis is to reduce the dimensionality of a data set by seeking latent patterns or structure. Broadly speaking, unsupervised procedures can be sub-divided into those that seek continuous latent variables and those that seek categorical ones. The former type includes principal components analysis (PCA) and factor analysis (FA), while the latter includes mixture models and nonparametric clustering. A thorough treatment of these methods is beyond the scope of this text, but given extensive consideration in Haste et al. [21]. However, we briefly provide an overview. We assume that the methylation measurement is represented as a d-dimensional vector of continuous values $\mathbf{y}_i$, where i indexes subjects (or specimens) 1 through n.

The goal of PCA and FA is to represent a large number of continuous measures by a much smaller number continuous measures that capture the majority of the true signal present in the data. PCA and FA have been used in a number of recent methylation

studies (e.g., [24–26]). In PCA and FA, it is assumed that an F-dimensional vector $\mathbf{u}_i$ is related to $\mathbf{y}_i$ by the following linear equation:

$$\mathbf{y}_i = \Lambda \mathbf{u}_i + \varepsilon_i \quad (5.1)$$

where Λ is $d \times F$ matrix of parameters and ε_i is a d-dimensional vector of zero-mean uncorrelated errors (constraint 1). Usually the F *factors* or *components* of $\mathbf{u}_i$ are also assumed to be independent (constraint 2). In this model, $\mathbf{u}_i$ "explains" the correlation among the features of $\mathbf{y}_i$, because all such correlations must be induced by their common dependence on $\mathbf{u}_i$. For PCA, $d=n$ and ε_i is identically zero, in which case Λ can be obtained by concatenating rescaled eigenvectors of the matrix $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$. On the other hand, FA assumes $F << n$, in which case there are numerous procedures for obtaining Λ, including maximum likelihood [27, 28] and the principal factor method which employs an eigen-decomposition similar to that used for PCA [28]. Note that the matrix Λ can be right-multiplied any orthogonal $f \times f$ matrix $\mathbf{Q}$ (corresponding to a geometric rotation, which alters the interpretations of Λ) to obtain a new $d \times F$ matrix satisfying constraints 1 and 2 but with a different interpretation. When $F = n$ and all eigenvalues are unique, then $\mathbf{Q}$ is limited to the set of reflections and permutations (essentially maintaining original interpretations aside from sign considerations), so that PCA yields an essentially unique solution in applied settings. Recent work has demonstrated the utility of factor-analytic techniques to account for confounders [72]. However, when $F < n$, there are nontrivial choices of $\mathbf{Q}$, leading to the familiar "rotation problem" of factor analysis [28], where factor interpretation can be subjective. Nonlinear versions of equation (1) are also possible. For example, an item response theory (IRT) model, a form of FA appropriate for binary variables (i.e. instances where the entries of $\mathbf{y}_i$ are binary), has been used to model methylation [29, 30]. Alternatively, one can substitute a nonlinear relation for $\Lambda \mathbf{u}_i$ in (5.1) [31], although this approach has not appeared in any DNA methylation study to date. Once Λ and $\text{cov}(\varepsilon_i)$ have been estimated, the latent vectors $\mathbf{u}_i$ can subsequently be estimated (or, formally speaking, *predicted*), and associations between overall methylation (represented by the prediction $\hat{\mathbf{u}}_i$) and a phenotype can be assessed through common regression techniques. Note that the number F must be determined in advance; various methods [27, 28] exist, many of them based on the comparison of appropriate goodness-of-fit statistics, but are beyond the scope of this text.

One fundamental problem with PCA and FA is the difficulty in interpreting the factors $\hat{\mathbf{u}}_i$, especially when the dimension F is much greater than 1. As an alternative to methods employing continuous latent variables, one can instead use a method that infers a categorical latent variable. Such methods are commonly referred to as *clustering*. Parametric clustering is typically conducted using a *mixture model* [32], where each subject i is assumed to belong to one of K classes (where K is assumed fixed), and that conditional on membership in a particular class k, $\mathbf{y}_i$ is assumed to have a distribution f_k belonging to a known family, usually normal, lognormal [33] or beta [14]. In other words, the patterns within the data set are assumed to be driven by a fixed number of classes, within which each of the individual features comprising $\mathbf{y}$ is identically distributed (and usually assumed to be independent). As a nonparametric

alternative, one can use *K-means clustering*, where the K classes are determined only by their mean. Examples of the use of mixture models in the DNA methylation literature include [34–37, 73], while K-means clustering was used in Shen et al. [38].

The most familiar form of nonparametric clustering to most geneticists is *hierarchical clustering*. In hierarchical clustering, the data set of n subjects is repeatedly partitioned based on dissimilarity (*divisive* clustering), or the n subjects are agglomerated based on similarity (*agglomerative* clustering), yielding a *dendrogram*, or "tree", that conveys a sense of similarity by analogy with phylogenetics. The dendrogram can then be "cut" in such a way that K distinct classes are obtained, each class represented by a branch on the tree. Dissimilarity between two methylation vectors $\mathbf{y}_i$ and $\mathbf{y}_h$ is measured by *metric* $\langle \cdot,\cdot \rangle$. The most common choice is the Euclidean metric, which posits

$$\langle \mathbf{y}_i, \mathbf{y}_h \rangle = \|\mathbf{y}_i - \mathbf{y}_h\| = \left[\sum_{j=1}^{d} (y_{ij} - y_{hj})^2 \right]^{1/2},$$

(i.e. the conventional understanding of "distance" as the bird flies), but the Manhattan metric

$$\langle \mathbf{y}_i, \mathbf{y}_h \rangle = \sum_{j=1}^{d} \left| y_{ij} - y_{hj} \right|,$$

could also be used (with an analogy to distance provided by the number of square city blocks required to traverse). In order to conduct hierarchical clustering, dissimilarity between two *sets* of vectors ("linkage") must also be specified. Common choices are *complete linkage*, based on maximum distance between pairs of vectors, each from one of the two sets; *single linkage*, based on minimum distance, and *average linkage*, based on average distance. A variant of average linkage is *Ward's linkage* [39], which tends to produce compact, spherical clusters. There are also techniques based on medians [40]. Methods for selecting the optimal cut point, equivalent to selecting the number K of classes, include the *Hierarchical Ordered Partitioning and Collapsing Hybrid* [41], or *HOPACH*, and *Dynamic Tree Cutting* [42].

A common problem in clustering is the choice of K. Methods for choosing K are discussed extensively in Hastie et al. [21], but typically involve varying the choice of K and comparing an appropriate goodness-of-fit statistic for each separate fit. Thus, clustering can become computationally intensive when K is unknown and larger than 3. An alternative in the mixture model setting is *Recursively Partitioned Mixture Model* (RPMM) [14], which obtains the value of K simultaneously as the mixture model is fit, leading to greater computational efficiency. An additional benefit to RPMM is the arrangement of classes in a hierarchy similar to that obtained from hierarchical clustering. Mixture models in general and RPMM in particular have been shown empirically to perform better than non-parametric alternatives for DNA methylation data [14, 33, 43].

Note that the motivation for clustering is inherent in the idea of the *CpG Island Methylator Phenotype* [44–46], or CIMP, which was the explicit motivation for clustering in Shen et al. [38]. Thus, from an interpretation standpoint, clustering may represent an alternative that is superior to FA and PCA. Following the choice

of K and completion of the clustering procedure, subjects are assigned to clusters, either explicitly in nonparametric algorithms such as K-means and hierarchical clustering, or implicitly using empirical Bayes procedures subsequent to the fitting of mixture models. The latter allows "fuzzy" clustering, where a subject can be allocated fractionally to more than one of the K clusters. *Fanny* is another nonparametric clustering procedure that also allows fuzzy clustering [40]. Once study subjects have been assigned to clusters, associations with a phenotype can be assessed by common regression techniques where class membership is included as a categorical variable, or, more simply, by Fisher's Exact Test, t-test, Mann-Whitney test, ANOVA, or Kruskal-Wallis test.

In both types of unsupervised multivariate analysis, regression can be used to investigate the association between a phenotype and overall methylation, represented in reduced form by either a latent vector or a class membership. Simple testing procedures that compare distributions across categories can be used with cluster analysis, permitting exact inference through Fisher's Exact test or by permutation. Permutation tests are conducted by first choosing a test statistic T appropriate for the desired hypothesis, followed by the determination of a null distribution for the test statistic by shuffling the cluster assignments among the subjects and recalculating T, with the latter operation repeated a large but fixed number of times. For example, a test of association between a continuous phenotype (e.g. age) and K methylation classes is obtained by choosing T to be the ANOVA F-test statistic or the Kruskal-Wallis test statistic; recalculating T a large number of times, each time shuffling the K methylation classes with respect to their phenotype; and calculating the p-value as the proportion of shuffled data sets that produce values of T larger than the one calculated from the data in their original configuration. In contrast, PCA and FA obligate the use of a regression method, rendering the selection of an appropriate test statistic T, as well as its interpretation, slightly more difficult when the number of chosen factors, F, is greater than 1. Note that in either case, the testing of associations between a phenotype and the latent variables obtained by unsupervised analysis constitutes a form of analysis closely linked with the naïve Bayes procedure described below, and thus may require validation in an independent data set if the study objective is confirmation.

5.5 Supervised Analysis

In contrast to unsupervised multivariate analysis, the goal of supervised analysis in the context of DNA methylation is to predict a phenotype Z on the basis of a DNA methylation profile $\mathbf{y}$ obtained from a microarray or candidate gene study with moderate number of interrogated loci. This is sometimes called *classification*, with the understanding that classification of $\mathbf{y}$ is "supervised" by Z. Examples of supervised analysis in DNA methylation studies are numerous (e.g. [26, 35, 36]). In this context, there is usually a distinction made between a *training set* used to build a model and a *test* or *validation set* used to verify the properties of the trained model. (Note that

the formal determination of the relative size of the training and validation sets is complex, dependent on numerous factors such as the multivariate structure of **y** as well as the nature of the association between **y** and Z, though typically 1:1 or 2:1 training:test ratios are used.) To distinguish validation from training, for exposition, it often useful to imagine a single additional observation of paired phenotype and methylation $(Z^*, \mathbf{y}^*)$, which has been excluded from the data set used for modeling. In this section, we will assume that Z is continuous or Bernoulli to simplify notation, although the case of a polytomous Z is not much more difficult.

A simple example of a supervised analysis is the regression of Z on **Y**, which provides predictions of the form $E[Z^* \mid \mathbf{y}^*]$. Another simple example is *Naïve Bayes* [32, 47], where unsupervised analysis is first used to cluster **y** into K classes; with $C(\mathbf{y})$ denoting the class to which **y** belongs, $E[Z \mid C(\mathbf{y})]$ can easily be computed by regression or $2 \times K$ table, producing $E[Z^* \mid C(\mathbf{y}^*)]$ from the resulting fitted model. For dichotomous Z, *Linear Discriminant Analysis* (LDA) seeks a vector β of coefficients such that $\Pr(Z = 1 \mid \mathbf{y}_i^T \beta > 0)$ is maximized. On the surface, LDA seems similar to logistic regression, but common estimation techniques require more stringent normality assumptions on **y** than does logistic regression. A closely related procedure is the *Support Vector Machine* [21, 48], or SVM, which provides computational efficiency when the dimension d of **y** is large. *Kernel Machines* [49] relax the strong linearity assumptions of the SVM.

Another common procedure is the *Classification and Regression Tree* [50], or CART, where the data set is recursively partitioned in such a way as to maximize the distinctions between $E[Z \mid R]$ among the partitions R of the data set $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$. The output is a dendrogram, similar to that produced by hierarchical clustering, the branches of which represent the data set partitions R and their corresponding predictions of Z. Typically, each branch is determined by a single feature of **y** (i.e. a single DNA methylation measurement). In principle, one could partition all the way to single observations, though the utility of such a model is limited; consequently, there are various strategies employed for *pruning* the tree. As in the choice of K for clustering, these typically involve the comparison of model fit statistics (e.g. BIC). $E[Z^* \mid \mathbf{y}^*]$ is obtained by using $\mathbf{y}^*$ to traverse the tree to a desired depth, arriving at a prediction for Z^*. For large dimensions d, the enhancement provided by *Random Forests* [51] can be used. The Random Forest algorithm iterates through a large number of CART analyses, using a bootstrap (i.e., sampling the n subjects with replacement) for each individual analysis. In addition, for each node of each individual tree, only a subset of the dimensions of **y** are considered for partitioning, determined by randomly sampling (without replacement) a fixed number of components. For each iteration, both *variable importance* and *prediction error* are assessed by comparing $E[Z^* \mid \mathbf{y}^*]$ to Z^* in the "out-of-bag" sample (i.e. among subjects not selected in the bootstrap). Prediction error is simply the average of a reasonable comparison statistic ("loss" function or "risk function"), while variable importance is assessed by the change in loss before and after shuffling the variable in the out-of-bag sample.

The Random Forests algorithm reflects a general principle in supervised analysis: a need to assess the predictive performance of the trained model. If there is a natural

separation between the training and test sets (e.g. temporal or geographic separation of populations), then the comparison of $E[Z^* \mid \mathbf{y}^*]$ with Z^* among the subjects of the test set provides a convincing assessment of the performance of the classifier in question. On the other hand, if no such natural separation exists, a partitioning into training and test sets can be made randomly. Often, in this context, such partitioning is done repeatedly. Cross-validation techniques randomly partition the data set into several subsets, each of which functions as a test set for an individual analysis that leaves that subset out. An extreme form of cross validation is *Leave-One-Out-Cross-Validation* (LOOCV), where the partition consists of n singletons. Marsit et al. [44], provides an example of various supervised analyses, where 15 dichotomous methylation variables are used to classify solid tumors into one of four types.

One drawback of supervised analysis is that (except for Naïve Bayes), no insight is provided on the correlation of the features of $\mathbf{y}$. For DNA methylation studies, such insight may have biological value. Another drawback is the possibility, when d is large, of *over-fitting*, where spurious associations between $\mathbf{y}$ and Z are apparent only because of the large dimension of $\mathbf{y}$; this is a greater problem for microarrays than for candidate gene approaches. To some extent, validation strategies protect against this phenomenon, but validation is most biologically convincing when there is a natural separation between the training and test sets.

5.6 Special Considerations for the Analysis of Microarray Data

Almost all of the univariate and multivariate techniques described above are applicable to the statistical analysis of DNA methylation microarray and miRNA microarrays. However, there are some additional considerations that are specific to the analysis of these high-dimensional, high-throughput data. The first is normalization. There is now a vast literature on normalization of mRNA microarrays [52], some of which may apply to normalization of miRNA arrays, though there is some controversy on the subject [2, 3]. There is also some controversy on whether explicit normalization should be applied to Illumina DNA methylation arrays. Although Illumina makes normalization tools available in its software, the company is neutral on the position of whether they should be used. Some authors have elected to use ad-hoc methods based on bisulfite conversion efficiency [26], while others have elected not to normalize [34–37]. Recent work has focused on accurate normalization of DNA methylation microarrays [53, 70].

Assuming that decisions on normalization have been made, data exploration is often facilitated by the construction of a *clustering heatmap*, a color image of the data matrix $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$ where rows represent individual subjects, columns represent individual features, and the value of y_{ij} (methylation extent or miRNA expression) is indicated by color. Furthermore, hierarchical clustering is used to cluster the columns and cluster the rows, so that similar array profiles are grouped together and similarly expressing loci are grouped together, allowing for patterns to be recognized by eye. Frequently the dendrograms resulting from the hierarchical clustering are also shown.

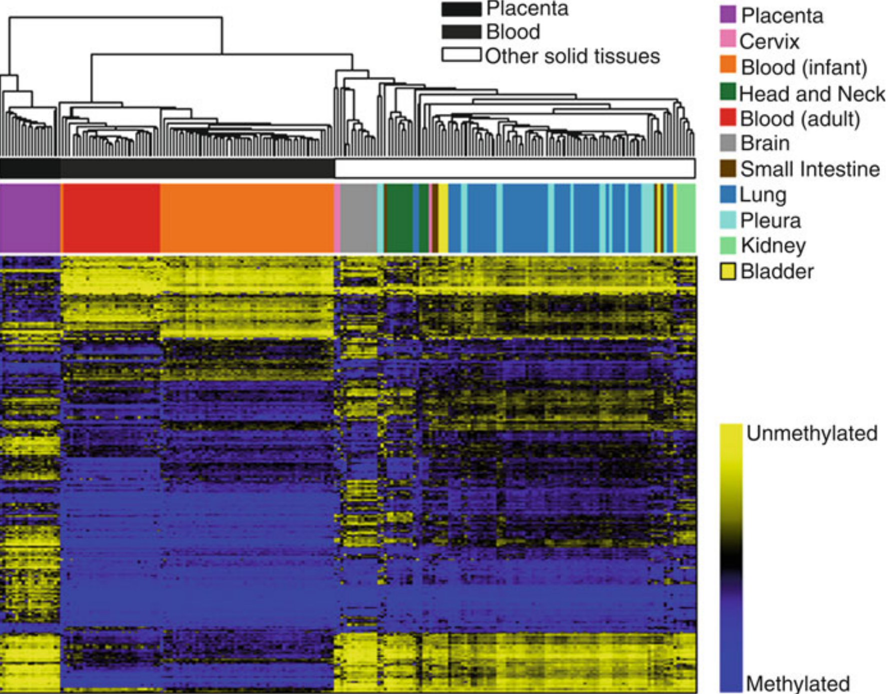


Fig. 5.2 Example of a Heatmap showing data from a DNA methylation microarray (Reproduced from [34])

Figure 5.2 shows an example of a heatmap constructed from DNA methylation data obtained from several different types of non-pathological tissues [34].

Locus-by-locus analysis is often conducted by *moderated* t-tests and ANOVA procedures, available through methods such as *Limma* [52]. These use a Bayesian formulation to motivate a procedure for estimating variance components, resulting in slightly more efficient tests. Because of the large number of loci tested, and the potential for substantial correlation between loci, a Bonferroni adjustment for multiple comparisons can be inefficient. Consequently, control of FDR is typically preferred over the control of FWER. However, the more liberal procedures based on FDR imply a subsequent confirmation or validation of a small number of loci, at least philosophically. This is often done with a follow-up study or study phase, often with a distinct assay of the more labor-intensive variety, with the goal of confirming the biochemical results as much as the epidemiologic associations. This issue is discussed in more detail below. When there is reasonable confidence in the microarray platform, an entire array can be used to test a formal hypothesis. When this occurs, data reduction strategies such as those used in Christensen et al. [34] may sometimes be employed. In this study, CpGs were clustered based on their methylation profile across specimens; subsequently, for each specimen, methylation measurements

were averaged over each cluster, resulting in a small number of average methylation values, each somewhat analogous to LINE1 methylation (a biological average of methylation of a specific repeat sequence). These averages were then tested for association with other variables such as age, with appropriately rigorous multiple comparisons adjustment based on FWER. A similar, more comprehensive study [74].

Microarrays imply a high dimension d of $\mathbf{y}$. Therefore, computationally efficient methods are needed for some of the techniques described in the previous section. For example, RPMM may be preferred over conventional mixture model techniques. In addition, when d is extremely large (e.g. the Infinium array produced by Illumina), preprocessing of the array may be needed, selecting a moderate number of loci to be used in subsequent analysis. One simple approach selects the M most variable loci (with a predetermined choice of M), though this method may inadvertently choose biochemically noisy loci that are not necessarily predictive of phenotype. Consequently, alternatives such as semi-supervised analysis [54, 55] can be used to select predictive loci. These methods employ cross-validation strategies to protect against over-fitting. A semi-supervised version of RPMM [43] has recently been demonstrated to have good performance.

A final consideration in the analysis of microarray data concerns the type of locus on which to focus analysis. A number of studies have found that the association between DNA methylation and phenotype may vary according to the local genomic geography of the target locus (e.g. CpG Island status) [34, 56]. There may be value in stratifying analysis based on such considerations.

5.7 Multiple Comparisons and Study Objectives

Univariate analysis, whether of the low or high dimensional variety, whether employing simple nonparametric tests of association or moderated linear models, will produce a number of test statistics (one for each feature), thus requiring an adjustment for multiple comparisons. The fundamental problem can be illustrated by a simple example. Suppose a study is designed to investigate the association of age with DNA methylation at 100 different CpG sites. Even if none of the 100 sites has a true association with age, approximately 5 of the 100 univariately generated p-values will nominally fall below 0.05, thus misleading the naïve researcher.

As mentioned above, the choice of adjustment depends on the goals of the study. Confirmation studies demand rigorous control of the familywise error rate, i.e. the overall probability of any false conclusion, because their objective is to test an overall hypothesis about a set of loci. FWER, akin to Type I error, i.e. Pr("significant"| truly null), represents the probability of any test resulting in significance under the assumption that *all* hypotheses are truly null. On the other hand, exploratory studies do not demand such rigor, since their design implies a subsequent confirmation step (whether or not it is actually conducted in the context of the study). Errors are therefore tolerable, and the objective is to control the *rate* of such errors. False

discovery rate, which represents the fraction of significant tests for which the hypothesis is truly null, is increasingly being used to address multiple comparisons in exploratory studies, especially those using microarrays. FDR is based on the probability that an individual locus, already having been deemed significant, represents a truly null association, i.e. Pr(truly null | "significant"). Note the difference in denominator between the formulation of FWER and FDR! Methods based on FDR are inappropriate when the intention is confirmation, i.e. the measurements represent a "final" data set that will not be followed up with a validation phase. Confirmation studies generate data with the purpose of providing definitive evidence for or against a hypothesis, and therefore do not represent an exploratory analysis with the goal of "discovery". Consequently, focus on FDR is inappropriate for such studies.

A great many classical methods for controlling FWER, such as those of Tukey/Kramer [57], Scheffé [58], and Duncan [59] are ultimately based on Gaussian statistics, and therefore are appropriate only for Gaussian outcomes or large sample sizes. The classical and straightforward method of Bonferroni and its more powerful variants [60, 61] are popular alternatives, although they have been widely criticized as extremely conservative (i.e. having actual FWER much lower than their nominal value). In the classical approach, the association measured at a single locus is deemed significant at level α if its nominal (unadjusted) p-value is less than α / M, where M is the number of loci being interrogated. A less conservative approach, proposed by Holm, first ranks the loci by p-value, assesses significance for the lowest p-value by comparing it to α / M; if the first test is rejected, the second-lowest p-value is assessed by comparing it to $\alpha / (M-1)$, etc., continuing until the first failure to reject. This approach is related to a testing procedure proposed by Simes [61], where the j'th ordered p-value is rejected if it is less than $j\alpha / M$, and is equivalent to a conservative approach based on the control of FDR, described below. When the methylation variables are independent of one other, Bonferroni adjustment typically represents "the best one can do", in that its true FWER is close to the nominal value α. However, when the methylation values are substantially correlated, as is typical, Bonferroni procedures may be overly conservative and therefore inefficient. In these settings, permutation-based adjustments offer a remedy. These are beyond the scope of the current chapter, but are covered extensively in Westfall and Young [62]. An alternative to permutation-based adjustments is the use of unsupervised techniques to reduce dimensionality before assessing an association with phenotype.

Control of FDR typically utilizes the methodology of *q-values*. Nominal p-values arising from individual tests are converted to their corresponding q-values, which have the following interpretation: for a given threshold q_0, the set of $M(q_0)$ features having q-value less than q_0 (the "discoveries") will contain approximately $q_0 M(q_0)$ features whose associations are in fact null. The practical interpretation is that q_0 represents the proportion of loci that will fail a subsequent attempt to independently validate the $M(q_0)$ most significantly associated loci. Note that, via Bayes' Rule,

$$\Pr(null \mid significant) = \frac{\Pr(significant \mid null)\Pr(null)}{\Pr(significant)} = p_0 \frac{\Pr(null)}{\Pr(significant)},$$

the latter equality justified with an understanding that a fixed p-value threshold p_0 determines significance, and therefore Pr(*significant*) is obtained empirically from the set of generated p-values as the fraction of significant results at threshold p_0. Consequently, a critical quantity in the construction of q-values is the proportion $\pi_0 = \Pr(null)$ of truly null loci on the entire array, and various methods have been proposed for its estimation [23, 63]. Note that Simes' modification of Bonferroni adjustment is equivalent to the construction of q-values with $\pi_0 = 1$, i.e. no truly associated loci. It is important to note that q-values do not provide a method for testing a formal hypothesis in the way that p-values have been used traditionally. Their use is more appropriate when the objective is to select a small number of loci for subsequent validation, where p-values appropriately adjusted for multiple comparisons and control of FWER can be used more efficiently to test a scientific hypothesis. Traditionally, q-values between 0.05 and 0.15 have been considered "significant", though the final determination of significance rests on the subsequent validation, and the selected q-value threshold should really reflect the investigator's tolerance for failure in validation.

A final note is in order on the use of microarrays for confirmatory studies. The analysis presented in Christensen et al. [34] used a multiple comparisons adjustment based on FWER, but was reasonably powered with a moderate sample size because of its data reduction strategy: for each subject, methylation measurements were averaged over CpGs clustered into a small set of classes. This strategy is more efficient than testing each CpG separately with subsequent adjustment by control of FWER. For example, suppose that in a particular study, M CpG sites have similar DNA methylation distributions and are similarly associated with a phenotype Z; then, for each individual CpG, the test statistic for its association with Z will have standard error $\sqrt{M}$ times the standard error of the test statistic for association with the average over the M CpG sites. However, both statistics will have similar effect size. If the individual statistics are not strongly dependent, there will be an additional cost in efficiency after adjusting for multiple comparisons. For example, in the limiting case of independent test statistics, adequate correction for multiple comparisons would effectively require, for each CpG, a critical value corresponding to the nominal level α / M. However, we note that the use of externally derived clusters might be more convincing than the data-driven clustering used in this study, though the latter is no worse than the typical two-stage analyses used with unsupervised analysis and subsequent testing of associations between a phenotype and predicted latent variables. Recent work has explored this topic in detail [74]

5.8 Study Design

In addition to proper microarray design considerations, such as randomizing chip arrangement with respect to target phenotype as much as possible [64, 65], careful attention should be paid to epidemiologic study design. Again, an important consideration is whether the study is designed for exploration or for confirmation. If the data are to be used for exploration, control of FDR is the key consideration, and methods such as *Significance Analysis of Microarrays* [66], or SAM, *PowerAtlas*

[67], or the methods of Liu and Hwang [68] should be used. If the objective is to confirm associations for M specific loci, then power and sample size considerations can be addressed conservatively by assuming Bonferroni correction, i.e. using standard methods for univariate analysis with control of Type I error set to α / M, where α is the desired overall FWER.

For the more complex multivariate models, design considerations can be difficult both analytically and computationally, and may require ad-hoc methods such as simulation. For such models, it may be useful to pose the problem in a simpler framework in order to approximate power. For example, if the design of a study where classification is a goal, two classes with specific prevalences may be assumed, and the problem thus converted to a conventional two-sample comparison.

5.9 Useful Software

The R Project for Statistical Computing (www.R-project.org) offers open-source statistical software with user-updatable libraries, called *packages*. In addition, it serves as the foundation for the *Bioconductor* suite of bioinformatic tools (www.bioconductor.org). Many of the procedures mentioned in this chapter are available through R or Bioconductor packages. The R functions *fisher.test*, *wilcox.test*, and *kruskal.test* allow for nonparametric testing that can be easily programmed in a loop. The R function *glm* can be used to fit quasi-likelihood models appropriate for beta distributions. The R function *hclust* and package *cluster* offer tools for hierarchical clustering, and the function *heatmap* constructs heatmaps. The Bioconductor package *limma* fits moderated versions of linear models. The R packes *qvalue* calculates q-values for controlling FDR. The packages *mclust*, and *RPMM* fit mixture models, *tree* and *rpart* fit CART models, *kernlab* implements SVMs and kernel machines, and *randomForests* implements Random Forests. The R package *samr* implements SAM power analysis as well as many of the univariate test options presented above.

SAS has fewer tools available specifically for the analysis of high-dimensional data arising from microarrays. However, some of the simpler procedures described above are available. Univariate tests can be conducted using the *ttest* and *npar1way* procedures (the latter supporting nonparametric tests). Regression can be conducted using the *reg*, *glm*, *mixed*, or *genmod* procedures; note, however, that the *glm* procedure is not analogous to the *glm* function in R, which is more similar to the *genmod* procedure. Some more complex procedures are also available; e.g. unsupervised multivariate analysis is supported by the *factor* and *cluster* procedures. SAS macros for some more sophisticated procedures can be found by searching online.

References

1. Park PJ (2009) ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet 10:669–680

2. Pradervand SJ, Weber J, Thomas J, Bueno M, Wirapati P, Lefort K et al (2009) Impact of normalization on miRNA microarray expression profiling. RNA 15:493–501
3. Tricoli JV, Jacobson JW (2007) MicroRNA: potential for cancer detection, diagnosis, and prognosis. Cancer Res 67:4553–4555
4. Laird P (2010) Principles and challenges of genome-wide DNA methylation analysis. Nat Rev Genet 11:191–203
5. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB (1996) Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci USA 93:9821–9826
6. Xiong Z, Laird PW (1997) COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res 25:2532–2534
7. Tost J, El abdalaoui H, Gut IG (2006) Serial pyrosequencing for quantitative DNA methylation analysis. Biotechniques 40:721–722, 724, 726
8. Eads CA, Danenberg KD, Kawakami K, Saltz LB, Blake C, Shibata D et al (2000) MethyLight: a high-throughput assay to measure DNA methylation. Nucleic Acids Res 28:E32
9. Jurinke C, Denissenko MF, Oeth P, Ehrich M, van den Boom D, Cantor CR (2005) A single nucleotide polymorphism based approach for the identification and characterization of gene expression modulation using MassARRAY. Mutat Res 573:83–95
10. Bibikova M, Fan JB (2009) GoldenGate assay for DNA methylation profiling. Methods Mol Biol 507:149–163
11. Bibikova M, Le J, Barnes B, Saedinia-Melnyk S, Zhou L, Shen R et al (2009) Genome-wide DNA methylation profiling using Infinium assay. Epigenomics 1:177–200
12. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B et al (2006) High-throughput DNA methylation profiling using universal bead arrays. Genome Res 16:383–393
13. Shen Y, Fouse SD, Fan G (2009) Genome-wide DNA methylation profiling: the mDIP-chip technology. Methods Mol Biol 568:203–216
14. Houseman EA, Christensen BC, Marsit CJ, Karagas MR, Wrensch MR, Yeh RF et al (2008) Model-based clustering of DNA methylation array data: a recursive-partitioning algorithm for high-dimensional data arising as a mixture of beta distributions. BMC Bioinformatics 9:365
15. Hsiung DT, Marsit CJ, Houseman EA, Eddy K, Furniss CS, McClean MD et al (2007) Global DNA methylation level in whole blood as a biomarker in head and neck squamous cell carcinoma. Cancer Epidemiol Biomarkers Prev 16:108–114
16. Houseman EA, Christensen BC, Karagas MR, Wrensch MR, Nelson HH, Wiemels JL et al (2009) Copy number variation has little impact on bead-array-based measures of DNA methylation. Bioinformatics 25:1999–2005
17. Lambertini L, Diplas AI, Lee MJ, Sperling R, Chen J, Wetmur J (2008) A sensitive functional assay reveals frequent loss of genomic imprinting in human placenta. Epigenetics 3:261–269
18. Little RJA, Rubin DB (2002) Statistical analysis with missing data, 2nd edn. Wiley, Hoboken
19. Baker SG, Kramer BS (2008) Using microarrays to study the microenvironment in tumor biology: the crucial role of statistics. Semin Cancer Biol 18:305–310
20. Siegmund KD, Laird PW (2002) Analysis of complex methylation data. Methods 27:170–178
21. Hastie T, Tibshirani R, Friedman J (2001) The elements of statistical learning: data mining, inference, and prediction. Springer, New York
22. Storey J, Tibshirani R (2003) Statistical significance for genome-wide experiments. Proc Natl Acad Sci USA 100:9440–9445
23. Storey JD (2002) A direct approach to false discovery rates. J R Stat Soc Ser B 64:479–498
24. Davidsson J, Lilljebjorn H, Andersson A, Veerla S, Heldrup J, Behrendtz M et al (2009) The DNA methylome of pediatric acute lymphoblastic leukemia. Hum Mol Genet 18:4054–4065
25. Sun YV, Turner ST, Smith JA, Hammond PI, Lazarus A, Van De Rostyne JL et al (2010) Comparison of the DNA methylation profiles of human peripheral blood cells and transformed B-lymphocytes. Hum Genet 127:651–658
26. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Gayther SA, Apostolidou S et al (2009) An epigenetic signature in peripheral blood predicts active ovarian cancer. PLoS One 4:e8274

27. Lawley DN, Maxwell AE (1971) Factor analysis as a statistical method, 2nd edn. Butterworths, London
28. Bartholomew DJ, Knott M (1999) Latent variable models and factor analysis, 2nd edn. Hodder Arnold, London
29. Houseman EA, Marsit C, Karagas M, Ryan LM (2007) Penalized item response theory models: application to epigenetic alterations in bladder cancer. Biometrics 63:1269–1277
30. Marsit CJ, Houseman EA, Schned AR, Karagas MR, Kelsey KT (2007) Promoter hypermethylation is associated with current smoking, age, gender and survival in bladder cancer. Carcinogenesis 28:1745–1751
31. Hastie T, Stuetzle W (1989) Principal Curves. J Am Stat Assoc 84:502–516
32. Fraley C, Raftery AE (2002) Model-based clustering, discriminant analysis, and density estimation. J Am Stat Assoc 97:611–631
33. Siegmund KD, Laird PW, Laird-Offringa IA (2004) A comparison of cluster analysis methods using DNA methylation data. Bioinformatics 20:1896–1904
34. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL et al (2009) Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet 5:e1000602
35. Christensen BC, Kelsey KT, Zheng S, Houseman EA, Marsit CJ, Wrensch MR et al (2010) Breast cancer DNA methylation profiles are associated with tumor size and alcohol and folate intake. PLoS Genet 6:e1001043
36. Christensen BC, Marsit CJ, Houseman EA, Godleski JJ, Longacker JL, Zheng S et al (2009) Differentiation of lung adenocarcinoma, pleural mesothelioma, and nonmalignant pulmonary tissues using DNA methylation profiles. Cancer Res 69:6315–6321
37. Marsit CJ, Christensen BC, Houseman EA, Karagas MR, Wrensch MR, Yeh RF et al (2009) Epigenetic profiling reveals etiologically distinct patterns of DNA methylation in head and neck squamous cell carcinoma. Carcinogenesis 30:416–422
38. Shen L, Toyota M, Kondo Y, Lin E, Zhang L, Guo Y, Hernandez NS, Chen X, Ahmed S, Konishi K, Hamilton S R, Issa JP (2007). Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer, PNAS 104:18654–18659
39. Ward JH (1963) Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236–244
40. Kaufman L, Rousseeuw PJ (1990) Finding groups in data: an introduction to cluster analysis. Wiley, New York
41. van der Laan MJ, Pollard KS (2003) A new algorithm for hybrid hierarchical clustering with visualization and the bootstrap. J Stat Plan Infer 117:275–303
42. Langfelder P, Zhang B, Horvath S (2008) Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut library for R. Bioinformatics 24:719–720
43. Koestler DC, Marsit CJ, Christensen BC, Karagas MR, Bueno R, Sugarbaker DJ et al (2010) Semi-supervised recursively partitioned mixture models for identifying cancer subtypes. Bioinformatics. doi:10.1093/bioinformatics/btq470
44. Marsit CJ, Houseman EA, Christensen BC, Eddy K, Bueno R, Sugarbaker DJ et al (2006) Examination of a CpG island methylator phenotype and implications of methylation profiles in solid tumors. Cancer Res 66:10621–10629
45. Issa JP (2004) CpG island methylator phenotype in cancer. Nat Rev Cancer 4:988–993
46. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP (1999) CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci USA 96:8681–8686
47. Domingos P, Pazzani M (1997) On the optimality of the simple Bayesian classifier under zero–one loss. Mach Lear 29:103–137
48. Vapnik V, Kotz S (2006) Estimation of dependences based on empirical data. Springer, New York
49. Liu D, Lin X, Ghosh D (2007) Semiparametric regression of multidimensional genetic pathway data: Least-squares kernel machines and linear mixed models. Biometrics 63:1079–1088
50. Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Chapman & Hall, Boca Raton

51. Breiman L (2001) Random forests. Mach Learn 45:5–32
52. Smyth GK, Yang Y-H, Speed TP (2003) Statistical issues in microarray data analysis. Methods Mol Biol 224:111–136
53. Aryee MJ, Wu Z, Ladd-Acosta C, Herb B, Feinberg AP, Yegnasubramanian S et al (2011) Accurate genome-scale percentage DNA methylation estimates from microarray data. Biostatistics 12: 197-210.
54. Bair E, Hastie T, Paul D, Tibshirani R (2006) Prediction by supervised principal components. J Am Stat Assoc 101:119–137
55. Bair E, Tibshirani R (2004) Semi-supervised methods to predict patient survival from gene expression data. PLoS Biol 2:1544–9173
56. Doi A, Park IH, Wen B, Murakami P, Aryee MJ, Irizarry R et al (2009) Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nat Genet 41:1350–1353
57. Kramer CY (1956) Extension of multiple range tests to group means with unequal numbers of replications. Biometrics 12:307–310
58. Scheffé H (1953) A method for judging all contrasts in the analysis of variance. Biometrika 40:87–104
59. Duncan DB (1975) Tests and intervals for comparisons suggested by the data. Biometrics 31:339–359
60. Holm S (1979) A simple sequentially rejective multiple test procedure. Scan J Stat 6:65–70
61. Simes RJ (1986) An improved Bonferroni procedure for multiple tests of significance. Biometrika 73:751–754
62. Westfall PH, Young SS (1993) Resampling-based multiple testing: examples and methods for p-value adjustment. Wiley, New York
63. Storey JD (2003) The positive false discovery rate: a Bayesian interpretation and the q-value. Ann Stat 31:2013–2035
64. Wang S-C, Petronis A (2008) DNA methylation microarrays: experimental design and statistical analysis. Chapman & Hall, Boca Raton
65. Simon RM, Korn EL, Simon RM, Korn EL, McShane LM, Radmacher MD et al (2004) Design and analysis of DNA microarray investigations. Springer, New York
66. Tusher V, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 98:5116–5121
67. Page GP, Edwards JW, Gadbury GL, Yelisetti P, Wang J, Trivedi P et al (2006) The PowerAtlas: a power and sample size atlas for microarray experimental design and research. BMC Bioinformatics 7:84
68. Liu P, Hwang JT (2007) Quick calculation for sample size while controlling false discovery rate with application to microarray analysis. Bioinformatics 23:739–746
69. Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddeloh JA, Wen B, Feinberg AP (2008). Comprehensive high-throughput arrays for relative methylation (CHARM). Genome Res. 18:780–790
70. Kuan PF, Wang S, Zhou X, and Chu H (2010). A statistical framework for Illumina DNA methylation arrays. Bioinformatics 26:2849–2855
71. Benjamini Y, Hochberg Y, (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc Ser B 57:289–300
72. Teschendorff AE, Zhuang J, Widschwendter M (2011). Independent surrogate variable analysis to deconvolve confounding factors in large-scale microarray profiling studies. Bioinformatics 27:1496–1505
73. Hinoue T, Weisenberger DJ, Lange CPE, Byun HM, Van Den Berg D, Malik S, Pan F, Noushmehr H, van Dijk CM, Tollenaar RA, Laird PW (2011). Genome-scale analysis of aberrant DNA methylation in colorectal cancer. Genome Res. in press
74. Langevin SM, Houseman EA, Christensen BC, Wiencke JK, Nelson HH, Karagas MR, Marsit CJ, Kelsey KT (2011). The influence of aging, environmental exposures and local sequence features on the variation of DNA methylation in blood. Epigenetics 6:908–19

Chapter 6
Epigenome Changes During Development

Gavin Kelsey

Contents

6.1 Introduction 79
6.2 Epigenetic Reprogramming Events in Mammalian Germ Cells 80
    6.2.1 Germ Cell Specification and Erasure of the Somatic Epigenome 81
    6.2.2 Establishment of Gamete-Specific Epigenomes 82
6.3 Epigenetic Reprogramming Events in Mammalian Embryos 86
    6.3.1 Reprogramming of Paternal and Maternal Genomes After Fertilisation 86
    6.3.2 Establishment of New Epigenomes with the First Lineage Specification Events 88
6.4 Tissue-Specific Epigenomes: Plasticity and Stability 89
    6.4.1 Epigenome Changes During Differentiation 89
    6.4.2 Longevity of Epigenetic Marks 92
6.5 Transgenerational Epigenetic Effects 94
References 97

Abstract Epigenetic modifications accomplish the functional compartmentalisation of our genome. They ensure a high level of compaction of our DNA in a manner that nonetheless allows genes vital to given cell types to be expressed appropriately whilst sequestering away silent genes. The stability of epigenetic modifications provides long term memory in phenomena such as X-chromosome inactivation in females and genomic imprinting, but epigenetic states must also be dynamic as they are intimately involved in establishing the gene expression programmes that define cell lineage and are required to register changes in the environment. In this chapter, I shall describe the major epigenomic events that occur during mammalian development,

G. Kelsey (✉)
Epigenetics Programme, The Babraham Institute, Babraham Research Campus, Cambridge CB22 3AT, UK

Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EG, UK
e-mail: gavin.kelsey@babraham.ac.uk

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_6,
© Springer Science+Business Media B.V. 2012

from the specification of germ cells, to how the epigenome differences of the gametes are resolved at fertilisation, and how epigenomic events contribute to and reinforce lineage determination events. The advent of genome-wide profiling technologies is providing us with an unprecedented opportunity to investigate the scale of epigenomic changes during development and differentiation and how epigenomes are altered in disease.

Abbreviations

A^{vy}	*agouti viable yellow* allele
Bi-Seq	bisulphite whole genome sequencing
BMP4	bone morphogenic protein 4
c-Kit	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
CARM1	co-activator associated arginine methyltransferase 1
Cdx2	caudal type homeobox 2
Cfp1	CXXC finger protein 1
ChIP-Seq	chromatin immunopreciptation combined with next generation sequencing
Dnmt1	DNA methyltransferase 1
Dnmt3a	DNA methyltransferase 3a
Dnmt3b	DNA methyltransferase 3b
Dnmt3L	DNA methyltransferase 3-like
DMR	differentially methylated region (of an imprinted gene)
EHMT2	euchromatic histone methyltransferase 2
ELF5	ETS-related family transcription factor 5
EOMES	eomesodermin
ES	embryonic stem
Ezh2	Enhancer of zeste homologue 2
GCs	germ cells
H1T2	histone 1 variant T2
H1LS1	histone 1 variant LS1
H2	histone H2
H2A.Z	histone 2a variant Z
H2AR3me2	di-methylated H2A arginine 3
H3	histone H3
H3.3	histone 3 variant 3
H3K4	H3 lysine 4
H3K4me1/2/3	mono-/di- or tri-methylated H3 lysine 4
H3K9ac	acetylated H3 lysine 9
H3K9me1/2/3	mono-/di- or tri-methylated H3 lysine 9
H3K18ac	acetylated H3 lysine 18
H3K27me3	tri-methylated H3 lysine 37
H3K36me3	tri-methylated H3 lysine 36

H4K20me3	tri-methylated H4 lysine 20
H4	histone H4
H4R3me2	di-methylated H4 arginine 3
IAP	intracisternal A particle
ICM	inner cell mass
ICR	imprinting control region
Igf2	insulin-like growth factor 2 gene
iPSCs	induced pluripotential stem cells
KAP1	KRAB (Krüppel-associated box)-associated protein 1
KDM1B	lysine (K)-specific demethyase 1B
KDM2A	lysine (K)-specific demethyase 2A
Klf2	Krüppel-like factor 2
LTR	long terminal repeat
MeDIP-chip	methylcytosine immunoprecipitation combined with microarray hybridisation
NIH	National Institutes of Health
NLRP2	NLR family Pyrin domain containing protein 2
NLRP7	NLR family Pyrin domain containing protein 7
PcG	Polycomb group
PGCs	primordial germ cells
Pparα	peroxisome proliferator-activated receptor alpha
PRC2	Polycomb group (PcG) repressor complex 2
Prdm1(Blimp1)	PR domain containing 1
PRDM9	PR domain containing protein 9
Prdm14	PR domain containing protein 14
Prmt5	protein arginine methyltransferase 5
RdDM	RNA-directed DNA methylation
Sox2	SRY (sex determining region Y)-box 2
Suv39h1/2	suppressor of variegation 3-9 homolog 1 & 2
TE	trophoectoderm
Tet1	ten-eleven translocation 5mc-hydrolase 1
Tet3	ten-eleven translocation 5mc-hydrolase 3
TNP1 & 2	transition protein 1 & 2
Uhrf1	ubiquitin-like containing PHD and RING finger domains 1
ZFP57	zinc-finger protein 57
5mC	5-methylcytosine
5hmC	5-hydroxymethylcytosine

6.1 Introduction

Cellular phenotype is the outcome of the gene expression programme of a cell, which may be defined by a combination of the transcription factor repertoire, signalling events and the developmental history of the cell. The memory of earlier

events that impact on cell identity is epigenetic in nature. At a molecular level, epigenetics refers to the range of chromatin modifications, including DNA methylation, post-translation modifications of histones, the stable association of some non-histone chromatin proteins and higher levels of organisation of chromatin and the chromosomes. Collectively, these have become known as the epigenome. During mammalian development, the epigenome experiences a number of major reprogramming events that have the effect of establishing or restricting totipotency and contributing to and consolidating lineage specification decisions. These reprogramming events are best understood in the mouse (and much of the following description refers to what is known in mouse), but it is assumed that similar events occur in all placental mammals, including human. The key reprogramming events occur during germ cell specification and differentiation, in the embryo immediately after fertilisation, and in the preimplantation embryo as the first lineage specification occurs [1–3]. Subsequently, it is assumed that the epigenomic profiles of the various lineages and cell-types are progressively established throughout embryogenesis, as cell lineages arise and undergo terminal differentiation; each cell type will have its own epigenome, which may be further modulated by extrinsic factors. We also expect a lot of the epigenome: that it retains a memory of earlier development decisions *and* that it can respond appropriately to extrinsic signals and adapt to environmental changes. In truth, the epigenome will probably comprise 'hard-wired' regions, which might include imprinted genes that permanently remember their parental origin throughout life [4], and 'plastic' regions, comprising developmentally regulated genes and sequences exhibiting nutritional, environmental or experience-based adaptations. It has also become apparent that that some sequences escape full reprogramming in the germline, leading to the possibility of transgenerational epigenetic inheritance. We are now enjoying an unprecedented ability to map the repertoire of epigenetic features across the whole genome, either on high resolution microarray platforms or, increasingly, by next generation sequencing based approaches [5, 6]. In this chapter, I shall consider how the epigenome changes during development in mammals to provide a framework for understanding its vulnerability to factors that can cause disease.

6.2 Epigenetic Reprogramming Events in Mammalian Germ Cells

To consider the epigenomic journey experienced by the organism during normal development, let us start with the germline. In mammals, germ cells arise in the gastrulating embryo from a cohort of cells that has initiated a somatic cell programme. Germ cells will undergone two major epigenetic reprogramming events before mature gametes are produced. The first of these is to reverse the epigenotype of the embryonic cells from which the germ cells arise in order to restore a pluripotent state similar to the cells of the inner cell mass (ICM) of the blastocyst. The second is during the differentiation and maturation events that result in production

of the distinctive male and female gametes. Knowing the extent of DNA methylation and other epigenetic modifications in the oocyte and sperm genomes and the processes by which they are established is essential for a full understanding of the possibilities of transgenerational inheritance of epigenetic states and their possible associated phenotypic (mal)adaptations.

6.2.1 Germ Cell Specification and Erasure of the Somatic Epigenome

In the mouse, primordial germ cells (PGCs) are specified from cells destined to form embryonic and extra-embryonic mesoderm situated in the proximal epiblast of the egg cylinder stage embryo at embryonic day (E) 6.0–6.5 [7]. A small number (~40) of cells becomes specified, in part in response to BMP4 signalling, and are visible by prominent expression of the protein Fragilis; PGCs themselves will develop from a subpopulation of this initial pool and are distinguished by expression of the pluripotency factor Stella [8]. A number of epigenetic modifiers play dominant roles in PGC specification, including the PR/SET domain zinc-finger protein Prdm1 (also known as Blimp1) [9], which forms a repressive complex with the arginine methyltransferase Prmt5 that methylates histones H2a and H4 on the arginine 3 residue [10]. As well as suppressing the ongoing somatic cell programme, Prdm1 initiates the germ cell specification programme, including induction of expression of the pluripotency factors Nanog, c-Kit, Sox2 and Klf2. A second PR domain zinc-finger protein Prdm14 is induced independently of and cooperates with Prdm1 [11]. Amongst the factors Prdm1 and Prdm14 repress are Glp, which cooperates with the H3K9 methyltransferase EHMT2 (G9a), the *de novo* DNA methyltransferases Dnmt3a and Dnmt3b, and Uhrf1, which is required to target the maintenance methyltransferase Dnmt1 to replicating DNA. Ezh2, a component of the Polycomb group (PcG) repressor complex 2 (PRC2), is also up-regulated. This combination of changes results in PGCs exhibiting reduced levels of H3K9me2, but a compensatory increase in the repressive mark H3K27me3 [12, 13], and having a DNA demethylating environment. PGCs can be distinguished from surrounding somatic cells at this stage on the basis of their distinctive chromatin state by immunofluorescence for appropriate histone modifications.

From their site of specification in the proximal epiblast, PGCs migrate towards the genital ridges, which develop into the embryonic gonads, arriving around E10.5. During migration, PGCs proliferate such that they number ~8,000 (per embryo) by the time they reach the genital ridges. They continue to proliferate until ~ E13.5, when they enter mitotic arrest (in the male germline) or meiotic arrest (in the female germline), by which time each gonad contains ~25,000 germ cells (GCs). From E11.5, post-migratory PGCs undergo further, dramatic epigenomic changes. The most intriguing of these is global demethylation of DNA (Fig. 6.1a), which has been observed for repetitive sequences and some single copy genes, including imprinted genes [15, 16]. Because DNA demethylation has been observed to occur in a very

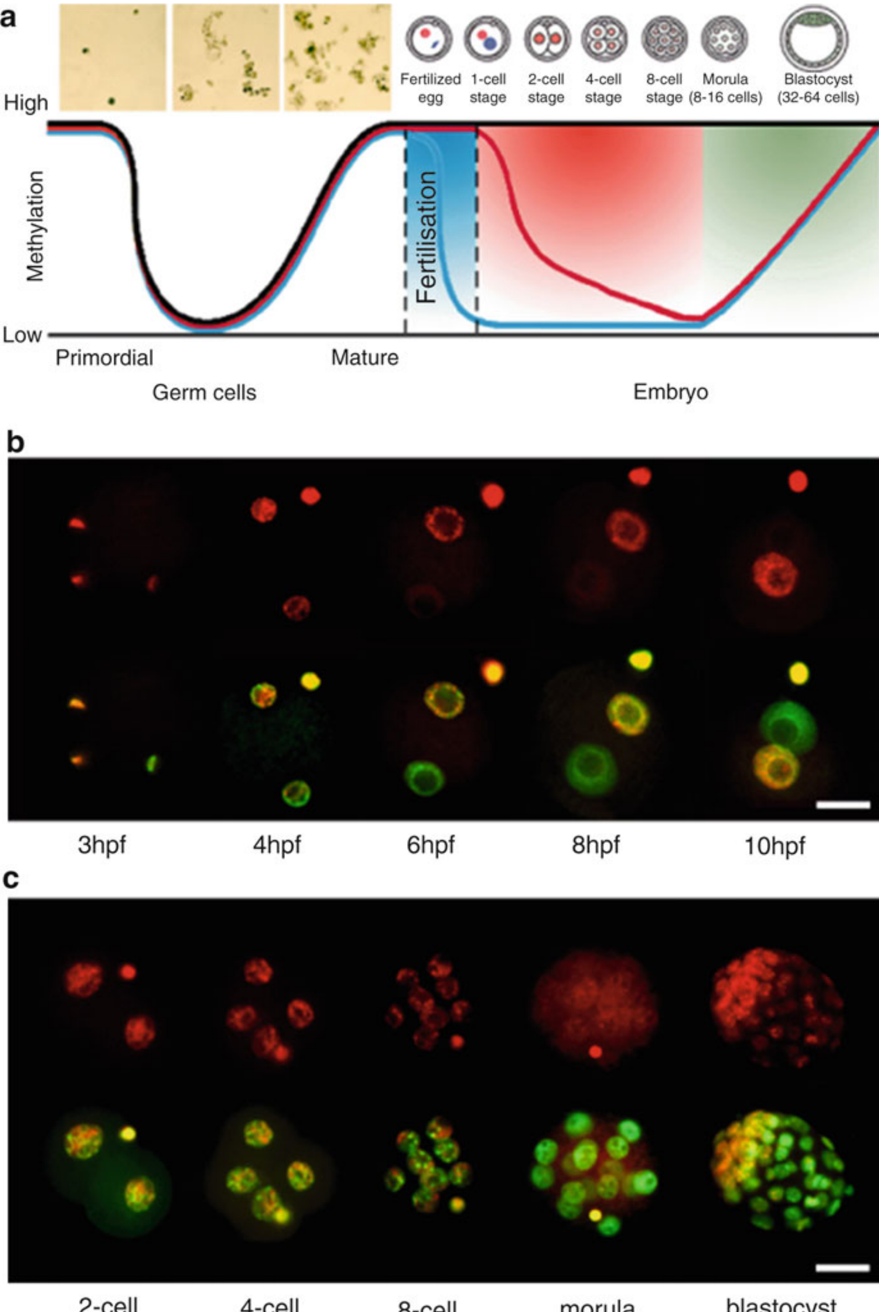
narrow time window (in a single G2 phase of the cell cycle), it is assumed to be in part an 'active' process, rather than simply by dilution of existing DNA methylation patterns owing to the failure to maintain methylation at replication. The mechanism of such dramatic DNA demethylation is not fully understood: it is thought to involve deamination of methylcytosine followed by base excision repair (BER) [17], but other mechanisms have not been discounted. After the onset of DNA demethylation, there is also substantial remodelling of chromatin, as revealed by immunofluorescence, with transient losses in signal for the linker histone H1 and the repressive marks H3K9me3 and H3K27me3, but long-term reduction in H3K9Ac, H2AR3me2 and H4R3me2 [13].

Recent whole genome shutgun bisulphite sequencing (Bi-Seq) of GCs at E13.5 has underscored the comprehensive nature of demethylation [18], although the depth of sequencing in that analysis was not able to detail methylation at individual loci. It is thought, however, that not all sequences are fully demethylated: repetitive elements of the intra-cisternal A particle (IAP) retrotransposon class have been found to be substantially resistant to demethylation [16, 18]. This observation provides a basis for some degree of transgenerational inheritance of epigenetic states. From this nadir, DNA methylation is re-established as GC differentiate (Fig. 6.1a), but follows quite different trajectories in male and female GCs. The dynamics of, and factors involved in, re-establishment are best understood for imprinted genes and repetitive sequences [1], although it might be assumed that other genomic elements that become methylated obey the same principles.

6.2.2 Establishment of Gamete-Specific Epigenomes

In female GCs in the mouse, *de novo* methylation occurs in oocytes arrested at the diplotene stage of meiotic prophase I. It initiates after birth as oocytes have been recruited into growth and is dependent on oocytes attaining a specific size [19, 20].

Fig. 6.1 Methylation reprogramming during mouse development. (**a**) The dynamics of DNA methylation, starting from PGCs as they enter the genital ridge through to blastocyst stage embryos is illustrated. The diagram depicts methylation levels generally of the maternal (*red*) and paternal (*blue*) genomes and of imprinted genes (*black*), which are resistant to demethylation in the zygote. Examples of PGCs at E11.5, 13.5 and 14.5 stained for the PGC marker alkaline phosphatase are pictured above. (**b**) Genome-wide DNA methylation dynamics illustrated by indirect immunofluorescence of fertilised oocytes using an antibody to 5-methylcytosine. Active demethylation of sperm-derived chromosomes during the first cell cycle is revealed by loss of DNA methylation (*red* signal) from the male pronucleus (green in lower merge panel) but not the female pronucleus. The lower panels show a merge (yellow) between the 5-methylcytosine (*red*) and DNA stain (*green*). hpf, hours post fertilisation. Scale bar 25 μm. (**c**) Passive demethylation, thought largely to affect the oocyte-derived chromosomes, is visualized from the 2-cell stage to the morula by reduction in 5-methylcytosine staining (*red*). *De novo* methylation initiates by the blastocyst stage and preferentially in the ICM. The lower panels show a merge (*yellow*) between the 5-methylcytosine (*red*) and DNA stain (*green*). Scale bar 25 μm. (Reprinted from Santos and Dean [14]. © Society for Reproduction and Fertility (2004), reproduced by permission)


a
Fertilized egg
1-cell stage
2-cell stage
4-cell stage
8-cell stage
Morula (8-16 cells)
Blastocyst (32-64 cells)
High
Methylation
Low
Fertilisation
Primordial
Mature
Germ cells
Embryo
b
3hpf
4hpf
6hpf
8hpf
10hpf
c
2-cell
4-cell
8-cell
morula
blastocyst

Full methylation of imprinted sequences may only be completed in fully-grown germinal vesicle stage oocytes, shortly before ovulation. Therefore, the whole cycle from erasure to full resetting of epigenetic marks in the female germline spans the period from germ cell specification in the embryo to pre-pubertal oocyte growth and perhaps beyond; the extent to which this applies in human is not known. Methylation of imprinted sequences and repetitive elements depends upon the *de novo* methyltransferase Dnmt3A and its co-factor Dnmt3L [21–23]; the second *de novo* activity, Dnmt3B, is not highly expressed and is dispensible in oocytes [24]. Recent genome-wide analysis reveals that ~7% of CpG islands (CpG islands are associated with many gene promoters) are highly methylated in mature oocytes in the mouse and are similarly dependent for methylation on Dnmt3a and 3L [25]. Many of these methylated CpG islands retain significant methylation in preimplantation embryos and may thus influence gene expression patterns after fertilisation. Oocytes appear to be unusual (together with embryonic stem cells), in exhibiting a substantial amount of cytosine methylation outside of the normal 'symmetrical' CpG dinucleotide context [26]. This might reflect the fact the DNMT3A does not exclusively methylate cytosines in a CpG context, its high level of activity in oocytes and the absence of replication-dependent maintenance which is only active at CpG sites. It is not clear whether non-CpG methylation has any biological significance. What other factors contribute to *de novo* methylation in oocytes is not fully clear. That the interaction of Dnmt3L and Dnmt3A with chromatin is inhibited by the presence of methylation at H3K4 [27, 28] indicates an important interplay between histone modification state and DNA methylation and is consistent with methylation being directed towards transcriptional inert sequences (H3K4 methylation is a marker of active genes). Consistent with this possibility, methylation of some imprinted sequences is disrupted in oocytes lacking the H3K4 demethylase KDM1B [29]. In addition, however, methylation of some imprint control regions (ICRs) has been shown to depend on transcription through ICRs in growing oocytes [30]. Studies of human imprinted gene syndromes, including biparental complete hydatidiform moles, in which conceptuses lack methylation at imprinted genes determined in oocytes, have identified mutations in two members of the NLRP family of inflammation-related proteins, NLRP2 and NLRP7 [31, 32], but the mechanism of their involvement in DNA methylation establishment is not known.

DNA methylation in the male germline, in contrast, occurs in GCs (referred to as gonocytes or prospermatogonia) arrested in G1 of mitosis. At imprinted sequences and repetitive sequences, *de novo* methylation initiates in fetal gonocytes from E14.5, but may continue until birth and beyond [1, 33]. Genome-wide remethylation is completed in primary spermatocytes, prior to entry into meiosis, although there is evidence that some loci may become methylated late in spermiogenesis [34, 35]. As in oocytes, Dnmt3A and Dnmt3L are both required, but Dnmt3B also plays a role at one of the four imprinted genes known to be methylated in sperm [23, 36]. The mechanisms targeting DNA methylation in male GCs are less well understood, but might be anticipated also to involve interactions with permissive histone modifications [37]. One mechanism, which might be specific to male GCs, is the involvement of a specific class of small RNAs, the Piwi RNAs, in a process called

RNA-directed DNA methylation (RdDM) [38, 39]. This mechanism appears to be directed particularly towards methylation of transposable elements, but also operates at at least one imprinted locus [40].

The outcome of these distinct *de novo* methylation processes are radically different DNA methylation landscapes in the mature oocyte and sperm [24]. Differences in DNA methylation between male and female gametes are exemplified by the germline differentially methylated regions (DMRs) of imprinted genes. Imprinting is controlled by such differences, which manifest as regions of dense CpG methylation that are fully methylated in one germline but not the other; currently, a substantially larger number of DMRs are known that acquire methylation in oocytes than in sperm [26]. These methylation differences, which are faithfully maintained as a memory of parental origin of the alleles in embryos, will determine parental-allele specific expression states, in many cases monoallelic silencing, of the associated genes. The extent to which other sites of differential methylation in gametes impact on gene expression patterns in embryos is not known, but should also emerge from genome-wide methylation studies [25, 41]. One reason for the late onset of methylation is that methylcytosine is the base most prone to mutation: deamination of a methylcytosine results in conversion to thymidine and occurrence of a T:G mismatch. Repair of this mismatch may restore the correct C:G base pair but could also lead to replacement by a T:A pair and an ensuing C→T transversion. Because of this vulnerability, it is suggested that remethylation is left as late as possible during gamete development to minimise the fixation of these transversion events [42, 43]; nevertheless, C→T transversions are the most frequent mutation events in the human genome. There is evidence of natural variation, both intra- and inter-individual, in DNA methylation in human sperm samples [44]; in addition, sub-fertile males appear to have increased incidence of DNA methylation defects at imprinted loci [45, 46]. Variations in sperm methlyation patterns could have an influence on offspring phenotype, if such differences were maintained beyond the genome-wide demethylation experienced by sperm-derived chromosomes in the preimplantation embryo (see Sect. 6.3.1).

There are also changes in chromatin architecture during germ cell differentiation and maturation to accommodate the requirements of meiosis and, in the mature sperm, genome compaction. Both male germ cells and growing oocytes contain variants of the canonical histones [47]. Disruption of genes for the H3K4 trimethyltransferase PRDM9 and the H3K9 mono- and di-methyltransferase EHMT2 cause meiotic arrest in oocytes [48, 49], suggesting that H3K4me3 and H3K9me1/2 are required for chromosome structures for chromosome synapsis and/or recombination, or that these activities are essential for the correct meiosis-specific gene expression. Otherwise, there is little change to oocyte histones during meiosis, except for deacetylation of H3 and H4 at the stage of germinal-vesicle breakdown and the resumption of meiosis prior to maturation. These, and other chromatin modifiers such as the H3K9 trimethyltransferases Suv39h1 and 2, are also required for progress through early meiosis in male GCs [48–50]. Subsequently, during spermiogenesis, chromatin becomes highly condensed through a sequence of events including replacement of testis-specific linker histones H1T2 and H1LS1, replacement of histone

by the transition proteins TNP1 and 2 and, finally, in spermatids their replacement by protamines [47] to complete the compaction of DNA in the spermatozoa. However, sperm chromatin is not devoid of histones, but up to 15% remain (depending on the species). Histone retention is locus-specific, including at imprinted loci and genes with functions in early embryogenesis, and retained histones carry activating and repressive modifications, such as H3K4me3 and H3K27me3, suggesting that developmentally important genes may be poised for activation in the early embryo [51, 52], if these histones and their associated modifications are maintained during the global replacement of protamines by histones in the zygote. Understanding the extent of epigenetic information in sperm and its persistence during embryogenesis is fundamentally important in considering how environmental effects experienced by males can be transmitted to influence offspring phenotype [53].

6.3 Epigenetic Reprogramming Events in Mammalian Embryos

Fertilisation brings with it a new wave of reprogramming of the epigenome that seeks to achieve a homogenisation of the differences between sperm- and oocyte-derived chromosomes, preparation of the DNA for the process of zygotic genome activation, the restoration of totipotency and preparation for the first lineage specification. The major reprogramming events have been elegantly demonstrated by immunofluorescence studies with antibodies directed against methylcytosine (Fig. 6.1b), specific histone modifications and the associated histone modifier complexes [54, 55]; in addition, methylation of a few genes has been analysed directly by bisulphite sequencing [25, 41].

6.3.1 Reprogramming of Paternal and Maternal Genomes After Fertilisation

The first of these events is the decondensation of sperm chromatin and replacement of protamines with histones stored in the oocyte, which occurs in the male pronucleus before formation of the zygote is complete with the fusion with the female pronucleus. This replication-independent deposition of histones results in the chromatin of the sperm-derived chromosomes being enriched in the variant histone H3.3 and depleted in modifications such as H3K9me2, H3K9me3 and H3K27me3, which are apparent in the female pronucleus. The differences in histone modifications between paternal and maternal genomes can be detected at last until the four cell stage [55]. Such differences are prominent at the highly condensed regions of pericentromeric heterochromatin, but whether they also reside at single copy sequences in euchromatic regions and whether they influence how the two parental alleles of

genes are expressed at that stage is not clear. Some imprinted genes have been examined and found to exhibit monoallelic expression in preimplantation embryos dependent on the activity of PcG proteins, but this might depend on the pre-existing DNA methylation differences inherited from the gametes [56].

Soon after protamine replacement, there is a rapid and comprehensive reduction in DNA methylation of the paternal pronucleus, which has been demonstrated by antibody staining and by bisulphite sequencing of specific genes [57–59]. It occurs prior to DNA replication, so is thought to be an active process. The activities responsible for this demethylation have long been sought, and have included members of the transcription elongation complex [60], whose mechanism of action is unclear at this time. Recent studies have identified 5-hydroxymethylcytosine (5hmC) as a possible intermediate [61]. 5hmC has been detected in the zygote by immunofluorescence and one of the enzymes that convert 5mC to 5hmC, Tet3, is present in the fertilised oocyte [62]. 5hmC can be detected at least until the 8-cell stage, suggesting that it remains on the paternal genome without further metabolism, but diluted by DNA replication [62, 63]. What the subsequent fate of this 5hmC is and the extent to which it has an impact on gene expression and on determining other epigenetic modifications is still unknown. 5hmC was first identified as a minor modification in DNA in neuronal tissue and mouse ES cells [61, 64]; more recent studies suggest that it can act as an epigenetic mark it is own right beyond a role as a by-product of cytosine demethylation, as Tet protein activity is required for normal ES cell self renewal and early lineage decisions [65–67].

Remarkably, the female pronucleus appears to be protected from active demethylation in the fertilised oocyte, and the reasons for this are unknown at present, although one oocyte-derived protein, Stella, has been identified without which sequences from the maternal genome also become demethylated [68]. However, there is a substantial reduction of DNA methylation signal overall during the first embryonic cleavage divisions, suggesting that the maternal genome also experiences demethylation, and this is assumed to be a ‘passive’ process, owing to the absence of maintenance methylation at DNA replication: 5hmC appears to be absent from the maternal genome and the Tet3 enzyme rapidly disappears after fertilisation [62]. Indeed, the maintenance methyltransferase Dnmt1, which has been concentrated in the oocyte cytoplasm during oocyte growth, appears to be largely excluded from the nucleus of preimplantation embryos [69, 70]. Thus, by the time the embryo comprises 16–32 cells (morula stage), genome-wide DNA methylation levels have reached a minimum. This picture has depended largely upon 5-methylcytosine (5mC) staining studies and, until recently, the number of single copy genes examined for methylation was limited. It is clear, however, that some sequences do faithfully maintain CpG methylation whilst the genome as a whole is losing methylation and the classic example are the DMRs of imprinted genes [69]. Recent genome-wide analysis of methylation in mouse blastocysts has, however, revealed a greater diversity in the behaviour of gametically-derived DNA methylation and identified additional gene promoters that retain at least some degree of methylation from oocytes throughout preimplantation development [25, 41]. What governs which sequences maintain or lose methylation is not fully understood, nor whether

the existence of gamete-derived CpG island methylation influences early lineage decision events. In addition to the dependence on Dnmt1 and its associated factors [69, 71, 72] a number, but not all, DMRs also require the presence of the zinc-finger protein ZFP57: mouse embryos lacking Zfp57 fail to maintain methylation at a number of DMRs [73] and homozygous mutation of ZFP57 has also been found as a common cause of 'maternal hypomethylation syndrome' in humans, in which patients have loss of methylation of a variable number of imprinted DMRs [74]. ZFP57 is a member of the Krüppel-like family of transcriptional repressors and an interacting protein, KAP1, has been found to be necessary for maintenance of methylation of endogenous retroviral elements (IAPs) in early mouse embryos [75]. The mechanisms by which ZFP57 and KAP1 interact directly or indirectly with Dnmt1 to ensure methylation are not known.

6.3.2 Establishment of New Epigenomes with the First Lineage Specification Events

From a low point in the morula, DNA methylation begins to be restored by the blastocyst stage (Fig. 6.1c), but differences in methylation level between the incipient embryonic lineage, represented by the ICM and extra-embryonic lineage, represented by the trophectoderm (TE) soon become apparent, with the TE showing reduced 5-methylcytosine staining [54]. This global difference likely reflects generally reduced methylation of repetitive sequences in TE [76]. Concomitant with the difference in DNA methylation between ICM and TE, gross differences in some histone modifications have also been observed, including acetylated H3K9 and methylated H3K27 [77]. The ICM and TE lineages are defined by distinct networks of pluripotency and lineage-determining transcription factors and there is considerable interplay between epigenetic modifiers and these factors in the preimplantation embryo (as discussed extensively in [2]). This is exemplified by the co-activator associated arginine methyltransferase 1 (CARM1), which is active against the R2, R17 and R26 residues of histone H3. Blastomeres of the 8-cell embryo expressing higher levels of CARM1 are destined preferentially for the ICM [78]. Of the genes directly or indirectly regulated by the action of CARM1, the caudal-type homeobox 2 gene *Cdx2*, whose expression is reduced in blastomeres with higher CARM1 levels [79], is one of the major determining factors of TE [80]. Epigenetic mechanisms are seen to reinforce lineage specification decisions. The ETS-related family transcription factor ELF5 cooperates with CDX2 and another TE determinant eomesodermin (EOMES) to stabilise the TE decision and once ELF5 is activated the cell fate choice becomes fixed. ELF5 is epigenetically regulated as its promoter is silenced by DNA methylation in the embryonic lineage [81]. Similar epigenetic fixation of lineage decisions can be seen for the pluripotency factor Stella, which is expressed in the ICM but becomes repressed and methylated in the epiblast [82]. Individual blastomeres are not thought to be irreversibly committed initially to either ICM or TE fate but the expression of factors such as ELF5 and their epigenetic silencing form a barrier to switching and consolidate cell fate choice.

Recent genome-wide investigations of promoter DNA methylation performed by MeDIP-chip on DNA from preimplantation (blastocysts, E3.5) and post-implantation embryos (E6.5 and E9.5) reinforce the notion that DNA methylation is used to control the activity of genes in a lineage-specific manner [41]. That study confirmed that trophoblast and pluripotency related genes, including *Elf5* and *Stella*, had become *de novo* methylated in the epiblast by E6.5; it also identified that many germ-cell specific gene promoters are already methylated in blastocysts. In addition, several genes specific to differentiated tissues, including brain, eye and the haematopoietic lineages, were found to be methylated in the epiblast, suggesting that the promoters of these genes must become demethylated specifically in their respective lineages later in embryogenesis. This was indeed found to be the case by testing promoter methylation of selected genes during adult haematopoietic stem cell differentiation into B cells (and has been corroborated in an independent analysis; [83]). Therefore, epigenetic mechanisms and factors are intimately involved in regulating lineage-specifying genes, in reinforcing cell fate choices by disallowing the expression of genes inappropriate to a given lineage and, through demethylation events, in allowing activation of differentiation-specific gene sets.

6.4 Tissue-Specific Epigenomes: Plasticity and Stability

The advent of genome-wide methods to map DNA methylation and histone modifications is providing full epigenomic profiles of specific cell types, how they develop during differentiation and how epigenomes are altered in disease. A goal of current international effort, for example, as coordinated by the International Human Epigenome Consortium and NIH Roadmap Epigenomics Mapping Consortium [84], is to generate standardised reference epigenomes for a common set of epigenomic marks in a defined set of human tissues and cell lines. With the explosion of epigenome mapping studies in the last couple of years it is impossible to give a comprehensive account, so I will restrict the discussion to a few examples to illustrate general principles and some new insights that have emerged.

6.4.1 Epigenome Changes During Differentiation

Embryonic stem (ES) cells have been used widely as a proxy for the ICM to understand the gene expression programmes and epigenomic profiles associated with pluripotency. In an early application of next generation sequencing to chromatin immunoprecipitation of specific histone modifications (ChIP-Seq), Mikkelsen and colleagues profiled mouse ES cells in comparison to neural progenitor cells and embryonic fibroblasts [5]. The power of sequencing based-approaches is the unambiguous way in which data on histone modifications, DNA methylation and mRNA expression can be correlated. In this way, it was possible to detect how the patterns

of enrichment for an activating mark (H3K4me3) and repressive mark (H3K27me3) were able to discriminate genes expressed, poised for expression or stably repressed. This, and other studies (e.g., [85]), suggested that genes poised for expression, in particular, were found to be marked by 'bivalent' chromatin, a paradoxical combination of the active and repressive modification, with bivalent marks resolving into active or repressed states in parallel to changes in gene expression in differentiated cells. In addition, ChIP-Seq was able to confirm the distributions of H3K36me3 as a marker of active transcription units, and H3K9me3 and H4K20me3 as markers of constitutive heterochromatin at centromeric satellite sequences, telomeres and other repetitive element classes. The reference epigenome of ES cells is an essential asset in seeking to determine the extent and fidelity of reprogramming of differentiated somatic cells to induced pluripotent stem cells (iPSCs) [86, 87].

In a systematic investigation of a single cell type, human CD4+ T lymphocytes, Zhao and colleagues mapped the distribution of 39 histone modifications (18 acetylation states and 19 methylation states) in comparison with gene expression data [88]. Although a large number of different combinations was found, a common pattern comprising 17 modifications was detected at over 3000 promoters. Genes with no or low levels of expression were depleted in histone acetylation and marked with H3K27me3 and/or other repressive histone methylation marks (H3K9me2, H3K9me3, H4K20me3). Expressed genes had the 'backbone' set of 17 modifications; additional modifications were associated with the most highly expressed genes, with the number of histone acetylations strongly associated with expression. Different acetylated histones were distributed differentially at promoter regions and gene bodies. Enhancers also appeared to have distinct and limited combinations of modifications, with more than 20% characterised by presence of the variant histone H2A.Z, as well as H3K4me1, H3K4me2, H3K4me3, H3K9me1 and H3K18ac modifications. This diversity of histone modification patterns seems bewildering and leaves one wondering what marks to focus on if one wished to compare epigenomic profiles in case-control studies for a disorder of interest and future insight into the dependencies of histone modifications will be important. To consider how histone modifications differ between cell types, Ren and colleagues mapped a limited number of histone methylation states in five human cell lines. The study found that the chromatin state at promoters was relatively invariant, but that there were highly cell-type-specific histone modification patterns at enhancers that correlated strongly with cell-type-specific gene expression [89]. Similar genome-wide histone modification and transcription factor binding analyses have been applied to *in vitro* differentiation systems, such as adipogenesis in both human and mouse models [90]. One outcome of the study was the identification of novel, stage-specific alternative promoters and distal enhancer elements, which were inferred from changes in chromatin signatures during differentiation (Fig. 6.2).

Whereas histone modification status emerges from such studies as being potentially dynamic, DNA methylation has traditionally been thought of as a more stable mark associated with long-term gene repression. CpG islands, short C:G-rich regions that are associated with a large proportion of gene promoters, are generally viewed as being constitutively unmethylated, except for the rare cases of those in

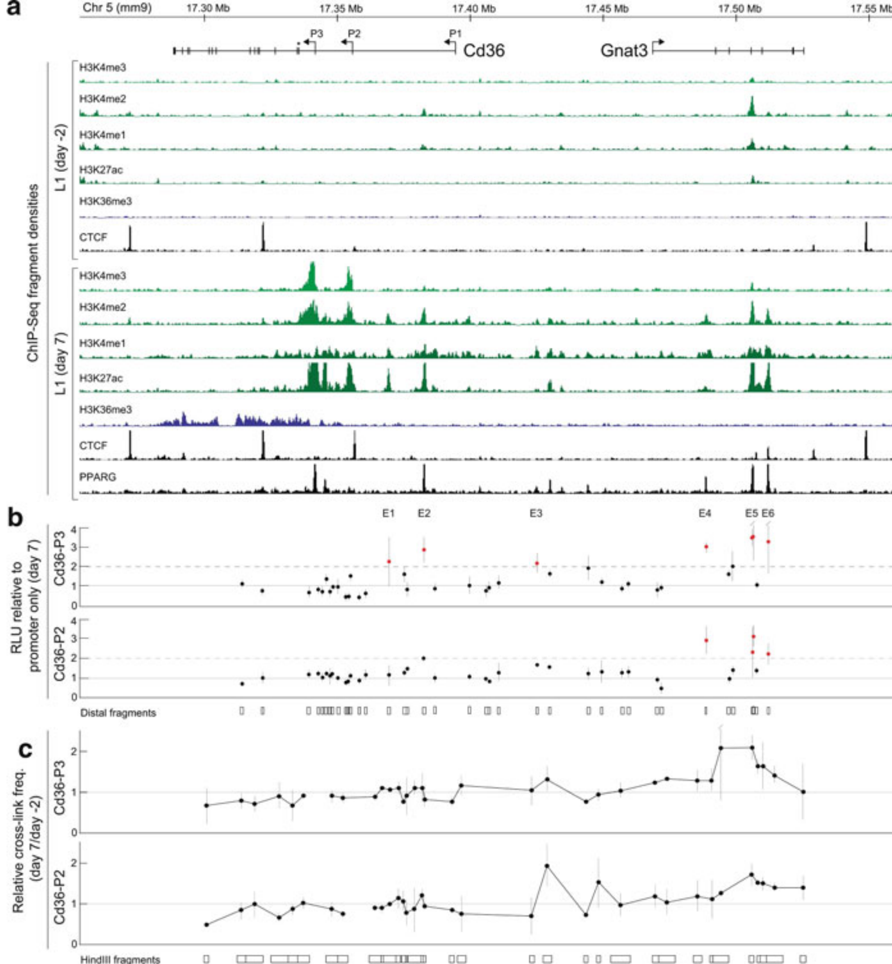


Fig. 6.2 Chromatin changes during adipogenic differentiation. Chromatin state maps of the ~300 kb *Cd36-Gnat3* locus on mouse chromosome 5 mouse L1 preadipocytes (day −2) and mature adipocytes (day 7) obtained by ChIP-Seq. Profiles for modified histone H3K4me1, me2 and me3, H3K27ac1 and M3K36me3 are shown, as well as binding sites for CTCF and Pparγ. Of particular note, enrichment peaks for H3Kme3 reveal the recruitment of new promoters of *Cd36* (P2 and P3) during differentiation and enrichment in H3K36me3 over the *Cd36* gene body is a marker of transcription elongation. Pparγ is a transcriptional regulator during adipogenesis and CTCF is a multifunctional DNA binding factor that has chromatin insulator function. (Reprinted from Mikkelsen et al. [90]. Copyright (2010), with permission from Elsevier)

imprinted genes and those on the inactive X chromosome in females [91]. Aberrant CpG island methylation is a common phenomenon in tumours, in which it is one mechanism by which tumour suppressor genes are inactivated. Several studies have now revealed that around 10% of CpG islands are methylated normally in a

tissue-specific manner [92–94]; in many cases, tissue-specifically methylated CpG islands demarcate alternative and tissue-specific promoters. CpG island methylation is strongly inhibitory to promoter activity, and erasure of methylation for promoter de-repression is likely to require substantial chromatin remodelling events, but absence of methylation by itself is not a strong predictor of expression level (e.g., [42]). Some observations have indicated that the less CpG-dense regions flanking CpG islands, for which the term 'CpG island shores' has been coined, exhibit tissue-specific differences in methylation that more closely correlate with gene expression status [83, 95]. Although DNA methylation has long been suspected as having a role in controlling gene expression programmes during differentiation, very few studies, up to now, have looked globally at DNA methylation status during differentiation of a specific lineage. One recent study has now done so during haematopoietic differentiation in the mouse, comparing methylation at 4.6 million CpG sites in eight sorted populations ranging from multipotent progenitors to progressively restricted progenitors of the lymphoid (T-cell), and myeloid (granulocyte and macrophage) lineages [83]. In this system, reduced methylation is associated with myeloid rather than lymphoid differentiation. Numerous methylation differences were found between the various populations, with differences at CpG islands shores generally predominating except, interestingly, in the earlier differentiation stages. As noted above, CpG island shore methylation correlated more closely with gene expression.

There is assumed to be a considerable interplay between DNA methylation and histone modificadtions. DNA methylation is inversely correlated with H3K4me3 at promoters and CpG islands [42, 93, 96], and this is consistent with the observation that the *de novo* methylation co-factor Dnmt3L is inhibited from binding histone H3 when lysine 4 is methylated [27], so has some mechanistic basis. Other factors are likely to contribute to the general protection of CpG islands against DNA methylation. A protein bound at the majority of CpG islands, Cfp1, interacts with the H3K4 methytransferase Setd1 and CpG islands lose H3K4me3 if Cfp1 is ablated [97]. An H3K36me2 demethylase, KDM2A, is also constitutively bound at unmethylated CpG islands, which might help to ensure a chromatin environment non-permissive for DNA methylation [98]. Cell-type-specific gain of methylation at CpG islands is thus likely to involve destabilisation of binding of several protective factors. Conversely, interactions between Dnmt proteins and Ezh2, a component of the PRC2 complex that deposits the repressive mark H3K27me3, have been found [99, 100]. Interestingly, a number of genes identified with a bivalent chromatin signature in ES cells, and thus targets of the PRC2 complex, undergo aberrant DNA methylation in human cancers [101, 102].

6.4.2 Longevity of Epigenetic Marks

A key question is to what extent the epigenome is 'hard-wired' and fixed in perpetuity once established. Certainly, there is an elegantly simple mechanism by which

DNA methylation can be inherited at each DNA replication cycle, because Dnmt1 is present at the replication fork and methylates the nascent DNA strand at hemimethylated CpG sites. It is assumed that DNA methylation is maintained by default, unless active mechanisms remove it (as in PGCs and the zygote) or prevent access of Dnmt1. This model is likely to be simplistic, as multiple factors, such as the Uhrf1 protein, are required to target Dnmt1 to newly synthesised, hemimethylated DNA at replication [103, 104], and these dependencies offer additional levels of regulation (or scope for error) that could allow methylation states to change. Histone modification states are also seen to be mitotically heritable, as in the blocks of constitutive heterochromatin at centromeres and telomeres. But how histone modifications are propagated at DNA replication is still hotly debated – whether there is semi-conservative inheritance of the dissociated core nucleosomes to the daughter strands that may then template modifications of the newly incorporated histones, or whether there needs to be wholesale re-establishment of histone modifications [105]. Persistence of chromatin states might also involve a cross-talk between DNA methylation and histone modifying factors, for example, Dnmt1 may contribute to maintenance of H3K9 methylation [106].

It is certainly the case that some epigenetic marks are extremely stable, such as the methylation of CpG islands on the inactive X chromosomes of females or in imprinted genes. In the case of X-inactivation, the choice of which chromosome to inactivate and methylate is taken randomly at a cellular level in female embryos around implantation and then all daughter cells maintain this choice throughout the lifetime of the organism. CpG island methylation is just one of the large set of epigenetic properties that differentiate active and inactive X chromosomes, and these epigenetic marks are likely to reinforce one another. For imprinted genes, methylation differences of the two alleles are determined in the parental gametes and these two states persist as a faithful memory of parental origin in all tissues (it is believed) and throughout life. Again, the parental alleles are also differentiated by multiple histone modifications that may act in concert with DNA methylation [107]. DMRs generally extend several kilobases and comprise hundreds of CpGs fully methylated on one parental allele and fully unmethylated on the other [26], so effects on allelic expression may require large scale changes in methylation. However, if lost in somatic tissues, there is no mechanism by which to re-impose methylation imprint marks except by passage through the germline [108]. It might be the case that these represent one extreme for which robust mechanisms have evolved to ensure the maintenance of methylation, because of the extreme adverse effects on X chromosome balance or imprinted gene expression dosage should the mechanisms fail. Nevertheless, there is genetic and molecular evidence that X-inactivation weakens with ageing [109, 110], and a substantial number of genes on the human X chromosome partially escapes inactivation, and this appears also to vary between individuals [111]. Several reports have indicated a relaxation of *Igf2* imprinting with ageing, particularly because of its association with cancer (e.g. [112]). One study reported considerable inter-individual variation in DMR methylation and found that skewed methylation clustered in families, suggesting a genetic or shared environmental component [113]. In contrast, another group found that imprinted genes showed

relatively little change in methylation associated with aging in adults [114]. There may be technical reasons for these apparent discrepancies, including the tissues analysed, methods used and whether the limited number of CpG sites assayed in any one study is representative, and more extensive analyses are warranted. In addition, the inability to assay expression in the most relevant tissues may also explain why imprinted gene expression is sometimes found to correlate poorly with DNA methylation [115]. Beyond X-inactivation and imprinting, twin studies have indicated a drift in DNA methylation patterns even in genetically identical individuals with age: older monozygotic twin pairs exhibited greater differences in DNA methylation profiles [116]. But, again, other investigations suggest a limited number of methylation changes associated with ageing: in microarray analysis of skin biopsies, DNA from older individuals was characterised by a specific hypermethylation pattern, however, this affected less than 1% of the markers analysed [117]. It may be anticipated that additional, carefully designed genome-wide studies with emerging epigenomic techniques will lead to a consensus on the degree to which our epigenomes age with us, the factors that promote change and the genomic loci most vulnerable.

6.5 Transgenerational Epigenetic Effects

There is a legitimate interest, but also some confusion or controversy, over the existence of transgenerational epigenetic inheritance. Epigenetic inheritance could be interpreted from variation in a phenotype present in one generation influencing the penetrance of the phenotype in subsequent generations in the absence of any DNA sequence change, or because environmental conditions (in the broadest sense) present in parents have an impact on subsequent, unexposed generations, bearing in mind that the germ cells that will give rise to grandchildren are developing in a female exposed during pregnancy [118, 119]. Evidence for transgenerational effects in humans and mice comes from the study of metastable epialleles, such as *agouti viable yellow* (A^{vy}) [120, 121], dietary challenges particularly before conception [53], exposure to environmental toxins, such as endocrine disruptors [122, 123], and epidemiological studies showing disease outcomes in grandchildren of affected populations [124, 125]. It is important to set such observations against the knowledge of epigenomic reprogramming in germ cells before concluding whether they satisfy criteria for transgenerational epigenetic inheritance.

As we have seen earlier in this chapter, one of the prerequisites of germ cell development is to erase the pre-existing epigenome of somatic cells to return to a state of pluripotency. However, it is also evident that germ cells are specified and differentiate in the developing embryo, so they could be vulnerable to adverse conditions *in utero*. Gametes have their own epigenomic landscapes, so there is the potential that modified or aberrant epigenetic marking of gametes can be transmitted to the embryo, providing that such alterations can survive the next phase of genome-wide epigenetic reprogramming in the preimplantation embryo and are

not further corrected as lineages are specified. Epigenetic mechanisms are energy consuming processes, so could be affected by the metabolic status of the cell and nutritional conditions of the organism [126] and, indeed, it is expected that there will be at least acute gene expression responses to nutritional changes that depend on epigenetic mechanisms. Finally, exposing pregnant females to diets that significantly influence one carbon metabolism that provides the methyl groups for methylation reactions could affect the maintenance of DNA methylation patterns potentially at loci common in offspring and their progeny (depending on the timing of exposure), because the germ cells for the next generation are developing in the exposed embryos and because maintenance depends universally and ultimately on the activity of Dnmt1.

Amongst the most extensively characterised examples of epigenetic inheritance is the A^{vy} allele in mice [120]. A^{vy} is a naturally occurring allele of the *agouti* locus in which an IAP element has inserted upstream of the *agouti* gene resulting in its misregulation (ectopic expression). IAPs are normal targets of epigenetic silencing but, in this case, methylation is variable between individuals resulting in variation in ectopic *agouti* expression, which manifests in coat colour phenotype, with genetically identically littermates ranging from completely yellow, through mottling to fully agouti (brown). Importantly, the range of coat colour phenotypes amongst offspring is influenced by maternal coat colour. Direct assessment of methylation at the IAP long terminal repeat (LTR) shows not only a strong inverse correlation between coat colour and degree of methylation, both in mothers and progeny, but also that DNA methylation in oocytes and sperm is related to coat colour and methylation type of parents [120, 121]. Embryo transfer experiments revealed that the variation was intrinsic to the A^{vy} allele and not a consequence of physiological differences between the different classes of females [120]. It is probably best to regard this example as a case of incomplete erasure of a somatic methylation pattern in germ cells and, indeed, IAPs are known to resist partially the genome-wide reprogramming in germ cells and early embryos [16, 18]. Intriguingly, however, sperm-derived methylation at the A^{vy} allele is largely erased in the zygote (consistent with absence of paternal transmission), but so is oocyte-derived methylation by the time of implantation [121], suggesting that additional epigenetic marks are responsible for the probability of re-establishing the methylation type of the mother. Owing to the sensitivity of the A^{vy} system, it has been used as a model to study dietary influences on epigenotype. For example, supplementation of the diet of wild-type female mice with methyl donors increases the methylation of the A^{vy} allele of their a/A^{vy} progeny [127]. However, there are conflicting reports whether these acquired methylation changes are passed on to subsequent generations [128, 129].

Analogous metastable epialleles may exist throughout the genome and influence expression of genes in their proximity, including in humans. A recent study identified such epialleles in humans: methylation level was similar across tissues of an individual but varied even between monozygotic twins, demonstrating the early and stochastic nature of methylation establishment at these loci. Moreover, it was found that in a sub-Saharan African population, which experiences profound seasonal fluctuations in food quality and availability, levels of DNA methylation varied in

individuals according to the season of their conception [130]. It was noted that methylation at these loci may be particularly sensitive to this effect, and more so than the imprinted loci tested. Although this observation establishes the principle of nutritionally conditioned epigenetic variation, it will be important to determine whether such epigenetic differences contribute to physiological differences and whether there is any vestige of methylation level in subsequent generations. Perhaps the most extensively analysed example of the impact of extreme maternal nutrition on health outcomes in offspring in humans is the Dutch hunger winter. DNA methylation analysis of several imprinted genes has detected differences at some, but not all, loci in individuals born to women pregnant at the time, with the alterations differing according to whether malnutrition was experienced in the first or third trimester. However, methylation was determined in whole blood DNA of offspring 60 years after the famine (by necessity) and the differences were marginal, so it is not possible to conclude whether they correspond to epigenetic alterations that were induced during gestation or are a consequence of later developmental or physiological differences [131, 132].

Transgenerational effects through the paternal line should be easier to interpret, because exposed fathers merely transmit genetic material and have no other direct influence on development of offspring. Sperm carry epigenetic information in the form of DNA methylation patterns, in as much as these survive early embryonic reprogramming, and the observation that sperm chromosomes retain some modified histones non-randomly adds impetus to investigate the possibility of paternal epigenetic transmission. In this context, it was recently reported that gene expression alterations can be detected in offspring of male mice fed a low protein diet for 3 months before mating [53]. Changes in hepatic DNA methylation were detected, exemplified by a modest but reproducible reduction in methylation at an upstream enhancer for the *Pparα* gene that correlated with gene expression (generally, promoter methylation was unaffected). Gene expression changes found in livers of offspring differed from those of the treated males, but the study did not investigate effects in the next generation. At this time, it is not clear how nutritional intervention in male mice resulted in gene expression changes in offspring, but it appears not to be explained by developmental compromise or immaturity of sperm. Also, there did not appear to be a simple relationship between altered DNA methylation profiles in liver of offspring and sperm, as the *Pparα* enhancer was highly methylated in sperm of control and dietary restricted males alike. Nevertheless, it would be important to ascertain the ontogeny and tissue-specificity of DNA methylation alterations and examine whether the effects noted are treatment specific (and potentially adaptive) or reflect a particularly vulnerable set of loci that could be deregulated by any adverse condition. Of alternative epigenetic marks as potential means of transmission, ChIP analysis detected differences in H3K27me3 at an unrelated locus and differences in mRNA loading of sperm were reported. Although not able to provide an explanation for the effects observed, the study does illustrate that we do now have the techniques to conduct sensitive, genome-wide studies into transgenerational inheritance. The most extensively investigated example of a transgenerational effect through the male germline is after the exposure to an endocrine disruptor, the

fungicide vinclozolin. Exposure of pregnant female rats by a course of intraperitoneal injection during the time of gonadal sex determination in embryos has been shown to cause impaired fertility in male offspring until the F4 generation, and a variety of other effects [119–122]. Microarray analysis of embryonic testis RNA from the F1 to F3 generations detected altered gene expression profiles, and although these tended to normalise with generation, ~200 transcripts remained deregulated in the F3 generation (being the first fully unexposed generation). Such gene expression changes could in part be a reflection of impaired spermatogenesis, although in the F3 generation no histological abnormality was detected in the embryonic testis. Altered expression of epigenetic modifiers including DNA methyltransferases was highlighted [133]. Recent genome-wide studies in sperm of F3 progeny identified a number of genomic loci displaying altered DNA methylation patterns, a number of which were validated in independent assays [123]. Further analysis is needed to follow the history of these alterations in sperm in the previous, exposed generations, as well as to investigate whether these loci transmit altered methylation to the somatic tissues and germline of their progeny.

In conclusion, there is a need to exercise caution and we should, perhaps, be more careful with definitions. An epigenetic change in offspring or grandchildren on account of parental exposure may be easy enough to explain from existing knowledge of the reprogramming events in germ cells and the timing of establishment of the final epigenetic profiles of gametes. This is not to say that such altered epigenetic patterning is unimportant, and can indeed have significant phenotypic or disease outcomes in offspring, but the expectation is that most epigenetic alterations should be reprogrammed in the germline of the unexposed generation and not permanently fixed. The term transgenerational inheritance should be reserved for transmission of an altered epigenetic state perpetuated across multiple generations in the absence of the inducer, for which there is evidence in plants [134]). In particular, it seems premature to entertain the possibility that there is an adaptive process at work, in the sense of Lamarckian-type transmission of acquired epigenetic states.

References

1. Sasaki H, Matsui Y (2008) Epigenetic events in mammalian germ-cell development: reprogramming and beyond. Nat Rev Genet 9:129–140
2. Hemberger M, Dean W, Reik W (2009) Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. Nat Rev Mol Cell Biol 10:526–537
3. Hajkova P (2010) Epigenetic reprogramming–taking a lesson from the embryo. Curr Opin Cell Biol 22:342–350
4. Ferguson-Smith AC, Surani MA (2001) Imprinting and the epigenetic asymmetry between parental genomes. Science 293:1086–1089
5. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G et al (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448:553–560
6. Bock C, Tomazou EM, Brinkman AB, Müller F, Simmer F, Gu H et al (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat Biotechnol 28:1106–1114

7. Hayashi K, de Sousa Lopes SM, Surani MA (2007) Germ cell specification in mice. Science 316:394–396
8. Saitou M, Barton SC, Surani MA (2002) A molecular programme for the specification of germ cell fate in mice. Nature 418:293–300
9. Ohinata Y, Payer B, O'Carroll D, Ohinata Y, Surani MA (2005) Blimp1 is a critical determinant of the germ cell lineage in mice. Nature 436:207–213
10. Ancelin K, Lange UC, Hajkova P, Schneider R, Bannister AJ, Kouzarides T et al (2006) Blimp1 associates with Prmt5 and directs histone arginine methylation in mouse germ cells. Nat Cell Biol 8:623–630
11. Yamaji M, Seki Y, Kurimoto K, Yabuta Y, Yuasa M, Shigeta M et al (2008) Critical function of Prdm14 for the establishment of the germ cell lineage in mice. Nat Genet 40:1016–1022
12. Seki Y, Hayashi K, Itoh K, Mizugaki M, Saitou M, Matsui Y (2005) Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. Dev Biol 278:440–458
13. Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F et al (2008) Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature 452:877–881
14. Santos F, Dean W (2004) Epigenetic reprogramming during early development in mammals. Reproduction 127(6):643–651
15. Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W et al (2002) Epigenetic reprogramming in mouse primordial germ cells. Mech Dev 117:15–23
16. Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J et al (2003) Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis 35:88–93
17. Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, Surani MA (2010) Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. Science 329:78–82
18. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M et al (2010) Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature 463:1101–1105
19. Lucifero D, Mann MR, Bartolomei MS, Trasler JM (2004) Gene-specific timing and epigenetic memory in oocyte imprinting. Hum Mol Genet 13:839–849
20. Hiura H, Obata Y, Komiyama J, Shirai M, Kono T (2006) Oocyte growth-dependent progression of maternal imprinting in mice. Genes Cells 11:353–361
21. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH (2001) Dnmt3L and the establishment of maternal genomic imprints. Science 294:2536–2539
22. Hata K, Okano M, Lei H, Li E (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development 129:1983–1993
23. Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E et al (2004) Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature 429:900–903
24. Kaneda M, Hirasawa R, Chiba H, Okano M, Li E, Sasaki H (2010) Genetic evidence for Dnmt3a-dependent imprinting during oocyte growth obtained by conditional knockout with Zp3-Cre and complete exclusion of Dnmt3b by chimera formation. Genes Cells 15:169–179
25. Smallwood SA, Tomizawa S, Krueger F, Ruf N, Carli N, Sato S et al (2011) Dynamic CpG island methylation landscape in oocytes and its fate in preimplantation embryos. Nat Genet. 43:811–814
26. Tomizawa S, Kobayashi H, Watanabe T, Andrews S, Hata K, Kelsey G et al (2011) Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. Development 138:811–820
27. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z et al (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 448:714–71

28. Zhang Y, Jurkowska R, Soeroes S, Rajavelu A, Dhayalan A, Bock I et al (2010) Chromatin methylation activity of Dnmt3a and Dnmt3a/3 L is guided by interaction of the ADD domain with the histone H3 tail. Nucleic Acids Res 38:4246–4253
29. Ciccone DN, Su H, Hevi S, Gay F, Lei H, Bajko J et al (2009) KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. Nature 461:415–418
30. Chotalia M, Smallwood SA, Ruf N, Dawson C, Lucifero D, Frontera M et al (2009) Transcription is required for establishment of germline methylation marks at imprinted genes. Genes Dev 23:105–117
31. Murdoch S, Djuric U, Mazhar B, Seoud M, Khan R, Kuick R et al (2006) Mutations in NALP7 cause recurrent hydatidiform moles and reproductive wastage in humans. Nat Genet 38:300–302
32. Meyer E, Lim D, Pasha S, Tee LJ, Rahman F, Yates JRW et al (2009) Germline mutation in NLRP2 (NALP2) in a familial imprinting disorder (Beckwith-Wiedemann Syndrome). PLoS Genet 5:e1000423
33. Davis TL, Yang GJ, McCarrey JR, Bartolomei MS (2000) The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. Hum Mol Genet 9:2885–2894
34. Ariel M, Cedar H, McCarrey J (1994) Developmental changes in methylation of spermatogenesis-specific genes include reprogramming in the epididymis. Nat Genet 7:59–63
35. Oakes CC, La Salle S, Smiraglia DJ, Robaire B, Trasler JM (2007) Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. Dev Biol 307:368–379
36. Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y et al (2007) Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. Hum Mol Genet 16:2272–2280
37. Kota SK, Feil R (2010) Epigenetic transitions in germ cell development and meiosis. Dev Cell 19:675–686
38. Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF et al (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell 31:785–799
39. Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Takamatsu K, Chuma S, Kojima-Kita K et al (2010) DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev 22:908–917
40. Watanabe T, Tomizawa S, Mitsuya K, Totoki Y, Yamamoto Y, Kuramochi-Miyagawa S et al (2011) Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse Rasgrf1 locus. Science 332:848–852
41. Borgel J, Guibert S, Li Y, Chiba H, Schübeler D, Sasaki H et al (2010) Targets and dynamics of promoter DNA methylation during early mouse development. Nat Genet 42:1093–1100
42. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M et al (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 39:457–466
43. Schulz R, Proudhon C, Bestor TH, Woodfine K, Lin CS, Lin SP et al (2010) The parental non-equivalence of imprinting control regions during mammalian development and evolution. PLoS Genet 6:e1001214
44. Flanagan JM, Popendikyte V, Pozdniakovaite N, Sobolev M, Assadzadeh A, Schumacher A et al (2006) Intra- and interindividual epigenetic variation in human germ cells. Am J Hum Genet 79:67–84
45. Kobayashi H, Hiura H, John RM, Sato A, Otsu E, Kobayashi N et al (2009) DNA methylation errors at imprinted loci after assisted conception originate in the parental sperm. Eur J Hum Genet 17:1582–1591
46. Poplinski A, Tüttelmann F, Kanber D, Horsthemke B, Gromoll J (2010) Idiopathic male infertility is strongly associated with aberrant methylation of MEST and IGF2/H19 ICR1. Int J Androl 33:642–649

47. Kimmins S, Sassone-Corsi P (2005) Chromatin remodelling and epigenetic features of germ cells. Nature 434:583–589
48. Hayashi K, Yoshida K, Matsui Y (2005) A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature 438:374–378
49. Tachibana M, Nozaki M, Takeda N, Shinkai Y (2007) Functional dynamics of H3K9 methylation during meiotic prophase progression. EMBO J 26:3346–3359
50. Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schöfer C et al (2001) Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell 107:323–337
51. Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR (2009) Distinctive chromatin in human sperm packages genes for embryo development. Nature 460:473–478
52. Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC et al (2010) Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat Struct Mol Biol 17:679–687
53. Carone BR, Fauquier L, Habib N, Shea JM, Hart CE, Li R et al (2010) Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. Cell 143:1084–1096
54. Santos F, Peters AH, Otte AP, Reik W, Dean W (2005) Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. Dev Biol 280:225–236
55. Puschendorf M, Terranova R, Boutsma E, Mao X, Isono K, Brykczynska U et al (2008) PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. Nat Genet 40:411–420
56. Terranova R, Yokobayashi S, Stadler MB, Otte AP, van Lohuizen M, Orkin SH et al (2008) Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. Dev Cell 15:668–679
57. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T (2000) Demethylation of the zygotic paternal genome. Nature 403:501–502
58. Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R et al (2000) Active demethylation of the paternal genome in the mouse zygote. Curr Biol 10:475–478
59. Santos F, Hendrich B, Reik W, Dean W (2002) Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol 241:172–182
60. Okada Y, Yamagata K, Hong K, Wakayama T, Zhang Y (2010) A role for the elongator complex in zygotic paternal genome demethylation. Nature 463:554–558
61. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324:930–935
62. Iqbal K, Jin SG, Pfeifer GP, Szabó PE (2011) Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. Proc Natl Acad Sci USA 108:3642–3647
63. Inoue A, Zhang Y (2011) Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science 334:194 doi: 10.1126/science.1212483
64. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324:929–930
65. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466:1129–1133
66. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA et al (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature 473:398–402
67. Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J et al (2011) Tet1 and tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell 8:200–213
68. Nakamura T, Arai Y, Umehara H, Masuhara M, Kimura T, Taniguchi H et al (2007) PGC7/Stella protects against DNA demethylation in early embryogenesis. Nat Cell Biol 9:64–71

69. Howell CY, Bestor TH, Ding F, Latham KE, Mertineit C, Trasler JM et al (2001) Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. Cell 104:829–838
70. Ratnam S, Mertineit C, Ding F, Howell CY, Clarke HJ, Bestor TH et al (2002) Dynamics of Dnmt1 methyltransferase expression and intracellular localization during oogenesis and preimplantation development. Dev Biol 245:304–314
71. Hirasawa R, Chiba H, Kaneda M, Tajima S, Li E, Jaenisch R et al (2008) Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. Genes Dev 15:1607–1616
72. Mohan KN, Ding F, Chaillet JR (2011) Distinct roles of DMAP1 in mouse development. Mol Cell Biol 31:1861–1869
73. Li X, Ito M, Zhou F, Youngson N, Zuo X, Leder P et al (2008) A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. Dev Cell 15:547–557
74. Mackay DJ, Callaway JL, Marks SM, White HE, Acerini CL, Boonen SE et al (2008) Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. Nat Genet 40:949–951
75. Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, Aktas T et al (2010) KAP1 controls endogenous retroviruses in embryonic stem cells. Nature 463:237–240
76. Chapman V, Forrester L, Sanford J, Hastie N, Rossant J (1984) Cell lineage-specific undermethylation of mouse repetitive DNA. Nature 307:284–286
77. Sarmento OF, Digilio LC, Wang Y, Perlin J, Herr JC, Allis CD et al (2004) Dynamic alterations of specific histone modifications during early murine development. J Cell Sci 117:4449–4459
78. Torres-Padilla ME, Parfitt DE, Kouzarides T, Zernicka-Goetz M (2007) Histone arginine methylation regulates pluripotency in the early mouse embryo. Nature 445:214–218
79. Parfitt DE, Zernicka-Goetz M (2010) Epigenetic modification affecting expression of cell polarity and cell fate genes to regulate lineage specification in the early mouse embryo. Mol Biol Cell 21:2649–2660
80. Jedrusik A, Parfitt DE, Guo G et al (2008) Role of Cdx2 and cell polarity in cell allocation and specification of trophectoderm and inner cell mass in the mouse embryo. Genes Dev 22:2692–2706
81. Ng RK, Dean W, Dawson C, Lucifero D, Madeja Z, Reik W et al (2008) Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. Nat Cell Biol 10:1280–1290
82. Hayashi K, Lopes SM, Tang F, Surani MA (2008) Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell Stem Cell 3:391–401
83. Ji H, Ehrlich LI, Seita J, Murakami P, Doi A, Lindau P et al (2010) Comprehensive methylome map of lineage commitment from haematopoietic progenitors. Nature 467:338–342
84. Satterlee JS, Schübeler D, Ng HH (2010) Tackling the epigenome: challenges and opportunities for collaboration. Nat Biotechnol 28:1039–1044
85. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J et al (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125:315–326
86. Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P et al (2008) Dissecting direct reprogramming through integrative genomic analysis. Nature 454:49–55
87. Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G et al (2011) Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature 471:68–73
88. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S et al (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 40:897–903
89. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF et al (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 459:108–112
90. Mikkelsen TS, Xu Z, Zhang X, Wang L, Gimble JM, Lander ES et al (2010) Comparative epigenomic analysis of murine and human adipogenesis. Cell 143:156–169

91. Illingworth RS, Bird AP (2009) CpG islands–'a rough guide'. FEBS Lett 583:1713–1120
92. Illingworth R, Kerr A, Desousa D, Jørgensen H, Ellis P, Stalker J et al (2008) A novel CpG island set identifies tissue-specific methylation at developmental gene loci. PLoS Biol 6:e22
93. Illingworth RS, Gruenewald-Schneider U, Webb S, Kerr AR, James KD, Turner DJ et al (2010) Orphan CpG islands identify numerous conserved promoters in the mammalian genome. PLoS Genet 6:e1001134
94. Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD et al (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 466:253–257
95. Irizarry RA, Ladd-Acosta C, Wen B et al (2008) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 41:178–186
96. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A et al (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454:766–770
97. Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, Webb S et al (2010) CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature 464:1082–1086
98. Blackledge NP, Zhou JC, Tolstorukov MY, Farcas AM, Park PJ, Klose RJ (2010) CpG islands recruit a histone H3 lysine 36 demethylase. Mol Cell 38:179–190
99. Hernandez-Munoz I, Taghavi P, Kuijl C, Neefjes J, van Lohuizen M (2005) Association of BMI1 with polycomb bodies is dynamic and requires PRC2/EZH2 and the maintenance DNA methyltransferase DNMT1. Mol Cell Biol 25:11047–11058
100. Viré E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C et al (2006) The Polycomb group protein EZH2 directly controls DNA methylation. Nature 439:871–874
101. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L et al (2007) A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 39:237–242
102. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J et al (2007) Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet 39:232–236
103. Bostick M, Kim JK, Estève PO, Clark A, Pradhan S, Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science 317:1760–1764
104. Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA et al (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 450:908–912
105. Margueron R, Reinberg D (2010) Chromatin structure and the inheritance of epigenetic information. Nat Rev Genet 11:285–296
106. Espada J, Ballestar E, Fraga MF, Villar-Garea A, Juarranz A, Stockert JC et al (2004) Human DNA methyltransferase 1 is required for maintenance of the histone H3 modification pattern. J Biol Chem 279:37175–37184
107. Kacem S, Feil R (2009) Chromatin mechanisms in genomic imprinting. Mamm Genome 20:544–556
108. Tucker KL, Beard C, Dausmann J, Jackson-Grusby L, Laird PW, Lei H et al (1996) Germline passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. Genes Dev 10:1008–1020
109. Wareham KA, Lyon MF, Glenister PH, Williams ED (1987) Age-related reactivation of an X-linked gene. Nature 327:725–727
110. Bennett-Baker PE, Wilkowski J, Burke DT (2003) Age-associated activation of epigenetically repressed genes in the mouse. Genetics 165:2055–2062
111. Carrel L, Willard HF (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434:400–404
112. Fu VX, Dobosy JR, Desotelle JA, Almassi N, Ewald JA, Srinivasan R et al (2008) Aging and cancer-related loss of insulin-like growth factor 2 imprinting in the mouse and human prostate. Cancer Res 68:6797–6802

113. Sandovici I, Leppert M, Hawk PR, Suarez A, Linares Y, Sapienza C (2003) Familial aggregation of abnormal methylation of parental alleles at the IGF2/H19 and IGF2R differentially methylated regions. Hum Mol Genet 12:1569–1578
114. Schneider E, Pliushch G, El Hajj N, Galetzka D, Puhl A, Schorsch M et al (2010) Spatial, temporal and interindividual epigenetic variation of functionally important DNA methylation patterns. Nucleic Acids Res 38:3880–3890
115. Frost JM, Monk D, Stojilkovic-Mikic T, Woodfine K, Chitty LS, Murrell A et al (2010) Evaluation of allelic expression of imprinted genes in adult human blood. PLoS One 5:e13556
116. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML et al (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 102:10604–10609
117. Grönniger E, Weber B, Heil O, Peters N, Stäb F, Wenck H et al (2010) Aging and chronic sun exposure cause distinct epigenetic changes in human skin. PLoS Genet 6:e1000971
118. Daxinger L, Whitelaw E (2010) Transgenerational epigenetic inheritance: more questions than answers. Genome Res 20:1623–1628
119. Skinner MK, Manikkam M, Guerrero-Bosagna C (2011) Epigenetic transgenerational actions of endocrine disruptors. Reprod Toxicol 31:337–343
120. Morgan HD, Sutherland HG, Martin DI, Whitelaw E (1999) Epigenetic inheritance at the agouti locus in the mouse. Nat Genet 23:314–318
121. Blewitt ME, Vickaryous NK, Paldi A, Koseki H, Whitelaw E (2006) Dynamic reprogramming of DNA methylation at an epigenetically sensitive allele in mice. PLoS Genet 2:e49
122. Anway MD, Cupp AS, Uzumcu M, Skinner MK (2005) Epigenetic transgenerational actions of endocrine disruptors and male fertility. Science 308:1466–1469
123. Guerrero-Bosagna C, Settles M, Lucker B, Skinner MK (2010) Epigenetic transgenerational actions of vinclozolin on promoter regions of the sperm epigenome. PLoS One 5:e13100
124. Kaati G, Bygren LO, Edvinsson S (2002) Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period. Eur J Hum Genet 10:682–688
125. Pembrey ME, Bygren LO, Kaati G, Edvinsson S, Northstone K, Sjöström M et al (2006) Sex-specific, male-line transgenerational responses in humans. Eur J Hum Genet 14:159–166
126. Teperino R, Schoonjans K, Auwerx J (2010) Histone methyl transferases and demethylases; can they link metabolism and transcription? Cell Metab 1:321–327
127. Waterland RA, Jirtle RL (2003) Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol 23:5293–5300
128. Cropley JE, Suter CM, Beckman KB, Martin DI (2006) Germ-line epigenetic modification of the murine Avy allele by nutritional supplementation. Proc Natl Acad Sci USA 103:17308–17312
129. Waterland RA, Travisano M, Tahiliani KG (2007) Diet-induced hypermethylation at agouti viable yellow is not inherited transgenerationally through the female. FASEB J 21:3380–3385
130. Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M et al (2010) Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. PLoS Genet 6:e1001252
131. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES et al (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci USA 105:17046–17049
132. Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD et al (2009) DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Hum Mol Genet 18:4046–4053
133. Anway MD, Rekow SS, Skinner MK (2008) Transgenerational epigenetic programming of the embryonic testis transcriptome. Genomics 91:30–40
134. Richards EJ (2009) Quantitative epigenetics: DNA sequence variation need not apply. Genes Dev 23:1601–1605

Chapter 7
The Role of Epigenetics in the Developmental Origins of Health and Disease

Karin B. Michels and Robert A. Waterland

Contents

7.1 The Developmental Origins of Health and Disease 106
7.2 DOHaD Mechanisms 107
7.3 Potential Critical Periods for Developmental Epigenetics 107
7.4 First Clues from the Agouti Mouse 110
7.5 Epigenetic Epidemiology Data on DOHaD 110
7.6 Intrauterine Exposures and Transgenerational Epigenetic Inheritance 111
7.7 Outlook 112
7.8 Conclusions 112
References 113

Abstract The Developmental Origins of Health and Disease (DOHaD) hypothesis suggests that the prenatal and early postnatal environments shape the future probability of physical and mental wellbeing and risk of disease. A wealth of epidemiologic data supports the associations between maternal malnutrition, intrauterine growth retardation, and birthweight, and the risk of chronic disease including cardiovascular disease, hypertension, type 2 diabetes mellitus, obesity, neuropsychiatric

K.B. Michels (✉)
Obstetrics and Gynecology Epidemiology Center, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Avenue, Boston, MA 02115, USA

Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA
e-mail: kmichels@rics.bwh.harvard.edu

R.A. Waterland
Department of Pediatrics, USDA/ARS Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX 77030, USA

Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA
e-mail: waterland@bcm.edu

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_7,
© Springer Science+Business Media B.V. 2012

disorders, and cancer. While the mechanisms underlying these observations remain unresolved, the DOHaD model assumes a developmental plasticity that allows adaptive regulation of the embryonic, fetal, and/or early postnatal metabolism in response to nutritional and environmental perturbations. Establishment of the epigenome coincides with vulnerable phases in development and provides one potential mechanism for long-lasting responses to transient environmental stimuli. Future studies in epigenetic epidemiology will seek to understand the role of various epigenetic mechanisms in DOHaD.

7.1 The Developmental Origins of Health and Disease

The search for the origins of chronic diseases has shifted focus toward the earliest phases of the life course. While classic epidemiology has targeted lifestyle patterns of adults at various ages, the importance of early life for lifelong health has been increasingly recognized over the past decades. Following the seminal work of Rose [1], Forsdahl [2], and Barker [3, 4], the period from conception to birth as well as the first few years of life are considered critical in the imprinting of disease susceptibility. This shift in thinking and research gave birth to the "Developmental Origins of Health and Disease" (DOHaD) hypothesis.

Epidemiologic studies support the hypothesis that chronic diseases have their roots in early life. Barker's work linked low birthweight to a number of cardiovascular diseases (including ischemic heart disease), hypertension, cholesterol levels, stroke, and impaired glucose tolerance [4–8]. His findings have been confirmed by other groups in different populations [9–11]. Data from the Dutch Famine in 1944/45, when food rations dropped below 1,000 kcal/day for 6 months, suggest an increased risk of obesity among offspring of mothers exposed to the famine during the first and second trimester [12], glucose intolerance if exposure peaked during late gestation [13], and schizophrenia if conception occurred during the famine [14]. Other maternal characteristics such as maternal weight and malnutrition also increase the risk of coronary heart disease in the offspring [15].

However, there is trouble at both ends of the birthweight spectrum. Like low birthweight, high birthweight is also associated with adult obesity [16]. Similarly, women are more likely to become obese in adulthood if their mothers were obese prior to pregnancy and/or had very high or very low gestational weight gain [17]. Gestational diabetes increases the risk of childhood and adult obesity in the offspring [18]. Moreover, a high birthweight is associated with an elevated risk of several cancers. Numerous epidemiologic studies support the association between a high birthweight and an increased risk of premenopausal breast cancer [19, 20]. In addition, a high birthweight has been linked to childhood leukemia [21], childhood brain tumors [22], and testicular cancer [23].

7.2 DOHaD Mechanisms

The fetal origins hypothesis suggests that perturbations at a critical period of development induce persistent alterations with potentially lifelong consequences. These epidemiologic observations led Hales and Barker to suggest a "thrifty phenotype" hypothesis, which proposes that poor fetal nutrition and growth leads to metabolic reprogramming of the glucose-insulin metabolism [24]. This adaptive developmental plasticity allows the fetus to adjust to and survive adverse environments. A limited supply of transplacental nutrients compels the fetus to channel the nutrients to the most vital organs, namely brain and heart, at the expense of other organs, which may remain underdeveloped and compromised in growth and function [25]. Moreover, permanent insulin resistance may be induced during development and reduce basal metabolic requirements; this permits survival under sub-optimal prenatal and predicted postnatal conditions [26]. While the environmental conditions *in utero* are transient, they induce permanent changes in metabolism and alter chronic disease susceptibility. Indeed, environmental perturbations may have a long-lasting impact at times of greatest plasticity during growth and development.

The potential benefits of such so-called 'predictive adaptive responses' depend on the accuracy of the prediction; the cost of inaccurate predictions is high [26]. If developmental conditions that induce intrauterine growth retardation are followed by a resource-rich postnatal environment, high plasma glucose levels will coincide with insulin resistance, greatly increasing the risk for metabolic disease in later life [26]. This "mismatch" between predicted and actual postnatal environment may explain profound long-lasting implications for chronic disease among individuals prenatally exposed to the Dutch Hunger Winter, which lasted only 9 months and was followed by normal nutritional availability [27]. Individual variation in sensitivity to mismatch and in turn to disease susceptibility is likely due to a variety of factors including genetic variation and the degree of developmental plasticity [25].

7.3 Potential Critical Periods for Developmental Epigenetics

The molecular mechanisms underlying DOHaD are not well understood. Developmental plasticity allows a specific genotype to create alternate phenotypes depending on intrauterine and early postnatal conditions, which may induce lasting changes in chronic disease susceptibility. Among the various potential biologic mechanisms underlying developmental plasticity [28], environmental influences on developmental epigenetics are receiving increasing attention [29]. Epigenetics describes the study of non-genetic mitotically heritable states in gene expression potential [30]. Epigenetic regulation is achieved by a layer of marks added onto the DNA that govern selective access to the genetic information and are perpetuated in differentiated tissues. The specific molecular mechanisms that function interactively

Fig. 7.1 Sources of interindividual variation in the epigenome. Environmental influences on the epigenome are likely most important during establishment of the epigenetic marks in prenatal and early postnatal development. (Reprinted with permission from [31])

to heritably regulate chromatin conformation include DNA methylation, various modifications of the histone proteins that package DNA in the nucleus, and auto-regulatory DNA binding proteins [31]. Interindividual variation in epigenetic states provide a basis for epidemiologic studies (Fig. 7.1).

Many studies of epigenetics in DOHaD have focused on imprinted gene regions. Genomic imprinting is the epigenetic silencing of one allele of specific genes by DNA methylation in a parent-of-origin-specific manner. Imprinted genes normally transcribe and express either the maternal or the paternal allele, while the remaining allele is silenced by DNA methylation of CpG islands in the promoter region. Loss of imprinting results in the aberrant biallelic expression of an imprinted gene. In humans, approximately 100 imprinted genes have been identified. Since most imprinted genes play a role in intrauterine and early-life growth, they have been proposed as good candidates to translate early nutritional and environmental influences into fetal development. Moreover, loss of imprinting of fetal growth genes, in particular *IGF2*, has been associated with childhood disorders such as Beckwith-Wiedemann Syndrome [32, 33], Silver-Russell Syndrome [34], and Wilms' Tumor [35, 36], as well as with adult-onset diseases [37, 38]. However, imprinted genes have not convincingly been found to be particularly susceptible to early developmental influences in the few informative human studies that have been conducted [39].

Importantly, the epigenome is established at crucial developmental time points that coincide with vulnerable periods of adaptive plasticity. In the mouse model, the parental epigenome is largely erased during gametogenesis, followed by epigenetic reprogramming as primordial germ cells differentiate into oocyte and sperm [40]. After fertilization, the non-imprinted gene regions in the zygotic genome undergo another round of epigenetic reprogramming that restores totipotency. This permits cell fate commitment of the first cell lineages during the preimplantation phase facilitated by a genome-wide *de novo* methylation [41–44] (discussed in more detail in Chap. 6). These dramatic waves of epigenetic reprogramming likely make early embryonic

development a critical period during which nutritional, environmental, and metabolic factors affect the developmental establishment of epigenetic regulation.

To understand the role of epigenetic mechanisms in DOHaD, defining the window of susceptibility is crucial. Based on the mouse model, the time point of *de novo* methylation differs for imprinted and non-imprinted genes and for the female and male germline [42] (discussed in more detail in Chap. 6). The DNA methylation signature of non-imprinted genes may be most amenable to environmental stimuli after fertilization, when the totipotent blastocyst becomes remethylated prior to implantation. Nutritional and metabolic factors early during the first trimester therefore may have the greatest potential to disrupt the introduction of cytosine methylation. The timing of remethylation of imprinted genes is less clear. By extrapolation from the mouse model, cytosine methylation of the differentially methylated regions (DMRs) of one of the two parental chromosomes is established at different time points for different imprinted genes [45–47]. In the mouse, maternal imprints are established at some point between oocyte development and ovulation [47], and paternal imprints are completed by the time spermatocytes enter meiosis [48]. Whether the establishment of imprinting marks is similar in humans remains to be established.

While stimuli in the periconceptional environment may affect several tissues if epimutations emerge prior to cell differentiation, perturbations during late gestation are more likely to induce tissue-specific epigenetic changes [31]. Further, epigenetic development is not limited to prenatal life; for example, the early postnatal period appears to be a critical period for establishment of DNA methylation in the brain [49].

We have previously proposed two mechanisms to explain environmental influences on the developmental establishment of DNA methylation [31]. First, an imbalance in dietary methyl donors and/or activity of DNA methyltransferases may affect developmental establishment of DNA methylation. Metastable epialleles show large interindividual differences in DNA methylation and gene regulation and appear particularly labile in response to environmental stimuli during developmental establishment of the epigenome. While most transposable elements in the mammalian genome are silenced by CpG methylation [50, 51], some are metastable and can also affect expression of neighboring genes [52].

Secondly, nutritional or environmental stimuli may alter transcriptional activity during periods of *de novo* DNA methylation, which may permanently alter epigenetic regulation and corresponding phenotypes. Genes actively transcribed during *de novo* methylation are protected from methylation and remain hypomethylated [53]. Interference with active transcription renders these promoters susceptible to *de novo* methylation and alters their function [54].

The placenta's critical role in nutrient transfer from mother to fetus makes it particularly vulnerable to adverse intrauterine conditions. Whereas induced epigenetic changes in the soma persist to influence later phenotype, maternal nutrition may also induce epigenetic changes in the placenta, affecting nutrient transport and fetal growth [25]. Imprinted genes are highly expressed in the placenta, which may make them vulnerable to variation in maternal nutrition [55, 56].

7.4 First Clues from the Agouti Mouse

Seminal experiments with the agouti viable yellow (A^{vy}) mouse model support the importance of maternal nutrition for developmental epigenetics [57]. The *agouti* gene codes for yellow pigment in fur. Transposition of an IAP retrotransposon upstream of *agouti* resulted in the A^{vy} metastable epiallele. DNA methylation of the retrotransposon exhibits spontaneous interindividual variation, controlling expression of the *agouti* gene and therefore the coat color of A^{vy} mice. Moreover, supplementation of mouse dams during pregnancy with the dietary methyl donors and cofactors folic acid, vitamin B12, betaine, and choline shifts the coat color distribution of A^{vy} heterozygous offspring from yellow to brown [58] by increasing DNA methylation at the A^{vy} locus [57].

Similarly, supplementation of dams with the phytoestrogen genistein resulted in an analogous coat color shift also mediated through A^{vy} hypermethylation [59]. Maternal methyl donor supplementation studies in another murine metastable epiallele model, the *axin fused* mouse, essentially corroborated the findings in the A^{vy} model [60], indicating that epigenetic regulation at metastable epialleles is generally susceptible to early environmental influences. Putative metastable epialleles were recently identified in humans [61]; as in the mouse models, these human loci showed dramatic and systemic interindividual epigenetic variation that was influenced by maternal environment around the time of conception.

7.5 Epigenetic Epidemiology Data on DOHaD

Data from the Dutch Famine suggest that individuals prenatally exposed to the famine had somewhat lower DNA methylation of the IGF2 genes six decades later compared to their unexposed siblings [62]. In this study of 60 same-sex sib pairs, the authors examined five CpGs in the *IGF2*DMR0 and found in average 2.7% lower methylation among individuals exposed to the famine *in utero*. Whether this small difference in methylation has any functional consequence remains unclear, in particular since the authors did not examine IGF2 expression levels. Methylation differences of even smaller magnitude were observed for some other genes including *IL10, GNASAS, INSIGF, LEP*, and *MEG3* [63].

The association between an epigenetic difference assessed in adulthood and a prenatal exposure does not allow causal inference about the induction of that change by the prenatal factor, unless the change is already present directly after the exposure period [28, 31]. Of course, collecting relevant samples in humans to test such concepts is logistically challenging.

A few studies have considered prenatal exposures and gene-specific DNA methylation at birth. Self-reported high maternal folic acid supplementation (defined as an average 1,200 μg/day or more) was associated with 2.4% lower methylation at the *H19* differentially methylated region (DMR) in umbilical cord blood leukocytes if initiated before pregnancy and 3.7% less methylation at that locus if initiated

during pregnancy, compared to non-users; no difference was found for the *IGF2*DMR0 [64]. A study from the Netherlands reported a 4% higher methylation level in one CpG of *IGF2*DMR0 in the blood of 17-month-old children whose mothers took 400 μg folic acid during pregnancy compared to those whose mothers took no folic acid supplements; no difference was found for four other CpGs, however, and *IGF2* expression levels were not examined [65]. Prenatal exposure to tobacco smoke was associated with marginally lower methylation in AluYb8b retrotransposons (but no difference in LINE-1 methylation) and slightly higher methylation in *AXL* and *PTPRO* in young children [66]. Higher maternal bone lead, but not cord blood lead levels, were associated with marginally lower methylation at *Alu* retrotransposons in cord blood DNA of Mexican infants [67].

A number of studies have explored the epigenetic profile of newborns with low birthweight or intrauterine growth retardation (IUGR). Using a DNA methylation microarray approach, Einstein and colleagues compared cord blood samples from five IUGR and five normal pregnancies and identified methylation differences at a restricted number of loci [68]. A few small studies identified differences in methylation or expression of selected imprinted genes in placenta and cord blood of IUGR or low birthweight compared to normal-weight infants [69–72] and in selected non-imprinted genes [73–75]. High birthweight has been associated with increased promoter methylation of the glucocorticoid receptor gene in human placenta [68–72, 76]. Overall, differences in methylation in these studies were small.

In one recent study, CpG methylation of 68 CpGs in five candidate genes was assessed in umbilical cord tissue from healthy neonates in two prospective cohorts [77]. DNA methylation of one CpG was consistently associated with adiposity at age 9 years but statistical analyses were not adjusted for multiple comparisons.

In summary, it remains unclear whether small differences in DNA methylation at birth associated with intrauterine exposures have functional relevance and are maintained into adulthood. No data are available on perinatal perturbation and epigenetic modifications, which translate into adult disease. Whether the intrauterine environment affects adult disease susceptibility in humans through epigenetic mechanisms remains to be established.

7.6 Intrauterine Exposures and Transgenerational Epigenetic Inheritance

DOHaD is mostly concerned with intrauterine experiences that affect adult disease susceptibility. Much interest in epigenetics has centered around the possibility of transmitting effects of environmental exposures across generations [78]. There has been considerable confusion in the literature about the difference between intrauterine exposures that affect the epigenetic profile in the offspring and transgenerational epigenetic inheritance [79, 80]. If environmentally induced epigenetic changes are inherited transgenerationally, then by definition the putatively inherited epigenetic change must be present in both the contributed gametes and the soma of the offspring.

For epigenetic marks to be inherited, they would have to escape erasure during developmental reprogramming of the epigenome. Such incomplete erasure has been described for epigenetic silencing of the A^{vy} allele [81]. IAP transposons seem to resist developmental reprogramming in the mouse, and their influence on the expression state of neighboring genes may be passed on to subsequent generations [82]. Genomic imprinting is another example of transgenerational epigenetic inheritance in mammals including humans. Following clearance of imprinting marks in primordial germ cells, parents-of-origin-specific allelic methylation and expression of imprinted genes is transmitted from parent to offspring in every generation. The molecular marks insuring such memory remain to be identified. It has been suggested that histone modifications or other marks, perhaps microRNAs, at imprinted gene loci are only incompletely erased and provide signals for *de novo* methylation at DMRs [40].

Whether transgenerational inheritance of acquired epigenetic changes exists in humans remains elusive. These issues are discussed in more detail in Chap. 6.

7.7 Outlook

Future epidemiologic studies exploring the role of epigenetics in DOHaD may target the effect of maternal nutrition on the epigenome of the newborn. Furthermore, studies exploring the time periods of DMR establishment in human gametes will aid our understanding of the window of susceptibility. The NIH Roadmap Epigenomics project and other international epigenome mapping projects will provide important insights into the establishment of the epigenome in humans [83].

Studies on the epigenetic epidemiology of DOHaD have to rely on loci with large interindividual epigenetic variability in easily obtainable tissues of physiologic relevance, preferably at various ages [31]. Metastable epialleles are candidates because their epimutations make them particularly labile to environmental influences during development, accommodating change and contributing to plasticity. Since developmental establishment of DNA methylation at metastable epialleles occurs probabilistically, interindividual variation at these loci is particularly large and thus informative in epidemiologic studies [61]. Moreover, metastable epialleles are less tissue-specific than other epigenetic marks, making them logistically easier to study [57, 60].

7.8 Conclusions

The concept of DOHaD arose from epidemiologic studies. Developmental plasticity allows adaptation of the fetal metabolism to transient nutritional and environmental experiences, which can result in lasting changes in chronic disease susceptibility. While our understanding of the underlying mechanisms is rudimentary, alterations

in the epigenome are likely contributors. We must emphasize, however, that epigenetics is only one of several potential mechanisms explaining developmental plasticity [28]. A better understanding of the mechanisms contributing to DOHaD may make it possible to reduce risk of disease by both preventive strategies targeted at early life, and corrective approaches designed to normalize malleable molecular mechanisms set askew by early adverse exposures.

References

1. Rose G (1964) Familial patterns in ischaemic heart disease. Br J Prev Soc Med 18:75–80
2. Forsdahl A (1977) Are poor living conditions in childhood and adolescence an important risk factor for arteriosclerotic heart disease? Br J Prev Soc Med 31:91–95
3. Barker DJ (1995) Fetal origins of coronary heart disease. BMJ 311:171–174
4. Barker DJ, Winter PD, Osmond C, Margetts B, Simmonds SJ (1989) Weight in infancy and death from ischaemic heart disease. Lancet 2:577–580
5. Barker DJ (1997) Intrauterine programming of coronary heart disease and stroke. Acta Paediatr Suppl 423:178–182, discussion 183
6. Barker DJ, Bull AR, Osmond C, Simmonds SJ (1990) Fetal and placental size and risk of hypertension in adult life. BMJ 301:259–262
7. Barker DJ, Osmond C (1986) Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. Lancet 1:1077–1081
8. Hales CN, Barker DJ, Clark PM, Cox LJ, Fall C, Osmond C et al (1991) Fetal and infant growth and impaired glucose tolerance at age 64. BMJ 303:1019–1022
9. Curhan GC, Chertow GM, Willett WC, Spiegelman D, Colditz GA, Manson JE et al (1996) Birth weight and adult hypertension and obesity in women. Circulation 94:1310–1315
10. Rich-Edwards JW, Stampfer MJ, Manson JE, Rosner B, Hankinson SE, Colditz GA et al (1997) Birth weight and risk of cardiovascular disease in a cohort of women followed up since 1976. BMJ 315:396–400
11. Andersson SW, Lapidus L, Niklasson A, Hallberg L, Bengtsson C, Hulthen L (2000) Blood pressure and hypertension in middle-aged women in relation to weight and length at birth: a follow-up study. J Hypertens 18:1753–1761
12. Ravelli GP, Stein ZA, Susser MW (1976) Obesity in young men after famine exposure in utero and early infancy. N Engl J Med 295:349–353
13. Ravelli AC, van der Meulen JH, Michels RP, Osmond C, Barker DJ, Hales CN et al (1998) Glucose tolerance in adults after prenatal exposure to famine. Lancet 351:173–177
14. Susser ES, Lin SP (1992) Schizophrenia after prenatal exposure to the Dutch Hunger Winter of 1944–1945. Arch Gen Psychiatry 49:983–988
15. Forsen T, Eriksson JG, Tuomilehto J, Teramo K, Osmond C, Barker DJ (1997) Mother's weight in pregnancy and coronary heart disease in a cohort of Finnish men: follow up study. BMJ 315:837–840
16. Sorensen HT, Sabroe S, Rothman KJ, Gillman M, Fischer P, Sorensen TI (1997) Relation between weight and length at birth and body mass index in young adulthood: cohort study. BMJ 315:1137
17. Stuebe AM, Forman MR, Michels KB (2009) Maternal-recalled gestational weight gain, pre-pregnancy body mass index, and obesity in the daughter. Int J Obes (Lond) 33:743–752
18. Lawlor DA, Lichtenstein P, Langstrom N (2011) Association of maternal diabetes mellitus in pregnancy with offspring adiposity into early adulthood: sibling study in a prospective cohort of 280,866 men from 248,293 families. Circulation 123:258–265
19. Michels KB, Xue F (2006) Role of birthweight in the etiology of breast cancer. Int J Cancer 119:2007–2025

20. Michels KM, Trichopoulos D, Robins JM, Rosner B, Manson JE, Hunter DJ et al (1996) Birthweight as a risk factor for breast cancer. Lancet 348:1542–1546
21. Caughey RW, Michels KB (2009) Birth weight and childhood leukemia: a meta-analysis and review of the current evidence. Int J Cancer 124:2658–2670
22. Harder T, Plagemann A, Harder A (2008) Birth weight and subsequent risk of childhood primary brain tumors: a meta-analysis. Am J Epidemiol 168:366–373
23. Michos A, Xue F, Michels KB (2007) Birth weight and the risk of testicular cancer: a meta-analysis. Int J Cancer 121:1123–1131
24. Hales CN, Barker DJ (2001) The thrifty phenotype hypothesis. Br Med Bull 60:5–20
25. Hochberg Z, Feil R, Constancia M, Fraga M, Junien C, Carel JC et al (2011) Child health, developmental plasticity, and epigenetic programming. Endocr Rev 32:159–224
26. Rickard IJ, Lummaa V (2007) The predictive adaptive response and metabolic syndrome: challenges for the hypothesis. Trends Endocrinol Metab 18:94–99
27. Gluckman P, Hanson M (2006) Mismatch. Why our world no longer fits our bodies. Oxford University Press, Oxford
28. Waterland RA, Garza C (1999) Potential mechanisms of metabolic imprinting that lead to chronic disease. Am J Clin Nutr 69:179–197
29. Waterland RA, Jirtle RL (2004) Early nutrition, epigenetic changes at transposons and imprinted genes, and enhanced susceptibility to adult chronic diseases. Nutrition 20:63–68
30. Jaenisch R, Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 33(Suppl):245–254
31. Waterland RA, Michels KB (2007) Epigenetic epidemiology of the developmental origins hypothesis. Annu Rev Nutr 27:363–388
32. Ohlsson R, Nystrom A, Pfeifer-Ohlsson S, Tohonen V, Hedborg F, Schofield P et al (1993) IGF2 is parentally imprinted during human embryogenesis and in the Beckwith-Wiedemann syndrome. Nat Genet 4:94–97
33. Weksberg R, Shen DR, Fei YL, Song QL, Squire J (1993) Disruption of insulin-like growth factor 2 imprinting in Beckwith-Wiedemann syndrome. Nat Genet 5:143–150
34. Gicquel C, Rossignol S, Cabrol S, Houang M, Steunou V, Barbu V et al (2005) Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. Nat Genet 37:1003–1007
35. Ogawa O, Eccles MR, Szeto J, McNoe LA, Yun K, Maw MA et al (1993) Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. Nature 362:749–751
36. Steenman MJ, Rainier S, Dobry CJ, Grundy P, Horon IL, Feinberg AP (1994) Loss of imprinting of IGF2 is linked to reduced expression and abnormal methylation of H19 in Wilms' tumour. Nat Genet 7:433–439
37. Cui H, Cruz-Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S et al (2003) Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. Science 299: 1753–1755
38. Jelinic P, Shaw P (2007) Loss of imprinting and cancer. J Pathol 211:261–268
39. Waterland RA (2009) Early environmental effects on epigenetic regulation in humans. Epigenetics 4:523–525
40. Morgan HD, Santos F, Green K, Dean W, Reik W (2005) Epigenetic reprogramming in mammals. Hum Mol Genet 14 Spec No 1:R47–58
41. Razin A, Riggs AD (1980) DNA methylation and gene function. Science 210:604–610
42. Reik W, Dean W, Walter J (2001) Epigenetic reprogramming in mammalian development. Science 293:1089–1093
43. Rideout WM 3rd, Eggan K, Jaenisch R (2001) Nuclear cloning and epigenetic reprogramming of the genome. Science 293:1093–1098
44. Santos F, Dean W (2004) Epigenetic reprogramming during early development in mammals. Reproduction 127:643–651
45. Obata Y, Kono T (2002) Maternal primary imprinting is established at a specific time for each gene throughout oocyte growth. J Biol Chem 277:5285–5289
46. Lucifero D, Mann MR, Bartolomei MS, Trasler JM (2004) Gene-specific timing and epigenetic memory in oocyte imprinting. Hum Mol Genet 13:839–849

47. Tomizawa S, Kobayashi H, Watanabe T, Andrews S, Hata K, Kelsey G et al (2011) Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. Development 138:811–820
48. Allegrucci C, Thurston A, Lucas E, Young L (2005) Epigenetics and the germline. Reproduction 129:137–149
49. Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR et al (2004) Epigenetic programming by maternal behavior. Nat Neurosci 7:847–854
50. Yoder JA, Walsh CP, Bestor TH (1997) Cytosine methylation and the ecology of intragenomic parasites. Trends Genet 13:335–340
51. Yoder JA, Soman NS, Verdine GL, Bestor TH (1997) DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. J Mol Biol 270:385–395
52. Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E (2002) Metastable epialleles in mammals. Trends Genet 18:348–351
53. Bird A (2002) DNA methylation patterns and epigenetic memory. Genes Dev 16:6–21
54. Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A et al (1994) Sp1 elements protect a CpG island from de novo methylation. Nature 371:435–438
55. Fowden AL, Sibley C, Reik W, Constancia M (2006) Imprinted Genes, Placental Development and Fetal Growth. Horm Res Paediatr 65:50–58
56. Coan PM, Burton GJ, Ferguson-Smith AC (2005) Imprinted genes in the placenta - a review. Placenta 26:S10–S20
57. Waterland RA, Jirtle RL (2003) Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol 23:5293–5300
58. Wolff GL, Kodell RL, Moore SR, Cooney CA (1998) Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. FASEB J 12:949–957
59. Dolinoy DC, Weidman JR, Waterland RA, Jirtle RL (2006) Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. Environ Health Perspect 114:567–572
60. Waterland RA, Dolinoy DC, Lin JR, Smith CA, Shi X, Tahiliani KG (2006) Maternal methyl supplements increase offspring DNA methylation at Axin Fused. Genesis 44:401–406
61. Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M et al (2010) Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. PLoS Genet 6:e1001252
62. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES et al (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci USA 105:17046–17049
63. Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD et al (2009) DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Hum Mol Genet 18:4046–4053
64. Hoyo C, Murtha AP, Schildkraut JM, Jirtle RL, Demark-Wahnefried W, Forman MR et al (2011) Methylation variation at IGF2 differentially methylated regions and maternal folic acid use before and during pregnancy. Epigenetics 6:928–936
65. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, Lindemans J, Siebel C, Steegers EA et al (2009) Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child. PLoS One 4:e7845
66. Breton CV, Byun HM, Wenten M, Pan F, Yang A, Gilliland FD (2009) Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. Am J Respir Crit Care Med 180:462–467
67. Pilsner JR, Hu H, Ettinger A, Sanchez BN, Wright RO, Cantonwine D et al (2009) Influence of prenatal lead exposure on genomic methylation of cord blood DNA. Environ Health Perspect 117:1466–1471
68. Einstein F, Thompson RF, Bhagat TD, Fazzari MJ, Verma A, Barzilai N et al (2010) Cytosine methylation dysregulation in neonates following intrauterine growth restriction. PLoS One 5:e8887
69. Apostolidou S, Abu-Amero S, O'Donoghue K, Frost J, Olafsdottir O, Chavele KM et al (2007) Elevated placental expression of the imprinted PHLDA2 gene is associated with low birth weight. J Mol Med 85:379–387

70. Diplas AI, Lambertini L, Lee MJ, Sperling R, Lee YL, Wetmur J et al (2009) Differential expression of imprinted genes in normal and IUGR human placentas. Epigenetics 4:235–240
71. Guo L, Choufani S, Ferreira J, Smith A, Chitayat D, Shuman C et al (2008) Altered gene expression and methylation of the human chromosome 11 imprinted region in small for gestational age (SGA) placentae. Dev Biol 320:79–91
72. McMinn J, Wei M, Schupf N, Cusmai J, Johnson EB, Smith AC et al (2006) Unbalanced placental expression of imprinted genes in human intrauterine growth restriction. Placenta 27:540–549
73. Chelbi ST, Doridot L, Mondon F, Dussour C, Rebourcet R, Busato F et al (2011) Combination of promoter hypomethylation and PDX1 overexpression leads to TBX15 decrease in vascular IUGR placentas. Epigenetics 6:247–255
74. Fryer AA, Emes RD, Ismail KM, Haworth KE, Mein C, Carroll WD et al (2011) Quantitative, high-resolution epigenetic profiling of CpG loci identifies associations with cord blood plasma homocysteine and birth weight in humans. Epigenetics 6:86–94
75. Ferreira JC, Choufani S, Grafodatskaya D, Butcher DT, Zhao C, Chitayat D et al (2011) WNT2 promoter methylation in human placenta is associated with low birthweight percentile in the neonate. Epigenetics 6:440–449
76. Filiberto AC, Maccani MA, Koestler DC, Wilhelm-Benartzi C, Avissar-Whiting M, Banister CE et al (2011) Birthweight is associated with DNA promoter methylation of the glucocorticoid receptor in human placenta. Epigenetics 6:566–572
77. Godfrey KM, Sheppard A, Gluckman PD, Lillycrop KA, Burdge GC, McLean C et al (2011) Epigenetic gene promoter methylation at birth is associated with child's later adiposity. Diabetes 60:1528–1534
78. Pembrey ME, Bygren LO, Kaati G, Edvinsson S, Northstone K, Sjostrom M et al (2006) Sex-specific, male-line transgenerational responses in humans. Eur J Hum Genet 14:159–166
79. Carone BR, Fauquier L, Habib N, Shea JM, Hart CE, Li R et al (2010) Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. Cell 143:1084–1096
80. Daxinger L, Whitelaw E (2010) Transgenerational epigenetic inheritance: more questions than answers. Genome Res 20:1623–1628
81. Morgan HD, Sutherland HG, Martin DI, Whitelaw E (1999) Epigenetic inheritance at the agouti locus in the mouse. Nat Genet 23:314–318
82. Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J et al (2003) Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis 35:88–93
83. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A et al (2010) The NIH Roadmap Epigenomics Mapping Consortium. Nat Biotechnol 28:1045–1048

Chapter 8
Epigenetics and Assisted Reproductive Technology*

Nicole Banks and James H. Segars

Contents

8.1 Introduction 118
8.2 Imprinting 119
8.3 Biologic Plausibility 120
  8.3.1 In Vitro Culture and Culture Medium 121
  8.3.2 Hormonal Stimulation 121
  8.3.3 ICSI 122
  8.3.4 Embryo Manipulation and General Effects 123
  8.3.5 Cryopreservation and Vitrification 123
  8.3.6 Histones 124
  8.3.7 Current Understanding and Limitations of Existing Knowledge 124
8.4 Beckwith-Wiedemann Syndrome 126
8.5 Angelman Syndrome 131
8.6 Retinoblastoma 132
8.7 Conclusion 133
References 133

Abstract With increasing use of assisted reproductive technology (ART), concern has arisen over possible risks and long-term health implications of ART. Studies in animals have revealed epigenetic alterations associated with ART, including altered DNA methylation after embryo manipulation, ovarian stimulation, and in vitro embryo culture. Of concern, studies in humans have suggested an increased prevalence of imprinting disorders in children born after ART, particularly Beckwith-Wiedemann syndrome (BWS) and Angelman syndrome (AS). Moreover, children

*Financial support: This research was supported, in part, by the Intramural Research Program in Reproductive and Adult Endocrinology, NICHD, NIH.

N. Banks • J.H. Segars (✉)
Program in Reproductive and Adult Endocrinology, NICHD, NIH, Building 10, CRC, Rm 1E-3140, 10 Center Drive, Bethesda, MD 20892, USA
e-mail: banksnk@mail.nih.gov; segarsj@mail.nih.gov

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_8,
© Springer Science+Business Media B.V. 2012

conceived using ART and affected by BWS or AS are more likely to have methylation abnormalities as the underlying molecular cause of the syndrome compared with affected offspring born after natural conception. Fortunately, both BWS and AS are very uncommon and the absolute risk remains low. Large prospective cohort studies of children conceived using ART are needed to better delineate long-term effects of potential epigenetic alterations possibly associated with ART.

Abbreviations

AS	Angelman Syndrome
ART	Assisted reproductive technology
BWS	Beckwith-Wiedemann syndrome
CpG	Cytosine-phosphate-guanine
DMD	Differentially methylated domain
DMRs	Differentially methylated regions
ICR	Imprinting control region
IVF	In vitro fertilization
ICSI	Intracytoplasmic sperm injection
IUI	Intrauterine insemination
LOS	Large offspring syndrome
PWS	Prader-Willi syndrome
PP	Proximal promoter

8.1 Introduction

Since the birth of the first "test-tube baby" in 1978, the use of assisted reproductive technology (ART) has grown exponentially. From 1990 to 2001, the annual number of ART cycles increased by 246% [1]. In 2007, 138,198 cycles of ART were performed in the United States alone, with the birth of 54,656 infants [2]. With increasing use of ART, concerns over possible risks and long-term health implications have also risen.

Animal models have suggested that embryo culture media can affect non-genomic inheritance, lending biologic plausibility to the concern that ART may increase the risk for imprinting disorders [3–5]. The first study to draw attention to a potential link between ART and imprinting disorders was a report by Cox et al. in 2002 [6] of two children born with Angelman syndrome (AS) after intracytoplasmic sperm injection (ICSI). Since 2002, several studies have suggested a higher prevalence of Beckwith-Wiedemann syndrome (BWS) and AS among children born after ART. In this chapter, we examine the current evidence regarding the possible association of imprinting disorders with ART.

8.2 Imprinting

Genomic imprinting is an epigenetic process which regulates gene transcription and may lead to expression of only one allele of a particular gene in a parental-specific manner. Epigenetic mechanisms are reviewed elsewhere in this book in detail, but two genomic mechanisms relevant to published studies involving ART are DNA methylation and histone modification. Sex-specific methylation patterns in gametes occur at DNA sites called differentially methylated regions (DMRs). Gene imprinting was first studied using mouse nuclear-transplant studies in which mouse embryos created with two female pronuclei or two male pronuclei failed to survive in utero [7–9]. Embryos derived from female pronuclei showed near-normal embryos, with abnormal extraembryonic tissue; conversely, embryos derived from male pronuclei showed poorly developed embryonic tissue, but robust development of extraembryonic tissue [7, 9–11]. These studies suggest that for normal embryogenesis to occur, both a maternal and a paternal genome were required. Currently, more than 80 imprinted genes have been identified in mammals, approximately half of which occur in humans [12]. Most imprinted genes are involved in embryonic and extraembryonic growth and development. Ten syndromes are known to be caused by imprinting defects in humans, including the most common conditions BWS, Prader-Willi syndrome (PWS), and AS.

ART interventions have the potential to affect epigenetic processes at multiple stages of gametogenesis and embryo development (Fig. 8.1). In particular, epigenetic reprogramming of DNA methylation occurs during germ cell and preimplantation development when ART could theoretically disturb the normal conception process [13, 14] (reviewed in [15]). Based on data in animal models (see below) there is concern that ART procedures performed on human embryos may also be vulnerable to manipulations. ART manipulations of conception include ovarian stimulation, in vitro maturation of oocytes, the use of ICSI, the use of immature sperm, in vitro culture of embryos, and cryopreservation of both embryos and gametes. A key factor is the *timing* of the manipulation relative to the erasure and establishment of imprinting marks. For instance, oocytes are vulnerable because re-establishment of methylation marks occurs just before ovulation, thus ovarian stimulation in the course of an ART cycle could conceivably affect oocyte imprinting. Likewise, imprinting marks undergo erasure actively in the male haplogenome and passively in the female haplogenome during early embryogenesis, but there is remethylation of the zygotic genome around the time of implantation. Thus, the timing of embryo manipulations at ART coincides with demethylation and remethylation normally occurring in the embryo [15].

A second equally important consideration is that some cells in the embryo appear to be more susceptible to alteration in imprint marks during in vitro manipulation. For example, the placenta exhibited a greater loss of imprinting than did the embryo in a mouse model of in vitro culture [16, 17]. This observation is noteworthy given the transgenerational effects of altered epigenetic marks in the placenta in animals ([18] and references therein) and their relevance to human diseases [19, 20]. Concern

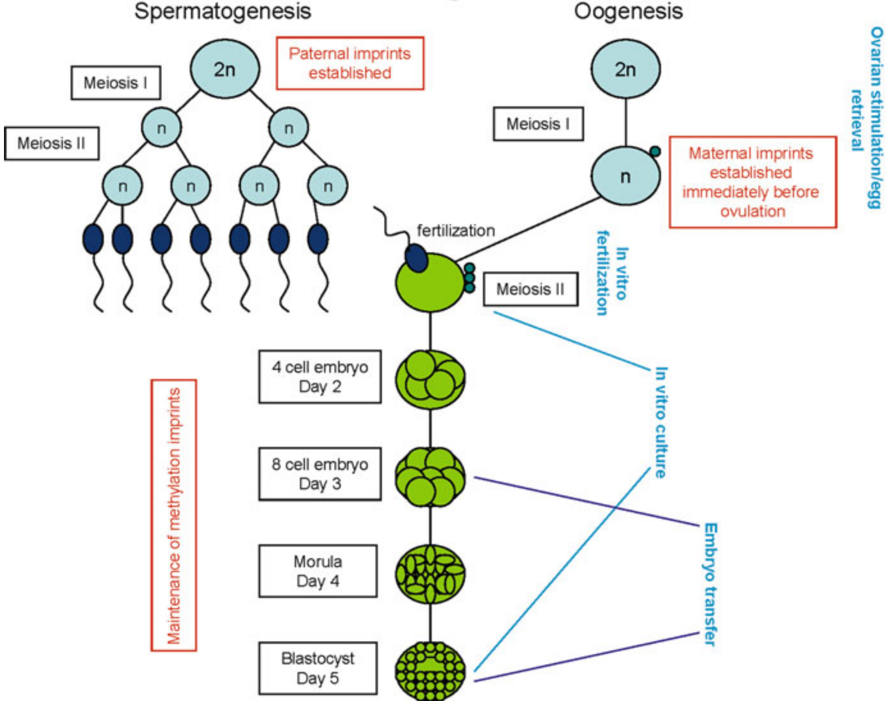


Fig. 8.1 ART influence on gametogenesis and preimplantation development. Steps in meiosis are shown schematically for sperm and oocytes. Maternal imprints finalized before ovulation. Normally, paternal imprints are established very early in spermatogenesis, and are not vulnerable to influences of ART. Following fertilization, imprint marks undergo differential demethylation for each haplogenome, parternal and maternal, and remethylation of the zygote genome occurs around the time of implantation

that ART might alter placental epigenetic marks might seem irrelevant since the placenta is discarded at birth, but even transiently altered placental function theoretically could affect future generations [21]. Nonetheless, the evidence in humans is scant and the concern is inferred, based on results in animal studies. Next, we discuss the biologic plausibility and expand the discussion and examine the evidence for effects of specific ART procedures on imprinting.

8.3 Biologic Plausibility

Many animal and some human studies have been conducted to address how various ART manipulations might affect gene imprinting. Strong evidence from animal studies indicates that in vitro culture can affect epigenetic inheritance. There is

evidence that hormonal stimulation and/or ovarian stimulation may also lead to methylation abnormalities. In addition to DNA methylation and histone modification, other mechanisms of imprinting might be affected by ART manipulations, but evidence for effects of ART on other mechanisms of epigenetic inheritance is sparse, particularly in human offspring. Finally, a few recent studies in humans suggest more widespread altered methylation patterns in children born after ART and support the notion of biologic plausibility (reviewed in Ref. [21]).

8.3.1 In Vitro Culture and Culture Media

The role of in vitro culture in abnormal gene imprinting has been investigated in several animal studies. Cattle born after in vitro culture of embryos are known to develop large offspring syndrome (LOS). Cattle affected by LOS are large at birth, have skeletal anomalies and organomegaly, and suffer from neonatal respiratory distress and increased rates of sudden neonatal death [22]. Sheep cultured in vitro with LOS showed hypomethylation leading to expression of *IGF2R* [23]. Zaitseva et al. [5] performed a study on mouse embryos and found significantly higher methylation levels in two cell embryos cultured in vitro, compared with embryos produced in vivo. The same study also showed a greater degree of demethylation in male pronuclei from in vivo-derived embryos compared to embryos cultured in vitro. This result differed from female pronuclei, which showed no difference in methylation whether developed in vitro or in vivo. The sex-based difference supports the hypothesis that altered patterns of methylation following in vitro culture could arise during culture.

Animal studies have also revealed that different embryo culture media differentially affect imprinting patterns of imprinted genes. One study of mouse embryos detected loss of methylation at *H19* in embryos cultured with Whitten's culture media, but not in embryos developed in KSOM+AA media [3]. Despite hypomethylation of *H19* in Whitten's culture media, no difference in DNA methyltransferase activity or distribution of *DNMT1* proteins was observed. Another study of growth-related imprinted genes in mice compared embryos produced in vivo to embryos cultured in either M16 media or M16 media with fetal calf serum [4]. Significant differences in gene expression were detected between the three groups. Collectively, these studies indicate that embryo culture media has the potential to affect gene imprinting.

8.3.2 Hormonal Stimulation

Ovarian stimulation is a very common practice in infertility treatment and is generally used in most in vitro fertilization (IVF) and ICSI cycles. In a study comparing two cell mouse embryos from superovulated mice and nonsuperovulated mice, a

higher incidence of methylation abnormalities in embryos was detected in mice that received ovarian stimulation [24]. This observation suggests that hormonal stimulation of the ovaries may affect gene imprinting. Market-Velker et al. [25] studied superovulation in a mouse model using embryos fertilized in vivo to examine effects of in vitro manipulations. In comparing superovulated mice versus control mice, loss of methylation was observed on the maternal allele at *SNRPN*, *PEG3*, and *KCNQ1OT1*. This effect was dose dependent and statistically significant at the higher dosage of hormone stimulation. Gain of methylation was present in the *H19* gene, but again in a dose-dependent manner that was significant at the "high dose." Interestingly, both the maternal and paternal *H19* alleles were perturbed by superovulation suggesting that maintenance of imprinting after fertilization may be affected. The authors [25] suggest that superovulation affects oogenesis in two ways, first by disrupting the acquisition of imprints in growing oocytes, and second by altering maternal gene products required for imprint maintenance during preimplantation development. Consistent with this observation, a human study suggested that hormonal stimulation of the ovary affected gene imprinting leading to a loss of methylation at *PEG1* and a gain of methylation at *H19* in superovulated immature human oocytes [26].

Fauque et al. [27] studied the effects of superovulation, IVF, and embryo culture media (M16 versus G1.2/G2.2) on the methylation status of the *H19* imprinting control region (ICR) and *H19* proximal promoter (PP). The study also tested gene expression level of *H19* in the blastocysts for the various conditions. The authors showed that methylation defects occurred only in the IVF group and the defects were more common when the embryo was cultured in M16 medium. Moreover, superovulation significantly disrupted gene expression of *H19* in blastocysts. The culture media again affected the degree of disruption, with *H19* gene expression higher in blastocysts cultured in G1.2/G2.2 media [27]. Taken together, these studies suggest that hormonal stimulation of ovaries may affect oocyte imprinting.

8.3.3 ICSI

Use of ICSI has increased dramatically in recent years and the proposed indications for ICSI are rapidly expanding. The ICSI procedure per se has not been definitively shown to play a direct role in imprinting disorders. Use of immature sperm is unlikely to affect paternal allele imprints because the imprints are normally established early in spermatogenesis [28, 29]. However, it is possible that abnormal spermatogenesis per se is associated with epigenetic anomalies. A study analyzing methylation patterns of *H19* in sperm found an association between low sperm counts and methylation abnormalities [30]. Specifically, 0%, 17%, and 30% of normozoospermia, moderate oligozoospermia, and severe oligozoospermia respectively showed altered *H19* methylation patterns. Another study found that 14% of infertile men had abnormal paternal imprints in their sperm and 20% had abnormal maternal imprints [31]. Kobayashi et al. [32] collected 78 aborted samples from IVF/ICSI procedures and examined seven autosomal imprinted genes and the X-linked *XIST* gene and matched parental sperm.

The study found 17 cases of abnormal DNA methylation [32]. Seven of the 17 cases (41%) had identical alterations in the parental sperm, suggesting that some imprinting errors associated with ART were inherited from the father. Thus, it does appear that abnormal spermatogenesis is associated with epigenetic changes and these changes may be heritable, but no direct effects of ICSI have been detected.

8.3.4 Embryo Manipulation and General Effects

Physical embryo manipulation has been shown to affect DNA methylation status in the mouse model. Embryo manipulation limited to blastocyst collection and culture for 1.5 h or less was correlated with a loss of methylation at the maternal *KvDMR1* locus [17]. The same study found that in vitro culture of embryos from the two cell stage to blastocysts resulted in further methylation abnormalities.

A recent human study analyzed DNA methylation patterns of patients delivered after IVF versus natural conception. Katari et al. [33] studied the methylation patterns of 10 children born after IVF and 13 children conceived in vivo. The investigators used a custom-designed methylation bead-array platform with probes for 1536 cytosine-phosphate-guanine (CpG) sites located in the promoters of more than 700 genes to compare DNA methylation of placenta and cord blood of the subjects and controls. The study found that in vitro conception was associated with lower mean methylation at CpG sites in placenta and higher mean methylation levels in cord blood. Of the CpG sites where mean methylation levels differed in placenta, 63% showed reduced methylation in children born after IVF. For cord blood, 77% of discordant CpG sites showed increased methylation in post-IVF children. The study also attempted to examine gene expression by selecting a sampling of genes for analysis of transcription activity and found that methylation differences in post-IVF children were associated with differences in gene expression at imprinted and non-imprinted genes. The authors suggest [33] that their findings could be explained by pre-existing differences in the gametes of couples undergoing IVF treatments, or alternatively, the differences could reflect changes caused by global patterns of DNA methylation and gene expression due to ART. The study does not address what ART procedures may be the cause of global DNA methylation changes.

8.3.5 Cryopreservation and Vitrification

Currently there are very limited data regarding the possible effects of cryopreservation of embryos or gametes on epigenetic imprints during ART. A recent study in a murine model focused on H19/Igf2 differentially methylated domain (DMD) reported that in vitro culture resulted in epigenetic changes in embryos, consistent with findings of previous investigators, but the process of embryo vitrification significantly augmented the loss of methylation of the H19 DMD [34]. Again, the

placenta appeared to be more affected than the embryo. Despite these data to suggest the process of cryopreservation alters imprint marks in animals, human offspring resulting from cryopreserved embryos have a better neonatal outcome than human offspring after a "fresh" ART cycle ([35], and references therein). The methods of embryo cryopreservation at ART have undergone considerable refinement in recent years and it is worth noting that procedural methodology involved in cryopreservation could greatly influence findings, therefore more studies are needed.

8.3.6 Histones

Very little evidence is available regarding histone modification and ART. One study examined histone modification in murine embryos, but found no difference in histone modification patterns between in vitro fertilized and in vivo fertilized embryos [36].

8.3.7 Current Understanding and Limitations of Existing Knowledge

There is overwhelming evidence from well-designed, rigorously controlled studies of ART procedures in "normal" animals that the processes associated with ART can influence gene imprinting, leading to specific syndromes and phenotypes. Furthermore, evidence in animals clearly demonstrates that different cell types have a differential susceptibility to alteration of imprinted genes. The stress of ART (Fig. 8.2) can be considered similar to environmental stresses that are well-described, but in general, understanding of the mechanisms mediating altered expression on imprinted genes as a direct result of ART procedures are not as advanced as understanding of the mechanisms of chemical or environmental toxins. Understanding of the molecular and cellular processes involved in re-programming, and more specifically the fidelity and preservation of the imprinted marks during their erasure and re-establishment, is incomplete.

While there is no logical reason to assume the human embryos would be immune to the effects of in vitro culture, there are clear species-dependent differences in vulnerability to the alteration of imprint marks. In general, the evidence in humans is not as compelling as in other species, and several additional conditions may confound the relationship. For instance, unlike experiments in animals, ART is not generally conducted on normal controls (other than oocyte donors), and some data suggest the diagnosis of infertility might be associated with altered imprinting. Analysis of imprinted genes in groups such as oocyte donors or "natural cycle" IVF would be of interest, since such studies might control for factors such as infertility and ovarian stimulation, respectively. Based on current evidence, there *does* appear to be an effect of in vitro culture or ART procedures upon imprinted genes. More specifically, a slight effect on a number of imprinted genes has been detected in some ART offspring [33], although the degree of the risk and consequences of this

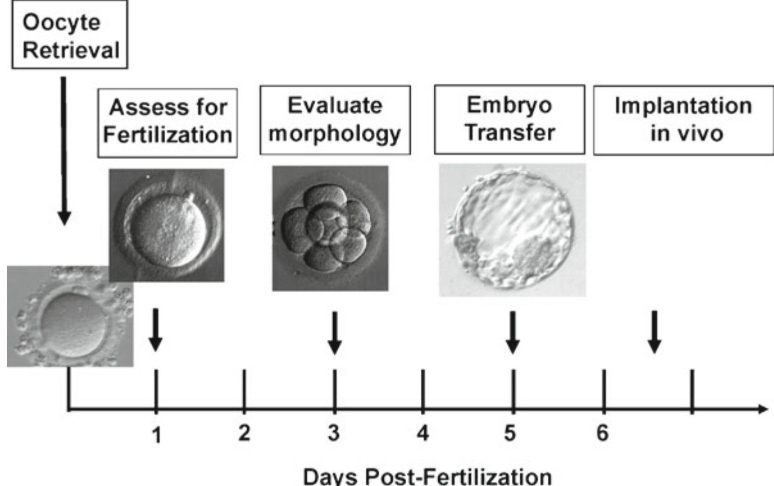


Fig. 8.2 Stages of early embryo development during assisted reproduction. Oocytes are retrieved following stimulation and are evaluated in the laboratory. Following fertilization by either addition of sperm (in vitro fertilization) or ICSI, fertilization is assessed on day 1 following oocyte retrieval (*first down arrow* from the *left*). On day 2 or 3 of culture embryos are re-assessed for development (*second down arrow*), and either transferred to the endometrial cavity, or maintained in culture until day 5, when blastocyst development has occurred (*third down arrow*). Embryos are transferred to the endometrial cavity and implantation occurs in vivo (*fourth down arrow*). Epigenetic imprints are normally completed in oocytes just prior to retrieval (day 0). Epigenetic imprints in male gametes normally occurs in the gonadal ridge and have been completed well before use of sperm for fertilization (evening of day 0). Imprinting marks undergo erasure actively in the male haplogenome and passively in the female haplogenome during early embryogenesis (day 1–5). The timing of embryo manipulations at ART coincides with re-programming of epigenetic marks that occur during early embryo development

level of alteration in imprinting marks is not clear [21]. Two questions before the field are: how often are epigenetic marks altered due to ART, and what, if any, are the long-term consequence of any alterations that may occur?

Based on the analysis of specific diseases, such as Beckwith-Wiedemann Syndrome (BWS) as discussed in the next section, the *absolute level of risk* for profoundly affected offspring is very low. Comparison of ART offspring with BWS, to spontaneously occurring BWS cases where ART was not performed, reveals there is often similarity in the mechanism involved, but notably, loss of methylation of the maternal allele is more frequent in the ART-conceived offspring [15]. A significant current limitation in the level of evidence is that most studies in humans have reported differences in gene imprinting, but have not proven the case by showing altered levels of mRNA transcripts and/or protein expression. Another limitation is that studies have focused primarily on DNA methylation, and other possible epigenetic mechanisms, such as miRNAs, have not been assessed.

In the following sections we examine the association of ART with specific syndromes. Because of the rarity of imprinting disorders that arise spontaneously, many reports are either retrospective observational or case-control design. Accordingly, the

strength of association between ART and the specific condition in question often hinges upon appropriate control groups, absence of recall and ascertainment bias, appropriate control for confounding factors, and adequate statistical power. Unfortunately, many studies have weaknesses that bring into question many of these important considerations, and thus weaken the level of certainty between the specific condition and ART.

8.4 Beckwith-Wiedemann Syndrome

Beckwith-Wiedemann syndrome is a congenital overgrowth syndrome that occurs in approximately one in 13,700 births [37]. Children with BWS exhibit macroglossia, macrosomia, midline abdominal wall defects, neonatal hypoglycemia, as well as other developmental anomalies [38]. BWS is also notable for an increased risk of developing embryonal tumors in childhood, most commonly Wilm's tumor, rhabdomyosarcoma, and hepatoblastoma [39].

BWS is a multi-genic disorder with both known genetic and epigenetic causes. Paternally expressed genes *IGF2* and *KCNQ1OT1* and maternally expressed genes *H19, CDKN1C*, and *KCNQ1* at the 11p15 imprinted region are implicated in BWS [40]. *H19* and *IGF2* are usually jointly regulated by *DMR1* to assure that *H19* is only expressed from the maternal allele and *IGF2* is only expressed from the paternal allele. One cause of BWS involves hypermethylation at the maternal *DMR1*, leading to silencing of the maternal *H19* and bi-allelic expression of *IGF2* [41]. Another cause implicates the imprinting center *KvDMR1,* which controls *KCNQ1OT1* (also known as *LIT1*), *KCNQ1*, and *CDKN1C* expression. Loss of maternal methylation at *KvDMR1* leads to biallelic expression of *KCNQ1OT1* and decreased expression of *CDKN1C*. Approximately 50–60% of spontaneous BWS cases are due to loss of maternal methylation at *KvDMR1* [41]. There is an increased incidence of BWS in monozygotic twins, with the majority of twins being discordant for BWS [42, 43]. The most common defect among these discordant twins was an imprinting defect at *KCNQ1OT1* [42, 43]. Because of this, it has been hypothesized that *KCNQ1OT1* is especially vulnerable to imprinting errors during the preimplantation developmental period, a stage frequently involved in manipulation during ART [43].

Concerns for an association between BWS and ART first arose with the publication of three case series from BWS registries in the United States and Europe in 2003 [44–46]. DeBaun et al. [44] reported a case series of seven offspring with BWS after ART (two IVF and five ICSI). The authors calculated the prevalence of ART in the BWS cohort as 4.6%. This included data after 2001 only, because there was no formalized data collection within the cohort regarding ART prior to that year. Compared to the 0.76% total live births after ART in the general population in 1999, this suggested a six-fold increased risk of BWS with ART. Five of six patients tested for the molecular basis of their disease were found to have an imprinting defect (Table 8.1), specifically hypomethylation of *LIT1* (*KCNQ1OT1/KvDMR1*).

Similarly, Maher et al. [46] reviewed 149 sporadic BWS cases in Birmingham, United Kingdom, and found six BWS children were born after ART (4%),

Table 8.1 Beckwith-Wiedemann syndrome in ART: prevalence and molecular defect

Author	Study type	Number of BWS cases	Prevalence of ART in cases (%)	Prevalence of ART in general population (%)	Percentage of BWS patients with *KvDMR1* hypomethylation[a]
Debaun et al. [44]	Case series	65	4.6	0.76	83.3
Maher et al. [46]	Case series	149	4	1.2	100
Gicquel et al. [45]	Case series	149	4	1.3	100
Halliday et al. [47]	Case-control	37	10.81	0.67	100
Lidegaard et al. [55]	Cohort	0	0	1.3%	0
Chang et al. [51]	Case series[b]	341	5.6	NR	NR
Sutcliffe et al. [52]	Survey	209	2.9	0.8	100
Rossignol et al. [49]	Cohort[c]	11	27.5[c]	NA	100
Bowdin et al. [53]	ART cohort	4	NA	NA	25
Doornbos et al. [54]	Survey	71	5.6	0.92	100
Lim et al. [48]	ART case-control	24	NA	NA	96

NA not applicable based on study design, *NR* not reported

[a]Number of patients with *KvDMR1* hypomethylation/number of tested patients

[b]BWS registry

[c]Cohort with known imprinting defect at *KvDMR1*

significantly higher than the estimated 1.2% ART use in the general population. Two of the patients were tested for *KvDMR1* methylation and both showed a loss of methylation on the maternal allele. The third case series by Gicquel et al. [45] reported a series of 149 BWS cases in France. Six children from that cohort were born after ART (4%), as compared to 1.3% of the general population. Based on these data, an odds ratio (OR) of 3.2 (95% CI, 1.4–7.3) was calculated for the risk of BWS after ART. All six patients were tested for the molecular defect causing BWS and all demonstrated hypomethylation of *KvDMR1*.

One case-control study from Victoria, Australia, studied 37 cases of BWS and 148 non-BWS matched controls [47]. Using data from 1983 to 2003, all cases of BWS in Victoria were identified and for each BWS case, four live-born controls were randomly selected from babies born within 1 month of the index case with maternal parity of one and maternal age within 1 year of the BWS mother. In that study, four BWS patients were conceived via ART (10.81%) and only one infant from the control group was conceived via ART (0.67%); OR = 17.8 (95% CI, 1.8–432.9) [47]. Three of the patients born with BWS after ART were tested for and found to have *KvDMR1* hypomethylation, compared with only 46% of children with BWS conceived without ART.

Another case-control study analyzed the molecular genetic and clinical features of 25 cases of BWS born after ART compared with 87 BWS children conceived spontaneously [48]. The control group was composed of patients with *KvDMR1* loss of methylation (IC2 defect). Of the 25 post-ART BWS cases, 24 had loss of methylation at *KvDMR1* (no molecular cause was identified in one post-ART BWS case). A molecular analysis of DMRs (*ZAC, PEG1, SNRPN*, and *DLK1*) showed loss of maternal allele methylation in 37.5% of ART versus 6.4% of non-ART BWS patients with IC2 defects. The authors concluded that wide-spread hypomethylation of DMRs was more common in post-ART BWS cases than in non-ART BWS patients with IC2 defects. Additionally, this study examined clinical differences between post-ART and non-ART BWS patients and found a significant difference in incidence of facial nevus flammeus, exomphalos and neoplastic lesions. Post-ART BWS patients were more likely to have facial nevus flammeus (90 vs. 46%, $p = 0.0004$) and less likely to have exomphalos (43 vs. 69%, $p = 0.029$). Of the controls, no children developed a neoplasm, but two of the BWS post-ART cases developed an embryonal tumor, one developed a hepatoblastoma and the other a rhabdomyosarcoma, a significant difference ($p = 0.0014$) [48].

Rossignol et al. [49] analyzed the methylation status of 40 patients with BWS and known hypomethylation of *KvDMR1* in three additional DMRs but the findings differed from those of Lim et al. [48]. Eleven patients were born after ART and 29 were conceived naturally [49]. The authors found that 27% of ART-conceived patients and 24% of naturally conceived patients had hypomethylation of at least one additional locus. The authors concluded that the involvement of other loci in hypomethylation in BWS patients was not restricted to those conceived with ART. Additionally, a cohort study from Italy and the Netherlands of 149 BWS patients identified 17 with hypomethylation of multiple imprinted loci [50].

Only one of the 17 patients was born after ART, thus the authors concluded that the use of ART was not the cause of multi-local hypomethylation that has been identified in some BWS patients.

Chang et al. [51] reviewed the same BWS registry used by DeBaun et al. of 341 patients and identified 19 children with BWS born after ART (including ovulation induction and artificial insemination). Twelve of the patient's records were obtained and showed five IVF pregnancies, five ICSI pregnancies, and two with ovulation induction (one gonadotrophin with intrauterine insemination [IUI] and one clomiphene citrate with artificial insemination). The authors reviewed the treatment data for each of the 12 patients to determine what ART treatments might be associated with BWS. There were no observed associations between cause of infertility, type of culture media, type of ART procedure, or day of embryo transfer and BWS. The only commonality in treatment among the 12 patients was the use of ovarian stimulation medications [51].

A study from the United Kingdom was published in 2006 based on a survey of infertility treatments in 213 families with children affected by BWS [52]. Eighty-three responses were received and four familial cases were excluded. The survey found that six BWS children were conceived with IVF or ICSI, for a calculated prevalence of 2.9% of ART use in children with BWS, using the 209 eligible surveyed families as the total number of BWS cases. This was significantly higher than the 0.8% prevalence of children in the general population born after IVF/ICSI during the same time frame. Of the 11 children born after any form of ART (one IVF, five ICSI, five ovarian stimulation medications), molecular analysis was performed on eight. The authors found that all eight children tested had loss of maternal methylation at *KvDMR1*. The study [52] concluded that there was an association between BWS and ART, but also observed that BWS occurred in less than 1% of patients conceived by ART.

Another British questionnaire-based cohort study examined all children conceived by ART from two United Kingdom ART treatment centers [53]. The questionnaire was designed to identify children with a possible diagnosis of BWS or AS. The study had a 61% response rate (1,524 responses out of 2,492). From the questionnaire, 70 children warranted further evaluation. Forty-seven agreed to an evaluation and one case of BWS was identified. This patient was tested for and had *KvDMR1* hypomethylation. The authors concluded [53] that there was a higher prevalence of BWS in children born after ART, but the absolute risk was less than 0.1%.

A questionnaire-based study in the Netherlands assessed families with children affected by imprinting disorders born between January 1983 and December 2003 [54]. The questionnaire addressed infertility and ART (including ovulation induction and artificial insemination). Of 71 eligible cases of BWS identified, six patients with BWS were born after ART (four IVF/ICSI, one IUI with donor sperm, one ovarian stimulation). All six cases were tested for and found to have hypomethylation of *KvDMR1*. The prevalence of BWS after IVF/ICSI was 5.6%, compared to 0.92% in the general Dutch population over the same time period for a relative risk of 6.1 for BWS and ART. However, the authors concluded [54] that after correcting

for infertility problems in parents there was no increased incidence of imprinting disorders (BWS, AS, or PWS) in children born after ART.

Two population based studies arrived at similar conclusions. A cohort study in Denmark from January 1995 to December 2001 used the National IVF registry (6,052 singletons) and the National Birth Registry (442,349 singletons) and found no cases of BWS [55]. The authors therefore did not identify an increased risk of BWS with ART. A Swedish population-based study from 1982 to 2001 tracked the presence of congenital malformations based on data from the Swedish Medical Birth Register, Swedish Register of Congenital Malformations, and the Swedish Hospital Discharge Register [56]. Of 16,280 children born after ART, no cases of BWS were identified [56]. Of note, cryopreservation of embryos in Scandinavia has recently increased, and it is helpful to bear in mind that embryo handling procedures and ART practices in general differ greatly between countries.

The consistent finding of hypomethylation at *KvDMR1* in most BWS patients born after ART raised the question whether clinically normal children conceived by ART might have a higher rate of subclinical methylation abnormalities. To address this question, Gomes et al. [57] tested 18 healthy children born after IVF/ICSI for methylation abnormalities in *KvDMR1*. Three of 18 children were found to have hypomethylation at *KvDMR1* (two IVF, one ICSI), compared to zero of 30 natural conception children used as controls. In contrast, Teirling et al. [58] studied 77 ICSI, 35 IVF, and 73 spontaneously conceived clinically normal children for methylation abnormalities at ten separate DMRs. Nine of the DMRs, including *KvDMR1*, showed no difference in methylation indices between ART and natural conception children. The *MEST* DMR showed a higher methylation index in IVF compared to ICSI and natural conception children. The authors concluded that children conceived by ART do not have a higher degree of imprint variability, but did note that the sample size was too small to exclude rare imprinting disorders [58].

In summary, several studies have examined the association of BWS and ART with conflicting findings. Seven studies among children with BWS showed a higher prevalence of ART use than in the general population (Table 8.1). Findings from large cohort studies were mixed. Two large national studies did not identify any cases of BWS. The third study reported a higher prevalence of BWS with the use of ART. Because BWS is a rare disorder, a large cooperative international study would be required to further study the prevalence of BWS in naturally conceived children and children born after ART. In particular, some findings may be subject to recall or ascertainment bias, because BWS may be underestimated in the general population but more frequently diagnosed in ART-conceived children. Virtually all (90–100%) of BWS cases after ART have imprinting aberration, specifically hypomethylation of the maternal allele at *KvDMR1* (Table 8.1). This is significantly greater than the 50–60% of natural conception BWS patients with the same defect. At this time, there is conflicting evidence whether widespread (generalized) hypomethylation occurs in children born after ART. Overall, the data suggest an association between ART and BWS but the prevalence in ART children is low at approximately one in 4,000.

8.5 Angelman Syndrome

Angelman syndrome is a neurogenetic disorder with a prevalence of approximately 1:12,000 in the general population [59]. It is characterized by severe mental retardation, frequent laughter, seizure disorder, and gait disturbances [60]. The genetic defect in AS is failure of maternal gene expression in the brain of the *UBE3A* gene, located at 15q11-13 [60, 61]. This deficiency can be caused by deletion of maternal *UBE3A* (68%), mutation (13%), uniparental disomy (3%), or imprinting defect causing under-expression of the maternal allele (6%) (Table 8.2) [15, 62]. Imprinting defects as a cause of AS are estimated to occur in 1:300,000 newborns [6].

Cox et al. [6] first raised concerns about AS and ART in 2002 by their report of two cases of AS after ICSI (for male factor infertility). Both affected children were tested and found to have hypomethylation of chromosome 15q11-13. This finding was alarming because of the rarity of AS due to an imprinting defect (1:300,000). A third case of AS after ICSI was reported in 2003. This patient was found to have a complete absence of methylated maternal pattern for the *SNRPN* locus [63]. The case was unique as the father had a normal semen analysis, but the mother and grandmother of the patient had significant reproductive difficulties.

Ludwig et al. [64] reported in 2005 that 20% of AS patients were born to subfertile couples. Using a questionnaire sent to 270 members of the German Angelman Syndrome Support group addressing infertility issues and treatments, the study identified 16 of 79 respondents as having subfertility. Of these 16 cases, four children were found to have imprinting defects. Three of the 16 cases were after ICSI, and only one post-ICSI AS child had an imprinting defect as the cause. The relative risk of an imprinting defect in a child with AS born to a subfertile couple was 6.25. Couples with subfertility who received infertility treatment to conceive had a relative risk of 12.5 of having an imprinting defect as the cause of their child's AS. The authors concluded [64] that imprinting disorders and subfertility may have a common cause and that superovulation, not ICSI, may further increase the risk of imprinting defects in offspring.

The previously mentioned British survey by Sutcliffe et al. [52] queried 384 families with children affected by AS. Eighty-one surveys were returned (21% response rate) and 75 sporadic cases of AS were identified. Three of these cases were born after ART (IUI, donor sperm artificial insemination). The patient born after IUI

Table 8.2 Molecular defect of *UBE3A* in Angelman syndrome

Molecular defect	Approximate percentage
Deletion of maternal *UBE3A*	68
Intragenic mutation on maternal allele	13
Imprinting defect	6
Uniparental disomy	3
Unknown/normal genetic testing	10

to a couple who had previously used IVF was found to have hypomethylation of the maternal *SNRPN* locus. A nationwide survey in the Netherlands [54] found 12 of 63 children with AS (19%) were born to parents with fertility problems, significantly more than the 5.6% of children in the general population. Three of the children were conceived with the use of ovulation induction medications (4.8%), compared to 0.39% of the general population. No cases of AS children born after IVF/ICSI were identified. The authors concluded [54] that there was not an increased risk of imprinting disorders with the use of ART after controlling for subfertility. A survey study of 2,492 children born after ART in the United Kingdom also failed to identify any cases of AS from the 1,524 responses [53]. The authors estimated the incidence in that population of AS from an imprinting defect to be 1:750,000 and determined that the study was not adequately powered to find a 50-fold increased risk.

A small prospective study of 92 healthy children born after ICSI analyzed methylation patterns at chromosome 15q11-13 [65]. A normal methylation pattern was found in all of the children studied. Another study published in 2008 assessed methylation patterns at chromosome 15q11-13 in 53 children born after ICSI, and also found no methylation abnormalities [66]. Both studies pointed out that due to the rarity of imprinting defects and AS, they lacked sufficient power to detect an increased prevalence of such defects.

In summary, seven reported cases of AS after ART were identified, and of these, five suggested an association of ART with altered imprinting. The rarity of an imprinting defect cause in AS, coupled with the high proportion of post-ART AS children with such an imprinting defect, suggests an association between AS and ART. Further study is needed to better examine the potential link between AS and ART.

8.6 Retinoblastoma

Retinoblastoma is a rare childhood malignant tumor of the retina. It occurs in approximately 1:17,000 children [67]. Retinoblastoma is most often caused by a mutation of the tumor suppressor gene *RB1* combined with a loss or deletion of the second allele [68]. Approximately 40% of cases are familial. Hypermethylation of the RB gene has been shown to play a role in retinoblastoma [69, 70]. A case series from the Netherlands reported five children with retinoblastoma conceived with ART between November 2000 and February 2002 [71]. The authors calculated an increased risk of 4.9–7.2. As methylation status was not studied, the authors could not comment on whether imprinting defects were involved. A national cohort study in Denmark of all singleton births from 1995 to 2001 found no cases of retinoblastoma among 6,052 singleton children born after IVF, followed for an average of 4.1 years [55]. The finding of five cases of retinoblastoma post-IVF is concerning due to the rarity of the disease; however, no other studies to date have shown any association between retinoblastoma and ART.

8.7 Conclusion

ART has been associated with an increased prevalence of the imprinting disorders BWS and AS in published reports. Epigenetic alterations due to ART are biologically plausible based on animal studies that have shown DNA methylation defects after embryo manipulation, ovarian stimulation, and in vitro embryo culture. Furthermore, a higher proportion of children born post-ART with BWS or AS have imprinting defects as the cause of the condition, compared to affected BWS or AS children not conceived using ART. Imprinting disorders are exceedingly rare, complicating study of the subject. The existing literature is limited by under-powered studies, and reliance on retrospective observational analyses and survey based studies, both vulnerable to bias. Despite these limitations, multiple case series and case-control studies have detected an association between BWS and ART, especially hypomethylation of *KvDMR1*. While not as extensively studied, five cases in the literature of AS caused by sporadic imprinting disorder raises concern due to the rarity of AS. Some recent reports suggest a generalized alteration in DNA methylation associated with ART, which, given the rarity of both BWS and AS, could obfuscate the findings. More studies are needed to explore whether infertility or any facet of ART treatment may increase the likelihood of diseases associated with imprinting.

References

1. Henne MB, Bundorf MK (2008) Insurance mandates and trends in infertility treatments. Fertil Steril 89:66–73
2. Centers for Disease Control and Prevention, A. S. R. M., Society for Assisted Reproductive Technology (2007) Assisted Reproductive Technology Success Rates: National Summary and Fertility Clinic Reports. Atlanta, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention 2009
3. Doherty AS, Mann MR, Tremblay KD, Bartolomei MS, Schultz RM (2000) Differential effects of culture on imprinted H19 expression in the preimplantation mouse embryo. Biol Reprod 62:1526–1535
4. Khosla S, Dean W, Brown D, Reik W, Feil R (2001) Culture of preimplantation mouse embryos affects fetal development and the expression of imprinted genes. Biol Reprod 64:918–926
5. Zaitseva I, Zaitsev S, Alenina N, Bader M, Krivokharchenko A (2007) Dynamics of DNA-demethylation in early mouse and rat embryos developed in vivo and in vitro. Mol Reprod Dev 74:1255–1261
6. Cox GF, Bürger J, Lip V, Mau UA, Sperling K, Wu B-L, Horsthemke B (2002) Intracytoplasmic sperm injection may increase the risk of imprinting defects. Am J Hum Genet 71:162–164
7. Barton SC, Adams CA, Norris ML, Surani MA (1985) Development of gynogenetic and parthenogenetic inner cell mass and trophectoderm tissues in reconstituted blastocysts in the mouse. J Embryol Exp Morpho 90:267–285
8. McGrath J, Solter D (1984) Completion of mouse embryogenesis requires both the maternal and paternal genomes. Cell 37:179–183
9. Surani MA, Barton SC, Norris ML (1984) Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature 308:548–550
10. Surani MA, Barton SC (1983) Development of gynogenetic eggs in the mouse: implications for parthenogenetic embryos. Science 222:1034–1036

11. Barton SC, Surani MA, Norris ML (1984) Role of paternal and maternal genomes in mouse development. Nature 311:374–376
12. Morison IM, Ramsay JP, Spencer HG (2005) A census of mammalian imprinting. Trends Genet 21:457–465
13. Reik W, Dean W, Walter J (2001) Epigenetic reprogramming in mammalian development. Science 293:1089–1093
14. Wilkins-Haug L (2008) Assisted reproductive technology, congenital malformations, and epigenetic disease. Clin Obstet Gynecol 51:96–105
15. Owen C, Segars J (2009) Imprinting disorders and assisted reproductive technology. Semin Reprod Med 27:417–428
16. Mann M, Lee S, Doherty A, Verona R, Nolen L, Schultz R et al (2004) Selective loss of imprinting in the placenta following preimplantation development in culture. Development 131:3727–3735
17. Rivera RM, Stein P, Weaver J, Mager J, Schultz R, Bartolomei M (2008) Manipulations of mouse embryos prior to implantation result in aberrant expression of imprinted genes on day 9.5 of development. Hum Mol Genet 17:1–14
18. Bocock PN, Aagaard-Tillery K (2009) Animal models of epigenetic inheritance. Semin Reprod Med 27:369–379
19. Barker D (2007) The origins of the developmental origins theory. J Intern Med 261:412–417
20. Wadhwa PD, Buss C, Entringer S, Swanson JM (2009) Developmental origins of health and disease: a brief history of the approach and current focus on epigenetic mechanisms. Semin Reprod Med 27:358–368
21. Batcheller A, Maguire M, DeCherney AH, Segars JH (2011) Are there subtle, genome-wide epigenetic alterations in normal offspring conceived from Assisted Reproductive Technologies? Submitted. Fertility and Sterility, Article in Press. Oct 26. [Epub ahead of print]
22. Young LE, Sinclair KD, Wilmut I (1998) Large offspring syndrome in cattle and sheep. Rev Reprod 3:155–163
23. Young LE, Fernandes K, McVoy TG, Butterwith SC, Guiterrez CG, Carolan C, Broadbent PJ, Robinson JJ, Wilmut I, Sinclair KD (2001) Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture. Nat Genet 27:153–154
24. Shi W, Haaf T (2002) Aberrant methylation patterns at the two-cell stage as an indicator of early developmental failure. Mol Reprod Dev 63:329–334
25. Market-Velker BA, Zhang L, Magri LS, Bonivissuto AC, Mann MR (2010) Dual effects of superovulation: loss of maternal and paternal imprinted methylation in a dose-dependent manner. Hum Mol Genet 19:36–51
26. Sato A, Otsu E, Negishi H, Utsunomiya T, Arima T (2007) Aberrant DNA methylation of imprinted loci in superovulated oocytes. Hum Reprod 22:26–35
27. Fauque P, Jouannet P, Lesaffre C, Ripoche MA, Dandolo L, Vaiman D, Jammes H (2007) Assisted Reproductive Technology affects developmental kinetics, H19 Imprinting Control Region methylation and H19 gene expression in individual mouse embryos. BMC Dev Biol 7:116
28. Hartmann S, Bergmann M, Bohle RM, Weidner W, Steger K (2006) Genetic imprinting during impaired spermatogenesis. Mol Hum Reprod 12:407–411
29. Manipalviratn S, DeCherney A, Segars J (2009) Imprinting disorders and assisted reproductive technology. Fertil Steril 91:305–315
30. Marques CJ, Carvalho F, Sousa M, Barros A (2004) Genomic imprinting in disruptive spermatogenesis. Lancet 363:1700–1702
31. Kobayashi H, Sato A, Otsu E, Hiura H, Tomatsu C, Utsunomiya T, Sasaki H, Yaegashi N, Arima T (2007) Aberrant DNA methylation of imprinted loci in sperm from oligospermic patients. Hum Mol Genet 16:2542–2551
32. Kobayashi H, Hiura H, John R, Sato A, Otsu E, Kobayashi N, Suzuki R, Hayashi C, Utsunomiya T, Yaegashi N, Arima T (2009) DNA methylation errors at imprinted loci after assisted conception originate in the parental sperm. Eur J Hum Genet 17:1582–1591
33. Katari S, Turan N, Bibikova M, Erinie O, Chalian R, Foster M, Gaughan JP, Coutifaris C, Sapienza C (2009) DNA methylation and gene expression differences in children conceived in vitro or in vivo. Hum Mol Genet 18:3769–3778

34. Wang Z, Wang Z, Xu L, He F (2010) Embryo vitrification affects the methylation of the H19/IGF2 differentially methylated domain and the expression of H19 and Igf2. Fertil Steril 93(8):2729–2733
35. Pinborg A, Loft A, Henningsen A-A A, Rasmussen S, Andersen AN (2010) nfant outcome of 957 singletons born after frozen embryo replacement: the Danish National cohort study 1995–2006. Fertil Steril 94:1320–1327
36. Huang JC, Lei ZL, Shi LH, Miao YL, Yang JW, Ouyang YC, Sun QY, Chen DY (2007) Comparison of histone modifications in in vivo and in vitro fertilization mouse embryos. Biochem Biophys Res Commun 354:77–83
37. Thorburn MJ, Wright ES, Miller CG, Smith-Read EH (1970) Exomphalos-macroglossia-gigantism syndrome in Jamaican infants. Am J Dis Child 119:316–321
38. Elliott M, Bayly R, Cole T, Temple IK, Maher ER (1994) Clinical features and natural history of Beckwith-Wiedemann syndrome: presentation of 74 new cases. Clin Genet 46:168–174
39. Junien C (1992) Beckwith-Wiedemann syndrome, tumourigenesis and imprinting. Curr Opin Genet Dev 2:431–438
40. Weksberg R, Smith AC, Squire J, Sadowski P (2003) Beckwith-Wiedemann syndrome demonstrates a role for epigenetic control of normal development. Hum Mol Genet 12(Spec No 1):R61–8
41. Weksberg R, Shuman C, Smith AC (2005) Beckwith-Wiedemann syndrome. Am J Med Genet C Semin Med Genet 137:12–23
42. Clayton-Smith J, Read AP, Donnai D (1992) Monozygotic twinning and Wiedemann-Beckwith syndrome. Am J Med Genet 42:633–637
43. Weksberg R, Shuman C, Caluseriu O, Smith AC, Fei YL, Nishikawa J, Stockley TL, Best L, Chitayat D, Olney A, Ives E, Schneider A, Bestor TH, Li M, Sadowski P, Squire J (2002) Discordant KCNQ1OT1 imprinting in sets of monozygotic twins discordant for Beckwith-Wiedemann syndrome. Hum Mol Genet 11:1317–1325
44. DeBaun MR, Niemitz EL, Feinberg AP (2003) Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. Am J Hum Genet 72:156–160
45. Gicquel C, Gaston V, Mandelbaum J, Siffroi J-P, Flahault A, Le Bouc Y (2003) In vitro fertilization may increase the risk of Beckwith-Wiedemann syndrome related to the abnormal imprinting of the KCN1OT gene. Am J Hum Genet 72:1338–1341
46. Maher ER, Brueton LA, Bowdin SC, Luharia A, Cooper W, Cole TR, Macdonald F, Sampson JR, Barratt CL, Reik W, Hawkins MM (2003) Beckwith-Wiedemann syndrome and assisted reproduction technology (ART). J Med Genet 40:62–64
47. Halliday J, Oke K, Breheny S, Algar E, Amor D (2004) Beckwith-Wiedemann syndrome and IVF: a case-control study. Am J Hum Genet 75:526–528
48. Lim D, Bowdin SC, Tee L, Kirby GA, Blair E, Fryer A, Lam W, Oley C, Cole T, Brueton LA, Reik W, Macdonald F, Maher ER (2009) Clinical and molecular genetic features of Beckwith-Wiedemann syndrome associated with assisted reproductive technologies. Hum Reprod 24:741–747
49. Rossignol S, Steunou V, Chalas C, Kerjean A, Rigolet M, Viegas-Pequignot E, Jouannet P, Le Bouc Y, Gicquel C (2006) The epigenetic imprinting defect of patients with Beckwith-Wiedemann syndrome born after assisted reproductive technology is not restricted to the 11p15 region. J Med Genet 43:902–907
50. Bliek J, Verde G, Callaway J, Maas SM, De Crescenzo A, Sparago A, Cerrato F, Russo S, Ferraiuolo S, Rinaldi MM, Fischetto R, Lalatta F, Giordano L, Ferrari P, Cubellis MV, Larizza L, Temple IK, Mannens MM, Mackay DJ, Riccio A (2009) Hypomethylation at multiple maternally methylated imprinted regions including PLAGL1 and GNAS loci in Beckwith-Wiedemann syndrome. Eur J Hum Genet 17:611–619
51. Chang AS, Moley KH, Wangler M, Feinberg AP, DeBaun MR (2005) Association between Beckwith-Wiedemann syndrome and assisted reproductive technology: a case series of 19 patients. Fertil Steril 83:349–354
52. Sutcliffe AG, Peters CJ, Bowdin S, Temple K, Reardon W, Wilson L, Clayton-Smith J, Brueton LA, Bannister W, Maher ER (2006) Assisted reproductive therapies and imprinting disorders–a preliminary British survey. Hum Reprod 21:1009–1011

53. Bowdin S, Allen C, Kirby G, Brueton L, Afnan M, Barratt C, Kirkman-Brown J, Harrison R, Maher ER, Reardon W (2007) A survey of assisted reproductive technology births and imprinting disorders. Hum Reprod 22:3237–3240
54. Doornbos ME, Maas SM, McDonnell J, Vermeiden JP, Hennekam RC (2007) Infertility, assisted reproduction technologies and imprinting disturbances: a Dutch study. Hum Reprod 22:2476–2480
55. Lidegaard Ø, Pinborg A, Andersen AN (2005) Imprinting diseases and IVF: Danish National IVF cohort study. Hum Reprod 20:950–954
56. Kallen B, Finnström O, Nygren KG, Olausson PO (2005) In vitro fertilization (IVF) in Sweden: risk for congenital malformations after different IVF methods. Birth Defects Res A Clin Mol Teratol 73:162–169
57. Gomes MV, Huber J, Ferriana R, Amaral Neto A, Ramos E (2009) Abnormal methylation at the KvDMR1 imprinting control region in clinically normal children conceived by assisted reproductive technologies. Mol Hum Reprod 15:471– 477
58. Tierling S, Souren NY, Gries J, Loporto C, Groth M, Lutsik P, Neitzel H, Utz-Billing I, Gillessen-Kaesbach G, Kentenich H, Griesinger G, Sperling K, Schwinger E, Walter J (2010) Assisted reproductive technologies do not enhance the variability of DNA methylation imprints in human. J Med Genet 47:371–376
59. Steffenburg S, Gillberg CL, Steffenburg U, Kyllerman M (1996) Autism in Angelman syndrome: a population-based study. Pediatr Neurol 14:131–136
60. Kishino T, Lalande M, Wagstaff J (1997) UBE3A/E6-AP mutations cause Angelman syndrome. Nat Genet 15:70–73
61. Mann MR, Bartolomei MS (1999) Towards a molecular understanding of Prader-Willi and Angelman syndromes. Hum Mol Genet 8:1867–1873
62. Williams CA (2005) Neurological aspects of the Angelman syndrome. Brain Dev 7:88–94
63. Ørstavik KH, Eiklid K, van der Hagen CB, Spetalen S, Kierulf K, Skjeldal O, Buiting K (2003) Another case of imprinting defect in a girl with Angelman syndrome who was conceived by intracytoplasmic semen injection. Am J Hum Genet 72:218–219
64. Ludwig M, Katalinic A, Gross S, Sutcliffe A, Varon R, Horsthemke B (2005) Increased prevalence of imprinting defects in patients with Angelman syndrome born to subfertile couples. J Med Genet 42:289–291
65. Manning M, Lissens W, Bonduelle M, Camus M, De Rijcke M, Liebaers I, Van Steirteghem A (2000) Study of DNA-methylation patterns at chromosome 15q11-q13 in children born after ICSI reveals no imprinting defects. Mol Hum Reprod 6:1049–1053
66. Neri QV, Takeuchi T, Palermo GD (2008) An update of assisted reproductive technologies results in the United States. Ann N Y Acad Sci 1127:41–48
67. Moll AC, Kuik DJ, Bouter LM, Den Otter W, Bezemer PD, Koten JW, Imhof SM, Kuyt BP, Tan KE (1997) Incidence and survival of retinoblastoma in the Netherlands: a register based study 1862–1995. Br J Ophthalmol 81:559–562
68. Thompson JR, Williams CJ (2005) Genomic imprinting and assisted reproductive technology: connections and potential risks. Semin Reprod Med 23:285–295
69. Greger V, Passarge E, Höpping W, Messmer E, Horsthemke B (1989) Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. Hum Genet 83:155–158
70. Ohtani-Fujita N, Dryja TP, Rapaport JM, Fujita T, Matsumura S, Ozasa K, Watanabe Y, Hayashi K, Maeda K, Kinoshita S, Matsumura T, Ohnishi Y, Hotta Y, Takahashi R, Kato MV, Ishizaki K, Sasaki MS, Horsthemke B, Minoda K, Sakai T (1997) Hypermethylation in the retinoblastoma gene is associated with unilateral, sporadic retinoblastoma. Cancer Genet Cytogenet 98:43–49
71. Moll AC, Imhof SM, Cruysberg JR, Schouten-van Meeteren AY, Boers M, van Leeuwen FE (2003) Incidence of retinoblastoma in children born after in-vitro fertilisation. Lancet 361:309–310

Chapter 9
Imprinting Disorders of Early Childhood

I. Karen Temple, Jill Clayton-Smith, and Deborah J.G. Mackay

Contents

9.1 Introduction 138
9.1.1 Genetic Imprinting 138
9.2 Human Imprinting Disorders 139
9.2.1 Beckwith-Wiedemann Syndrome (BWS) 143
9.2.2 Silver Russell Syndrome (SRS) 146
9.2.3 Angelman and Prader-Willi Syndromes (Chromosome 15q11-13 Associated Imprinting Disorders) 147
9.2.4 Angelman Syndrome (AS) 149
9.2.5 Prader-Willi Syndrome 150
9.2.6 Transient Neonatal Diabetes (TND) 151
9.2.7 Chromosome 14 Related Imprinting Disorders 153
9.2.8 Pseudohypoparathyroidism Type 1b PHP1b 155
9.3 Conclusion 156
References 157

Abstract Imprinted genes are exceptional in that one allele is silenced (imprinted) in a parent of origin specific manner, making the two parental alleles functionally different. Imprinted genes are known to play a vital role in fetal growth and normal metabolism and most of the medical conditions caused by aberrant imprinting result in problems with growth, neurodevelopment and glycaemic control. There are eight known human disorders; Beckwith Wiedemann syndrome (fetal overgrowth, umbilical abnormalities and macroglossia), Silver Russell Syndrome (poor fetal and post natal

---

I.K. Temple (✉) • D.J.G. Mackay
Division of Human Genetics, Faculty of Medicine, University of Southampton, Southampton, UK
e-mail: i.k.temple@soton.ac.uk; D.J.G.Mackay@soton.ac.uk

J. Clayton-Smith
Genetic Medicine, Manchester Academic Health Sciences Centre, University of Manchester
e-mail: Jill.clayton-smith@cmft.nhs.uk

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_9,
© Springer Science+Business Media B.V. 2012

growth and short stature), Prader Willi syndrome (hypotonia, developmental delay, excessive appetite and obesity), Angelman syndrome (severe developmental delay, no speech and ataxia), Transient Neonatal Diabetes (neonatal diabetes and low birth weight), Temple syndrome (hypotonia short stature and early puberty), Wang syndrome (hypotonia, developmental delay and small thorax) and Pseudohypoparathyroidism type 1B (hypoparathyroidism). At any of the imprinted loci involved in these conditions, abnormal gene expression is due to three mechanisms: epigenetic loss or gain of DNA methylation within control regions, uniparental disomy and structural chromosome abnormalities. The known prevalence of each condition depends on the distinctiveness of symptoms and thus ease of diagnosis and the susceptibility of the different loci to genetic and epigenetic aberrations; the reasons for the difference in susceptibility at different loci is not yet known. This chapter describes the imprinted loci and the medical consequences of aberrant imprinted gene expression.

Abbreviations

AS Angelman syndrome
BWS Beckwith Wiedemann syndrome
DMR Differentially methylated region
ICR Imprinting control region
PHP1B Pseudohypoparathyroidism type 1B
PWS Prader Willi syndrome
SDS Standard deviation scores
SRS Silver Russell syndrome
TND Transient Neonatal Diabetes
TS Temple syndrome
WS Wang syndrome
UPD Uniparental disomy

9.1 Introduction

Genetic imprinting is often considered a perplexing and complex phenomenon with limited practical application to mainstream paediatric diagnosis. However, the exquisite molecular mechanisms conserved across placental mammals to maintain tissue- and stage-specific expression of imprinted genes argue powerfully for their importance in normal human development.

9.1.1 Genetic Imprinting

Imprinted genes are exceptional in that one allele is silenced (imprinted) in a parent of origin specific manner, making the two parental alleles functionally different

(for review see [1]). The best-characterised mechanism of genetic imprinting is DNA methylation, where gene expression is regulated by allele-specific methylation of cytosine nucleotides. In imprinted loci critical DNA sequences are ‘marked’ by cytosine methylation in the developing sperm or oocyte, forming primary differentially methylated regions (primary DMRs). These primary DMRs subsequently evade the normal resetting of DNA methylation that occurs in the embryo soon after fertilisation, and remain ubiquitous and permanent in the somatic DNA. There are more maternally derived genomic imprints reported than paternal ones. The DMRs function in conjunction with various associated molecules and processes to alter expression of genes within the influence of the imprint, although the pattern of expression is often complex, differing for each imprinted locus, and can be highly isoform and tissue-specific.

Imprinted gene products, which comprise <1% of the genome (http://igc.otago.ac.nz/home.html) include factors that play key roles in fetal growth and development, from RNA transcription regulators to growth factors. The kinship theory of genomic imprinting predicts that imprinted gene expression is balanced to promote optimal growth of the fetus, and specifically that paternally-expressed genes are fetal growth promoters while maternally-expressed imprinted genes are growth restrictors [2].

Critically, their functional hemizygosity makes imprinted genes susceptible to both genetic and epigenetic mutations. Errors in imprint establishment, erasure or maintenance can lead to aberrant gene expression and disease. Similarly genomic deletions, duplications or uniparental disomy (UPD) that involve imprinted regions can have a similar impact on gene expression. The relative proportion of each subtype of genetic or epigenetic aberration differs for each condition and this has a significant impact on recurrence risks for families. For example if Transient neonatal diabetes (TND) is due to an inherited paternal duplication at 6q24 then subsequent children, born to the father, will have a 50% risk of developing TND. However if the condition is due to uniparental disomy of chromosome 6, the risk to other relatives is very low.

9.2 Human Imprinting Disorders

There are eight recognised human disorders associated with imprinting aberrations (genetic and epigenetic) although it is likely that more will be reported as new imprinted loci are discovered; Angelman syndrome, Prader Willi syndrome, Beckwith Wiedemann Syndrome, Silver Russell syndrome, Transient neonatal diabetes (type 1), Temple syndrome, Wang syndrome and Pseudohypoparathyroidism type 1B. The main clinical features are shown in Table 9.1.

Taken together the conditions have a prevalence of at least 1 in 5,000 but the accurate figure for each disorder is only an estimate and likely to be higher than stated (Table 9.1). For example, one of the studies quoted on the prevalence of Angelman syndrome is from Steffenburg et al. in 1996 [3] who clinically evaluated prepubertal children in Sweden with developmental delay and epilepsy and identified four patients from a total population of 49,000 with clinical features fitting the

Table 9.1 Eight known imprinting disorders with a recognised clinical phenotype

Disease	Prevalence	Main diagnostic clinical features	Additional clinical features (may develop with time)	Frequency of 'epigenetic' aberration	Reference
Prader Willi syndrome	1 in 17,500	Low birth weight Hypotonia Hyperphagia Developmental delay	Hypogonadism Diabetes Obesity	<<1%	[35]
Angelman syndrome	1 in 16,000	Severe developmental delay No speech Epilepsy Ataxia	Microcephaly	4%	[4]
Beckwith Wiedemann syndrome	1 in 13,700	Macrosomia/overgrowth Macroglossia Umbilical defect	Increased risk of Wilms tumor Hypoglycemia	60%	[8]
Silver Russell syndrome	1 in 100,000 likely underestimate	Intrauterine growth retardation Faltering growth Short stature	Relative macrocephaly Genital abnormalities Hypoglycemia	50%	[12]
Transient neonatal diabetes	1 in 400,000	Intrauterine growth retardation Neonatal diabetes with remission	Macroglossia Umbilical hernia Developmental delay Diabetes	30%[a]	[38]

Temple syndrome (maternal UPD 14 associated syndrome)	Unknown	Intrauterine growth retardation Hypotonia Scolisosis Developmental delay Early puberty Short stature	Hydrocephalus Cleft palate	Uncertain	[46]
Wang syndrome (paternal UPD 14 associated syndrome)	Unknown	Bell shaped chest Hypotonia Developmental delay	Umbilical defects	Uncertain	[48]
Pseudohypoparathyroidism 1B	Unknown	Hypocalcaemia due to parathryoid resistance (tetany/parasthesia)		>90%+	[54]

[a]50% due to *ZFP57* mutations + Familial cases often have a deletion at *STX16*

criteria for Angelman syndrome. The diagnosis was confirmed molecularly in only two of the cases, but the study was performed before the advent of new technologies to identify epigenetic aberrations and without the benefit of gene sequencing of *UBE3A*. It is thus likely that this is an underestimate, as the features of Angelman can be variable and the differential diagnosis is broad. The prevalence figure used for Angelman syndrome in Table 9.1, 1 in 16,000, is based on an amalgamation of studies that included earlier estimates of a prevalence of 1 in 20,000 [4], almost certainly based on clinical diagnosis only. This is typical of the research on prevalence for these relatively rare disorders. For many of the other conditions the figures used are even less well substantiated.

Among the difficulties in determining the prevalence of imprinting disorders is the clinical presentation of patients, which is heterogeneous and often non specific. Furthermore, a lack of widely-adopted cost-effective high-throughput testing for epimutations, coupled with the expense of single gene sequencing and the lack of uniform methodology or quality control, puts the diagnosis of an imprinting disorder outside normal diagnostic practice. There is a real need for up-to-date surveys of many imprinting disorders, integrating clinical and molecular information, in order to establish their prevalence characteristics and future therapeutic interventions.

The same problems also confound estimation of the relative contribution of epigenetic versus genetic causes for each imprinting syndrome shown in Table 9.1. For example, the older literature on Transient neonatal diabetes [5] underestimates the prevalence of epigenetically-caused disease, due to small case series and a lack of available technology at the time of the investigation. Another significant confounder is somatic mosaicism, which is well recognised in imprinting disorders. For example, paternal UPD 11 in Beckwith Wiedemann syndrome is usually a post zygotic error and involves only a proportion of cells (see below). This causes problems on two levels: firstly, clinical findings may be restricted to certain regions of the body, making diagnosis difficult; secondly, testing of DNA from leukocytes derived from either blood or saliva, may not detect an abnormality that is restricted to non-hematologic tissues. Another possible confounding factor is the potential for epigenetic mutations to change over a lifetime, although this has never been demonstrated in imprinting disorders where most of the mutations demonstrated to date, are primary and stable.

There is some clinical overlap between imprinting disorders, irrespective of the imprinted locus involved (Table 9.1). Most are associated with disordered growth in fetal and early life. Developmental delay and feeding difficulties are features of some of the conditions but not all, and long term health problems include diabetes and obesity. Hypotonia is a significant component of many of the conditions in early childhood and while congenital abnormalities are not common in general, when they occur they tend to involve abdominal wall development, the genitalia, the tongue and palate. Unusual onset of puberty is another finding, but waiting for this clue delays diagnosis, and early management may be key to prevention of later health sequelae. One explanation for this overlap in clinical presentation is that some of the underlying aetiologies for imprinting disorders cause epimutations at more than one imprinted locus simultaneously. This is well recognised in Transient Neonatal Diabetes where 50% of the subgroup with epigenetic mutations at 6q24

also have epimutations at other imprinted loci such as 11p15.5. These cases are often due to mutations in *ZFP57* (see below) and are described as having Transient Neonatal Diabetes and Hypomethylation at Imprinted Loci (TND HIL) [6].

Despite all of these caveats, the diagnosis of each imprinting disorder discussed below can often be made by the astute clinician, while improving technologies such as routine array CGH and pyrosequencing mean that increasing numbers of cases are being recognised. It should also be noted that epigenetic research is fast-moving, and a patient in whom a positive molecular diagnosis was impossible 5 years ago may now be shown to have a well-defined molecular anomaly. An example of this is the identification in 2005 of a molecular anomaly accounting for ~50% of Silver-Russell Syndrome patients [7], where previously only ~10% of cases where amenable to positive diagnosis. For these reasons, a high-quality diagnostic suite, incorporating molecular analysis of DNA methylation, genetic sequencing, array analysis and cytogenetics, would be an invaluable adjunct to clinical diagnosis.

9.2.1 Beckwith-Wiedemann Syndrome (BWS)

The three cardinal clinical features originally described are macroglossia, gigantism at birth, and exomphalos (for review see [8]); additionally there is a predisposition to tumor development. In childhood there is a facial gestalt for BWS with a midline nevus flammeus of the forehead, infraorbital creases, a large jaw, mouth and tongue, and creases or pits in the ear lobes. Congenital abnormalities reported with BWS include cleft palate, visceromegaly, particularly of the kidneys, cardiac malformations and umbilical defects including exomphalos, umbilical hernia or diastasis recti. Weight and height can continue to increase and cross centiles for the first few years but then normalise and adult height is often within the normal range. Hypoglycemia has been reported in up to 50% in the neonatal period although it usually improves in the first few months. Feeding and respiratory difficulties are common, partially related to macroglossia. Intellectual development is usually normal.

Children with BWS are at increased risk of embryonal tumors, the most common being Wilms tumor followed by hepatoblastoma. This means that it is important to consider the diagnosis even if the features are minimal. Different molecular etiologies have a different risk of tumor development but overall the risk is in the region of 7.5% and significantly reduces over the age of 10 years [9]. This makes diagnosis and molecular classification of medical importance as it alters the surveillance programmes. It should also be noted that many individuals with BWS have a mild phenotype, and asymmetric limb hypertrophy may be the only manifesting feature [10]. This is typical for imprinting disorders where the extent of tissue involvement in a mosaic process makes the presentation variable. Indeed, Scott et al. [10] have reported 3% of individuals with isolated Wilms tumor to have a molecular diagnosis compatible with BWS but not associated with any growth abnormalities. This makes it impossible to instigate effective screening programs to eradicate this tumor, but means that 11p15.5 aberrations must always be tested for in such cases.

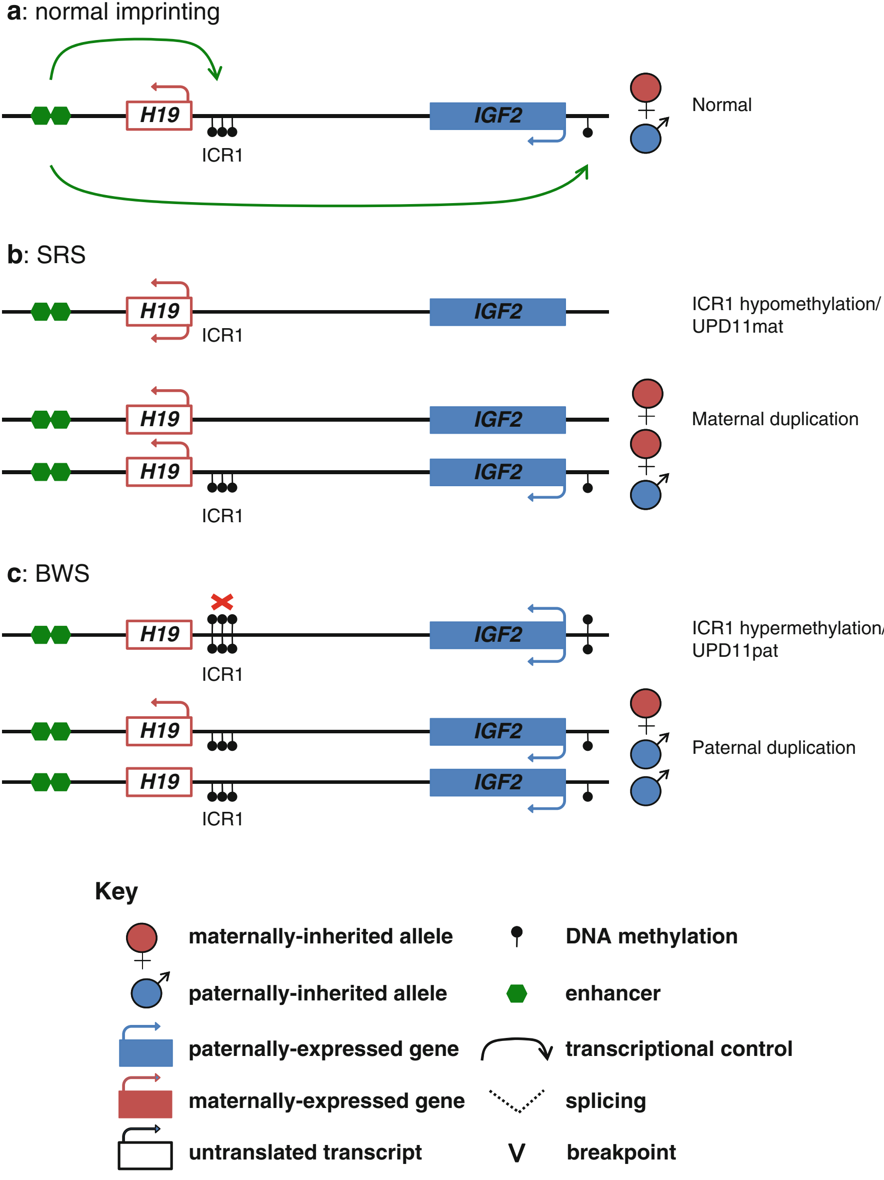
BWS is associated with genetic and epigenetic aberrations at 11p15.5, a region first implicated in patients with paternal duplications, although most such cases also had developmental delay, presumably due to duplication of non BWS genes in the region. 11p15 contains two independent imprinted domains, controlled by the imprinting control regions (ICRs) ICR1 and ICR2. Remarkably, BWS can result from aberrations of either or both. ICR1 is the more telomeric domain and controls two imprinted genes: *IGF2*, a fetal growth factor which is paternally expressed and *H19*, an untranslated RNA that is maternally expressed – see Fig. 9.1. Normally only the paternal allele is methylated at ICR1. The unmethylated maternal allele enables the binding of an insulator protein, CTCF, which prevents common enhancers acting on *IGF2*; hence *IGF2* is expressed only from the paternal allele. Hypermethylation of the maternal ICR1, resulting in maternal IGF2 expression, accounts for 5% of BWS cases.

ICR2, normally methylated on the maternal allele, regulates expression of the maternally-expressed growth restricting gene *CDKN1C* (see Fig. 9.2). ICR2 lies within a non imprinted gene *KCNQ1* and is the promoter for *KCNQ1OT1*, a paternally-expressed RNA whose expression regulates *CDKN1C* in *cis*. Hypomethylation at ICR2, resulting in reduced maternal expression of *CDKN1C*, is the commonest molecular finding (50%) in BWS. In 10% of cases there is a gene mutation in *CDKN1C*.

Partial paternal uniparental disomy (UPD) 11 (20%), usually as a result of a post zygotic error at mitosis, impacts on both domains and results in increased expression of *IGF2* rather than causing a known impact on *CDKN1C* expression.

There is a relation between molecular etiology and BWS features [11]. Most cases with exomphalos have an ICR2 abnormality while Wilms tumor is associated with ICR1 aberrations or paternal UPD 11, both of which result in overexpression of IGF2.

Fig. 9.1 H19/IGF2 Imprinted region on chromosome 11p15. Arrows and symbols above and below the line represent maternal and paternal alleles respectively. (**a**) Schematic of normal imprinting. The paternally-methylated imprinting control region (ICR: *black* "lollipops") lies within the promoter of the untranslated maternally-expressed transcript H19 (*red* unfilled oblong). Repetitive motifs within ICR1 bind the chromatin regulator CTCF in the unmethylated state, constituting a chromatin boundary. Other repeats in ICR1 bind OCT-family transcriptional regulators that are required to maintain the methylated state. The ICR forms a boundary element, regulating the expression of IGF2 (*blue* oblong) under the control of distant enhancer elements (*green* hexagons). The methylated ICR of the paternal allele does not bind CTCF, and therefore the distant enhancers support IGF2 expression. The binding of CTCF to the unmethylated maternal ICR prevents enhancer operation on IGF2, and instead supports expression of H19 (**b**) Mutations and epimutations involving ICR1 in SRS. Inappropriate hypomethylation of the paternal allele of ICR1 makes it behave effectively as a maternal allele, with H19 transcription but reduced IGF2 expression. Maternal uniparental disomy or duplication of the maternal allele of chr11 has a similar effect by causing under-representation of paternal alleles. (**c**) Mutations and epimutations involving ICR1 in BWS. Inappropriate methylation of the paternal allele of ICR1 makes it behave effectively as a paternal allele, with reduced H19 transcription but increased IGF2 expression; a proportion of such cases have microdeletions or mutations within ICR2 disrupting establishment or maintenance of methylation on the maternal allele. Paternal uniparental disomy or duplication of the paternal allele of chr11 has a similar effect by causing under-representation of maternal alleles


a: normal imprinting
H19
ICR1
IGF2
Normal
b: SRS
H19
ICR1
IGF2
ICR1 hypomethylation/
UPD11mat
H19
IGF2
H19
ICR1
IGF2
Maternal duplication
c: BWS
H19
ICR1
IGF2
ICR1 hypermethylation/
UPD11pat
H19
IGF2
H19
ICR1
IGF2
Paternal duplication
Key
maternally-inherited allele
paternally-inherited allele
paternally-expressed gene
maternally-expressed gene
untranslated transcript
mutation
DNA methylation
enhancer
transcriptional control
splicing
V breakpoint

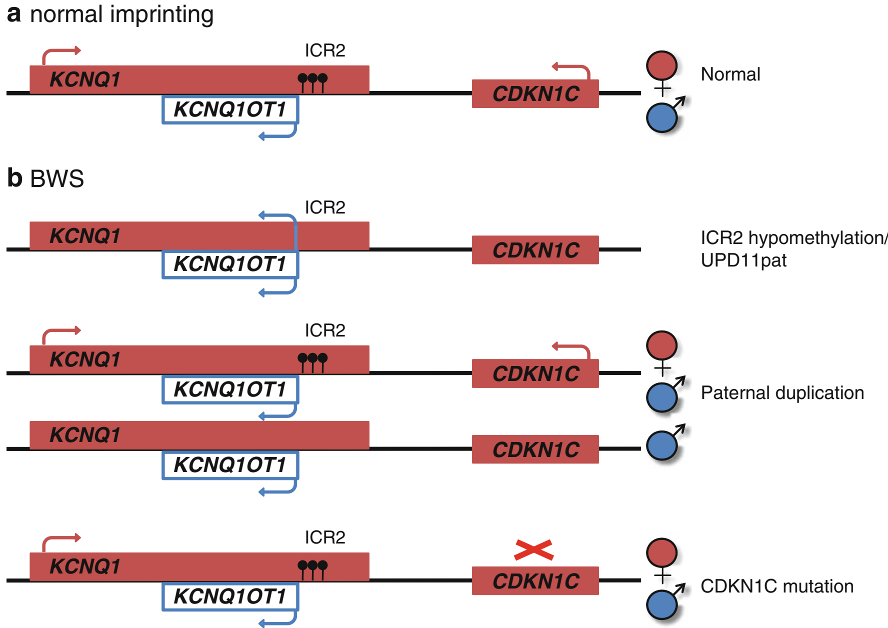


Fig. 9.2 KCNQ1OT1/CDKN1C imprinted region on chromosome 11p15. Arrows and symbols above and below the line represent maternal and paternal alleles respectively. (**a**) Schematic of normal imprinting. The maternally-methylated imprinting control region (ICR2: *black* "lollipops") lies within the promoter of the untranslated paternally-expressed transcript KCNQ1OT1 (*blue* unfilled oblong), which in turn lies within an intron of the maternally-expressed KCNQ1 gene (*red* filled oblong). Paternal expression of KCNQ1OT1 suppresses paternal expression of KCNQ1, and also of the cell-cycle regulator CDKN1C (*red* oblong), such that these proteins are maternally expressed. (**b**) Mutations and epimutations involving ICR2 in BWS. Inappropriate hypomethylation of the maternal allele of ICR2 makes it behave effectively as a paternal allele, resulting in increased transcription of KCNQ1OT1 and reduced CDKN1C expression. Paternal uniparental disomy or duplication of the paternal allele of chr11 has a similar effect by causing under-representation of maternal alleles. Some patients have coding mutations of the maternally-inherited allele of CDKN1C

9.2.2 *Silver Russell Syndrome (SRS)*

This is a condition characterised by intrauterine growth restriction and poor post natal growth with sparing of head growth [12]. Price et al. [13] described five key features:– a birth weight < −2 SD, height < −2 SD, preservation of head circumference, classic triangular facial features with frontal bossing, micrognathia and evidence of body asymmetry. Additional associated features may include patches of increased skin pigmentation and 5th finger clinodactyly, and less commonly cleft palate, genital abnormalities, heart defects and camptodactyly with limitation of peripheral joint mobility. Severe developmental delay is rare but 30% were found to have some mild delay in a recent review of 64 patients with a confirmed diagnosis [12]. The natural history of this condition includes significant feeding difficulties and early motor delay.

Speech problems are a feature in some patients. Postnatal growth is somewhat variable and it is not yet understood why some continue to fall from the centiles in early childhood whereas in others there is no change or an improvement in height standard deviation score (SDS) postnatally [14]. Few studies are available on adult height but of interest some achieve a final height within normal centiles. SRS is an indication for growth hormone treatment but this remains controversial. There is a tendency to central weight gain in older children and adults, and some facial features become less obvious with age. The relatively non specific clinical findings make this condition difficult to diagnose and differentiate from other causes of short stature.

SRS is usually sporadic and for many years after the condition was discovered no consistent cause could be identified. It was not certain that there was a genetic aetiology as the majority of monozygotic twins are discordant for SRS [15] and only a few families are reported with multiple affected members [16]. It is now clear, however, that imprinted genes on chromosomes 7 and 11 are implicated in the SRS phenotype and account for approximately 50–60% of cases with typical SRS.

ICR1 on chromosome 11p15.5 was first implicated in SRS through rare cases with maternal 11p duplication [17] (Fig. 9.1). In 2005, Gicquel et al. [7] showed that up to 40% of cases with classic SRS had hypomethylation at ICR1 (i.e. resulting in silencing of IGF2 and the inverse molecular anomaly to the ICR1 hypermethylation seen in BWS (characterized by biallelic IGF2 expression). This is now recognised as the major molecular cause of SRS, and has been confirmed in many other studies [18, 19]. Mosaic segmental maternal uniparental disomy of 11p15 has also been reported in SRS [20]. Interestingly, epigenetic and genetic mutations at ICR2 are not commonly associated with SRS although there is one reported case with isolated duplication restricted to ICR2 [21].

Kotzot [22] was the first to demonstrate maternal uniparental disomy (UPD) of chromosome 7, the inheritance of both chromosome 7s for the mother with no contribution from the father, in SRS. Maternal UPD7 is believed to account for 5–10% of SRS [23], and is nearly always heterodisomic, leading to speculation that it results from trisomy rescue and therefore that a trisomic cell line may contribute to the phenotype. It can be difficult to differentiate matUPD7 patients from those with other aetiologies [24] although Binder et al. showed a greater likelihood of having reduced postnatal height [14], and Wakeling et al. [12] noted an increased likelihood of developmental delay and a less classical presentation. There are at least three candidate imprinted regions on chromosome 7 at 7p12 (*GRB10*); 7q21 (*SGCE)* and 7q31-2 (*PEG1/MEST)* but there is no clarity on which of these loci may contribute most to SRS in matUPD7.

9.2.3 Angelman and Prader-Willi Syndromes (Chromosome 15q11-13 Associated Imprinting Disorders)

The imprinted chromosome 15q11-13 region contains both maternally and paternally expressed genes, summarised in Fig. 9.3. Interference with expression of the maternally expressed gene *UBE3A* gives rise to Angelman syndrome (AS) whereas a

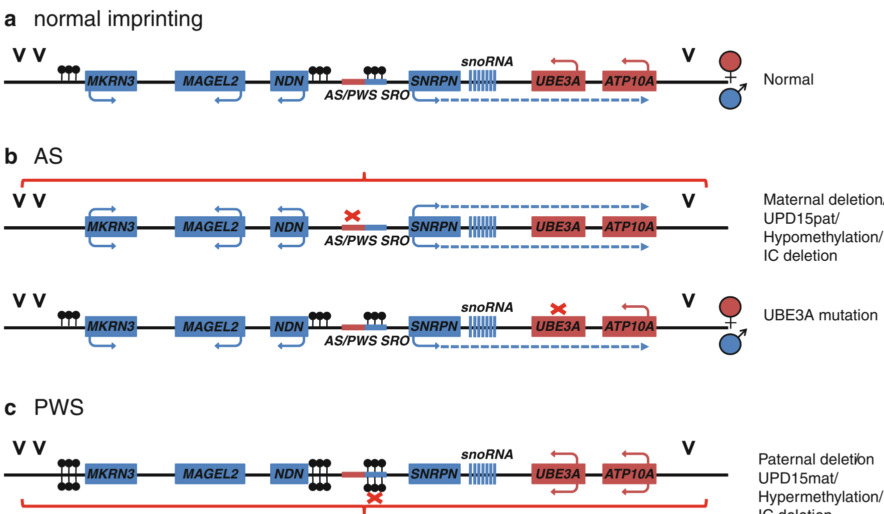


Fig. 9.3 PWS/AS imprinted region on chromosome 15q13. Arrows and symbols above and below the line represent maternal and paternal alleles respectively. (**a**) Schematic of normal imprinting. The maternally-methylated imprinting control region (ICR: *black* "lollipops") lies within the promoter of the SNRPN gene (*blue* filled oblong), which also forms the origin of a long transcript that generates multiple non-coding RNAs. Paternal expression from this promoter suppresses paternal expression of UBE3A and ATP10, such that these proteins are maternally expressed (*red* oblongs) while proximal to SNRPN are several paternally expressed genes (*blue* oblongs) with differential methylation (*black* lollipops) that develops secondarily to the germline methylation of SNRPN. Arrowheads to the left and right of the schematic indicate regions of genomic instability that are frequently observed as breakpoints in chromosomal rearrangements in PWS and AS. The narrow red and blue bars abutting the SNRPN imprinting centre represent the AS/PWS shortest region of overlap (AS/PWS SRO), a minimal region in deletions, and therefore believed to be minimal promoter elements for genes involved in these disorders. (**b**) Mutations and epimutations in PWS. Inappropriate methylation of the paternal ICR makes it behave effectively as a maternal allele, resulting in reduced expression of SNRPN and associated RNAs; some hypermethylation cases are caused by microdeletion involving the PWS-SRO. Maternal uniparental disomy or duplication of the maternal allele of chr15 has a similar effect by causing under-representation of paternal alleles. (**c**) Mutations and epimutations in AS. Inappropriate hypomethylation of the maternal ICR makes it behave effectively as a paternal allele, abrogating transcription of UBE3A; some hypomethylation cases are caused by microdeletion involving the AS-SRO. Paternal uniparental disomy or duplication of the paternal allele of chr15 has a similar effect by causing under-representation of maternal alleles. AS is also caused by coding mutations of the maternally-inherited allele of UBE3A

clinically different disorder, Prader Willi Syndrome (PWS), results if genes which are normally paternally expressed are affected. The expression of imprinted genes within 15q11-13 is controlled by a bipartite imprinting centre (IC) [25]. Part of this IC, termed AS-SRO, is involved with setting of the maternal imprint of genes within the region and the adjoining part, PWS-SRO, affects setting of the paternal imprint. Angelman and Prader-Willi syndrome can result from deletion of 15q11-13, uniparental disomy (UPD) for this region, or from impaired imprinting. Most deletions occur *de novo*, though a predisposing parental chromosomal rearrangement must be excluded.

Uniparental disomy is associated with increasing parental age, and with increased risk if a parent carries a Robertsonian translocation involving chromosome 15. The frequencies of all of the genetic mechanisms and the parental origin of the chromosome involved differ between the two conditions and are discussed further below. An excellent review of genetic testing for AS and PWS was recently published by Ramsden et al. [26].

9.2.4 Angelman Syndrome (AS)

The main characteristics of AS include severe learning disability, ataxia, seizures with a characteristic high voltage EEG, acquired microcephaly and a sociable affect with frequent smiling and laughing [4, 27]. Subtle dysmorphic features include brachycephaly, deep set eyes, a wide mouth which is usually held open and a prominent chin. Infants with AS have significant feeding difficulties. Developmental milestones are delayed and speech rarely develops. Most children achieve walking with a characteristic wide-based, stiff legged gait. The behavioural features of AS include hand flapping, especially when excited, jerky voluntary movements, love of water, sleep disturbance and easily provoked laughter. Though head circumference at birth is usually normal, the majority of children develop microcephaly. Seizures, variable in type and difficult to control, typically develop in the second year of life, lessen in adolescence and return during adulthood. Puberty occurs normally and though final adult height is usually within the normal range truncal obesity may develop, along with joint contractures, spasticity and scoliosis. Life span may be shortened if problems develop secondary to immobility, scoliosis or seizures.

AS results from impaired expression of the maternally imprinted *UBE3A* gene, whose expression is restricted to the brain, and this can be caused by gene mutations, UPD, deletions, and imprinting anomalies.

A small proportion of AS patients have point mutations within the coding region of *UBE3A* [28]. Though the majority are de novo, around 25% are inherited from a phenotypically normal mother and there is also a high risk of maternal gonadal mosaicism. Recurrence risks may therefore be high and study of the extended family is indicated. UBE3A functions as both a ubiquitin protein ligase within the brain, targeting proteins for degradation, and as a transcriptional activator. It is still not clear, however, just how *UBE3A* is involved in the pathogenesis of AS. Animal studies have shown that *Ube3a* is preferentially expressed in neurones where it appears to affect dendritic spine development [29]. It may also affect the number of available synaptic glutamate receptors [30]. Thus, the symptomatology of AS could be due to abnormalities of synaptic development and plasticity.

The majority of AS individuals (70–75%) have a maternally inherited deletion of 15q11-13, while 3–7% have paternal UPD15 [31]. Maternal hypomethylation is present in 4% of AS patients. In 10–15% of these cases, which may be familial, the hypomethylation is associated with small deletions in the AS IC; in the remaining 85–90% the abnormal imprint is presumed to be a chance epigenetic event [31].

An increased incidence of AS with imprinting defects has been reported after assisted conception, possibly arising from the assisted conception process itself or from underlying infertility, though the small number of cases reported makes it difficult to draw any definite conclusions [32]. Imprinting defects frequently occur in mosaic form, in which case the phenotype may be subtly different with more pronounced hypotonia, larger head size, better developed language skills and predisposition to obesity [33].

There are minor phenotypic differences between individuals with AS due to different genetic mechanisms [34]. Those with deletions are more likely to have seizures and microcephaly and usually have hypopigmentation of hair, skin and irides as they lack one allele of the *P* gene, implicated in Type II oculocutaneous albinism. Those with *UBE3A* mutations are normally-pigmented with a lower incidence of seizures and microcephaly. Patients with UPD are the most able and may develop some speech. There is a relatively broad spectrum of severity in individuals with imprinting mutations, reflecting the fact that many are mosaics.

9.2.5 Prader-Willi Syndrome

Prader-Willi syndrome (PWS) is one of the commoner imprinting disorders [with an incidence of 1 in 17,500 (for review see [35]). The high incidence is due to the susceptibility of this locus to undergo deletion and duplication errors rather than epigenetic mutations which make up a low proportion of cases. Its main characteristics are profound neonatal hypotonia and initial feeding difficulties followed by the development of hyperphagia and obesity; moreover, most affected individuals have mild to moderate learning disability. Hypogonadism is usually present and individuals with PWS have specific behavioural features including a tendency to pick at their skin, temper tantrums and obsessive compulsive disorder. Other complications include seizures and scoliosis. Small hands and feet have been noted in some patients and hypopigmentation is a feature in those with 15q11-13 deletions. Characteristic facial features include bitemporal narrowing of the face, almond-shaped palpebral fissures and a downturned mouth. The management of PWS is aimed at weight regulation using behavioural modification techniques, healthy eating and exercise programmes. In recent years, growth hormone treatment has improved the outlook for children with PWS by improving linear growth, reducing obesity and improving muscle mass so that exercise is easier for affected individuals [36]. Lifespan in PWS can, however, be significantly reduced by cardiovascular and respiratory problems occurring as a result of morbid obesity.

Around 75–80% of individuals with PWS have a paternally inherited deletion of the 15q11-13 region (Fig. 9.3). The remaining 20–25% virtually all have maternal uniparental disomy (UPD) for chromosome 15. UPD usually arises due to maternal non-dysjunction of the chromosome 15s during meiosis and is more common in older mothers. Imprinting defects of 15q11-13 occurs in less than 1% of PWS cases.

The 15q11-13 region contains five protein-coding genes and five snoRNA gene clusters with highly repetitive sequences (Fig. 9.3). Sahoo et al. [37] identified a microdeletion affecting only one of the snoRNA clusters, *SNORD116* (HB11-85) in a patient with a fairly typical PWS phenotype and suggested that this may be the critical region for PWS. There were, however, some atypical features including better growth than expected. It seems likely that other genes in the 15q11-13 region contribute to the PWS phenotype and certainly, the hypopigmentation seen in deletion patients is influenced, as in AS, by haploinsufficiency for the Type II oculocutaneous albinism gene. The *SNORD116* snoRNA has been shown to affect splicing of a serotonin receptor but there is no evidence to support this as a mechanism in PWS and the cause of the condition remains unknown.

9.2.6 Transient Neonatal Diabetes (TND)

Transient neonatal diabetes (type 1) (TND1) is a disorder of fetal growth and glycaemic control (for review see [5, 38, 39]). Babies present with hyperglycaemia, typically in the first week of life, and require insulin replacement for the first few months. Insulin levels are often undetectable at birth although ketoacidosis is rare implying that endogenous insulin is being produced. These babies are characteristically growth retarded with a paucity of body fat and may have other associated features including macroglossia and umbilical hernia. Although diabetes remits by 18 months of age these children are susceptible to episodes of hyperglycaemia during childhood, particularly during intercurrent illness, and a high proportion develop type 2 diabetes in later life.

The imprinted locus on 6q24 is one of the least complex human imprinted domains and contains just two known imprinted genes *PLAGL1* and *HYMAI,* where only the paternally inherited allele is normally expressed in the majority of tissues (Fig. 9.4). This expression pattern results from a primary methylation imprint established in the oocyte. The resulting DMR overlaps the main shared promoter of *PLAGL1* and *HYMAI*, and maternal promoter methylation silences the maternal alleles. *PLAGL1* has a second promoter 55Kb upstream from the imprinted promoter which is not imprinted and results in biallelic expression in some tissues including blood and adult pancreas in normal individuals. *PLAGL1* is a zinc finger DNA binding protein with tumor suppressor activity, involved in cell cycle control and apoptosis. Loss of *PLAGL1* expression has been described in a number of cancers. Murine studies have shown that overexpression of *PLAGL1* results in reduced fetal pancreatic beta cell mass [40]. The other imprinted gene, *HYMAI,* is a non-coding RNA gene transcribed in the same orientation as *PLAGL1,* and its function is not known.

TND is caused not by loss of function but by overexpression of imprinted genes. Three mechanisms are recognised: (i) paternal uniparental disomy of chromosome 6 involving 6q24 (35% of cases), (ii) paternal duplication of 6q24 (35% of cases), (iii) maternal hypomethylation of the TNDM1 DMR (in the absence of a genomic rearrangement; 30% of cases). Maternal hypomethylation is usually complete in

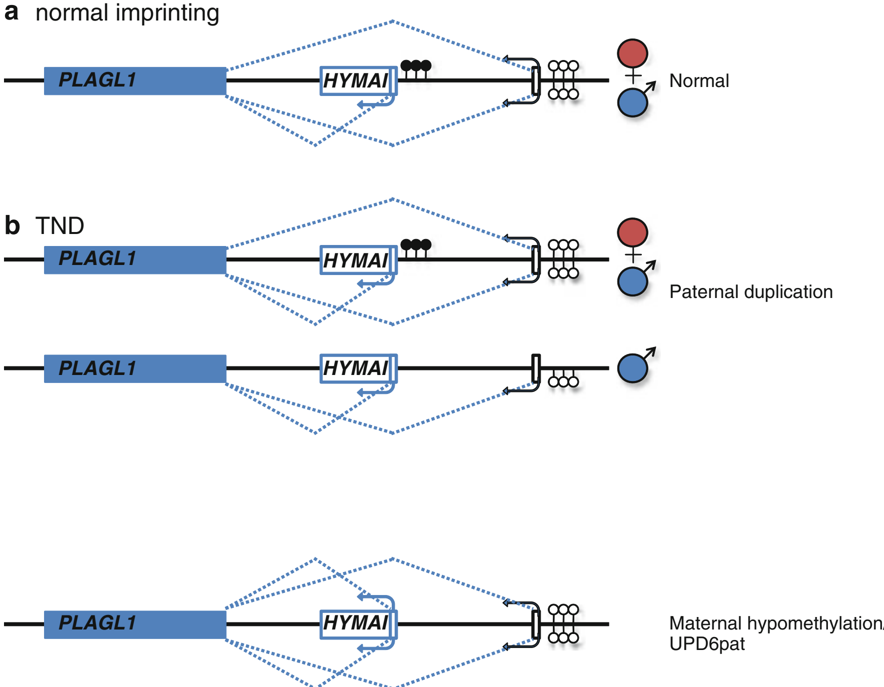


Fig. 9.4 PLAGL1 imprinted region on chromosome 6q24. Arrows and symbols above and below the line represent maternal and paternal alleles respectively. (**a**) schematic of normal imprinting. The maternally-methylated imprinting control region (ICR: *black* "lollipops") lies within an internal promoter of PLAGL1 (*blue* filled oblong), which is also the transcription origin of the untranslated transcript HYMAI (blue unfilled oblong). Maternal methylation of the ICR suppresses expression from that allele, such that PLAGL1 is expressed only from the paternal allele. The upstream promoter of PLAGL1 is unmethylated and supports biallelic expression. (**b**) Mutations and epimutations in TND. Inappropriate hypomethylation of the maternal ICR makes it behave effectively as a paternal allele, resulting in increased transcription of PLAGL1. In some patients this hypomethylation is secondary to homozygous mutation of *ZFP57*. 6q24 duplication of paternal origin has a similar effect by causing under-representation of maternal alleles

blood DNA and its cause is unknown, but as-yet undiscovered genetic mutations in cis may cause the failure of methylation of the maternal allele. In 50% of cases, hypomethylation affects other imprinted loci in addition to 6q24, and in many of these the pattern of loci involved is consistent. Mackay et al. [6] found that these cases were associated with homozygous/compound heterozygous mutation of *ZFP57,* a zinc finger transcription factor expressed in very early embryo development. Therefore this is an example of a genetic cause of global epigenetic aberrations at imprinted loci. Murine studies demonstrated that loss of *zfp57* function disturbs both establishment and maintenance of DNA methylation at a subset of imprinted loci [41].

The clinical phenotype of this autosomal recessive inherited global imprinting disorder has been very variable. While probands have presented with neonatal diabetes,

reflecting the involvement of 6q24, the additional features have ranged from severe developmental delay with brain malformations and early childhood death to normal development with no associated features. Non-probands have been reported with no health issues. The mechanism underlying the relation between *ZFP57* and differential DNA methylation is not known but the presence of a KRAB domain has linked it to the maintenance of the molecular scaffold required for DNA methylation. The cause of the other cases with hypomethylation at multiple loci but presenting with TND1 is not known although monozygous twinning and patients where the parents have received in-vitro fertilisation (IVF) are over-represented in this group [6].

9.2.7 Chromosome 14 Related Imprinting Disorders

The first reports of imprinting disorders in humans associated with chromosome 14 were those of Wang et al. [42] and Temple et al. [43] who reported individual cases with paternal and maternal uniparental disomy of chromosome 14, respectively. The phenotype of both disorders is now well characterised [44]. It has subsequently become clear that the phenotypes are indistinguishable from genetic and epigenetic aberrations confined to 14q32 which variously alter expression of the imprinted genes within the imprinted domain (see Fig. 9.5). Hence it has been suggested that these two conditions should be named after the first authors of the first reports rather than the somewhat cumbersome name of 'uniparental disomy of chromosome 14 associated conditions' [45].

9.2.7.1 Wang Syndrome (Paternal UPD 14 Associated Conditions)

The condition may be diagnosed antenatally with ultrasound evidence of a small thorax, polyhydramnios and large abdominal circumference that can be confused with exomphalos. Babies have normal or relatively high birth weight for gestational age but often have serious respiratory difficulties associated with a bell shaped chest and characteristic chest X-ray findings described as a 'coat hanger' appearance of the ribs. This can be very helpful for diagnosis of the condition. The children may be hypotonic and have abdominal wall defects, most commonly diastasis recti. Dysmorphic features include a high hirsute forehead, blepharophimosis, a small chin and finger contractures. The overall prognosis of this condition is poor and many children die in infancy. The children are reported to have developmental delay but there are relatively few reports of older children.

9.2.7.2 Temple Syndrome (Maternal UPD 14 Associated Conditions)

In contrast this disorder has a good prognosis, and is characterised by pre- and postnatal growth restriction, hypotonia leading to joint laxity, scoliosis which can be severe, and delayed motor milestones [46]. The children may initially be diagnosed

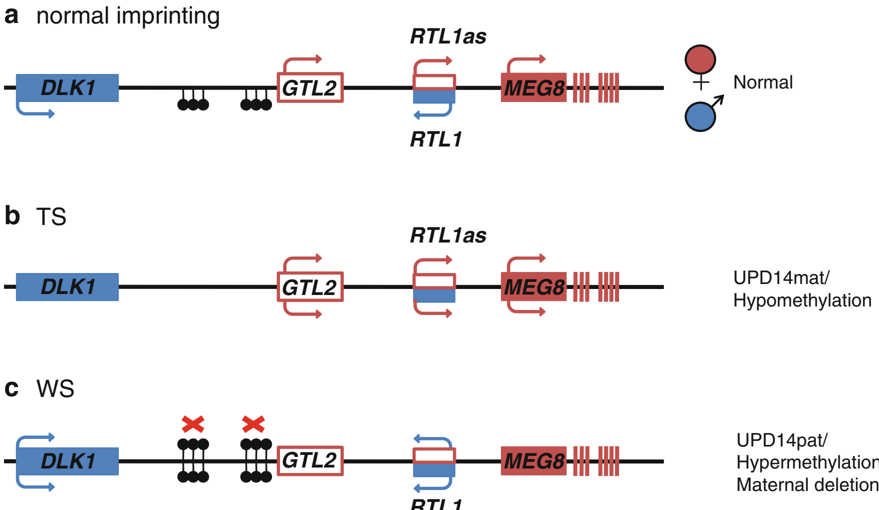


Fig. 9.5 The imprinted region on chromosome 14q32. Arrows and symbols above and below the line represent maternal and paternal alleles respectively. (**a**) schematic of normal imprinting. The paternally-methylated ICR (*black* "lollipops") is intergenic, lies between the untranslated maternally-expressed transcript GTL2 (*red* unfilled oblong) and the paternally-expressed gene DLK1 (blue filled oblong). Differential methylation of the GTL2 promoter is secondary to that of the intergenic DMR (IG-DMR). The IG-DMR regulates parent of origin-specific expression of multiple genes including numerous non-coding RNAs (red lines). (**b**) Mutations and epimutations of the 14q32 ICR in TS. Inappropriate hypomethylation of the paternal ICR makes it behave effectively as a maternal allele. Maternal uniparental disomy or paternal deletion of 14q32 have a similar effect by causing under-representation of paternal alleles. (**c**) Mutations and epimutations of the 14q32 ICR in WS. Inappropriate hypermethylation of the maternal ICR makes it behave effectively as a paternal allele. Paternal uniparental disomy or maternal microdeletion of 14q32 have a similar effect by causing under-representation of maternal alleles. Microdeletions of the IG-DMR and the secondary GTL2-DMR have different effects on placental and fetal development

as possible Prader Willi syndrome (PWS) and there have been several studies diagnosing Temple syndrome in cases negative for PWS [47]. The facial features are however distinct, with a high broad forehead and wide nasal base with prominent lips and a high, narrow or cleft palate. The head circumference is often large in proportion to height, and hands and feet are particularly small. Early puberty is well described although can be normal. The final height is often reduced and can be well below the 0.4th centile.

The imprinted locus at 14q32 (Fig. 9.5) has both maternally (*GTL2/MEG3, MEG8, RTL1as*) and paternally (*DLK1, RTL1*) expressed genes and an intergenic primary imprinting control region (IG-DMR). Hypomethylation at IG-DMR is associated with expression of *GTL2* and *RTL1as* which repress the expression of *DLK1* and *RTL1* in cis. The methylated paternal copy of IG-DMR is therefore associated with repression of *GTL2* and *RTL1as* and expression of *DLK1* and *RTL1* from the paternal allele. A series of microdeletions in rare families has shown that

the Wang syndrome is likely due to over-expression of *RTL1* and the smallest region implicated in either deletions or epimutations (hypermethylation) is the IG- DMR [48], whereas evidence from microdeletion families showed that Temple syndrome was due to a loss of expression of paternally expressed 14q32 domain genes. Kagami et al. [48] hypothesised that deletions involving *RTL1* and *DLK1* had a more severe phenotype with regard to height than those with a *DLK1* deletion alone. This is in keeping with other cases that have been reported with isolated epimutations and a similar phenotype [49]. *RTL1* (retrotransposon-like-1) expression is therefore key in both phenotypes. The paternally-expressed *DLK1* (delta, Drosophila homolog-like1) is a transmembrane signalling protein and regulator of growth. It has homology to proteins in the Notch/delta pathway.

9.2.8 Pseudohypoparathyroidism Type 1b PHP1b

This condition comprises hormone resistance (parathyroid and thyroid stimulating hormone) but only minor physical features of the related condition Pseudohypoparathryoidism type 1A which includes short stature, skeletal changes and variable developmental delay [50–52].

The imprinted *GNAS* cluster (Fig. 9.6) encodes several transcripts with different promoters (reviewed in [53, 54]); the main transcript, Gsα, is biallelic in most tissues but maternally expressed in certain tissues such as the anterior pituitary, the thyroid and the ovary. The protein stimulates adenylate cyclase in response to ligand binding at G-protein coupled receptors. Other transcripts at the locus include (i) *GNASXL,* paternally-expressed in some neuroendocrine tissues, whose protein product XLαs stimulates adenylate cyclase, (ii) *NESP55,* expressed from the maternal allele in some neuroendocrine tissues and encoding neuroendocrine secretory protein 55, a chromogranin-like protein, (iii) *GNAS exon 1A*, a paternally-expressed non-coding transcript originating upstream of exon 1of Gsα , in exon 1A, and E (iv) *NESPAS* another paternally-expressed non-coding transcript transcribed antisense to NESP55 and involved in its silencing.

The tissue-specific imprinting of genes in the *GNAS* cluster accounts for the various clinical presentations and explains why aberrations result in hormone resistance in only some tissues. Imprinted expression is controlled by 3 DMRs; exon 1A DMR (methylated on the maternal allele) regulating the *GNAS* promoter; *GNASXL/NESPAS* promoter DMR (methylated on the maternal allele) responsible for the imprinted expression pattern of both transcripts and the ICR for the locus, and the *NESP55* promoter DMR (methylated on the paternal allele).

PHP1B is nearly always due to a loss of methylation at the exon 1A DMR and no identified mutation in *GNAS*. DNA methylation analysis is therefore the test of choice when diagnosing this condition. In most sporadic PHP1B cases the cause is unknown and there are variable methylation defects at the other *GNAS* DMRs; however, in familial cases, hypomethylation restricted to the exon 1a DMR is often associated with microdeletion of Syntaxin-16 (*STX16*), up-steam of the *GNAS* locus

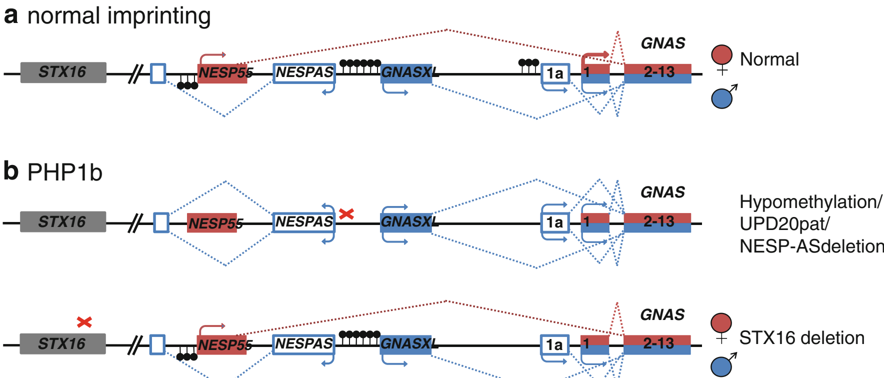


Fig. 9.6 The GNAS imprinted region on chromosome 20q13. Arrows and symbols above and below the line represent maternal and paternal alleles respectively. (**a**) schematic of normal imprinting. The maternally-methylated germline ICR (*black* "lollipops") lies within the promoter of the untranslated paternally-expressed transcript NESP-AS (*blue* unfilled oblong). Secondary differential methylation is established on the maternal allele in the promoter of the adjacent gene, GNASXL (*blue* filled oblong), and on the paternal allele intergenically between NESP-AS and NESP55 (*red* filled oblong). The GNAS exon 1a promoter harbours another ICR whose maternal methylation can be disrupted either in tandem with or independently of that of NESP-AS. Expression of GNAS is biallelic, but predominantly maternal in some tissues. (**b**) Mutations and epimutations in PHP1b. Inappropriate hypomethylation of the maternal NESP-AS ICR makes it behave effectively as a paternal allele, causing paternalisation of all other DMRs in the cluster and leading to reduced GNAS expression. In a small number of cases NESP-AS hypomethylation is caused by microdeletion within the NESP-55 DMR. Paternal uniparental disomy of chr20 has a similar effect by causing effective paternalisation of the region. Maternal deletion of the distant STX16 gene disrupts local chromatin organisation and leads to hypomethylation of the GNAS exon 1a DMR and resultant loss of GNAS expression. Other DMRs in the region are unaffected

[55–57]. This is therefore a poignant reminder that epigenetic aberrations can be related to genetic mutations and can thus be inherited. It is thought that the *STX16* locus harbours a *cis*-acting DNA element necessary for methylation at the exon 1A DMR. It is not yet known whether DNA hypomethylation in some sporadic PHP1B cases are caused by mutations in *cis* or mutations of regulatory genes in trans, and this makes genetic counselling difficult in families with isolated epigenetic aberrations.

9.3 Conclusion

Imprinting disorders are a diverse and fascinating group of conditions, that differ with respect to their relative prevalence. To date all imprinting disorders have been found to be due to a combination of gene mutations, epimutations and cytogenetic errors at specific imprinted loci (Table 9.1). While each imprinting disorder has its

own unique clinical and molecular characteristics, there can be substantial overlap between their phenotypic associations and fetal growth disturbance is the major finding common to many of the disorders. Clinical overlap may be due to common cellular pathways for many of the known imprinted genes but may also reflect the involvement of more than one imprinted locus at a time, so called hypomethylation at multiple imprinted loci. An increasing case literature shows that the clinical and molecular heterogeneity of these disorders currently requires imaginative and insightful diagnosis. Furthermore it is likely that additional phenotypes will be recognised as new imprinted loci are discovered and as diagnosis improves.

The complexity at the molecular level, means that expert interpretation is required when predicting prognosis and recurrence risks for families.

References

1. Reik W, Walter J (2001) Genomic imprinting: parental influence on the genome. Nat Rev Genet 2(1):21–32
2. Haig D (2004) Genomic imprinting and kinship: how good is the evidence? Annu Rev Genet 38:553–585
3. Steffenburg S, Gillberg CL, Steffenburg U, Lyllerman M (1996) Autism in Angelman syndrome: a population-based study. Pediat Neurol 14(2):131–136
4. Clayton-Smith J, Pembrey ME (1992) Angelman syndrome. J Med Genet 29(6):412–415
5. Temple IK, Gardner RJ, Mackay DJG, Barber JCK, Robinson DO, Shield JPH (2000) Transient neonatal diabetes mellitus: widening our understanding of the aetiopathogenesis of diabetes. Diabetes 49(8):1359–1366
6. Mackay DJG, Callaway JLA, Marks SM, White HE, Acerini CL, Boonen SE, Dayanikli P, Firth HV, Goodship JA, Haemers AP, Hahnemann JMD, Kordonouri O, Masoud AF, Oestergaard E, Storr J, Ellard S, Hattersley AT, Robinson DO, Temple IK (2008) Hypomethylation of multiple imprinted loci in patients with transient neonatal diabetes is associated with mutations in ZFP57. Nat Genet 40(8):949–951
7. Gicquel C, Rossignol S, Cabrol S, Houang M, Steunou V, Barbu V, Danton F, Thibaud N, Le Merrer M, Burglen L, Bertrand AM, Netchine I, Le Bouc Y (2005) Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. Nat Genet 37(9):1003–1007
8. Weksberg R, Shuman C, Beckwith JB (2010) Beckwith-Wiedemann syndrome. Eur J Hum Genet 18(1):8–14
9. Tan TY, Amor DJ (2006) Tumour surveillance in Beckwith-Wiedemann syndrome and hemi-hyperplasia: a critical review of the evidence and suggested guidelines for local practice. J Paediatr Child Health 42(9):486–490
10. Scott RH, Douglas J, Baskcomb L, Huxter N, Barker K, Hanks S, Craft A, Gerrard M, Kohler JA, Levitt GA, Picton S, Pizer B, Ronghe MD, Williams D, Factors Associated with Childhood Tumours (FACT) Collaboration, Cook JA, Pujol P, Maher ER, Birch JM, Stiller CA, Pritchard-Jones K, Rahman N (2008) Constitutional 11p15 abnormalities, including heritable imprinting center mutations, cause nonsyndromic Wilms tumor. Nat Genet 40(11):1329–1334
11. Cooper WN, Luharia A, Evans GA, Raza H, Haire AC, Grundy R, Bowdin SC, Riccio A, Sebastio G, Bliek J, Schofield PN, Reik W, Macdonald F, Maher ER (2005) Molecular subtypes and phenotypic expression of Beckwith-Wiedemann syndrome. Eur J Hum Genet 13(9):1025–1032
12. Wakeling EL, Abu Amero S, Alders M, Bliek J, Forsythe E, Kumar S, Lim DH, MacDonald F, Mackay DJ, Maher ER, Moore GE, Poole RL, Price SM, Tangeraas T, Turner CLS, Van Haelst

MM, Willoughby C, Temple IK, Cobben JM (2010) Epigenotype-phenotype correlations in Silver-Russell syndrome. J Med Genet 47(11):760–768

13. Price SM, Stanhope R, Garrett C, Preece MA, Trembath RC (1999) The spectrum of Silver-Russell syndrome: a clinical and molecular genetic study and new diagnostic criteria. J Med Genet 36(11):837–842
14. Binder G, Seidel AK, Martin DD, Schweizer R, Schwarze CP, Wollmann HA, Eggermann T, Ranke MB (2008) The endocrine phenotype in Silver-Russell syndrome is defined by the underlying epigenetic alteration. J Clin Endocrinol Metab 93(4):1402–1407
15. Yamazawa K, Kagami M, Fukami M, Matsubara K, Ogata T (2008) Monozygotic female twins discordant for Silver-Russell syndrome and hypomethylation of the H19-DMR. J Hum Genet 53(10):950–955
16. Duncan PA, Hall JG, Shapiro LR, Vibert BK (1990) Three-generation dominant transmission of the Silver-Russell syndrome. Am J Med Genet 35(2):245–250
17. Eggermann T, Meyer E, Obermann C, Heil I, Schüler H, Ranke MB, Eggermann K, Wollmann HA (2005) Is maternal duplication of 11p15 associated with Silver-Russell syndrome? J Med Genet 42(5):e26
18. Bliek J, Terhal P, van den Bogaard MJ, Maas S, Hamel B, Salieb-Beugelaar G, Simon M, Letteboer T, van der Smagt J, Kroes H, Mannens M (2006) Hypomethylation of the H19 gene causes not only Silver-Russell syndrome (SRS) but also isolated asymmetry or an SRS-like phenotype. Am J Hum Genet 78(4):604–614
19. Bruce S, Hannula-Jouppi K, Puoskari M, Fransson I, Simola KO, Lipsanen-Nyman M, Kere J (2009) Submicroscopic genomic alterations in Silver-Russell syndrome and Silver-Russell-like patients. J Med Genet 47(12):816–822
20. Horike S, Ferreira JC, Meguro-Horike M, Choufani S, Smith AC, Shuman C, Meschino W, Chitayat D, Zackai E, Scherer SW, Weksberg R (2009) Screening of DNA methylation at the H19 promoter or the distal region of its ICR1 ensures efficient detection of chromosome 11p15 epimutations in Russell-Silver syndrome. Am J Med Genet A 149A(11):2415–2423
21. Schönherr N, Meyer E, Roos A, Schmidt A, Wollmann HA, Eggermann T (2007) The centromeric 11p15 imprinting centre is also involved in Silver-Russell syndrome. J Med Genet 44(1):59–63
22. Kotzot D, Schmitt S, Bernasconi F, Robinson WP, Lurie IW, Ilyina H, Méhes K, Hamel BC, Otten BJ, Hergersberg M et al (1995) Uniparental disomy 7 in Silver-Russell syndrome and primordial growth retardation. Hum Mol Genet 4(4):583–587
23. Abu-Amero S, Monk D, Frost J, Preece M, Stanier P, Moore GE (2008) The genetic aetiology of Silver-Russell syndrome. J Med Genet 45(4):193–199
24. Kotzot D (2008) Maternal uniparental disomy 7 and Silver-Russell syndrome – clinical update and comparison with other subgroups. Eur J Med Genet 51(5):444–451
25. Buiting K, Saitoh S, Gross S, Dittrich B, Schwartz S, Nicholls RD, Horsthemke B (1995) Inherited microdeletions in the Angelman and Prader-Willi syndromes define an imprinting centre on human chromosome 15. Nat Genet 9(4):395–400
26. Ramsden SC, Clayton-Smith J, Birch R, Buiting K (2010) Practice guidelines for the molecular analysis of Prader-Willi and Angelman syndromes. BMC Med Genet 11:70
27. Williams CA, Angelman H, Clayton-Smith J, Driscoll DJ, Hendrickson JE, Knoll JH, Magenis RE, Schinzel A, Wagstaff J, Whidden EM et al (1995) Angelman syndrome: consensus for diagnostic criteria. Angelman Syndrome Foundation. Am J Med Genet 56(2):237–238
28. Fang P, Lev-Lehman E, Tsai TF, Matsuura T, Benton CS, Sutcliffe JS, Christian SL, Kubota T, Halley DJ, Meijers-Heijboer H, Langlois S, Graham JM Jr, Beuten J, Willems PJ, Ledbetter DH, Beaudet AL (1999) The spectrum of mutations in UBE3A causing Angelman syndrome. Hum Mol Genet 8(1):129–135
29. Dindot SV, Antalffy BA, Bhattacharjee MB, Beaudet AL (2008) The Angelman syndrome ubiquitin ligase localizes to the synapse and nucleus, and maternal deficiency results in abnormal dendritic spine morphology. Hum Mol Genet 17(1):111–118
30. Greer PL, Hanayama R, Bloodgood BL, Mardinly AR, Lipton DM, Flavell SW, Kim TK, Griffith EC, Waldon Z, Maehr R, Ploegh HL, Chowdhury S, Worley PF, Steen J, Greenberg

ME (2010) The Angelman Syndrome protein Ube3A regulates synapse development by ubiquitinating arc. Cell 140(5):704–716
31. Williams CA, Driscoll DJ, Dagli AI (2010) Clinical and genetic aspects of Angelman syndrome. Genet Med 12(7):385–395
32. Manipalviratn S, DeCherney A, Segars J (2009) Imprinting disorders and assisted reproductive technology. Fertil Steril 91(2):305–315
33. Gillessen-Kaesbach G, Demuth S, Thiele H, Theile U, Lich C, Horsthemke B (1999) A previously unrecognised phenotype characterised by obesity, muscular hypotonia, and ability to speak in patients with Angelman syndrome caused by an imprinting defect. Eur J Hum Genet 7(6):638–644
34. Lossie AC, Whitney MM, Amidon D, Dong HJ, Chen P, Theriaque D, Hutson A, Nicholls RD, Zori RT, Williams CA, Driscoll DJ (2001) Distinct phenotypes distinguish the molecular classes of Angelman syndrome. J Med Genet 38(12):834–845
35. Cassidy SB, Driscoll DJ (2009) Prader-Willi syndrome. Eur J Hum Genet 17(1):3–13
36. Mogul HR, Lee PD, Whitman BY, Zipf WB, Frey M, Myers S, Cahan M, Pinyerd B, Southren AL (2008) Growth hormone treatment of adults with Prader-Willi syndrome and growth hormone deficiency improves lean body mass, fractional body fat, and serum triiodothyronine without glucose impairment: results from the United States multicenter trial. J Clin Endocrinol Metab 93(4):1238–1245
37. Sahoo T, del Gaudio D, German JR, Shinawi M, Peters SU, Person RE, Garnica A, Cheung SW, Beaudet AL (2008) Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. Nat Genet 40(6):719–721
38. Temple IK, Shield JP (2002) Transient neonatal diabetes, a disorder of imprinting. J Med Genet 39(12):872–875
39. Polak M, Cave H (2007) Neonatal diabetes mellitus: a disease linked to multiple mechanisms. Orphanet J Rare Dis 2:12
40. Ma D, Shield JP, Dean W, Leclerc I, Knauf C, Burcelin RR, Rutter GA, Kelsey G (2004) Impaired glucose homeostasis in transgenic mice expressing the human transient neonatal diabetes mellitus locus, TNDM. J Clin Invest 114((3):339–348
41. Li X, Ito M, Zhou F, Youngson N, Zuo X, Leder P, Ferguson-Smith AC (2008) A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. Dev Cell 15(4):547–557
42. Wang JC, Passage MB, Yen PH, Shapiro LJ, Mohandas TK (1991) Uniparental heterodisomy for chromosome 14 in a phenotypically abnormal familial balanced 13/14 Robertsonian translocation carrier. Am J Hum Genet 48(6):1069–1074
43. Temple IK, Cockwell A, Hassold T, Pettay D, Jacobs P (1991) Maternal uniparental disomy for chromosome 14. J Med Genet 28(8):511–514
44. Ogata TM, Kagami M, Ferguson-Smith AC (2008) Molecular mechanisms regulating phenotypic outcome in paternal and maternal uniparental disomy for chromosome 14. Epigenetics 3(4):181–187
45. Buiting K, Kanber D, Martín-Subero JI, Lieb W, Terhal P, Albrecht B, Purmann S, Gross S, Lich C, Siebert R, Horsthemke B, Gillessen-Kaesbach G (2008) Clinical features of maternal uniparental disomy 14 in patients with an epimutation and a deletion of the imprinted DLK1/GTL2 gene cluster. Hum Mutat 29(9):1141–1146
46. Kotzot D (2004) Maternal uniparental disomy 14 dissection of the phenotype with respect to rare autosomal recessively inherited traits, trisomy mosaicism, and genomic imprinting. Ann Genet 47(3):251–260
47. Mitter D, Buiting K, von Eggeling F, Kuechler A, Liehr T, Mau-Holzmann UA, Prott EC, Wieczorek D, Gillessen-Kaesbach G (2006) Is there a higher incidence of maternal uniparental disomy 14 [upd(14)mat]? Detection of 10 new patients by methylation-specific PCR. Am J Med Genet A 140(19):2039–2049
48. Kagami M, Sekita Y, Nishimura G, Irie M, Kato F, Okada M, Yamamori S, Kishimoto H, Nakayama M, Tanaka Y, Matsuoka K, Takahashi T, Noguchi M, Tanaka Y, Masumoto K, Utsunomiya T, Kouzan H, Komatsu Y, Ohashi H, Kurosawa K, Kosaki K, Ferguson-Smith AC,

Ishino F, Ogata T (2008) Deletions and epimutations affecting the human 14q32.2 imprinted region in individuals with paternal and maternal upd(14)-like phenotypes. Nat Genet 40(2):237–242
49. Temple IK, Shrubb V, Lever M, Bullman H, Mackay DJ (2007) Isolated imprinting mutation of the DLK1/GTL2 locus associated with a clinical presentation of maternal uniparental disomy of chromosome 14. J Med Genet 44(10):637–640
50. Wilson LC, Trembath RC (1994) Albright's hereditary osteodystrophy. J Med Genet 31(10):779–784
51. Wilson LC (2006) Albright's hereditary osteodystrophy. J Pediatr Endocrinol Metab 19(Suppl 2):671–673
52. Dekelbab BH, Aughton DJ, Levine MA (2009) Pseudohypoparathyroidism type 1A and morbid obesity in infancy. Endocr Pract 15(3):249–253
53. Kelsey G (2009) Epigenetics and imprinted genes: insights from the imprinted Gnas locus. Horm Res 71(Suppl 2):22–29
54. Bastepe M (2008) The GNAS locus and pseudohypoparathyroidism. Adv Exp Med Biol 626:27–40
55. Bastepe M, Fröhlich LF, Linglart A, Abu-Zahra HS, Tojo K, Ward LM, Jüppner H (2005) Deletion of the NESP55 differentially methylated region causes loss of maternal GNAS imprints and pseudohypoparathyroidism type Ib. Nat Genet 37(1):25–27
56. Bastepe M, Lane AH, Juppner H (2001) Paternal uniparental isodisomy of chromosome 20q–and the resulting changes in GNAS1 methylation–as a plausible cause of pseudohypoparathyroidism. Am J Hum Genet 68(5):1283–1289
57. Bastepe M, Pincus JE, Sugimoto T, Tojo K, Kanatani M, Azuma Y, Kruse K, Rosenbloom AL, Koshiyama H, Jüppner H (2001) Positional dissociation between the genetic mutation responsible for pseudohypoparathyroidism type Ib and the associated methylation defect at exon A/B: evidence for a long-range regulatory element within the imprinted GNAS1 locus. Hum Mol Genet 10(12):1231–1241

Chapter 10
The Utility of Twins for Epigenetic Analysis

Richard Saffery, Ruth Morley, and Debra L. Foley

Contents

10.1 Why Study Epigenetics in Twins 163
10.2 The Value of Twin Studies in the Understanding Regulation of Epigenetic Profile 164
10.2.1 Twin Studies Highlight the Role of Genetic (Heritable) Factors in the Regulation of Epigenetic Profile 164
10.2.2 Twin Studies Reveal Cumulative Environmental Contributors to 'Epigenetic Drift' Over Time 165
10.3 Epigenetics and the Discordant Twin Model 167
10.3.1 The Assumption of Genetic Identity 167
10.3.2 Inferring Causation: Epigenetic vs. Genetic Analysis 167
10.3.3 Localised Epigenetic Variation in Phenotypically Discordant Twins 168
10.3.4 Skewed X-Chromosome Inactivation in Phenotypically Discordant Twins 168
10.3.5 Genome-Scale Epigenetic Investigations in Twins 172
10.4 Caveats and Cautions 179
References 180

R. Saffery (✉) • R. Morley
Cancer, Disease and Developmental Epigenetics, Murdoch Childrens Research Institute (MCRI), Australia and Department of Paediatrics University of Melbourne, Parkville, VIC 3052, Australia
e-mail: richard.saffery@mcri.edu.au; ruth.morley@mcri.edu.au

D.L. Foley
Applied Genetics and Biostatistics, Orygen Youth Health Research Centre, Parkville, VIC, Australia

Centre for Youth Mental Health, The University of Melbourne, Parkville, VIC, Australia
e-mail: dfoley@unimelb.edu.au

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_10,

© Springer Science+Business Media B.V. 2012

Abstract Twin studies have played an important role in our understanding of individual variation for over a century. The strength of these studies lies in the capacity to perfectly control for inter-individual genetic variation through the use of monozygotic (MZ) twin pairs. Despite their genetic identity, MZ twins often show phenotypic variability, presumably in response to different environmental exposures. Given that epigenetic mechanisms are widely believed to be the mediators of the influence of environmental factors on the underlying genome, it is not surprising that the use of twins in epigenetic research is increasingly recognised as an important approach to help unravel the complexities associated with human development and disease. In addition, the strategic use of twins in epigenetic studies has revealed the importance of genetic factors and both *in utero* and postnatal environments to the establishment and maintenance of the human epigenome. Finally, twin studies are generating compelling evidence linking epigenetic disruption to the disease-associated risk in humans.

Abbreviations

MZ	monozygotic
DZ	dizygotic
DNA	deoxyribonucleic acid
EEA	equal environments assumption
2D	two dimensional
ANOVA	analysis of variance
ICC	intraclass correlation
RRBS	reduced representation bisulphite sequencing
SLE	systemic lupus erythematosus
RA	rheumatoid arthritis
DM	dermatomyositis
RNA	ribonucleic acid
5MeC	5- methylcytosine
CBMC	cord blood mononuclear cells
HUVEC	human umbilical vein endothelial cells
AIMS	amplification of inter-methylated sites
DC	dichorionic (two placentas)
MC	monochorionic
CGI	CpG island region (enriched for CpG dinucleotides)
3′UTR	non protein coding 3′ untranslated region of RNA transcript
DMR	differentially methylated regions
DRD2	dopaminergic receptor D2
COMT	catechol-O-methyltransferase
MSRDA	methylation sensitive representational difference analysis
PPIEL	Peptidylprolyl isomerise E-like
SMS	Spermine synthase

10.1 Why Study Epigenetics in Twins

In today's post genomic era there has been a shift in our understanding of complex disease to include a more functionally relevant genomics approach, where the dynamic patterns of gene expression are analysed, not in isolation, but jointly as parts of functional clusters or networks. Similarly, more and more studies are integrating both genetic and environmental data into analyses with the aim of providing a more 'holistic' understanding of how diseases arise. The complexities are manifold and require the development of novel methodological approaches. The strategic use of twin samples, especially monozygotic (MZ) twin pairs who share the same genetic material, offers a unique opportunity to circumvent some of the complexities and reduce the required sample size for studies examining the aetiology of complex disease.

Twin studies have played an important role in our understanding of individual variation for over a century. The similarity of MZ and dizygotic (DZ) twins allows an estimation of heritability (the proportion of the total variance in a trait that is attributable to genetic factors). MZ twins are generally considered to be genetically identical, although this is not always true, (see Sect. 9.1 below), whereas DZ twins share on average 50% of their genetic (DNA sequence) variation (0–100% theoretical range). MZ twins exhibit a wide range of concordance rates for complex disease. Whereas MZ twin concordance approaching 100% indicates coinheritance of genetic variants that are likely dominant and highly penetrant, many diseases or traits show concordance rates far lower than this indicating a role of environmental and/or stochastic epigenetic factors in modulating phenotypes associated with underlying genotypes. Given that epigenetic mechanisms are widely believed to be the mediators of the influence of environmental factors on the underlying genome, it is not surprising that the use of twins in epigenetic research is becoming increasingly popular.

Previous studies examining gene expression in MZ twins have identified many hundreds of genes whose within-pair expression discordance [the absolute value of the ratio of (log) gene expression of twin 1: twin 2] is greater than an arbitrarily defined threshold (set above the level of variation attributable to experimental noise associated with a particular array platform) [1, 2]. This supports divergence in gene expression between MZ twins in response to environmental and stochastic factors as the driver of subsequent phenotypic divergence. In accordance with expectations, array based studies have also demonstrated a higher degree of similarity between specific gene expression patterns in MZ than in DZ pairs, highlighting the contribution of genetic factors to overall expression profile [1, 3]. This relationship is even apparent in multiple tissues taken from newborn twin pairs. Nevertheless, MZ pairs also show clear discordance in gene expression profiles at birth, highlighting the importance of the *in utero* environment in determining overall gene expression (and most likely epigenetic) profile [4].

Major molecular mechanisms underpinning gene expression and associated phenotypic discordance in MZ twins are likely epigenetic in nature. Epigenetic status

of a specific genomic region is determined by the sum of genetic, cumulative environmental and stochastic (probabilistic or random) factors. Measurement of epigenetic differences in phenotypically discordant MZ pairs therefore provides a unique opportunity to identify genes sensitive to the environment that are associated with complex disease.

10.2 The Value of Twin Studies in the Understanding Regulation of Epigenetic Profile

Elegant studies in mice have demonstrated that the phenotype in genetically identical animals reared in identical environmental conditions is normally distributed, highlighting the role of probabilistic (stochastic) factors in phenotype development [5]. DNA methylation levels at specific genomic loci in MZ twins are also normally distributed. However, at present the relative contributions of environmental, genetic, and stochastic factors to the establishment and maintenance of epigenetic profile, and its relative stability over time, remain poorly understood.

10.2.1 Twin Studies Highlight the Role of Genetic (Heritable) Factors in the Regulation of Epigenetic Profile

MZ twins share maternal, obstetric and genetic factors, but differences in epigenetic profile will accumulate in response to differences in environmental exposures and stochastic factors, accumulated both *in utero* and postnatally. In contrast DZ pairs share maternal and obstetric factors but the likelihood of sharing specific genetic variations is the same as that of non-twin siblings. Given the large number of genes and/or proteins implicated in the establishment and maintenance of the epigenetic profile, and the demonstrated link between genetic variation and epigenetic profile [6, 7], it reasonable to speculate that DZ twins as a group will be more epigenetically divergent than MZ twins in accordance with gene expression data. A powerful method for estimating the relative contribution of heritable and environmental/stochastic influences on variation in any quantitative trait such as DNA methylation is to compare the degree of within pair discordance in MZ pairs to that in DZ twin pairs. Several studies have now carried out this type of heritability analysis, using twins of varying ages and phenotypes. The underlying assumption of such studies is that genetic factors are the primary driver of cumulative MZ vs DZ disparity. Although highly likely, it is equally feasible that DZ twins as a group exhibit more epigenetic differences than MZ twins because the former originate from different zygotes, potentially already epigenetically divergent. In contrast, MZ twins derive from a single zygotic epigenome.

The major strength of the classical twin study is the lack of confounding by age effects between individuals within a twin pair, and the ability to distinguish between genetic and common environmental effects. However, this is predicated on the equal environments assumption (EEA) that no interaction exists between genes and environment [i.e. no twin-type (DZ versus MZ) differences in response to salient environmental effects]. If this is not the case, then the classical model lacks utility. Both MZ and DZ twins are required to estimate the relative contributions of genetic versus common environmental contributors. Age differences between groups of MZ and DZ twin pairs may be a potential confounding factor specific to epigenetic analyses [8, 9]. As such, care must be taken to match as closely as possible in this regard.

Heritability is an estimate of the proportion of the total phenotypic (outcome) variance in a population that is attributable to additive genetic effects. In twin studies heritability is estimated by comparing the degree of phenotypic similarity between groups of MZ and DZ twins, either as the concordance rate or intra-class correlation. In the context of twins, heritability (h^2) is determined as twice the difference between MZ and DZ concordance rates (correlation; $h^2 = 2(r_{MZ} - r_{DZ})$, where r is the concordance or intraclass correlation between each type of twin [10]. Heritability estimates are population specific unless the environment is constant. Additionally, the contribution of stochastic factors to epigenetic divergence is equally assumed to be constant across populations in these analyses.

10.2.2 Twin Studies Reveal Cumulative Environmental Contributors to 'Epigenetic Drift' Over Time

The cumulative effects of environmental and stochastic variation on changing epigenetic profile were first illustrated by a widely cited study that examined both genome wide and locus specific DNA methylation variation in a small number of young and middle-aged MZ twins [9]. Whereas 3 year-old MZ twins showed relatively few epigenetic differences within pairs, those aged 50 years showed considerable variability within pairs, and this was greater if the twins had divergent lifestyles. A multi-level statistical analysis was performed that generated a single descriptive value for each type of epigenetic measure. For DNA methylation this included (i) the numbers of bands present on a gel (or dots on a 2D gel) between individuals A and B within a pair, and (ii) the absolute percentage of 5-methylcytosine present within an individual. Gene expression differences were measured as the number of genes showing differential expression (adjusted p-values following ANOVA) between siblings. Histone acetylation levels were measured as relative peak heights following high performance liquid chromatography.

The similarity of each descriptive value within pairs was then estimated as the Euclidean squared distance by subtracting their respective values. In order to examine the relation between phenotypic/environmental data and epigenetic variables, categorical principle component analysis was firstly applied to the original questionnaire variables to reduce them to two uncorrelated components termed

'aging' (encompassing age, weight and height) and 'health' (all variables of disease and pharmacological treatments). Mixed models were used to estimate the contribution of each random-effect to the variance of the dependent variable and the variance component procedure was assessed by ANOVA. The relation of single scores generated for questionnaire-derived ageing and health variables to epigenetic data was evaluated by Pearson's test (note that the specific type of test was not described). Using this approach the general conclusion was that epigenetic profile is in constant "drift", although very few young twins were studied and statistical precision was low [9]. In addition, this was a cross sectional study and did not evaluate methylation in the same individuals over time. Finally, the specific epigenetic analyses employed were all low resolution in that they did not examine the distribution of epigenetic differences within the genome, but rather provided a global 'snapshot' of different classes of epigenetic disparity within twin pairs.

The first study to examine epigenetic drift longitudinally in twins involved analysis of buccal cell DNA methylation in 3 genes in 46 MZ and 45 DZ at 5 and 10 years of age [11]. Longitudinal change within individuals was calculated by assessing the correlation in methylation at each age. To assess the relative contributions of heritable and environmental/stochastic components to methylation levels, correlations within MZ pairs were compared to correlations within DZ pairs. Finally, to assess the relative contributions of heritable and environmental/stochastic components to changes in DNA methylation over time, intra-individual change scores were calculated and correlations within MZ twin pairs were compared to correlations within DZ pairs. This study was the first to reveal the extent of epigenetic discordance in MZ twins in early life and highlighted the ongoing instability of methylation levels over time. Importantly different genomic regions were found to show varying levels of epigenetic divergence over time.

An examination of DNA methylation levels at several sites in multiple tissues from newborn twins confirmed that epigenetic drift between genetically identical individuals (MZ twins) begins *in utero* and in a tissue-specific manner. In this study, intraclass correlation coefficients (ICC) were higher in MZ than in DZ twins [12] (Fig. 10.1) supporting previous findings of a role for genetic/heritable factors in the establishment of epigenetic profile.

Epigenetic drift over time has been observed for singletons and it is also possible that maintenance of the epigenetic mark of DNA methylation is under genetic control [13]. However, longitudinal changes in epigenetic marks are most likely to occur in a subset of genes [14], possibly in those with intermediate methylation levels, rather than those that are entirely unmethylated or in a hypermethylated state.

Limited studies are now beginning to utilise the twin model to estimate the relative contribution of genetic and environmental factors in mediating the response of the developing epigenome to specific environmental exposures. An examination of the methylation levels at 3 independent imprinting-associated DMRS (KvDMR1, PEG1 and H19/IGF2) in 59 pairs of twins (including 29 pairs conceived through IVF) did not reveal significant differences at three DMRs between IVF-conceived and naturally conceived twins. However, both sample size and the genomic regions examined were too small for any general conclusions to be made and further analysis in larger groups of twins is warranted [15].

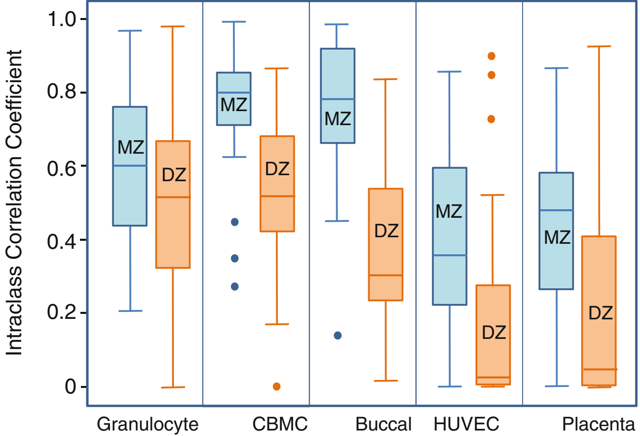


Fig. 10.1 Distribution of intraclass correlations reveals a higher median correlation coefficient for DNA methylation at the IGF2/H19 locus in MZ than DZ twin pairs in five tissues from newborns. The ICC measures the proportion of total variance attributable to within pair variation within MZ and DZ groups of twins. ICC analysis. CBMC- cord blood mononuclear cells; HUVEC- human umbilical vein endothelial cells. (Figure adapted from [12])

10.3 Epigenetics and the Discordant Twin Model

10.3.1 The Assumption of Genetic Identity

Many studies have reported phenotypic differences between MZ co-twins (reviewed in [16]). Discordant MZ pairs have proved very valuable in revealing the contribution of non-genetic variation to disease penetrance, aetiology, or effect of therapies. The underlying assumption in all such studies is that MZ twins are genetically identical. This has been directly disproved in several instances, where either specific point mutations [17–19], uniparental disomy [20], triplet repeat expansion [21, 22], chromosomal mosaicism [23] or heteroplasmy for mitochondrial-encoded mutations [24, 25] or chromosomal aneuploidies (reviewed in [16]) have all been linked to specific phenotypic differences in monozygotic twins. Importantly, differences in both copy number [26] and telomere length [27, 28] have been also described in phenotypically discordant and concordant MZ twin pairs. Nevertheless, such reports are relatively rare and it is likely that non-genetic (epigenetic) variation within MZ twin pairs underpins the majority of observed phenotypic discordance.

10.3.2 Inferring Causation: Epigenetic vs. Genetic Analysis

Since every cell in an individual derives from a single zygote with one genome, genetic studies generally require only a single DNA sample per individual, taken

from any tissue at any age. Cause and effect can be reliably predicted, or excluded, as genomic variation is not considered to be variable over the life course (except in relation to 'rare' somatic mutation). Such rules may not apply to most epigenetic studies where repeated samples from more than one tissue may be desirable. Even if biospecimens are collected very early in life (prior to phenotypic manifestation) unravelling cause and effect is problematic. Despite this and the many other caveats associated with inferring causation in any epigenetic association study (covered in other chapters of this book), several investigators have attempted to link specific epigenetic changes to disease phenotypes. Surprisingly few studies have attempted to utilise the discordant MZ twin pair model, despite the demonstrated utility of this approach in identifying specific risk factors of disease,

10.3.3 Localised Epigenetic Variation in Phenotypically Discordant Twins

Several studies have attempted to identify epigenetic 'causes' of complex human disorders through a low resolution, candidate gene hypothesis examination of epigenetic variation (primarily DNA methylation) at specific gene regions in phenotypically discordant MZ twins (summarised in Table 10.1). The contribution of epigenetic variation to the heritability of complex phenotypes is complex and likely regulated at many levels (Fig. 10.2). Generally speaking, effect sizes have been minimal and there have been few examples of independent replication of identified associations. As such, the biological relevance of such findings remains questionable.

10.3.4 Skewed X-Chromosome Inactivation in Phenotypically Discordant Twins

One of the most widely studied epigenetic phenomena in mammals involves the inactivation of the majority of genes on one X-chromosome in females. This 'dosage compensation' equalises the expression of most X-chromosome genes in males (XY) and females (XX) and usually occurs in a random manner, with roughly 50% of cells in a female showing inactivation of the maternally-derived X, whereas the other 50% show inactivation of the paternally-derived X chromosome. However, in some individuals a skewed pattern of X-inactivation is apparent, with one parental X, over represented in a particular tissue or cell type. In such instances, an otherwise recessive mutation can have profound adverse phenotypic effects. Thus, skewed X-inactivation, or more precisely the chance inactivation of a single functional allele in a physiologically-relevant tissue, has been associated with discordance for several disorders including haemophilia [29], Fragile-X syndrome [22, 30] and Duchenne Muscular Dystrophy [31].

Table 10.1 Non-disease-based epigenetic analyses in twins

Focus	Number of twin pairs	Biological samples	Target/methodology	Analysis methodology	Major finding	Main strengths/ Weakness	Reference
Identification of factors contributing to epigenetic variation	80 twins (age range 3–74 years)	Buccal	AIMS Global 5MeC levels	(see body text)	Increasing discordance in MZ twin epigenetic profile with age. First demonstration of 'epigenetic drift'. Some evidence of increasing drift with increasing cumulative environmental differences	S: multiple marks profiled. Global analysis W: small numbers. No longitudinal analysis	[9]
		Blood lymphocytes Skeletal muscle (n = 14) Abdominal muscle (n = 4)	Histone modifications				
Cataloguing range of interindividual methylation at differentially methylated regions	12 MZ and 14 DZ pairs (age range 27–73 years)	Whole blood	DNA methylation using bisuphite pyrosequencing at two paternal and four maternal DMRs. Gene promoters of *OCT4* and *APC* genes	Box plots to illustrate distribution of methylation. Statistical inference for eight studied regions was done using multivariate analysis of variance (MANOVA). Students *t*-test used to compare means of a particular gene between two groups. Linear regression was calculated to infer the dependence of methylation % on age. Goodness of fit (R^2) was used to quantify this relation. P value calculated by ANOVA. Multivariate discrimination using Fisher's linear discriminant analysis followed by non-parametric Wilcoxon test was used to compare the pairwise methylation differences of MZ vs DZ pairs.	For all DMRs, the median pairwise methylation difference between MZ was small than for DZ at the single gene level. DMR methylation is relatively stable between the ages of 27–72 years at the studied regions. Data are consistent with most methylation being established after fertilization in association with development of methylation mosaicism.	S: sensitive technique W: small numbers of twins and genes tested	[39]

(continued)

Table 10.1 (continued)

Focus	Number of twin pairs	Biological samples	Target/methodology	Analysis methodology	Major finding	Main strengths/ Weakness	Reference
Identification of factors regulating DNA methylation at *IGF2/H19* imprinted region	196 adolescent twins; 176 middle aged twins	Whole blood	Gene specific, bisulphite based Mass Array Epityping of previously described differentially methylated regions (DMRs)	Correlations between methylation at specific sites measured with GOLD [40]. Principle components analysis (PCA) SPSS 11. Heritability calculated using estimates of additive genetic V(A), common environmental V(C), unique environmental + stochastic V(E) obtained from variance components implemented in Mx 1.61 (www.vcu.edu/mx). Association of methylation with body size (BS) performed in variance components setting with BS as outcome measure, component scores for DMRs as variables, and age and sex as confounders.	PCA indicated that one major factor explaining between 53% and 73% of variance at either DMR. Heritability of methylation at individual CpG sites varied from 20 to 74% for H19 DMR and 57–97% for IGF2 DMR. No evidence for an age effect was found between the adolescent vs middle aged groups. Preliminary evidence for an association between H19 DMR methylation and body size ($p < 0.05$).	S: large numbers. Sensitive technique W: focus on single genomic region	[41]
Role of *in utero* environment in specifying methylation at *IGF2/H19* imprinted region in newborns	56 MZ and 35 same sex DZ collected at birth	Buccal CBMC HUVEC Placenta Granulocytes	Gene specific, bisulphite based Mass Array Epityping of previously described differentially methylated regions (DMRs)	Association between methylation at different CpG units, and between methylation and BW or methylation and GA, was measured using Spearman's rank correlation coefficient. The contribution of genetic factors outside the IGF2/H19 DMRs to methylation levels was investigated by fitting separatelinear mixed models for MZ and DZ twin pairs at each CpG unit. Within-twin pair correlation was modelled by including a random intercept specific to each twin pair; the models had no other covariates and were fitted using restricted maximum likelihood. The proportion of total variance attributable to 'between-twin pair' variation was calculated using intra class correlation.	Higher ICC values for MZ than DZ twin pairs. First evidence for epigenetic discordance in genetically identical humans at birth (ie. cumulative effect of environmental/stochastic factors). Additional evidence for a role of genetic factors in regulating the neonatal epigenetic profile. Evidence for tissue-specific effects.	S: multiple tissues from single individuals at birth. Relatively large numbers W: single genomic region	[12]

Annotation of degree of epigenetic metastability	114 MZ and 80 DZ teenage pairs	MZ: White blood cells (n = 19 pairs), buccal epithelia (N = 20 pairs), rectum (N = 18 pairs)	Enrichment of unmethylated DNA fraction and DNA methylation profiling at >12,000 CpG island fragments	Intraclass correlation for each unique genomic region (~6000 unique comparisons per pair; ICC range from +1 > −1 denotes very high > very low correlation). Wilcoxon rank-sum test used to compare variability of CGI vs non-CGI regions. P values corrected for multiple testing by Bonferroni method.	ICC profiles were more similar within different tissues from single individuals than between unrelated individuals. CGI regions less variable than non-CGI regions. ICC for Buccal cell DZ ICC values significantly lower than MZ ICC values; ICC for MZMC significantly lower MZDC pairs. But MZ v DZ ICC differences were tissue specific.	S: large numbers. Multiple technical replicates W: poor platform for methylation analysis	[34]
Effect of assisted reproductive technology on the stability of DNA methylation	29 IVF and 30 non-IVF twin pairs (no zygosity)	Umbilical cord tissue	Methyl-specific quantitative PCR	Fisher's exact test comparing mean methylation at each region between groups	No significant difference between IVF and non-IVF groups however some evidence for increasing variation in the IVF group.	W: limited genomic coverage and small numbers of each twin type. No zygosity information	[15]

5MeC 5- methylcytosine, CBMC cord blood mononuclear cells, *HUVEC* human umbilical vein endothelial cells, *AIMS* amplification of inter-methylated sites, *DC* dichorionic (two placentas), *MC* monochorionic, *CGI* CpG island region (enriched for CpG dinucleotides)

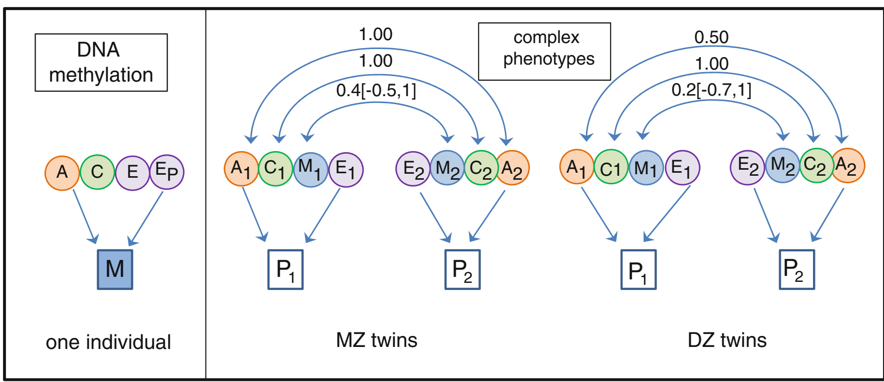


Fig. 10.2 The relation between epigenetic and phenotypic heritability. Proposed contribution of latent variables to the methylation status of an individual at a genomic region (*M*) and to their phenotype (*P*). The left pane represents shows the variables contribution to DNA methylation status at one genomic region in one individual; effects will be specific to age, sex, population (genetic factors), and the tissue sampled, and will also include stochastic factors. Methylation latent factors include additive genetic factors (*A*), common environmental factors (*C*), unique environment (*E*), and heritable and stable epigenetic factors that are not DNA sequence dependent (Ep). The right pane represents the path model in twins, depicting the contribution of DNA Methylation and other factors to the phenotype (*P*) in twin *i* with correlation estimates in MZ (*Left*) and DZ (*Right*) twins of latent variables including additive genetic effects (*Ai*), common environment (*Ci*), DNA methylation (*Mi*), and unique environment (*Ei*). Correlation estimates were obtained from previous genetic [38] and epigenetic studies [34] in twins. In siblings, the correlation in M will be lower than that observed in DZ twins due to age differences and associated increased level of cumulative stochastic change. (Figure adapted from [10])

10.3.5 Genome-Scale Epigenetic Investigations in Twins

Given the relative ease of simultaneously mapping DNA methylation profile across large regions of the genome relative to higher order epigenetic marks, it is not surprising that the limited twin studies carried out in this space have focused solely on DNA methylation (Table 10.2).

10.3.5.1 Medium Resolution (1–2 Sites per Gene in 1–10% of Genes in the Genome)

Recent studies have begun to reveal the true extent of genomic DNA methylation changes in MZ twins that could potentially explain phenotypic discordance. Perhaps the most successful combined application of the twin discordance model and epigenetic analysis has been investigation of the association of DNA methylation changes with immune related disorders, systemic lupus erythematosus (SLE), rheumatoid arthritis (RA) and dermatomyositis (DM) [32]. A tiered approach to analysis involved

Table 10.2 Examples of epigenetic analysis in discordant twin pairs

Focus	Number of twin pairs	Biological samples	Epigenetic target/ Methodology	Analysis methodology	Major finding	Ref.
Candidate gene approaches						
Schizophrenia (SCZ)	1 discordant + 1 concordant affected	Lymphocytes	DNA methylation of *DRD2* gene by Bisulphite Sequencing	Purely observational No statistical analysis	Discordant twin with schizophrenia shows more similar *DRD2* methylation profile to affected concordant twins than to unaffected co-twin	[42]
Birth weight	12 highly discordant MZ pairs	Buccal cell DNA at age 5	DNA methylation at 2 CpG sites of *COMT* gene by bisulphite pyrosequencing	Not stated	Highly variable methylation concordance rates between pairs. Average within pair methylation discordance was 10.3% at site 1 and 16.1% at site 2. Strong correlation in differences between both sites ($r = 0.87$; $p < 0.001$). No correlation between birth weight and methylation level	[43]
Primary biliary cirrhosis (PBC)	4 discordant MZ, 1 concordant affected	Peripheral blood	DNA methylation by bisulphite sequencing	Not stated	Decreased expression of *CLIC2* and *PIN4* in ¾ PBC affected individuals in discordant pairs. No evidence for DNA methylation changes driving expression.	[44]
Beckwith Wiedemann Syndrome (BWS)	10 discordant MZ 5 control MZ	Skin fibroblasts and/or peripheral blood	DNA methylation by Southern Blotting	Purely observational No statistical analysis	BWS affected individuals show loss of imprinting at *KVDMR1* with biallelic expression of *KCNQ1OT1*	[45]

(continued)

Table 10.2 (continued)

Focus	Number of twin pairs	Biological samples	Epigenetic target/ Methodology	Analysis methodology	Major finding	Ref.
Silver-Russell Syndrome (SRS)	1 discordant MZ pair	Peripheral blood leukocytes	Combined bisulfite restriction analysis (COBRA) and bisulfite sequencing for two H19 DMRs	Purely observational No statistical analysis	Loss of *H19*-DMR in roughly half of cells in the affected twin only with associated decreased expression of *IGF2*	[46]
Caudal Duplication Anomaly (CDA)	1 discordant MZ pair 9 unaffected MZ pairs	Peripheral blood mononuclear cells	DNA methylation of *AXIN-1* by Bisulphite Sequencing	Binomial generalised linear mixed model (lme4 package)	Higher methylation at *AXIN-1* promoter in affected vs unaffected co-twin ($p<0.0001$). Higher methylation in both twins than in controls ($p=0.02$). ICC correlation of 0.76 for all 10 MZ pairs.	[47]
Alzheimer's disease (AD)	1 discordant MZ	Post mortem temporal neocortex	Various epigenetic markers by immunohistochemistry	Two tailed *t*-test of staining intensity	Significantly reduced levels of DNA methylation in temporal neocortex neuronal nuclei in AD	[48]

Body Mass Index (BMI)	16 discordant MZ pairs	Saliva	DNA methylation at 9 DMR regions implicated in growth, by bisulphite based, amplification and primer extension and HPLC. Methylation index calculated as h(C)/h(C)+h(T) where h=peak height on HPLC	Paired *t*-test and Wilcoxon's signed rank test to test for association between phenotype and MI for each discordant pair. Pearson's and Spearman's correlation coefficients calculated to test for associations of intrapair BMI and MI differences	Only small intrapair differences in methylation observed. No significant correlations between intrapair BMI differences and intrapair methylation levels.	[49]
Genome scale (hypothesis free) approaches						
1. Systemic Lupus Erythramatousis(SLE) **2**. Rheumatoid arthritis (RA) **3**. Dermatomyositis (DM)	5 discordant for each disorder −17 SLE sib pairs including 1MZ, 4DZ for validation only	Whole blood: white cell fraction	DNA methylation with Illumina Golden Gate Bead Arrays (>1000 measurements) Validation by Bisulphite Sequencing & pyrosequencing Global 5MeC measurement	Student's *t*-test with FDR correction for multiple testing	No methylation differences associated with DM and RA Consistent methylation changes in SLE twins relative to unaffected co-twins in immune functioning genes Global decrease in 5-methylcytosine in SLE and hypo-methylation of 28 S and 18 S *rDNA* genes	[32]

(continued)

Table 10.2 (continued)

Focus	Number of twin pairs	Biological samples	Epigenetic target/ Methodology	Analysis methodology	Major finding	Ref.
Multiple Sclerosis (MS)	3 discordant MZ pairs,	CD4+ T cells	Next Generation sequencing of genomic DNA, mRNA (50–68 million reads) and RRBS (50–90 million reads covering >two million CpG sites)	RBBS sequence aligned using GSNAP [50]. No statistical testing used to determine significance (or otherwise) of observed within pair methylation differences	Between 2 and 176 significant methylation differences (of 2 million CpGs tested) between MZ co-twins (> 800 methylation differences in unrelated individuals). No evidence for genetic, epigenetic or transcriptome differences underlying MS discordance.	[35]
Bipolar disorder (BPD)	1 discordant MZ pair, 16 unrelated singletons with BPD. Independent set of 14 unrelated BPD singletons	Lymphoblastoid cell lines	MS-RDA with validation by bisulphite DNA sequencing and pyrosequencing	Mann–Whitney *U*-test for analysis of expression and methylation levels. One way ANOVA with cofactors of age and sex for effect of diagnosis controlled for confounders.	10 genomic regions showing differential methylation between co-twins by MS-RDA. 4 of these confirmed by bisulphite sequencing of co-twin DNA. 2 of these (*PPIEL*, *SMS*) confirmed as significant by pyrosequencing in case–control study design (25 BPD cases and 18 controls) with $p<0.05$. Methylation status of *PPIEL* inversely correlated with gene expression levels	[51]

Risk taking behaviour	1 discordant plus 9 MZ control pairs	Peripheral blood	DNA methylation profiling at >12,000 CpG island fragments	Multiple technical replicates. FDR correction for multiple testing. 1.15 fold ($log_2(1.15)=0.2$) used as experimental /technical variance threshold.	No technical comparisons survived FDR correction. Differential methylation of *DLX1*[a] gene (3′UTR) identified specifically in discordant MZ pair ($p<0.0004$)	[37]

[a]Transcription factor involved in the formation of GABAergic interneurons [52].
3′UTR non protein coding 3′ untranslated region of RNA transcript
DMR differentially methylated regions
ICC intraclass correlation
DRD2 dopaminergic receptor D2
COMT catechol-O-methyltransferase
RRBS reduced representation bisulphite sequencing
MSRDA methylation sensitive representational difference analysis
GSNAP [50]
PPIEL Peptidylprolyl isomerise E-like
SMS Spermine synthase

genome scale DNA methylation profiling of over 1500 CpG sites in five pairs of discordant MZ pairs, in combination with gene specific validation in further discordant twins and matched control subjects, and global measurement of 5-Methylcytosine levels. A *t*-test was used to identify array probes differentially methylated between affected individuals and their respective healthy twins. The resulting p-values were corrected for multiple testing using False Discovery Rate methodology [33]. As with most current methylation studies of this nature, an arbitrary mean methylation change between affected and non-affected groups was used as a cut-off (in this instance 10%) to generate a list of candidate genes for subsequent validation.

This approach revealed 49 genes, enriched for immune functioning, showing consistent changes in DNA methylation specifically in SLE. No consistent methylation changes were found in RA or DM. Subsequent validation in matched controls and a further set of discordant twin pairs (again with corrections for multiple testing) confirmed the array result. Interestingly, an examination of global 5MeC levels identified a generalised loss of DNA methylation in the SLE affected individuals, and this was found to be (at least in part) due to a specific decrease in methylation at repetitive 28S and 18S ribosomal RNA genes.

10.3.5.2 High Resolution (1–20 Sites in 10–50% of All Genes)

The most compelling analysis of factors regulating overall methylation profile was obtained through the analysis of three collections of DZ and MZ pairs [34]. This consisted of white blood cells (WBCs) from 19 MZ twin pairs who were dichorionic (each twin had their own placenta) and 20 DZ twin pairs matched for age, sex and blood count, plus buccal epithelial cells from 10 monochorionic MZ and 20 DZ age- and sex-matched. Additionally WBCs and buccal cells were obtained from each of 10 dichorionic MZ and 10 DZ pairs from another independent twin cohort and gut biopsies were collected from 18 MZ pairs from a third twin population. Intra-class correlations for MZ and DZ pairs were calculated for pair wise comparisons of over 6,000 individual data points corresponding to methylation levels at individual genomic regions. These compellingly demonstrated a higher degree of similarity between the MZ twins as a group, relative to the DZ group ($p = 1.2 \times 10^{-294}$). However, in studies carried out in parallel to the human analyses, wide ranging epigenetic differences were also found in nearly identical mouse strains, leading the authors to speculate that a differential zygotic epigenetic profile in DZ versus MZ twins was the primary driver of the ICC differences – not genetic variability between the different individuals [34]. However, this could not be directly tested because of the inability to directly measure epigenetic profile on single celled zygotes.

10.3.5.3 Ultra-High Resolution (Every Site and Every Gene)

The future of both genomic and epigenomic analysis of twins lies in the use of high throughput nucleotide sequencing. Baranzini et al. [35] searched for genetic,

expression and DNA methylation differences in purified CD4+ T cells in a small number of MZ twin pairs discordant for multiple sclerosis (MS). Between 50 and 68 million different sequencing reads of messenger RNAs were counted for each pair. Broadly speaking, this represents the expression 'output' of the original sample. The more reads present for a specific gene as a proportion of the total read number, the more highly expressed it is within the starting cell population. Surprisingly, a diagnosis of MS accounted for only 9.4% of the total variance; 57.3% was attributable to between pair differences, and 26.3% due to day-to-day variation within single individuals. This data alone highlights the highly dynamic nature of gene expression that raises considerable challenges when trying to identify disease-specific gene expression changes. Not surprisingly, no robust gene expression differences could be ascribed to the MS phenotype alone. Despite some evidence for allelic imbalance (where one of an individual's two copies of a gene is expressed more highly than the other) within discordant twin pairs, no such changes were common to all three MS discordant twin pairs examined. An examination of 50–90 million, high-quality reads of reduced representation bisulphite sequencing (RRBS) from CD4+ T cells of the 3 three discordant twins similarly showed little evidence of DNA methylation changes that could be specifically associated with the MS phenotype. Whereas several hundred distinct differences (at specific CpG sites) were apparent between males and females, or between primary and unrelated cancerous tissues, the magnitude of epigenetic differences between MZ twins discordant for MS was at least an order of magnitude lower than non-related individual differences, and ~3 orders of magnitude lower than between primary and malignant tissue. Unfortunately there was no independent validation of observed within pair differences using an alternative methodology, so the robustness of the few observed DNA methylation changes within MZ pairs remains unclear [35].

10.4 Caveats and Cautions

1. There may be epigenetic effects specific to twins that limit their utility as a model to reveal contributors to disease causation more broadly applicable in the wider singleton population. For example, the processes leading to MZ twinning remain to be determined but have traditionally been considered to result from random processes. However, more recent data have suggested that epigenetic factors themselves may be implicated in the embryo splitting that leads to MZ twins [36]. If specific to MZ twins, this would have the effect of confounding any MZ vs DZ comparisons aimed at identifying heritable factors. This could result in differential epigenomes of two newly created MZ embryos independently of later onset of phenotypic discordance.
2. Few studies have attempted to estimate statistical power required to falsify epigenetic hypotheses. The link between the number of technical replicates (independent measures from the same biospecimen, aimed at identifying the contribution of technical 'noise') and the power to detect biologically meaningful

DNA methylation differences in twin pairs has been elegantly examined by Kaminsky et al. [37]. As predicted a larger number of array technical replicates results in a smaller overall level of technical variance at individual array data points. Thus, as few as two technical replicates provided >80% power to detect a 1.2 fold change at ~90% of array probes in only nine twin pairs. This is increased from ~60% of probes with only a single technical replicate. Further power calculations predicted that with two technical replicates and 21, 14 and 6 twin pairs, there was an 80% chance of detecting true DNA methylation difference of 1.15, 1.2 and 1.6 fold, at >95% of data points tested, between groups [37]. One caveat to this calculation is the potential for variation in DNA methylation over time [8, 9]. This would have the effect of increasing population variance in methylation levels at age-sensitive loci. In order to deal with this it would be desirable to use age matched twin pair where possible for identification of epigenetic biomarkers specifically associated with phenotypic discordance. Despite this potential problem, limited attempts at power calculations so far undertaken are encouraging for the application of genome-scale methylation analysis to large twin-based studies.

It is hard to believe that just over 10 years ago the human genome had not been characterised fully, micro RNAs and other non-coding RNAs were considered largely artefactual by some, and epigenetics was considered of 'limited' importance biologically. Ten years from now it is likely that a full characterisation of both genomic and epigenomic data will be within the reach of most researchers; in the case of epigenomic data, this will be contingent on the provision of all manner of tissues from willing donors. In many cases (such as neural tissue in neuropsychiatric disorders), this will need to be done post mortem. It is likely that the level of complexity and variation revealed within individual cells/tissues/organs of any individual, *and* between individuals will be immense, greatly hindering the identification of biologically meaningful differences in epigenetic profile. Despite this, the inherent capacity to control for genetic variation in MZ twin pairs (and to a lesser extent environmental factors) should prove incredibly valuable in helping to unravel the mind boggling complexity of gene:environment:epigenetic interactions likely to contribute to human health and disease.

References

1. Hasler R, Begun A, Freitag-Wolf S, Kerick M, Mah N, Zvirbliene A et al (2009) Genetic control of global gene expression levels in the intestinal mucosa: a human twin study. Physiol Genomics 38:73–79
2. Sharma K (2005) Genetic epidemiology of epilepsy: a twin study. Neurol India 53:93–98
3. Tan Q, Christensen K, Christiansen L, Frederiksen H, Bathum L, Dahlgaard J et al (2005) Genetic dissection of gene expression observed in whole blood samples of elderly Danish twins. Hum Genet 117:267–274
4. Gordon L, Joo JH, Andronikos R, Ollikainen M, Wallace EM, Umstad MP et al (2011) Clues to the effect of intrauterine milieu on gene expression from a study of two tissues from neonatal monozygotic twins. Epigenetics 6:579–592

5. Gartner K (1990) A third component causing random variability beside environment and genotype. A reason for the limited success of a 30 year long effort to standardize laboratory animals? Lab Anim 24:71–77
6. Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, Hod E et al (2008) Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. Nat Genet 40:904–908
7. Schalkwyk LC, Meaburn EL, Smith R, Dempster EL, Jeffries AR, Davies MN et al (2010) Allelic skewing of DNA methylation is widespread across the genome. Am J Hum Genet 86:196–212
8. Fuke C, Shimabukuro M, Petronis A, Sugimoto J, Oda T, Miura K et al (2004) Age related changes in 5-methylcytosine content in human peripheral leukocytes and placentas: an HPLC-based study. Ann Hum Genet 68:196–204
9. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML et al (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 102:10604–10609
10. Bell JT, Spector TD (2011) A twin approach to unraveling epigenetics. Trends Genet 27:116–125
11. Wong CC, Caspi A, Williams B, Craig IW, Houts R, Ambler A et al (2010) A longitudinal study of epigenetic variation in twins. Epigenetics 5:516–526
12. Ollikainen M, Smith KR, Joo EJ, Ng HK, Andronikos R, Novakovic B et al (2010) DNA methylation analysis of multiple tissues from newborn twins reveals both genetic and intrauterine components to variation in the human neonatal epigenome. Hum Mol Genet 19:4176–4188
13. Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H et al (2008) Intra-individual change over time in DNA methylation with familial clustering. Jama 299:2877–2883
14. Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G et al (2010) Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. FASEB J 24:3135–3144
15. Li L, Wang L, Le F, Liu X, Yu P, Sheng J et al (2011) Evaluation of DNA methylation status at differentially methylated regions in IVF-conceived newborn twins. Fertil Steril 95:1975–1979
16. Zwijnenburg PJ, Meijers-Heijboer H, Boomsma DI (2010) Identical but not the same: the value of discordant monozygotic twins in genetic research. Am J Med Genet B Neuropsychiatr Genet 153B:1134–1149
17. Sakuntabhai A, Ruiz-Perez V, Carter S, Jacobsen N, Burge S, Monk S et al (1999) Mutations in ATP2A2, encoding a Ca2+ pump, cause Darier disease. Nat Genet 21:271–277
18. Robertson SP, Thompson S, Morgan T, Holder-Espinasse M, Martinot-Duquenoy V, Wilkie AO et al (2006) Postzygotic mutation and germline mosaicism in the otopalatodigital syndrome spectrum disorders. Eur J Hum Genet 14:549–554
19. Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y et al (2002) Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. Nat Genet 32:285–289
20. Smith AC, Rubin T, Shuman C, Estabrooks L, Aylsworth AS, McDonald MT et al (2006) New chromosome 11p15 epigenotypes identified in male monozygotic twins with Beckwith-Wiedemann syndrome. Cytogenet Genome Res 113:313–317
21. den Enden AT Helderman-van, Maaswinkel-Mooij PD, Hoogendoorn E, Willemsen R, Maat-Kievit JA, Losekoot M et al (1999) Monozygotic twin brothers with the fragile X syndrome: different CGG repeats and different mental capacities. J Med Genet 36:253–257
22. Kruyer H, Mila M, Glover G, Carbonell P, Ballesta F, Estivill X (1994) Fragile X syndrome and the (CGG)n mutation: two families with discordant MZ twins. Am J Hum Genet 54:437–442
23. Bonilla E, Younger DS, Chang HW, Tantravahi U, Miranda AF, Medori R et al (1990) Partial dystrophin deficiency in monozygous twin carriers of the Duchenne gene discordant for clinical myopathy. Neurology 40:1267–1270

24. Biousse V, Brown MD, Newman NJ, Allen JC, Rosenfeld J, Meola G et al (1997) De novo 14484 mitochondrial DNA mutation in monozygotic twins discordant for Leber's hereditary optic neuropathy. Neurology 49:1136–1138
25. Blakely EL, He L, Taylor RW, Chinnery PF, Lightowlers RN, Schaefer AM et al (2004) Mitochondrial DNA deletion in "identical" twin brothers. J Med Genet 41:e19
26. Bruder CE, Piotrowski A, Gijsbers AA, Andersson R, Erickson S, de Diaz ST et al (2008) Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. Am J Hum Genet 82:763–771
27. Bakaysa SL, Mucci LA, Slagboom PE, Boomsma DI, McClearn GE, Johansson B et al (2007) Telomere length predicts survival independent of genetic influences. Aging Cell 6:769–774
28. Slagboom PE, Droog S, Boomsma DI (1994) Genetic determination of telomere size in humans: a twin study of three age groups. Am J Hum Genet 55:876–882
29. Bennett CM, Boye E, Neufeld EJ (2008) Female monozygotic twins discordant for hemophilia A due to nonrandom X-chromosome inactivation. Am J Hematol 83:778–780
30. Willemsen R, Olmer R, De Diego Otero Y, Oostra BA (2000) Twin sisters, monozygotic with the fragile X mutation, but with a different phenotype. J Med Genet 37:603–604
31. Zneimer SM, Schneider NR, Richards CS (1993) In situ hybridization shows direct evidence of skewed X inactivation in one of monozygotic twin females manifesting Duchenne muscular dystrophy. Am J Med Genet 45:601–605
32. Javierre BM, Fernandez AF, Richter J, Al-Shahrour F, Martin-Subero JI, Rodriguez-Ubreva J et al (2010) Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. Genome Res 20:170–179
33. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc 57:289–300
34. Kaminsky ZA, Tang T, Wang SC, Ptak C, Oh GH, Wong AH et al (2009) DNA methylation profiles in monozygotic and dizygotic twins. Nat Genet 41:240–245
35. Baranzini SE, Mudge J, van Velkinburgh JC, Khankhanian P, Khrebtukova I, Miller NA et al (2010) Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. Nature 464:1351–1356
36. Shur N (2009) The genetics of twinning: from splitting eggs to breaking paradigms. Am J Med Genet C Semin Med Genet 151C:105–109
37. Kaminsky Z, Petronis A, Wang SC, Levine B, Ghaffar O, Floden D et al (2008) Epigenetics of personality traits: an illustrative study of identical twins discordant for risk-taking behavior. Twin Res Hum Genet 11:1–11
38. Spector TD (2000) Advances in twin and sib-pair analysis. Greenwich Medical Media
39. Schneider E, Pliushch G, El Hajj N, Galetzka D, Puhl A, Schorsch M et al (2010) Spatial, temporal and interindividual epigenetic variation of functionally important DNA methylation patterns. Nucleic Acids Res 38:3880–3890
40. Abecasis GR, Cookson WO (2000) GOLD–graphical overview of linkage disequilibrium. Bioinformatics 16:182–183
41. Heijmans BT, Kremer D, Tobi EW, Boomsma DI, Slagboom PE (2007) Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. Hum Mol Genet 16:547–554
42. Petronis A, Gottesman II, Kan P, Kennedy JL, Basile VS, Paterson AD et al (2003) Monozygotic twins exhibit numerous epigenetic differences: clues to twin discordance? Schizophr Bull 29:169–178
43. Mill J, Dempster E, Caspi A, Williams B, Moffitt T, Craig I (2006) Evidence for monozygotic twin (MZ) discordance in methylation level at two CpG sites in the promoter region of the catechol-O-methyltransferase (COMT) gene. Am J Med Genet B Neuropsychiatr Genet 141B:421–425
44. Mitchell MM, Lleo A, Zammataro L, Mayo MJ, Invernizzi P, Bach N et al (2011) Epigenetic investigation of variably X chromosome inactivated genes in monozygotic female twins discordant for primary biliary cirrhosis. Epigenetics 6:95–102

45. Weksberg R, Shuman C, Caluseriu O, Smith AC, Fei YL, Nishikawa J et al (2002) Discordant KCNQ1OT1 imprinting in sets of monozygotic twins discordant for Beckwith-Wiedemann syndrome. Hum Mol Genet 11:1317–1325
46. Yamazawa K, Kagami M, Fukami M, Matsubara K, Ogata T (2008) Monozygotic female twins discordant for Silver-Russell syndrome and hypomethylation of the H19-DMR. J Hum Genet 53:950–955
47. Oates NA, van Vliet J, Duffy DL, Kroes HY, Martin NG, Boomsma DI et al (2006) Increased DNA methylation at the AXIN1 gene in a monozygotic twin from a pair discordant for a caudal duplication anomaly. Am J Hum Genet 79:155–162
48. Mastroeni D, McKee A, Grover A, Rogers J, Coleman PD (2009) Epigenetic differences in cortical neurons from a pair of monozygotic twins discordant for Alzheimer's disease. PLoS One 4:e6617
49. Souren NY, Tierling S, Fryns JP, Derom C, Walter J, MP Zeegers (2011) DNA methylation variability at growth-related imprints does not contribute to overweight in monozygotic twins discordant for BMI. Obesity (Silver Spring) 19:1519–1522
50. Wu TD, Nacu S (2010) Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 26:873–881
51. Kuratomi G, Iwamoto K, Bundo M, Kusumi I, Kato N, Iwata N et al (2008) Aberrant DNA methylation associated with bipolar disorder identified from discordant monozygotic twins. Mol Psychiatry 13:429–441
52. Cobos I, Calcagnotto ME, Vilaythong AJ, Thwin MT, Noebels JL, Baraban SC et al (2005) Mice lacking Dlx1 show subtype-specific loss of interneurons, reduced inhibition and epilepsy. Nat Neurosci 8:1059–1068

Chapter 11
Age-Related Variation in DNA Methylation

Jean-Pierre Issa

Contents

11.1 DNA Methylation and the Regulation of Gene Expression 186
11.2 DNA Methylation Changes with Age 187
11.3 Causes of Age-Related DNA Methylation Changes 188
11.4 Consequences of Age-Related DNA Methylation Alterations 190
11.5 Age-Related Methylation and Cancer 192
11.6 Conclusions 193
References 194

Abstract Methylation is a ubiquitous, naturally occurring modification of DNA in mammalian cells that mediates stable repression of gene expression in epigenetically regulated genes, and plays diverse other roles such as regulation of chromosome structure and silencing of endogenous retrotransposons. DNA methylation patterns are generally stable in the short term but show prominent changes in aging cells and tissues including gains of methylation at previously protected promoter regions and losses of methylation genome wide. These age-related methylation changes remain mechanistically mysterious but are conserved from mouse to man, and are likely caused by infidelity in replication of the epigenome over time. Because of the link between DNA methylation and gene expression, these changes result in a mosaic epigenome in aged cells that could underlie diseases of aging such as cancer and atherosclerosis. Maintaining the health of the epigenome is worth investigating as a strategy to prevent age-related diseases.

J.-P. Issa (✉)
Fels Institute for Cancer and Molecular Biology, Temple University, 3307 North Broad Street, Rm 154 – PAHB, Philadelphia, PA 19140
e-mail: jpissa@temple.edu

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_11,
© Springer Science+Business Media B.V. 2012

11.1 DNA Methylation and the Regulation of Gene Expression

Methylation refers to the covalent addition of a methyl group and can be done to multiple biological molecules including DNA, RNA and proteins. DNA methylation is normally limited to the cytosine base, primarily in the context of cytosine followed by guanosine (the CpG dinucleotide) in adult cells [1]. The methylation reaction takes place after DNA synthesis, usually in the first few minutes post replication. It is catalyzed by DNA methyltransferases (DNMTs). DNMT3a and DNMT3b are primarily expressed during embryogenesis and in stem cells, and establish initial methylation patterns while DNMT1 is responsible for maintaining DNA methylation by copying the patterns seen on the original strand to the newly synthesized strand [2]. Thus, past embryogenesis, DNA methylation tends to be very stable in adult cells [3, 4].

The impact of DNA methylation on gene expression is highly variable depending on location of the involved CpG sites [5]. Thus, while it is enzymatically a single process, DNA methylation is best understood as multiple biological processes that utilize a common set of enzymes. The prime determinant of the function of a methylated CpG site is its location in the genome. One should distinguish sites located in CpG islands (short stretches of CpG rich DNA) or outside CpG islands. Similarly, one should differentiate promoter sites, genic sites and intergenic sites. Table 11.1

Table 11.1 DNA methylation in different genomic compartments

Location	Normal methylation patterns	Impact on expression	Change with age
Promoter CpG island	<2% of promoters are methylated	Strong silencing	Net increase, though a decrease in a few methylated promoters can also be seen
Non-promoter CpG island	20–30% are methylated	Some are silenced alternate promoters; most have an unclear impact	Net increase, though a decrease in a few methylated CpG islands can also be seen
Promoter non-CpG island	20–30% of promoters are methylated	Associated with lower expression, but can be reversed by transcriptional activation	Variable
Intragenic non-CpG island	Most CpG sites are highly methylated	Weak positive association with expression	Decrease
Intergenic non-CpG island	Most CpG sites are highly methylated	Most are in silenced retrotransposons and have an unclear impact on genes	Decrease

summarizes the methylation compartments, normal methylation patterns and the impact of DNA methylation on gene expression. The most compelling relation is at promoter CpG islands where the presence of methylation is rare, but associated with very stable and irreversible (in adult cells) silencing [1]. Most of the affected promoters are on the inactive X-chromosome in women, on monoallelically expressed genes (e.g. imprinted genes) and on a class of genes expressed exclusively in germ cells [6]. The impact of methylation in other genomic compartments is more variable, from reversible gene silencing (e.g. the IL2 gene promoter in resting lymphocytes [7]) to no clear impact on expression (e.g. intergenic sites, with the possible exception of enhancer sites where methylation remains poorly studied).

The mechanisms that underlie establishment and maintenance of DNA methylation at specific CpG sites remain somewhat mysterious. DNMTs are not known to have particular sequence specificity and it is assumed that the decision to methylate or not a given CpG site depends on co-factors that remain unknown. Recently, the TET family of enzymes was found to catalyze the formation of hydroxy-methylation at sites of cytosine methylation [8, 9]. There is no mechanism of maintaining hydroxymethylation post replication; thus, this reaction may be an intermediate step in catalyzing DNA demethylation. In some cases, strong gene activation is associated with localized removal of DNA methylation, best exemplified at non-CpG island promoters [7]. Thus, there is some dynamic regulation of DNA methylation in adult cells, but overall the patterns are very stable. Indeed, epigenetic processes marked by promoter CpG island methylation such as X-inactivation and imprinting are maintained through numerous cell division cycles in-vitro or in-vivo.

11.2 DNA Methylation Changes with Age

The earliest studies on DNA methylation with age focused on global measurement of 5-methylcytosine, which primarily reflects non-promoter, non-CpG island DNA. Initial studies showed progressive depletion of 5-methylcytosine associated with passaging normal fibroblasts in-vitro [10], a surrogate for age-dependent senescence. Subsequently, similar changes were observed in aging mouse tissues as well as human bronchial epithelial cells [11]. A study in human leukocytes demonstrated a relatively small but statistically significant decline (from ~4.05% to ~3.95%) over seven decades [12]. As expected given the location of most CpG sites, this decline can also be seen in repetitive elements such as retrotransposons, though the small changes noted require a large number of patients and a range of ages of several decades to be reliably detected [13]. It is important to note that genome wide studies have recently shown shifts in global methylation associated with stem cell differentiation[14]. Some studies have also reported shifts in 5-mC content associated with physiologic changes in proliferation (e.g. post-hepatectomy [15, 16]). Thus it remains possible that the shifts in 5-mC content with age (or passage number in-vitro) could be related to small differences in differentiation or proliferation states rather than a true aging defect.

When technology evolved to allow quantitative analysis of DNA methylation at specific sites, it became apparent that human aging was accompanied by small but measurable changes in many genes. Initial studies focused on a handful of genes with promoter CpG islands that were thought to be unmethylated in normal tissues but hypermethylated in cancer cells. In fact, careful studies unveiled a small degree of methylation in normal tissues that increased linearly with age in the Western populations studied. The first gene where this was described was *ERα*, where methylation increases at a rate of 1% every 3 years in human colon [17], but it quickly became apparent that many genes behaved similarly [18]. An unbiased, genome wide study showed that about 70–80% of promoter CpG islands that show hypermethylation in cancer also demonstrate this phenomenon of age-related hypermethylation [19]. CpG island hypermethylation is not a feature of differentiation. When present, it is established in embryogenesis and remains constant post-development [6]. Thus, changes in this genomic compartment truly represent a pathologic event that accompanies aging.

Multiple studies have now confirmed age-related promoter methylation in human tissues. An autopsy study showed that some degree of hypermethylation could be observed in most tissues examined, with significant tissue to tissue variability [20]. In fact, *ERα* methylation is highest in normal liver, followed by colon epithelium, followed by colon stroma, and lowest in normal breast and brain tissues [18]. In the colon, careful crypt by crypt analysis showed considerable variation in age-related methylation, with much higher inter-crypt than intra-crypt heterogeneity [21]. Given that each colonic crypt in older individuals is derived from a single or very few stem cells, the data suggested that age-related methylation was a property of stem cells that is conserved upon differentiation.

Genome wide technologies to study DNA methylation have evolved considerably recently and facilitated studies of aging. Examination of mouse tissues showed the same degree of age-related CpG island hypermethylation in the GI tract as that seen in humans [22]. In addition, age-related hypomethylation was also seen in a few normally methylated promoter CpG islands. Interestingly, methylation changes in mice reach the same degree in 3 years as is reached in 80 years in humans. Thus, methylation changes reflect physiologic age rather that chronologic age. Mouse models also allowed a careful examination of the tissue-specificity of the process. While aberrant methylation could be seen in all tissues, the degree of change varied considerably, with the most prominent changes seen in the GI tract, followed by the spleen, and relatively modest changes in other tissues such as kidneys. Genome wide studies of human tissues also revealed numerous changes (both hyper and hypomethylation) in many tissues [23–26].

11.3 Causes of Age-Related DNA Methylation Changes

In considering causes for age-related changes in methylation, one has to keep in mind that it is bidirectional – i.e. increases at some sites, and decreases at others [22]. Thus, one cannot invoke simple explanations such as changes in DNMTs with

age, which in any case are minor. Unifying features of the process include conservation over evolution and tissue-specificity, with the most proliferative tissues (GI epithelium) showing the highest degree of change. Thus, a simple explanation for the process is replication errors in maintenance of epigenetic states. If one simply postulates that every stem cell replication event is associated with a finite chance of a DNA methylation error at a given CpG site, then methylation may serve as a mitotic clock [21] in the same way as telomere shortening does.

The replication error/mitotic clock hypothesis for age-related methylation does not explain all the features of the process. In particular, there is considerable gene and genomic compartment specificity to the process; many unmethylated promoter CpG islands are completely resistant to age-related changes, and normally methylated genes (e.g. imprinted loci) also show substantial variation in susceptibility to age-related change. Genomic factors linked to this variation include the density of retrotransposons surrounding a given promoter [27] (a high density is associated with fewer changes with age) and the occupancy of the promoter by polycomb group proteins in embryonic cells [26] (with polycomb group targets showing more age-related changes). In addition, baseline gene expression very likely influences the process as well [28]. One can resolve the discrepancy between a random error hypothesis and the observed gene specificity by postulating that the fidelity of transmission of epigenetic patterns can be variable across the genome, in the same way as DNA repair can vary between transcribed and non-transcribed regions [29]. Thus, it is likely that epigenetic errors are repaired (or prevented) more readily in certain regions, perhaps as a result of local features such as transcription factor binding, chromatin characteristics and nucleosome density. Strong evidence for this came from studying a length polymorphism in the RIL gene promoter that creates an extra SP1/SP3 binding site [30]. Individuals who carry the longer allele show less age-related methylation at that site than those with a short allele. The regulation of methylation maintenance and errors remains somewhat mysterious, but there is increasing evidence for such a process. Indeed, the TET family of proteins may play a role in preventing DNA methylation errors [31], and it is possible that their relative targeting to genomic sites determines that site's susceptibility to age-related changes. A "methylation error" hypothesis could account for both hypermethylation and hypomethylation with age at promoter CpG islands. However, because few genes are normally methylated in promoter CpG islands [6], hypermethylation is quantitatively the predominant feature of aging.

The mitotic clock hypothesis, modified to account for uneven methylation maintenance across the genome, is very consistent with both species conservation and tissue-specificity of age-related methylation [22]. It also provides a useful framework to interpret studies on natural variation in the process. Perhaps the single dominant factor in modulating age-related methylation is chronic inflammation. In the colon [32], esophagus [33], stomach [34] and liver [35], chronic inflammation is associated with substantially increased methylation in normal appearing tissues. In a gerbil model of Helicobacter Pylori stomach infection, a methylation increase after infection required chronic inflammation[36], and while bacterial eradication reduced methylation, it did not return to baseline. In a mouse model of inflammatory

bowel disease, inflammation was associated with a marked increase of methylation of genes targeted by polycomb in embryonic stem cells [37]. Thus, inflammation is a universal accelerator of age-related methylation. While it is possible that inflammation interacts with DNMTs or DNA itself, the changes observed are also bi-directional (increases at some genes, decreases at others) just as observed in aging. Thus, it seems more likely that the inflammation effect is to accelerate the mitotic clock through stimulated stem cell proliferation due to cycles of tissue injury and repair.

From a public health perspective, it is important to learn whether environmental exposures and lifestyle factors can also modulate age-related methylation independent of inflammation. A possible player here is dietary habits. One carbon metabolism influences the availability of the methyl donor S-adenosyl-methionine, which could theoretically drive the enzymatic DNA methylation forward [38]. Indeed, gestational exposure to varying levels of folate and vitamin B12 can influence DNA methylation patterns in newborn mice [39]; it is not known whether this extends to diets in childhood or adulthood. Human studies have yielded mixed results so far in this respect, but it may take very large studies of individuals followed for decades to be sure of this. A recent study showed that RBC-folate levels, a measure of chronic exposure, positively correlated with age-related methylation [40]. The top and bottom quartiles in RBC-folate had the same difference in age-related methylation as 10 years of age. On the other hand, folate ingestion was also shown to induce an inflammatory like gene expression profile in the colon [41], and it is therefore possible that the observed association with RBC-folate (which should be validated in other studies) is also related to inflammation. The diet-DNA methylation link and a potential link between environmental chemical exposures and altered epigenetics in humans will only be resolved by extremely precise studies of carefully selected loci in large populations.

11.4 Consequences of Age-Related DNA Methylation Alterations

The consequences of age-related methylation changes remain a matter of speculation. Functionally, changes in promoter methylation, particularly at CpG islands, can clearly impact expression of the gene in affected cells [1]. To fully appreciate the impact of aging, one has first to consider the degree of methylation changes. Quantitative measurements reveal differences in the order of 5–25% methylation (at susceptible genes) between young and old mammals [22]. The impact of this depends on whether methylation is uniformly altered to a small degree in every cell, or greatly altered in a small proportion of cells. This can be tested in two ways – by single stem cell or single crypt analysis, or by high resolution analysis using a technique known as bisulfite cloning-sequencing. Both types of experiments arrive at the same conclusion – that aged cells are epigenetic mosaics, with some cells having a very low degree of methylation at a given gene and others having a very high

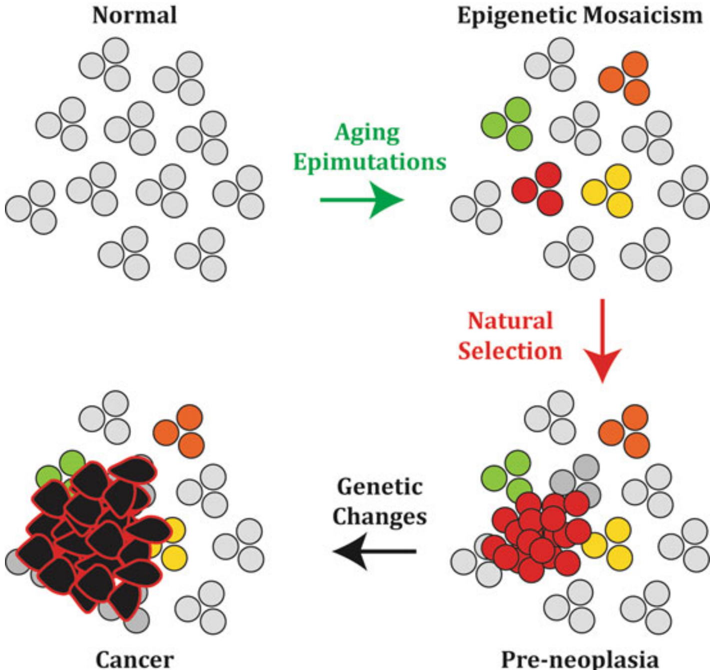


Fig. 11.1 A model of the contributions of age-related DNA methylation alterations to focal diseases such as cancer. In this model, it is proposed that our normal stem cells, represented by clusters of three cells with uniform borders in the top left quadrant, have uniform patterns of epigenetic regulation (represented by the *filled color*). Age-related DNA methylation results in epigenetic mosaicism (*top right*) creating slight stem cell to stem cell variation in gene expression. This variation serves as the engine driving natural selection that promotes the overgrowth of the stem cell that randomly acquires the appropriate gene expression pattern (*red cell cluster, bottom right*). This overgrowth leads to focal problems such as benign tumors (pre-neoplasia) and atherosclerotic lesions. Because of over-proliferation, these cells are also susceptible to genetic damage, either random or carcinogen induced, and the combination of epigenetic priming and genetic damage allows oncogenic mutations (e.g. KRAS mutation) to transform the cells into full malignancy (*black filled cells* with jagged edges, *bottom left*)

degree of methylation at the same gene [21]. These stark differences, which are underestimated when one studies averages of hundreds of thousands of cells, would be expected to result in large differences in gene expression between different stem cells in the same tissue. If one multiplies this effect by the hundreds of genes that show age-related methylation, a picture emerges whereby tissues in older individuals have considerable stem cell to stem cell variability in epigenetic profiles (Fig. 11.1). This variability has been demonstrated at the level of gene expression in intestinal crypts.

Does epigenetic mosaicism impact on the function of aged organs? At first glance, one would speculate that it is unlikely, given the vast reserves we are born with. For several organs, losing up to 90% of the cells has no appreciable impact on function.

However, there are circumstances where this may not be true. Epigenetic mosaicism may constrain the plasticity of stem cells. Indeed, many of the genes affected by age-related methylation are polycomb targets in embryonic stem cells, and are enriched for genes involved in differentiation and organ development [22, 27, 37]. Thus, it is possible that constraints on plasticity compromises the function of some aged stem cells in subtle ways that still could be contributing to declines in tissue function. For example, aged hematopoietic stem cells show skewing in differentiation, with more myeloid and fewer lymphoid progeny [42]; one wonders whether diminished stem cell plasticity and skewed differentiation could contribute to the known functional declines associated with aging.

While whole tissue function may be minimally affected by age-related methylation, the situation is very different for focal changes. Epigenetic mosaicism equates to gene expression variation which, in turn, is a powerful engine for natural selection. Thus, Darwinian mechanisms dictate that a high degree of epigenetic variation would predispose to emergence of cells with a selective growth advantage compared to neighboring cells in a given tissue. The simplest model (Fig. 11.1) would predict that variation plus natural selection results in the emergence of hyperproliferative clonal lesions in affected tissues. This may have different consequences depending on the tissue, the location of the lesion and the nature of existing co-factors. Thus, in tissues exposed to carcinogens, the initial clonal proliferation (and additional selective pressures) might allow oncogenic mutations to transform cells, resulting in full blown malignancy. Keeping in mind that many oncogenic mutations lead to senescence or death of normal cells, the convergence of age-related epigenetic changes and mutations may be the key factor in explaining the dramatic rise in incidence of cancer as humans age. Indeed, a high degree of DNA methylation abnormalities is observed in cancer [43], and often precedes full blown malignancy. In tissues where carcinogen exposure and spontaneous mutations are less prominent, focal proliferation itself may still lead to pathology, as in coronary atherosclerosis. In support of this, aberrant methylation of a few genes such as $ER\alpha$ and $ER\beta$ has been observed in atherosclerotic plaques [44, 45]. It also possible to imagine other situations where alterations in a small percentage of cells could result in significant pathology, such as through induction of a local inflammatory or fibrotic condition. Thus, several age-related pathologies from benign conditions such as loss of skin integrity to more serious diseases such as neurodegeneration could have an epigenetic component.

11.5 Age-Related Methylation and Cancer

Cancer provides an interesting example of how age-related methylation could interact with disease. Neoplastic cells show a dramatic accumulation of aberrant methylation (and other epigenetic alterations) [43]. Increased promoter CpG island

methylation in older cells was discovered by studying normal-appearing tissues adjacent to cancer, and aging changes in normal tissues were suggested to be an early step in carcinogenesis [17]. Indeed, much of the DNA methylation abnormalities in cancer can be traced to aging defects [19], and cancer itself can be seen as reflecting the state of very old cells, if one considers stem cell replication as the ultimate mark of the biologic clock. But do age-related methylation defects cause cancer? Causation is difficult to establish in humans. In mouse models, reducing methylation in normal colon by partial deletion of DNMT1 prevents the development of polyps [46], while increasing methylation by over-expression of DNMT3b increases polyp formation [47]. These data would support the model in Fig. 11.1 that predicts that age-related methylation defects may be rate-limiting for the development of pre-neoplastic lesions or "primed" cells. In human carcinogenesis, cancers are often found to arise in fields of aberrant DNA methylation [48], which is also consistent with the model. On the other hand, it has not been conclusively shown that the degree of methylation abnormalities in normal tissues correlate with the risk of cancer development, or of recurrence of intestinal polyps after resection [40]. Indeed, age-related changes in the colon are ubiquitous in the population, yet the lifetime prevalence of colon cancer is only about 5% in the US. Thus, age-related methylation alterations are clearly not sufficient to bring about neoplasia. As discussed earlier, a useful framework to think of cancer development then would be that age-related methylation subtly changes the physiology of a few stem cells, making them susceptible to neoplastic transformation, perhaps by acquisition of oncogenic mutations [49]. In this model, neither epigenetic changes nor genetic changes in isolation are sufficient for transformation, and the rate limiting step for neoplasia would be the co-occurrence of both processes. Implications for cancer prevention and risk assessment can be hypothesized from this model. Long term epigenetic intervention would be effective for prevention (as seen in mice), but short term studies would not. Conversely, a high degree of age-related changes would predict for the earlier occurrence of pre-neoplastic lesions, but the propensity to acquire genetic changes might be a more effective indicator of cancer risk in older individuals. As we learn more about potential epigenetic interventions and technology to detect rare mutations evolves, these hypotheses could become testable in human studies.

11.6 Conclusions

Aging is accompanied by progressive, proliferation dependent DNA methylation deregulation that creates epigenetic mosaicism in older tissues. These changes may compromise stem cell function and contribute to diseases of aging, particularly focal diseases such as neoplasia. Interventions aimed at reducing age-related methylation deregulation have the potential to attenuate the occurrence and severity of these diseases.

References

1. Bird A (2002) DNA methylation patterns and epigenetic memory. Genes Dev 16:6–21
2. Jaenisch R, Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 33(Suppl):245–254
3. Migeon BR, Axelman J, Beggs AH (1988) Effect of ageing on reactivation of the human x-linked hprt locus [see comments]. Nature 335:93–96
4. Pagani F, Toniolo D, Vergani C (1990) Stability of DNA methylation of X-chromosome genes during aging. Somat Cell Mol Genet 16:79–84
5. Estecio MR, Issa JP (2010) Dissecting DNA hypermethylation in cancer. FEBS Lett 585:2078–2086
6. Shen L, Kondo Y, Guo Y, Zhang J, Zhang L, Ahmed S et al (2007) Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. PLoS Genet 3:2023–2036
7. Bruniquel D, Schwartz RH (2003) Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. Nat Immunol 4:235–240
8. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324:930–935
9. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466:1129–1133
10. Wilson VL, Jones PA (1983) Dna methylation decreases in aging but not in immortal cells. Science 220:1055–1057
11. Wilson VL, Smith RA, Ma S, Cutler RG (1987) Genomic 5-methyldeoxycytidine decreases with age. J Biol Chem 262:9948–9951
12. Fuke C, Shimabukuro M, Petronis A, Sugimoto J, Oda T, Miura K et al (2004) Age related changes in 5-methylcytosine content in human peripheral leukocytes and placentas: an HPLC-based study. Ann Hum Genet 68:196–204
13. Bollati V, Schwartz J, Wright R, Litonjua A, Tarantini L, Suh H et al (2009) Decline in genomic DNA methylation through aging in a cohort of elderly subjects. Mech Ageing Dev 130:234–239
14. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J et al (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462:315–322
15. Lapeyre JN, Becker FF (1979) 5-Methylcytosine content of nuclear DNA during chemical hepatocarcinogenesis and in carcinomas which result. Biochem Biophys Res Commun 87:698–705
16. Kren BT, Trembley JH, Steer CJ (1996) Alterations in mRNA stability during rat liver regeneration. Am J Physiol 270:G763–G777
17. Issa JP, Ottaviano YL, Celano P, Hamilton SR, Davidson NE, Baylin SB (1994) Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. Nat Genet 7:536–540
18. Ahuja N, Li Q, Mohan AL, Baylin SB, Issa JPJ (1998) Aging and DNA methylation in colorectal mucosa and cancer. Cancer Res 58:5489–5494
19. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JPJ (1999) CpG Island methylator phenotype in colorectal cancer. Proc Natl Acad Sci USA 96:8681–8686
20. Waki T, Tamura G, Sato M, Motoyama T (2003) Age-related methylation of tumor suppressor and tumor-related genes: an analysis of autopsy samples. Oncogene 22:4128–4133
21. Shibata D (2009) Inferring human stem cell behaviour from epigenetic drift. J Pathol 217:199–205
22. Maegawa S, Hinkal G, Kim HS, Shen L, Zhang L, Zhang J et al (2010) Widespread and tissue specific age-related DNA methylation changes in mice. Genome Res 20:332–340

23. Hernandez DG, Nalls MA, Gibbs JR, Arepalli S, Van der Brug M, Chong S et al (2011) Distinct DNA methylation changes highly correlated with chronological age in the human brain. Hum Mol Genet 20:1164–1172
24. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL et al (2009) Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet 5:e1000602
25. Gronniger E, Weber B, Heil O, Peters N, Stab F, Wenck H et al (2010) Aging and chronic sun exposure cause distinct epigenetic changes in human skin. PLoS Genet 6:e1000971
26. Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H et al (2010) Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome Res 20:434–439
27. Estecio MR, Gallegos J, Vallot C, Castoro RJ, Chung W, Maegawa S et al (2010) Genome architecture marked by retrotransposons modulates predisposition to DNA methylation in cancer. Genome Res 20:1369–1382
28. Takeshima H, Yamashita S, Shimazu T, Niwa T, Ushijima T (2009) The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands. Genome Res 19:1974–1982
29. Hanawalt PC, Spivak G (2008) Transcription-coupled DNA repair: two decades of progress and surprises. Nat Rev Mol Cell Biol 9:958–970
30. Boumber YA, Kondo Y, Chen X, Shen L, Guo Y, Tellez C et al (2008) An Sp1/Sp3 binding polymorphism confers methylation protection. PLoS Genet 4:e1000162
31. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J et al (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 473:343–348
32. Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA (2001) Accelerated age-related CpG island methylation in ulcerative colitis. Cancer Res 61:3573–3577
33. Eads CA, Lord RV, Kurumboor SK, Wickramasinghe K, Skinner ML, Long TI et al (2000) Fields of aberrant CpG island hypermethylation in Barrett's esophagus and associated adenocarcinoma. Cancer Res 60:5021–5026
34. Maekita T, Nakazawa K, Mihara M, Nakajima T, Yanaoka K, Iguchi M et al (2006) High levels of aberrant DNA methylation in Helicobacter pylori-infected gastric mucosae and its possible association with gastric cancer risk. Clin Cancer Res 12:989–995
35. Shen L, Ahuja N, Shen Y, Habib NA, Toyota M, Rashid A et al (2002) DNA methylation and environmental exposures in human hepatocellular carcinoma. J Natl Cancer Inst 94:755–761
36. Niwa T, Tsukamoto T, Toyoda T, Mori A, Tanaka H, Maekita T et al (2010) Inflammatory Processes Triggered by Helicobacter pylori Infection Cause Aberrant DNA Methylation in Gastric Epithelial Cells. Cancer Res 70:1430–1440
37. Hahn MA, Hahn T, Lee DH, Esworthy RS, Kim BW, Riggs AD et al (2008) Methylation of polycomb target genes in intestinal cancer is mediated by inflammation. Cancer Res 68:10280–10289
38. Mason JB (2003) Biomarkers of nutrient exposure and status in one-carbon (methyl) metabolism. J Nutr 133:941S–947S
39. Waterland RA, Jirtle RL (2003) Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol 23:5293–5300
40. Wallace K, Grau MV, Levine AJ, Shen L, Hamdan R, Chen X et al (2010) Association between folate levels and CpG Island hypermethylation in normal colorectal mucosa. Cancer Prev Res (Phila) 3:1552–1564
41. Protiva P, Mason JB, Liu Z, Hopkins ME, Nelson C, Marshall JR et al (2011) Altered folate availability modifies the molecular environment of the human colorectum: implications for colorectal carcinogenesis. Cancer Prev Res (Phila) 4:530–543
42. Chambers SM, Shaw CA, Gatza C, Fisk CJ, Donehower LA, Goodell MA (2007) Aging hematopoietic stem cells decline in function and exhibit epigenetic dysregulation. PLoS Biol 5:e201
43. Taby R, Issa JP (2010) Cancer epigenetics. CA Cancer J Clin 60:376–392

44. Kim J, Kim JY, Song KS, Lee YH, Seo JS, Jelinek J et al (2007) Epigenetic changes in estrogen receptor beta gene in atherosclerotic cardiovascular tissues and in-vitro vascular senescence. Biochim Biophys Acta 1772:72–80
45. Post WS, Goldschmidt-Clermont PJ, Wilhide CC, Heldman AW, Sussman MS, Ouyang P et al (1999) Methylation of the estrogen receptor gene is associated with aging and atherosclerosis in the cardiovascular system. Cardiovasc Res 43:985–991
46. Laird PW, Jackson-Grusby L, Fazeli A, Dickinson SL, Jung WE, Li E et al (1995) Suppression of intestinal neoplasia by DNA hypomethylation. Cell 81:197–205
47. Linhart HG, Lin H, Yamada Y, Moran E, Steine EJ, Gokhale S et al (2007) Dnmt3b promotes tumorigenesis in vivo by gene-specific de novo methylation and transcriptional silencing. Genes Dev 21:3110–3122
48. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J et al (2005) MGMT promoter methylation and field defect in sporadic colorectal cancer. J Natl Cancer Inst 97:1330–1338
49. Issa JP (2008) Cancer prevention: epigenetics steps up to the plate. Cancer Prev Res (Phila Pa) 1:219–222

Chapter 12
Influence of Environmental Factors on the Epigenome

Brock C. Christensen, Carmen J. Marsit, and Karl T. Kelsey

Contents

12.1 Introduction 198
12.2 Environmental Exposures of Diet and Lifestyle and Epigenetics 199
  12.2.1 Folate, One-Carbon Metabolism, Diet, and Body Mass Index 199
  12.2.2 Alcohol 202
  12.2.3 Aging 203
  12.2.4 Tobacco 204
12.3 Environmental Toxicants and Epigenetic Alterations 206
  12.3.1 Endocrine Disruptors 206
  12.3.2 Benzene 207
  12.3.3 Radiation 207
  12.3.4 Arsenic 208
  12.3.5 Other Metals (Nickel, Lead, Cadmium, Chromium) 209
  12.3.6 Particulates/Air Pollution 209
  12.3.7 Asbestos 210
12.4 Environmental Exposures and Imprinting and Development 211
12.5 Mechanisms 211
12.6 Development of Environmental Epigenomics 214
References 215

B.C. Christensen (✉)
Department of Community and Family Medicine, Section of Biostatistics and Epidemiology, Department of Pharmacology and Toxicology, Dartmouth Medical School Hanover, Hanover, NH, USA
e-mail: Brock.C.Christensen@Dartmouth.edu

C.J. Marsit
Department of Pharmacology and Toxicology, Department of Community and Family Medicine, Section of Biostatistics and Epidemiology, Dartmouth Medical School, Hanover, NH, USA
e-mail: carmen.marsit@dartmouth.edu

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_12,
© Springer Science+Business Media B.V. 2012

Abstract In this chapter we will consider the role of environmental factors on the epigenome. The importance of rapidly emerging research into the types of exposures that may alter epigenetic marks is increasingly being recognized. A large portion of epigenetic research to date has focused on epigenetic alterations in cancer and therefore, exposures such as tobacco, alcohol, radiation, arsenic, and air pollution will be covered. Additionally, we review research on other exposures that may affect epigenetic states such as folate and diet, aging, and endocrine disruptors. We will briefly cover environmental exposures and imprinting and development, as well as discuss potential mechanisms for exposures to modify epigenetic states. Appropriate epidemiologic studies will be crucial to understanding the true effect of environmental exposures on the human epigenome and this work is urgently needed in order to better understand the biology of epigenetic alterations. With a more comprehensive understanding of the affects of exposures on the epigenome (including consideration of genetic background), not only will the prediction of the toxic potential of new compounds be more readily achieved, but more personalized disease prevention and treatment strategies may be developed.

12.1 Introduction

The epigenome is the landscape of mitotically heritable changes in gene expression and gene expression potential that are mediated without altering genetic sequence. A role for environmental exposures in epigenetic alteration is increasingly recognized and investigated, and much of the existing evidence implicating a role of environmental factors on the epigenome comes from studies of disease outcomes such as cancer and adverse reproductive/developmental events. Initially, epigenetic alterations were identified in various human tumors and consequently, environmental exposures known to have an etiologic role in cancer such as tobacco, air pollutants, and metals have been studied and implicated in the modification of epigenetic marks. Of course, this has highlighted the need to consider how environmental factors may influence the epigenome in pathologically normal tissues, perhaps representing

K.T. Kelsey
Department of Community Health, Center for Environmental Health and Technology, Providence, RI, USA

Pathology and Laboratory Medicine, The Warren Alpert Medical School of Brown University, Providence, RI, USA
e-mail: Karl_Kelsey@brown.edu

alterations necessary to initiate or predispose to disease phenotypes. The majority of epigenetic alteration work conducted to date has focused on DNA methylation, though emerging work is beginning to address the affect of environmental exposures on histone modifications and microRNA (miRNA) expression. This chapter highlights the current evidence that epigenetic alterations are associated with environmental exposures, focusing on human studies and drawing primarily on studies of human cancer. The potential mechanisms behind environmentally related epigenetic alterations and methods for studying the relation between the epigenome and the environment that are amenable to epidemiologic research are also presented. As the data from human studies has, for the most part, centered on alterations to DNA methylation, this chapter will focus on the examination of DNA methylation and environment.

12.2 Environmental Exposures of Diet and Lifestyle and Epigenetics

12.2.1 Folate, One-Carbon Metabolism, Diet, and Body Mass Index

One-carbon metabolism is the network of biochemical reactions essential to both DNA synthesis and all cellular biomolecule (nucleic acids, proteins, lipids) methylation reactions that involve the transfer of one-carbon groups. A critical nutrient central to one-carbon metabolism is folate (vitamin B_9), which donates its methyl group for homocysteine remethylation to methionine. Subsequently, methionine is the methyl donor for all cellular methylation reactions, most notable for this discussion being DNA and histone methylation via S-adenosyl methionine (SAM), and other B vitamins (B_2, B_6, B_{12}) that act as enzymatic cofactors in the network. Collectively, B vitamins, homocysteine, and methionine are important contributors to the maintenance of DNA integrity and DNA methylation. Figure 12.1 represents the links between diet, one-carbon metabolism, DNA synthesis, DNA methylation, and toxicant metabolism.

In mammals, establishment of somatic cell epigenetic patterns occurs early in fetal development subsequent to genome-wide reprogramming of epigenetic patterns at the multiple developmental stages (fertilization, implantation, blastocyst) when totipotency and pluripotency requirements are dynamic [1]. In addition, DNA methylation is necessary for the mitotic inheritance process of genomic imprinting (monoallelic gene expression which can also be tissue-specific). Embryonic and fetal nutrient availability have great potential to affect the epigenetic reprogramming and patterning phenomena, which has implications for proper development and perhaps even life-long conditioning and health [2, 3]. Transmission of both nutrients and environmental exposures to developing mammals in utero proceeds through the placenta, and maternal folate status is recognizably associated with

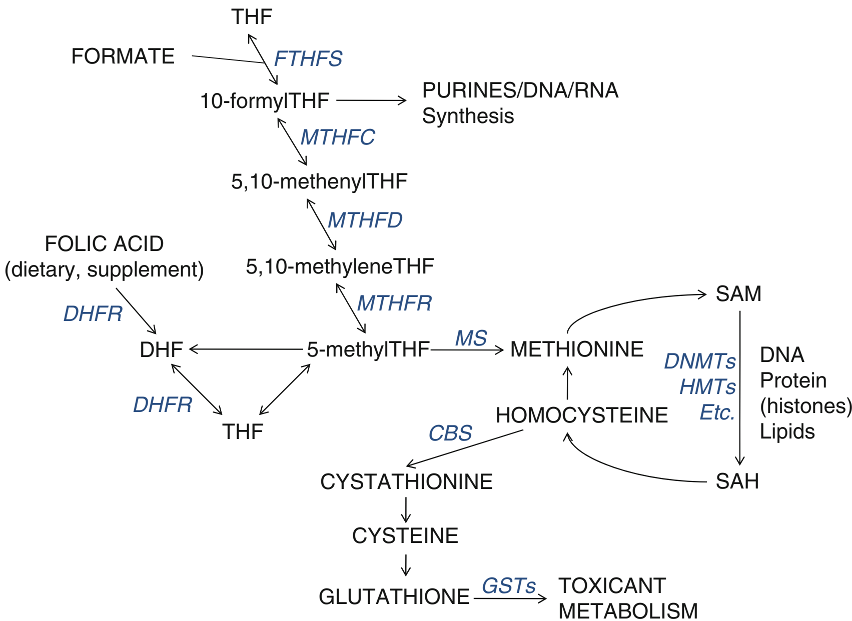


Fig. 12.1 Diet, one-carbon metabolism, DNA synthesis, DNA methylation, and toxicant metabolism

fetal development. For instance, it is well-known that folate deficiency is associated with neural tube defects (NTD) [4] and animal studies have indicated the necessity of sufficient methyl group availability for proper neural tube development [5]. Further, hypomethylation of long interspersed nucleotide elements (LINE) and genomic DNA has been associated with increased risk of NTDs in humans [6]. Beyond this, Waterland and Michels [7] have proposed that epigenetic mechanisms may contribute to the developmental origins of health and disease hypothesis, which states that nutrition and environmental stimuli at critical developmental periods (pre- and postnatal) can induce permanent alterations in susceptibility to metabolic and chronic diseases [8]. Additional details on epigenetic changes during development are available in Chap. 5, and the developmental origins of health and disease hypothesis is covered in Chap. 6.

There have been several animal studies that directly link folate intake, diet, and epigenetic alterations though similar evidence from human studies is only beginning to emerge. Many human studies have, however, described associations between folate, B vitamins, or diet and disease risk and suggested the potential for epigenetic mechanisms to account for these observations. For instance, elevated serum vitamin B_6 and methionine have been associated with reduced risk of lung cancer [9], folate intake has been associated with reduced risk of breast cancer (especially among women with high alcohol consumption) [10, 11] and a protective role for folate in

colon cancer [12] has been consistently reported. In a pooled analysis of 13 prospective cohort studies on folate intake and colon cancer Kim and colleagues reinforced the significant, though modest protective role of folate for reducing risk of colon cancer by approximately 2% per 100 μg/day of folate [13]. A recent work from the Netherlands Cohort Study on diet and cancer investigated risk of colorectal cancer according to methyl-donor (folate, methionine, and vitamins B_2 and B_6) intake and considering genotypes of enzymes in the one-carbon metabolism pathway [14]. These authors did not find significant diet-gene interactions, but did observe a significant inverse association between methionine intake and risk of colon cancer among individuals with common *DNMT3B* (*de novo* DNA methyltransferase) genotypes as well between B_2 intake and risk of colon cancer in individuals with one or less rare allele among the folate enzyme genes *MTHFR, MTRR,* and *MTR* [14]. In addition, a study of serum concentrations of folate and plasma SAM levels in over 330 women showed that body mass index (BMI) was the strongest determinant of SAM concentration, suggesting that nutrient intake and nutrient availability are distinguishable, and that BMI is a potential confounder of folate-methylation associations [15]. Yet, far fewer examples of studies providing direct lines of evidence for associations between folate and/or diet and epigenetic alterations in humans exist [16]. In one such study, Ingrosso et al. demonstrated that global DNA hypomethylation (measured through cytosine extension and Southern blot) was higher in uremia patients with hyperhomocysteinemia, and that folate treatment restored DNA methylation to normal levels [17]. Hypomethylation of non-diseased colon tissue DNA from disease-free controls, measured through ^{3}H-methyl incorporation, was found to be greater in samples amongst individuals with adenomas and colon tumors compared to controls, and this hypomethylation of both colon tissue and leukocyte DNA was associated with risk for colon cancer [18]. Folate status has been inversely correlated with colon tissue DNA methylation extent [19]. In head and neck squamous cell carcinoma (HNSCC) repetitive element methylation extent was associated with folate intake, was reduced amongst individuals with a variant *MTHFR* genotype, and was a significant risk factor for disease [20]. A study of tumor DNA methylation in over 160 breast tumors at over 750 cancer-related genes demonstrated a significant association between methylation profile and dietary folate when controlling for potential confounders such as alcohol intake [21].

Outside of cancer, although there is a large body of literature examining the relation of diet and disease risk, particularly examining nutrients involved in one-carbon metabolism, there is less direct evidence linking these associations through epigenetic mechanisms. As an example, there is a reasonable literature that suggests altered methylation status is associated with cardiovascular disease and this may be linked with folate status [22–24]; however, the causal nature of these associations remains to be shown. Further study in humans investigating the associations between dietary factors and methylation alterations (in normal and diseased states), including potential confounders such as BMI and alcohol are necessary.

Other areas of interest for future study include the potential for differential contribution of nutrient factors dependent upon disease context (healthy vs. diseased), the potential risks of high-dose nutrient supplementation practices, and interactions

between one-carbon network participants. Folate supplementation is associated with a reduced risk of many forms of cancer, though once disease is present it is possible that a relative increase in one-carbon nutrient levels may speed disease progression or invasiveness dependent upon the epigenetic character of the disease-state. Some evidence for this potential comes from our recent study of breast tumor methylation where increased folate intake (controlling for potential confounders) was associated with a methylation profile that was independently associated with increased tumor size [21]. Excessive folate supplementation may have potential adverse effects including masking of B_{12} deficiency, disruption of zinc function, or interfering with one-carbon homeostasis and additional study is required to elucidate these potential effects [16, 25]. The one-carbon pathway links nutrient availability with not only DNA methylation, but also with toxicant metabolism through glutathione and the glutathione transferase (GST) enzyme family (Fig. 12.1). Therefore, interactions between one-carbon pathway participants may modify associations between exposures, pathway participants, and methylation alterations. In fact, a study of reduced GST enzyme activity in a mouse model of Alzheimer's disease and neuronal health has shown that SAM can mediate the activity of GST enzymes [26]. This suggests that reactions dependent upon SAM are necessary for GST activity and that SAM may be a critical mediator of neuronal health [26]. Furthermore, it could be predicted that GST enzyme family genotypes (among other genotypes) will be shown to modify associations between certain exposures and DNA methylation alterations. Hence integrative studies that incorporate genetics (genotype and/or alterations) with measures of epigenetic alterations will allow a more comprehensive understanding of the relation among exposures, epigenetic alterations, and genetic states. However, at this time it is unclear the extent to which genotype may influence epigenetic states. More research is needed to ascertain best practices for assessing folate/methyl availability, perhaps comparing food-frequency-questionnaire data to homocysteine and/or SAM/SAH levels in a prospective manner.

12.2.2 Alcohol

Alcohol is known to interfere with folate absorption in the intestine and hepatic release of folate, and hence, supply to tissues [27]. Although it is a weak mutagen [28], by inhibiting the one-carbon metabolism network, alcohol may contribute to increased risk of disease in an epigenetic fashion. More specifically, ethanol can interfere with several steps of methionine metabolism and can lead to activation of betaine homocysteine methyltransferase to activate a compensatory pathway for maintenance of SAM levels [29]. However, with continued exposure, the compensatory pathway cannot be maintained and liver injury will result. In addition to liver injury, an etiologic role for alcohol in several cancers is well established [30]. For example, alcohol consumption increases colon cancer risk [31], though the association may be modified by dietary folate, and may be specific to LINE-1 hypomethylated

tumors [32]. More specifically, subjects with high folate intake were less likely to develop LINE-1 hypomethylated colon cancers but subjects with high alcohol consumption had a significantly increased risk of developing LINE-1 hypomethylated colon cancers with no association for more highly methylated LINE-1 tumors [32]. In addition, alcohol has a well-recognized etiologic role in HNSCC, and in these tumors, alcohol consumption has been significantly associated with reduced LINE methylation [33]. HNSCC tumor methylation profiles (based on 1413 CpG loci) are significantly and independently associated with alcohol intake [34], and beyond DNA methylation, miR-375 expression levels in HNSCC have been shown to be significantly increased with increasing alcohol consumption [35]. An excess risk for each alcoholic drink per day of about 10% has been reported in multiple meta-analyses of prospective and case-control studies of breast cancer [36, 37], and a study of breast tumor DNA methylation demonstrated a significant, independent association (controlling for age, dietary folate and other variables) between alcohol intake and tumor DNA methylation profile (based on 1413 CpG loci) [21]. In addition to DNA methylation and miRNA expression alterations, ethanol exposure has been shown to selectively acetylate histone H3 at lysine 9 (H3K9) in primary culture of rat hepatocytes [38] and in rats in vivo [39], though studies in humans are necessary. Collectively, these studies suggest that a major carcinogenic mechanism of action of alcohol is interference with epigenetic regulation, in part through disruption of one-carbon metabolism.

12.2.3 Aging

The aging process itself and differences in environment have been hypothesized to influence clinically significant changes in methylation profiles as individuals accumulate varying exposures with age. Epigenetic variation has been hypothesized to cause underlying differences in disease susceptibility among monozygotic twins, and in one study young twin-pairs seemed more epigenetically similar than older monozygotic twins, though there were a limited number of twin pairs studied [40]. An early report from Issa et al. described an association between aging colonic mucosa and estrogen receptor methylation [41]. Cancer is a disease of aging, and initial studies of age-related methylation in normal tissues were motivated in large part by studies of methylation in cancer [42]. Beyond diseased tissues, recent work has shown an overall trend of increased methylation associated with older age in normal human prostate and colon tissues in several genes [43, 44]. Although an increase in promoter methylation with aging is generally accepted, recent evidence from Bjornsson et al. suggests a more complex picture. These authors measured intra-individual global methylation changes over >10 years and found both increased and decreased methylation levels dependent on the individual, with over 50% of participants exhibiting >5% change in methylation in peripheral blood cell DNA over time [45].

In general, trends of global (repeat element) hypomethylation and promoter hypermethylation found in cancer also have been observed in normal tissues with aging [46]. A study from Agrawal et al. recently linked age-related DNA hypomethylation to chronic inflammation via increased reactivity to self-antigens revealed by the hypomethylation *per se* [47]. This is particularly interesting because it links age-related hypomethylation to immunogenicity and inflammation which has been proposed as a mechanism for aberrant hypermethylation. In recent reports of age-related methylation in normal human prostate and colon tissues, several CpG-island-containing genes were reported to have age-related increases in methylation [43, 44]. A study of over 200 non-pathologic human tissues from ten anatomic sites using the GoldenGate methylation array platform (1413 autosomal CpG loci) demonstrated consistent associations between methylation and age at previously reported candidate genes, and in conjunction with other reports, suggested that the relation between aging and promoter CpG methylation is complex [48]. The direction and strength of correlation between age and methylation were largely dependent upon CpG island status, and specifically a propensity for CpG-island loci to gain methylation with age, and non-island CpGs to lose methylation with age was found, a result consistent with [43, 44, 49], as well as with the findings of Tra et al. and Bjornsson et al. who showed bi-modal age-related methylation in normal tissues. Using restriction-landmark genome scanning of over 2000 CpG loci in T lymphocytes comparing newborns, middle age, and elderly people, Tra et al. reported that 29 loci had age-related methylation alterations, with 23 loci displaying increased methylation with age and 6 decreasing with age [50]. In addition Teschendorff et al. have suggested that age-dependent methylation of polycomb-group target genes (genes suppressed in stem cells allowing differentiation) is a hallmark of cancer and is independent of gender, tissue type, or disease state, suggesting a mechanism for aging to predispose to carcinogenesis [51]. The observed pattern of age associated methylation in [48] was also irrespective of tissue-type, suggesting a common mechanism or dysregulation to explain these alterations. Reduced fidelity of maintenance methyltransferases with aging is one potential explanation for age related decreases in methylation; while age related increases in methylation could potentially reflect the accumulation of stochastic methylation events over time. Additional details on aging and changes in epigenetic profiles are available in Chap. 10.

12.2.4 Tobacco

Epigenetic alterations, particularly DNA methylation, have been thoroughly examined in human lung cancer [52] and have been recently reviewed in regards to its contribution to carcinogenesis in rodent lung cancer models and human disease [53]. Of course, the major etiological contributor to lung cancer is tobacco smoking, and this exposure also plays an important carcinogenic role in a number of other solid tumors, such as oral and bladder cancers [54]. The *CDKN2A* gene (an extensively studied tumor suppressor gene encoding p16INK4A), which is involved in

the Rb cell cycle control checkpoint pathway, is one of the most widely described targets of methylation-induced silencing in cancers, and is thought to play an early and important role in tumor development [55]. Methylation silencing of this gene occurs more often in squamous cell lung cancer, and has been associated, in these tumors, with an increasing duration of tobacco smoking [56, 57]. Even in adenocarcinomas of the lung, methylation of the *CDKN2A* promoter occurs more often in smokers [58] and has been associated with a higher degree of background anthracosis [59]. Methylation of the *RASSF1A* gene has been associated with an earlier age at starting smoking in lung cancer patients [60, 61], suggesting that still growing adolescent lungs may be particularly susceptible to tobacco-related epigenetic events [60, 62]. Interestingly, in mouse models of lung adenocarcinoma induced by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) exposure, tumors exhibit *Cdkn2a* deletion, but also exhibit hypermethylation of *Dapk* and *Rarb* [63–66]. NNK exposure has also been linked to *Cdkn2a* methylation in rat liver adenomas and hepatocellular carcinomas [67].

In oral cancer, methylation of *CDKN2A*, *DAPK*, *RASSF1A*, or *CDH1* (encoding E-cadherin) was significantly associated with smoking, with *CDKN2A* methylation occurring more often in patients who started smoking at a younger age, and *CDH1* methylation occurring in patients with higher pack-years smoked [68]. More recently, profiling of over 750 cancer-related genes with a bead-array approach in 68 tumors of the head and neck denoted that methylation profiles of these tumors were significantly associated with smoking intensity (packs per day) though not significantly associated with smoking duration or cumulative exposure (packyears) [34]. Additionally, in a study of colorectal cancer (Iowa Women's Health Study), smoking variables including age at initiation, duration, and packyears were associated with CpG island methylator phenotype positive tumors [69].

In bladder cancer, which is also associated with tobacco smoke exposures, although to a lesser extent than lung cancer, associations have been observed between smoking and *CDKN2A* methylation, with increased risk of this alteration in current and former smokers, particularly former smokers of less than 10 years [70]. Also in bladder cancer, methylation of any of the *SFRP* genes was associated with having a smoking history [71]. Collectively, these reports suggest that in bladder cancer, less important than intensity or duration is simply the continuous exposure of target cells to tobacco smoke-derived carcinogens, which selects for these alterations.

Building on the mechanistic link between tobacco exposure mediating its cancer risk through epigenetic alterations are a number of studies demonstrating a relation between tobacco exposure and promoter hypermethylation detected in DNA from sputum samples and pre-malignant lung epithelium from heavy smokers [72–78] as well as in non-malignant oral epithelium of smokeless tobacco users [79]. Beyond DNA methylation alterations, recent evidence has suggested that maternal cigarette smoking during pregnancy is associated with downregulation of multiple miRNAs in placental tissues [80].

It should be noted, though, that although tobacco exposure has been associated with specific profiles of tumor methylation including a methylator phenotype, it

does not necessarily lead to global hypermethylation of tumor suppressor genes, suggesting instead a targeted mechanism for the carcinogen's role in epigenetic alteration. One example is *DAPK* methylation which is associated with advanced stage lung cancer, was not associated with tobacco smoke exposure [81]. Promoter methylation of *MGMT* occurs more commonly amongst never-smokers than ever smokers in lung adenocarcinomas [82], and *FANCF* methylation occurs more often in patients with a shorter duration of tobacco use [83]. Methylation of the *ESR1* gene (encoding estrogen receptor-alpha) also occurs at a lower prevalence in NNK-induced lung tumors in mice and rats and in human lung cancer in smokers compared to non-smokers [84]. Together, these results suggest that for some genes, susceptibilities or exposures other than tobacco-smoke may be important in the targeting of these loci for epigenetic alteration.

12.3 Environmental Toxicants and Epigenetic Alterations

12.3.1 Endocrine Disruptors

Exposures to endocrine disrupting chemicals such as diethylstilbestrol (DES) and bisphenol A (BPA) are of particular concern in the context of development. DES is a non-genotoxic carcinogen with developmental stage-specificity, that in the past was administered to women during pregnancy to prevent miscarriages [85]. In utero exposure of mice to DES has been shown to result in the hypermethylation of the developmentally critical (specifically to uterine organogenesis) *Hoxa10* gene [86]. Epidemiologic evidence from individuals exposed to DES during the first 3 months in utero indicates an increase in vaginal clear cell carcinoma incidence and reproductive disorders [87]. In addition, grandchildren of DES exposed women reported higher incidences of rare reproductive disorders; whether this reflects detection bias or possibly implicates a role for epigenetic transgenerational inheritance remains to be clarified [87]. Though the mechanism through which DES establishes altered epigenetic marks capable of transgenerational inheritance remains unclear, a model for the epigenetic effects of DES has been proposed by Ruden et al. drawing similarities between DES and Hsp90, which acts to play a role in modifying H3K4 methylation by increasing the activity of the H3K4 methyltransferase *SMYD3*, thereby altering epigenetic control of various genes [88].

BPA is a monomer used in the production of polycarbonate plastic, a flame retardant, a fungicide and a surface-coating for everyday objects [89], and human exposure to BPA has been increasing in recent years [90]. BPA has been shown to readily cross the placental barrier and accumulate both in the placenta and the fetus [91]. Animal studies have demonstrated that developmental exposure to BPA can alter epigenetic profiles. Briefly, Dolinoy et al. showed that in utero BPA exposure

decreases CpG methylation in agouti mice and that methyl-donor supplementation negated BPA related hypomethylation [92]. In human placental cell lines BPA exposure has been shown to alter miRNA expression levels [93]. More specifically, miR-146a was strongly induced by BPA treatment and resulted in both slower proliferation rate and higher sensitivity to the DNA damaging agent bleomycin [93]. BPA, like DES, has also been shown to alter the methylation status of the *Hoxa10* gene in rodent in-utero exposure models [94], and breast cancer cell lines exposed to BPA showed altered DNA methylation of specific gene promoters [95]. Although awareness of BPA exposure has increased dramatically and some manufacturers have stopped using it, additional study of BPA (as well as emerging substitutes) is necessary to better understand its potential epigenetic effects.

12.3.2 Benzene

Benzene is a ubiquitous chemical, has been shown to be toxic to blood-forming systems, is an established carcinogen [96], and is known to be a cause of acute myeloid leukemia (AML) and myelodysplastic syndromes [97]. Exposed to low dose benzene, gasoline attendants and traffic officers have demonstrated significant leukocyte LINE-1 and Alu DNA hypomethylation as well as gene specific hypermethylation of *CDKN2B* and hypomethylation of *MAGE1* [98]. One preliminary study of benzene exposure using buffy coat DNA from six exposed workers and four unexposed workers has shown some evidence for relatively small changes in methylation [99]. Another preliminary study analyzed miRNA expression in seven exposed-control matched pairs and identified upregulated miR-154*, miR-487a, miR-493-3p and miR-668 in exposed subjects [99]. Additional study of the potential epigenetic effects of benzene exposure is needed.

12.3.3 Radiation

Exposures to ionizing radiation have long been linked to cancer, through DNA damage in the form of large deletions and in some cases point mutation [100–102]. On the other hand, silencing of *Cdkn2a* was detected in lung tumors of rats induced by exposure to ^{239}plutonium [103] and in human lung adenocarcinoma, *CDKN2A* methylation occurred more often in workers compared to non-worker controls of the Russian MAYAK weapons-grade plutonium plant, and the prevalence of this methylation exhibited a dose–response with radiation internal exposure dose [104]. Methylation of *CDKN2A* has also been linked to reactive oxygen species produced by radiation exposures [105]. Murine models of radiation-induced lymphoma have also demonstrated hypermethylation of *Cdkn2b* (encoding p15ink4b [106–108]).

12.3.4 Arsenic

Arsenic exposure has been associated with DNA methylation alterations in non-pathologic as well as tumor tissues and there is some data suggesting that changes in miRNA expression and histone tail modifications are also associated with exposure to arsenic. Although the mode(s) of arsenic's carcinogenicity is not completely clear, there is some speculation about arsenite-generated free radicals and reactive oxygen species lead to genotoxic damage [109]. However, in vitro exposures to inorganic arsenic species have demonstrated dose-dependent increases in promoter region hypermethylation of CpG sites, although not to those resulting in altered gene expression, as well as to the occurrence of genome wide hypomethylation [110–112]. In human bladder cancer, relatively low levels of inorganic arsenic exposure has been associated with methylation of *RASSF1A* and *PRSS3* (but not *CDKN2A*) [70], and in mouse models of methyl- or folate-deficient diets, arsenic exposure through water supply led to hypomethylation in hepatic-derived DNA [113], as well as to increases in chromosomal aberrations in blood lymphocytes [114]. Thus, perhaps a more plausible mechanism for arsenic-related carcinogenicity is via the depletion of S-adenosyl methionine (SAM, the universal methyl donor) due to the metabolism of inorganic arsenic to its methylated forms resulting in altered DNA methylation.

As more studies are conducted a complex picture of dose-dependent DNA methylation alterations with arsenic exposure is beginning to emerge. Chanda et al. have shown that arsenic exposure can result in *CDKN2A* hypermethylation in human blood DNA, but that a subgroup of cases had hypomethylation with high arsenic exposure [115]. More recently, a group of some of the same authors has better quantified the relation showing that exposures to 250-500ug/L of arsenic in drinking water results in global hypermethylation (^{3}H methyl group uptake), but that >500ug/L of arsenic results in global hypomethylation [116]. In another study of peripheral blood DNA methylation by ^{3}H methyl group uptake, Pilsner et al. showed increased DNA methylation to be associated with urinary and plasma arsenic and plasma folate, and that the association between arsenic and methylation was modified by folate in that it was restricted to individuals with high plasma folate [117]. This same group has shown that folic acid supplementation lowers blood arsenic [118], maternal and cord blood pairs have highly correlated arsenic levels [119], and folate deficiency, hyperhomocysteinemia, and leukocyte hypomethylation are associated with arsenic-induced skin lesions [120]. Beyond interfering with one-carbon metabolism and affecting DNA methylation, treatment of human lymphoblastoid cells with sodium arsenite led to global increases in miRNA expression [121]. The genotoxic effects of arsenic may also be mediated by altered chromatin and Ramirez et al. have shown in human hepatocarcinoma cells that treatment with sodium arsenite resulted in global increases in histone acetylation [122]. Thus, it is reasonable to posit that increased expression of miRNAs associated with arsenic exposure may be due to increased histone acetylation. Nonetheless, additional studies are necessary to further elucidate the associations between arsenic and all major forms of epigenetic alteration with particular attention being given to dose and modification by one-carbon metabolism pathway participants.

12.3.5 Other Metals (Nickel, Lead, Cadmium, Chromium)

Although there is currently very little evidence for epigenetic alterations due to exposure to other metals in humans, several studies have suggested that exposures to metals such as nickel, lead, cadmium and others may alter epigenetic marks. For example, nickel exposure has been associated with various histone modifications and exposure of mammalian cells to nickel resulted in decreased histone H4 acetylation [123, 124]. Although it is an established carcinogen, cadmium is a weak mutagen. However, exposure of a rat cell line to cadmium resulted in a concentration-dependent reduction in DNA methyltransferase activity [125]. A study of metal-rich (lead, cadmium, chromium) particulate matter exposure among electric-furnace steel plant workers has shown significant alterations of micro-RNA expression in peripheral blood leukocytes [126]. In the Normative Aging Study, patella lead levels were associated with reduced global DNA methylation (LINE-1 elements) though not Alu repeat regions [127], suggesting a role for lead exposure in epigenetic alterations. Interestingly, a sexually dimorphic global methylation (Luminometric methylation assay for CCGG) level decrease was associated with brain mercury levels in male polar bears [128].

Chromium has industrial applications including chrome plating and stainless steel welding and is implicated in occupationally related lung cancers [129]. Mechanisms of chromium carcinogenicity include adduct formation [130] and DNA strand breakage [131], and some limited evidence suggests the potential for chromium to have epigenetic effects. In cases of chromate-induced lung cancer Kondo et al. showed that although prevalence of *CDKN2A* methylation was similar between chromate and non-chromate lung cancers, chromate lung cancers were more likely to have methylated *CDKN2A* with increasing duration of occupational exposure [132]. An *in vitro* exposure of mouse hepatoma cells to chromium showed transcriptional repression of *Cyp1a1* by local cross-linking of Hdac and Dnmt1 and altered histone marks [133]. However, additional study of this diminishing exposure is necessary to validate potential chromium-induced epigenetic alterations.

12.3.6 Particulates/Air Pollution

Environmental particulate exposure has been linked to increased risk of lung cancer [134] and increased morbidity and mortality from cardiovascular and respiratory illnesses [135], though few studies have addressed the potential for these exposures to alter DNA methylation. In rat lung cancer models of tumors induced by diesel exhaust, 59% were methylated at *Cdkn2a* and 14% at *Esr1*, in those induced by carbon black 45% were methylated at *Cdkn2a* and 14% at *Esr1*, and in those induced by beryllium metal strikingly 80% were methylated at *Cdkn2a* and 50% at *Esr1* [136]. Extending to miRNA alterations, a study using primary bronchial epithelial cells from nonsmokers has shown dramatic alterations in miRNA expression

following treatment with diesel exhaust particles (197 of 313 measured were up or down-regulated more than 1.5 fold), and suggested that the alterations were associated with inflammatory response pathways [137]. In humans, exposure to particulate matter of <10 μm (PM_{10}) in steel plant workers was associated with significantly lower peripheral blood *NOS2* promoter methylation [138]. In addition, this same study reported significantly reduced methylation of LINE-1 and Alu repeat elements in blood DNA associated with long-term PM_{10} exposure [138]. Consistent with this, a recent report examining the affect of ambient particulate pollutants on repeat element methylation in subjects from the Boston area Normative Aging Study found significantly decreased LINE-1 methylation following recent exposure to higher black carbon [139].

Ren et al. have recently provided some insight into a potential mechanism of air pollution associated epigenetic alterations, and evidence for genetic interactions that modify effects of exposure [140]. These authors found that interquartile range increases in $PM_{2.5}$ and black carbon were associated with increases in plasma homocysteine (part of one-carbon metabolism), and glutathione transferase theta genotype (an antioxidant dependent upon homocysteine levels) [140]. This same group of authors also reported that increases in mean air pollution PM_{10} concentrations did not significantly alter fasting or postmethionine-load total homocystiene in non-smokers, but was associated with significantly increased homocysteine levels in smokers, suggesting interactions between exposures may contribute to epigenetic alterations [141]. Another line of evidence for modification of exposure effects on epigenetics by genetics comes from Wilker et al. who have reported that increases in black carbon associated with increased blood pressure are modified by SNPs in microRNA processing machinery genes such as *DICER* and *GEMIN* genes [142].

12.3.7 Asbestos

Exposure to asbestos is the main risk factor for malignant pleural mesothelioma with approximately 80% of cases reporting a known exposure to asbestos [143]. Importantly, in contrast to tobacco smoke and radiation, asbestos is known to be a weak mutagen [144], and there have been several reports of tumor suppressor gene methylation in mesothelioma [145–148]. In patients with mesothelioma, Tsou et al. described a significant association between self-reported asbestos exposure and methylation at the *MT1A*, and *MT2A* gene loci among 28 genes examined [149]. A significant association has been reported between an increasing number of methylated cell cycle control genes and a quantitative asbestos burden measure [150]. This was followed by examination using an array-based approach, which demonstrated that quantitative measure of asbestos exposure was associated with altered methylation at over 100 discrete CpG loci, and that in almost all cases (94%) there was increased methylation associated with increased exposure [151]. Further, overall methylation profiles for mesotheliomas were significantly associated with asbestos exposure burden [151].

12.4 Environmental Exposures and Imprinting and Development

Genomic imprinting involves epigenetic silencing of a single allele distinguishing the maternal or paternal inheritance pattern of a specific locus [152]. Imprinting and its importance in development as well as alterations to imprinting and its effect on disease, including cancer, have been thoroughly reviewed by Feinberg and colleagues [152–154]. Here, we would like to point out some interesting work that suggests that there are environmental exposures that can alter imprinting. The allele-specific expression pattern associated with imprinting is based upon CpG island methylation of differential methylation regions (DMR) which occur in the upstream regions of a number of genes and act as both transcriptional repressive regions as well as insulators to neighboring repressive elements [155, 156]. Alterations to the imprinted status of DMRs has been observed after prolonged culture of mouse embryonic stem cells, leading to a number of phenotypic abnormalities, through both gains and losses of appropriate methylation of these regions [157]. These alterations to the imprinting pattern are believed to be related to exposure to particular media or serum [158, 159]. In humans, it has recently been suggested that assisted reproductive technology (ART) may affect the epigenetics of development thus altering imprinting status. In sheep and cattle, large offspring syndrome has been linked to culture of pre-implantation embryos [160] and in humans, associations between Angelman syndrome [161], and Beckwith-Wiedemann [162] syndrome and in-vitro fertilization have been reported, and are related to changes in imprinting status of specific genes. The exact exposures encountered by these pre-implantation embryos and leading to these epigenetic alterations are unclear. The relation between environmental exposures and imprinting alterations has been examined in a cross-sectional study of patients referred to an academic gastrointestinal practice, and although no associations were observed with exposure to tobacco or alcohol nor with measure dietary factors and NSAID drug use, the authors themselves acknowledge that the study was designed such that it could detect only quite large effects of these exposures on any loss of imprinting [163]. The animal as well as human studies point to the need for additional work to characterize whether alterations in imprinting occur somatically and if these changes are related to environmental exposures, as well as to determine if and how in-utero and neo-natal exposures may affect the epigenetic marks at these imprinted loci. More on epigenome changes during development can be found in Chap. 5, details of imprinting disorders are available in Chap. 8, and assisted reproductive technologies are covered in Chap. 7.

12.5 Mechanisms

Ascribing a mode by which environmental exposures drive or select epigenetic alteration is difficult due in part to the lack of understanding of the mechanism by which *de novo* methylation occurs in disease. In fact, dependent upon context and

exposure, different mechanisms can contribute to epigenetic alterations. Although it is beyond our scope to consider specific contextual details such as relative location (repeat, promoter, other), existing methylation state, nucleosome position, local and regional CpG density and transcription factor binding sites; it is important to recognize that contextual elements may differentially affect the likelihood of exposure-related alterations. For example, a recent report from Rakyan et al. showed that aging-associated hypermethylation was more likely to occur at promoters associated with bivalent chromatin domains [164]. Additionally, distinct mechanisms may be responsible for altered repeat element (global) methylation and altered promoter methylation. Although traditional dogma indicates separate methyltransferase enzymes for maintenance (*DNMT1*) and de novo methylation (*DNMT3A/B*), a recent perspective from Jones and Liang proposes that maintenance methylation in CpG islands may be a cooperative process [165]. When considering the mechanisms of environmental exposure related methylation alterations it is important to remain conscious of the fact that we do not have a complete understanding of *de novo* and maintenance methylation mechanisms. Further, similar to histone modifications, environmental and intrinsic signaling mechanisms that alter miRNA expression levels are poorly understood. Nonetheless, several mechanisms for environmental exposure related epigenetic alterations should be considered.

One mechanism by which an exposure may result in altered DNA methylation is increased reactive oxygen species as a result of inflammatory response. More specifically, it has been reported that 5-hydroxymethylcytosine can be generated by oxidation of 5-methylcytosine [166], and both 5-methylcytosine adjacent to 8-oxoguanine, and 5-hydroxymethylcytosine have been shown to inhibit binding of methyl-CpG binding protein 2, a critical epigenetic regulator that recruits cytosine methyltransferases and histone deacetylases [167]. It is also known that 5-hydroxymethylcytosine is not recognized as 5-methylcytosine by the maintenance methyltransferase DNMT1, and hence, may lead to aberrant loss of methylation during cell replication [168]. Additional base alterations occur via neutrophil and eosinophil peroxidase-derived HOCl and HOBr which can react with DNA to form 5-chlorocytosine and 5-bromocytosine respectively [169]. These halogenated cytosines can be mistaken by DNMT1 as 5-methylcytosine during replication, thus providing a potential mechanism for inflammation-induced aberrant hypermethylation [168].

Another mechanism that has been hypothesized is a spreading process, whereby CpG methylation from within the gene migrates into the promoter due to the loss of protective boundary elements which normally protect promoters from CpG island methylation [170–172]. Alternatively, for example in aging, a spreading process subsequent to a stochastic methylation alteration (in pathologically normal tissue) without functional consequences for gene expression, may allow the initiation of a spreading process that eventually confers an altered expression phenotype. Such age-related drift of normal epigenomes (without prominent changes in gene expression) may confer significantly increased risk of conversion to a pathologic phenotype by enhancing both the likelihood and frequency of methylation events that ultimately result in altered expression or genomic stability. For example, in the

context of acquired "non-functional" CpG methylation in the promoter region of an aged individual, continued stochastic methylation events (e.g. "methylation spreading") increase the chance of methylation induced silencing at that promoter (or silencing of another locus through action at a distance via silencing of other important regions such as enhancers), and hence, progression to a pathologic phenotype. Certainly, this hypothesis is especially plausible for the many diseases of aging. Alternatively, aberrant CpG methylation that silences a gene on a single allele may not appear to have a functional consequence if the complementary allele can provide compensatory expression. As a result, for example, clusters of cellular clones with mono-allelic gene expression could contribute to an increased risk of progression to a pathologic phenotype (e.g. loss of the 2nd functional allele). Future population-based studies addressing the potential of quantifying age and/or exposure associated methylation alterations indicative of disease risk are necessary.

It also remains unclear if changes to chromatin, such as histone H3 Lys 9 methylation occurs prior to or because of CpG island methylation in the promoter, although in tumors these alterations are found together [173–175]. The totality of epigenetic marks (DNA and chromatin) may also be important, especially as the alterations are selected for throughout the carcinogenic process. Several questions remain: is there selective pressure for only some CpG sites to become methylated or must an entire region undergo this alteration? Is the selection based on a binary event, expressing or not expressing the target gene, or may there be reduced expression states important for some periods of disease development? How does this alteration interact with genetic alterations, such as deletion or mutation, in driving carcinogenesis, and is the timing and/or order of the alterations important? These questions not only highlight the need for additional biochemical characterization of the process of *de novo* epigenetic silencing but also suggest how an understanding of the impact of carcinogenic exposure on these alterations may play an important, albeit complex role.

For a number of the exposures discussed above, much is known about their role in the creation of genotoxic damage and thus mutation. For example, tobacco smoke is known to lead to the production of bulky DNA adducts, which are repaired via base excision repair pathways [176–178], but consistent reports have linked tobacco smoke exposure to increasing prevalences of methylation of specific genes, for example *CDKN2A* [56, 57]. These results suggest that there may be a link between DNA damage or DNA repair pathways and methylation in at least this one instance for this one gene locus. These DNA damage events may lead to alterations in the ability of proteins, such as transcription factors or insulating elements, from binding DNA, and may therefore trigger epigenetic alterations at these sites [174, 179]. In addition, genetic alterations in epigenetic regulatory genes may lead to widespread epigenetic alterations. Evidence for this has been shown in mesotheliomas where tumors with deletion of the DNA maintenance methyltransferase gene *DNMT1* had significantly reduced methylation (array-wide) compared to tumors without *DNMT1* deletion [180].

Genomic hypomethylation, in addition to specific gene hypermethylation, is also an important form of epigenetic alteration and a common feature of human tumors, although it is unclear if there is any interplay between these alterations [181].

DNA hypomethylation may lead to the expression of normally silenced transposons, leading to an increase in transposition events and increases in genomic instability [182]. In response, hypermethylation may be occurring as a defense for retrotransposition and the resulting instability, as methylation is thought to suppress homologous recombination [183–185]. An overall decrease in cellular SAM levels through poor dietary folate intake or through alternative use of SAM for carcinogen metabolism (such as that seen with arsenic) may lead to hypomethylation [18, 186–189]. This hypomethylation may also allow activation of normally silenced non-coding RNA sequences, which, as we have discussed previously, may play an important role in the epigenetic programming of specific genes.

Common to all carcinogenic exposures, both mutagenic and non-mutagenic compounds, is their ability to impart selective pressures on cellular clones. Some researchers believe that it is this effect, and not genotoxicity, that provides the carcinogenicity attributable to some exposures, by selecting pre-existing genetic or possibly epigenetic alterations [189]. If *de novo* methylation in the developing tumor occurred stochastically, cells with a silenced tumor suppressor may be selected for growth or expansion due to exposure to specific carcinogens. In time these pressures will allow then for clonal expansion of the epigenetically altered cells into fields which would have an increased susceptibility for tumorigenesis [62, 190].

There is, as we have presented, a growing body of literature beginning to link epigenetic alterations to carcinogen exposures from in vitro work, animal models, and human observational studies. It is also clear that the modes by which exposures drive the clonal selection of genetically or epigenetically altered cells or modulate the epigenetic landscape of individual cells are yet to be elucidated and require further exploration.

12.6 Development of Environmental Epigenomics

The fields of environmental genomics and toxicogenomics are bringing the wealth of knowledge gained through the human genome project to the fields of environmental health and toxicology. This will stimulate novel approaches to understanding the modes of action of various chemical and physical agents. It has also been suggested that collaborative efforts between the various disciplines of environmental science, genomics, systems biology, medicine, and basic biology may hold great promise to elucidate more thoroughly the effect of exposures on human health [191]. In addition to studying the effects at the genotoxic and transcriptional levels, it is important to also consider the role of carcinogenic, teratogenic and developmentally toxic exposures on the epigenome, and on clonal selection, again requiring multi-disciplinary work from in-vitro and model systems to epidemiologic studies of human disease. New technologies now allow for global analysis of epigenetic alterations and these may provide insight into the extent and patterns of alteration in normal and diseased tissue. Notably, the latest generation of CpG methylation arrays interrogates hundreds of thousands of CpG loci, includes all designable RefSeq

genes, and is low-cost enough for application in population-level studies. Next generation sequencing approaches, as their cost decreases and throughput increases will also become increasingly used as tools for annotating the epigenome and its relation to exposures. Mammalian stem cells may also prove extremely useful in bettering the understanding of epigenetics in development, providing insightful models for demonstrating how these pathways become dysregulated in cancer. Finally, appropriate epidemiologic studies, including both case-series and prospective designs, requiring defined and consistent methodologies and data collection will be crucial to understanding the true effect of environmental exposures on the human epigenome. This work is urgently needed in order to better understand the biology of epigenetic alterations and the effects of toxic exposures on these disease-associated somatic alterations. Further, this will lead to better prediction of the toxic potential of new compounds introduced into the environment and allow for more targeted, and appropriate, disease prevention strategies.

References

1. Morgan HD, Santos F, Green K, Dean W, Reik W (2005) Epigenetic reprogramming in mammals. Hum Mol Genet 14:R47–58
2. Li E (2002) Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet 3:662–673
3. Gluckman PD, Hanson MA, Cooper C, Thornburg KL (2008) Effect of in utero and early-life conditions on adult health and disease. N Engl J Med 359:61–73
4. Daly LE, Kirke PN, Molloy A, Weir DG, Scott JM (1995) Folate levels and neural tube defects. Implications for prevention. JAMA 274:1698–1702
5. Dunlevy LP, Burren KA, Mills K, Chitty LS, Copp AJ, Greene ND (2006) Integrity of the methylation cycle is essential for mammalian neural tube closure. Birth Defects Res A Clin Mol Teratol 76:544–552
6. Wang L, Wang F, Guan J, Le J, Wu L, Zou J et al (2010) Relation between hypomethylation of long interspersed nucleotide elements and risk of neural tube defects. Am J Clin Nutr 91:1359–1367
7. Waterland RA, Michels KB (2007) Epigenetic epidemiology of the developmental origins hypothesis. Annu Rev Nutr 27:363–388
8. Gluckman PD, Hanson MA (2004) Developmental origins of disease paradigm: a mechanistic and evolutionary perspective. Pediatr Res 56:311–317
9. Johansson M, Relton C, Ueland PM, Vollset SE, Midttun O, Nygard O et al (2010) Serum B vitamin levels and risk of lung cancer. JAMA 303:2377–2385
10. Freudenheim JL, Marshall JR, Vena JE, Laughlin R, Brasure JR, Swanson MK et al (1996) Premenopausal breast cancer risk and intake of vegetables, fruits, and related nutrients. J Natl Cancer Inst 88:340–348
11. Graham S, Hellmann R, Marshall J, Freudenheim J, Vena J, Swanson M et al (1991) Nutritional epidemiology of postmenopausal breast cancer in western New York. Am J Epidemiol 134:552–566
12. Martinez ME, Marshall JR, Giovannucci E (2008) Diet and cancer prevention: the roles of observation and experimentation. Nat Rev Cancer 8:694–703
13. Kim DH, Smith-Warner SA, Spiegelman D, Yaun SS, Colditz GA, Freudenheim JL et al (2010) Pooled analyses of 13 prospective cohort studies on folate intake and colon cancer. Cancer Causes Control 21:1919–1930

14. de Vogel S, Wouters KA, Gottschalk RW, van Schooten FJ, de Goeij AF, de Bruine AP et al (2011) Dietary methyl donors, methyl metabolizing enzymes, and epigenetic regulators: diet-gene interactions and promoter CpG island hypermethylation in colorectal cancer. Cancer Causes Control 22:1–12
15. van Driel LM, Eijkemans MJ, de Jonge R, de Vries JH, van Meurs JB, Steegers EA et al (2009) Body mass index is an important determinant of methylation biomarkers in women of reproductive ages. J Nutr 139:2315–2321
16. Sauer J, Mason JB, Choi SW (2009) Too much folate: a risk factor for cancer and cardiovascular disease? Curr Opin Clin Nutr Metab Care 12:30–36
17. Ingrosso D, Cimmino A, Perna AF, Masella L, De Santo NG, De Bonis ML et al (2003) Folate treatment and unbalanced methylation and changes of allelic expression induced by hyperhomocysteinaemia in patients with uraemia. Lancet 361:1693–1699
18. Pufulete M, Al-Ghnaniem R, Leather AJ, Appleby P, Gout S, Terry C et al (2003) Folate status, genomic DNA hypomethylation, and risk of colorectal adenoma and cancer: a case control study. Gastroenterology 124:1240–1248
19. Pufulete M, Al-Ghnaniem R, Rennie JA, Appleby P, Harris N, Gout S et al (2005) Influence of folate status on genomic DNA methylation in colonic mucosa of subjects without colorectal adenoma or cancer. Br J Cancer 92:838–842
20. Hsiung DT, Marsit CJ, Houseman EA, Eddy K, Furniss CS, McClean MD et al (2007) Global DNA methylation level in whole blood as a biomarker in head and neck squamous cell carcinoma. Cancer Epidemiol Biomarkers Prev 16:108–114
21. Christensen BC, Kelsey KT, Zheng S, Houseman EA, Marsit CJ, Wrensch MR et al (2010) Breast cancer DNA methylation profiles are associated with tumor size and alcohol and folate intake. PLoS Genet 6:e1001043
22. Gluckman PD, Hanson MA, Buklijas T, Low FM, Beedle AS (2009) Epigenetic mechanisms that underpin metabolic and cardiovascular diseases. Nat Rev Endocrinol 5:401–408
23. Kim M, Long TI, Arakawa K, Wang R, Yu MC, Laird PW (2010) DNA methylation as a biomarker for cardiovascular disease risk. PLoS One 5:e9692
24. Papait R, Condorelli G (2010) Epigenetics in heart failure. Ann N Y Acad Sci 1188:159–164
25. Hathcock JN (1997) Vitamins and minerals: efficacy and safety. Am J Clin Nutr 66:427–437
26. Tchantchou F, Graves M, Falcone D, Shea TB (2008) S-adenosylmethionine mediates glutathione efficacy by increasing glutathione S-transferase activity: implications for S-adenosyl methionine as a neuroprotective dietary supplement. J Alzheimers Dis 14:323–328
27. Hillman RS, Steinberg SE (1982) The effects of alcohol on folate metabolism. Annu Rev Med 33:345–354
28. Dumitrescu RG, Shields PG (2005) The etiology of alcohol-induced breast cancer. Alcohol 35:213–225
29. Kharbanda KK (2009) Alcoholic liver disease and methionine metabolism. Semin Liver Dis 29:155–165
30. International Agency for Research on Cancer (1988) Alcohol drinking, vol 44. IARC, Lyon
31. Giovannucci E, Rimm EB, Ascherio A, Stampfer MJ, Colditz GA, Willett WC (1995) Alcohol, low-methionine–low-folate diets, and risk of colon cancer in men. J Natl Cancer Inst 87:265–273
32. Schernhammer ES, Giovannucci E, Kawasaki T, Rosner B, Fuchs CS, Ogino S (2010) Dietary folate, alcohol and B vitamins in relation to LINE-1 hypomethylation in colon cancer. Gut 59:794–799
33. Smith IM, Mydlarz WK, Mithani SK, Califano JA (2007) DNA global hypomethylation in squamous cell head and neck cancer associated with smoking, alcohol consumption and stage. Int J Cancer 121:1724–1728
34. Marsit CJ, Christensen BC, Houseman EA, Karagas MR, Wrensch MR, Yeh RF et al (2009) Epigenetic profiling reveals etiologically distinct patterns of DNA methylation in head and neck squamous cell carcinoma. Carcinogenesis 30:416–422

35. Avissar M, McClean MD, Kelsey KT, Marsit CJ (2009) MicroRNA expression in head and neck cancer associates with alcohol consumption and survival. Carcinogenesis 30:2059–2063
36. Hamajima N, Hirose K, Tajima K, Rohan T, Calle EE, Heath CW Jr et al (2002) Alcohol, tobacco and breast cancer–collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. Br J Cancer 87:1234–1245
37. Key J, Hodgson S, Omar RZ, Jensen TK, Thompson SG, Boobis AR et al (2006) Meta-analysis of studies of alcohol and breast cancer with consideration of the methodological issues. Cancer Causes Control 17:759–770
38. Park PH, Miller R, Shukla SD (2003) Acetylation of histone H3 at lysine 9 by ethanol in rat hepatocytes. Biochem Biophys Res Commun 306:501–504
39. Kim JS, Shukla SD (2006) Acute in vivo effect of ethanol (binge drinking) on histone H3 modifications in rat tissues. Alcohol Alcohol 41:126–132
40. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML et al (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 102:10604–10609
41. Issa JP, Ottaviano YL, Celano P, Hamilton SR, Davidson NE, Baylin SB (1994) Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. Nat Genet 7:536–540
42. Richardson B (2003) Impact of aging on DNA methylation. Ageing Res Rev 2:245–261
43. Kwabi-Addo B, Chung W, Shen L, Ittmann M, Wheeler T, Jelinek J et al (2007) Age-related DNA methylation changes in normal human prostate tissues. Clin Cancer Res 13:3796–3802
44. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J et al (2005) MGMT promoter methylation and field defect in sporadic colorectal cancer. J Natl Cancer Inst 97:1330–1338
45. Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H et al (2008) Intra-individual change over time in DNA methylation with familial clustering. Jama 299:2877–2883
46. Fraga MF, Agrelo R, Esteller M (2007) Cross-talk between aging and cancer: the epigenetic language. Ann N Y Acad Sci 1100:60–74
47. Agrawal A, Tay J, Yang GE, Agrawal S, Gupta S (2010) Age-associated epigenetic modifications in human DNA increase its immunogenicity. Aging (Albany NY) 2:93–100
48. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL et al (2009) Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet 5:e1000602
49. Issa JP (2003) Age-related epigenetic changes and the immune system. Clin Immunol 109:103–108
50. Tra J, Kondo T, Lu Q, Kuick R, Hanash S, Richardson B (2002) Infrequent occurrence of age-dependent changes in CpG island methylation as detected by restriction landmark genome scanning. Mech Ageing Dev 123:1487–1503
51. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H et al (2010) Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome Res 20:440–446
52. Minna JD, Roth JA, Gazdar AF (2002) Focus on lung cancer. Cancer Cell 1:49–52
53. Belinsky SA (2005) Silencing of genes by promoter hypermethylation: key event in rodent and human lung cancer. Carcinogenesis 26:1481–1487
54. Services USDoHaH (2004) The health consequences of smoking: a report of the surgeon general. Atlanta
55. Rocco JW, Sidransky D (2001) p16(MTS-1/CDKN2/INK4a) in cancer progression. Exp Cell Res 264:42–55
56. Kim DH, Nelson HH, Wiencke JK, Zheng S, Christiani DC, Wain JC et al (2001) p16(INK4a) and histology-specific methylation of CpG islands by exposure to tobacco smoke in non-small cell lung cancer. Cancer Res 61:3419–3424

57. Toyooka S, Maruyama R, Toyooka KO, McLerran D, Feng Z, Fukuyama Y et al (2003) Smoke exposure, histologic type and geography-related differences in the methylation profiles of non-small cell lung cancer. Int J Cancer 103:153–160
58. Divine KK, Pulling LC, Marron-Terada PG, Liechty KC, Kang T, Schwartz AG et al (2005) Multiplicity of abnormal promoter methylation in lung adenocarcinomas from smokers and never smokers. Int J Cancer 114:400–405
59. Hou M, Morishita Y, Iljima T, Inadome Y, Mase K, Dai Y et al (1999) DNA methylation and expression of p16(INK4A) gene in pulmonary adenocarcinoma and anthracosis in background lung. Int J Cancer 84:609–613
60. Marsit CJ, Kim DH, Liu M, Hinds PW, Wiencke JK, Nelson HH et al (2005) Hypermethylation of RASSF1A and BLU tumor suppressor genes in non-small cell lung cancer: implications for tobacco smoking during adolescence. Int J Cancer 114:219–223
61. Kim DH, Kim JS, Ji YI, Shim YM, Kim H, Han J et al (2003) Hypermethylation of RASSF1A promoter is associated with the age at starting smoking and a poor prognosis in primary non-small cell lung cancer. Cancer Res 63:3743–3746
62. Wiencke JK, Kelsey KT (2002) Teen smoking, field cancerization, and a "critical period" hypothesis for lung cancer susceptibility. Environ Health Perspect 110:555–558
63. Belinsky SA, Swafford DS, Middleton SK, Kennedy CH, Tesfaigzi J (1997) Deletion and differential expression of p16INK4a in mouse lung tumors. Carcinogenesis 18:115–120
64. Herzog CR, Soloff EV, McDoniels AL, Tyson FL, Malkinson AM, Haugen-Strano A et al (1996) Homozygous codeletion and differential decreased expression of p15INK4b, p16INK4a-alpha and p16INK4a-beta in mouse lung tumor cells. Oncogene 13:1885–1891
65. Vuillemenot BR, Pulling LC, Palmisano WA, Hutt JA, Belinsky SA (2004) Carcinogen exposure differentially modulates RAR-beta promoter hypermethylation, an early and frequent event in mouse lung carcinogenesis. Carcinogenesis 25:623–629
66. Pulling LC, Vuillemenot BR, Hutt JA, Devereux TR, Belinsky SA (2004) Aberrant promoter hypermethylation of the death-associated protein kinase gene is early and frequent in murine lung tumors induced by cigarette smoke and tobacco carcinogens. Cancer Res 64:3844–3848
67. Pulling LC, Klinge DM, Belinsky SA (2001) p16INK4a and beta-catenin alterations in rat liver tumors induced by NNK. Carcinogenesis 22:461–466
68. Hasegawa M, Nelson HH, Peters E, Ringstrom E, Posner M, Kelsey KT (2002) Patterns of gene promoter methylation in squamous cell cancer of the head and neck. Oncogene 21:4231–4236
69. Limsui D, Vierkant RA, Tillmans LS, Wang AH, Weisenberger DJ, Laird PW et al (2010) Cigarette smoking and colorectal cancer risk by molecularly defined subtypes. J Natl Cancer Inst 102:1012–1022
70. Marsit CJ, Karagas MR, Danaee H, Liu M, Andrew A, Schned A et al (2006) Carcinogen exposure and gene promoter hypermethylation in bladder cancer. Carcinogenesis 27:112–116
71. Marsit CJ, Karagas MR, Andrew A, Liu M, Danaee H, Schned A, et al (2005) Epigenetic inactivation of SFRP genes and TP53 alteration act jointly as markers of invasive bladder cancer. Cancer Res 65:7081–7085
72. Zochbauer-Muller S, Lam S, Toyooka S, Virmani AK, Toyooka KO, Seidl S et al (2003) Aberrant methylation of multiple genes in the upper aerodigestive tract epithelium of heavy smokers. Int J Cancer 107:612–616
73. Palmisano WA, Divine KK, Saccomanno G, Gilliland FD, Baylin SB, Herman JG et al (2000) Predicting lung cancer by detecting aberrant promoter methylation in sputum. Cancer Res 60:5954–5958
74. Belinsky SA, Nikula KJ, Palmisano WA, Michels R, Saccomanno G, Gabrielson E et al (1998) Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. Proc Natl Acad Sci USA 95:11891–11896
75. Zochbauer-Muller S, Fong KM, Maitra A, Lam S, Geradts J, Ashfaq R et al (2001) 5' CpG island methylation of the FHIT gene is correlated with loss of gene expression in lung and breast cancer. Cancer Res 61:3581–3585

76. von Zeidler SV, Miracca EC, Nagai MA, Birman EG (2004) Hypermethylation of the p16 gene in normal oral mucosa of smokers. Int J Mol Med 14:807–811
77. Kersting M, Friedl C, Kraus A, Behn M, Pankow W, Schuermann M (2000) Differential frequencies of p16(INK4a) promoter hypermethylation, p53 mutation, and K-ras mutation in exfoliative material mark the development of lung cancer in symptomatic chronic smokers. J Clin Oncol 18:3221–3229
78. Sozzi G, Pastorino U, Moiraghi L, Tagliabue E, Pezzella F, Ghirelli C et al (1998) Loss of FHIT function in lung cancer and preinvasive bronchial lesions. Cancer Res 58:5032–5037
79. Kulkarni V, Saranath D (2004) Concurrent hypermethylation of multiple regulatory genes in chewing tobacco associated oral squamous cell carcinomas and adjacent normal tissues. Oral Oncol 40:145–153
80. Maccani MA, Avissar-Whiting M, Banister CE, McGonnigal B, Padbury JF, Marsit CJ (2010) Maternal cigarette smoking during pregnancy is associated with downregulation of miR-16, miR-21 and miR-146a in the placenta. Epigenetics 5:583–589
81. Kim DH, Nelson HH, Wiencke JK, Christiani DC, Wain JC, Mark EJ et al (2001) Promoter methylation of DAP-kinase: association with advanced stage in non-small cell lung cancer. Oncogene 20:1765–1770
82. Pulling LC, Divine KK, Klinge DM, Gilliland FD, Kang T, Schwartz AG et al (2003) Promoter hypermethylation of the O6-methylguanine-DNA methyltransferase gene: more common in lung adenocarcinomas from never-smokers than smokers and associated with tumor progression. Cancer Res 63:4842–4848
83. Marsit CJ, Liu M, Nelson HH, Posner M, Suzuki M, Kelsey KT (2004) Inactivation of the Fanconi anemia/BRCA pathway in lung and oral cancers: implications for treatment and survival. Oncogene 23:1000–1004
84. Issa JP, Baylin SB, Belinsky SA (1996) Methylation of the estrogen receptor CpG island in lung tumors is related to the specific type of carcinogen exposure. Cancer Res 56:3655–3658
85. Herbst AL, Ulfelder H, Poskanzer DC (1971) Adenocarcinoma of the vagina. Association of maternal stilbestrol therapy with tumor appearance in young women. N Engl J Med 284:878–881
86. Bromer JG, Wu J, Zhou Y, Taylor HS (2009) Hypermethylation of homeobox A10 by in utero diethylstilbestrol exposure: an epigenetic mechanism for altered developmental programming. Endocrinology 150:3376–3382
87. Newbold RR (2004) Lessons learned from perinatal exposure to diethylstilbestrol. Toxicol Appl Pharmacol 199:142–150
88. Ruden DM, Xiao L, Garfinkel MD, Lu X (2005) Hsp90 and environmental impacts on epigenetic states: a model for the trans-generational effects of diethylstibesterol on uterine development and cancer. Hum Mol Genet 14:R149–155
89. Vom Saal FS, Hughes C (2005) An extensive new literature concerning low-dose effects of bisphenol A shows the need for a new risk assessment. Environ Health Perspect 113:926–933
90. Kang JH, Kondo F, Katayama Y (2006) Human exposure to bisphenol A. Toxicology 226:79–89
91. Schonfelder G, Wittfoht W, Hopp H, Talsness CE, Paul M, Chahoud I (2002) Parent bisphenol A accumulation in the human maternal-fetal-placental unit. Environ Health Perspect 110:A703–707
92. Dolinoy DC, Huang D, Jirtle RL (2007) Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. Proc Natl Acad Sci USA 104:13056–13061
93. Avissar-Whiting M, Veiga KR, Uhl KM, Maccani MA, Gagne LA, Moen EL et al (2010) Bisphenol A exposure leads to specific microRNA alterations in placental cells. Reprod Toxicol 29:401–406
94. Bromer JG, Zhou Y, Taylor MB, Doherty L, Taylor HS (2010) Bisphenol-A exposure in utero leads to epigenetic alterations in the developmental programming of uterine estrogen response. FASEB J 24:2273–2280

95. Weng YI, Hsu PY, Liyanarachchi S, Liu J, Deatherage DE, Huang YW et al (2010) Epigenetic influences of low-dose bisphenol A in primary human breast epithelial cells. Toxicol Appl Pharmacol 248:111–121
96. International Agency for Research on Cancer (1999) Benzene, vol 29. IARC, Lyon
97. Ye Y, Wang KK, Gu J, Yang H, Lin J, Ajani JA et al (2008) Genetic variations in microRNA-related genes are novel susceptibility loci for esophageal cancer risk. Cancer Prev Res (Phila Pa) 1:460–469
98. Bollati V, Baccarelli A, Hou L, Bonzini M, Fustinoni S, Cavallo D et al (2007) Changes in DNA methylation patterns in subjects exposed to low-dose benzene. Cancer Res 67:876–880
99. Zhang L, McHale CM, Rothman N, Li G, Ji Z, Vermeulen R et al (2010) Systems biology of human benzene exposure. Chem Biol Interact 184:86–93
100. McDonald JW, Taylor JA, Watson MA, Saccomanno G, Devereux TR (1995) p53 and K-ras in radon-associated lung adenocarcinoma. Cancer Epidemiol Biomarkers Prev 4:791–793
101. Prise KM, Pinto M, Newman HC, Michael BD (2001) A review of studies of ionizing radiation-induced double-strand break clustering. Radiat Res 156:572–576
102. Little JB (2000) Radiation carcinogenesis. Carcinogenesis 21:397–404
103. Swafford DS, Middleton SK, Palmisano WA, Nikula KJ, Tesfaigzi J, Baylin SB et al (1997) Frequent aberrant methylation of p16INK4a in primary rat lung tumors. Mol Cell Biol 17:1366–1374
104. Belinsky SA, Klinge DM, Liechty KC, March TH, Kang T, Gilliland FD et al (2004) Plutonium targets the p16 gene for inactivation by promoter hypermethylation in human lung adenocarcinoma. Carcinogenesis 25:1063–1067
105. Romanenko A, Morell-Quadreny L, Lopez-Guerrero JA, Pellin A, Nepomnyaschy V, Vozianov A et al (2002) P16INK4A and p15INK4B gene alteration associated with oxidative stress in renal cell carcinomas after the chernobyl accident (pilot study). Diagn Mol Pathol 11:163–169
106. Malumbres M, Perez de Castro I, Santos J, Melendez B, Mangues R, Serrano M et al (1997) Inactivation of the cyclin-dependent kinase inhibitor p15INK4b by deletion and de novo methylation with independence of p16INK4a alterations in murine primary T-cell lymphomas. Oncogene 14:1361–1370
107. Malumbres M, Perez de Castro I, Santos J, Fernandez Piqueras J, Pellicer A (1999) Hypermethylation of the cell cycle inhibitor p15INK4b 3’-untranslated region interferes with its transcriptional regulation in primary lymphomas. Oncogene 18:385–396
108. Cleary HJ, Boulton E, Plumb M (1999) Allelic loss and promoter hypermethylation of the p15INK4b gene features in mouse radiation-induced lymphoid – but not myeloid – leukaemias. Leukemia 13:2049–2052
109. Rossman TG (2003) Mechanism of arsenic carcinogenesis: an integrated approach. Mutat Res 533:37–65
110. Zhong CX, Mass MJ (2001) Both hypomethylation and hypermethylation of DNA associated with arsenite exposure in cultures of human cells identified by methylation-sensitive arbitrarily-primed PCR. Toxicol Lett 122:223–234
111. Mass MJ, Wang L (1997) Arsenic alters cytosine methylation patterns of the promoter of the tumor suppressor gene p53 in human lung cells: a model for a mechanism of carcinogenesis. Mutat Res 386:263–277
112. Zhao CQ, Young MR, Diwan BA, Coogan TP, Waalkes MP (1997) Association of arsenic-induced malignant transformation with DNA hypomethylation and aberrant gene expression. Proc Natl Acad Sci USA 94:10907–10912
113. Okoji RS, Yu RC, Maronpot RR, Froines JR (2002) Sodium arsenite administration via drinking water increases genome-wide and Ha-ras DNA hypomethylation in methyl-deficient C57BL/6J mice. Carcinogenesis 23:777–785
114. McDorman EW, Collins BW, Allen JW (2002) Dietary folate deficiency enhances induction of micronuclei by arsenic in mice. Environ Mol Mutagen 40:71–77

115. Chanda S, Dasgupta UB, Guhamazumder D, Gupta M, Chaudhuri U, Lahiri S et al (2006) DNA hypermethylation of promoter of gene p53 and p16 in arsenic-exposed people with and without malignancy. Toxicol Sci 89:431–437
116. Majumdar S, Chanda S, Ganguli B, Mazumder DN, Lahiri S, Dasgupta UB (2010) Arsenic exposure induces genomic hypermethylation. Environ Toxicol 25:315–318
117. Pilsner JR, Liu X, Ahsan H, Ilievski V, Slavkovich V, Levy D et al (2007) Genomic methylation of peripheral blood leukocyte DNA: influences of arsenic and folate in Bangladeshi adults. Am J Clin Nutr 86:1179–1186
118. Gamble MV, Liu X, Slavkovich V, Pilsner JR, Ilievski V, Factor-Litvak P et al (2007) Folic acid supplementation lowers blood arsenic. Am J Clin Nutr 86:1202–1209
119. Hall M, Gamble M, Slavkovich V, Liu X, Levy D, Cheng Z et al (2007) Determinants of arsenic metabolism: blood arsenic metabolites, plasma folate, cobalamin, and homocysteine concentrations in maternal-newborn pairs. Environ Health Perspect 115:1503–1509
120. Pilsner JR, Liu X, Ahsan H, Ilievski V, Slavkovich V, Levy D et al (2009) Folate deficiency, hyperhomocysteinemia, low urinary creatinine, and hypomethylation of leukocyte DNA are risk factors for arsenic-induced skin lesions. Environ Health Perspect 117:254–260
121. Marsit CJ, Eddy K, Kelsey KT (2006) MicroRNA responses to cellular stress. Cancer Res 66:10843–10848
122. Ramirez T, Brocher J, Stopper H, Hock R (2008) Sodium arsenite modulates histone acetylation, histone deacetylase activity and HMGN protein dynamics in human cells. Chromosoma 117:147–157
123. Broday L, Peng W, Kuo MH, Salnikow K, Zoroddu M, Costa M (2000) Nickel compounds are novel inhibitors of histone H4 acetylation. Cancer Res 60:238–241
124. Ke Q, Davidson T, Chen H, Kluz T, Costa M (2006) Alterations of histone modifications and transgene silencing by nickel chloride. Carcinogenesis 27:1481–1488
125. Takiguchi M, Achanzar WE, Qu W, Li G, Waalkes MP (2003) Effects of cadmium on DNA-(Cytosine-5) methyltransferase activity and DNA methylation status during cadmium-induced cellular transformation. Exp Cell Res 286:355–365
126. Bollati V, Marinelli B, Apostoli P, Bonzini M, Nordio F, Hoxha M et al (2010) Exposure to metal-rich particulate matter modifies the expression of candidate microRNAs in peripheral blood leukocytes. Environ Health Perspect 118:763–768
127. Wright RO, Schwartz J, Wright RJ, Bollati V, Tarantini L, Park SK et al (2010) Biomarkers of lead exposure and DNA methylation within retrotransposons. Environ Health Perspect 118:790–795
128. Pilsner JR, Lazarus AL, Nam DH, Letcher RJ, Sonne C, Dietz R et al (2010) Mercury-associated DNA hypomethylation in polar bear brains via the LUminometric Methylation Assay: a sensitive method to study epigenetics in wildlife. Mol Ecol 19:307–314
129. Gibb HJ, Lees PS, Pinsky PF, Rooney BC (2000) Lung cancer among workers in chromium chemical production. Am J Ind Med 38:115–126
130. Zhitkovich A, Shrager S, Messer J (2000) Reductive metabolism of Cr(VI) by cysteine leads to the formation of binary and ternary Cr–DNA adducts in the absence of oxidative DNA damage. Chem Res Toxicol 13:1114–1124
131. Shi H, Hudson LG, Liu KJ (2004) Oxidative stress and apoptosis in metal ion-induced carcinogenesis. Free Radic Biol Med 37:582–593
132. Kondo K, Takahashi Y, Hirose Y, Nagao T, Tsuyuguchi M, Hashimoto M et al (2006) The reduced expression and aberrant methylation of p16(INK4a) in chromate workers with lung cancer. Lung Cancer 53:295–302
133. Schnekenburger M, Talaska G, Puga A (2007) Chromium cross-links histone deacetylase 1-DNA methyltransferase 1 complexes to chromatin, inhibiting histone-remodeling marks critical for transcriptional activation. Mol Cell Biol 27:7089–7101
134. Vineis P, Husgafvel-Pursiainen K (2005) Air pollution and cancer: biomarker studies in human populations. Carcinogenesis 26:1846–1855
135. Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL (1987–1994) Fine particulate air pollution and mortality in 20 U.S. cities, 1987–1994. N Engl J Med 343:1742–1749

136. Belinsky SA, Snow SS, Nikula KJ, Finch GL, Tellez CS, Palmisano WA (2002) Aberrant CpG island methylation of the p16(INK4a) and estrogen receptor genes in rat lung tumors induced by particulate carcinogens. Carcinogenesis 23:335–339
137. Jardim MJ, Fry RC, Jaspers I, Dailey L, Diaz-Sanchez D (2009) Disruption of microRNA expression in human airway cells by diesel exhaust particles is linked to tumorigenesis-associated pathways. Environ Health Perspect 117:1745–1751
138. Tarantini L, Bonzini M, Apostoli P, Pegoraro V, Bollati V, Marinelli B et al (2009) Effects of particulate matter on genomic DNA methylation content and iNOS promoter methylation. Environ Health Perspect 117:217–222
139. Baccarelli A, Wright RO, Bollati V, Tarantini L, Litonjua AA, Suh HH et al (2009) Rapid DNA methylation changes after exposure to traffic particles. Am J Respir Crit Care Med 179:572–578
140. Ren C, Park SK, Vokonas PS, Sparrow D, Wilker E, Baccarelli A et al (2010) Air pollution and homocysteine: more evidence that oxidative stress-related genes modify effects of particulate air pollution. Epidemiology 21:198–206
141. Baccarelli A, Zanobetti A, Martinelli I, Grillo P, Hou L, Lanzani G et al (2007) Air pollution, smoking, and plasma homocysteine. Environ Health Perspect 115:176–181
142. Wilker EH, Baccarelli A, Suh H, Vokonas P, Wright RO, Schwartz J (2010) Black carbon exposures, blood pressure, and interactions with single nucleotide polymorphisms in MicroRNA processing genes. Environ Health Perspect 118:943–948
143. Robinson BW, Lake RA (2005) Advances in malignant mesothelioma. N Engl J Med 353:1591–1603
144. Sugarbaker DJ, Richards WG, Gordon GJ, Dong L, De Rienzo A, Maulik G et al (2008) Transcriptome sequencing of malignant pleural mesothelioma tumors. Proc Natl Acad Sci USA 105:3521–3526
145. He B, Lee AY, Dadfarmay S, You L, Xu Z, Reguart N et al (2005) Secreted frizzled-related protein 4 is silenced by hypermethylation and induces apoptosis in beta-catenin-deficient human mesothelioma cells. Cancer Res 65:743–748
146. Hirao T, Bueno R, Chen CJ, Gordon GJ, Heilig E, Kelsey KT (2002) Alterations of the p16(INK4) locus in human malignant mesothelial tumors. Carcinogenesis 23:1127–1130
147. Lee AY, He B, You L, Dadfarmay S, Xu Z, Mazieres J et al (2004) Expression of the secreted frizzled-related protein gene family is downregulated in human mesothelioma. Oncogene 23:6672–6676
148. Ohta Y, Shridhar V, Kalemkerian GP, Bright RK, Watanabe Y, Pass HI (1999) Thrombospondin-1 expression and clinical implications in malignant pleural mesothelioma. Cancer 85:2570–2576
149. Tsou JA, Galler JS, Wali A, Ye W, Siegmund KD, Groshen S et al (2007) DNA methylation profile of 28 potential marker loci in malignant mesothelioma. Lung Cancer 58:220–230
150. Christensen BC, Godleski JJ, Marsit CJ, Houseman EA, Lopez-Fagundo CY, Longacker JL et al (2008) Asbestos exposure predicts cell cycle control gene promoter methylation in pleural mesothelioma. Carcinogenesis 29:1555–1559
151. Christensen BC, Houseman EA, Godleski JJ, Marsit CJ, Longacker JL, Roelofs CR et al (2009) Epigenetic profiles distinguish pleural mesothelioma from normal pleura and predict lung asbestos burden and clinical outcome. Cancer Res 69:227–234
152. Feinberg AP, Cui H, Ohlsson R (2002) DNA methylation and genomic imprinting: insights from cancer into epigenetic mechanisms. Semin Cancer Biol 12:389–398
153. Feinberg AP (2000) DNA methylation, genomic imprinting and cancer. Curr Top Microbiol Immunol 249:87–99
154. Feinberg AP (2004) The epigenetics of cancer etiology. Semin Cancer Biol 14:427–432
155. Wolffe AP (2000) Transcriptional control: imprinting insulation. Curr Biol 10:R463–465
156. Tilghman SM (1999) The sins of the fathers and mothers: genomic imprinting in mammalian development. Cell 96:185–193

157. Dean W, Bowden L, Aitchison A, Klose J, Moore T, Meneses JJ et al (1998) Altered imprinted gene methylation and expression in completely ES cell-derived mouse fetuses: association with aberrant phenotypes. Development 125:2273–2282
158. Doherty AS, Mann MR, Tremblay KD, Bartolomei MS, Schultz RM (2000) Differential effects of culture on imprinted H19 expression in the preimplantation mouse embryo. Biol Reprod 62:1526–1535
159. Khosla S, Dean W, Brown D, Reik W, Feil R (2001) Culture of preimplantation mouse embryos affects fetal development and the expression of imprinted genes. Biol Reprod 64:918–926
160. Young LE, Sinclair KD, Wilmut I (1998) Large offspring syndrome in cattle and sheep. Rev Reprod 3:155–163
161. Cox GF, Burger J, Lip V, Mau UA, Sperling K, Wu BL et al (2002) Intracytoplasmic sperm injection may increase the risk of imprinting defects. Am J Hum Genet 71:162–164
162. DeBaun MR, Niemitz EL, Feinberg AP (2003) Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. Am J Hum Genet 72:156–160
163. Cruz-Correa M, Cui H, Giardiello FM, Powe NR, Hylind L, Robinson A et al (2004) Loss of imprinting of insulin growth factor II gene: a potential heritable biomarker for colon neoplasia predisposition. Gastroenterology 126:964–970
164. Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H et al (2010) Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome Res 20:434–439
165. Jones PA, Liang G (2009) Rethinking how DNA methylation patterns are maintained. Nat Rev Genet 10:805–811
166. Masuda T, Shinoara H, Kondo M (1975) Reactions of hydroxyl radicals with nucleic acid bases and the related compounds in gamma-irradiated aqueous solution. J Radiat Res (Tokyo) 16:153–161
167. Valinluck V, Tsai HH, Rogstad DK, Burdzy A, Bird A, Sowers LC (2004) Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). Nucleic Acids Res 32:4100–4108
168. Valinluck V, Sowers LC (2007) Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. Cancer Res 67:946–950
169. Henderson JP, Byun J, Takeshita J, Heinecke JW (2003) Phagocytes produce 5-chlorouracil and 5-bromouracil, two mutagenic products of myeloperoxidase, in human inflammatory tissue. J Biol Chem 278:23522–23528
170. Turker MS (2002) Gene silencing in mammalian cells and the spread of DNA methylation. Oncogene 21:5388–5393
171. Turker MS, Bestor TH (1997) Formation of methylation patterns in the mammalian genome. Mutat Res 386:119–130
172. Sun FL, Elgin SC (1999) Putting boundaries on silence. Cell 99:459–462
173. Robertson KD (2002) DNA methylation and chromatin – unraveling the tangled web. Oncogene 21:5361–5379
174. Rountree MR, Bachman KE, Herman JG, Baylin SB (2001) DNA methylation, chromatin inheritance, and cancer. Oncogene 20:3156–3165
175. Nguyen CT, Weisenberger DJ, Velicescu M, Gonzales FA, Lin JC, Liang G et al (2002) Histone H3-lysine 9 methylation is associated with aberrant gene silencing in cancer cells and is rapidly reversed by 5-aza-2'-deoxycytidine. Cancer Res 62:6456–6461
176. Wiencke JK (2002) DNA adduct burden and tobacco carcinogenesis. Oncogene 21:7376–7391
177. Memisoglu A, Samson L (2000) Base excision repair in yeast and mammals. Mutat Res 451:39–51
178. Demple B, DeMott MS (2002) Dynamics and diversions in base excision DNA repair of oxidized abasic lesions. Oncogene 21:8926–8934
179. Parsian AJ, Funk MC, Tao TY, Hunt CR (2002) The effect of DNA damage on the formation of protein/DNA complexes. Mutat Res 501:105–113

180. Christensen BC, Houseman EA, Poage GM, Godleski JJ, Bueno R, Sugarbaker DJ et al (2010) Integrated profiling reveals a global correlation between epigenetic and genetic alterations in mesothelioma. Cancer Res 70:5686–5689
181. Ehrlich M (2002) DNA methylation in cancer: too much, but also too little. Oncogene 21:5400–5413
182. Florl AR, Lower R, Schmitz-Drager BJ, Schulz WA (1999) DNA methylation and expression of LINE-1 and HERV-K provirus sequences in urothelial and renal cell carcinomas. Br J Cancer 80:1312–1321
183. Maloisel L, Rossignol JL (1998) Suppression of crossing-over by DNA methylation in Ascobolus. Genes Dev 12:1381–1389
184. Dominguez-Bendala J, McWhir J (2004) Enhanced gene targeting frequency in ES cells with low genomic methylation levels. Transgenic Res 13:69–74
185. Colot V, Rossignol JL (1999) Eukaryotic DNA methylation as an evolutionary device. Bioessays 21:402–411
186. Pogribny IP, Basnakian AG, Miller BJ, Lopatina NG, Poirier LA, James SJ (1995) Breaks in genomic DNA and within the p53 gene are associated with hypomethylation in livers of folate/methyl-deficient rats. Cancer Res 55:1894–1901
187. Machover D, Zittoun J, Saffroy R, Broet P, Giraudier S, Magnaldo T et al (2002) Treatment of cancer cells with methioninase produces DNA hypomethylation and increases DNA synthesis. Cancer Res 62:4685–4689
188. Kim YI (2004) Folate and DNA methylation: a mechanistic link between folate deficiency and colorectal cancer? Cancer Epidemiol Biomarkers Prev 13:511–519
189. Thilly WG (2003) Have environmental mutagens caused oncomutations in people? Nat Genet 34:255–259
190. Slaughter DP, Southwick HW, Smejkal W (1953) Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer 6:963–968
191. Schwartz DA, Freedman JH, Linney EA (2004) Environmental genomics: a key to understanding biology, pathophysiology and disease. Hum Mol Genet 13:R217–224

Chapter 13
Epigenetic Epidemiology of Cancer

Angela Risch, Narek Sarkisyan, David Scherf, Hanna Jacobsson, Wolfgang Hagmann, and Christoph Plass

Contents

13.1 Epigenetic Markers in Tumor Tissue and Surrogate Tissue as Diagnostic and Predictive/Prognostic Markers ... 227
13.1.1 DNA Methylation ... 227
13.1.2 Histone Modifications and Chromatin Structure ... 229
13.1.3 miRNAs ... 229
13.2 Epigenetic Mechanisms Play a Key Role in Mediating Environmental Influences on Gene Expression and Tumor Development ... 230
13.2.1 One-Carbon Metabolism ... 230
13.2.2 Smoking, Other Environmental Factors, Age, and Epigenetics ... 232
13.3 Mechanisms of Interaction Between Genetic and Epigenetic Factors in Cancer (Risk) ... 233
13.4 The Potential of Epigenetic Biomarkers in Cancer Epidemiology ... 234
13.4.1 Gene-Epigene-Environment Interaction ... 235
13.4.2 Transgenerational Inheritance ... 235
13.5 State of the Art ... 236
13.5.1 Epigenetic Epidemiology of Colorectal Cancer ... 236
13.5.2 Epigenetic Epidemiology of Upper Aerodigestive Tract Tumors ... 238
13.5.3 Epigenetic Epidemiology of Hormone-Related Cancers ... 240
13.5.4 Epigenetic Epidemiology in Leukemia ... 245
13.5.5 Epigenetic Epidemiology of Other Tumors ... 247
13.6 Outlook ... 250
References ... 251

A. Risch (✉) • N. Sarkisyan • D. Scherf • H. Jacobsson • W. Hagmann • C. Plass
Division C010, Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
e-mail: a.risch@dkfz.de; n.sarkisyan@dkfz.de; d.scherf@dkfz.de; h.Jacobsson@dkfz.de; w.hagmann@dkfz.de; c.plass@dkfz.de

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_13,

© Springer Science+Business Media B.V. 2012

Abstract Epigenetic changes occur in all types of cancers and include alterations in deoxyribonucleic acid (DNA) methylation, histone modification, or expression changes of non-coding ribonucleic acid (RNA). There is now evidence that epigenetic alterations influence cancer risk and could be developed into useful biomarkers. While it is clear that epigenetic mechanisms can play a key role in mediating environmental influences including dietary folate intake, smoking, other environmental stressors, and age, many of the mechanisms involved remain obscure. There are numerous mechanisms by which genetic and epigenetic factors interact and thus the field of epigenetic epidemiology of cancer faces considerable challenges in tearing apart gene-epigene-environment interactions and in determining how they modulate individual cancer risk. An important question for epigenetic epidemiology of cancer is whether epigenetic changes can be identified in surrogate tissues; e.g., DNA isolated from peripheral lymphocytes, plasma samples, urine, or buccal swabs. The current state of the art of cancer epigenetic epidemiology studies as relating to epigenetic cancer risk factors and markers of diagnostic and prognostic significance is reviewed. In the past, epigenetic work has been carried out in comparative studies of primary human normal/tumor tissue pairs, but recent advances in technologies have made large scale analyses on surrogate tissues within defined cohorts possible. Interindividual variations exist in epigenetic markers, but their origin and relevance needs to be determined. In future studies it will be important to characterize the interaction between environmental, dietary, genetic, and epigenetic factors. Echoing the development in the field of genetics, candidate gene approaches employed thus far will be replaced by genome-wide epigenome analyses in the future. Approaches integrating genetic, epigenetic, and well-documented epidemiological and clinical data are most promising and may improve the currently available risk models for cancer in the future.

Abbreviations

DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
GBM	Glioblastoma multiforme
HCC	Hepatocellular carcinoma
miRNA	microRNA
NSCLC	Non-small cell lung cancer
PDAC	Pancreatic adenocarcinoma
RCC	Renal clear cell carcinoma
RNA	Ribonucleic acid
SAH	S-adenosylhomocysteine
SAM	S-adenosylmethionine
SCC	Squamous cell carcinoma
SCLC	Small cell lung cancer
SNP	Single nucleotide polymorphism

13.1 Epigenetic Markers in Tumor Tissue and Surrogate Tissue as Diagnostic and Predictive/Prognostic Markers

Epigenetic questions have reached center stage in cancer research. Here research has focused on the mechanisms that lead to epigenetic modifications (DNA methylation, histone modifications, and alterations of non-coding RNA expression) and how these modifications relate to cancer initiation and progression mediated through their effects on gene expression. However, it has become obvious that epigenetic alterations are also attractive targets for cancer therapy as spearheaded in leukemias. Furthermore, due to the stable nature of modifications, in particular the covalent binding of a methylgroup to position 5 of the cytosine ring (DNA methylation), this epigenetic modification is under intense investigation as a possible target to be developed into a biomarker for early detection or as a marker with prognostic or predictive value. Focus in this review will be on the potential of DNA methylation as a biomarker for risk assessment and early detection with relevance for cancer epidemiology.

13.1.1 DNA Methylation

DNA methylation is probably the most intensively studied epigenetic mark in cancers and is now recognized as a crucial component in cancer initiation and progression. Site specific hypermethylation of the usually unmethylated CpG islands in gene promoters is found in almost every human malignancy and is associated with transcriptional silencing of adjacent genes. Hypermethylation mediated silencing in cancer cells often affects classical tumor suppressor genes involved in DNA repair, cell-cycle control, anti-apoptotic signaling, and genes that regulate developmental pathways. Interestingly, aberrant CpG island methylation in cancer has non-random, tumor-type-specific patterns indicating the presence of molecular mechanisms leading to selected gene silencing. Similarly, global loss of DNA methylation (hypomethylation) in cancer results in aberrant transcription via genomic instability and activation of silenced oncogenes [72, 79].

One of the most promising approaches for the detection of early malignant or even premalignant lesions with the chance of diagnosing early stages of many tumors lies in the development of new biomarkers. Aberrant promoter methylation of tumor suppressor genes is believed to be an early event in carcinogenesis and to occur before the clinical onset of disease, and hence, in principle, holds great promise as early detection marker. Tumor biomarkers are indispensable not only for early diagnosis of cancer but also for prognosis, prediction of therapeutic response, monitoring therapy, or assessment of risk of recurrence after surgery. Given that a number of different cancers frequently harbor hypermethylated tumor suppressor genes together with the observation that tumor-derived genomic DNAs are present in various body fluids, e.g., in sputum [21], urine [33], and bronchial lavage [55], DNA methylation analysis is a promising minimally invasive tool for early diagnosis from

surrogate tissue. The epigenetic biomarkers particularly have the potential to improve the accuracy of diagnosis in targeted sub-populations, such as men at high risk of prostate cancer with a persistently increased level of prostate specific antigen (PSA) and negative biopsies, for the detection of bladder cancer in patients with hematuria, populations with occupational carcinogen exposure, with family history of cancer, etc. Although key studies showed sensitive and specific detection of hypermethylated genes in urine from patients with early stage prostate, bladder, and renal cancer, the sensitivity of detection of methylated genes in urine of those patients varies a lot from 49% to 91% [32]. In cancers of the aerodigestive tract, detection of aberrant promoter methylation of tumor suppressor genes in sputum accurately represents the methylation status of those genes in the tumor. However, disadvantages of sputum analysis are the variability in composition of the specimen (epithelial vs. inflammatory components) and the fact that up to 30% of current or former smokers do not have increased bronchial secretions and will not produce the sputum. Moreover, studies so far have not exclusively examined early stage tumors, for which the sensitivity of detection might be lower. Another important consideration in the development of early detection epigenetic biomarkers is their specificity for tumor DNA compared with DNA from normal or non-neoplastic cells. Studies of gene methylation in body fluids have not examined a sufficient numbers of controls to clearly demonstrate a high specificity for cancer. Differences in sensitivity and specificity between different studies probably derive from different genes analysed, primer design and different methods for methylation analysis.

Recently, cell-free DNA has been identified in the serum and plasma from patients with various cancers [21, 184, 192, 227, 238, 267]. These circulating nucleic acids have demonstrated similar epigenetic alterations and characteristics as those found in the primary tumor providing a potential surrogate marker for diagnosis, monitoring, and prognosis. Their levels appear to be elevated during disease progression. Detection of aberrant promoter methylation of several tumor suppressor genes (*APC*, *GSTP1*, *RASSF1A*, *RARβ2*, *p16*, *BLU*, *FHIT*, *CDH13*, *Pax5α*, *DAPK1*, *MGMT*, *GATA5*, etc.) in blood has also been intensively studied as an early detection and prognostic marker in various cancers, including lung, breast, and colorectal cancer [19, 20, 22, 106, 108, 195, 243]. Although the exact mechanism of the release of circulating DNA remains to be proven [7], an active release of tumor gDNA from highly proliferating tumor cells during its growth into the bloodstream has been proposed [8]. However, the number of such tumor circulating DNA fragments is small compared to the number of normal circulating DNA fragments, making it challenging to detect and quantify them with sufficient sensitivity for meaningful clinical use. Highly sensitive assays will have to be developed to assess gene-promoter methylation in blood [251].

There are already diagnostic assays for routine epigenetic testing used in the clinic; among them tests for hypermethylation of the *MGMT* gene promoter for prediction of response to temozolomide in glioma tissue, *GSTP1* and *APC* hypermethylation in prostate biopsy tissue for detection of cancer, methylation of vimentin in stool-based screen for colorectal cancer, *SEPT9* methylation in serum DNA for colorectal cancer, and *MLH1* methylation for mismatch-repair–deficient colon cancer, etc. [58].

Instead of having one epigenetic biomarker, diagnostic coverage for early stage tumors may require a larger panel of genes. Challenges to the clinical implementation of epigenetic early detection include the need for validation in larger populations with optimized and standardized methodology. Up to now no cut-off point in quantitative methylation values between normal, benign, and cancer DNA is established that could reliably indicate the presence of cancer [32]. The timing of gene methylation during the early stages of cancer must also be better understood.

13.1.2 Histone Modifications and Chromatin Structure

Histones, strongly alkaline proteins found in eukaryotic cell nuclei, which package and order DNA into structural units called nucleosomes, also play an important role in epigenetic gene regulation. In contrast to DNA methylation, the patterns of histone modifications are much more complex. Histones are subject to various post-synthetic modifications, including methylation, acetylation, ubiquitylation, phosphorylation, and sumoylation that play important roles in gene regulation in normal cells. In cancer these patterns can change and lead to altered gene expression. Acetylated histones associate preferentially with transcriptionally activated chromatin and may also facilitate binding of transcription factors to the promoters and disrupt higher order of chromosome structure, promoting transcription. Histone deacetylation is often found to be overrepresented around promoter regions of cancer genes resulting in the formation of a more compact structure to silence genes. Histone modifications can also be used to predict prognosis and response to treatment in subsets of cancer patients. For example, decreased global histone modification levels in tumor tissue (H3K4me2 or H3K18Ac) are predictive for prostate cancer recurrence [213], and global histone modifications of multiple histones have prognostic relevance in glioma patients [147]. A role for global histone deacetylation as a prognostic marker was shown in various tumors including breast cancer (H3K9Ac, H3K18Ac , H4K16Ac), lung cancer (H3K9Ac), and renal cell carcinoma (H3K18Ac) [16, 62, 170]. Low cellular levels of H3K27me3 were also recently associated with poor outcome in pancreatic cancer [155, 259]. Since histone modifications also exhibit abnormal regulation in cancer, and those abnormalities are reversible, they have attracted great interest in cancer research as potential drug targets.

13.1.3 miRNAs

MicroRNAs (miRNAs) are short, single stranded RNA molecules that do not encode proteins but instead restrict the production of proteins by inhibiting translation of coding or messenger RNAs, or by promoting their degradation. The first miRNA was discovered in 1993 in the nematode *Caenorhabditis elegans*, where it was found to regulate postembryonic development [135]. More than 1,000 human miRNAs

have been discovered until now, and this number is increasing rapidly. One-third of all human genes may be regulated by miRNAs and a considerable number of deregulated miRNAs are involved in cancer progression [130]. MicroRNA expression levels differ between normal and tumor tissue and vary among tumor types [186, 209]. During the last decade, a unique set of cancer regulator miRNAs has emerged and they are divided into oncomiRNAs and anti-oncomiRNAs. OncomiRNAs and anti-oncomiRNAs negatively regulate tumor suppressor genes and oncogenes, respectively. Results from expression analysis indicate that specific miRNAs are associated with disease stage and survival in several tumor types, such as chronic lymphocytic leukemia, prostate cancer, breast cancer, lung cancer, and colorectal cancer [34, 78, 211, 244]. For instance, expression levels of miRNA-486, -30d, -1, and -499 in serum are significantly associated with overall survival in non-small cell lung cancer (NSCLC) patients [109]. MiRNA-21 is also overexpressed in sputum of NSCLC patients compared to normal subjects [272]. MiRNAs are stable and differentially present in serum and plasma of cancer patients, as well as in extracellular biological materials such as stool and urine [41, 91, 167] and thus have potential of serving as biomarkers for diagnosis, prognosis, and monitoring of cancer via noninvasive methods. Overexpression of several miRNAs in tumor tissues of various cancer types has already been successfully shown to serve as an independent predictive marker of overall survival. Besides plasma and serum, miRNAs can also be identified in other body fluids such as saliva, urine, and semen. The utility of miRNAs in body fluids as biomarkers for various cancer types is being actively investigated; however, an important obstacle for quantification of miRNAs in body fluids is the relatively low amount of RNA in such fluids, and the lack of a suitable RNA for normalization purposes.

13.2 Epigenetic Mechanisms Play a Key Role in Mediating Environmental Influences on Gene Expression and Tumor Development

Epidemiologic studies have shown that certain environmental exposures are associated with various steps during cancer development. Epigenetic mechanisms play a crucial role in tumor development, also by mediating environmental influences on gene expression (Fig. 13.1). Environmental stress including toxins, as well as microbial and viral exposures, nutrition, and aging can change epigenetic patterns, potentially increase or decrease the levels of transcription, thereby effecting changes in gene activation and cell phenotype (see [95] for review).

13.2.1 One-Carbon Metabolism

The one-carbon metabolism is centered on S-adenosylmethionine (SAM) and S-adenosylhomocysteine (SAH), and the SAM to SAH conversion frees up methyl

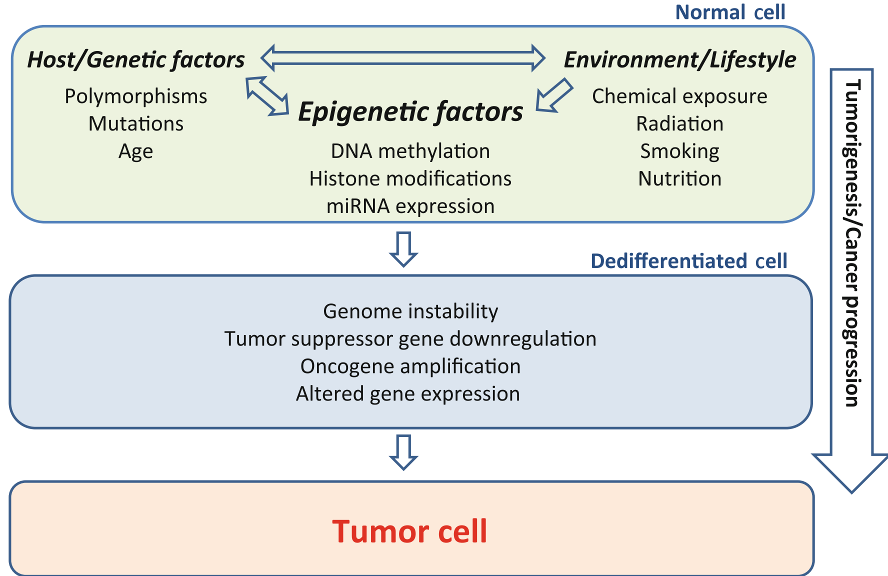


Fig. 13.1 The interplay of host-, environment- and epigenetic factors in tumorigenesis

groups for DNA methylation. This biochemical pathway plays a critical role in DNA methylation and DNA synthesis, and thus facilitates epigenetic and genetic carcinogenic processes. The association between one-carbon metabolism and cancer has been extensively studied and epidemiological studies have shown the contribution of one-carbon metabolism in cancer etiology, prognosis, treatment, and prevention. One-carbon metabolism is not only capable of influencing the pathogenesis of cancer, but it can also be the target pathway for chemotherapy as well as chemoprevention. The major external source of methyl groups comes from folate, which plays an essential role in one-carbon transfer, as well as methionine, and choline in diet. Other vitamins, such as B_2, B_6, and B_{12} are also key factors for one-carbon metabolism. Alcohol is a folate antagonist, and excess consumption of alcohol affects folate absorption and decreases the hepatic uptake and renal conservation of circulating folate. Several case-control and cohort association studies, clinical, and animal studies suggest inverse associations between folate intake and risk of developing various cancers while some more recent studies support a positive association between very high intake from diet with lung, premenopausal and post-menopausal breast cancer and can promote the progression of already existing preneoplasms [60, 145, 183, 208, 230]. Evidence on folate and alcohol intake in relation to methylation in colorectal cancer brings strong support that diet is capable of influencing carcinogenesis through epigenetic pathways by showing that prevalence of promoter hypermethylation of a certain panel of genes is higher in tumors derived from patients with low folate and high alcohol intake compared to those

with high folate and low alcohol intake [252]. Evidence also exists that functional polymorphisms in one-carbon metabolizing genes can modify the DNA methylation status. For example, a study investigating three key genes involved in the metabolism of the methyl group, methylene-tetrahydrofolate reductase (*MTHFR*), methionine synthase, and cystathionine beta-synthase, reported higher DNA hypomethylation in people with the *MTHFR*677TT genotype compared to those with the *MTHFR*677CC genotype in colorectal, breast, and lung cancers [189]. It is also important to mention that unlike genetic changes, epigenetic modifications are potentially reversible, and the concept of modulating DNA methylation by nutrition is a highly attractive proposition in public health.

13.2.2 Smoking, Other Environmental Factors, Age, and Epigenetics

Cigarette smoke contains more than 60 carcinogens and there is a strong link between some of these carcinogens and various smoking-induced cancers, in particular lung cancer, esophageal, oral, pharyngeal, and laryngeal. The mechanisms of actions have been shown to include induction of DNA adducts, gene methylation, mutations and chromosomal translocations. Epigenetic changes, such as methylation patterns of CpG islands, have been linked with smoking, and inactivation of key genes through promoter hypermethylation is particularly believed to be one of the mechanisms by which tobacco smoke promotes the development of lung cancer. In vitro and animal studies showed that carcinogenic components of cigarette smoke induce DNA methylation changes in normal tissue as well as in malignant cells (see also Chap. 12). Human studies to investigate the effect of low-dose benzene, one component of cigarette smoke, on DNA methylation showed that *p15* promoter methylation in exposed individuals is twice as high as that observed in unexposed ones. Also, *MAGE1*, which is usually highly methylated in normal tissue, tends to have lower methylation levels in association with benzene [27], which all together indicates the involvement of benzene in the initial activation of the biological process of carcinogenesis. In a large prospective study it was shown that cigarette smoking is strongly associated with promoter hypermethylation of *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3*, and *SOCS1* tumor suppressor genes in colorectal cancer [143]. In summary, cigarette smoking represents a potentially modifiable risk factor for many cancers, but modulation of epigenetic susceptibility to cancer by cigarette smoke must be further investigated, since the data and results are not comprehensive and coherent between different types of cancers.

A variety of other environmental factors has also been linked with epigenetic phenomena (reviewed in Chap. 12), mainly in model organisms. In order to evaluate the relevance of epigene-environment interactions in human cancer, carefully conducted epidemiologic studies will be required.

Epigenetic investigation of cancer and aging has recently emerged as a crucial area of study to determine the impact of age on epigenetic changes and alterations.

Both genome-wide and specific methylation profiles change with age (see Chap. 11). Age-related differences in tumor methylation have been observed in cancers [124]. Aging is one of the greatest risk factors for cancer and the highest cancer rates occur in people over the age of 70 years. This association between cancer and aging has led many investigators to point to the accumulation of genetic mutations as a primary factor in tumorigenesis. However, a number of investigations have revealed epigenetic, rather than genetic, explanations for loss of gene expression involved in tumorigenesis [129]. In the context of epigenetic epidemiology of cancer it will be of crucial importance to elucidate the mechanisms and patterns of age-related methylation, and to take this phenomenon into account in careful study design or analysis.

13.3 Mechanisms of Interaction Between Genetic and Epigenetic Factors in Cancer (Risk)

Genetic variations as well as epigenetic modifications are known to be involved in tumorigenesis by modifying the expression of cancer genes [18, 25, 121, 203]. Besides the individual impact of these factors recent publications have drawn attention to the interplay between genetic and epigenetic factors [103]. Advanced, sequencing based technologies, now allow the thorough analysis of the dependencies among epigenetic modifications as well as the influence of genetic on epigenetic factors, an example of which are genetic variants in epigenetic genes.

DNA methyltransferases (DNMTs) are the main players in methylating DNA in animals. In humans DNMT1 is mainly responsible for maintaining methylation and DNMT3A/B for de novo DNA methylation. Recurrent mutations recently identified by whole genome sequencing in acute myeloid leukemia highlight the link between genetic alterations and global epigenetic dysregulation [141]. DNMT overexpression has been detected in cancerous tissue and various miRNAs targeting the 3′-UTR of DNMT are downregulated in cancers [253]. Genetic variants in DNMTs are known to modify risk of various cancers and can be linked to deregulation of DNMTs [70, 137, 150, 168, 214, 220, 270]. Epidemiologic studies that combined haplotype and mutagen sensitivity analysis identified *DNMT1* and *DNMT3B* variants which associate with enhanced sensitivity to tobacco carcinogens in cancer-free smokers [139]. In addition to genes affecting DNA methylation further genes involved in histone tail modification are highlighted by other studies. In an association study, consisting of ~4,500 cases and controls, genetic variants in several histone methyltransferases (HMTs) were detected. After genetic and epigenetic follow up in cell lines those HMT polymorphisms could be linked to levels of H3K9 and H3K27 trimethylation [37]. By single nucleotide polymorphism (SNP) analysis in combination with expression, methylation, or ChIP analysis a lot of information could be gained how epigenetic modifications are manifested by germline SNPs and reflected in an expression phenotype [59, 121, 123, 219]. The phenomenon of mitotically inherited germline aberration in methylation has also been described and has been termed “epimutation” [98, 235].

Genetic common variants which have been shown to predispose to cancer by modifying binding of miRNAs either by presence of SNPs in the seeding or the binding site have been detected [128]. Further, a meta-analysis could show that one SNP in pre-miRNA hsa-miR-196a2 was associated with breast cancer risk by potential modification of processing and subsequent expression of the mature miRNA [80].

In addition to genetic factors such as polymorphisms, it could be shown that deregulation of miRNAs by aberrant DNA demethylation can contribute to tumorigenesis either as tumor suppressor or as oncogenes. Loss of DNA methylation in gastric cancer leads to a high level of miRNA-196b [249]. Transfection of neuroblastoma cells with miRNA-152 which targets DNMTs together with all-trans retinoic acid treatment leads to global demethylation, upregulation of the nitric oxide synthase (NOS1), and differentiation [54]. Taken together, there are many examples of how genetic features like mutations or polymorphisms can modify epigenetic events, which eventually could be relevant for the development of malignancies. Hence multifactorial analysis of different regulatory layers, including genetic and epigenetic factors, are necessary to reveal a conclusive picture of disease associations.

13.4 The Potential of Epigenetic Biomarkers in Cancer Epidemiology

As described above, epigenetic changes, in particular DNA methylation, can be detected in body fluids in various cancer types. Aberrant methylation is thus a good candidate biomarker both as clinical (early diagnostic and/or prognostic) marker and for cancer risk assessment. For diagnostic purposes methylation aberrations can be detected in surrogate tissues or body fluids such as in urine for bladder cancer, pancreatic juice for pancreatic cancer, sputum for lung cancer, and stool for colorectal cancer. An impressive panel of genes has already been shown to be aberrantly hypermethylated in tumors, and those same epigenetic aberrations are also being heavily investigated in such body fluids in order to implement specific epigenetic biomarkers as a tool for risk assessment. Epidemiological studies based on biomarkers are crucial to predict and identify high risk populations. For example, in a case-control study it has been suggested that methylation of either *p16* or *RASSF1A* in plasma is associated with an increased risk of lung cancer (odds ratio (OR) 10.2 and 9.9, respectively) [108]. In a bigger cohort study, Belinsky with colleagues observed that aberrant methylation of six genes in sputum is associated with more than 50% increased lung cancer risk [23]. Other case-controls studies revealed nine genes whose methylation status in plasma is a significant and independent predictive factor for colorectal cancer [133], and that epigenetic inactivation of *MGMT* is also a useful tool in risk assessment of that disease [216]. Investigation of epigenetic biomarkers in body fluids for cancer risk assessment in a population-based manner is technically challenging, since only a small amount of DNA is present and thus highly sensitive assays must be performed. However, current rapid technological

progress will make it possible to investigate in more depth the potential of epigenetic biomarkers as risk markers for developing various cancers in big population studies using high-throughput quantitative methylation analysis.

Recently another area in the field of cancer research has emerged, namely to determine epigenetic risk biomarkers in blood, in particular, peripheral blood cells where epigenetic alterations are not expected to be related to the epigenetics of the tumor. Instead, epigenetic variation detected here, if detected before the clinical onset of cancer, may be risk biomarkers and might even be heritable. Challenges to the implementation of epigenetic risk biomarker detection in blood include the need for large population based cohort studies and the fact that those heritable epigenetic risk biomarkers, if so found, might be also mediated by various environmental exposure during the lifetime.

13.4.1 Gene-Epigene-Environment Interaction

Environmental agents can modify gene expression. Epigenetic modifications are the likely mediators of epi-gene-environment interaction because genetic factors can modify the epigenetic response to the environment, and certain epigenetic aberrations can have downstream genetic consequences [75, 115, 248]. There is accumulating evidence that epigenetic alterations may play an important role in determining individual susceptibility to cancer, and that in order to assess the role of epigenetic susceptibility in cancer risk, one should consider the effect of epi-gene-environment interaction in disease etiology. It is imperative to properly design epigenetic epidemiology studies taking into consideration a potential confounding effect of environmental factors (diet, smoking, lifestyle, etc.) and genetic alterations in order to comprehensively elucidate the epi-gene-environment interaction. Conducting epigenetic cancer case-control studies using blood samples one should be aware of the fact that blood might be contaminated by tumor DNA. It is also possible that the composition of cell types in blood is affected by the disease (e.g., due to concomitant inflammation), which may, in turn, affect the epigenetic patterns observed in peripheral blood lymphocytes and could lead to misinterpretation of results.

13.4.2 Transgenerational Inheritance

Although epigenetic states are maintained for the life of the organism, it is rare for these epigenetic states to be passed to the next generation. However, there is strong evidence for transmission of the epigenetic alterations through the gametes to the next generation in the mice. This is known as transgenerational epigenetic inheritance and was first demonstrated at the agouti locus in the mouse by Morgan and colleagues [169]. Transgenerational epigenetic inheritance has then been reported at another murine *Axin-fused* allele [199], at a genetically modified locus [96], and at *c-kit* [201].

Sensitivity to environment, combined with transgenerational epigenetic inheritance, suggests that the diet of a pregnant female could affect not only her offspring's coat color, but also that of next generations. Cropley with colleagues showed that methyl donor dietary supplementation can change the epigenetic state of the A^{vy} allele in the germline and that these modifications are retained through the epigenetic reprogramming during early embryogenesis [50]. Although DNA sequence with epigenetic information can be inherited from parent to offspring in mice, there is little direct evidence for transgenerational epigenetic inheritance in humans.

To summarize, epigenetic biomarkers have repeatedly been shown to have potential for risk assessment, prevention, early detection, diagnosis, treatment, prognosis, and recurrence in cancer epidemiology, and epidemiological studies based on epigenetic biomarkers, behavior, and lifestyle are crucial to predict and identify high risk populations.

13.5 State of the Art

13.5.1 Epigenetic Epidemiology of Colorectal Cancer

Colorectal cancer (CRC) is considered to be one of the major public health concerns and is probably the best studied cancer with respect to epigenetic alterations. This is facilitated by established screening approaches and availability of cancer tissue for epidemiological analysis. Environmental exposure may result in persistent epigenetic changes that can influence cancer development. In the frame of the Netherlands Cohort Study on diet and cancer, the association between early life energy intake and risk of subsequent CRC by promoter methylation of CpG island methylator phenotype (CIMP) genes was investigated [112]. Severe energy restriction during the Hunger Winter was associated with a decreased risk of developing CRC having characterized by CIMP compared to those not exposed (hazard ratio (HR) = 0.65) and exposure to the Hunger Winter was associated with the degree of promoter hypermethylation when further categorize individuals by an index "0–1", "2–3" or "4–7" genes methylated in the promoter region (HR = 1.01, when up to one gene is methylated; HR = 0.83, when 2–3 genes; and HR = 0.72, when 4–7 genes are methylated). Another case-control study within the Netherlands Cohort Study demonstrated that alcohol consumption increases methylation at promoters of *APC-1A*, *p14*, *p16*, *MLH1*, *MGMT*, and *RASSF1A* genes and is associated with methylation-induced silencing in colorectal cancer [252]. Additionally, the degree of promoter hypermethylation of all these genes was higher in cancer tissue from patients with low folate and high alcohol intake compared to those with high folate and low alcohol intake, consistent with the fact that alcohol is a folate antagonist. In another study, within the same Netherlands Cohort Study, Stefan de Vogel and colleagues investigated the association between folate intake, vitamin B2, vitamin B6, methionine, and alcohol and the risk of developing CRC with or without *MLH1*

hypermethylation, absence of *MLH1* protein expression, or *BRAF* mutations. Dietary intake of vitamin B6 was positively associated with CRC in men (RR = 1.54) and the association of vitamin B6 with CRC risk was even stronger with *MLH1* hypermethylation (RR = 3.23) [56]. In a population-based study of 951 cases and 1205 controls, the association of rectal tumors characterized by *CIMP*, *KRAS2*, and *TP53* mutations with dietary and supplemental folate, riboflavin, vitamins B6 and B12, and methionine were examined. Results indicate that folate supplementation effects may differ by gender and might be important in the initiation and progression of CIMP positive rectal cancers in men (OR = 3.2) [52]. A large population-based cohort study, the Iowa Women's Health Study, found that cigarette smoking is strongly associated with promoter hypermethylation of *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3*, and *SOCS1* tumor suppressor genes (RR = 1.88) in CRC [143], whereas no association was found between never-smokers and promoter hypermethylation. Compared with never-smokers, smokers have increased risk of CRC with increased promoter methylation (RR = 1.46). Taking all evidence together there is a strong indication that epigenetic changes may be mechanistically involved in environment-related carcinogenesis of CRC. In two prospective cohort studies (the Nurses' Health Study and the Health Professionals Follow-up Study) it has been reported that synchronous CRC patients have worse survival than solitary CRC patients, and, compared with solitary cancers, synchronous cancer patients showed more frequent promoter hypermethylation in a panel of eight genes [181]. Cui and colleagues investigated loss of imprinting (LOI), a form of epigenetic gene silencing which is found in the tumors and matched normal colonic mucosa of 30% CRC patients compared with only 10% of individuals without CRC, in the insulin-like growth factor II gene (*IGF2*), as a marker of CRC risk [51]. Results of this study indicate that LOI in blood is strongly associated with family history of CRC, compared to those without family history (OR = 5.15), strongly associated with past or present adenomas but no CRC, compared to those without colorectal neoplasia (OR = 3.46), and strongly associated with past or present CRC, compared to those without colorectal neoplasia (OR = 21.7), strongly indicating that epigenetic deregulation of IGF2 expression by LOI is associated with initiation and progression of CRC.

Regarding the field of epigenetic aberrations in CRC, in addition to cohort and case-control studies, it is imperative to mention the breakthrough studies that showed transgenerational epigenetic inheritance in a family with hereditary non-polyposis colorectal cancer (HNPCC). Chan et al. reported heritable germline methylation of *MSH2* associated with cancer development [39]. Germline hypermethylation in *MLH1* observed in two individuals was also found in normal somatic tissue and in spermatozoa of one individual, and there was no evidence of genetic mutations in *MLH1* gene [235], which indicates a germline epigenetic alteration that passes to the next generations. In tumors of patients with HNPCC, also known as Lynch syndrome, with *MSH2* deficiency, Nagasaka and colleagues also reported frequent somatic *MSH2* hypermethylation. In the context of Knudsons's "two hit" hypothesis this likely constitutes an epigenetic "second hit," affecting the wild-type allele. None of the sporadic tumors showed *MSH2* promoter methylation [174].

In comparison with DNA methylation, current knowledge regarding histone modifications in CRC patients is less advanced, but this field is being increasingly investigated. While thus far there is a lack of population-based association studies, aberrant histone modifications in CRC have been reported. CRC patients' DNMT1mRNA levels were significantly different between patients with elevated mRNA levels of SUV39H1 histone methyltransferase compared with those without, and patients with an elevated mRNA level of SUV39H1 had a 2.71 times greater risk of an elevated DNMT1 mRNA level [119], suggesting that SUV39H1 and DNMT1 may work together for transcriptional repression. Overexpression of a putative histone methyltransferase SMYD3 that methylates the lysine 4 residue of histone H3 in colorectal cancer has been also reported [90].

Among the first miRNAs to be studied in CRC were miRNA-143 and miRNA-145, which were both reported to be down-regulated in cancers compared to corresponding normal tissue [164, 257]. The most comprehensive analysis of miRNA expression in CRC performed to date showed that, in particular, miRNA-20a, miRNA-21, miRNA 106a, mi-RNA 181b, and miRNA-203 were found to be over expressed in tumor tissue with high tumor-normal ratios, as well as being associated with poor survival [211, 212]. High expression of miRNA-17-3-p and miRNA-92a was associated with the presence of CRC with quite a promising sensitivity of 89% and specificity of 70% [179]. Landi and colleagues investigated the association between miRNA-binding SNPs sequences and the risk of developing CRC within a hospital-based case-control study [132]. Thirty-seven among 104 candidate genes for CRC showed 57 SNPs in the predicted miRNA-binding sites. Heterozygosity and homozygosity for the variant allele rs17281995 within CD86 was associated with an increased risk of CRC (OR = 1.33 and OR = 2.93, respectively) and an increased risk of CRC was observed for homozygotes for the variant allele rs1051690 within *INSR* (OR = 1.86).

In summary, CRC is one of the best studied cancers in the area of epigenetic population-based association studies, and a role for epigenetic regulation in CRC initiation and progression is reported by numerous studies. The most widely applied screening method for CRC to date is the fecal occult blood test, and the principle of detecting epigenetically altered products shed from a tumor located in the colon and rectum into the stool and blood offers great opportunities for increased diagnostic sensitivity and specificity, and for identifying a population at high risk of developing CRC.

13.5.2 Epigenetic Epidemiology of Upper Aerodigestive Tract Tumors

13.5.2.1 Lung Cancer

Smoking is the most important risk factor for lung cancer. Short-term exposure of mice to tobacco smoke condensate has resulted in epigenetic changes in lung

tissue [191]. In vitro experiments in lung cancer cell lines point to effects of tobacco smoke exposure on expression of methyltransferases [149] and on histone modifications [113, 148]. Promoter hypermethylation and other epigenetic changes have been identified in lung tumors (reviewed in [203]). Examples of tumor suppressor genes upregulated in lung cancer include *p53*, *p16*, and *RB*. Silencing of the *p16* tumor suppressor gene by promoter methylation plays an important role in smoking-related lung cancer, and radionuclides in cigarette smoke may explain the phenomenon of *p16* inactivation by promoter hypermethylation [197]. Inactivation of another tumor suppressor gene, *FHIT*, also occurs very often in aerodigestive tumors, in particular, lung and esophageal cancer, and methylation of the *FHIT* gene promoter is associated with transcriptional inactivation and is strongly linked to cigarette smoking [190]. Belinsky and colleagues showed repeatedly that promoter hypermethylation in *p16* and *DAPK1* is frequently seen in bronchial epithelium from lung cancer current smokers but not in bronchial epithelium from never-smoker lung cancer patients [20, 24]. Aberrant tumor methylation has also been associated with changes in prognosis [236]. Promoter hypermethylation of retinoid X receptors (RXRs) is observed in NSCLC. Methylation did not significantly correlate with prognosis; however, the effect of *RXRG* methylation on prognosis of NSCLC patients was significantly different between never- and ever-smokers [138]. Methylation differences have also been reported between NSCLC tissue from patients with or without COPD [237], and between adenocarcinomas with *EGFR* or *K-RAS* mutations, suggesting specific interactions of genetic and epigenetic alterations in lung tumorigenesis [247]. DNA methylation in sputum is considered a promising early detection [19, 21, 232, 251] and risk marker [23]. A first cohort-based study on dietary factors associated with reduced promoter methylation in cells exfoliated from the airway epithelium of smokers identified multivitamins, folate, and green vegetables as protective [229]. Plasma methylation biomarkers are also showing promise for early detection [13, 68]. Long-term exposure to inhalable nickel and arsenic, which predisposes to lung cancer, has recently been shown to be associated with histone H3K4 dimethylation and H3K9 acetylation in a cross-sectional study of steel workers [36]. Importantly a recent study comparing lung cancer patients and controls from a prospective cohort has shown blood levels of 1-carbon metabolism markers to affect methylation levels in blood cells, but failed to find a clear association between DNA methylation of the five genes investigated and case-control status [255].

13.5.2.2 Head and Neck Cancer

There is evidence from case-only studies that gene specific hypermethylation has also been observed in head and neck squamous cell carcinoma (HNSCC) and has been associated with smoking [122, 157] as well as clinical prognosis [1]. One recent study pointed to age-dependence of an observed association between *p16CDKN2A* methylation and *DNMT3B* polymorphism in HNSCC specimens [71]. Recent evidence from methylation analysis of HNSCC suggests that global

hypomethylation observed in tumors and gene specific methylation is associated in a sequence dependent manner [194]. In a population-based case-control study ($n = 1{,}039$) MIR196A2 genotype was implicated in both susceptibility and prognosis of HNSCC [45].

13.5.3 Epigenetic Epidemiology of Hormone-Related Cancers

13.5.3.1 Prostate Cancer

There is increasing evidence that epigenetic aberrations could be used for risk assessment and early detection for prostate cancer. The best studied epigenetically modified gene in prostate cancer is *GSTP1*, encoding a protein involved in xenobiotic metabolism, protecting DNA against free radical damage. More than 50 independent studies have shown that *GSTP1* is hypermethylated in up to 90% of all prostate cancers, thus making epigenetic modifications a more frequent event than genetic alterations in prostate cancer (reviewed in [17, 187]).

Evidence that epigenetic aberrations may serve as biomarkers for risk or early detection comes from several studies. For example, DNA hypermethylation of commonly methylated genes in prostate cancer was investigated in DNA isolated from whole blood from prostate cancer patients with and without disease progression and compared in age-matched controls. This study suggested that *GSTP1, RASSF1A, APC*, and *RAR beta* are more likely hypermethylated in circulating cell DNA from cancer patients compared with matched controls [205]. Additionally, an increase in methylation for all four genes was seen from the time of diagnosis to the time of relapse but not at the corresponding follow-up visit of patients without relapse. This indicates that the combination of *GSTP1, RASSF1A, APC*, and *RAR beta* methylation could serve as a marker for early detection with potential relevance for prostate cancer recurrence and progression. Interestingly, *GSTP1* methylation has additionally been found in urine, prostatic secretion, and serum, suggesting that GSTP1 methylation is a promising minimally invasive tool for early detection and diagnosis [85, 86].

miRNAs regulate a wide range of biological processes and thereby play a critical role in the pathogenesis of human cancer including prostate carcinogenesis [6, 196]. Several studies have reported that genetic variants within precursor miRNA may influence the maturation of these miRNAs and thus affect a loss or gain of function. This has been reported for miRNA-146, which was shown to be overexpressed in breast, pancreatic, and prostate cancer [256]. A recent case-control study in a Chinese-Han population showed an association between an SNP (rs2910164) in the miR-146a precursor and prostate cancer ($p = 0.03$) [266], a finding also reported for familial breast cancer susceptibility [250]. In contrast, another population based association study in North Indians could not show a significant association with rs2910164 and prostate cancer but instead found that two other SNPs in hsa-miR499 and hsa-mir-196a2 showed significant association with risk [82].

Prostate tumorigenesis depends upon an interplay between genetic and epigenetic changes. In depth investigations of aberrant methylation patterns present in blood or other body fluids on a genome wide scale in prostate cancer patients and investigation of common variants in precursor miRNAs associated with risk are required to clarify the link between epigenetics, genetics, and predisposition to prostate cancer.

13.5.3.2 Breast Cancer

Epigenetic alterations including promoter specific hypermethylation, global hypomethylation, and deregulation of miRNAs have been well studied in breast carcinogenesis (reviewed in [253]). Decreased methylation of downstream target genes of estrogen receptor (ER) alpha is associated with active ER alpha, a risk factor for breast cancer [263]. DNA methylation status of ER alpha targeting genes (ERT) and Poly comb group target genes (PCGT) has additionally been studied in peripheral blood cell DNA in a case-control study where decreased methylation of ERT genes and increased methylation of PCGT genes were associated with breast cancer risk for 7 of 25 genes investigated [262].

Another recent association study observed a relation between gene body methylation of *ATM* and breast cancer risk, suggesting that the methylation variability is highest at ~1 kb downstream of transcription start site [76]. This study conducted a DNA methylation high density tiling array analysis to identify methylation patterns in peripheral blood DNA of 17 breast cancer susceptibility genes in 14 bilateral breast cancer patients and disease-free age-matched controls. Another case-control study performed in Jewish BRCA1/1 mutation carriers with ($n=41$) and without ($n=48$) breast cancer and controls ($n=89$) investigated promoter methylation in several *BRCA1*/*BRCA2* associated genes including *ATM*, and found no significant methylation of *ATM* at the promoter site [127]. This suggests that differentially methylated intergenic regions rather than *ATM* promoter regions may be suitable as a biomarker for breast cancer risk.

Global hypomethylation is an important feature in breast cancer development and has been proposed to contribute to genomic instability and activation of oncogenes [26, 225]. A study by Choi et al., investigated whether 5-methylcytosine at CpG sites and LINE 1 methylation present in leukocyte DNA from 179 breast cancer cases and 180 controls could mirror the overall methylation activity and the association with breast cancer risk [44]. They found a significantly lower methylation level in leukocyte DNA from cases than in healthy controls ($P=0.001$), suggesting that the global genomic hypomethylation may be a risk factor due to genomic instability, as has been seen for many cancer types. This finding provides evidence that loss of methylation may modify breast cancer risk and thus serve as a non invasive method for identification of women at high risk of developing breast cancer.

Three independent studies investigated the association between SNP genotypes in pre-miRNAs and breast cancer risk in Chinese, German, and Italian populations [80]. The first study was carried out in a case-control study of 1,009 breast cancer

cases and 1,093 controls in Chinese women, evaluating the association of four SNPs: rs2910164, rs2292832, rs11614913, and rs3746444, located in pre-miRNA hsa-mir-146a, hsa-mir-149, hsa-mir-196a, and hsa-mir-499, respectively. The SNPs located in hsa-mir-196a2 and hsa-mir-499 were associated with increased risk of breast cancer [110]. In contradiction to this, Catucci et al., found no association between the variants and risk in German and Italian familial breast cancer cases and the common variant of rs11614913 in miR-196a-2 was associated with decreased breast cancer risk in a case-control study conducted in Connecticut, US [104].

Kontorovich and colleagues investigated SNPs within predicted binding sites for miRNAs expressed in breast cancer/ovarian cancer in a case-control study with BRCA1 mutation carriers. They could reveal an association between CC homozygosity for rs6505162 with increased risk for ovarian cancer and heterozygote SNP carriers of rs11169571 had a 2-fold increase of breast and ovarian cancer [127]. Polymorphisms located in miRNAs may change their property through altering miRNA expression or maturation. Since the sequence complementarity of miRNA target regions is crucial for interaction and proper downregulation, variations within the targeting binding sites by SNPs could disrupt the miRNA-mRNA interaction and thus result in diverse functional consequences.

The epigenome of breast cancer has been intensively studied resulting in the identification of several hypermethylated genes that may serve as biomarkers for disease progression, survival, and early detection. Global hypomethylation of repetitive elements as well as polymorphisms in miRNA may serve as risk markers. However, in the future large-scale epidemiological studies are required to clarify if these epigenetic events identified in retrospective case-control studies do indeed have predisposing potential and thus could be used to identify individuals at high risk.

13.5.3.3 Ovarian Cancer

Ovarian carcinoma is often detected at a late and aggressive stage since there are no obvious symptoms at early stages. This makes sensitive screening markers for early detection and risk highly desirable. The Hox gene family has been studied in detail in many different cancers and DNA promoter methylation of *HOXA9, HOXA10*, and *HOXA11* is a frequent event in ovarian cancer [74, 200]. In one study methylation of *HOXA9* was highly significantly ($p < 0.0001$) increased in normal endometrium from pre-menopausal Japanese women with ovarian cancer compared to an age-matched control group [261]. A promising study linking epidemiology with epigenetics was performed using genome-wide DNA methylation profiling in a case-control cohort, analyzing over 27,000 CpG sites in blood derived DNA from 148 healthy individuals, 113 pre-treatment, and 122 post-treatment ovarian cancer cases. Among the CpG sites investigated, 2,714 CpGs, classified as cancer diagnostic CpGs, were found to be differently methylated in blood between pre-treatment cases and healthy individuals. In order to identify CpG sites which may have predisposing features or early detection potential for ovarian cancer, they compared CpG methylation from 70 post-treatment ovarian cancer patients who did not have

recurrent disease at the time when blood was drawn with healthy controls and performed a Surrogate Variable Analysis framework. 84 CpG sites could discriminate between the post-treated cases and healthy controls. The CpGs were further deconvoluted from the diagnostic CpGs and resulted in 18 potentially predisposing risk CpG sites ($P=0.003$). Among them, 11 showed hypermethylation and 7 showed hypomethylation. *TSG101*, a reported tumor suppressor gene and breast cancer predisposing gene, was among the candidates [245]. These results suggest that aberrant DNA methylation patterns detected in peripheral blood cells from cancer patients could indicate a predisposing epigenetic event that could serve as a marker for risk, but careful confirmation of such findings in blood samples collected prior to disease onset will be necessary. DNA methylation as a prognostic and screening marker for ovarian cancer has been shown for a subset of genes where hypermethylation was found in tumor tissue and correlated with DNA methylation patterns in serum; *SFRP1,2,4,5, SOX1, PAX1, and LMX1A* [231]. Additionally, *BRCA1* hypermethylation has been found at elevated levels in serum, plasma, and peritoneal fluid from ovarian cancer patients [114], but the clinical usefulness of such a marker is limited by its lack of specificity.

13.5.3.4 Endometrial Cancer

There are two types of endometrial cancers. Type I is estrogen dependent, positive for both estrogen and progesterone receptors and occurs mainly in pre- or perimenopausal women, while Type II occurs mainly in postmenopausal women and is hormone independent. It is thought that the two types have different molecular mechanisms. Aberrant DNA methylation has been reported to be more common in Type I endometrial cancer than in Type II [264]. A possible explanation for this has been proposed by Zhou and coworkers who observed that Type I endometrial tumors have a 2–4-fold higher DNMT1 and DNMT3B level than in normal endometrium and Type II tumors have a 2–4-fold reduction in relation to normal endometrial tissue [116].

Type I endometrial cancer development depends partly on defects in the DNA repair system. *MLH1* is aberrantly methylated in atypical endometrial hyperplasia, but not in normal endometrium, suggesting that MLH1 may be involved in endometrial carcinogenesis [15].

Endometrial cancer risk factors include hormone exposure, in particular estrogen. The protein expression of the estrogen receptor and progesterone receptor is reduced in endometrial cancer and much effort has been put into elucidating the association between promoter methylation of the estrogen receptor and progesterone receptor and endometrial cancer, however with contradictory results [107, 193, 207, 218]. The relevance of chromatin structure and histone modifications on expression of the ER and PR genes have not been explored in detail and further studies are required to understand this epigenetic deregulation in endometrial cancer.

Epigenetic biomarkers for early detection of endometrial cancer have been suggested by a few studies. Fliegl et al., investigated the methylation pattern of

five genes known to be hypermethylated in tumor tissue (*CDH13, HSPA2, MLH1, RASSF1A*, and *SOCS2*) in DNA isolated from vaginal secretion using the MethylLight methylation detection method. The outcome of this small case-control study (15 cases, 109 controls) revealed that cases had three or more methylated genes and 91% of controls had no or fewer than three genes methylated [73]. However, a larger prospective study would be needed to determine whether these genes have prognostic value, and whether vaginal secretion could be used as a surrogate tissue for aberrant methylation and be applied in early detection screening for endometrial cancer.

13.5.3.5 Cervical Cancer

Aberrant promoter methylation of *FHIT* and *RASSF1A* gene was significantly associated with cervical cancer ($P<0.01$) in a study in a North Indian population with 60 cervical cancer tissue biopsies and 23 healthy control samples with normal cervical cytology [178]. Promoter DNA methylation of *MYOD1, CDH1*, and *CDH13* found in tumor derived DNA in serum from patients with cervical cancer has been associated with poorer disease outcome, indicating that aberrant methylation patterns may be used clinically as a prognostic factor [172].

Human papilloma virus (HPV) infection is the primary cervical cancer risk factor. However, HPV infection alone is insufficient for transforming a normal cell to malignancy, suggesting that other factors such as genetic and epigenetic modifications may contribute to disease progression. The function of the viral proteins HPV E6 and E7 oncoproteins as modulators of cell transformation includes inactivation of tumor suppressors thus making it possible for the virus to avoid cell cycle checkpoints and thereby turning a normal cell into a malignant cell. E6 interferes with and reduces the expression of miRNA-218 and therefore affects downstream targets such as Laminin-5 beta 3 (*LAMB3*), a marker for cervical invasiveness [159, 222]. In a large case-control study with 703 cases and 713 controls in Chinese women, 2 SNPs in pri-miR218 and *LAMB3* were associated with risk of cervical cancer with a dose response effect dependent on the risk alleles for these SNPs [275]. Furthermore, CpG methylation seems to target HPV16 DNA as a cellular defense mechanism against infection. A study by Badal et al., reported that methylation patterns in Long Control Regions (LCRs) and promoter regions of oncogene *E6* decreased with progression of the infection. This study suggests that decreased methylation of promoter and enhancer regions of the oncogenes *E6* and *E7*, and *LCR*, is associated with disease progression [12]. In contrast, another study did not find a correlation between increased methylation and disease progression with significantly higher methylation of the *LCR* region in cervical cancer in comparison to asymptomatic patients [117]. Polymorphisms in methyl group metabolizing enzymes such as MTHFR and MS have been suggested to influence DNA methylation in colorectal, lung, and breast cancer [189] and later also in cervical cancer [120]. This study revealed that hypermethylation of the repair gene *MGMT* was significantly decreased in cervical cancer patients with the *MTHFR* 677 CT genotype. Genetic inter-individual variations in combination with aberrant methylation of crucial genes

involved in cervical cancer development may play an important role for cervical cancer predisposition and progression and thus indicates a great potential for the use of epigenetic markers in a clinical setting and for risk assessment.

13.5.4 Epigenetic Epidemiology in Leukemia

Leukemia is a heterogenous disease with four main disease subtypes: acute lymphatic leukemia (ALL), chronic lymphatic leukemia (CLL), acute myeloid leukemia, (AML) and chronic myeloid leukemia. With the exception of ALL, the risk for development of leukemia increases with age. Several causes for the development of leukemias have been identified. The following sections will indicate advances and strategies to shed light on epigenetic modifications as potential risk factor and as treatment modifier in the different classes of leukemia and the difficulties in designing case-control studies evaluating the role of genetics and epigenetics.

13.5.4.1 Acute Leukemias

Chemical exposure (benzene, pesticides), radiation exposure (therapeutic, nontherapeutic), and chemotherapeutic treatment are known risk factors for developing acute leukemia. Discovery of genetic disorders as predisposition factors contributed to advances in understanding the underlying oncogenic processes. Abnormal karyotypes, caused by translocations and deletions, result in fusion genes, which influence the differentiation of the respective blood cells [40], and are used to determine clinical subtypes of AML and ALL [46, 48]. In addition aberrant silencing by epigenetic mechanisms leads to changed expression patterns interfering with hematopoiesis. The origin of the disease often appears to causally involve a combination of different layers of inherited information and the interplay between these is poorly understood.

Genome-wide association studies find polymorphisms in genes involving the epigenetic machinery as significantly associated with developing acute leukemia, especially during childhood. The translocation t(11;16)(q23;p13.3) of the mixed lineage leukemia (MLL) gene with the CREB-binding protein (CBP), which is a histone acetyl transferase (HAT), is a frequent rearrangement in acute leukemias and associated with poor prognosis [53, 142, 226]. Gene disruption or polymorphisms of epigenetic genes, like miRNAs and HATs, are very common in acute leukemias [28, 146]. Mutations in DNMT3a and EZH2, a polycomb repressor complex member, have been reported in AML [63, 141, 254]. Altogether this leads to the conclusion that DNA methylation and chromatin remodeling are strongly involved in etiology of acute leukemias.

Case-control studies aim to detect risk markers in non-affected tissues using non invasive methods. To avoid a bias, e.g., by clonal selection of epigenetically altered tumor cells, a reference tissue other than blood would be desirable for investigation of epigenetic host factors in acute leukemia. One additional problem is that the results in a blood methylation analysis depend highly on the assay used and the

distribution of blood cell types [265], which can vary. A population-based study determining levels of DNA methylation in a group of patients with a variety of leukemias compared blood methylation of the *bone morphogenic protein 6* (*BMP6*) in patients with peripheral blood monocyte (PBMC) methylation from healthy donors. *BMP6* was hypermethylated only in adult T-cell leukemias, especially an aggressive subtype, [242], which indicates the importance of stratifying for histology. The combination of epidemiologic factors with epigenetic factors appears promising as epigenetic alterations, especially age-dependent DNA methylation changes, have been detected in various cancer-related genes in a mouse model [153], and several studies suggest the potential influence of genetic factors on epigenetics and vice versa.

13.5.4.2 Chronic Leukemias

Chronic leukemia can be divided into the two subclasses: chronic myeloid lymphoma (CML) and chronic lymphocytic leukemia, which is the most frequent leukemia in the western world [228]. Both malignancies are characterized by slow progress during disease and a low incidence rate in children or young adults.

In a large genome-wide association study genetic polymorphisms were connected with CLL and several risk loci were identified [57]. Characterization of 506 CLL patient samples could show the impact of karyotypes, namely deletions in 6q21, 11q22.3, 13q14, and 17p13 (TP53 locus) and a trisomy of chromosome 12 on survival [89]. Additionally, those aberrations could be linked to Immunoglobulin heavy-chain variable-region (IgVh) mutation status which is related to disease development. ZAP-70 expression, which is associated with an unfavorable course of disease, is linked to the IgVh unmutated status [49]. *ZAP-70* promoter methylation is linked to the downregulation of ZAP-70 [47], and additionally ZAP-70 expression was shown to predict methylation levels of important genes, e.g., *Sall1*, *MLL*, *APP*, or of long interspersed nuclear element sequences (LINE-1) which serves as a marker for global methylation [14, 69, 246].

CML is strongly associated with a rearrangement in a 5.8-kilobase stretch of chromosome 22 which occurs in the vast majority of patients [250]. This translocation results in the *BCR/ABL1* fusion oncogene producing a gene product with tyrosine kinase activity. In healthy individuals two active copies of the *ABL1* gene are present. It could be shown in samples of 201 patients that the remaining functional copy in CML patients is not methylated during the chronic stage, is partially methylated in the accelerated, and completely methylated in the blast crisis phase [11].

Chronic leukemias are diseases which seem to be strongly influenced by epigenetic factors, especially DNA methylation. Epigenetic marks clearly determine etiology and course of disease. In the future, genome-wide screens in large cohorts will be powerful tools in identifying epigenetically silenced genes, which influence hematologic diseases. Such insights may also help in the development of effective treatments.

13.5.5 *Epigenetic Epidemiology of Other Tumors*

13.5.5.1 Liver

Hepatocellular carcinoma (HCC) ranks among the most common cancers worldwide, and various risk factors such as hepatitis B virus and hepatitis C virus infection are known to participate in its development [84]. Genome-wide or global hypomethylation, previously identified as a marker for HCC [144], and/or promoter CpG hypermethylation [134, 136, 233, 273] correlate with clinical outcome (survival) in HCC [35]. Specific methylation signatures in HCC exist which are associated with environmental risk factors, survival, and tumor staging; one such methylation signature obtained by using array technology, pyrosequencing, and quantitative polymerase chain reaction comprises various promoter CpG sites and correctly classified tumor and surrounding normal liver tissue, thus a small set of predictor CpG sites may be sufficient for future diagnostic and clinical use [97]. Interestingly, genome-wide DNA methylation profiling using bacterial artificial clone arrays also revealed that few such clones may suffice as an excellent indicator set for HCC risk estimation and prognosis (reviewed in: [118]). Further epigenetic markers linked to HCC include: *LINE-1* hypomethylation in tissue or serum of HCC patients [125] as independent marker of poor prognosis [241]; altered promoter methylation of tumor suppressor genes and oncogenes [126]; histone deacetylase SNPs as potential predictors for tumor recurrence [269]; epigenetic silencing of *Tip30* expression by CpG island DNA hypermethylation is associated with poor prognosis in HCC patients [151]; in rodents, epigenetically silenced *methionine adenosyltransferase* (*MAT1A*) expression is associated with HCC development [163]; and methylation of *dickkopf-3 (DKK3)* was shown to occur early in HCC development and may qualify as prognostic risk factor in HCC [268]. Detection of epigenetically modified biomarkers in serum, e.g., of promoters of tumor suppressor genes [274] are attractive goals for early diagnosis of HCC especially since an epigenetically dysregulated promoter of one gene can influence methylation changes in other genes as shown in a mouse model of HCC [99, 140, 198]. Given evidence that epigenetic aberrations are involved in HCC development (see [94] for review) it appears that investigation of epigenetic markers as possible modulators also of risk of hepatic carcinoma holds promise.

13.5.5.2 Brain

Global DNA hypomethylation is frequent (80%) in primary glioblastomas [31], either in repetitive sequences *(Sat2)* possibly due to DNA methyltransferase 3b (DNMT3b) dysfunction [61] or in single-copy genes such as *MAGEA1* [31, 271]. Mouse models indicate that global hypomethylation is sufficient to initiate tumorigenesis [81]. For glioblastoma multiforme (GBM, reviewed in [160, 173]), promoter CpG island hypermethylation is frequent, e.g., in tumor suppressor (*EMP3*), DNA

repair enzymes (*MGMT1*; [67]), and growth factor genes ($TGF\beta - Smad$ and *PDGF-B*; [30]) and can be associated with poor prognosis [30]. *MGMT* function is lost upon CpG island promoter methylation in brain tumors [67], and this epigenetic silencing is linked to survival time in GBM [66, 92, 93]. *MGMT* promoter methylation is also linked to a hypermutator phenotype during glioblastoma therapy, which has potential clinical implications [177], and its methylation status may change during development of GBM or relapses [29, 161]. In addition, *MGMT* methylation was recently shown to be an independent prognostic factor for overall and progression-free survival in elderly GBM patients [165]. More recently, in secondary glioblastomas *RASSF10* methylation was reported as independent prognostic factor associated with worst-progression-free and overall survival [100]. Further, a direct or indirect link between genetic and epigenetic modifications was recently noted: *MGMT* silencing by hypermethylation facilitates G to A mutations [65, 66], and such a mutated *IDH1* status was found to be an independent prognostic factor in gliomas [206]. Recently, promoter methylation of the suppressor of cytokine signalling gene *SOC3* was found to be associated with outcome in GBM patients [162]. For medullablastoma, aberrant promoter methylation of several promoter sequences also correlated with poor prognosis [77, 171].

In summary, existing evidence links defined epigenetic alterations with brain tumor incidence thus supporting the idea of an association of epigenetics and epidemiology of brain tumors.

13.5.5.3 Stomach

In a recent study, up regulation of *SOX4* was found to be inversely associated with the epigenetic silencing of *miR-129-2* in gastric cancer and inversely correlated with patient survival time (70 cases; [217]). Further, the methylation level in several gastric cancer candidate CpG islands was higher in gastric cancer mucosa suggesting a possible association with gastric cancer risk [154, 175, 176]. Recent studies also linked differential methylation status of *Cystatin M, BNIP3,* and *DAPK* promoters with survival of gastric cancer patients [42, 234].

13.5.5.4 Skin

Many epigenetic alterations are known for skin cancerous tissue (reviewed in [87]). In SCC, hypermethylation of *RASSF2A, MGMT*, and *TSP-1* gene promoters was identified, and *TSP-1* hypermethylation was reported as a bad prognosis factor in vulvar SCC [88]. In malignant melanoma, aberrant CpG island methylation of tumor-related genes is associated with methylation of noncoding methylated-in-tumor (*MINT*) loci, one of which (*MINT31*) was characterized as a predicting factor for tumor progression and survival [240]. In contrast, in cutaneous T-cell lymphomas, methylation profiles of candidate promoters ($p14^{ARF}$, $p16^{INK4A}$, *and* $p15^{INK4B}$) did not correlate with prognosis [131]. In oral SCCs, promoter hypermethylation silenced the

expression of *IKKα*, one of the IkB kinases linked to NFκB-mediated transcriptional activity, and this epigenetic alteration was associated with an unfavorable prognosis [152]. Recent evidence linked microRNAs, which downregulate DNA methyltransferases *DNMT3A* and *DNMT3B*, with survival of skin cancer patients [180].

13.5.5.5 Kidney

In renal cancer, epigenetic silencing of various Wnt pathway genes has been reported. Polymorphisms within the Wnt antagonists of the *Dickkopf (DKK)* gene family have been associated with this cancer, and a specific *DKK2* polymorphism may be a predictor for survival of kidney cancer patients [101, 102]. Other studies on renal clear cell carcinoma (RCC) cases found that clustering is associated with regional DNA hypermethylation suggesting such epigenetic cluster analysis as a prognostic indicator for RCC patients [9, 10]. At present it is controversial whether epigenetic alterations in the von *Hippel-Lindau (VHL)* gene have prognostic value in clear cell RCC [188, 224]. In addition, histone modifications such as histone acetylation as epigenetic regulators of gene activity have been shown to be of prognostic value in kidney cancer [170].

13.5.5.6 Bladder

The candidate tumor suppressor gene *myopodin* is aberrantly methylated in several cancer types including bladder cancer [38] and its hypermethylation was found to be associated with increased recurrence and shorter survival in bladder cancer patients [5]. Also, *polyamine modulated factor-1 (PMF*1*)* was identified to be epigenetically modified in bladder cancer, and its promoter methylation status correlates with tumor progression of bladder cancer [2]. A more recent population-based incident case-control association study analyzing peripheral blood DNA from 285 cases and 465 controls indicated that *LINE1* hypomethylation may be an important if unspecific marker of bladder cancer risk, and the association was stronger among women (OR = 2.48) than men (OR = 1.47) [264]. Further epigenetically modified candidate targets associated with bladder tumor stage or patient survival include *P16INK4a, P14ARF*, and several other promoters [3, 156].

13.5.5.7 Testis

LINE-1 methylation patterns may be associated with cancer risk [166]. Other epigenetic candidate markers for testicular cancer include (reviewed in: [83]): *RASSF1A* (tumor suppressor, promoter frequently hypermethylated in testis cancer) [105]; altered global methylation profiles (seminomas: hypomethylated, practically no CpG hypermethylated, whereas nonseminomas are less hypomethylated and have variable CpG island methylation) [223]; increased expression of *DNMT3A; and DNMT3B* in embryonal carcinoma subtypes [4, 221].

13.5.5.8 Pancreas

Epigenetic epidemiology studies have thus far not been reported for pancreatic carcinoma. However, methylation profiling of CpG island promoter sites in normal pancreas, pancreatic tumors, and cancer cell lines indicate frequent aberrant DNA methylation in pancreatic cancer [239]. Many of these sites were known from other cancer types [64] and include genes for, e.g., cadherins, growth factor binding proteins, imprinted genes, and transcription factors. Another study comparing the genome-wide promoter methylation profile of a human pancreatic cancer cell line (PANC-1) with that of a non-neoplastic cell line (HPDE) also detected highly significant differences in methylation which were validated for some genes in normal and pancreatic cancer tissue [182]. An earlier prospective study with cases and controls found that plasma levels of folate, vitamin B6, vitamin B12, and homocysteine are not associated with the risk of pancreatic cancer [210], which may argue against a role of methylation in epigenetic epidemiology of pancreatic adenocarcinoma (PDAC). However, epigenetic silencers such as polycomb proteins were found upregulated in PDAC tissue which may link them to pancreatic tumor development [158]. In addition, varying expressions of epigenetic modifiers including *DNMT1, HDAC1*, methyl-CpG binding proteins, and miRNAs such as miR-373 are associated with staging or survival in cases of this tumor type [43, 258].

13.6 Outlook

Important technological developments will advance and considerably change the field of epigenetic epidemiology of cancer. The fact that quantitative instead of qualitative methylation measurements are becoming standard even for high-throughput analyses [202] will much refine future analyses. Methylation next generation sequencing approaches, while not yet affordable in high throughput, will provide important new targets.

While the majority of studies to date relating to epigenetic cancer biomarkers have been focussing on CpG island promoter methylation, in the future whole genome methylation and histone modifications will become more prominent. Repetitive element hypomethylation in leukocyte DNA was recently investigated within the Normative Aging Study cohort, and associated with increased incidence and mortality from cancer [276]. While this is interesting the clinical use of this phenomenon is limited due to its lack of specificity. The development of quantitative detection of DNA methylation in minute amounts of DNA [185] will be particularly important in the context of epigenetic early diagnosis markers.

The fact that within case-control studies, sample collection occurs after disease development makes it impossible to determine whether epigenetic associations observed are causal or possibly related to tumor development. Cohort studies will thus be invaluable in future investigations to confirm any associations observed. Population based cohorts with good follow-up allow the investigation of samples

from individuals with tumors that were diagnosed after sample collection. Another important challenge for future investigations will be to identify the doses and time-window in an individual lifetime when particular exposures can lead to epigenetic changes [260]. Given the reversibility of epigenetic alterations and the importance of epigenetic mechanisms for gene regulation, cancer chemoprevention by targeting the epigenome is an obvious goal, and there is evidence that this is becoming a real possibility [111]. In this context especially, the identification of epigenetic risk factors and refinement of currently available cancer risk models has real public health potential.

References

1. Ai L, Vo QN, Zuo C, Li L, Ling W, Suen JY, Hanna E, Brown KD, Fan CY (2004) Ataxia-telangiectasia-mutated (ATM) gene in head and neck squamous cell carcinoma: promoter hypermethylation with clinical correlation in 100 cases. Cancer Epidemiol Biomarkers Prev 13:150–156
2. Aleman A, Cebrian V, Alvarez M, Lopez V, Orenes E, Lopez-Serra L, Algaba F, Bellmunt J, Lopez-Beltran A, Gonzalez-Peramato P, Cordon-Cardo C, Garcia J, del Muro JG, Esteller M, Sanchez-Carbayo M (2008) Identification of PMF1 methylation in association with bladder cancer progression. Clin Cancer Res 14:8236–8243
3. Ali Hosseini S, Sobti RC, Malekzadeh K, Singh SK, Joshi K (2010) Frequency of P16INK4a and P14ARF genes methylation and its impact on bladder cancer cases in north Indian population. Dis Markers 28:361–368
4. Almstrup K, Hoei-Hansen CE, Nielsen JE, Wirkner U, Ansorge W, Skakkebaek NE, Rajpert-De Meyts E, Leffers H (2005) Genome-wide gene expression profiling of testicular carcinoma in situ progression into overt tumours. Br J Cancer 92:1934–1941
5. Alvarez-Mugica M, Cebrian V, Fernandez-Gomez JM, Fresno F, Escaf S, Sanchez-Carbayo M (2010) Myopodin methylation is associated with clinical outcome in patients with T1G3 bladder cancer. J Urol 184:1507–1513
6. Ambs S, Prueitt RL, Yi M, Hudson RS, Howe TM, Petrocca F, Wallace TA, Liu CG, Volinia S, Calin GA, Yfantis HG, Stephens RM, Croce CM (2008) Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. Cancer Res 68:6162–6170
7. Anker P, Lyautey J, Lederrey C, Stroun M (2001) Circulating nucleic acids in plasma or serum. Clin Chim Acta 313:143–146
8. Anker P, Mulcahy H, Chen XQ, Stroun M (1999) Detection of circulating tumour DNA in the blood (plasma/serum) of cancer patients. Cancer Metastasis Rev 18:65–73
9. Arai E, Ushijima S, Fujimoto H, Hosoda F, Shibata T, Kondo T, Yokoi S, Imoto I, Inazawa J, Hirohashi S, Kanai Y (2009) Genome-wide DNA methylation profiles in both precancerous conditions and clear cell renal cell carcinomas are correlated with malignant potential and patient outcome. Carcinogenesis 30:214–221
10. Arai E, Ushijima S, Tsuda H, Fujimoto H, Hosoda F, Shibata T, Kondo T, Imoto I, Inazawa J, Hirohashi S, Kanai Y (2008) Genetic clustering of clear cell renal cell carcinoma based on array-comparative genomic hybridization: its association with DNA methylation alteration and patient outcome. Clin Cancer Res 14:5531–5539
11. Asimakopoulos FA, Shteper PJ, Krichevsky S, Fibach E, Polliack A, Rachmilewitz E, Ben-Neriah Y, Ben-Yehuda D (1999) ABL1 methylation is a distinct molecular event associated with clonal evolution of chronic myeloid leukemia. Blood 94:2452–2460
12. Badal V, Chuang LS, Tan EH, Badal S, Villa LL, Wheeler CM, Li BF, Bernard HU (2003) CpG methylation of human papillomavirus type 16 DNA in cervical cancer cell lines and in

clinical specimens: genomic hypomethylation correlates with carcinogenic progression. J Virol 77:6227–6234

13. Bailey VJ, Zhang Y, Keeley BP, Yin C, Pelosky KL, Brock M, Baylin SB, Herman JG, Wang TH (2010) Single-tube analysis of DNA methylation with silica superparamagnetic beads. Clin Chem 56:1022–1025
14. Baird DM (2010) Variation at the TERT locus and predisposition for cancer. Expert Rev Mol Med 12:e16
15. Banno K, Yanokura M, Susumu N, Kawaguchi M, Hirao N, Hirasawa A, Tsukazaki K, Aoki D (2006) Relationship of the aberrant DNA hypermethylation of cancer-related genes with carcinogenesis of endometrial cancer. Oncol Rep 16:1189–1196
16. Barlesi F, Giaccone G, Gallegos-Ruiz MI, Loundou A, Span SW, Lefesvre P, Kruyt FA, Rodriguez JA (2007) Global histone modifications predict prognosis of resected non small-cell lung cancer. J Clin Oncol 25:4358–4364
17. Bastian PJ, Yegnasubramanian S, Palapattu GS, Rogers CG, Lin X, De Marzo AM, Nelson WG (2004) Molecular biomarker in prostate cancer: the role of CpG island hypermethylation. Eur Urol 46:698–708
18. Baylin SB, Ohm JE (2006) Epigenetic gene silencing in cancer – a mechanism for early oncogenic pathway addiction? Nat Rev Cancer 6:107–116
19. Belinsky SA (2004) Gene-promoter hypermethylation as a biomarker in lung cancer. Nat Rev Cancer 4:707–717
20. Belinsky SA (2005) Silencing of genes by promoter hypermethylation: key event in rodent and human lung cancer. Carcinogenesis 26:1481–1487
21. Belinsky SA, Grimes MJ, Casas E, Stidley CA, Franklin WA, Bocklage TJ, Johnson DH, Schiller JH (2007) Predicting gene promoter methylation in non-small-cell lung cancer by evaluating sputum and serum. Br J Cancer 96:1278–1283
22. Belinsky SA, Klinge DM, Dekker JD, Smith MW, Bocklage TJ, Gilliland FD, Crowell RE, Karp DD, Stidley CA, Picchi MA (2005) Gene promoter methylation in plasma and sputum increases with lung cancer risk. Clin Cancer Res 11:6505–6511
23. Belinsky SA, Liechty KC, Gentry FD, Wolf HJ, Rogers J, Vu K, Haney J, Kennedy TC, Hirsch FR, Miller Y, Franklin WA, Herman JG, Baylin SB, Bunn PA, Byers T (2006) Promoter hypermethylation of multiple genes in sputum precedes lung cancer incidence in a high-risk cohort. Cancer Res 66:3338–3344
24. Belinsky SA, Palmisano WA, Gilliland FD, Crooks LA, Divine KK, Winters SA, Grimes MJ, Harms HJ, Tellez CS, Smith TM, Moots PP, Lechner JF, Stidley CA, Crowell RE (2002) Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. Cancer Res 62:2370–2377
25. Bennett KL, Romigh T, Arab K, Teresi RE, Tada Y, Eng C, Plass C (2009) Activator protein 2 alpha (AP2alpha) suppresses 42 kDa C/CAAT enhancer binding protein alpha (p42(C/EBPalpha)) in head and neck squamous cell carcinoma. Int J Cancer 124:1285–1292
26. Bernardino J, Roux C, Almeida A, Vogt N, Gibaud A, Gerbault-Seureau M, Magdelenat H, Bourgeois CA, Malfoy B, Dutrillaux B (1997) DNA hypomethylation in breast cancer: an independent parameter of tumor progression? Cancer Genet Cytogenet 97:83–89
27. Bollati V, Baccarelli A, Hou L, Bonzini M, Fustinoni S, Cavallo D, Byun HM, Jiang J, Marinelli B, Pesatori AC, Bertazzi PA, Yang AS (2007) Changes in DNA methylation patterns in subjects exposed to low-dose benzene. Cancer Res 67:876–880
28. Borrow J, Stanton VP Jr, Andresen JM, Becher R, Behm FG, Chaganti RS, Civin CI, Disteche C, Dube I, Frischauf AM, Horsman D, Mitelman F, Volinia S, Watmore AE, Housman DE (1996) The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. Nat Genet 14:33–41
29. Brandes AA, Franceschi E, Tosoni A, Bartolini S, Bacci A, Agati R, Ghimenton C, Turazzi S, Talacchi A, Skrap M, Marucci G, Volpin L, Morandi L, Pizzolitto S, Gardiman M, Andreoli A, Calbucci F, Ermani M (2010) O(6)-methylguanine DNA-methyltransferase methylation status can change between first surgery for newly diagnosed glioblastoma and second surgery for recurrence: clinical implications. Neuro Oncol 12:283–288

30. Bruna A, Darken RS, Rojo F, Ocana A, Penuelas S, Arias A, Paris R, Tortosa A, Mora J, Baselga J, Seoane J (2007) High TGFbeta-Smad activity confers poor prognosis in glioma patients and promotes cell proliferation depending on the methylation of the PDGF-B gene. Cancer Cell 11:147–160
31. Cadieux B, Ching TT, VandenBerg SR, Costello JF (2006) Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetrahydrofolate reductase allele status, and increased proliferation. Cancer Res 66:8469–8476
32. Cairns P (2007) Gene methylation and early detection of genitourinary cancer: the road ahead. Nat Rev Cancer 7:531–543
33. Cairns P, Esteller M, Herman JG, Schoenberg M, Jeronimo C, Sanchez-Cespedes M, Chow NH, Grasso M, Wu L, Westra WB, Sidransky D (2001) Molecular detection of prostate cancer in urine by GSTP1 hypermethylation. Clin Cancer Res 7:2727–2730
34. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, Aldler H, Rattan S, Keating M, Rai K, Rassenti L, Kipps T, Negrini M, Bullrich F, Croce CM (2002) Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci USA 99:15524–15529
35. Calvisi DF, Ladu S, Gorden A, Farina M, Lee JS, Conner EA, Schroeder I, Factor VM, Thorgeirsson SS (2007) Mechanistic and prognostic significance of aberrant methylation in the molecular pathogenesis of human hepatocellular carcinoma. J Clin Invest 117:2713–2722
36. Cantone L, Nordio F, Hou L, Apostoli P, Bonzini M, Tarantini L, Angelici L, Bollati V, Zanobetti A, Schwartz J, Bertazzi PA, Baccarelli A (2011) Inhalable metal-rich air particles and histone H3K4 dimethylation and H3K9 acetylation in a cross-sectional study of steel workers. Environ Health Perspect 119:964–969
37. Cebrian A, Pharoah PD, Ahmed S, Ropero S, Fraga MF, Smith PL, Conroy D, Luben R, Perkins B, Easton DF, Dunning AM, Esteller M, Ponder BA (2006) Genetic variants in epigenetic genes and breast cancer risk. Carcinogenesis 27:1661–1669
38. Cebrian V, Alvarez M, Aleman A, Palou J, Bellmunt J, Gonzalez-Peramato P, Cordon-Cardo C, Garcia J, Piulats JM, Sanchez-Carbayo M (2008) Discovery of myopodin methylation in bladder cancer. J Pathol 216:111–119
39. Chan TL, Yuen ST, Kong CK, Chan YW, Chan AS, Ng WF, Tsui WY, Lo MW, Tam WY, Li VS, Leung SY (2006) Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. Nat Genet 38:1178–1183
40. Chen J, Odenike O, Rowley JD (2010) Leukaemogenesis: more than mutant genes. Nat Rev Cancer 10:23–36
41. Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K, Guo J, Zhang Y, Chen J, Guo X, Li Q, Li X, Wang W, Wang J, Jiang X, Xiang Y, Xu C, Zheng P, Zhang J, Li R, Zhang H, Shang X, Gong T, Ning G, Zen K, Zhang CY (2008) Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res 18:997–1006
42. Chen X, Cao X, Dong W, Xia M, Luo S, Fan Q, Xie J. (2010) Cystatin M expression is reduced in gastric carcinoma and is associated with promoter hypermethylation. Biochem Biophys Res Commun 391(1):1070–1074
43. Chen Y, Gao W, Luo J, Tian R, Sun H, Zou S (2011) Methyl-CpG binding protein MBD2 is implicated in methylation-mediated suppression of miR-373 in hilar cholangiocarcinoma. Oncol Rep 25:443–451
44. Choi JY, James SR, Link PA, McCann SE, Hong CC, Davis W, Nesline MK, Ambrosone CB, Karpf AR (2009) Association between global DNA hypomethylation in leukocytes and risk of breast cancer. Carcinogenesis 30:1889–1897
45. Christensen BC, Avissar-Whiting M, Ouellet LG, Butler RA, Nelson HH, McClean MD, Marsit CJ, Kelsey KT (2010) Mature microRNA sequence polymorphism in MIR196A2 is associated with risk and prognosis of head and neck cancer. Clin Cancer Res 16:3713–3720
46. Claus R, Plass C, Armstrong SA, Bullinger L (2010) DNA methylation profiling in acute myeloid leukemia: from recent technological advances to biological and clinical insights. Future Oncol 6:1415–1431

47. Corcoran M, Parker A, Orchard J, Davis Z, Wirtz M, Schmitz OJ, Oscier D (2005) ZAP-70 methylation status is associated with ZAP-70 expression status in chronic lymphocytic leukemia. Haematologica 90:1078–1088
48. Coustan-Smith E, Mullighan CG, Onciu M, Behm FG, Raimondi SC, Pei D, Cheng C, Su X, Rubnitz JE, Basso G, Biondi A, Pui CH, Downing JR, Campana D (2009) Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. Lancet Oncol 10:147–156
49. Crespo M, Bosch F, Villamor N, Bellosillo B, Colomer D, Rozman M, Marce S, Lopez-Guillermo A, Campo E, Montserrat E (2003) ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. N Engl J Med 348:1764–1775
50. Cropley JE, Suter CM, Beckman KB, Martin DI (2006) Germ-line epigenetic modification of the murine Avy allele by nutritional supplementation. Proc Natl Acad Sci USA 103:17308–17312
51. Cui H, Cruz-Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S, Wu Y, He X, Powe NR, Feinberg AP (2003) Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. Science 299:1753–1755
52. Curtin K, Samowitz WS, Ulrich CM, Wolff RK, Herrick JS, Caan BJ, Slattery ML (2011) Nutrients in folate-mediated, one-carbon metabolism and the risk of rectal tumors in men and women. Nutr Cancer 63(3):357–366
53. Damnjanovic T, Milicevic R, Novkovic T, Jovicic O, Bunjevacki V, Jekic B, Lukovic L, Novakovic I, Redzic D, Milasin J (2010) Association between the methylenetetrahydrofolate reductase polymorphisms and risk of acute lymphoblastic leukemia in Serbian children. J Pediatr Hematol Oncol 32:e148–e150
54. Das S, Foley N, Bryan K, Watters KM, Bray I, Murphy DM, Buckley PG, Stallings RL (2010) MicroRNA mediates DNA demethylation events triggered by retinoic acid during neuroblastoma cell differentiation. Cancer Res 70:7874–7881
55. de Fraipont F, Moro-Sibilot D, Michelland S, Brambilla E, Brambilla C, Favrot MC (2005) Promoter methylation of genes in bronchial lavages: a marker for early diagnosis of primary and relapsing non-small cell lung cancer? Lung Cancer 50:199–209
56. de Vogel S, Bongaerts BW, Wouters KA, Kester AD, Schouten LJ, de Goeij AF, de Bruine AP, Goldbohm RA, van den Brandt PA, van Engeland M, Weijenberg MP (2008) Associations of dietary methyl donor intake with MLH1 promoter hypermethylation and related molecular phenotypes in sporadic colorectal cancer. Carcinogenesis 29:1765–1773
57. Di Bernardo MC, Crowther-Swanepoel D, Broderick P, Webb E, Sellick G, Wild R, Sullivan K, Vijayakrishnan J, Wang Y, Pittman AM, Sunter NJ, Hall AG, Dyer MJS, Matutes E, Dearden C, Mainou-Fowler T, Jackson GH, Summerfield G, Harris RJ, Pettitt AR, Hillmen P, Allsup DJ, Bailey JR, Pratt G, Pepper C, Fegan C, Allan JM, Catovsky D, Houlston RS (2008) A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. Nat Genet 40:1204–1210
58. Diamandis EP, Sidransky D, Laird PW, Cairns P, Bapat B (2010) Epigenomics-based diagnostics. Clin Chem 56:1216–1219
59. Dindot SV, Person R, Strivens M, Garcia R, Beaudet AL (2009) Epigenetic profiling at mouse imprinted gene clusters reveals novel epigenetic and genetic features at differentially methylated regions. Genome Res 19:1374–1383
60. Ebbing M, Bonaa KH, Nygard O, Arnesen E, Ueland PM, Nordrehaug JE, Rasmussen K, Njolstad I, Refsum H, Nilsen DW, Tverdal A, Meyer K, Vollset SE (2009) Cancer incidence and mortality after treatment with folic acid and vitamin B12. JAMA 302:2119–2126
61. Ehrlich M (2003) The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. Clin Immunol 109:17–28
62. Elsheikh SE, Green AR, Rakha EA, Powe DG, Ahmed RA, Collins HM, Soria D, Garibaldi JM, Paish CE, Ammar AA, Grainge MJ, Ball GR, Abdelghany MK, Martinez-Pomares L, Heery DM, Ellis IO (2009) Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. Cancer Res 69:3802–3809

63. Ernst T, Chase AJ, Score J, Hidalgo-Curtis CE, Bryant C, Jones AV, Waghorn K, Zoi K, Ross FM, Reiter A, Hochhaus A, Drexler HG, Duncombe A, Cervantes F, Oscier D, Boultwood J, Grand FH, Cross NC (2010) Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. Nat Genet 42:722–726
64. Esteller M (2007) Epigenetic gene silencing in cancer: the DNA hypermethylome. Hum Mol Genet 16(Spec No 1):R50–R59
65. Esteller M, Corn PG, Baylin SB, Herman JG (2001) A gene hypermethylation profile of human cancer. Cancer Res 61:3225–3229
66. Esteller M, Garcia-Foncillas J, Andion E, Goodman SN, Hidalgo OF, Vanaclocha V, Baylin SB, Herman JG (2000) Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. N Engl J Med 343:1350–1354
67. Esteller M, Hamilton SR, Burger PC, Baylin SB, Herman JG (1999) Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. Cancer Res 59:793–797
68. Esteller M, Sanchez-Cespedes M, Rosell R, Sidransky D, Baylin SB, Herman JG (1999) Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. Cancer Res 59:67–70
69. Fabris S, Bollati V, Agnelli L, Morabito F, Motta V, Cutrona G, Matis S, Grazia Recchia A, Gigliotti V, Gentile M, Deliliers GL, Bertazzi PA, Ferrarini M, Neri A, Baccarelli A (2011) Biological and clinical relevance of quantitative global methylation of repetitive DNA sequences in chronic lymphocytic leukemia. Epigenetics 6:188–194
70. Fan H, Zhang F, Hu J, Liu D, Zhao Z (2008) Promoter polymorphisms of DNMT3B and the risk of colorectal cancer in Chinese: a case-control study. J Exp Clin Cancer Res 27:24
71. Farias LC, Fraga CA, De Oliveira MV, Silva TF, Marques-Silva L, Moreira PR, De-Paula AM, Gomez RS, Guimaraes AL (2010) Effect of age on the association between p16CD-KN2A methylation and DNMT3B polymorphism in head and neck carcinoma and patient survival. Int J Oncol 37:167–176
72. Feinberg AP, Vogelstein B (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature 301:89–92
73. Fiegl H, Gattringer C, Widschwendter A, Schneitter A, Ramoni A, Sarlay D, Gaugg I, Goebel G, Muller HM, Mueller-Holzner E, Marth C, Widschwendter M (2004) Methylated DNA collected by tampons – a new tool to detect endometrial cancer. Cancer Epidemiol Biomarkers Prev 13:882–888
74. Fiegl H, Windbichler G, Mueller-Holzner E, Goebel G, Lechner M, Jacobs IJ, Widschwendter M (2008) HOXA11 DNA methylation – a novel prognostic biomarker in ovarian cancer. Int J Cancer 123:725–729
75. Filkowski JN, Ilnytskyy Y, Tamminga J, Koturbash I, Golubov A, Bagnyukova T, Pogribny IP, Kovalchuk O (2010) Hypomethylation and genome instability in the germline of exposed parents and their progeny is associated with altered miRNA expression. Carcinogenesis 31:1110–1115
76. Flanagan JM, Munoz-Alegre M, Henderson S, Tang T, Sun P, Johnson N, Fletcher O, Dos Santos SI, Peto J, Boshoff C, Narod S, Petronis A (2009) Gene-body hypermethylation of ATM in peripheral blood DNA of bilateral breast cancer patients. Hum Mol Genet 18:1332–1342
77. Fruhwald MC, O'Dorisio MS, Dai Z, Tanner SM, Balster DA, Gao X, Wright FA, Plass C (2001) Aberrant promoter methylation of previously unidentified target genes is a common abnormality in medulloblastomas – implications for tumor biology and potential clinical utility. Oncogene 20:5033–5042
78. Fujita Y, Kojima K, Hamada N, Ohhashi R, Akao Y, Nozawa Y, Deguchi T, Ito M (2008) Effects of miR-34a on cell growth and chemoresistance in prostate cancer PC3 cells. Biochem Biophys Res Commun 377:114–119
79. Gama-Sosa MA, Slagel VA, Trewyn RW, Oxenhandler R, Kuo KC, Gehrke CW, Ehrlich M (1983) The 5-methylcytosine content of DNA from human tumors. Nucleic Acids Res 11:6883–6894

80. Gao LB, Bai P, Pan XM, Jia J, Li LJ, Liang WB, Tang M, Zhang LS, Wei YG, Zhang L (2011) The association between two polymorphisms in pre-miRNAs and breast cancer risk: a meta-analysis. Breast Cancer Res Treat 125:571–574
81. Gaudet F, Hodgson JG, Eden A, Jackson-Grusby L, Dausman J, Gray JW, Leonhardt H, Jaenisch R (2003) Induction of tumors in mice by genomic hypomethylation. Science 300:489–492
82. George GP, Gangwar R, Mandal RK, Sankhwar SN, Mittal RD (2011) Genetic variation in microRNA genes and prostate cancer risk in North Indian population. Mol Biol Rep 38:1609–1615
83. Godmann M, Lambrot R, Kimmins S (2009) The dynamic epigenetic program in male germ cells: its role in spermatogenesis, testis cancer, and its response to the environment. Microsc Res Tech 72:603–619
84. Gomaa AI, Khan SA, Toledano MB, Waked I, Taylor-Robinson SD (2008) Hepatocellular carcinoma: epidemiology, risk factors and pathogenesis. World J Gastroenterol 14:4300–4308
85. Gonzalgo ML, Nakayama M, Lee SM, De Marzo AM, Nelson WG (2004) Detection of GSTP1 methylation in prostatic secretions using combinatorial MSP analysis. Urology 63:414–418
86. Gonzalgo ML, Pavlovich CP, Lee SM, Nelson WG (2003) Prostate cancer detection by GSTP1 methylation analysis of postbiopsy urine specimens. Clin Cancer Res 9:2673–2677
87. Greinert R (2009) Skin cancer: new markers for better prevention. Pathobiology 76:64–81
88. Guerrero D, Guarch R, Ojer A, Casas JM, Mendez-Meca C, Esteller M, Barba-Ramos E, Garcia-Bragado F, Puras A (2011) Differential hypermethylation of genes in vulvar cancer and lichen sclerosus coexisting or not with vulvar cancer. Int J Cancer 128:2853–2864
89. Haferlach C, Dicker F, Schnittger S, Kern W, Haferlach T (2007) Comprehensive genetic characterization of CLL: a study on 506 cases analysed with chromosome banding analysis, interphase FISH, IgV(H) status and immunophenotyping. Leukemia 21:2442–2451
90. Hamamoto R, Furukawa Y, Morita M, Iimura Y, Silva FP, Li M, Yagyu R, Nakamura Y (2004) SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. Nat Cell Biol 6:731–740
91. Hanke M, Hoefig K, Merz H, Feller AC, Kausch I, Jocham D, Warnecke JM, Sczakiel G (2009) A robust methodology to study urine microRNA as tumor marker: microRNA-126 and microRNA-182 are related to urinary bladder cancer. Urol Oncol 28:655–661
92. Hegi ME, Diserens AC, Godard S, Dietrich PY, Regli L, Ostermann S, Otten P, Van Melle G, de Tribolet N, Stupp R (2004) Clinical trial substantiates the predictive value of O-6-methylguanine-DNA methyltransferase promoter methylation in glioblastoma patients treated with temozolomide. Clin Cancer Res 10:1871–1874
93. Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, Kros JM, Hainfellner JA, Mason W, Mariani L, Bromberg JE, Hau P, Mirimanoff RO, Cairncross JG, Janzer RC, Stupp R (2005) MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med 352:997–1003
94. Herceg Z, Paliwal A (2011) Epigenetic mechanisms in hepatocellular carcinoma: how environmental factors influence the epigenome. Mutat Res 727:55–61
95. Herceg Z, Vaissière T (2011) Epigenetic mechanisms and cancer: an interface between the environment and the genome. Epigenetics 6:804–819
96. Herman H, Lu M, Anggraini M, Sikora A, Chang Y, Yoon BJ, Soloway PD (2003) Trans allele methylation and paramutation-like effects in mice. Nat Genet 34:199–202
97. Hernandez-Vargas H, Lambert MP, Le Calvez-Kelm F, Gouysse G, McKay-Chopin S, Tavtigian SV, Scoazec JY, Herceg Z (2010) Hepatocellular carcinoma displays distinct DNA methylation signatures with potential as clinical predictors. PLoS One 5:e9749
98. Hesson LB, Hitchins MP, Ward RL (2010) Epimutations and cancer predisposition: importance and mechanisms. Curr Opin Genet Dev 20:290–298
99. Higgs MR, Lerat H, Pawlotsky JM (2010) Downregulation of Gadd45beta expression by hepatitis C virus leads to defective cell cycle arrest. Cancer Res 70:4901–4911
100. Hill VK, Underhill-Day N, Krex D, Robel K, Sangan CB, Summersgill HR, Morris M, Gentle D, Chalmers AD, Maher ER, Latif F (2011) Epigenetic inactivation of the RASSF10 candidate

tumor suppressor gene is a frequent and an early event in gliomagenesis. Oncogene 30(8): 978–989

101. Hirata H, Hinoda Y, Nakajima K, Kawamoto K, Kikuno N, Kawakami K, Yamamura S, Ueno K, Majid S, Saini S, Ishii N, Dahiya R (2009) Wnt antagonist gene DKK2 is epigenetically silenced and inhibits renal cancer progression through apoptotic and cell cycle pathways. Clin Cancer Res 15:5678–5687
102. Hirata H, Hinoda Y, Nakajima K, Kikuno N, Yamamura S, Kawakami K, Suehiro Y, Tabatabai ZL, Ishii N, Dahiya R (2009) Wnt antagonist gene polymorphisms and renal cancer. Cancer 115:4488–4503
103. Hitchins MP, Rapkins RW, Kwok CT, Srivastava S, Wong JJ, Khachigian LM, Polly P, Goldblatt J, Ward RL (2011) Dominantly inherited constitutional epigenetic silencing of MLH1 in a cancer-affected family is linked to a single nucleotide variant within the 5'UTR. Cancer Cell 20(2):200–213
104. Hoffman AE, Zheng T, Yi C, Leaderer D, Weidhaas J, Slack F, Zhang Y, Paranjape T, Zhu Y (2009) microRNA miR-196a-2 and breast cancer: a genetic and epigenetic association study and functional analysis. Cancer Res 69:5970–5977
105. Honorio S, Agathanggelou A, Wernert N, Rothe M, Maher ER, Latif F (2003) Frequent epigenetic inactivation of the RASSF1A tumour suppressor gene in testicular tumours and distinct methylation profiles of seminoma and nonseminoma testicular germ cell tumours. Oncogene 22:461–466
106. Hoque MO, Feng Q, Toure P, Dem A, Critchlow CW, Hawes SE, Wood T, Jeronimo C, Rosenbaum E, Stern J, Yu M, Trink B, Kiviat NB, Sidransky D (2006) Detection of aberrant methylation of four genes in plasma DNA for the detection of breast cancer. J Clin Oncol 24:4262–4269
107. Hori M, Iwasaki M, Shimazaki J, Inagawa S, Itabashi M (2000) Assessment of hypermethylated DNA in two promoter regions of the estrogen receptor alpha gene in human endometrial diseases. Gynecol Oncol 76:89–96
108. Hsu HS, Chen TP, Hung CH, Wen CK, Lin RK, Lee HC, Wang YC (2007) Characterization of a multiple epigenetic marker panel for lung cancer detection and risk assessment in plasma. Cancer 110:2019–2026
109. Hu Z, Chen X, Zhao Y, Tian T, Jin G, Shu Y, Chen Y, Xu L, Zen K, Zhang C, Shen H (2010) Serum microRNA signatures identified in a genome-wide serum microRNA expression profiling predict survival of non-small-cell lung cancer. J Clin Oncol 28:1721–1726
110. Hu Z, Liang J, Wang Z, Tian T, Zhou X, Chen J, Miao R, Wang Y, Wang X, Shen H (2009) Common genetic variants in pre-microRNAs were associated with increased risk of breast cancer in Chinese women. Hum Mutat 30:79–84
111. Huang J, Plass C, Gerhauser C (2010) Cancer chemoprevention by targeting the epigenome. Curr Drug Targets. Dec 15. [Epub ahead of print]
112. Hughes LA, van den Brandt PA, de Bruine AP, Wouters KA, Hulsmans S, Spiertz A, Goldbohm RA, de Goeij AF, Herman JG, Weijenberg MP, van Engeland M (2009) Early life exposure to famine and colorectal cancer risk: a role for epigenetic mechanisms. PLoS One 4:e7951
113. Hussain M, Rao M, Humphries AE, Hong JA, Liu F, Yang M, Caragacianu D, Schrump DS (2009) Tobacco smoke induces polycomb-mediated repression of Dickkopf-1 in lung cancer cells. Cancer Res 69:3570–3578
114. Ibanez de Caceres I, Battagli C, Esteller M, Herman JG, Dulaimi E, Edelson MI, Bergman C, Ehya H, Eisenberg BL, Cairns P (2004) Tumor cell-specific BRCA1 and RASSF1A hypermethylation in serum, plasma, and peritoneal fluid from ovarian cancer patients. Cancer Res 64:6476–6481
115. Ji W, Hernandez R, Zhang XY, Qu GZ, Frady A, Varela M, Ehrlich M (1997) DNA demethylation and pericentromeric rearrangements of chromosome 1. Mutat Res 379:33–41
116. Jin F, Dowdy SC, Xiong Y, Eberhardt NL, Podratz KC, Jiang SW (2005) Up-regulation of DNA methyltransferase 3B expression in endometrial cancers. Gynecol Oncol 96:531–538

117. Kalantari M, Calleja-Macias IE, Tewari D, Hagmar B, Lie K, Barrera-Saldana HA, Wiley DJ, Bernard HU (2004) Conserved methylation patterns of human papillomavirus type 16 DNA in asymptomatic infection and cervical neoplasia. J Virol 78:12762–12772
118. Kanai Y (2010) Genome-wide DNA methylation profiles in precancerous conditions and cancers. Cancer Sci 101:36–45
119. Kang MY, Lee BB, Kim YH, Chang DK, Kyu Park S, Chun HK, Song SY, Park J, Kim DH (2007) Association of the SUV39H1 histone methyltransferase with the DNA methyltransferase 1 at mRNA expression level in primary colorectal cancer. Int J Cancer 121:2192–2197
120. Kang S, Kim JW, Kang GH, Park NH, Song YS, Kang SB, Lee HP (2005) Polymorphism in folate- and methionine-metabolizing enzyme and aberrant CpG island hypermethylation in uterine cervical cancer. Gynecol Oncol 96:173–180
121. Kasowski M, Grubert F, Heffelfinger C, Hariharan M, Asabere A, Waszak SM, Habegger L, Rozowsky J, Shi M, Urban AE, Hong MY, Karczewski KJ, Huber W, Weissman SM, Gerstein MB, Korbel JO, Snyder M (2010) Variation in transcription factor binding among humans. Science 328:232–235
122. Kaur J, Demokan S, Tripathi SC, Macha MA, Begum S, Califano JA, Ralhan R (2010) Promoter hypermethylation in Indian primary oral squamous cell carcinoma. Int J Cancer 127:2367–2373
123. Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, Hod E, Li K, Murty VV, Schupf N, Vilain E, Morris M, Haghighi F, Tycko B (2008) Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. Nat Genet 40:904–908
124. Kim HG, Lee S, Kim DY, Ryu SY, Joo JK, Kim JC, Lee KH, Lee JH (2010) Aberrant methylation of DNA mismatch repair genes in elderly patients with sporadic gastric carcinoma: a comparison with younger patients. J Surg Oncol 101:28–35
125. Kim MJ, White-Cross JA, Shen L, Issa JP, Rashid A (2009) Hypomethylation of long interspersed nuclear element-1 in hepatocellular carcinomas. Mod Pathol 22:442–449
126. Kiran M, Chawla YK, Kaur J (2009) Methylation profiling of tumor suppressor genes and oncogenes in hepatitis virus-related hepatocellular carcinoma in northern India. Cancer Genet Cytogenet 195:112–119
127. Kontorovich T, Cohen Y, Nir U, Friedman E (2009) Promoter methylation patterns of ATM, ATR, BRCA1, BRCA2 and p53 as putative cancer risk modifiers in Jewish BRCA1/BRCA2 mutation carriers. Breast Cancer Res Treat 116:195–200
128. Kontorovich T, Levy A, Korostishevsky M, Nir U, Friedman E (2010) Single nucleotide polymorphisms in miRNA binding sites and miRNA genes as breast/ovarian cancer risk modifiers in Jewish high-risk women. Int J Cancer 127:589–597
129. Kulis M, Esteller M (2010) DNA methylation and cancer. Adv Genet 70:27–56
130. Kwak PB, Iwasaki S, Tomari Y (2010) The microRNA pathway and cancer. Cancer Sci 101:2309–2315
131. Laharanne E, Chevret E, Idrissi Y, Gentil C, Longy M, Ferrer J, Dubus P, Jouary T, Vergier B, Beylot-Barry M, Merlio JP (2010) CDKN2A-CDKN2B deletion defines an aggressive subset of cutaneous T-cell lymphoma. Mod Pathol 23:547–558
132. Landi D, Gemignani F, Naccarati A, Pardini B, Vodicka P, Vodickova L, Novotny J, Forsti A, Hemminki K, Canzian F, Landi S (2008) Polymorphisms within micro-RNA-binding sites and risk of sporadic colorectal cancer. Carcinogenesis 29:579–584
133. Lee BB, Lee EJ, Jung EH, Chun HK, Chang DK, Song SY, Park J, Kim DH (2009) Aberrant methylation of APC, MGMT, RASSF2A, and Wif-1 genes in plasma as a biomarker for early detection of colorectal cancer. Clin Cancer Res 15:6185–6191
134. Lee HS, Kim BH, Cho NY, Yoo EJ, Choi M, Shin SH, Jang JJ, Suh KS, Kim YS, Kang GH (2009) Prognostic implications of and relationship between CpG island hypermethylation and repetitive DNA hypomethylation in hepatocellular carcinoma. Clin Cancer Res 15:812–820
135. Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75:843–854

136. Lee S, Lee HJ, Kim JH, Lee HS, Jang JJ, Kang GH (2003) Aberrant CpG island hypermethylation along multistep hepatocarcinogenesis. Am J Pathol 163:1371–1378
137. Lee SJ, Jeon HS, Jang JS, Park SH, Lee GY, Lee BH, Kim CH, Kang YM, Lee WK, Kam S, Park RW, Kim IS, Cho YL, Jung TH, Park JY (2005) DNMT3B polymorphisms and risk of primary lung cancer. Carcinogenesis 26:403–409
138. Lee SM, Lee JY, Choi JE, Lee SY, Park JY, Kim DS (2010) Epigenetic inactivation of retinoid X receptor genes in non-small cell lung cancer and the relationship with clinicopathologic features. Cancer Genet Cytogenet 197:39–45
139. Leng S, Stidley CA, Bernauer AM, Picchi MA, Sheng X, Frasco MA, Van Den Berg D, Gilliland FD, Crowell RE, Belinsky SA (2008) Haplotypes of DNMT1 and DNMT3B are associated with mutagen sensitivity induced by benzo[a]pyrene diol epoxide among smokers. Carcinogenesis 29:1380–1385
140. Lerat H, Honda M, Beard MR, Loesch K, Sun J, Yang Y, Okuda M, Gosert R, Xiao SY, Weinman SA, Lemon SM (2002) Steatosis and liver cancer in transgenic mice expressing the structural and nonstructural proteins of hepatitis C virus. Gastroenterology 122:352–365
141. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, Kandoth C, Payton JE, Baty J, Welch J, Harris CC, Lichti CF, Townsend RR, Fulton RS, Dooling DJ, Koboldt DC, Schmidt H, Zhang Q, Osborne JR, Lin L, O'Laughlin M, McMichael JF, Delehaunty KD, McGrath SD, Fulton LA, Magrini VJ, Vickery TL, Hundal J, Cook LL, Conyers JJ, Swift GW, Reed JP, Alldredge PA, Wylie T, Walker J, Kalicki J, Watson MA, Heath S, Shannon WD, Varghese N, Nagarajan R, Westervelt P, Tomasson MH, Link DC, Graubert TA, DiPersio JF, Mardis ER, Wilson RK (2010) DNMT3A mutations in acute myeloid leukemia. N Engl J Med 363:2424–2433
142. Lightfoot TJ, Johnston WT, Painter D, Simpson J, Roman E, Skibola CF, Smith MT, Allan JM, Taylor GM (2010) Genetic variation in the folate metabolic pathway and risk of childhood leukemia. Blood 115:3923–3929
143. Limsui D, Vierkant RA, Tillmans LS, Wang AH, Weisenberger DJ, Laird PW, Lynch CF, Anderson KE, French AJ, Haile RW, Harnack LJ, Potter JD, Slager SL, Smyrk TC, Thibodeau SN, Cerhan JR, Limburg PJ (2010) Cigarette smoking and colorectal cancer risk by molecularly defined subtypes. J Natl Cancer Inst 102:1012–1022
144. Lin CH, Hsieh SY, Sheen IS, Lee WC, Chen TC, Shyu WC, Liaw YF (2001) Genome-wide hypomethylation in hepatocellular carcinogenesis. Cancer Res 61:4238–4243
145. Lin J, Lee IM, Cook NR, Selhub J, Manson JE, Buring JE, Zhang SM (2008) Plasma folate, vitamin B-6, vitamin B-12, and risk of breast cancer in women. Am J Clin Nutr 87:734–743
146. Linggi BE, Brandt SJ, Sun ZW, Hiebert SW (2005) Translating the histone code into leukemia. J Cell Biochem 96:938–950
147. Liu BL, Cheng JX, Zhang X, Wang R, Zhang W, Lin H, Xiao X, Cai S, Chen XY, Cheng H (2010) Global histone modification patterns as prognostic markers to classify glioma patients. Cancer Epidemiol Biomarkers Prev 19:2888–2896
148. Liu F, Killian JK, Yang M, Walker RL, Hong JA, Zhang M, Davis S, Zhang Y, Hussain M, Xi S, Rao M, Meltzer PA, Schrump DS (2010) Epigenomic alterations and gene expression profiles in respiratory epithelia exposed to cigarette smoke condensate. Oncogene 29:3650–3664
149. Liu H, Zhou Y, Boggs SE, Belinsky SA, Liu J (2007) Cigarette smoke induces demethylation of prometastatic oncogene synuclein-gamma in lung cancer cells by downregulation of DNMT3B. Oncogene 26:5900–5910
150. Liu Z, Wang L, Wang LE, Sturgis EM, Wei Q (2008) Polymorphisms of the DNMT3B gene and risk of squamous cell carcinoma of the head and neck: a case-control study. Cancer Lett 268:158–165
151. Lu B, Ma Y, Wu G, Tong X, Guo H, Liang A, Cong W, Liu C, Wang H, Wu M, Zhao J, Guo Y (2008) Methylation of Tip30 promoter is associated with poor prognosis in human hepatocellular carcinoma. Clin Cancer Res 14:7405–7412
152. Maeda G, Chiba T, Kawashiri S, Satoh T, Imai K (2007) Epigenetic inactivation of IkappaB Kinase-alpha in oral carcinomas and tumor progression. Clin Cancer Res 13:5041–5047

153. Maegawa S, Hinkal G, Kim HS, Shen L, Zhang L, Zhang J, Zhang N, Liang S, Donehower LA, Issa JP (2010) Widespread and tissue specific age-related DNA methylation changes in mice. Genome Res 20:332–340
154. Maekita T, Nakazawa K, Mihara M, Nakajima T, Yanaoka K, Iguchi M, Arii K, Kaneda A, Tsukamoto T, Tatematsu M, Tamura G, Saito D, Sugimura T, Ichinose M, Ushijima T (2006) High levels of aberrant DNA methylation in Helicobacter pylori-infected gastric mucosae and its possible association with gastric cancer risk. Clin Cancer Res 12:989–995
155. Manuyakorn A, Paulus R, Farrell J, Dawson NA, Tze S, Cheung-Lau G, Hines OJ, Reber H, Seligson DB, Horvath S, Kurdistani SK, Guha C, Dawson DW (2010) Cellular histone modification patterns predict prognosis and treatment response in resectable pancreatic adenocarcinoma: results from RTOG 9704. J Clin Oncol 28:1358–1365
156. Marsit CJ, Houseman EA, Schned AR, Karagas MR, Kelsey KT (2007) Promoter hypermethylation is associated with current smoking, age, gender and survival in bladder cancer. Carcinogenesis 28(8):1745–1751
157. Marsit CJ, Christensen BC, Houseman EA, Karagas MR, Wrensch MR, Yeh RF, Nelson HH, Wiemels JL, Zheng S, Posner MR, McClean MD, Wiencke JK, Kelsey KT (2009) Epigenetic profiling reveals etiologically distinct patterns of DNA methylation in head and neck squamous cell carcinoma. Carcinogenesis 30:416–422
158. Martinez-Romero C, Rooman I, Skoudy A, Guerra C, Molero X, Gonzalez A, Iglesias M, Lobato T, Bosch A, Barbacid M, Real FX, Hernandez-Munoz I (2009) The epigenetic regulators Bmi1 and Ring1B are differentially regulated in pancreatitis and pancreatic ductal adenocarcinoma. J Pathol 219:205–213
159. Martinez I, Gardiner AS, Board KF, Monzon FA, Edwards RP, Khan SA (2008) Human papillomavirus type 16 reduces the expression of microRNA-218 in cervical carcinoma cells. Oncogene 27:2575–2582
160. Martinez R, Esteller M (2010) The DNA methylome of glioblastoma multiforme. Neurobiol Dis 39:40–46
161. Martinez R, Setien F, Voelter C, Casado S, Quesada MP, Schackert G, Esteller M (2007) CpG island promoter hypermethylation of the pro-apoptotic gene caspase-8 is a common hallmark of relapsed glioblastoma multiforme. Carcinogenesis 28:1264–1268
162. Martini M, Pallini R, Luongo G, Cenci T, Lucantoni C, Larocca LM (2008) Prognostic relevance of SOCS3 hypermethylation in patients with glioblastoma multiforme. Int J Cancer 123:2955–2960
163. Mato JM, Corrales FJ, Lu SC, Avila MA (2002) S-Adenosylmethionine: a control switch that regulates liver function. FASEB J 16:15–26
164. Michael MZ, O'Connor SM, van Holst Pellekaan NG, Young GP, James RJ (2003) Reduced accumulation of specific microRNAs in colorectal neoplasia. Mol Cancer Res 1:882–891
165. Minniti G, Salvati M, Arcella A, Buttarelli F, D'Elia A, Lanzetta G, Esposito V, Scarpino S, Maurizi Enrici R, Giangaspero F (2011) Correlation between O6-methylguanine-DNA methyltransferase and survival in elderly patients with glioblastoma treated with radiotherapy plus concomitant and adjuvant temozolomide. J Neurooncol 102:311–316
166. Mirabello L, Savage SA, Korde L, Gadalla SM, Greene MH (2010) LINE-1 methylation is inherited in familial testicular cancer kindreds. BMC Med Genet 11:77
167. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, Lin DW, Urban N, Drescher CW, Knudsen BS, Stirewalt DL, Gentleman R, Vessella RL, Nelson PS, Martin DB, Tewari M (2008) Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci USA 105:10513–10518
168. Montgomery KG, Liu MC, Eccles DM, Campbell IG (2004) The DNMT3B C–>T promoter polymorphism and risk of breast cancer in a British population: a case-control study. Breast Cancer Res 6:R390–R394
169. Morgan HD, Sutherland HG, Martin DI, Whitelaw E (1999) Epigenetic inheritance at the agouti locus in the mouse. Nat Genet 23:314–318

170. Mosashvilli D, Kahl P, Mertens C, Holzapfel S, Rogenhofer S, Hauser S, Buttner R, von Ruecker A, Muller SC, Ellinger J (2010) Global histone acetylation levels: prognostic relevance in patients with renal cell carcinoma. Cancer Sci 101(12):2664–2669
171. Muhlisch J, Bajanowski T, Rickert CH, Roggendorf W, Wurthwein G, Jurgens H, Fruhwald MC (2007) Frequent but borderline methylation of p16 (INK4a) and TIMP3 in medulloblastoma and sPNET revealed by quantitative analyses. J Neurooncol 83:17–29
172. Muller HM, Fiegl H, Widschwendter A, Widschwendter M (2004) Prognostic DNA methylation marker in serum of cancer patients. Ann N Y Acad Sci 1022:44–49
173. Nagarajan RP, Costello JF (2009) Epigenetic mechanisms in glioblastoma multiforme. Semin Cancer Biol 19:188–197
174. Nagasaka T, Rhees J, Kloor M, Gebert J, Naomoto Y, Boland CR, Goel A (2010) Somatic hypermethylation of MSH2 is a frequent event in lynch syndrome colorectal cancers. Cancer Res 70:3098–3108
175. Nakajima T, Enomoto S, Ushijima T (2008) DNA methylation: a marker for carcinogen exposure and cancer risk. Environ Health Prev Med 13:8–15
176. Nakajima T, Maekita T, Oda I, Gotoda T, Yamamoto S, Umemura S, Ichinose M, Sugimura T, Ushijima T, Saito D (2006) Higher methylation levels in gastric mucosae significantly correlate with higher risk of gastric cancers. Cancer Epidemiol Biomarkers Prev 15:2317–2321
177. Network CGAR (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455:1061–1068
178. Neyaz MK, Kumar RS, Hussain S, Naqvi SH, Kohaar I, Thakur N, Kashyap V, Das BC, Husain SA, Bharadwaj M (2008) Effect of aberrant promoter methylation of FHIT and RASSF1A genes on susceptibility to cervical cancer in a North Indian population. Biomarkers 13:597–606
179. Ng EK, Chong WW, Jin H, Lam EK, Shin VY, Yu J, Poon TC, Ng SS, Sung JJ (2009) Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening. Gut 58:1375–1381
180. Nguyen T, Kuo C, Nicholl MB, Sim MS, Turner RR, Morton DL, Hoon DS (2011) Downregulation of microRNA-29c is associated with hypermethylation of tumor-related genes and disease outcome in cutaneous melanoma. Epigenetics 6:388–394
181. Nosho K, Kure S, Irahara N, Shima K, Baba Y, Spiegelman D, Meyerhardt JA, Giovannucci EL, Fuchs CS, Ogino S (2009) A prospective cohort study shows unique epigenetic, genetic, and prognostic features of synchronous colorectal cancers. Gastroenterology 137:1609–1620, e1-3
182. Omura N, Li CP, Li A, Hong SM, Walter K, Jimeno A, Hidalgo M, Goggins M (2008) Genome-wide profiling of methylated promoters in pancreatic adenocarcinoma. Cancer Biol Ther 7:1146–1156
183. Osterhues A, Holzgreve W, Michels KB (2009) Shall we put the world on folate? Lancet 374:959–961
184. Paci M, Maramotti S, Bellesia E, Formisano D, Albertazzi L, Ricchetti T, Ferrari G, Annessi V, Lasagni D, Carbonelli C, De FS, Brini M, Sgarbi G, Lodi R (2009) Circulating plasma DNA as diagnostic biomarker in non-small cell lung cancer. Lung Cancer 64:92–97
185. Paliwal A, Vaissiere T, Herceg Z (2010) Quantitative detection of DNA methylation states in minute amounts of DNA from body fluids. Methods 52:242–247
186. Palmero EI, de Campos SG, Campos M, de Souza NC, Guerreiro ID, Carvalho AL, Marques MM (2011) Mechanisms and role of microRNA deregulation in cancer onset and progression. Genet Mol Biol 34(3):363–370
187. Park JY (2010) Promoter hypermethylation in prostate cancer. Cancer Control 17:245–255
188. Patard JJ, Rioux-Leclercq N, Masson D, Zerrouki S, Jouan F, Collet N, Dubourg C, Lobel B, Denis M, Fergelot P (2009) Absence of VHL gene alteration and high VEGF expression are associated with tumour aggressiveness and poor survival of renal-cell carcinoma. Br J Cancer 101:1417–1424

189. Paz MF, Avila S, Fraga MF, Pollan M, Capella G, Peinado MA, Sanchez-Cespedes M, Herman JG, Esteller M (2002) Germ-line variants in methyl-group metabolism genes and susceptibility to DNA methylation in normal tissues and human primary tumors. Cancer Res 62:4519–4524
190. Pekarsky Y, Zanesi N, Palamarchuk A, Huebner K, Croce CM (2002) FHIT: from gene discovery to cancer treatment and prevention. Lancet Oncol 3:748–754
191. Phillips JM, Goodman JI (2009) Inhalation of cigarette smoke induces regions of altered DNA methylation (RAMs) in SENCAR mouse lung. Toxicology 260:7–15
192. Pinzani P, Salvianti F, Zaccara S, Massi D, De Giorgi V, Pazzagli M, Orlando C (2011) Circulating cell-free DNA in plasma of melanoma patients: qualitative and quantitative considerations. Clin Chim Acta 412(23–24):2141–2145
193. Piva R, Kumar VL, Hanau S, Rimondi AP, Pansini S, Mollica G, del Senno L (1989) Abnormal methylation of estrogen receptor gene and reduced estrogen receptor RNA levels in human endometrial carcinomas. J Steroid Biochem 32:1–4
194. Poage GM, Christensen BC, Houseman EA, McClean MD, Wiencke JK, Posner MR, Clark JR, Nelson HH, Marsit CJ, Kelsey KT (2010) Genetic and epigenetic somatic alterations in head and neck squamous cell carcinomas are globally coordinated but not locally targeted. PLoS One 5:e9651
195. Ponomaryova AA, Rykova EY, Cherdyntseva NV, Skvortsova TE, Dobrodeev AY, Zav'yalov AA, Tuzikov SA, Vlassov VV, Laktionov PP (2011) RARB2 gene methylation level in the circulating DNA from blood of patients with lung cancer. Eur J Cancer Prev 20:453–455
196. Porkka KP, Pfeiffer MJ, Waltering KK, Vessella RL, Tammela TL, Visakorpi T (2007) MicroRNA expression profiling in prostate cancer. Cancer Res 67:6130–6135
197. Prueitt RL, Goodman JE, Valberg PA (2009) Radionuclides in cigarettes may lead to carcinogenesis via p16(INK4a) inactivation. J Environ Radioact 100:157–161
198. Rai K, Huggins IJ, James SR, Karpf AR, Jones DA, Cairns BR (2008) DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. Cell 135:1201–1212
199. Rakyan VK, Chong S, Champ ME, Cuthbert PC, Morgan HD, Luu KV, Whitelaw E (2003) Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. Proc Natl Acad Sci USA 100:2538–2543
200. Raman V, Martensen SA, Reisman D, Evron E, Odenwald WF, Jaffee E, Marks J, Sukumar S (2000) Compromised HOXA5 function can limit p53 expression in human breast tumours. Nature 405:974–978
201. Rassoulzadegan M, Grandjean V, Gounon P, Vincent S, Gillot I, Cuzin F (2006) RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. Nature 441:469–474
202. Raval A, Tanner SM, Byrd JC, Angerman EB, Perko JD, Chen SS, Hackanson B, Grever MR, Lucas DM, Matkovic JJ, Lin TS, Kipps TJ, Murray F, Weisenburger D, Sanger W, Lynch J, Watson P, Jansen M, Yoshinaga Y, Rosenquist R, de Jong PJ, Coggill P, Beck S, Lynch H, de la Chapelle A, Plass C (2007) Downregulation of death-associated protein kinase 1 (DAPK1) in chronic lymphocytic leukemia. Cell 129:879–890
203. Risch A, Plass C (2008) Lung cancer epigenetics and genetics. Int J Cancer 123:1–7
204. Risinger JI, Maxwell GL, Berchuck A, Barrett JC (2003) Promoter hypermethylation as an epigenetic component in type I and type II endometrial cancers. Ann N Y Acad Sci 983:208–212
205. Roupret M, Hupertan V, Catto JW, Yates DR, Rehman I, Proctor LM, Phillips J, Meuth M, Cussenot O, Hamdy FC (2008) Promoter hypermethylation in circulating blood cells identifies prostate cancer progression. Int J Cancer 122:952–956
206. Sanson M, Marie Y, Paris S, Idbaih A, Laffaire J, Ducray F, El Hallani S, Boisselier B, Mokhtari K, Hoang-Xuan K, Delattre JY (2009) Isocitrate dehydrogenase 1 codon 132 mutation is an important prognostic biomarker in gliomas. J Clin Oncol 27:4150–4154
207. Sasaki M, Kotcherguina L, Dharia A, Fujimoto S, Dahiya R (2001) Cytosine-phosphoguanine methylation of estrogen receptors in endometrial cancer. Cancer Res 61:3262–3266

208. Sauer J, Mason JB, Choi SW (2009) Too much folate: a risk factor for cancer and cardiovascular disease? Curr Opin Clin Nutr Metab Care 12:30–36
209. Schee K, Fodstad O, Flatmark K (2010) MicroRNAs as biomarkers in colorectal cancer. Am J Pathol 777:1592–1599
210. Schernhammer E, Wolpin B, Rifai N, Cochrane B, Manson JA, Ma J, Giovannucci E, Thomson C, Stampfer MJ, Fuchs C (2007) Plasma folate, vitamin B6, vitamin B12, and homocysteine and pancreatic cancer risk in four large cohorts. Cancer Res 67:5553–5560
211. Schetter AJ, Leung SY, Sohn JJ, Zanetti KA, Bowman ED, Yanaihara N, Yuen ST, Chan TL, Kwong DL, Au GK, Liu CG, Calin GA, Croce CM, Harris CC (2008) MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. JAMA 299:425–436
212. Schetter AJ, Nguyen GH, Bowman ED, Mathe EA, Yuen ST, Hawkes JE, Croce CM, Leung SY, Harris CC (2009) Association of inflammation-related and microRNA gene expression with cancer-specific mortality of colon adenocarcinoma. Clin Cancer Res 15:5878–5887
213. Seligson DB, Horvath S, Shi T, Yu H, Tze S, Grunstein M, Kurdistani SK (2005) Global histone modification patterns predict risk of prostate cancer recurrence. Nature 435:1262–1266
214. Shen H, Wang L, Spitz MR, Hong WK, Mao L, Wei Q (2002) A novel polymorphism in human cytosine DNA-methyltransferase-3B promoter is associated with an increased risk of lung cancer. Cancer Res 62:4992–4995
215. Shen J, Ambrosone CB, DiCioccio RA, Odunsi K, Lele SB, Zhao H (2008) A functional polymorphism in the miR-146a gene and age of familial breast/ovarian cancer diagnosis. Carcinogenesis 29:1963–1966
216. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J, Houlihan PS, Krouse RS, Prasad AR, Einspahr JG, Buckmeier J, Alberts DS, Hamilton SR, Issa JP (2005) MGMT promoter methylation and field defect in sporadic colorectal cancer. J Natl Cancer Inst 97:1330–1338
217. Shen R, Pan S, Qi S, Lin X, Cheng S (2010) Epigenetic repression of microRNA-129-2 leads to overexpression of SOX4 in gastric cancer. Biochem Biophys Res Commun 394:1047–1052
218. Shiozawa T, Itoh K, Horiuchi A, Konishi I, Fujii S, Nikaido T (2002) Down-regulation of estrogen receptor by the methylation of the estrogen receptor gene in endometrial carcinoma. Anticancer Res 22:139–143
219. Shoemaker R, Deng J, Wang W, Zhang K (2010) Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. Genome Res 20:883–889
220. Singal R, Das PM, Manoharan M, Reis IM, Schlesselman JJ (2005) Polymorphisms in the DNA methyltransferase 3b gene and prostate cancer risk. Oncol Rep 14:569–573
221. Skotheim RI, Lind GE, Monni O, Nesland JM, Abeler VM, Fossa SD, Duale N, Brunborg G, Kallioniemi O, Andrews PW, Lothe RA (2005) Differentiation of human embryonal carcinomas in vitro and in vivo reveals expression profiles relevant to normal development. Cancer Res 65:5588–5598
222. Skyldberg B, Salo S, Eriksson E, Aspenblad U, Moberger B, Tryggvason K, Auer G (1999) Laminin-5 as a marker of invasiveness in cervical lesions. J Natl Cancer Inst 91:1882–1887
223. Smiraglia DJ, Szymanska J, Kraggerud SM, Lothe RA, Peltomaki P, Plass C (2002) Distinct epigenetic phenotypes in seminomatous and nonseminomatous testicular germ cell tumors. Oncogene 21:3909–3916
224. Smits KM, Schouten LJ, van Dijk BA, Hulsbergen-van de Kaa CA, Wouters KA, Oosterwijk E, van Engeland M, van den Brandt PA (2008) Genetic and epigenetic alterations in the von hippel-lindau gene: the influence on renal cancer prognosis. Clin Cancer Res 14:782–787
225. Soares J, Pinto AE, Cunha CV, Andre S, Barao I, Sousa JM, Cravo M (1999) Global DNA hypomethylation in breast carcinoma: correlation with prognostic factors and tumor progression. Cancer 85:112–118

226. Sobulo OM, Borrow J, Tomek R, Reshmi S, Harden A, Schlegelberger B, Housman D, Doggett NA, Rowley JD, Zeleznik-Le NJ (1997) MLL is fused to CBP, a histone acetyltransferase, in therapy-related acute myeloid leukemia with a t(11;16)(q23;p13.3). Proc Natl Acad Sci USA 94:8732–8737
227. Sozzi G, Conte D, Mariani L, Lo VS, Roz L, Lombardo C, Pierotti MA, Tavecchio L (2001) Analysis of circulating tumor DNA in plasma at diagnosis and during follow-up of lung cancer patients. Cancer Res 61:4675–4678
228. Stevenson FK, Caligaris-Cappio F (2004) Chronic lymphocytic leukemia: revelations from the B-cell receptor. Blood 103:4389–4395
229. Stidley CA, Picchi MA, Leng S, Willink R, Crowell RE, Flores KG, Kang H, Byers T, Gilliland FD, Belinsky SA (2010) Multivitamins, folate, and green vegetables protect against gene promoter methylation in the aerodigestive tract of smokers. Cancer Res 70:568–574
230. Stolzenberg-Solomon RZ, Chang SC, Leitzmann MF, Johnson KA, Johnson C, Buys SS, Hoover RN, Ziegler RG (2006) Folate intake, alcohol use, and postmenopausal breast cancer risk in the prostate, lung, colorectal, and ovarian cancer screening trial. Am J Clin Nutr 83:895–904
231. Su HY, Lai HC, Lin YW, Chou YC, Liu CY, Yu MH (2009) An epigenetic marker panel for screening and prognostic prediction of ovarian cancer. Int J Cancer 124:387–393
232. Suga Y, Miyajima K, Oikawa T, Maeda J, Usuda J, Kajiwara N, Ohira T, Uchida O, Tsuboi M, Hirano T, Kato H, Ikeda N (2008) Quantitative p16 and ESR1 methylation in the peripheral blood of patients with non-small cell lung cancer. Oncol Rep 20:1137–1142
233. Sugawara W, Haruta M, Sasaki F, Watanabe N, Tsunematsu Y, Kikuta A, Kaneko Y (2007) Promoter hypermethylation of the RASSF1A gene predicts the poor outcome of patients with hepatoblastoma. Pediatr Blood Cancer 49:240–249
234. Sugita H, Iida S, Inokuchi M, Kato K, Ishiguro M, Ishikawa T, Takagi Y, Enjoji M, Yamada H, Uetake H, Kojima K, Sugihara K (2011) Methylation of BNIP3 and DAPK indicates lower response to chemotherapy and poor prognosis in gastric cancer. Oncol Rep 25:513–518
235. Suter CM, Martin DI, Ward RL (2004) Germline epimutation of MLH1 in individuals with multiple cancers. Nat Genet 36:497–501
236. Suzuki M, Mohamed S, Nakajima T, Kubo R, Tian L, Fujiwara T, Suzuki H, Nagato K, Chiyo M, Motohashi S, Yasufuku K, Iyoda A, Yoshida S, Sekine Y, Shibuya K, Hiroshima K, Nakatani Y, Yoshino I, Fujisawa T (2008) Aberrant methylation of CXCL12 in non-small cell lung cancer is associated with an unfavorable prognosis. Int J Oncol 33:113–119
237. Suzuki M, Wada H, Yoshino M, Tian L, Shigematsu H, Suzuki H, Alaa M, Tamura H, Fujiwara T, Nagato K, Motohashi S, Moriya Y, Hoshino H, Yoshida S, Shibuya K, Hiroshima K, Nakatani Y, Yoshino I (2010) Molecular characterization of chronic obstructive pulmonary disease-related non-small cell lung cancer through aberrant methylation and alterations of EGFR signaling. Ann Surg Oncol 17:878–888
238. Taback B, Giuliano AE, Lai R, Hansen N, Singer FR, Pantel K, Hoon DS (2006) Epigenetic analysis of body fluids and tumor tissues: application of a comprehensive molecular assessment for early-stage breast cancer patients. Ann N Y Acad Sci 1075:211–221
239. Tan AC, Jimeno A, Lin SH, Wheelhouse J, Chan F, Solomon A, Rajeshkumar NV, Rubio-Viqueira B, Hidalgo M (2009) Characterizing DNA methylation patterns in pancreatic cancer genome. Mol Oncol 3:425–438
240. Tanemura A, Terando AM, Sim MS, van Hoesel AQ, de Maat MF, Morton DL, Hoon DS (2009) CpG island methylator phenotype predicts progression of malignant melanoma. Clin Cancer Res 15:1801–1807
241. Tangkijvanich P, Hourpai N, Rattanatanyong P, Wisedopas N, Mahachai V, Mutirangura A (2007) Serum LINE-1 hypomethylation as a potential prognostic marker for hepatocellular carcinoma. Clin Chim Acta 379:127–133

242. Taniguchi A, Nemoto Y, Yokoyama A, Kotani N, Imai S, Shuin T, Daibata M (2008) Promoter methylation of the bone morphogenetic protein-6 gene in association with adult T-cell leukemia. Int J Cancer 123:1824–1831
243. Tanzer M, Balluff B, Distler J, Hale K, Leodolter A, Rocken C, Molnar B, Schmid R, Lofton-Day C, Schuster T, Ebert MP (2010) Performance of epigenetic markers SEPT9 and ALX4 in plasma for detection of colorectal precancerous lesions. PLoS One 5:e9061
244. Tazawa H, Tsuchiya N, Izumiya M, Nakagama H (2007) Tumor-suppressive miR-34a induces senescence-like growth arrest through modulation of the E2F pathway in human colon cancer cells. Proc Natl Acad Sci USA 104:15472–15477
245. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Gayther SA, Apostolidou S, Jones A, Lechner M, Beck S, Jacobs IJ, Widschwendter M (2009) An epigenetic signature in peripheral blood predicts active ovarian cancer. PLoS One 4:e8274
246. Tong WG, Wierda WG, Lin E, Kuang SQ, Bekele BN, Estrov Z, Wei Y, Yang H, Keating MJ, Garcia-Manero G (2010) Genome-wide DNA methylation profiling of chronic lymphocytic leukemia allows identification of epigenetically repressed molecular pathways with clinical impact. Epigenetics 5:499–508
247. Toyooka S, Tokumo M, Shigematsu H, Matsuo K, Asano H, Tomii K, Ichihara S, Suzuki M, Aoe M, Date H, Gazdar AF, Shimizu N (2006) Mutational and epigenetic evidence for independent pathways for lung adenocarcinomas arising in smokers and never smokers. Cancer Res 66:1371–1375
248. Toyota M, Suzuki H (2010) Epigenetic drivers of genetic alterations. Adv Genet 70:309–323
249. Tsai KW, Hu LY, Wu CW, Li SC, Lai CH, Kao HW, Fang WL, Lin WC (2010) Epigenetic regulation of miR-196b expression in gastric cancer. Genes Chromosomes Cancer 49:969–980
250. Turhan AG, Eaves CJ, Kalousek DK, Eaves AC, Humphries RK (1988) Molecular analysis of clonality and bcr rearrangements in Philadelphia chromosome-positive acute lymphoblastic leukemia. Blood 71:1495–1498
251. Vaissiere T, Cuenin C, Paliwal A, Vineis P, Hoek G, Krzyzanowski M, Airoldi L, Dunning A, Garte S, Hainaut P, Malaveille C, Overvad K, Clavel-Chapelon F, Linseisen J, Boeing H, Trichopoulou A, Trichopoulos D, Kaladidi A, Palli D, Krogh V, Tumino R, Panico S, Bueno-De-Mesquita HB, Peeters PH, Kumle M, Gonzalez CA, Martinez C, Dorronsoro M, Barricarte A, Navarro C, Quiros JR, Berglund G, Janzon L, Jarvholm B, Day NE, Key TJ, Saracci R, Kaaks R, Riboli E, Herceg Z (2009) Quantitative analysis of DNA methylation after whole bisulfitome amplification of a minute amount of DNA from body fluids. Epigenetics 4:221–230
252. van Engeland M, Weijenberg MP, Roemen GM, Brink M, de Bruine AP, Goldbohm RA, van den Brandt PA, Baylin SB, de Goeij AF, Herman JG (2003) Effects of dietary folate and alcohol intake on promoter methylation in sporadic colorectal cancer: the Netherlands cohort study on diet and cancer. Cancer Res 63:3133–3137
253. Veeck J, Esteller M (2010) Breast cancer epigenetics: from DNA methylation to microRNAs. J Mammary Gland Biol Neoplasia 15:5–17
254. Vilas-Zornoza A, Agirre X, Martín-Palanco V, Martín-Subero JI, San José-Eneriz E, Garate L, Álvarez S, Miranda E, Rodríguez-Otero P, Rifón J, Torres A, Calasanz MJ, Cruz Cigudosa J, Román-Gómez J, Prósper F (2011) Frequent and simultaneous epigenetic inactivation of TP53 pathway genes in acute lymphoblastic leukemia. PLoS One 6:e17012
255. Vineis P, Chuang SC, Vaissiere T, Cuenin C, Ricceri F, Johansson M, Ueland P, Brennan P, Herceg Z (2011) DNA methylation changes associated with cancer risk factors and blood levels of vitamin metabolites in a prospective study. Epigenetics 6:195–201
256. Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, Prueitt RL, Yanaihara N, Lanza G, Scarpa A, Vecchione A, Negrini M, Harris CC, Croce CM (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci USA 103:2257–2261

257. Wang CJ, Zhou ZG, Wang L, Yang L, Zhou B, Gu J, Chen HY, Sun XF (2009) Clinicopathological significance of microRNA-31, -143 and −145 expression in colorectal cancer. Dis Markers 26:27–34
258. Wang W, Gao J, Man XH, Li ZS, Gong YF (2009) Significance of DNA methyltransferase-1 and histone deacetylase-1 in pancreatic cancer. Oncol Rep 21(6):1439–1447
259. Wei Y, Xia W, Zhang Z, Liu J, Wang H, Adsay NV, Albarracin C, Yu D, Abbruzzese JL, Mills GB, Bast RC Jr, Hortobagyi GN, Hung MC (2008) Loss of trimethylation at lysine 27 of histone H3 is a predictor of poor outcome in breast, ovarian, and pancreatic cancers. Mol Carcinog 47:701–706
260. Weidman JR, Dolinoy DC, Murphy SK, Jirtle RL (2007) Cancer susceptibility: epigenetic manifestation of environmental exposures. Cancer J 13:9–16
261. Widschwendter M, Apostolidou S, Jones AA, Fourkala EO, Arora R, Pearce CL, Frasco MA, Ayhan A, Zikan M, Cibula D, Iyibozkurt CA, Yavuz E, Hauser-Kronberger C, Dubeau L, Menon U, Jacobs IJ (2009) HOXA methylation in normal endometrium from premenopausal women is associated with the presence of ovarian cancer: a proof of principle study. Int J Cancer 125:2214–2218
262. Widschwendter M, Apostolidou S, Raum E, Rothenbacher D, Fiegl H, Menon U, Stegmaier C, Jacobs IJ, Brenner H (2008) Epigenotyping in peripheral blood cell DNA and breast cancer risk: a proof of principle study. PLoS One 3:e2656
263. Widschwendter M, Siegmund KD, Muller HM, Fiegl H, Marth C, Muller-Holzner E, Jones PA, Laird PW (2004) Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. Cancer Res 64:3807–3813
264. Wilhelm CS, Kelsey KT, Butler R, Plaza S, Gagne L, Zens MS, Andrew AS, Morris S, Nelson HH, Schned AR, Karagas MR, Marsit CJ (2010) Implications of LINE1 methylation for bladder cancer risk in women. Clin Cancer Res 16:1682–1689
265. Wu HC, Delgado-Cruzata L, Flom JD, Kappil M, Ferris JS, Liao Y, Santella RM, Terry MB (2010) Global methylation profiles in DNA from different blood cell types. Epigenetics 6:76–85
266. Xu B, Feng NH, Li PC, Tao J, Wu D, Zhang ZD, Tong N, Wang JF, Song NH, Zhang W, Hua LX, Wu HF (2010) A functional polymorphism in Pre-miR-146a gene is associated with prostate cancer risk and mature miR-146a expression in vivo. Prostate 70:467–472
267. Xue X, Zhu YM, Woll PJ (2006) Circulating DNA and lung cancer. Ann N Y Acad Sci 1075:154–164
268. Yang B, Du Z, Gao YT, Lou C, Zhang SG, Bai T, Wang YJ, Song WQ (2010) Methylation of Dickkopf-3 as a prognostic factor in cirrhosis-related hepatocellular carcinoma. World J Gastroenterol 16:755–763
269. Yang Z, Zhou L, Wu LM, Xie HY, Zhang F, Zheng SS (2010) Combination of polymorphisms within the HDAC1 and HDAC3 gene predict tumor recurrence in hepatocellular carcinoma patients that have undergone transplant therapy. Clin Chem Lab Med 48:1785–1791
270. Ye C, Beeghly-Fadiel A, Lu W, Long J, Shu XO, Gao YT, Zheng W, Cai Q (2010) Two-stage case-control study of DNMT-1 and DNMT-3B gene variants and breast cancer risk. Breast Cancer Res Treat 121:765–769
271. Yu J, Zhang H, Gu J, Lin S, Li J, Lu W, Wang Y, Zhu J (2004) Methylation profiles of thirty four promoter-CpG islands and concordant methylation behaviours of sixteen genes that may contribute to carcinogenesis of astrocytoma. BMC Cancer 4:65
272. Yu L, Todd NW, Xing L, Xie Y, Zhang H, Liu Z, Fang H, Zhang J, Katz RL, Jiang F (2010) Early detection of lung adenocarcinoma in sputum by a panel of microRNA markers. Int J Cancer 127:2870–2878
273. Yuan Y, Wang J, Li J, Wang L, Li M, Yang Z, Zhang C, Dai JL (2006) Frequent epigenetic inactivation of spleen tyrosine kinase gene in human hepatocellular carcinoma. Clin Cancer Res 12:6687–6695
274. Zhang YJ, Wu HC, Shen J, Ahsan H, Tsai WY, Yang HI, Wang LY, Chen SY, Chen CJ, Santella RM (2007) Predicting hepatocellular carcinoma by detection of aberrant promoter methylation in serum DNA. Clin Cancer Res 13:2378–2384

275. Zhou X, Chen X, Hu L, Han S, Qiang F, Wu Y, Pan L, Shen H, Li Y, Hu Z (2010) Polymorphisms involved in the miR-218-LAMB3 pathway and susceptibility of cervical cancer, a case-control study in Chinese women. Gynecol Oncol 117:287–290
276. Zhu ZZ, Sparrow D, Hou L, Tarantini L, Bollati V, Litonjua AA, Zanobetti A, Vokonas P, Wright RO, Baccarelli A, Schwartz J (2011) Repetitive element hypomethylation in blood leukocyte DNA and cancer incidence, prevalence, and mortality in elderly individuals: the normative aging study. Cancer Causes Control 22:437–447

Chapter 14
Epigenetic Epidemiology of Infectious Diseases

Toshikazu Ushijima and Hideyuki Takeshima

Contents

14.1 Introduction 270
14.2 Infectious Agents and Methylation Induction 271
 14.2.1 The Significance of Methylation in Cancers and in Non-cancerous Tissues 271
 14.2.2 *Helicobacter pylori* Infection 273
 14.2.3 Hepatitis Viruses 274
 14.2.4 Epstein-Barr Virus 275
 14.2.5 Papilloma Virus 275
 14.2.6 Liver Flukes 276
 14.2.7 Schistosoma 276
14.3 Mechanisms for Methylation Induction by Infectious Agents 276
 14.3.1 Role of Chronic Inflammation 276
 14.3.2 Direct Effect of Infectious Agents 277
14.4 Methylation Fingerprints Produced by Specific Inducers 277
 14.4.1 Target Gene Specificity of DNA Methylation in Cancers 278
 14.4.2 Strategy to Analyze Methylation Fingerprints in Non-cancerous Tissues 278
 14.4.3 Methylation Fingerprint in Gastric Mucosae Exposed to *H. pylori* Infection 278
 14.4.4 Methylation Fingerprint in Esophageal Mucosae of Long-Term Smokers 280
14.5 Mechanisms for Formation of a Methylation Fingerprint 280
 14.5.1 Low Transcription Levels and DNA Methylation Susceptibility 280
 14.5.2 Trimethylation of Histone H3 Lysine 27 and DNA Methylation Susceptibility 281
 14.5.3 Stalled RNA Polymerase II and DNA Methylation Resistance 282
 14.5.4 Role of Genomic Position Relative to Repetitive Elements 282
 14.5.5 A Model for Formation of a Methylation Fingerprint of an Agent 282
14.6 Epilogue 283
References 284

T. Ushijima (✉) • H. Takeshima
Division of Epigenomics, National Cancer Center Research Institute, 1-1 Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan
e-mail: tushijim@ncc.go.jp; hitakesh@ncc.go.jp

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_14,

© Springer Science+Business Media B.V. 2012

Abstract Cancers induced by infectious agents, such as *Helicobacter pylori*, hepatitis viruses, Epstein-Barr virus, papilloma virus, liver fluke, and Schistosoma, are often associated with frequent DNA methylation of tumor-suppressor genes. Chronic inflammation and direct effects of some infectious agents are responsible for methylation induction. Analysis of non-cancerous, thus polyclonal, tissues provides unbiased information on which genes are methylated by specific agents. Gastric mucosae with *Helicobacter pylori* infection show methylation of specific promoter CpG islands. Promoter CpG islands of genes with low transcription levels and trimethylation of histone H3 lysine 27 are susceptible to methylation induction, and those with RNA polymerase II are resistant, even if the genes are not transcribed. Methylation of specific genes is expected to remain even after its inducing agent has vanished and these cytosine methylation fingerprints may prove to be good markers of past exposure to specific agents in molecular epidemiology.

Abbreviations

CGI	CpG island
CIMP	CpG island methylator phenotype
DNMT	DNA methyltransferase
EBV	Epstein-Barr virus
ESCC	Esophageal squamous cell carcinoma
H. pylori	*Helicobacter pylori*
H3Ac	Acetylation of histone H3
H3K27me3	Trimethylation of histone H3 lysine 27
H3K4me3	Trimethylation of histone H3 lysine 4
H3K9me3	Trimethylation of histone H3 lysine 9
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HPV	Human papilloma virus
LMP1	Latent membrane protein 1
MeDIP	Methylated DNA immunoprecipitation
MSP	Methylation-specific PCR
Pol II	RNA polymerase II
PRC	Polycomb repressive complex

14.1 Introduction

Infectious agents are some of the most well-known inducers of aberrant DNA methylation. Infection with *Helicobacter pylori* (*H. pylori*), a bacterial strain causally involved in gastric carcinogenesis, is known to induce aberrant DNA methylation in

gastric epithelial cells [1–3]. Infection with hepatitis C and B viruses (HCV and HBV), both of which are involved in the development of hepatocellular carcinomas, is associated with aberrant DNA methylation in cancer tissues and surrounding non-cancerous tissues [4–6]. Infection with Epstein-Barr virus (EBV), associated with lymphomas, nasopharyngeal cancers, and gastric cancers, is also associated with frequent DNA methylation in tumor tissue. Aberrant methylation induced by an infectious agent is present not in random genes, but in a group of genes affected by the agent [7]. Such gene-specific methylation provides an excellent tool for molecular epidemiology revealing past exposure to specific infectious agents, even if these agents are no longer present.

In this chapter, we will introduce methylation induction by infectious agents, its mechanisms, methylation of specific genes by specific inducers (a methylation fingerprint), and its mechanisms. We will conclude with the potential application to epidemiology.

14.2 Infectious Agents and Methylation Induction

Cancers induced by infectious agents, such as *H. pylori*, hepatitis viruses, EBV, papilloma virus, liver fluke, and Schistosoma, tend to have frequent aberrant DNA methylation of promoter CpG islands (CGIs) of tumor-suppressor and other genes (Table 14.1). In addition, non-cancerous tissues exposed to some infectious agents also show methylation of some genes, many of which are carried over into cancer cells.

14.2.1 The Significance of Methylation in Cancers and in Non-cancerous Tissues

Methylation of promoter CGIs of tumor-suppressor genes is frequently present in cancers induced by an infectious agent. This can be a result of frequent induction of methylation in the infected tissue by the agent, or a result of growth advantage of a cell with methylation of a tumor-suppressor gene and its clonal expansion. If methylation is present in non-cancerous (thus polyclonal) tissue after infection, it is unlikely that a cell with a rare event of methylation of a tumor-suppressor gene, which itself is unrelated to the infection, has been selected and expanded. Therefore, when tumor-suppressor genes or other genes whose methylation can confer growth advantage to a cell are analyzed, use of non-cancerous tissues provides more unbiased information on whether or not an infectious agent is capable of inducing DNA methylation.

Since non-cancerous tissues are polyclonal, high methylation levels of a specific gene means that more cells acquired methylation of this gene in the past. Therefore, high methylation levels can be associated with stronger tissue damage in the past

Table 14.1 Infectious agents and genes whose methylation is associated with an agent

	Methylated genes in cancer tissue	Methylated genes in non-cancerous tissue	Ref
Helicobacter pylori			
Gastric cancer			
	CDKN2A	*CDKN2A*	[2]
	FLNc	*FLNc*	[2]
	LOX	*LOX*	[2, 8]
	THBD	*THBD*	[2]
	miR-124a	*miR-124a*	[8]
	26 genes, including BDNF and IGFBP3	*26 genes, including BDNF and IGFBP3*	[9]
Hepatitis virus B (HBV)			
Hepatocellular carcinoma			
	APC	*APC*	[6]
	CDKN2A	*CDKN2A*	[6]
	RASSF1A	*RASSF1A*	[6]
Hepatitis virus C (HCV)			
Hepatocellular carcinoma			
	APC	*APC*	[5, 6]
	CDKN2A	*CDKN2A*	[6]
	RASSF1A	*RASSF1A*	[6]
Epstein-Barr virus (EBV)			
Gastric cancer			
	CDH1		[10]
	CDKN2A		[11, 12]
	MGMT		[11]
	PTEN		[13]
Papilloma virus (HPV)			
Cervical cancer			
	RASSF1A (unmethylated)[a]		[14, 15]
Head and neck cancer			
	RASSF1A (unmethylated)[a]		[16]
	SFRP4		[17]
Liver fluke			
Cholangiocarcinoma			
	MLH1		[18]
Schistosoma			
Bladder cancer			
	APC		[19]
	CDH1		[19]
	CDKN2A		[19]
	RASSF1A		[19]

[a]Considered to be due to the functional interaction between *RASSF1A* inactivation and HPV infection

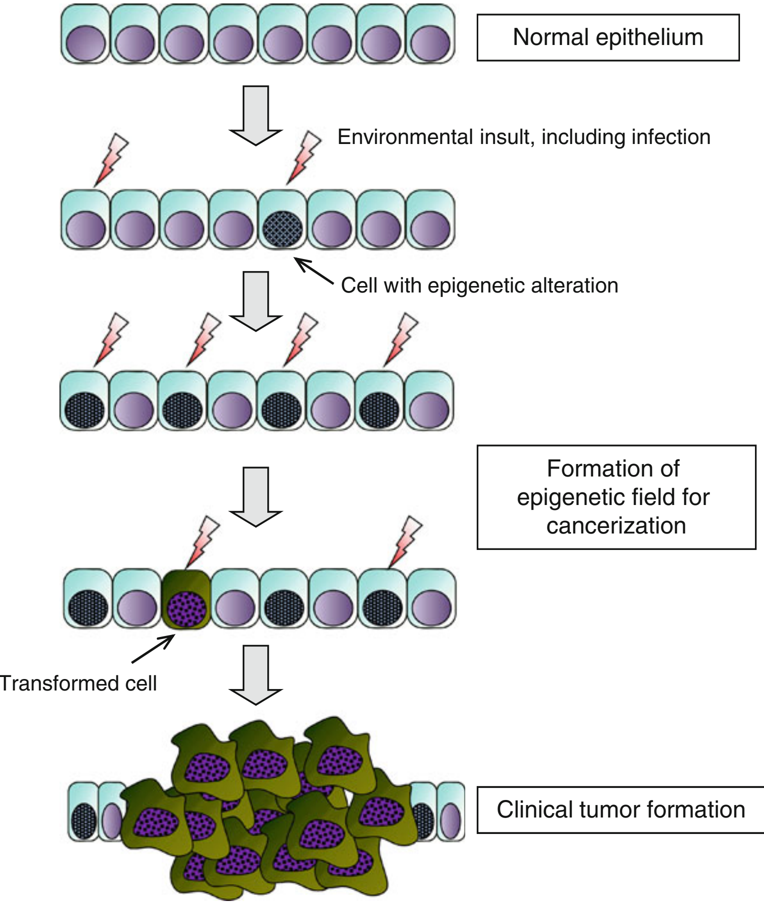


Fig. 14.1 Formation of an epigenetic field for cancerization. Exposure to environmental stimuli, including infection, induces epigenetic alterations in a significant fraction of cells in normal-appearing tissues. When such stimuli continue, epigenetic alterations accumulate in a tissue, and an epigenetic field for cancerization is produced

and accumulation of epigenetic alterations. Indeed, for cancers like gastric cancers and hepatocellular carcinomas (HCCs), methylation accumulation in non-cancerous tissues is known to form a field from which cancers tend to develop (an epigenetic field for cancerization; Fig. 14.1) [20]. Methylation levels in gastric mucosae are associated with risk of developing gastric cancers [20, 21].

14.2.2 Helicobacter pylori Infection

Gastric cancer is one of the leading causes of cancer deaths in the world, and most gastric cancer patients have a history of *H. pylori* infection [22]. *H. pylori* is a

Gram-negative bacterium, and is involved not only in gastric cancers but also in peptic ulcers and specific types of gastric B-cell lymphomas [23]. The infection itself is very prevalent, affecting nearly half of the world's population, although the prevalence is consistently decreasing [24]. *H. pylori* infection usually leads to acute and chronic gastritis, but only a small fraction of the infected individuals develop gastric cancer. Differences in *H. pylori* strains and host genetic backgrounds are considered to be responsible for the diverse clinical outcomes [25].

In gastric cancers, tumor-suppressor genes, such as *CDKN2A*, *MLH1*, *CDH1*, and *RUNX3*, are inactivated more frequently by methylation of their promoter CpG islands than by mutations [1]. Recent genome-wide analyses revealed that in addition to tumor-suppressor genes, hundreds of other genes are also methylated in gastric cancers [26, 27]. The possible association between *H. pylori* infection and methylation suggested by analysis of gastric cancers is solidified by analysis of non-cancerous gastric tissues with and without *H. pylori* infection [2, 9]. Methylation levels of eight specific CGIs in individuals with *H. pylori* infection were 5.4- to 303-fold as high as those in individuals without *H. pylori* infection ($P < 0.0001$). In addition to protein coding genes, microRNAs are also methylated in *H. pylori* infected tissues [8].

The accumulation level of aberrant DNA methylation in gastric mucosae is known to correlate with gastric cancer risk [2, 21]. Among individuals without current *H. pylori* infection, methylation levels in gastric mucosae were lowest in healthy individuals, high in cancer patients with a single gastric cancer, and highest in cancer patients with multiple gastric cancers (both metachronous and synchronous gastric cancers) [21]. The correlation strongly supports that the accumulation of aberrant methylation in gastric mucosae produces an epigenetic field for cancerization.

An animal model is available for gastric cancers induced by *H. pylori* infection [28]. Mongolian gerbils infected with *H. pylori* showed increased methylation levels compared with age-matched non-infected gerbils, demonstrating a causal role of *H. pylori* infection in methylation induction [3].

14.2.3 Hepatitis Viruses

HCCs are also a major cause of cancer deaths and are known to arise from liver tissues that heavily suffered from chronic inflammation, such as chronic hepatitis and liver cirrhosis [29]. As etiological agents, HCV and HBV play major roles, as well as some additional agents, such as excessive alcohol consumption and aflatoxin.

HCC tissues also have aberrant methylation of promoter CGIs of multiple tumor-suppressor genes [5, 6, 30]. Some genes, such as *CDKN2A*, are specifically methylated in HCCs while others, such as *RASSF1A*, are already methylated in surrounding liver tissues with chronic hepatitis or liver cirrhosis [30]. HCCs arising from liver cirrhosis have more methylated genes than those arising from chronic hepatitis, indicating that background tissue damage is reflected in HCCs [30]. When methylation

of 19 genes, mostly tumor-suppressor or tumor-associated genes, was analyzed, HCV-positive HCCs had more frequent methylation than had HBV-positive cancers and virus-negative HCCs [6].

Aberrant methylation of various genes is also present in surrounding non-cancerous liver tissues of patients with HCCs [6, 30, 31]. HCV-positive liver tissues show more frequent methylation than HBV-positive liver tissues and virus-negative liver tissues [6], strongly supporting HCV infection induces aberrant DNA methylation. Non-cancerous liver tissues of individuals with HBV and alcoholic hepatitis also had higher methylation levels than normal liver tissues of healthy individuals, indicating that these agents can induce aberrant DNA methylation in human liver tissues.

14.2.4 Epstein-Barr Virus

EBV infection is highly prevalent in the world, and is known to affect mainly B cells [32]. Its infection is implicated as an etiology of EBV-associated B-cell, T-cell, and NK-cell lymphomas and nasopharyngeal carcinomas. A minor fraction of gastric cancers is also associated with EBV infection, and such cancers are accompanied with strong infiltration of inflammatory cells [11, 33]. Gastric cancers with EBV infection are known to be associated with highly frequent methylation of multiple genes, including *CDH1*, *MGMT*, *PTEN*, and *p16*, demonstrating the CGI methylator phenotype (CIMP) [10–13, 33]. Surprisingly, non-cancerous gastric mucosae of patients with gastric cancers with EBV infection tended to have lower levels of methylation than those of patients with gastric cancers but without EBV infection, most of whom were infected by *H. pylori* [12].

14.2.5 Papilloma Virus

Human papilloma virus (HPV) is associated with cervical cancer, head and neck cancers, oral cancers, and possibly lung cancer [34–36]. In cervical cancers, HPV infection was associated with unmethylated status of *RASSF1A*, possibly due to their complementary functions in cervical carcinogenesis [14]. In contrast, in cervical dysplasia, from which cervical cancers develop, aberrant methylation of multiple genes, including *CCNA1*, *DAPK1*, *HS3ST2*, *PAX1*, and *TFPI2*, was present [37, 38], suggesting that HPV infection was associated with methylation induction.

In head and neck cancers, HPV infection was associated with unmethylated status of *RASSF1A*, again possibly due to their complementary functions in cervical carcinogenesis [16]. HPV infection was associated with *SFRP4* methylation in cancer tissues [17].

14.2.6 Liver Flukes

Liver flukes are parasites that infect the biliary tract, and are known to be a risk factor for cholangiocarcinoma, a rare malignant liver tumor arising from the intrahepatic biliary tract [39]. Like in other cancers, aberrant DNA methylation of tumor-suppressor genes, such as *APC*, *CDH1*, *CDKN2A*, and *MLH1* is frequently observed in cholangiocarcinoma [40]. Especially, *MLH1* methylation is frequently observed in liver fluke-related cholangiocarcinoma in Thailand [18].

14.2.7 Schistosoma

Schistosoma is a parasite that infects the urinary ducts and bladder, and Schistosoma infection is associated with risk of several malignancies, including bladder cancers [41]. Bladder cancers in general also show aberrant DNA methylation in multiple genes [42]. Notably, bladder cancers with Schistosoma infection had more methylated tumor-suppressor genes, including *APC*, *CDH1*, *CDKN2A*, and *RASSF1A*, than those without Schistosoma infection [19].

14.3 Mechanisms for Methylation Induction by Infectious Agents

Infectious agents are inducers of inflammation, which is known to be deeply involved in methylation induction [43, 44]. In addition, direct effects by infectious agents on cancer precursor cells need to be considered.

14.3.1 Role of Chronic Inflammation

The role of chronic inflammation in methylation induction was proposed based on the observation that aberrant methylation of specific genes was present in colonic mucosae of patients with ulcerative colitis [43, 45]. In addition, the infectious agents listed above are all associated with chronic inflammation. This strongly indicates that chronic inflammation is involved in methylation induction.

Direct evidence for the role of chronic inflammation in methylation induction was provided by an animal model of methylation induction by *H. pylori* infection. When inflammation induction by *H. pylori* infection was suppressed by treating Mongolian gerbils with cyclosporin A, an immunosuppressant, methylation induction was markedly suppressed without affecting colonization of *H. pylori* (Fig. 14.2). This showed that it is inflammation, not *H. pylori* itself, that is involved in methylation induction [3].

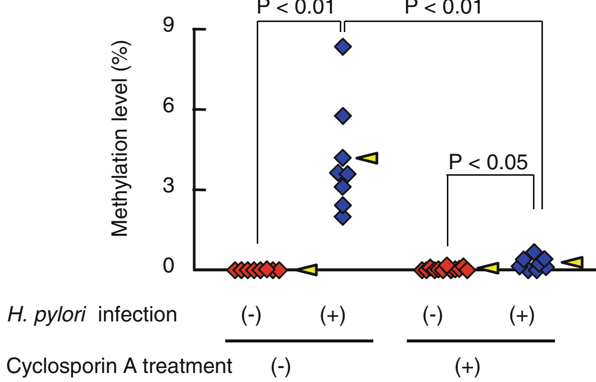


Fig. 14.2 A critical role of inflammation in methylation induction. The methylation level of a CGI (HE6) was measured by quantitative MSP in gastric mucosae of gerbils infected with *H. pylori* with or without cyclosporin A treatment. While a high level of methylation was induced by *H. pylori* infection without cyclosporin A, the induction was markedly suppressed with cyclosporin A. It was confirmed that cyclosporin A suppressed immune reaction, but did not affect colonization of *H. pylori*. Arrow head: mean methylation level in a group (Modified from Niwa et al. [3])

14.3.2 Direct Effect of Infectious Agents

In addition to signals from chronic inflammation to epithelial cells, some infectious agents have a direct effect on methylation induction. Especially, up-regulation of DNA methyltransferases by infectious agents has been reported for multiple viruses. The X protein of HBV (HBx) up-regulates DNMT1 and DNMT3A; the HCV core protein up-regulates DNMT1 and DNMT3B; HPV E7 and adenovirus E1A up-regulate DNMT1; and Kaposi's sarcoma-associated herpes virus LANA protein up-regulates DNMT3A [46–52]. The major EBV oncogene, latent membrane protein 1 (LMP1), up-regulates DNMT1, DNMT3A, and DNMT3B in nasopharyngeal carcinoma cell lines and induces methylation of the tumor-suppressor genes *RARB* and *CDH13* in these cell lines [49, 53]. LMP2A, another EBV latent gene, up-regulates DNMT1 and induces methylation of the tumor-suppressor gene, *PTEN*, in gastric cancer cell lines [11, 13]. In contrast with these viruses, DNMTs were not upregulated in human and gerbil gastric mucosae infected by *H. pylori* [3, 9].

14.4 Methylation Fingerprints Produced by Specific Inducers

It is likely that promoter CGIs of specific genes are methylated by a specific inducer, as described in this section. We here call this specific pattern of methylation "methylation fingerprint," as a mutation pattern specific to a mutagen is termed its

"mutational fingerprint" [54, 55]. The presence of methylation fingerprints was first suggested by analysis of cancer tissues, and was recently shown by analysis of non-cancerous tissues as well [56].

14.4.1 Target Gene Specificity of DNA Methylation in Cancers

The presence of target gene specificity in DNA methylation induction was initially indicated by the presence of methylation of specific genes in cancer cells [57–59]. A pioneering study of 1,184 CGIs using restriction landmark genomic scanning, an early-stage genome-wide scanning technique for differences in DNA methylation, revealed that certain CGIs were more frequently methylated in specific tumor types [57]. Analyses of promoter CGIs of mostly tumor-suppressor genes suggested that some CGIs were methylated at high incidences in specific tumor types [58]. Methylated DNA immunoprecipitation (MeDIP)-microarray analysis of colon cancer tissue revealed that most methylated genes were located within defined genomic clusters [59]. Comparison of 77 HCCs associated with either HBV or HCV revealed that some genes are preferentially methylated in HCCs, depending on the specific hepatitis virus [6].

14.4.2 Strategy to Analyze Methylation Fingerprints in Non-cancerous Tissues

Analysis of methylation in non-cancerous tissues should be conducted using a method with high sensitivity, such as methylation-specific PCR (MSP), because aberrant methylation of a gene is expected to be present only in a minor fraction of cells in a polyclonal tissue. Since methods with high sensitivity can analyze only selected genes, this requires pre-specification of genes to be studied. By a genome-wide methylation analysis of cancer tissues, we can obtain candidate genes methylated also in non-cancerous tissues. Using these pre-specified genes, we can perform a high-sensitivity methylation analysis in non-cancerous tissues [9, 60]. Recent methylome analyses yielded many promoter CGIs methylated in various cancers, including gastric, colon, lung, and breast cancers [26, 61–63]. High-sensitivity methylation analysis of these CGIs in non-cancerous tissues of patients with documented exposure to specific agents will lead to establishment of methylation fingerprints of the agents.

14.4.3 Methylation Fingerprint in Gastric Mucosae Exposed to *H. pylori* Infection

To clarify the presence of target gene specificity in methylation induction by *H. pylori*, methylation analysis of a set of candidate CGIs in gastric mucosae was

Fig. 14.3 The presence of a methylation fingerprint of *H. pylori* infection. Methylation of promoter CGIs of 48 genes was analyzed in gastric mucosae of individuals without and with *H. pylori* infection using a high-sensitivity method, MSP. Although genes 1–8 were not methylated at all, genes 22–47 were consistently methylated in individuals with *H. pylori* infection (Modified from Nakajima et al. [9])

conducted by searching for promoter CGIs methylated in gastric cancers [26]. MSP of 48 promoter CGIs in gastric mucosae with and without *H. pylori* infection suggested that specific genes were consistently methylated in gastric mucosae with *H. pylori* infection while others were not, forming a methylation fingerprint (Fig. 14.3) [9]. The susceptible genes had lower transcription levels in normal

gastric mucosae than had the resistant genes. Up-regulation of DNMTs was not observed in gastric mucosae with *H. pylori* infection.

14.4.4 Methylation Fingerprint in Esophageal Mucosae of Long-Term Smokers

Esophageal squamous cell carcinomas (ESCCs) are cancer types infamous for repeated occurrence, which is considered to be due to accumulation of genetic/epigenetic alterations in the background esophageal mucosae. We isolated 13 promoter CGIs methylated in ESCCs, and analyzed their methylation levels in esophageal mucosae. Methylation levels of 5 genes (*HOXA9*, *MT1M*, *NEFH*, *RSPO4*, and, *UCHL1*) were significantly correlated with smoking duration [60]. Although smoking is not an infectious agent, this finding in non-cancerous esophageal mucosae supports the theory that a specific agent induces methylation of specific genes, leaving a methylation fingerprint.

14.5 Mechanisms for Formation of a Methylation Fingerprint

A methylation fingerprint in non-cancerous tissues is formed because some specific genes are susceptible and others are resistant to methylation induction. Factors involved in such susceptibility and resistance have been studied for the past decade, and include low transcription levels, trimethylation of histone H3 lysine 27 (H3K27me3), and binding of RNA polymerase II (Pol II) (Fig. 14.4).

14.5.1 Low Transcription Levels and DNA Methylation Susceptibility

A low transcription level of a gene was proposed to be involved in methylation induction in the early 2000s [64–67]. Song et al. demonstrated that disruption of promoter activity (thus low transcription levels) of a transfected gene leads to aberrant DNA methylation of promoter CGIs in a cancer cell line [64]. de Smet et al. demonstrated that a gene demethylated by a DNA demethylating agent, 5-aza-2′-deoxycitidine, tends to be re-methylated when it is not transcribed [65]. The majority of genes methylated in cancer tend to have low transcription levels in normal cells [59, 66–68]. Among genes methylated in non-cancerous tissues, genes susceptible to methylation induction had lower transcription levels than resistant genes [9].

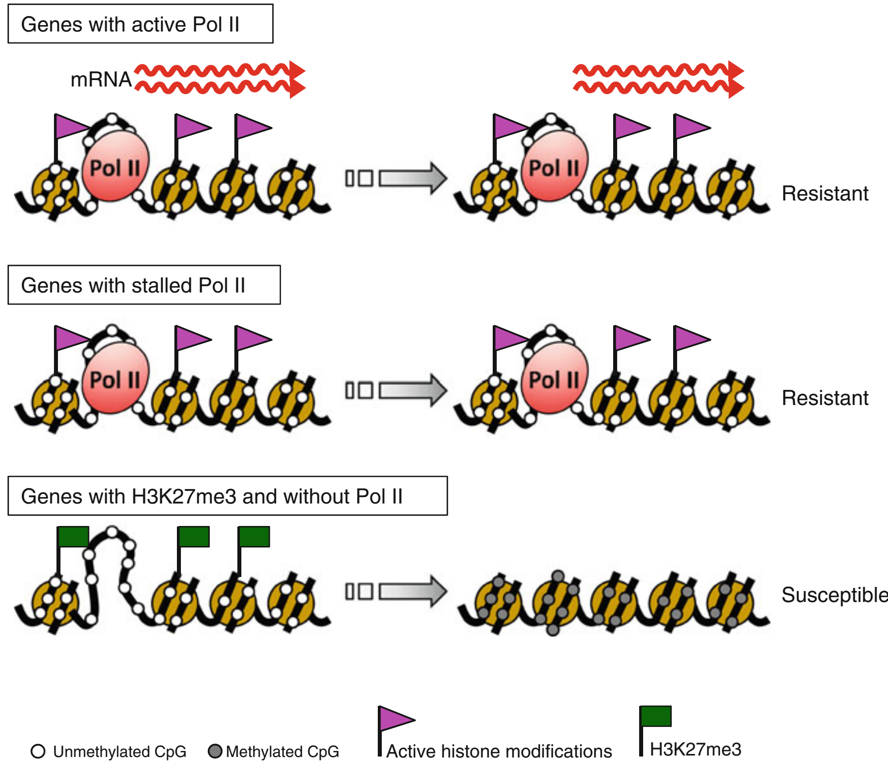


Fig. 14.4 Players involved in induction of DNA methylation in specific genes. Genes with active Pol II, 68% of which have active histone modifications (H3Ac), are resistant to DNA methylation induction. Genes with stalled Pol II, 19% of which also have active histone modifications (H3Ac), are also resistant. On the other hand, genes without Pol II, 90% of which are associated with H3K27me3 modification, are susceptible to methylation induction

14.5.2 Trimethylation of Histone H3 Lysine 27 and DNA Methylation Susceptibility

Histone modifications and DNA methylation depend upon each other. Using a limited number of genes, genes methylated in cancers were shown to be pre-marked by H3K27me3 in embryonic stem cells [69–71] and normal corresponding tissue [71]. This finding was also confirmed by genome-wide analyses [68, 72, 73]. H3K27me3 is recognized by a polycomb repressive complex (PRC) [74–76]. EZH2, a component of PRC2, and CBX7, a component of PRC1, are known to interact with DNA methyltransferases (DNMTs) [77, 78], and H3K27me3 may function as a recruiting signal for DNMTs. Another representative repressive histone modification, trimethylation of histone H3 lysine 9 (H3K9me3), was not associated with DNA methylation susceptibility [68]. Conversely, histone modifications

of active chromatin, such as acetylation of histone H3 (H3Ac) and trimethylation of histone H3 lysine 4 (H3K4me3), are weakly associated with resistance to DNA methylation [68].

14.5.3 Stalled RNA Polymerase II and DNA Methylation Resistance

Even among the genes with low transcription levels, some genes are still resistant to methylation induction. We have demonstrated that binding of Pol II to promoter CGIs (stalled Pol II) is associated with resistance to methylation induction [68]. Multivariate analysis of transcription levels, H3K27me3, H3Ac, and Pol II binding suggested that Pol II binding had a stronger effect on DNA methylation resistance than active histone modifications. Taken together with the fact that transcribed genes are resistant, binding of Pol II, active or stalled, is associated with resistance to methylation induction during carcinogenesis. Molecular mechanisms of how Pol II binding confers resistance to methylation induction remain to be clarified, but protection of CGIs from accession by DNMTs is one possible mechanism.

14.5.4 Role of Genomic Position Relative to Repetitive Elements

In addition, genomic architecture is also involved in gene-specificity of methylation induction. Compared with methylation-resistant genes, methylation-prone genes are located further apart from SINE and LINE retrotransposons [79]. However, since genome architecture does not change according to inducers of aberrant methylation, its role in the formation of a methylation fingerprint of an inducer is limited.

14.5.5 A Model for Formation of a Methylation Fingerprint of an Agent

As a mechanism of how a specific agent induces methylation of specific genes, we speculate that the agent first induces changes in transcription, H3K27me3, and binding of Pol II in its target genes, and that the changes then lead to methylation induction of the genes that acquired a susceptible epigenetic status (Fig. 14.5). Since the first changes in epigenetic status are consistently induced depending upon an agent, DNA methylation is expected to be induced in specific genes, forming a methylation fingerprint. This model is currently being validated using animal models.

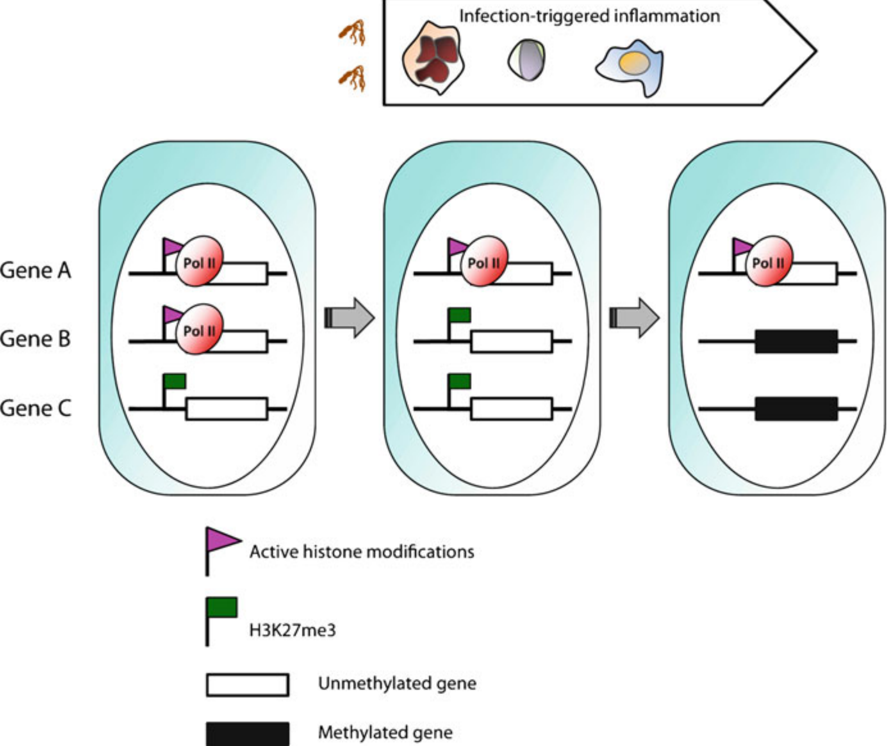


Fig. 14.5 A model for formation of a methylation fingerprint by an infectious agent. An agent is expected to induce changes in epigenetic statuses, including H3K27me3 and binding of Pol II. A gene that acquires a susceptible epigenetic status by exposure to an agent, such as *H. pylori* infection, is expected to become methylated (Gene B). Since the changes in epigenetic statuses are consistently induced by an agent, specific genes are expected to be methylated by the agent. A gene with Pol II (Gene A) and one with H3K27me3 (Gene C) regardless of the exposure to an agent are expected to be resistant and susceptible, respectively, to methylation induction. This model needs to be validated

14.6 Epilogue

Aberrant DNA methylation is induced by various infectious agents, and inflammation is an important element in the induction. Specific genes are methylated by specific inducers, forming methylation fingerprints. The potential application of methylation fingerprints in molecular epidemiology is to identify past exposure to infectious agents. An antibody to an agent usually gradually decreases after elimination of the agent, but a methylation fingerprint is expected to remain forever. So far, the number of known methylation fingerprints of specific agents is limited. However, a strategy to identify the fingerprint of an agent is already established if the agent induces cancer.

Once methylation fingerprints of various infectious agents become available, their use is expected to add new information to the involvement of infectious agents in various human disorders.

Acknowledgment The original research described here was supported by Grants-in-Aid for the Third-Term Comprehensive Cancer Control Strategy from the Ministry of Health, Labour and Welfare, Japan.

References

1. Ushijima T, Sasako M (2004) Focus on gastric cancer. Cancer Cell 5:121–125
2. Maekita T, Nakazawa K, Mihara M, Nakajima T, Yanaoka K, Iguchi M et al (2006) High levels of aberrant DNA methylation in *Helicobacter pylori*-infected gastric mucosae and its possible association with gastric cancer risk. Clin Cancer Res 12:989–995
3. Niwa T, Tsukamoto T, Toyoda T, Mori A, Tanaka H, Maekita T et al (2010) Inflammatory processes triggered by *Helicobacter pylori* infection cause aberrant DNA methylation in gastric epithelial cells. Cancer Res 70:1430–1440
4. Lee S, Lee HJ, Kim JH, Lee HS, Jang JJ, Kang GH (2003) Aberrant CpG island hypermethylation along multistep hepatocarcinogenesis. Am J Pathol 163:1371–1378
5. Yang B, Guo M, Herman JG, Clark DP (2003) Aberrant promoter methylation profiles of tumor suppressor genes in hepatocellular carcinoma. Am J Pathol 163:1101–1107
6. Nishida N, Nagasaka T, Nishimura T, Ikai I, Boland CR, Goel A (2008) Aberrant methylation of multiple tumor suppressor genes in aging liver, chronic hepatitis, and hepatocellular carcinoma. Hepatology 47:908–918
7. Takeshima H, Ushijima T (2010) Methylation destiny: Moira takes account of histones and RNA polymerase II. Epigenetics 5:89–95
8. Ando T, Yoshida T, Enomoto S, Asada K, Tatematsu M, Ichinose M et al (2009) DNA methylation of microRNA genes in gastric mucosae of gastric cancer patients: its possible involvement in the formation of epigenetic field defect. Int J Cancer 124:2367–2374
9. Nakajima T, Yamashita S, Maekita T, Niwa T, Nakazawa K, Ushijima T (2009) The presence of a methylation fingerprint of Helicobacter pylori infection in human gastric mucosae. Int J Cancer 124:905–910
10. Sudo M, Chong JM, Sakuma K, Ushiku T, Uozaki H, Nagai H et al (2004) Promoter hypermethylation of E-cadherin and its abnormal expression in Epstein-Barr virus-associated gastric carcinoma. Int J Cancer 109:194–199
11. Chang MS, Uozaki H, Chong JM, Ushiku T, Sakuma K, Ishikawa S et al (2006) CpG island methylation status in gastric carcinoma with and without infection of Epstein-Barr virus. Clin Cancer Res 12:2995–3002
12. Enomoto S, Maekita T, Tsukamoto T, Nakajima T, Nakazawa K, Tatematsu M et al (2007) Lack of association between CpG island methylator phenotype in human gastric cancers and methylation in their background non-cancerous gastric mucosae. Cancer Sci 98:1853–1861
13. Hino R, Uozaki H, Murakami N, Ushiku T, Shinozaki A, Ishikawa S et al (2009) Activation of DNA methyltransferase 1 by EBV latent membrane protein 2A leads to promoter hypermethylation of PTEN gene in gastric carcinoma. Cancer Res 69:2766–2774
14. Kuzmin I, Liu L, Dammann R, Geil L, Stanbridge EJ, Wilczynski SP et al (2003) Inactivation of RAS association domain family 1A gene in cervical carcinomas and the role of human papillomavirus infection. Cancer Res 63:1888–1893
15. Cohen Y, Singer G, Lavie O, Dong SM, Beller U, Sidransky D (2003) The RASSF1A tumor suppressor gene is commonly inactivated in adenocarcinoma of the uterine cervix. Clin Cancer Res 9:2981–2984

16. Dong SM, Sun DI, Benoit NE, Kuzmin I, Lerman MI, Sidransky D (2003) Epigenetic inactivation of RASSF1A in head and neck cancer. Clin Cancer Res 9:3635–3640
17. Marsit CJ, McClean MD, Furniss CS, Kelsey KT (2006) Epigenetic inactivation of the SFRP genes is associated with drinking, smoking and HPV in head and neck squamous cell carcinoma. Int J Cancer 119:1761–1766
18. Limpaiboon T, Khaenam P, Chinnasri P, Soonklang M, Jearanaikoon P, Sripa B et al (2005) Promoter hypermethylation is a major event of hMLH1 gene inactivation in liver fluke related cholangiocarcinoma. Cancer Lett 217:213–219
19. Gutierrez MI, Siraj AK, Khaled H, Koon N, El-Rifai W, Bhatia K (2004) CpG island methylation in Schistosoma- and non-Schistosoma-associated bladder cancer. Mod Pathol 17:1268–1274
20. Ushijima T (2007) Epigenetic field for cancerization. J Biochem Mol Biol 40:142–150
21. Nakajima T, Maekita T, Oda I, Gotoda T, Yamamoto S, Umemura S et al (2006) Higher methylation levels in gastric mucosae significantly correlate with higher risk of gastric cancers. Cancer Epidemiol Biomarkers Prev 15:2317–2321
22. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M et al (2001) Helicobacter pylori infection and the development of gastric cancer. N Engl J Med 345:784–789
23. Asaka M, Kato M, Takahashi S, Fukuda Y, Sugiyama T, Ota H et al (2010) Guidelines for the management of Helicobacter pylori infection in Japan: 2009 revised edition. Helicobacter 15:1–20
24. Peek RM Jr, Blaser MJ (2002) Helicobacter pylori and gastrointestinal tract adenocarcinomas. Nat Rev Cancer 2:28–37
25. McNamara D, El-Omar E (2008) Helicobacter pylori infection and the pathogenesis of gastric cancer: a paradigm for host-bacterial interactions. Dig Liver Dis 40:504–509
26. Yamashita S, Tsujino Y, Moriguchi K, Tatematsu M, Ushijima T (2006) Chemical genomic screening for methylation-silenced genes in gastric cancer cell lines using 5-aza-2′-deoxycytidine treatment and oligonucleotide microarray. Cancer Sci 97:64–71
27. Yamashita S, Hosoya K, Gyobu K, Takeshima H, Ushijima T (2009) Development of a novel output value for quantitative assessment in methylated DNA immunoprecipitation-CpG island microarray analysis. DNA Res 16: 275-286
28. Tsukamoto T, Mizoshita T, Tatematsu M (2007) Animal models of stomach carcinogenesis. Toxicol Pathol 35:636–648
29. Farazi PA, DePinho RA (2006) Hepatocellular carcinoma pathogenesis: from genes to environment. Nat Rev Cancer 6:674–687
30. Gao W, Kondo Y, Shen L, Shimizu Y, Sano T, Yamao K et al (2008) Variable DNA methylation patterns associated with progression of disease in hepatocellular carcinomas. Carcinogenesis 29:1901–1910
31. Kondo Y, Kanai Y, Sakamoto M, Mizokami M, Ueda R, Hirohashi S (2000) Genetic instability and aberrant DNA methylation in chronic hepatitis and cirrhosis – a comprehensive study of loss of heterozygosity and microsatellite instability at 39 loci and DNA hypermethylation on 8 CpG islands in microdissected specimens from patients with hepatocellular carcinoma. Hepatology 32:970–979
32. Young LS, Rickinson AB (2004) Epstein-Barr virus: 40 years on. Nat Rev Cancer 4:757–768
33. Kang GH, Lee S, Kim WH, Lee HW, Kim JC, Rhyu MG et al (2002) Epstein-barr virus-positive gastric carcinoma demonstrates frequent aberrant methylation of multiple genes and constitutes CpG island methylator phenotype-positive gastric carcinoma. Am J Pathol 160:787–794
34. Campisi G, Giovannelli L (2009) Controversies surrounding human papilloma virus infection, head & neck vs oral cancer, implications for prophylaxis and treatment. Head Neck Oncol 1:8
35. Deligeoroglou E, Christopoulos P, Aravantinos L, Papadias K (2009) Human papilloma virus molecular profile and mechanisms of cancerogenesis: a review. Eur J Gynaecol Oncol 30:128–132

36. Klein F, Amin Kotb WF, Petersen I (2009) Incidence of human papilloma virus in lung cancer. Lung Cancer 65:13–18
37. Wang KH, Liu HW, Lin SR, Ding DC, Chu TY (2009) Field methylation silencing of the protocadherin 10 gene in cervical carcinogenesis as a potential specific diagnostic test from cervical scrapings. Cancer Sci 100:2175–2180
38. Lim EH, Ng SL, Li JL, Chang AR, Ng J, Ilancheran A et al (2010) Cervical dysplasia: assessing methylation status (Methylight) of CCNA1, DAPK1, HS3ST2, PAX1 and TFPI2 to improve diagnostic accuracy. Gynecol Oncol 119:225–231
39. Chapman RW (1999) Risk factors for biliary tract carcinogenesis. Ann Oncol 10(Suppl 4):308–311
40. Tischoff I, Wittekind C, Tannapfel A (2006) Role of epigenetic alterations in cholangiocarcinoma. J Hepatobiliary Pancreat Surg 13:274–279
41. Bedwani R, Renganathan E, El Kwhsky F, Braga C, Abu Seif HH, Abul Azm T et al (1998) Schistosomiasis and the risk of bladder cancer in Alexandria, Egypt. Br J Cancer 77:1186–1189
42. Maruyama R, Toyooka S, Toyooka KO, Harada K, Virmani AK, Zochbauer-Muller S et al (2001) Aberrant promoter methylation profile of bladder cancer and its relationship to clinicopathological features. Cancer Res 61:8659–8663
43. Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA (2001) Accelerated age-related CpG island methylation in ulcerative colitis. Cancer Res 61:3573–3577
44. Ushijima T, Okochi-Takada E (2005) Aberrant methylations in cancer cells: where do they come from? Cancer Sci 96:206–211
45. Hsieh CJ, Klump B, Holzmann K, Borchard F, Gregor M, Porschen R (1998) Hypermethylation of the p16INK4a promoter in colectomy specimens of patients with long-standing and extensive ulcerative colitis. Cancer Res 58:3942–3945
46. Arora P, Kim EO, Jung JK, Jang KL (2008) Hepatitis C virus core protein downregulates E-cadherin expression via activation of DNA methyltransferase 1 and 3b. Cancer Lett 261:244–252
47. Burgers WA, Blanchon L, Pradhan S, de Launoit Y, Kouzarides T, Fuks F (2007) Viral oncoproteins target the DNA methyltransferases. Oncogene 26:1650–1655
48. Laurson J, Khan S, Chung R, Cross K, Raj K (2010) Epigenetic repression of E-cadherin by human papillomavirus 16 E7 protein. Carcinogenesis 31:918–926
49. Seo SY, Kim EO, Jang KL (2008) Epstein-Barr virus latent membrane protein 1 suppresses the growth-inhibitory effect of retinoic acid by inhibiting retinoic acid receptor-beta2 expression via DNA methylation. Cancer Lett 270:66–76
50. Shamay M, Krithivas A, Zhang J, Hayward SD (2006) Recruitment of the de novo DNA methyltransferase Dnmt3a by Kaposi's sarcoma-associated herpesvirus LANA. Proc Natl Acad Sci USA 103:14554–14559
51. Zheng DL, Zhang L, Cheng N, Xu X, Deng Q, Teng XM et al (2009) Epigenetic modification induced by hepatitis B virus X protein via interaction with de novo DNA methyltransferase DNMT3A. J Hepatol 50:377–387
52. Park IY, Sohn BH, Yu E, Suh DJ, Chung YH, Lee JH et al (2007) Aberrant epigenetic modifications in hepatocarcinogenesis induced by hepatitis B virus X protein. Gastroenterology 132:1476–1494
53. Tsai CN, Tsai CL, Tse KP, Chang HY, Chang YS (2002) The Epstein-Barr virus oncogene product, latent membrane protein 1, induces the downregulation of E-cadherin gene expression via activation of DNA methyltransferases. Proc Natl Acad Sci USA 99:10084–10089
54. Inga A, Iannone R, Campomenosi P, Molina F, Menichini P, Abbondandolo A et al (1995) Mutational fingerprint induced by the antineoplastic drug chloroethyl-cyclohexyl-nitrosourea in mammalian cells. Cancer Res 55:4658–4663
55. Kakiuchi H, Watanabe M, Ushijima T, Toyota M, Imai K, Weisburger JH et al (1995) Specific 5′-GGGA-3′–>5′-GGA-3′ mutation of the Apc gene in rat colon tumors induced by 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine. Proc Natl Acad Sci USA 92:910–914

56. Ushijima T, Asada K (2010) Aberrant DNA methylation in contrast with mutations. Cancer Sci 101:300–305
57. Costello JF, Fruhwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X et al (2000) Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. Nat Genet 24:132–138
58. Esteller M, Corn PG, Baylin SB, Herman JG (2001) A gene hypermethylation profile of human cancer. Cancer Res 61:3225–3229
59. Keshet I, Schlesinger Y, Farkash S, Rand E, Hecht M, Segal E et al (2006) Evidence for an instructive mechanism of de novo methylation in cancer cells. Nat Genet 38:149–153
60. Oka D, Yamashita S, Tomioka T, Nakanishi Y, Kato H, Kaminishi M et al (2009) The presence of aberrant DNA methylation in noncancerous esophageal mucosae in association with smoking history: a target for risk diagnosis and prevention of esophageal cancers. Cancer 115:3412–3426
61. Shen L, Toyota M, Kondo Y, Lin E, Zhang L, Guo Y et al (2007) Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. Proc Natl Acad Sci USA 104:18654–18659
62. Rauch TA, Zhong X, Wu X, Wang M, Kernstine KH, Wang Z et al (2008) High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. Proc Natl Acad Sci USA 105:252–257
63. Bediaga NG, Acha-Sagredo A, Guerra I, Viguri A, Albaina C, Ruiz Diaz I et al (2010) DNA methylation epigenotypes in breast cancer molecular subtypes. Breast Cancer Res (BCR) 12:R77
64. Song JZ, Stirzaker C, Harrison J, Melki JR, Clark SJ (2002) Hypermethylation trigger of the glutathione-S-transferase gene (GSTP1) in prostate cancer cells. Oncogene 21:1048–1061
65. de Smet C, Loriot A, Boon T (2004) Promoter-dependent mechanism leading to selective hypomethylation within the 5′ region of gene MAGE-A1 in tumor cells. Mol Cell Biol 24:4781–4790
66. Hagihara A, Miyamoto K, Furuta J, Hiraoka N, Wakazono K, Seki S et al (2004) Identification of 27 5′ CpG islands aberrantly methylated and 13 genes silenced in human pancreatic cancers. Oncogene 23:8705–8710
67. Furuta J, Nobeyama Y, Umebayashi Y, Otsuka F, Kikuchi K, Ushijima T (2006) Silencing of peroxiredoxin 2 and aberrant methylation of 33 CpG islands in putative promoter regions in human malignant melanomas. Cancer Res 66:6080–6086
68. Takeshima H, Yamashita S, Shimazu T, Niwa T, Ushijima T (2009) The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands. Genome Res 19:1974–1982
69. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C et al (2007) Epigenetic stem cell signature in cancer. Nat Genet 39:157–158
70. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L et al (2007) A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 39:237–242
71. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J et al (2007) Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet 39:232–236
72. Gal-Yam EN, Egger G, Iniguez L, Holster H, Einarsson S, Zhang X et al (2008) Frequent switching of polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. Proc Natl Acad Sci USA 105:12979–12984
73. Hahn MA, Hahn T, Lee DH, Esworthy RS, Kim BW, Riggs AD et al (2008) Methylation of polycomb target genes in intestinal cancer is mediated by inflammation. Cancer Res 68:10280–10289
74. Fischle W, Wang Y, Jacobs SA, Kim Y, Allis CD, Khorasanizadeh S (2003) Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by polycomb and HP1 chromodomains. Genes Dev 17:1870–1881

75. Min J, Zhang Y, Xu RM (2003) Structural basis for specific binding of polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev 17:1823–1828
76. Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, Monrad A et al (2008) A model for transmission of the H3K27me3 epigenetic mark. Nat Cell Biol 10:1291–1300
77. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C et al (2006) The polycomb group protein EZH2 directly controls DNA methylation. Nature 439:871–874
78. Mohammad HP, Cai Y, McGarvey KM, Easwaran H, Van Neste L, Ohm JE et al (2009) Polycomb CBX7 promotes initiation of heritable repression of genes frequently silenced with cancer-specific DNA hypermethylation. Cancer Res 69:6322–6330
79. Estecio MR, Gallegos J, Vallot C, Castoro RJ, Chung W, Maegawa S et al (2010) Genome architecture marked by retrotransposons modulates predisposition to DNA methylation in cancer. Genome Res 20:1369–1382

Chapter 15
Epigenetic Epidemiology of Inflammation and Rheumatoid Arthritis

Caroline Ospelt and Steffen Gay

Contents

15.1 Introduction 290
15.2 Epigenetics in Immune Function and Inflammation 291
15.3 Rheumatoid Arthritis 292
15.3.1 Epidemiology of RA 292
15.3.2 Epigenetics in RA 294
15.4 Conclusions and Perspectives 299
References 300

Abstract The field of epigenetic research has evolved dramatically in recent years and has given valuable insight in regulation and dysregulation of gene expression in health and disease. The knowledge of the mechanisms controlling epigenetic changes increased and in more and more diseases a role for epigenetics could be found. Rheumatoid arthritis is a chronic inflammatory, autoimmune disease which leads to joint destruction but also affects other tissues and organs. Its etiology is not yet clarified, but a combination of the genetic background and environmental factors is proposed to trigger the onset of the disease. Epigenetic changes might be the link between non-genetic risk factors and development of symptoms. Changes in DNA methylation and histone modifications have been detected in patients with RA and have been found to promote inflammation and joint destruction. Future studies will have to show which epigenetic changes are causative factors and which are induced at a later stage by the chronic inflammatory environment seen in the disease. This knowledge holds the potential for new preventive, predictive and therapeutic opportunities in RA.

C. Ospelt (✉) • S. Gay
Center of Experimental Rheumatology and Center of Human Integrative Physiology, University Hospital Zürich, Zürich, Switzerland
e-mail: caroline.ospelt@usz.ch; steffen.gay@usz.ch

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_15,

© Springer Science+Business Media B.V. 2012

Abbreviations

ACPA	anti-citrullinated peptide auto-antibodies
DNMT	DNA methyltransferase
EBV	Epstein Barr virus
H3	histone 3
H4	histone 4
HAT	histone acetylases
HDACs	histone deacetylases
HERV	human endogenous retrovirus
IGF	Insulin-like growth factor
IL	interleukin
MHC	major histocompatibility complex
PBMCs	peripheral blood mononuclear cells
RA	rheumatoid arthritis
SFN	Sulforaphane
SLE	systemic lupus erythematodes
TLRs	Toll-like receptors
Xi	inactivated X chromosome

15.1 Introduction

The classic definition of epigenetics describes epigenetic changes as mitotically heritable changes in gene function that in contrast to mutations do not alter the sequence of the DNA. However, the more epigenetics is studied in health and disease, the more facets and extensions have been added to this definition. Although an intriguing concept, the stable and mitotically heritable changes induced by the environment exclude a number of chromatin modifications and transcriptional regulators which, while short lived, profoundly influence gene function and might initiate changes that are passed on to the next generation of cells. In the current review, we follow the definition of Adrian Bird, describing epigenetic events as “the structural adaption of chromosomal regions so as to register, signal or perpetuate altered activity states” [1]. This definition not only describes the classic epigenetic systems such as methylation of DNA and of histone residues, but also includes more transient histone modifications such as acetylation and phosphorylation. Furthermore, it emphasizes the responsive nature of the epigenome, which is of particular interest in the context of the contribution of epigenetics to the development of diseases. Epigenetics may be the missing link between risk factors such as ageing, environmental toxins or co-morbidities and pathognomonic findings.

15.2 Epigenetics in Immune Function and Inflammation

The initiation of an immune response demands activation of regulatory processes to keep the frail balance between efficient defense and self destruction. Epigenetic modifications have been found to play an important role in the transcriptional regulation of genes modulating and coordinating the immune response at various levels.

The innate immune system gets activated via binding of pathogen-associated molecular patterns, so-called PAMPs by innate immune receptors. The best-studied innate immune receptors at the moment are Toll-like receptors (TLRs) which initiate a fast pro-inflammatory response against invading pathogens. It has long been known that repeated stimulation of TLR pathways, in particular of TLR4 pathways leads to the induction of tolerance against the inducing agent [2]. This mechanism controls the pro-inflammatory response and protects the organism from development of a septic shock. Foster et al. could show that there are two categories of genes that are induced by TLR4 stimulation [3]. The first set of genes is only induced at the first response, becomes then silenced and is not any more expressed after further TLR4 stimulations. This set mainly comprises pro-inflammatory mediators whose prolonged presence in tissues is detrimental and can induce tissue damage. The second set of genes is induced after every TLR4 stimulation. Even more, the magnitude and rapidness of their expression can even increase with repeated simulations. This set includes genes responsible for elimination of pathogens, expression of further immune receptors and activators of the adaptive immune response. After an initial TLR stimulation both sets of genes undergo specific epigenetic modifications such as histone acetylation and methylation which shape innate immune responses, improve the elimination of the pathogen and prevent inflammatory tissue damage.

Not only innate immunity, but also adaptive immunity gets modulated by epigenetic mechanisms. The specific cytokine pattern produced by Th1 and Th2 T cell subtypes for instance could be shown to be regulated by histone acetylation and methylation. While in Th1 cells expression of interferon γ and interleukin (IL)-4 is repressed by histone methylation, Th2 cells have no histone methylation at this locus but instead show hyperacetylation of histones at the *IL-4-IL-13* promoter site facilitating gene expression [4–7]. Also, expression of IL-2, a key regulator of T cell differentiation and survival is regulated by histone deacetylases (HDACs) [8].

Since in addition the expression of a variety of pro-inflammatory cytokines such as IL-1, IL-8 and IL-12 has been found to be controlled by HDACs, it can be assumed that modulation of the histone code is an important general mechanism to contain the inflammatory response [9–11]. Accordingly, disturbances in epigenetic control mechanisms might promote excessive, chronic or self-directed immune responses resulting in chronic inflammatory diseases and autoimmunity.

15.3 Rheumatoid Arthritis

To scale the influence of epigenetic changes on disease development, the case of monozygotic twins with discordant disease is often drawn upon. In celiac disease concordance in monozygotic twins is more than 80%, suggesting a strong genetic component in this disease [12, 13]. In rheumatoid arthritis (RA) the concordance rate between monozygotic twins is about 15% [14, 15]. However, concordance rates are lower among diseases with low prevalence. Heritability, which estimates the genetic contribution, is independent of disease prevalence and is about 60% for RA [16]. Nevertheless, these data suggest a substantial role of non-genetic factors in RA etiology. In the following, we will discuss current evidence on the contribution of non-genetic factors to the development of RA and the role of epigenetic mechanisms.

15.3.1 Epidemiology of RA

RA is a systemic autoimmune disease which simultaneously affects multiple joints and ultimately leads to irreversible damage of articular structures. Destructive arthritis primarily affects small joints of the hands and feet, but larger joints are also affected during the course of the disease. In a murine disease model circulating synovial fibroblasts in the blood were connected to this typical pattern of joint involvement [17]. The most common extra-articular manifestations of RA are subcutaneous nodules, serositis, and amyloidosis.

15.3.1.1 Geography

Worldwide, 0.5–1% of the population is affected by RA. Incidence rates vary between countries; RA was reported to newly affect 9/100,000 individuals per year in France and 45/100,000 individuals per year in the U.S. [18, 19]. However, these differences may be, at least partly, due to methodological differences, making it difficult to compare data from different studies. The prevalence of RA differs between regions and is generally lower in less economically developed countries [20, 21]. While genetic variance, environmental and socioeconomic factors affect disease prevalence, data from epidemiologic studies may reflect underdiagnosis due to lack of access to healthcare. In addition, lower life expectancy may also contribute to the low prevalence of RA in developing countries. Studies measuring disease rates within a country are particularly useful in separating the contribution of genetic and non-genetic factors. High prevalence rates in population subgroups point to genetic components, whereas higher rates of disease in specific geographical areas are indicators of environmental influences. Examples for the first scenario are studies indicating a very high RA prevalence in Native-Americans [22]. A larger genetic contribution can be assumed in these population groups, although social habits and living conditions may also play a role. Since the incidence and prevalence of RA in

Pima Indians has significantly decreased over a 25-year time period, an important environmental contribution cannot be excluded [23]. A Finnish study suggested geographical variation in the incidence of RA in Finland with more than 2.5 fold higher incidence rate in north-eastern compared to north-western Finland [24]. These variances could not be explained by genetic differences or by differences in health care or diagnostics. In central Finland, RA incidence rates declined from 45/100'000 in 1980 to 27/100'000 in 1995. Worldwide the incidence of RA has declined in many countries, e.g. in the United States or Japan [19, 25]. Since the genetic background in a population hardly changes in such a short time-period, these studies support a role of environmental factors affecting RA etiology.

15.3.1.2 Sex

Like most autoimmune diseases, RA is more common in women than in men, however, the ratio changes with age. While in women the RA incidence rate peaks at around 55 years of age and then remains stable, it increases in men up to an age of about 85 years. Accordingly, the sex ratio is highest in the mid-forties with 4:1, declines in the mid-fifties to about 2:1 and levels in women and men over the age of 70 years [19, 26]. A possible explanation is the role of sex hormones on inflammation and immune response, but mechanisms are insufficiently understood. Studies explored the influence of reproductive factors on RA risk, but provided inconsistent results. Some studies identified breastfeeding as possible inducer of postpartum RA flares [27, 28], while others could not confirm these results [29], or even found a protective role of breastfeeding on RA development [30]. Similarly, reports on the influence of age at menarche, use of oral contraceptives and exogenous hormones are inconclusive [30–32].

15.3.1.3 Lifestyle

Occupational risk to develop RA was mainly described in professions where workers were exposed to silica [33]. Also exposure to mineral oils and mineral dust seems to lead to an increased risk of RA [34].

It remains unresolved whether infectious agents may influence the development of RA. One of the most prominent candidates is the Epstein Barr virus (EBV), but studies analyzing the presence of EBV in RA patients were often limited by technical shortcomings and results were conflicting [35–37]. However, RA patients consistently have higher titers of EBV antibodies compared to healthy controls [38, 39]. These may result from a generally dysregulated immune response in RA patients, since a clear link between EBV infection and RA is missing [40]. A variety of other infectious agents, e.g. parvovirus B19, *Mycoplasma pneumoniae* or *Proteus mirabilis* have also been implicated in increased risk of RA, but none provide strong and consistent evidence for a causal role in RA [41].

Studies on the role of dietary factors in RA etiology are compromised by the complexity of diet, making it difficult to single out the possible contribution of

individual foods or nutrients and by the presence of unidentified confounding factors that cannot be corrected for. Conflicting results do not permit conclusions on the role of consumption of fruits and vegetables, dairy products, coffee or alcohol [42, 43]. More consistent evidence links anti-oxidants and RA. High dietary intake of antioxidants was found to decrease the risk for RA in the Iowa Women's Health Study cohort of older women [44], and several studies found lower serum levels of antioxidants in individuals who subsequently developed RA compared to healthy individuals [45–47]. However, no association was found between intake of antioxidants from foods and supplements and the risk of RA in the large prospective Nurses' Health Study [48].

The role of smoking in RA etiology is supported by several studies; two studies have also described an elevated risk in past smokers until at least 10 years after cessation [49–52]. A gene-environment interaction was identified in smokers that carry risk alleles within the *HLA-DRB1* locus. The *HLA-DRB1* gene codes for the β-chain of the major histocompatibility complex (MHC) class II receptor on antigen-presenting cells and is highly allelic diverse. *HLA-DRB1* RA-risk alleles have a common amino acid sequence, which is important in peptide binding and is therefore known as "shared epitope". Depending on the type of risk allele, individuals carrying shared-epitope alleles are 1.5–3 times as likely to develop RA compared to non-carriers. Sub-classification of RA patients into those with anti-citrullinated peptide auto-antibodies (ACPA) and without, revealed that shared epitope alleles mainly influence the development of ACPA in RA patients [53]. With an odds ratio 5.27 compared to shared epitope negative non-smokers, smoking is a strong risk factor for the development of ACPA positive RA in individuals with the shared epitope but not in individuals without any risk alleles [54, 55].

15.3.2 Epigenetics in RA

In RA, the field of epigenetics is only at its beginnings, but already now, interesting data about how epigenetics might contribute to the pathogenesis of RA and how it could be connected to the above mentioned risk factors for RA are emerging. One reason for the failure to identify environmental risk factors for RA in humans might be the focus of studies on the time around disease onset, whereas environmental factors might act much earlier and lead to epigenetic changes that predispose for the development of RA.

15.3.2.1 DNA Methylation

One of the main tasks of epigenetic research nowadays is the mapping of genes that are susceptible to environmentally induced changes in DNA methylation in humans. This knowledge is a prerequisite for the in-depth analysis of epigenetic dysregulation in disease.

Transposable elements:
It is well known that in humans the expression of transposable elements is regulated by DNA methylation [56]. Transposable elements are mobile DNA sequences that can move to different positions in the genome either by transcription/reverse transcription (retrotransposons) or by a cut and paste mechanism (DNA transposons). These remnants of ancient infections comprise around 45% of the human genome and have a sustained effect on human cell biology and evolution by influencing gene expression and genomic organization. However, insertion and/or expression of transposable elements can interfere with normal cellular processes and lead to diseases such as transposon induced coagulopathies. In this case insertion of transposons disrupts gene function. On the other hand, demethylation of transposal elements leads to their transcription which can increase expression of neighboring genes by providing promoter functions or can disturb cell physiology by translation of transposon gene products. The product of the retrotransposon *HRES-1*, which belongs to the family of human endogenous retrovirus (HERV), for instance was shown to suppress CD4 expression on human CD4+ T cells [57]. Most important for autoimmune diseases, transposon products, in particular HERV proteins can induce an antibody response and these antibodies can cross-react with auto-antigens [58]. In patients with the autoimmune disease systemic lupus erythematodes (SLE), expression of the HERV-E clone 4–1 *gag* transcripts correlated with the presence of the anti-nuclear antibodies anti-U1 ribonucleoprotein and anti-Sm [59]. Furthermore HERV-E clone 4–1 transcripts were only found to be expressed in peripheral blood mononuclear cells (PBMCs) from SLE patients but not from healthy controls and their expression could be induced by DNA demethylation with 5-azacytidine [60]. Similar results in RA suggested that the retrotransposable element LINE-1 is expressed in synovial fibroblasts from RA patients but not from patients with osteoarthritis (OA) and its expression *in vitro* was also induced by DNA demethylation [61]. Together these data clearly point to the fact that aberrant expression of transposable elements due to promoter hypomethylation plays a role in autoimmune diseases. Why they are re-activated and how they influence cell physiology however has yet to be clarified.

Imprinted genes:
In addition to transposable elements, imprinted genes are particularly sensitive to changes in DNA methylation. A recent study suggests that loss of imprinting may play a role in the pathogenesis of RA. A subset of RA synovial fibroblasts was found to have loss of imprinting at the *Insulin-like growth factor* (*IGF-2)* gene locus [62]. High expression of IGF-2 was further found to correlate with a low-inflammatory type of RA, suggesting that in a subgroup of RA patients, loss of imprinting of the *IGF-2* gene in synovial fibroblasts leads to higher expression of IGF-2 promoting synovial fibroblast proliferation and an inflammation-independent type of synovial invasion. It is not clear however, at which stage of the disease loss of imprinting occurs and whether this phenomenon really causes a specific disease subtype or is a mere bystander effect that occurs in a subgroup of patients.

Promoter specific changes in methylation:
Expression of transposable elements and loss of imprinting in RA patients both generally point to a hypomethylated state in this disease. And indeed in RA synovial tissues in particular in synovial fibroblasts DNA is hypomethylated which could be linked to decreased expression of DNA methyltransferase (DNMT)-1, the enzyme that promotes methylation of DNA [63]. In addition to global hypomethylation, also specific loci were found to be hypomethylated in synovial fibroblasts of RA patients. Increased production of the chemokine CXCL12 by RA synovial fibroblasts could be shown to be caused by hypomethylation of the CXCL12 promoter in these cells [93]. Decrease in DNA methylation was also found in PBMCs and in T cells from patients with RA [64]. Further analysis of the 'senescent' subset of T cells, namely CD4+ CD28-, which are mainly found in elderly people and in people with chronic inflammatory diseases such as RA, revealed that demethylation leads to the specific gene expression pattern seen in these cells. Due to repetitive stress, these T cells downregulate ERK and JNK signalling pathways which was connected to loss of methylation via downregulation of DNMT1 and DNMT3a [65]. In PBMCs it could furthermore been shown that in RA patients methylation is lost at a specific site in the *IL-6* gene promoter. Lack of methylation at this site was connected to higher levels of IL-6 transcripts after stimulation [66]. Even though most studies with RA patients found global and site-specific hypomethylation, single DNA sites can also be hypermethylated in RA. Takami et al. found higher methylation rates in the promoter of the death receptor *DR3* in synovial fibroblasts of RA patients compared to OA patients. Accordingly, protein levels of the DR3 receptor were found to be lower in RA synovial fibroblasts compared to OA synovial fibroblasts, which might be the explanation for the lack of apoptosis seen in RA synovial fibroblasts after application of the physiologic ligand of DR3 [67]. Even though a direct correlation between promoter methylation status and mRNA levels of DR3 is missing, these data provide evidence that despite of global demethylation single promoter regions might be hypermethylated as also seen in tumor cells [68].

In summary these studies suggest a dysregulated methylation pattern in peripheral blood cells and local cells in the joints of RA patients. However, since biospecimen were obtained from patients with manifested RA, the temporal relation between changes in DNA methylation and RA remains unclear, and it is possible that the methylation changes result from chronic inflammation. Some epigenetic changes as seen in CD4+ CD28- cells might result from repetitive stress of immune cells during chronic inflammation and may evolve during disease and perpetuate the inflammatory process. Since the enzymes transferring methylation marks after mitosis have a much higher error rate compared to DNA polymerases, differences in methylation patterns accumulate with every mitotic cycle; this may explain why the risk of developing a variety of diseases including RA increases with age [69]. Some epigenetic changes may already be induced years before onset of disease, since peri- and postnatal nutrition and environment impacts the epigenome in animal models. Even though some epidemiological studies suggest a connection between incidence rates of RA and year of birth, the underlying events are not clear and studies on the early life origins of RA patients are scarce and some results conflicting [19, 70]. Associations between

being breastfed and risk of RA could be shown in one study, but were not confirmed by another [71, 72]. Whereas no relation between birth weight and Rheumatoid Factor positivity could be found, there seems to be a higher risk of developing RA associated with high birth weight [71, 73, 74].

Another lifestyle factor that might affect methylation is smoking; a recent study detected changes in the methylation pattern in colorectal biopsies in smokers and proposed a potential causal relation between smoking and hypomethylation [75]. In conjunction with the genetic background, smoking-induced hypomethylation may lay the groundwork for the development of various diseases including tumors and autoimmune diseases.

15.3.2.2 Histone Modifications

In addition to DNA methylation, epigenetic regulation also occurs via histone modifications. These modifications are more dynamic and varied and include acetyl, methyl, phosphate and ubiquitin residues which can be placed at different sites on the histone tails. Not only influence the different histone modifications each other, but they are also influenced by DNA methylation creating an intricate system. The combination of different modifications at specific positions at a certain stage of transcription tips the scale for transcriptional repression or activation [76]. The complexity of this system has led investigators who explore altered histone modifications in disease to mostly concentrate on specific modifications at specific sites. Such data can only provide a limited picture of the complex interactions *in vivo*. Most intensely studied are acetylation and methylation of histone 3 (H3) and histone 4 (H4). Histone acetylation is commonly associated with transcriptional activation, whereas deacetylation by histone deacetylases (HDACs) leads to transcriptional repression. In CD4+ T cells of SLE patients global H3/H4 acetylation was decreased which suggests transcriptional repression. Accordingly levels of histone acetylases (HAT) were significantly lower in these patients [77]. However, HDAC2 and HDAC7 levels were also lower in SLE patients than in controls suggesting that not the expression of single HAT and HDAC but the balance between total HAT and HDAC activity may be relevant for global histone acetylation levels. In synovial tissues of RA patients this balance was found to be shifted towards higher levels of HAT with hyperacetylated histones promoting gene expression [78]. In contrast, another study found increased activity and expression of HDAC1 in RA synovial tissues [94]. These discrepancies might be caused by differences in the treatment and the lifestyle of the study patients, since levels of TNF and nutritional factors can influence the expression of HDACs. Also in peripheral blood cells of RA patients HDAC activity was higher than in healthy controls [95]. However, global histone acetylation has not been measured in RA up to now and given the intricate interactions between different histone modifications it is questionable whether a single modification impacts gene expression.

Another hypothesis linking histone modifications to the development of autoimmune diseases suggests that changes in the histone code lead to the appearance of

neo-epitopes which, after cell death, might be recognized by lymphocytes and induce an autoimmune response [79]. In SLE and in chronic inflammatory bowel disease auto-antibodies against the mono-ubiquitinated H2A are regularly found and also auto-antibodies against poly (ADP) ribose, another histone modification, are present in the serum of these patients [80–82]. However, the appearance of these auto-antibodies may be the consequence and not the cause of an aberrantly activated and disturbed immune system.

Histone modifications are much more dynamic than DNA methylation and a variety of external stimuli likely have an impact on the histone code. Sulforaphane (SFN) is an organosulfur compound that is present at high levels in cruciferous vegetables such as cabbage, broccoli and horseradish and is a naturally occurring HDAC inhibitor. Study participants were asked to consume one cup of broccoli sprouts after 48 h of a cruciferous vegetable-free diet and HDAC activity and H3/H4 acetylation was measured in PBMCs [83].Three hours after broccoli consumption a strong downregulation of HDAC activity was detected and H3 and H4 were hyper-acetylated accordingly. This effect was still detectable 48 h after consumption of the broccoli sprouts. SFN is not the only naturally occurring HDAC inhibitor and the authors speculate that long-term intake of such modulators of the histone code may prime cells for an appropriate response to exogenous insults [84]. A combination of different dietary modulators of chromatin remodelling may interact to induce constant subtle changes in the epigenetic state, and this dynamic state may reduce susceptibility to diseases with epigenetic components such as chronic inflammation, autoimmunity or cancer.

Certain histone modifications can be regulated by the activation of the transcription factor NF-κB. NF-κB is a key transcription factor in inflammation and cell growth and its activation leads to expression of a variety of pro-inflammatory cytokines and chemokines [85]. Saccani et al. observed that not all NF-κB target genes are accessible for NF-κB binding and that H3 phosphorylation/phosphoacetylation or H4 acetylation needs to be increased before NF-κB can bind to the promoter of *IL-6*. Their data also suggest that different stimuli induce histone modifications at different gene locations, thereby shaping the resulting inflammatory response [86, 87]. H3 methylation seems to have a crucial role in basal and post-inductional repression of a subgroup of NF-κB inducible genes, tightly regulating gene expression and leaving only a small time window for activation [88]. These examples of NF-κB regulation by histone modifications underline the importance of epigenetics in the regulation of inflammation and indicate that small but lasting changes in the epigenetic code may have a sustained effect on inflammatory and immune responses.

15.3.2.3 Sex Difference

One of the main questions regarding the occurrence of autoimmunity is why women are generally more frequently affected than men. Among different explanations including hormone status in particular estrogen levels and differences in environmental exposures, hypotheses relating to the sex chromosomes have recently drawn

increasing attention. During embryogenesis one of the two X chromosomes in the cells of female mammals gets inactivated. In humans the choice between paternal or maternal X chromosome is randomly, but once made, the choice is permanent for the cells. Silencing is achieved by different epigenetic mechanisms. The future inactivated X chromosome (Xi) expresses a large non-coding RNA from the *Xist* (X inactive specific transcript) gene which binds to Xi and suppresses gene expression. Inactivation of Xi is completed by histone ubiquitination, methylation and loss of acetylation and DNA methylation [89]. The combination of these epigenetic modifications makes Xi silencing irreversible albeit not complete. Only 75% of the genes on the Xi are constantly silenced, 15% are constantly expressed and another 10% are differential expressed between individuals [90]. The fact that there are genes that are constantly expressed at a higher level in females than in males and that the expression of some of these genes also varies between females might be of importance in the development of diseases. Furthermore, Xi chromosome reactivation may play a role. Methylation of the promoter region of the *CD40LG* gene on the X chromosome was diminished in women with SLE and transcript levels of this B cell co-stimulatory molecule were accordingly higher in affected women compared to healthy women, but interestingly also compared to affected men [91]. However, direct proof that silencing *in vivo* was lost at this specific site of Xi in female SLE patients is lacking. Finally, skewed X chromosome inactivation has been implicated in disease development. In this case, inactivation of the X chromosomes is not random, but either the maternal or the paternal X chromosome is preferentially silenced. The cause of skewed inactivation is not clear yet. Although skewed X inactivation has not been observed in RA, SLE, inflammatory bowel disease or multiple sclerosis patients, a higher frequency of skewed inactivation patterns was found in peripheral blood cells of patients with scleroderma compared to healthy controls [92].

15.4 Conclusions and Perspectives

In recent years, epigenetic research has drastically increased and provided interesting and exciting insights in gene regulation and transcriptional control mechanisms. However, we are only beginning to understand how cells react to an altered environment via the intricate network of epigenetic modifications. Depending on the cell type every individual gene has a particular pattern of DNA methylation and histone modifications which change with its state of activation and must be tightly regulated. Environmental factors such as nutrition, toxins, or infections that can interfere with the epigenome may induce epigenetic changes over time that persist even in the absence of the triggering factor and shape response mechanisms.

Some epigenetic changes that are seen in chronic inflammation and autoimmunity may be induced by consistently aberrant activation of certain signalling pathways and perpetuate disease by facilitating accessibility to these pathways like a well-trodden path. As illustrated in Fig. 15.1 epigenetic changes might also be induced before disease onset by seemingly disease-unrelated factors shifting cellular pathways towards a vulnerable state that together with a susceptible genetic

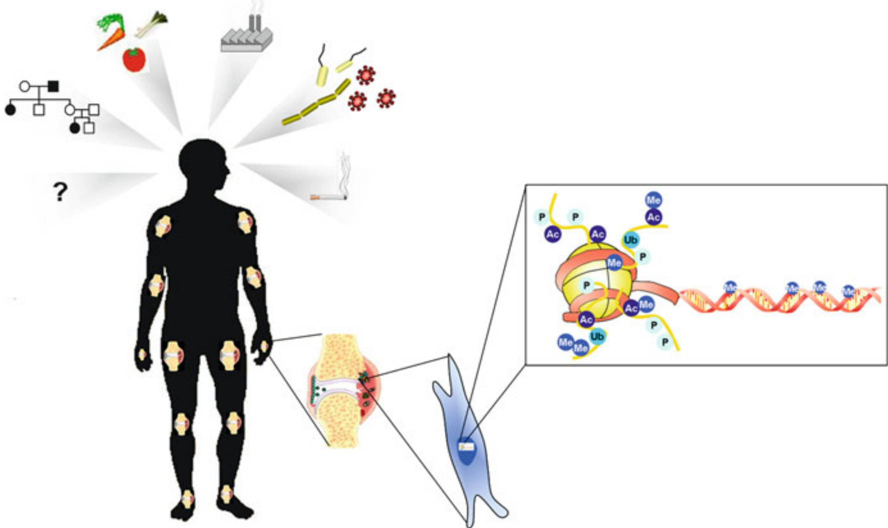


Fig. 15.1 The combination of a variety of different factors such as smoking, infections, pollution, nutrition, genetic predisposition and other factors may lead to changes in histone modifications and DNA methylation in local cells of the joints, thereby supporting the onset of chronic inflammation and autoimmunity (*P* phosphorylation, *Ac* acetylation, *Me* methylation, *Ub* ubiquitination)

background or additional exogenous factors triggers disease. One key question for future research will therefore be which epigenetic changes occur before disease onset and which occur as a result of disease. Answers to this question will not only offer novel insights in disease pathogenesis, but also provide new preventive, predictive and therapeutic opportunities.

Future epidemiologic studies will benefit from additional epigenetic insights and start to integrate epigenetic data. The combination of epidemiologic, genetic, and epigenetic data can be a powerful tool to clarify susceptibility, outcome, and therapeutic response not only in RA but in a variety of chronic inflammatory diseases.

References

1. Bird A (2007) Perceptions of epigenetics. Nature 447:396–398
2. West MA, Heagy W (2002) Endotoxin tolerance: a review. Crit Care Med 30:S64–S73
3. Foster SL, Hargreaves DC, Medzhitov R (2007) Gene-specific control of inflammation by TLR-induced chromatin modifications. Nature 447:972–978
4. Avni O, Lee D, Macian F, Szabo SJ, Glimcher LH, Rao A (2002) T(H) cell differentiation is accompanied by dynamic changes in histone acetylation of cytokine genes. Nat Immunol 3:643–651
5. Baguet A, Bix M (2004) Chromatin landscape dynamics of the Il4-Il13 locus during T helper 1 and 2 development. Proc Natl Acad Sci USA 101:11410–11415

6. Koyanagi M, Baguet A, Martens J, Margueron R, Jenuwein T, Bix M (2005) EZH2 and histone 3 trimethyl lysine 27 associated with Il4 and Il13 gene silencing in Th1 cells. J Biol Chem 280:31470–31477
7. Chang S, Aune TM (2007) Dynamic changes in histone-methylation 'marks' across the locus encoding interferon-gamma during the differentiation of T helper type 2 cells. Nat Immunol 8:723–731
8. Wang J, Lee S, Teh CE, Bunting K, Ma L, Shannon MF (2009) The transcription repressor, ZEB1, cooperates with CtBP2 and HDAC1 to suppress IL-2 gene activation in T cells. Int Immunol 21:227–235
9. Enya K, Hayashi H, Takii T, Ohoka N, Kanata S, Okamoto T et al (2008) The interaction with Sp1 and reduction in the activity of histone deacetylase 1 are critical for the constitutive gene expression of IL-1 alpha in human melanoma cells. J Leukoc Biol 83:190–199
10. Lu J, Sun H, Wang X, Liu C, Xu X, Li F et al (2005) Interleukin-12 p40 promoter activity is regulated by the reversible acetylation mediated by HDAC1 and p300. Cytokine 31:46–51
11. Schmeck B, Lorenz J, N'Guessan PD, Opitz B, van Laak V, Zahlten J et al (2008) Histone acetylation and flagellin are essential for Legionella pneumophila-induced cytokine expression. J Immunol 181:940–947
12. Greco L, Romino R, Coto I, Di Cosmo N, Percopo S, Maglio M et al (2002) The first large population based twin study of coeliac disease. Gut 50:624–628
13. Nistico L, Fagnani C, Coto I, Percopo S, Cotichini R, Limongelli MG et al (2006) Concordance, disease progression, and heritability of coeliac disease in Italian twins. Gut 55:803–808
14. Aho K, Koskenvuo M, Tuominen J, Kaprio J (1986) Occurrence of rheumatoid arthritis in a nationwide series of twins. J Rheumatol 13:899–902
15. Silman AJ, MacGregor AJ, Thomson W, Holligan S, Carthy D, Farhan A et al (1993) Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br J Rheumatol 32:903–907
16. MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K et al (2000) Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 43:30–37
17. Lefevre S, Knedla A, Tennie C, Kampmann A, Wunrau C, Dinser R et al (2009) Synovial fibroblasts spread rheumatoid arthritis to unaffected joints. Nat Med 15:1414–1420
18. Guillemin F, Briancon S, Klein JM, Sauleau E, Pourel J (1994) Low incidence of rheumatoid arthritis in France. Scand J Rheumatol 23:264–268
19. Doran MF, Pond GR, Crowson CS, O'Fallon WM, Gabriel SE (2002) Trends in incidence and mortality in rheumatoid arthritis in Rochester, Minnesota, over a forty-year period. Arthritis Rheum 46:625–631
20. Beighton P, Solomon L, Valkenburg HA (1975) Rheumatoid arthritis in a rural South African Negro population. Ann Rheum Dis 34:136–141
21. Silman AJ, Ollier W, Holligan S, Birrell F, Adebajo A, Asuzu MC et al (1993) Absence of rheumatoid arthritis in a rural Nigerian population. J Rheumatol 20:618–622
22. Peschken CA, Esdaile JM (1999) Rheumatic diseases in North America's indigenous peoples. Semin Arthritis Rheum 28:368–391
23. Jacobsson LT, Hanson RL, Knowler WC, Pillemer S, Pettitt DJ, McCance DR et al (1994) Decreasing incidence and prevalence of rheumatoid arthritis in Pima Indians over a twenty-five-year period. Arthritis Rheum 37:1158–1165
24. Kaipiainen-Seppanen O, Aho K, Nikkarinen M (2001) Regional differences in the incidence of rheumatoid arthritis in Finland in 1995. Ann Rheum Dis 60:128–132
25. Shichikawa K, Inoue K, Hirota S, Maeda A, Ota H, Kimura M et al (1999) Changes in the incidence and prevalence of rheumatoid arthritis in Kamitonda, Wakayama, Japan, 1965–1996. Ann Rheum Dis 58:751–756
26. Wiles N, Symmons DP, Harrison B, Barrett E, Barrett JH, Scott DG et al (1999) Estimating the incidence of rheumatoid arthritis: trying to hit a moving target? Arthritis Rheum 42:1339–1346

27. Barrett JH, Brennan P, Fiddler M, Silman A (2000) Breast-feeding and postpartum relapse in women with rheumatoid and inflammatory arthritis. Arthritis Rheum 43:1010–1015
28. Brennan P, Silman A (1994) Breast-feeding and the onset of rheumatoid arthritis. Arthritis Rheum 37:808–813
29. Jorgensen C, Picot MC, Bologna C, Sany J (1996) Oral contraception, parity, breast feeding, and severity of rheumatoid arthritis. Ann Rheum Dis 55:94–98
30. Karlson EW, Mandl LA, Hankinson SE, Grodstein F (2004) Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis? Results from the Nurses' health study. Arthritis Rheum 50:3458–3467
31. Linos A, Worthington JW, O'Fallon WM, Kurland LT (1983) Case-control study of rheumatoid arthritis and prior use of oral contraceptives. Lancet 1:1299–1300
32. Vandenbroucke JP, Witteman JC, Valkenburg HA, Boersma JW, Cats A, Festen JJ et al (1986) Noncontraceptive hormones and rheumatoid arthritis in perimenopausal and postmenopausal women. JAMA 255:1299–1303
33. Turner S, Cherry N (2000) Rheumatoid arthritis in workers exposed to silica in the pottery industry. Occup Environ Med 57:443–447
34. Olsson AR, Skogh T, Axelson O, Wingren G (2004) Occupations and exposures in the work environment as determinants for rheumatoid arthritis. Occup Environ Med 61:233–238
35. Mousavi-Jazi M, Bostrom L, Lovmark C, Linde A, Brytting M, Sundqvist VA (1998) Infrequent detection of cytomegalovirus and Epstein-Barr virus DNA in synovial membrane of patients with rheumatoid arthritis. J Rheumatol 25:623–628
36. Niedobitek G, Lisner R, Swoboda B, Rooney N, Fassbender HG, Kirchner T et al (2000) Lack of evidence for an involvement of Epstein-Barr virus infection of synovial membranes in the pathogenesis of rheumatoid arthritis. Arthritis Rheum 43:151–154
37. Newkirk MM, Watanabe Duffy KN, Leclerc J, Lambert N, Shiroky JB (1994) Detection of cytomegalovirus, Epstein-Barr virus and herpes virus-6 in patients with rheumatoid arthritis with or without Sjogren's syndrome. Br J Rheumatol 33:317–322
38. Kouri T, Petersen J, Rhodes G, Aho K, Palosuo T, Heliovaara M et al (1990) Antibodies to synthetic peptides from Epstein-Barr nuclear antigen-1 in sera of patients with early rheumatoid arthritis and in preillness sera. J Rheumatol 17:1442–1449
39. Fox RI, Luppi M, Pisa P, Kang HI (1992) Potential role of Epstein-Barr virus in Sjogren's syndrome and rheumatoid arthritis. J Rheumatol Suppl 32:18–24
40. Toussirot E, Roudier J (2007) Pathophysiological links between rheumatoid arthritis and the Epstein-Barr virus: an update. Joint Bone Spine 74:418–426
41. Carty SM, Snowden N, Silman AJ (2004) Should infection still be considered as the most likely triggering factor for rheumatoid arthritis? Ann Rheum Dis 63(Suppl 2):ii46–ii49
42. Pattison DJ, Harrison RA, Symmons DP (2004) The role of diet in susceptibility to rheumatoid arthritis: a systematic review. J Rheumatol 31:1310–1319
43. Pattison DJ, Winyard PG (2008) Dietary antioxidants in inflammatory arthritis: do they have any role in etiology or therapy? Nat Clin Pract Rheumatol 4:590–596
44. Cerhan JR, Saag KG, Merlino LA, Mikuls TR, Criswell LA (2003) Antioxidant micronutrients and risk of rheumatoid arthritis in a cohort of older women. Am J Epidemiol 157:345–354
45. Heliovaara M, Knekt P, Aho K, Aaran RK, Alfthan G, Aromaa A (1994) Serum antioxidants and risk of rheumatoid arthritis. Ann Rheum Dis 53:51–53
46. Comstock GW, Burke AE, Hoffman SC, Helzlsouer KJ, Bendich A, Masi AT et al (1997) Serum concentrations of alpha tocopherol, beta carotene, and retinol preceding the diagnosis of rheumatoid arthritis and systemic lupus erythematosus. Ann Rheum Dis 56:323–325
47. Knekt P, Heliovaara M, Aho K, Alfthan G, Marniemi J, Aromaa A (2000) Serum selenium, serum alpha-tocopherol, and the risk of rheumatoid arthritis. Epidemiology 11:402–405
48. Costenbader KH, Kang JH, Karlson EW (2010) Antioxidant intake and risks of rheumatoid arthritis and systemic lupus erythematosus in women. Am J Epidemiol 172:205–216
49. Vessey MP, Villard-Mackintosh L, Yeates D (1987) Oral contraceptives, cigarette smoking and other factors in relation to arthritis. Contraception 35:457–464

50. Hernandez Avila M, Liang MH, Willett WC, Stampfer MJ, Colditz GA, Rosner B et al (1990) Reproductive factors, smoking, and the risk for rheumatoid arthritis. Epidemiology 1:285–291
51. Criswell LA, Merlino LA, Cerhan JR, Mikuls TR, Mudano AS, Burma M et al (2002) Cigarette smoking and the risk of rheumatoid arthritis among postmenopausal women: results from the Iowa Women's health study. Am J Med 112:465–471
52. Stolt P, Bengtsson C, Nordmark B, Lindblad S, Lundberg I, Klareskog L et al (2003) Quantification of the influence of cigarette smoking on rheumatoid arthritis: results from a population based case-control study, using incident cases. Ann Rheum Dis 62:835–841
53. Huizinga TW, Amos CI, van der Helm-van Mil AH, Chen W, van Gaalen FA, Jawaheer D et al (2005) Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum 52:3433–3438
54. Linn-Rasker SP, van der Helm-van Mil AH, van Gaalen FA, Kloppenburg M, de Vries RR, le Cessie S et al (2006) Smoking is a risk factor for anti-CCP antibodies only in rheumatoid arthritis patients who carry HLA-DRB1 shared epitope alleles. Ann Rheum Dis 65:366–371
55. Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J et al (2006) A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum 54:38–46
56. Schulz WA, Steinhoff C, Florl AR (2006) Methylation of endogenous human retroelements in health and disease. Curr Top Microbiol Immunol 310:211–250
57. Nagy G, Ward J, Mosser DD, Koncz A, Gergely P Jr, Stancato C et al (2006) Regulation of CD4 expression via recycling by HRES-1/RAB4 controls susceptibility to HIV infection. J Biol Chem 281:34574–34591
58. Hishikawa T, Ogasawara H, Kaneko H, Shirasawa T, Matsuura Y, Sekigawa I et al (1997) Detection of antibodies to a recombinant gag protein derived from human endogenous retrovirus clone 4–1 in autoimmune diseases. Viral Immunol 10:137–147
59. Piotrowski PC, Duriagin S, Jagodzinski PP (2005) Expression of human endogenous retrovirus clone 4–1 may correlate with blood plasma concentration of anti-U1 RNP and anti-Sm nuclear antibodies. Clin Rheumatol 24:620–624
60. Okada M, Ogasawara H, Kaneko H, Hishikawa T, Sekigawa I, Hashimoto H et al (2002) Role of DNA methylation in transcription of human endogenous retrovirus in the pathogenesis of systemic lupus erythematosus. J Rheumatol 29:1678–1682
61. Neidhart M, Rethage J, Kuchen S, Kunzler P, Crowl RM, Billingham ME et al (2000) Retrotransposable L1 elements expressed in rheumatoid arthritis synovial tissue: association with genomic DNA hypomethylation and influence on gene expression. Arthritis Rheum 43:2634–2647
62. Martin-Trujillo A, van Rietschoten JG, Timmer TC, Rodriguez FM, Huizinga TW, Tak PP et al (2010) Loss of imprinting of IGF2 characterises high IGF2 mRNA-expressing type of fibroblast-like synoviocytes in rheumatoid arthritis. Ann Rheum Dis 69:1239–1242
63. Karouzakis E, Gay RE, Michel BA, Gay S, Neidhart M (2009) DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. Arthritis Rheum 60:3613–3622
64. Richardson B, Scheinbart L, Strahler J, Gross L, Hanash S, Johnson M (1990) Evidence for impaired T cell DNA methylation in systemic lupus erythematosus and rheumatoid arthritis. Arthritis Rheum 33:1665–1673
65. Chen Y, Gorelik GJ, Strickland FM, Richardson BC (2010) Decreased ERK and JNK signaling contribute to gene overexpression in "senescent" CD4+CD28- T cells through epigenetic mechanisms. J Leukoc Biol 87:137–145
66. Nile CJ, Read RC, Akil M, Duff GW, Wilson AG (2008) Methylation status of a single CpG site in the IL6 promoter is related to IL6 messenger RNA levels and rheumatoid arthritis. Arthritis Rheum 58:2686–2693
67. Takami N, Osawa K, Miura Y, Komai K, Taniguchi M, Shiraishi M et al (2006) Hypermethylated promoter region of DR3, the death receptor 3 gene, in rheumatoid arthritis synovial cells. Arthritis Rheum 54:779–787

68. Kulis M, Esteller M (2010) DNA methylation and cancer. Adv Genet 70:27–56
69. Wong AH, Gottesman II, Petronis A (2005) Phenotypic differences in genetically identical organisms: the epigenetic perspective. Hum Mol Genet 14(Spec No 1):R11–R18
70. Enzer I, Dunn G, Jacobsson L, Bennett PH, Knowler WC, Silman A (2002) An epidemiologic study of trends in prevalence of rheumatoid factor seropositivity in Pima Indians: evidence of a decline due to both secular and birth-cohort influences. Arthritis Rheum 46:1729–1734
71. Jacobsson LT, Jacobsson ME, Askling J, Knowler WC (2003) Perinatal characteristics and risk of rheumatoid arthritis. BMJ 326:1068–1069
72. Simard JF, Costenbader KH, Hernan MA, Liang MH, Mittleman MA, Karlson EW (2010) Early life factors and adult-onset rheumatoid arthritis. J Rheumatol 37:32–37
73. Edwards CJ, Goswami R, Goswami P, Syddall H, Dennison EM, Arden NK et al (2006) Growth and infectious exposure during infancy and the risk of rheumatoid factor in adult life. Ann Rheum Dis 65:401–404
74. Mandl LA, Costenbader KH, Simard JF, Karlson EW (2009) Is birthweight associated with risk of rheumatoid arthritis? Data from a large cohort study. Ann Rheum Dis 68:514–518
75. Paun BC, Kukuruga D, Jin Z, Mori Y, Cheng Y, Duncan M et al (2010) Relation between normal rectal methylation, smoking status, and the presence or absence of colorectal adenomas. Cancer 116:4495–4501
76. Strahl BD, Allis CD (2000) The language of covalent histone modifications. Nature 403:41–45
77. Hu N, Qiu X, Luo Y, Yuan J, Li Y, Lei W et al (2008) Abnormal histone modification patterns in lupus CD4+ T cells. J Rheumatol 35:804–810
78. Huber LC, Brock M, Hemmatazad H, Giger OT, Moritz F, Trenkmann M et al (2007) Histone deacetylase/acetylase activity in total synovial tissue derived from rheumatoid arthritis and osteoarthritis patients. Arthritis Rheum 56:1087–1093
79. Dieker J, Muller S (2010) Epigenetic histone code and autoimmunity. Clin Rev Allergy Immunol 39:78–84
80. Plaue S, Muller S, van Regenmortel MH (1989) A branched, synthetic octapeptide of ubiquitinated histone H2A as target of autoantibodies. J Exp Med 169:1607–1617
81. Reumaux D, Meziere C, Colombel JF, Duthilleul P, Mueller S (1995) Distinct production of autoantibodies to nuclear components in ulcerative colitis and in Crohn's disease. Clin Immunol Immunopathol 77:349–357
82. Kanai Y, Kawaminami Y, Miwa M, Matsushima T, Sugimura T (1977) Naturally-occurring antibodies to poly(ADP-ribose) in patients with systemic lupus erythematosus. Nature 265:175–177
83. Myzak MC, Tong P, Dashwood WM, Dashwood RH, Ho E (2007) Sulforaphane retards the growth of human PC-3 xenografts and inhibits HDAC activity in human subjects. Exp Biol Med (Maywood) 232:227–234
84. Dashwood RH, Ho E (2007) Dietary histone deacetylase inhibitors: from cells to mice to man. Semin Cancer Biol 17:363–369
85. Muller-Ladner U, Gay RE, Gay S (2002) Role of nuclear factor kappaB in synovial inflammation. Curr Rheumatol Rep 4:201–207
86. Saccani S, Pantano S, Natoli G (2001) Two waves of nuclear factor kappaB recruitment to target promoters. J Exp Med 193:1351–1359
87. Saccani S, Pantano S, Natoli G (2002) p38-Dependent marking of inflammatory genes for increased NF-kappa B recruitment. Nat Immunol 3:69–75
88. Saccani S, Natoli G (2002) Dynamic changes in histone H3 Lys 9 methylation occurring at tightly regulated inducible inflammatory genes. Genes Dev 16:2219–2224
89. Wutz A, Gribnau J (2007) X inactivation Xplained. Curr Opin Genet Dev 17:387–393
90. Carrel L, Willard HF (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434:400–404
91. Lu Q, Wu A, Tesmer L, Ray D, Yousif N, Richardson B (2007) Demethylation of CD40LG on the inactive X in T cells from women with lupus. J Immunol 179:6352–6358
92. Ozbalkan Z, Bagislar S, Kiraz S, Akyerli CB, Ozer HT, Yavuz S et al (2005) Skewed X chromosome inactivation in blood cells of women with scleroderma. Arthritis Rheum 52:1564–1570

93. Karouzakis El, Rengel Y, Jüngel A, Kolling C, Gay RE, Michel BA, Gay S, Neidhart M, Ospelt C (2011) DNA methylation regulates CXCL12 in rheumatoid arthritis synovial fibroblasts. Genes Immun. Jul 14. doi: 10.1038/gene.2011.45. [Epub ahead of print]
94. Kawabata T, Nishida K, Takasugi K, Ogawa H, Sada K, Kadota Y, Inagaki J, Hirohata S, Ninomiya Y, and Makino H (2010) Increased activity and expression of histone deacetylase 1 in relation to tumor necrosis factor-alpha in synovial tissue of rheumatoid arthritis. Arthritis Res Ther 12(4):R133
95. Gillespie J, Savic S, Wong C, Hempshall A, Inman M, Emery P, Grigg R, McDermott MF (2011) Histone deacetylases are dysregulated in rheumatoid arthritis and a novel HDAC3-selective inhibitor reduces IL-6 production by PBMC of RA patients. Arthritis Rheum. 2011 Sep 27. doi: 10.1002/art.33382. [Epub ahead of print]

Chapter 16
Asthma Epigenetics: Emergence of a New Paradigm?

Rachel L. Miller

Contents

16.1 Introduction 308
16.2 Common Epigenetic Mechanisms 309
16.3 Environmental Exposures Implicated in Both Asthma and Epigenetic Regulation 310
16.4 Environmental Exposures Implicated in Asthma Are Susceptible to Epigenetic Regulation 312
16.5 Epigenetic Regulation and Clinical Allergic Disease 314
16.6 Pharmacoepigenetics 316
16.7 Conclusion 316
References 317

Abstract Asthma is a disease that is influenced by environmental exposures, including those that occur prenatally. Recently, epigenetic regulation has been posited as an explanation for how environmental toxicants may induce asthma-related immune responses. However, our knowledge of the epigenetic regulation of asthma lags substantially behind our understanding of the epigenetic regulation of other complex diseases such as cancer. Fortunately new data are beginning to emerge. These include translational data from molecular experiments that implicate epigenetic regulation in T helper differentiation and/or the development of T regulatory cells, important in allergic immune responses. They also include a growing collection of cohort studies that associate epigenetic regulation with several components of the asthma clinical phenotype. So far these clinical studies are small, often

R.L. Miller (✉)
Division of Pulmonary, Allergy and Critical Care Medicine, Columbia University College of Physicians and Surgeons, New York, NY, USA
e-mail: rlm14@columbia.edu

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_16,
© Springer Science+Business Media B.V. 2012

unconfirmed, and only have started to address key issues relating to tissue specificity. Nonetheless, these studies provide a preview of what future research may reveal and raise the possibility that previously held paradigms for asthma pathogenesis may, and perhaps should, be changing.

Abbreviations

ACSL3	Acyl-CoA synthetase long-chain family member 3
COX	Cyclooxygenase
CCCEH	Columbia Center for Children's Environmental Health
CI	Confidence interval
CpG nucleotides	C–phosphate–G nucleotides
GRs	Glucocorticoid receptors
GSTM1	Glutathione S-transferase mu 1
HDAC	Histone deacetylases
HAT	Histone acetyltransferases
Ig	Immunoglobulin
iNOS	Inducible nitric oxide synthases
IGF	Insulin-like growth factor
IFN	Interferon
IL	Interleukin
LINE-1	Long interspersed nucleotide elements
MS4A2	Membrane-spanning 4-domains, subfamily A, member 2 gene
miRNAs	MicroRNAS
MyD88	Myeloid differentiation primary response gene (88)
nfkb	Nuclear factor kappa B
OR	Odds ratio
PM	Particulate matter
PAHs	Polycyclic aromatic hydrocarbons
RR	Relative risk
Runx3	Runt-related transcription factor 3
Th	T helper
TLR	Toll like receptor
Treg	T regulatory

16.1 Introduction

Our understanding of epigenetic regulation in asthma lags substantially behind our understanding of epigenetic regulation in other complex diseases such as cancer [1, 2]. But experimental work discovering new potential roles for epigenetic regulation in asthma-related immune pathways has grown exponentially in recent years, as

reviewed recently [3–5]. New epidemiological data that link molecular mechanisms or epigenetic biomarkers with asthma-related clinical outcomes are beginning to emerge.

Asthma is a disease that is influenced by environmental exposures, including those that occur prenatally [6, 7]. The longstanding presumption is that the development of asthma is plastic, i.e., not a foregone conclusion based on genetic makeup, and hence modifiable. Asthma also is a disease that exhibits a notoriously variable phenotype [8]. Epigenetic marks that occur prenatally or during other susceptible time periods may modify the clinical manifestations and variable nature of this complex disease. In this chapter, significant advances in the knowledge of epigenetic regulation in response to environmental exposures implicated in asthma will be reviewed. Emerging translational basic science that implicates epigenetic regulation associated with T helper differentiation and/or the development of T regulatory cells that may underlie clinical manifestations of asthma will be mentioned. The last several years have produced many molecular epidemiological reports that associate epigenetic regulation and several components of the clinical asthma phenotype. These studies provide a preview of what future research may reveal and raise the possibility that previously held paradigms for asthma pathogenesis may, and perhaps should, be changing.

16.2 Common Epigenetic Mechanisms

DNA methylation, histone modifications, and production of noncoding RNAs are epigenetic molecular changes that can alter gene transcription without changing the DNA coding sequence. As a result, the host response to environmental exposures, the downstream proinflammatory reaction, or even the therapeutic efficacy of pharmacological agents may be changed if epigenetic modifications ensue. Several animal studies have provided some evidence that the epigenetic state induced by environmental exposures may occur prenatally and influence the phenotype in the offspring [9, 10]. Alternatively, epigenetic modifications may occur postnatally, leading to sustained effects on gene transcription. Another feature of epigenetic regulation is that it may be modifiable [11–13]. There are specific time periods when asthma may be more susceptible to the effects of exposure to environmental toxicants (i.e., gestation, early childhood, adolescence). It has been postulated, but not yet demonstrated, that epigenetic modifications may be contributing to the greater susceptibility that seems to occur during these time periods [3] (Fig. 16.1).

DNA methylation refers to the process whereby methyl groups are added to the fifth carbon of the nucleotide cytosine. This process can suppress gene transcription by either inhibiting transcription factor binding to the recognition sites on CpG nucleotides or by aiding transcription inhibiting protein binding [14]. Importantly, DNA methylation may not be stable and exhibit variation over time [11, 13, 15]. In comparison, post-translational modifications of histones, key elements in the chromatin packaging of DNA, occur by means of acetylation, methylation, and

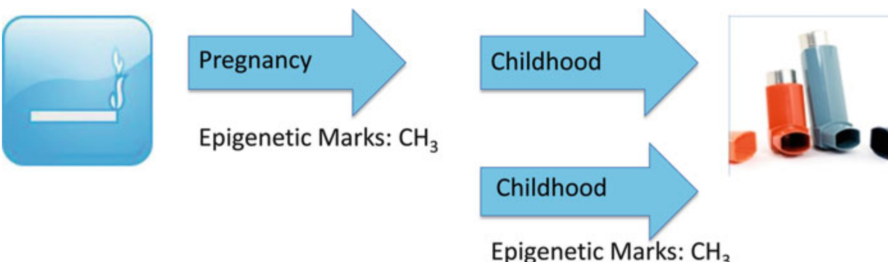


Fig. 16.1 Intrauterine exposures to cigarette smoke, air pollution, diet, and other environmental stressors may induce epigenetic modifications that may affect the risk for developing asthma. Alternately epigenetic changes may occur postnatally and lead to sustained effects on gene expression that as well may affect the risk for asthma development or be associated with its severity

phosphorylation. During acetylation by histone acetyltransferases (HAT), the DNA tightly packaged around the histone core unwinds, activators of transcription obtain access to DNA, and gene expression can then proceed [5]. Histone deacetylases (HDAC), of which there are at least 11 isoenzymes, reverse acetylation and turn off transcription [16]. In general, the extent of histone acetylation is a result of the amount of HAT compared to HDAC activity. The third prominent epigenetic mechanism involves the activity of microRNAs (miRNAs) that function as endogenous inhibitors of translation and thereby regulate protein production. They consist of single-stranded RNA molecules 21–23 nucleotides in length that induce the degradation of target mRNA. The activity of microRNAs' role in asthma or any complex disease is perhaps one of the least understood epigenetic mechanisms. But one exciting study by Mattes and colleagues found that inhaled dust mite allergens increased the expression of specific miRNAs in the airways through TLR4- or MyD88-dependent mechanisms in mouse models. The associated allergic phenotype was inhibited by selective blockade of microRNA (miR)-126 [17].

16.3 Environmental Exposures Implicated in Both Asthma and Epigenetic Regulation

One environmental exposure implicated in both asthma and epigenetic regulation is diet. In the former case, a number of cohort studies have implicated various dietary modifications with greater or reduced risk for asthma, such as supplementation with vitamin D, E, soy, and Mediterranean diet [18–20]. However, the dietary supplement most implicated for inducing epigenetic changes relevant to asthma is folic acid, a source of methyl donors. Hollingsworth and colleagues suggested such an association using elegant mouse models. In their experiments, folic acid supplementation administered during gestation and weaning was associated with greater airway hyperactivity and eosinophilic inflammation, as well as chemokines and immunoglobulin

(Ig) E production in the offspring. They argued that this association occurred due to folic acid's ability to hypermethylate DNA and hence suppress the expression of several genes including *Runx3*, a gene involved in CD4 silencing during T cell lineage decisions [21]. In human cohort work, Whitrow and colleagues assessed the use of supplemental folic acid in women during pregnancy, determined retrospectively by food frequency questionnaire, and found associations with a greater risk of reported physician-diagnosed asthma in children [22]. The strongest association occurred when women took supplemental folic acid later in their pregnancy (30–34 weeks; relative risk (RR) 1.26, confidence interval (CI) 1.08, 1.43), suggesting a possible time window during pregnancy when supplemental folic acid ingestion may induce more extensive epigenetic alterations associated with pediatric asthma. The association between prenatal folic acid supplementation and asthma in children was examined as well in the Norwegian Mother and Child Cohort Study. In Norway, it is recommended the women take supplemental 400 mg of folic acid daily around the time of conception and during the first 3 months of pregnancy. Because the food is not fortified, the authors argued that their assessment of folic acid intake by questionnaire may be relatively accurate. In this cohort, folic acid supplementation only in the first trimester was associated with a slightly greater relative risk of lower respiratory infections (LRTI) (adjusted RR 1.10, 95% CI 1.01–1.20) and hospitalizations for LRTIs (RR 1.28, CI 1.07–1.53) in the child through 18 months of age. Folic acid supplementation anytime during pregnancy was associated with more wheeze at ages 6–18 months (RR 1.07, CI 1.02–1.12). In other models, they adjusted the effects of exposure in the first trimester to exposure both later in pregnancy and in infancy, and the associations with exposure in the first trimester on LRTIs remained significant. In the follow-up nested study within the same cohort, the group reported on maternal folate levels obtained on nonfasting specimens collected during the second trimester of pregnancy (median 18 weeks) and reported that levels were higher among women who reported supplemental folic acid use. An increased risk of asthma at age 3 years for children with maternal plasma folate levels in pregnancy in the highest compared with the lowest quintile, and a trend of increasing risk across quintiles of folate levels also was reported. The authors acknowledged that exposure to folic acid supplements in pregnancy was associated as well with several other characteristics presumed to lower the risk of LTRI, including higher maternal educational level, longer duration of breast feeding, and less smoking, so residual confounding (that for these covariates could lead to a negative bias) could not be ruled out. They also acknowledged that nonfasting plasma samples may be influenced heavily by the most recent intake of folic acid and could have added preanalytical variation [23]. In contrast, Matsui and Matsui determined that folic acid supplementation, indicated by serum levels measured by the 2006–2006 National Health and Nutrition Examination Survey, was associated with a reduced risk for seroatopy and wheeze in individuals age 2 years and older [24]. It is unclear whether the apparent discrepancies among the three cohort studies may be related to exposure misclassification (i.e., from retrospective questionnaires vs measured sera folate levels) or timing of exposure (i.e., prenatal vs postnatal, adult), but some links between folic acid intake and asthma are suggested.

Furthermore, another group determined that periconceptional folic acid administration (400 μg/day) was associated with hypermethylation of the insulin-like growth factor *(IGF)2* gene. In this cohort study, 17-month-old children of mothers who used folic acid had a 4.5% higher relative methylation level of the differentially methylated region of *IGF2* compared to children who were not exposed to folic acid ($p=0.014$). Higher levels of methylation were observed for individual CpG dinucleotides comprising the area, but they were not always statistically significant. *IGF2* methylation of the child also was associated with the maternal (but not child) S-adenosylmethionine blood levels (+1.7% methylation per SD S-adenosylmethionine; $p=0.037$) [25]. Higher methylation levels were associated with decreased birth weight. However, the authors acknowledged that even the statistically significant differences were very small, suggesting that other genes or mechanisms may be involved as well. They also proposed that the extraction of DNA from whole blood composed of multiple cell types may have diluted some of the biological outcomes measured. Still, the paper remains one of the first studies to link periconceptional folic acid supplementation with DNA methylation and a clinically relevant phenotype in the child.

Another environmental exposure long considered a risk factor for the development of asthma in children is maternal smoking [7, 26–28]. One intriguing study by Li and colleagues suggested the mechanism of action on childhood asthma could be transmitted across generations. As part of a case-control study nested within the Children's Health Study, detailed maternal and household smoking histories and histories of other asthma risk factors were obtained by telephone interview. The authors found that a reported history of prenatal smoking was associated with increased risk for asthma diagnosed in the first 5 years of life (odds ratio [OR] 1.5; CI 1.0–2.3), and for persistent asthma (OR 1.5; 95% CI 1.0–2.3). The associations did not differ in children with early transient asthma compared to those with early persistent asthma. Interestingly, grandmaternal smoking during pregnancy was associated with a greater risk of asthma in grandchildren (OR 2.1; 95% CI 1.4–3.2) that remained borderline significant even if there was no report of maternal smoking during pregnancy (OR 1.8; 95% CI 1.0–3.3) [29]. The study was retrospective and did not test epigenetic regulation specifically. Yet this first report of a grandparental effect on asthma may lead to speculation about a contribution of epigenetic regulation to inheritance that would need to be tested rigorously.

16.4 Environmental Exposures Implicated in Asthma Are Susceptible to Epigenetic Regulation

Several environmental exposures implicated in asthma have been shown to induce epigenetic alterations, but perhaps the exposures best characterized in ex vivo or cohort studies are air pollutants. For example, Cao and colleagues documented that diesel exhaust particles induced chromatin modifications in an assay involving exposure of human bronchial epithelial cell lines. Increased histone H4 acetylation

and posttranslation degradation of histone deacetylase (HDAC1) was found. These events in turn were associated with activation of the cyclooxygenase *(COX)-2* promoter [30]. Baccarelli and colleagues recently compared exposure to traffic-related air pollution at multiple time points with repeated analyses for global methylation among a cohort of elderly males as part of the longitudinal Normative Aging Study. Ambient particulate matter $(PM)_{2.5}$, black carbon, and sulfate were measured by stationary site monitoring, and average pollutant concentrations 4 h to 7 days were calculated (moving averages) prior to phlebotomy. DNA methylation in long interspersed nucleotide elements (LINE-1), an indicator of global methylation, decreased in relation to higher black carbon and $PM_{2.5}$ levels, especially during the longer time windows (2–7 days) [13]. Tarantini and colleagues examined two indicators of DNA methylation, Alu and LINE-1, and one measure of asthma candidate gene-specific *(iNOS)* DNA methylation, after short- (after 2 consecutive days off from work in a steel production plant) and long-term (after 3 consecutive days of work) exposure to PM_{10} exposure among electric furnace steel plant workers ($n = 63$). They reported that ambient PM_{10} levels at work sites may be associated with the extent of DNA demethylation of the *iNOS* promoter. Also, PM_{10} levels correlated with global DNA demethylation as estimated in Alu repeated elements and LINE-1 [15].

The relative expression of several candidate miRNAs implicated in oxidative stress and inflammation (miR-222, miR-21, and miR-146a) also were measured in the blood prior to and following 3 days of work among the electric furnace steel plant workers. Levels were compared to individual exposures to fine and course PM and PM metal components (chromium, lead, cadmium, arsenic, nickel, manganese). Among the post-exposure samples, miR-222 expression was correlated positively with lead exposure ($\text{beta} = 0.41$, $p = 0.02$), but miR-21 expression was not associated with individual PM or metals. Some negative correlations between miRNA levels and measures of metals exposure were found [31]. At minimum, the study suggests that changes in miRNA expression may occur, albeit inconsistently, in association with exposure to PM and its metal components, pollutants associated with asthma and other respiratory diseases [32]. Our group at the Columbia Center for Children's Environmental Health (CCCEH) showed that higher levels of polycyclic aromatic hydrocarbons (PAHs), measured prenatally using personal monitors worn by pregnant women, was associated with greater DNA methylation of Acyl-CoA synthetase long-chain family member 3 (*ACSL3*) and other genes (see below) [33]. Finally, Breton and colleagues reported that a history of prenatal tobacco smoke exposure was associated with lower levels of DNA methylation for short interspersed nucleotide element (AluYb8; but not LINE-1) from buccal cells collected from children. Ilumina GoldenGate Bead Array DNA methylation assay of 1,031 gene-specific loci of similar samples yielded 9 candidate genes whose expression differed by prenatal tobacco smoke exposure. Furthermore, when children's data were stratified by the absence or presence of the common GSTM1 null genotype, prenatal tobacco smoke exposure was associated with lower LINE1 methylation in the GSTM1 null children but higher methylation in the GSTM1-present children. This work suggests a novel interaction between genotype and environmental exposure on DNA methylation [34].

16.5 Epigenetic Regulation and Clinical Allergic Disease

Reviewing the epigenetic epidemiology of asthma is plagued by the lack of cohort studies. But significant advances in basic molecular work that provide the biological support for several translational and clinical studies have been made in this area. These include mounting evidence suggesting that immune programming associated with the development of proallergic T helper (Th) 2 cytokine responses (e.g., interleukin (IL)-4) or the counterregulatory Th1 cytokine responses (e.g., interferon (IFN)γ) may be susceptible to epigenetic regulation [35–39]. The most progress to date has been made in determining how DNA methylation of CpG sites within the counterregulatory IFN-γ promoter may confer protection from proallergic immune activation [39]. The IFN γ promoter also has been shown to be susceptible to chromatin remodeling [40]. Susceptibility to proallergic Th2 immune activation in association with demethylation occurs at the IL-4 locus [36, 41] and at the proximal IL-13 promoter [42]. Moreover, epigenetic mechanisms controlling T regulatory (Treg) development are just beginning to be explored. Early results suggest that Treg suppressive function important to allergic sensitization may be dependent on demethylation of *Foxp3* [43, 44].

More clinical studies of epigenetic regulation in clinical asthma now test the etiology of the observed "maternal effect" of atopy. This paradigm refers to the greater predominance of asthma or allergy if the mother, as opposed to the father, is afflicted with the disease. Indeed this pattern has been observed for both asthma and production of proallergic IgE antibodies [5, 8]. To date, one cohort study by Ferreira and colleagues tried to associate such patterns with altered DNA methylation of either AluSp repeat or membrane-spanning 4-domains, subfamily A, member 2 gene (MS4A2) (β-chain of the IgE high-affinity receptor), a key gene in the allergic cascade. Their small study was essentially negative, failing to find differences in AuSp DNA methylation between cases and controls or according to atopic status of the mother or father [45].

As mentioned earlier, our group at CCCEH conducted one of the first proof of concept papers that tried to discover candidate asthma genes that may be susceptible to altered DNA methylation following air pollution exposure. In this case, the approach was to study cord blood DNA derived from high versus low prenatal PAH exposure groups. Genomic DNA was digested with methylation sensitive restriction enzymes, amplified, and aberrantly methylated bands on a gel eluted, reamplified, cloned, and sequenced. Using BLAST search or silicon database analysis of sequenced DNA that focused on promoter and CpG island searches, candidate clones were selected to undergo more intensive evaluations. These validation experiments included bisulfite conversion of DNA and methylation-specific PCR that revealed the 59-CpG island methylation status of *ACSL3* as associated with prenatal PAH exposure. In addition, methylation of *ACSL3* was associated significantly with a parental report of asthma symptoms in children prior to age 5 years [33]. This study is the first to associate a measured prenatal environmental exposure with

altered gene-specific DNA methylation in cord blood and then associate such alterations with the later development of disease. Subsequently, Nadeau and colleagues using a cross-sectional design of age and sex-matched participants, recruited children in Fresno and Stanford California. They compared regional levels of ambient air pollution exposure with peripheral blood DNA methylation of Forkhead box transcription factor 3 (Foxp3), a key transcription factor in Treg suppressive activity and asthma symptom scores. They found that asthmatics living in Fresno where levels of polycyclic aromatic hydrocarbons (PAH), fine particulate matter, and ozone are higher than in Stanford had worse asthma symptoms and higher levels of methylation in the CpG islands in the *Foxp3* promoter and one intronic region compared to nonasthmatics from Fresno and participants from Stanford. An association between the subject-specific estimated annual average PAH exposure in Fresno and the number of methylated CpG islands among Fresno asthmatics was detected as well [46]. While interesting, it should be noted that the study did not adjust the analyses levels for seasonal variation in PAH measures nor control for levels of other pollutants [47].

Other groups have reported provocative experimental results from small human studies, particularly in genes associated with Th polarization associated with allergic airway disease. As an example, ex vivo *IFN-γ* promoter methylation was reduced in CD8+ T cells, but not CD4+ T cells from atopic, but not healthy, children [48]. Kwon and colleagues measured the extent of DNA methylation of *IFNγ* and *IL-4* gene promoters following ex vivo stimulation of peripheral blood mononuclear cells with *Dermatophagoides pteronissinus/Dermatophagoides farinae* dust mite antigens. Among asthmatics, methylation at the *IL-4* promoter (CpG −80) increased following antigen stimulation and was strongly correlated with *IL-4* production. In contrast, demethylation at the *IFNγ* promoter increased following ex vivo stimulation with phytohemagglutinin [49].

Su and colleagues reported on the importance of endogenous HDAC activity in regulating Th1 versus Th2 differentiation. Their approach was to block endogenous HDAC activity with trichostatin A following ex vivo polyclonal stimulation of peripheral blood mononuclear cells derived from children. Not only did trichostatin markedly reduce endogenous HDAC activity, but this result was associated with greater Th2 polarization and increased expression of GATA-3, suggesting that endogenous HDAC activity is necessary for preserving Th1 versus Th2 balance. Finally, the same group compared levels of HAT activity and endogenous HDAC in atopic nonasthmatic ($n=27$) versus atopic asthmatic ($n=18$) children with the extent of bronchial hyperresponsiveness, an indicator of asthma severity. They found significant elevations of HAT activity and reductions in the level of HDAC activity (determined in nuclear peripheral blood mononuclear cell lysates) among asthmatic children compared to atopic nonasthmatic controls. Moreover, these changes were associated progressively with the degree of bronchial hyperreactivity becoming the first study to compare the extent of epigenetic alteration with clinically relevant physiological outcomes [50].

16.6 Pharmacoepigenetics

Given that epigenetic alterations may be reversible, at least in theory, it stands to reason that there may be substantial potential to develop or take advantage of epigenetically-based therapies in asthma. DNA methyltransferases and histone deacetylases have shown promising anti-tumorigenic effects for some malignancies (reviewed in [51, 52]). Inhibitors of enzymes controlling epigenetic modifications associated with inflammation or other immune processes that result in airway disease, specifically DNA methyltransferases, histone deacetylases, and inhibitory RNAs, have therapeutic potential.

One way to consider the therapeutic potential of epigenetic regulation is to consider mechanisms that underlie the efficacy of glucocorticoids. Steroids exert some of their anti-inflammatory properties by inducing the acetylation of anti-inflammatory genes important for the production of cytokines, chemokines, or adhesion molecules. Specifically, glucocorticoid receptors (GRs) bind to their DNA binding site following acetylation [53]. Under other circumstances, such as during suppression of Nfkβ activation, steroids may recruit histone deacetylase-2 (HDAC-2) to activated inflammatory gene complexes [53]. Indeed, alveolar macrophages have been shown to express elevated levels of HAT activity and reduced levels of HDAC activity [54, 55]. One could speculate that therapeutics that modify the relative HAT, HDAC levels or target miRNA expression levels in the airways or their alveolar macrophages may hold promise as novel anti-inflammatory treatments for allergic asthma [17].

16.7 Conclusion

Recent advances have been made in the last few years, particularly in the molecular epidemiological work that associates epigenetic regulation with asthma phenotypes. However, there are several areas where this field has barely begun to tackle important questions. These include translational studies that examine the differential effects of epigenetic regulation according to specific cell types. For example, HDAC activity sampled from human alveolar macrophages was lower than levels sampled from those measured simultaneously in peripheral blood mononuclear cells [54]. Promoter methylation of leukotriene B4 receptor and its associated gene expression also has been shown to vary according to cell type [56]. In addition, the cohort studies conducted to date have been small and possess limited statistical power to adjust for covariates or to measure small effect sizes associated with epigenetic changes. Large prospective cohort studies are needed to examine all of these issues and how they may apply to the complex disease of asthma. Despite these limitations, there is exciting promise that the "bench" will communicate with the "bedside" and direct appropriate translational studies in the molecular epidemiology of asthma epigenetics.

References

1. Shames DS, Girard L, Gao B, Sato M, Lewis CM, Shivapurkar N et al (2006) A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies. PLoS Med 3:e486
2. Chan AO, Rashid A (2006) CpG island methylation in precursors of gastrointestinal malignancies. Curr Mol Med 6:401–408
3. Miller RL, Ho S (2008) Environmental epigenetics and asthma: current concepts and call for studies. Am J Respir Crit Care Med 177:567–573
4. Adcock IM, Tsaprouni L, Bhavsar P, Ito K (2007) Epigenetic regulation of airway inflammation. Curr Opin Immunol 19:694–700
5. Miller RL, Herbstman J (2009) Epigenetic mechanisms in asthma. In: Epigenetics nutrition and environmental health. Wiley, Berlin, pp 253–263
6. Miller RL, Garfinkel R, Horton M, Camann D, Perera F, Whyatt RM et al (2004) Polycyclic aromatic hydrocarbons, environmental tobacco smoke, and respiratory symptoms in an inner-city birth cohort. Chest 126:1071–1078
7. Moshammer H, Hoek G, Luttmann-Gibson H, Neuberger MA, Antova T, Gehring U et al (2006) Parental smoking and lung function in children: an international study. Am J Respir Crit Care Med 173:1255
8. Mandhane PJ, Greene JM, Cowan JO, Taylor DR, Sears MR (2005) Sex differences in factors associated with childhood- and adolescent-onset wheeze. Am J Respir Crit Care Med 172:45–54
9. Rakyan VK, Chong S, Champ ME, Cuthbert PC, Morgan HD, Luu KVK et al (2003) Transgenerational inheritance of epigenetic states at the murine Axin (Fu) allele occurs after maternal and paternal transmission. Proc Natl Acad Sci USA 100:2538–2543
10. Anway MD, Cupp AS, Uzumcu M, Skinner MK (2005) Epigenetic transgenerational actions of endocrine disruptors and male fertility. Science 308:1466–1469
11. Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H et al (2008) Intra-individual change over time in DNA methylation with familial clustering. JAMA 299:2877–2883
12. Weaver I, Meaney M, Szyf M (2006) Maternal care effects on the hippocampal transcriptome and anxiety-mediated behaviors in the offspring that are reversible in adulthood. Proc Natl Acad Sci USA 103:3480–3485
13. Baccarelli A, Wright RO, Bollati V, Tarantini L, Litonjua AA, Suh HH et al (2009) Rapid DNA methylation changes after exposure to traffic particles. Am J Respir Crit Care Med 179:572–578
14. Isidoro-Garcia M, Davila-Gonzalez I, Pascual de Pedro M, Sanza-Lozano C, Lorente-Toledano F (2007) Interactions between genes and the environment. Epigenetics in allergy. Allergol Immunopathol 35:254–258
15. Tarantini L, Bonzini M, Apostoli P, Pegoraro V, Bollati V, Marinelli B et al (2009) Effects of particulate matter on genomic DNA methylation content and iNOS promoter methylation. Environ Health Perspect 117:217–222
16. Johnson CA, Turner BM (1999) Histone deacetylases: complex transducers of nuclear signals. Semin Cell Dev Biol 10:179–188
17. Mattes J, Collisona A, Planka M, Phipps S, Foster PS (2009) Antagonism of microRNA-126 suppresses the effector function of TH2 cells and the development of allergic airways disease. Proc Natl Acad Sci USA 106:18704–18709
18. Devereux G, Turner SW, Craig LC, McNeill G, Martindale S, Harbour PJ et al (2006) Low maternal vitamin E intake during pregnancy is associated with asthma in 5-year-old children. Am J Respir Crit Care Med 174:499–507
19. Devereux G, Litonjua A, Turner SW, Craig LC, McNeill G, Martindale S et al (2007) Maternal vitamin D intake during pregnancy and early childhood wheezing. Am J Clin Nutr 85:853–859

20. Castro-Rodriguez JA, Garcia-Marcos L, Sanchez-Solis M, Pérez-Fernández V, Martinez-Torres A, Mallol J (2010) Olive oil during pregnancy is associated with reduced wheezing during the first year of life of the offspring. Pediatr Pulmonol 45:395–402
21. Hollingsworth JW, 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 118:3462–3469
22. Whitrow M, Moore V, Rumbold A, Davies M (2009) Effect of supplemental folic acid in pregnancy on childhood asthma: a prospective birth cohort study. Am J Epidemiol 170:1486–1493
23. Haberg SE, London SJ, Nafstad P, Nilsen RM, Ueland PM, Vollset SE et al (2011) Maternal folate levels in pregnancy and asthma in children at age 3 years. J Allergy Clin Immunol 127:262–264, 264 e261
24. Matsui EC, Matsui W (2009) Higher serum folate levels are associated with a lower risk of atopy and wheeze. J Allergy Clin Immunol 123:1253–1259
25. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, Lindemans J, Siebel C, Steegers EA et al (2009) Periconceptional maternal folic acid use of 400 mg per day is related to increased methylation of the IGF2gene in the very young child. PLoS One 4:e7845
26. Stevenson LA, Gergen PJ, Hoover DR, Rosenstreich D, Mannino DM, Matte TD (2001) Sociodemographic correlates of indoor allergen sensitivity among United States children. J Allergy Clin Immunol 108:747–752
27. Alati R, Al-Mamun A, O'Callaghan M, Najman JM, Williams GM (2006) In utero and postnatal maternal smoking and asthma in adolescence. Epidemiology 17:138–144
28. Eisner MD (2002) Environmental tobacco smoke exposure and pulmonary function among adults in NHANES III: impact on the general population and adults with current asthma. Environ Health Perspect 110:765–770
29. Li Y-F, Langholz B, Salam MT, Gilliland FD (2005) Maternal and grandmaternal smoking patterns are associated with early childhood asthma. Chest 127:1232–1241
30. Cao D, Bromberg PA, Samet JM (2007) COX-2 expression induced by diesel particles involves chromatin modification and degradation of HDAC1. Am J Respir Cell Mol Biol 37:232–239
31. Bollati V, Marinelli B, Apostoli P, Bonzini M, Nordio F, Hoxha M et al (2010) Exposure to metal-rich particulate matter modifies the expression of candidate microRNAs in peripheral blood leukocytes. Environ Health Perspect 118:763–768
32. Patel MM, Hoepner L, Garfinkel R, Chillrud S, Reyes A, Quinn JW et al (2009) Ambient metals, elemental carbon, and wheeze and cough in New York city children through 24 months of age. Am J Respir Crit Care Med 180:1107–1113
33. Perera F, Tang WY, Herbstman J, Tang D, Levin L, Miller R et al (2009) Relation of DNA methylation of 5′-CpG island of ACSL3 to transplacental exposure to airborne polycyclic aromatic hydrocarbons and childhood asthma. PLoS One 4:e4488
34. Breton CV, Byun H-M, Wenten M, Pan F, Yang A, Gilliland FD (2009) Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. Am J Respir Crit Care Med 180:462–467
35. Agarwal S, Rao A (1998) Modulation of chromatin structure regulates cytokine gene expression during T cell differentiation. Immunity 9:765–775
36. Lee DU, Agarwal S, Rao A (2002) Th2 lineage commitment and efficient IL-4 production involves extended demethylation of the IL-4 gene. Immunity 16:649–660
37. Yano S, Ghosh P, Kusaba H, Buchholz M, Longo DL (2003) Effect of promoter methylation on the regulation of IFNg gene during in vitro differentiation of human peripheral blood T cells into a Th2 population. J Immunol 171:2510–2516
38. Tykocinski LO, Hajkova P, Chang HD, Stamm T, Sozeri O, Lohning M et al (2005) A critical control element for interleukin-4 memory expression in T helper lymphocytes. J Biol Chem 280:28177–28185
39. Jones B, Chen J (2006) Inhibition of IFN-g transcription by site-specific methylation during T helper cell development. EMBO J 25:2443–2452

40. Zhang F, Boothby M (2006) T helper type 1 specific Brb1 recruitment and remodeling of nucleosomes positioned at the IFN-gamma promoter are Stat4 dependent. J Exp Med 203:1493–1505
41. Janson P, Winerdal M, Winqvist O (2009) At the crossroads of T helper lineage commitment-epigenetics points the way. Biochim Biophys Acta 1790:906–919
42. Webster RB, Rodriguez Y, Klimecki W, Vercelli D (2007) The human IL-13 locus in neonatal CD4+ T cells is refractory to the acquisition of a repressive chromatin architecture. J Biol Chem 282:700–709
43. Kim HP, Leonar WJ (2007) CREB/ATF-dependent T cell receptor-induced FoxP3 gene expression: a role for DNA methylation. J Exp Med 204:1543–1551
44. Huehn J, Polansky JK, Hamann A (2009) Epigenetic control of FOXP3 expression: the key to a stable regulatory T-cell lineage? Nat Rev Immunol 9:83–89
45. Ferreira M, Oates N, Vliet Jv, Zhao ZZ, Ehrich M, Martin NG et al (2010) Characterization of the methylation patterns of MS4A2 in atopic cases and controls. Allergy 65:333–337
46. Nadeau K, McDonald-Hyman C, Noth E, Pratt B, Hammond S, Balmes J et al (2010) Ambient air pollution impairs regulatory T-cell function in asthma. J Allergy Clin Immunol 126:845–852
47. Jung KH, Patel MM, Moors K, Kinney PL, Chillrud SN, Whyatt R et al (2010) Effects of heating season on residential indoor and outdoor polycyclic aromatic hydrocarbons, black carbon, particulate matter in an urban birth cohort. Atmos Environ 44:4545–4552
48. White GP, Hollams EM, Yerkovich ST, Bosco A, Holt BJ, Bassami MR et al (2006) CpG methylation patterns in IFNγ promoter in naive T cells: variations during Th1 and Th2 differentiation and between atopics and non-atopics. Pediatr Allergy Immunol 17:557–564
49. Kwon N-H, Kim J-S, Lee J-Y, Oh M-J, Choi D-C (2008) DNA methylation and the expression of IL-4 and IFN-γ promoter genes in patients with bronchial asthma. J Clin Immunol 28:139–146
50. Su RC, Becker AB, Kozyrsyi AL, HayGlass KT (2009) Altered epigenetic regulation and increasing severity of bronchial hyperresponsiveness in atopic asthmatic children. J Allergy Clin Immunol 124:1116–1117
51. Reamon-Buettner SM, Borlak J (2007) A new paradigm in toxicology and teratology: altering gene activity in the absence of DNA sequence variation. Reprod Toxicol 24:20–30
52. Dekker FJ, Haisma HJ (2009) Histone acetyl transferases as emerging drug targets. Drug Discov Today 14:942–948
53. Barnes PJ (2009) Targeting the epigenome in the treatment of asthma and chronic obstructive pulmonary disease. Proc Am Thorac Soc 6:693–696
54. Cosío BG, Mann B, Ito K, Jazrawi E, Barnes PJ, Chung KF et al (2004) Histone acetylase and deacetylase activity in alveolar macrophages and blood monocytes in asthma. Am J Respir Crit Care Med 170:141–147
55. Ito K, Caramori G, Lim S, Oates T, Chung K, Barnes P et al (2002) Expression and activity of histone deacetylases in human asthmatic airway. Am J Respir Crit Care Med 166:392–396
56. Kato K, Yokomizo T, Izumi T, Shimizu T (2000) Cell specific transcriptional regulation of human leukotriene B4 receptor gene. J Exp Med 192:421–431

Chapter 17
Epigenetic Epidemiology of Autism and Other Neurodevelopmental Disorders

Rima Woods and Janine M. LaSalle

Contents

17.1 The Importance of Epigenetic Modifications in the Brain 322
17.2 Neurodevelopmental Disorders with Known Genetic Causes That Affect Epigenetic Pathways 323
17.2.1 Disorders Associated with Chromosome 15q11-13 Deletion and Duplication 323
17.2.2 Fragile-X Disorders 325
17.2.3 ICF Syndrome 327
17.2.4 X-Linked IDD and α-Thalassemia Mental Retardation, X-Linked 328
17.2.5 Rett Syndrome and the Spectrum of MECP2 Disorders 329
17.3 Autism Is a Spectrum of Neurodevelopmental Disorders with Complex Genetic and Environmental Contributors Affecting Epigenetic Pathways 331
17.4 Conclusions 334
References 335

Abstract Mammalian brain development continues postnatally, as neuronal circuits and pathways undergo maturational development in response to experiences in early life. Therefore, epigenetic mechanisms that act at the interface of genetic and environmental information are increasingly recognized for their essential roles in postnatal brain development, and they are thought to play an important role in learning and memory. Here we review some of the emerging examples of human genetic neurodevelopmental disorders with known epigenetic etiologies and use these examples to speculate on the potential epigenetic etiology of the more common

R. Woods (✉) • J.M. LaSalle
Medical Microbiology and Immunology, Genome Center, and Medical Institute of Neurodevelopmental Disorders, University of California, Davis, CA 95616, USA
e-mail: rwoods@ucdavis.edu; jmlasalle@ucdavis.edu

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_17,

© Springer Science+Business Media B.V. 2012

forms of idiopathic autisms. Understanding the epigenetic and biochemical pathways disrupted by genetic disorders can help direct the field of epigenetic epidemiology in its search for evidence of environmentally altered pathways, including investigations of how environmental factors such as nutrition and toxin exposures may contribute to epigenetic information essential for the higher cognitive functions associated with language and social behavior.

17.1 The Importance of Epigenetic Modifications in the Brain

The epigenomic landscape of the nervous system is dynamic and complex, and equally complex regulatory mechanisms direct the response of the neuronal system to environmental signals. Epigenetic mechanisms, including DNA methylation and histone modifications, are important for proper neuronal function through their regulation of chromatin structure, imprinting, and X-inactivation. The epigenome allows for varying degrees of regulation of the inherited genome, depending on the stimulus presented and the reaction required.

DNA methylation of CpG dinucleotides is orchestrated by DNA methyltransferases which "write" the methylome and methyl binding proteins which "read" the methylome. Despite being mitotically inactive, neurons have high levels of both *de novo* and maintenance DNA methyltransferases [1–4]. This suggests that neurons have a dynamic methylome that is capable of changing in response to stimuli. Both genetic and chemical methods of inhibiting DNA methylation significantly reduces learning and memory in rodents [5, 6], and hypo- and hypermethylation have been linked to various neurological disorders [7]. DNA methylation also plays a key role in imprinting and X-inactivation. Imprinted genes are silenced via DNA methylation based on the parent of origin in an allele specific manner. X-inactivation uses DNA methylation and chromatin modifications to randomly silence one of the two X chromosomes in female cells. Both imprinting and X-inactivation act to maintain proper gene dosage, and errors in these mechanisms can lead to neurological disorders. Maintaining the dynamics of DNA methylation is therefore crucial to proper postnatal neurological function.

The core histone proteins that package DNA into chromatin contain a number of post-translational modifications, which are collectively called the "histone code." The addition and removal of modifying marks on histone tails serves to alter the chromatin landscape and can globally and locally affect the expression of genes. For example, acetylation of lysine 9 on histone H3, a modification associated with gene activation, has been shown to change over time as neurons mature [8]. Histone deacetylases (HDACs) are enzymes that remove acetyl groups from histones and generally serve to inhibit transcription. The clinical use of HDAC inhibitors such as valproic acid (VPA) to treat neurological symptoms in schizophrenia, epilepsy, and mood disorders [9] highlights the importance of histone modifications in the brain.

Epigenetics is important for neuronal development and plasticity, which is achieved through regulation of the chromatin landscape. The level of chromatin

compaction greatly affects accessibility of DNA to various DNA-binding proteins, transcription factors, and cellular machinery. In normal cells, this serves to tightly regulate gene transcription and potentially serve as a molecular basis for memory. In cells with abnormal epigenetic marks, the effects on neurodevelopment can be profound. For instance, the combined contribution of both maternal and paternal epigentically marked genomes is required for normal neurodevelopment [10]. More subtle dysregulation of epigenetic mechanisms have been implicated in a variety of human neurodevelopmental disorders, which are discussed in greater detail in this review.

17.2 Neurodevelopmental Disorders with Known Genetic Causes That Affect Epigenetic Pathways

17.2.1 Disorders Associated with Chromosome 15q11-13 Deletion and Duplication

Among the first characterized "epigenetic" human diseases are two neurodevelopmental disorders caused by deletions of 15q11-13, Prader-Willi and Angelman syndromes (PWS and AS, respectively). Genetic rearrangements of human chromosomal locus 15q11-13 are common in a wide range of neurodevelopmental disorders, as this locus is a hotspot for low copy repeats in humans predisposing to deletions and duplications, as well as parentally imprinted genes [11–13]. The methylation status of a small imprinting control region (ICR), and its presence or absence, is central to the imprinting mechanisms in both humans and mouse models. Human PWS infants are recognized by hypotonia and failure to thrive in infancy, small gonads, small hands and feet, obsessive compulsive and hoarding behaviors, and an insatiable appetite leading to obesity. AS individuals are also hypotonic at birth, but develop distinct features from PWS, including a happy disposition, ataxic gait, absence of speech, and severe seizures.

Prader-Willi syndrome occurs as a result of loss of an unmethylated ICR from the paternal chromosome. Loss of the ICR is caused most frequently by paternal deletion of 15q11-13 (70%) or maternal uniparental disomy (UPD, 25–30%), in which both chromosome 15 s are of maternal origin. The remaining 3–5% of PWS cases are caused by small paternal deletions of the imprinting control region (PWS-IC). Angelman syndrome is caused by the reciprocal chromosomal imbalances as PWS, which are maternal 15q11-13 deletion (70%) or paternal uniparental disomy (<5%). The remaining cases of AS are caused by maternal mutation of the gene *UBE3A*, encoding a ubiquitin ligase E3, or imprinting mutations, which lack the maternal methylation imprint at the PWS-IC, resulting in reduced *UBE3A* expression. Mouse models of AS have implicated a role for *Ube3a* in experience dependent plasticity of the cortex and the maturation of synapses by ubiquitination of Arc [14–16].

Uniparental disomy in PWS and AS is predicted to derive from rare rescue events of trisomy for chromosome 15, and is thus influenced by increasing parental age, particularly maternal [17, 18]. Parental imprinting of the PWS/AS locus is established in the gametes and regulated through a bipartite imprinting control region (ICR), defined from rare small noncoding deletions in PWS and AS patients specifically affecting parental DNA methylation patterns [19, 20]. The maternally inherited AS-ICR is required during oocyte maturation in order to methylate the maternal PWS-ICR [21]. The maternal specific methylation of the PWS-ICR is diagnostic for the majority of PWS and AS patients and is stably maintained in clonal cell populations [22]. Maternal methylation at the PWS-ICR also serves to silence the maternal allele of an exceptionally long and complex transcriptional unit, starting at the PWS-ICR and containing *SNRPN* (encoding a splicing factor), two small nucleolar RNAs (snoRNAs) clusters, a spliced long noncoding RNA, and the antisense to the maternally expressed gene *UBE3A* [23]. The paternal allele *SNRPN* through *UBE3A* locus undergoes extensive chromatin decondensation in postnatal neurons that requires PWS-ICR transcription [24], suggesting that the local highly active chromatin environment of the paternal locus may be a critical feature lacking in PWS.

In addition to the 15q11-13 deletion syndromes, duplications of 15q11-13 are found in approximately 1–3% of children with autism-spectrum disorders, although the actual prevalence of 15q11-13 duplications in the general population has not been established. The phenotype of 15q duplication syndrome is heterogeneous, but it is becoming a clinically recognizable neurodevelopmental disorder characterized by hypotonia in infancy, followed by language and motor delays and frequently, severe seizures. Anxiety, hyperactivity, emotional lability, and sleep problems are also characteristic of 15q duplication syndrome in addition to the social behavioral deficits that enable a diagnosis of autism prior to genetic karyotype analysis (reviewed in [25]).

The chromosomal rearrangements leading to chromosome 15q duplication are due to meiosis U-type rearrangements between low copy repeat clusters in humans, and thus are more frequent in maternal meiosis and impacted by maternal age [18, 26]. 15q11-13 duplications leading to autism are almost exclusively maternal in origin, emphasizing the importance of parental imprinting in all of the 15q [27]. Paternal origin of 15q duplications have been described, but the phenotypes vary widely, including social and language delay, severe cognitive impairment, PWS-like behaviors, and an apparently normal phenotype [27–29]. Preliminary analyses of postmortem brain of 15q duplication samples suggest that transcript levels do not correlate to copy number and parental origin, suggesting that epigenetic dysregulation of the region has occurred [13].

Maturing postnatal neurons appear particularly sensitive to genetic and epigenetic alterations to both coding and noncoding regions of 15q11-13. The absence or duplication of a methylated or unmethylated ICR in 15q11-13 determines the phenotypic outcome, providing an example of loss of epigenetic control of a chromosomal region. Because of the occurrence of 15q11-13 rearrangements in autism spectrum disorders (~1%), further understanding of the epigenetic complexities as well as the functions of the critical imprinted genes is essential to future diagnosis and treatment of autism.

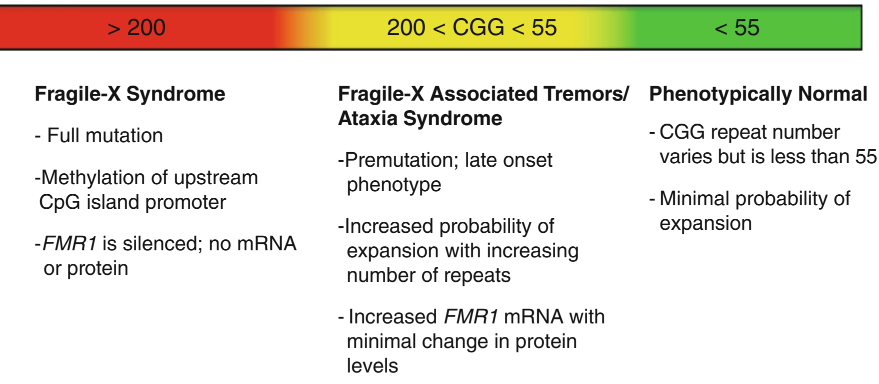


Fig. 17.1 Genotype-phenotype correlation of FMR1 CGG repeat number. The 5′ *FMR1* CGG polymorphic repeat is polymorphic in humans. Individuals with <55 repeats (*green*) are in the phenotypically normal range, those with intermediate expansions (*yellow*) carry a permutation and can develop Fragile-X Associated Tremors/Ataxia syndrome, while those with >200 repeats (*red*) generally silence *FMR1* and cause Fragile X syndrome

17.2.2 Fragile-X Disorders

The epigenetic mechanisms responsible for Fragile-X Syndrome (FXS) were discovered while the field of epigenetics was in its infancy. Fragile X-Associated Tremor Ataxia Syndrome (FXTAS), a closely related disorder involving the same genomic region but with a very different etiology, was discovered a decade later. The relation between FXS and FXTAS clearly shows the variable effects of epigenetic dysregulation in relation to neurological disorders. Aberrant epigenetic silencing through DNA methylation of a promoter causes FXS, whereas FXTAS results from uncontrolled gene expression. Both diseases stem from a genetic expansion of CGG repeats in the same X-linked gene, but their pathogenesis differs in the dysregulation of epigenetics that results from the expansion (Fig. 17.1).

Fragile X Syndrome is caused by aberrant silencing of *Fragile X Mental Retardation 1* (*FMR1*) and the loss of the *FMR1* product, Fragile X Mental Retardation Protein (FMRP). FMRP is an RNA binding protein and is involved in mRNA trafficking during synaptic activity [30, 31] and is important for neurodevelopment and neuronal function. Individuals with FXS often present with intellectual disabilities, autism spectrum disorder, ADHD, anxiety, and seizures [32, 33]. FXS occurs in approximately 1/2500 births [34], and is more often observed in males. *FMR1* is located on Xq27 with a CpG island at the promoter region that is unmethylated in normal individuals [35]. The 5′ UTR of *FMR1* contains a tandem array of up to 55 CGG repeats. The number of CGG repeats is variable in normal individuals, with the most common allele containing 29 repeats [36]. In FXS, the CGG repeat region is greatly expanded to over 200 copies (termed ‘full mutation’) during oogenesis, due to meiotic instability. This leads to methylation of the upstream CpG island, and silencing of *FMR1* [36–39]. Increased levels of repressive histone marks,

including H3K9me3 and H4K20me3, within the repeat region [40] may add an additional layer of silencing by altering the chromatin structure around *FMR1*, further silencing the expanded allele. Individual phenotypic severity varies greatly, as mosaicism and variation between individuals in repeat length and CpG island methylation has been observed. Repeat length mosaicism is due to mitotic instability at the repeat region, and differences in methylation are likely due to a combination of variables including environmental and stochastic factors [33, 37].

FXTAS, a late onset neurodegenerative disorder, is also due to expansions in the CGG repeat region in the 5′ UTR of *FMR1*. FXTAS is caused by a smaller expansion of between 55 and 200 CGGs (termed 'premutation') and is characterized by progressive cognitive decline and memory loss, action tremors, parkinsonism, and generalized brain atrophy [41]. Like FXS, the repeat expansion occurs in the female germline as a result of meiotic instability at the CGG repeat region. Often, a grandparent (who will go on to develop FXTAS later in life) will pass on a premutation to their daughter where it expands during oogenesis to a full mutation, resulting in a grandchild with FXS [39, 42]. Larger premutations result in an increased likelihood of expansion to a full mutation, with 100% likelihood of expansion to full mutation with a premutation of more than 90 CGG repeats [43]. Since FXTAS is an X-linked disorder, females can be less severely affected due to mosaicism or favorable skewing of X-inactivation [44]. It is estimated that 1/3000 males over the age of 50 will develop FXTAS [34].

A hallmark of FXTAS is elevated levels of *FMR1* mRNA in patients carrying premutations but near normal levels of FMRP. The increase in expression correlates with increased number of repeats, with a five-fold increase in expression in patients carrying more than 100 repeats [45]. An RNA toxicity model has emerged for FXTAS, where an increased number of CGG repeats leads to increased *FMR1* expression, possibly due to decreased translational efficiency, which in turn leads to the formation of *FMR1*-containing nuclear inclusions in brain and spinal cord tissues, and the neurological problems associated with FXTAS [46, 47]. A study of 93 male carriers of premutations found a significant correlation between the number of CGG repeats and the age of symptom onset [48]. A CGG knock-in mouse model carrying a human CGG repeat containing 98 CGGs showed inclusion size in neurons increased with age [49], suggesting the late onset of symptoms is due to the gradual accumulation of nuclear inclusions.

The Fragile-X related disorders provide a unique example of epigenetically regulated transgenerationally inherited neurological disorders. While the etiology is based on a genetic mutation (CGG repeat expansion), epigenetic changes to the X-linked gene *FMR1* are responsible for FXS. Aberrant silencing of *FMR1* via promoter methylation causes FXS. FXTAS results from abnormal increased expression of *FMR1* and serves as an intermediary between a normal phonotype and FXS. A premutation resulting in FXTAS often leads to a full mutation and FXS in as few as two generations. Although the mechanism responsible for the CGG repeat expansion is unknown, the vastly different effects of dysregulated epigenetic processes are evident in the two diseases. The Fragile-X related disorders highlight the importance of epigenetic control of gene dosage for proper neurological function.

17.2.3 ICF Syndrome

Immunodeficiency, centromeric region instability, and facial anomalies syndrome (ICF) is a rare autosomal recessive disorder linked to loss of DNA methylation. This decrease in global methylation is thought to lead to loss of heterochromatin and genomic instability at centromeric and telomeric regions. Fewer than 50 patients have been described worldwide. This rare syndrome is divided into two types. Approximately half of all identified ICF cases have mutations in *DNMT3B*, a *de novo* methyltransferase, which are responsible for ICF syndrome Type 1 and are thought to cause the global hypomethylation observed [50]. ICF syndrome Type 2 also shows decreased global methylation but lack DNMT3B mutations. Mutations in a gene coding for the zinc finger protein, *ZBTB24* are associated with ICF syndrome Type 2 [148]. Both types are characterized by facial dysmorphisms, including hypertelorism, flat nasal bridge, and low set ears, immune system deficiencies including decreased immunoglobulin in the presence of normal B and T cell counts, and cytogenetic abnormalities occurring at the centromeric regions of chromosomes 1, 9, and 16. Additionally, delayed development has been observed in 68% of cases, and mild to moderate intellectual developmental disability (IDD) has been observed in 61% of cases. The median age at death is 8 years old, with death usually occurring due to severe and/or recurrent infections [50, 51].

Both ICF Type 1 and 2 patients exhibit hypomethylation of pericentromeric satellite 2 and 3 repeats, telomeric, and subtelomeric regions resulting in decondensation of pericentromeric heterochromatin and telomeric shortening. ICF Type 2 also displays hypomethylation of alpha satellite DNA, indicating additional unknown factors are involved in ICF syndrome that affect methylation pathways. The hypomethylation characteristic of ICF syndrome leads to only a 7% reduction in global methylation, indicating that relatively small changes to global methylation levels can have a large impact on neurodevelopment.

Decondensation of pericentromeric heterochromatin has been observed in both metaphase and interphase chromosomes in ICF cells [52–54], and may be due to hypomethylation. These large-scale changes to chromatin structure likely affect transcription of a wide variety of genes as it is becoming increasingly clear that the spatial arrangement of the chromosomes within the nucleus plays a role in transcriptional regulation. Microarray expression analyses comparing ICF and control lymphoblast cells found both up and down regulation of various genes involved in immune function, development, neurogenesis, and apoptosis [51, 55]. A subset of these genes were examined for altered methylation at regulatory regions, in an attempt to explain the expression changes but no direct correlation was found. It therefore remains unclear if altered methylation patterns at regulatory regions of genes with altered expression profiles play a causative role in ICF syndrome and if the hypomethylation is responsible for the transcriptional changes. It is interesting to note that decreased H3K27 trimethylation, a repressive histone mark, was detected at the promoters of genes shown to be upregulated in ICF syndrome [55]. It seems likely that a combination of low-level changes in methylation along with altered histone marks at the regulatory regions of dysregulated genes contributes to the etiology of ICF syndrome.

In addition to the effects seen at the centromeric and pericentromeric regions, telomere shortening and hypomethylation at telomeric regions has been observed in ICF cells, despite levels of telomerase activity comparable to that found in control cells. The hypomethylation of the telomeric regions is accompanied by increased transcription of telomere RNA (telRNA or TERRA) [56]. These transcripts localize to the telomeric regions and inhibit activity of telomerase [57]. Decreased methylation at the telomeric regions may lead to increased expression of telRNA in ICF cells thereby inhibiting telomerase activity and resulting in shorter telomeres and increased chromosome instability [56]. Additionally, telRNA plays a role in telomeric heterochromatin formation and has been shown to associate with two key players in the epigenetic regulation of chromatin structure – Heterochromatin protein 1 (HP1) and methyl CpG binding protein 2 (MeCP2) [58–60].

ICF syndrome has a very complex epigenetic etiology involving altered DNA methylation patterns, chromatin structure, and histone modifications at dysregulated genes leading to wide scale changes in nuclear morphology, gene expression, and chromosome stability. Although the disorder is rare, it provides insight into the importance of DNA methylation during development, and strengthens the idea that epigenetic mechanisms are interconnected and essential for proper cellular function.

17.2.4 X-Linked IDD and α-Thalassemia Mental Retardation, X-Linked

X-linked mental retardation is the most common cause of intellectual developmental disability (XL-IDD, previously X-linked mental retardation) in males. Female carriers of XL-IDD mutations generally show favorable skewing of X-inactivation and are asymptomatic. Non-syndromic XL-IDD is caused by a number of different mutations, deletions, and duplications on the X chromosome. XL-IDD is associated with a varying degree of intellectual disability, facial abnormalities, and is often present with autism. Chromosomal regions of interest have been identified in several cases including Xp11, Xp22, and Xp28 [61, 62]. No single genetic mutation accounts for more than a few percent of cases of XL-IDD. Bioinformatic analyses of genes on the X chromosome revealed fewer than ten candidate genes and could not identify the genetic cause of XL-IDD in most families examined [63], suggesting a causative role for epigenetics.

Syndromic XL-IDD disorders, which have known genetic causes and an established set of phenotypic characteristics, are valuable tools for exploring the epigenetic mechanisms involved in X-linked neurological disorders. α-Thalassemia Mental Retardation, X-Linked (ATR-X) is a syndromic X-linked neurodevelopmental disorder that results in the dysregulation of epigenetic mechanisms. Over 200 cases of ATR-X have been identified with more than 100 different identified mutations [64]. ATR-X is caused by mutations in *ATRX* (also denoted as *XH2*), which codes for ATRX, a nuclear retained protein involved in chromatin remodeling. Deletions, duplications, missense mutations, and premature stop mutations have all

been identified as causing ATR-X [65, 66]. ATR-X is characterized by severe intellectual disability, developmental delay, genital abnormalities, microcephaly, seizures, and α-thalassemia [67]. ATRX levels increase with neuronal differentiation, and stay elevated in post-mitotic cells [68], much like the elevated levels of MeCP2 observed in post-mitotic, mature neurons. This evidence suggests an obligate role for ATRX in the epigenetic regulation of chromatin structure during neurodevelopment.

ATRX is a member of SNF2 family of chromatin remodeling factors that is targeted to heterochromatin via an interaction with methyl CpG binding protein 2 (MeCP2) [64, 69, 70]. This is of particular interest since MeCP2 is implicated in various other neurodevelopmental disorders. Most of the mutations that cause ATR-X disrupt heterochromatin targeting, highlighting the importance of chromatin structure in neuronal cells. Additional evidence indicates that ATRX interacts with the inactive X chromosome in both stem cells and somatic cells [71]. A possible role for ATRX in X-inactivation involves chromatin remodeling which promotes condensation of the inactive X chromosome, either directly or indirectly by recruiting additional factors, to effectively silence genes subject to X-inactivation.

In addition to its interactions with neuronally important proteins and its possible effects on chromatin structure, ATRX may also have an effect on DNA methylation. Although no change in methylation at imprinted genes has been found, changes in DNA methylation at repetitive regions has been observed including hypomethylation of ribosomal DNA arrays and hypermethylation of a Y-chromosome satellite repeat (DYZ2) in peripheral blood of ATR-X patients [72, 73]. The N-terminal PHD domain of ATRX, which targets it to heterochromatin, is sometimes referred to as the ATRX-DNMT3-DNMT3L (ADD) domain due to sequence and structural similarities with the *de novo* methyltransferases [4]; however, no methyltransferase activity has been observed, suggesting that ATRX does not directly alter methylation patterns but may recruit additional factors which then alter the methylation in a site-specific manner.

Mutations in *ATRX* can have severe effects on neurodevelopment. ATRX has been implicated in multiple epigenetically regulated events, including chromatin remodeling, DNA methylation, and X-inactivation, and interacts with other epigenetically important proteins. The variability in the ATR-X phenotype suggests a variety of factors influence the epigenetic mechanisms involved. As new causal genetic evidence is emerging for ATR-X and XL-IDD, it is likely that more emphasis will be placed on the importance of epigenetic regulation and chromatin remodeling in these, as well as other, neurodevelopmental disorders.

17.2.5 Rett Syndrome and the Spectrum of MECP2 Disorders

Rett syndrome (RTT) is an X-linked dominant disorder caused by mutations in *MECP2*. MeCP2 is at the center of shared pathways in human neurodevelopmental disorders and a central regulator of neuronal responses to activity-dependent DNA

methylation dynamics in the mammalian brain [74]. MeCP2 is a member of the family of methyl binding proteins that bind to methylated DNA [75]. MeCP2 is abundant in mature neurons of the mammalian brain [76–78] and DNA methylation appears to specify its chromosomal location [79, 80]. While originally thought to silence genes with methylated promoters [81, 82], the suggested functions of MeCP2 have increased in the last decade. Structural roles for MeCP2 in compacting and looping chromatin have been observed [83–85], as well as roles in RNA splicing [86], long range modulation of active genes [87], and transcriptional activation [88]. Neurons are exquisitely sensitive to MeCP2 levels and both *Mecp2* loss and gain have reciprocal detrimental effects on glutamatergic synapse strength and numbers [89]. Astrocytes also show alterations due to MeCP2 deficiency that are likely relevant to disease pathogenesis [90, 91]. Many downstream target genes have been characterized, but overall MeCP2 target genes depend on tissue and developmental context [92–96].

MECP2 mutations are the cause of RTT in >90% of the "classic" RTT cases [97]. Classic RTT is recognized clinically as affecting females who have a period of normal infancy, followed by a period of regression. The regressive period occurs between 6 and 18 months of age and is notable by the loss of language and motor skill, autistic features, characteristic hand wringing stereotypies, severe seizures, and autonomic dysfunctions such as breathing, cardiac, temperature, and gastrointestinal abnormalities. Some RTT girls develop Parkinsonian-like tremors and features as they reach adulthood [98]. Some cases of atypical RTT, characterized by infantile seizures, have mutations in the *CDKL5* [99], encoding a kinase predicted to phosphorylate MeCP2 [100], suggesting that post-translational modifications of MeCP2 are necessary for its function.

MECP2 is located at Xq28 and subject to X chromosome inactivation, thus making important differences in disease severity and pathogenesis between males and females. Females with RTT exhibit a mosaic pattern of MeCP2 expression, with some cells expressing the normal copy of MeCP2. Because they only have one X chromosome, males that inherit a RTT causing *MECP2* mutation have severe neonatal encephalopathy with death in infancy, rather than the classic RTT features seen in girls [101]. Exceptions to the severity of RTT-causing mutations in males have been observed with sex chromosome aneuploidy (47, XXY) or somatic mosaicism [102–104].

Chromosomal duplications in Xq28 that include *MECP2* are implicated in 1–2% of XL-IDD cases in males [103, 105–107]. Boys with *MECP2* duplication syndrome have clinical features including severe IDD, progressive spasticity, lack of speech, and immune abnormalities, including increased susceptibility to respiratory infections. In cases where males inherit *MECP2* duplications from their mothers, autism is also a common clinical feature [108].

Unlike the female carriers of *MECP2* mutations, females with *MECP2* duplications have highly skewed X inactivation in favor of the normal single copy allele. Mothers of these patients often have mild neuropsychiatric features including anxiety, depression, and compulsions [108]. These observations suggest that non-RTT causing *MECP2* mutations and altered expression levels may be more common than previously predicted.

MeCP2 seems to play a role in several neurodevelopmental disorders. This is likely due to its requirement during neurodevelopment and its role in neuronal maturation. This overlap is evident in postmortem cerebral cortex samples of individuals with autism and other neurodevelopmental disorders that displayed altered MeCP2 protein levels, suggesting a shared pathway among the disorders [109–111]. Epigenomic investigations of MeCP2 binding compared to DNA methylation and transcription genome-wide are likely going to be required to fully understand the diverse functions ascribed for MeCP2 and the importance of correct MeCP2 levels for the mammalian brain.

17.3 Autism Is a Spectrum of Neurodevelopmental Disorders with Complex Genetic and Environmental Contributors Affecting Epigenetic Pathways

The majority of neurodevelopmental disorders have complex genetic and environmental etiologies involving several different pathways, all of which manifest into an abnormal neurological phenotype. Autism spectrum disorder (ASD) is a prime example. ASD is termed a spectrum disorder due to the large variation in phenotype of affected individuals. ASD is characterized by abnormal social interactions and communication (both verbal and non-verbal), and the propensity for stereotyped and repetitive behaviors [112, 113]. ASD afflicts an estimated 1 in 150 individuals, with males 4 times more likely to be diagnosed than females [114, 115]. The phenotype of afflicted individuals is quite heterogeneous, with only a moderate percentage having known genetic or environmental causes.

Evidence suggests a strong genetic component since the incidence of ASD is strongly familial. Monozygotic twins have a 70–90% concordance rate, dizygotic twins have a 0–10% concordance rate, and siblings have an estimated concordance rate of 6% [116–118]. Conflicting estimates show anywhere between 10–25% of ASD cases are syndromic and have a known genetic or environmental cause [118–120]. Of these, no single genetic cause accounts for more than 2% of cases [120]. Because of the vast heterogeneity in sample populations, genome-wide association studies have had little success in identifying major genetic causes for ASD. Genetic association or linkage has been reported on every chromosome, but most genetic loci have not been replicated in subsequent studies. Chromosomal abnormalities, including duplications, deletions, and copy number variations (CNVs) together account for 5% of ASD cases, with the most common occurring on chromosome 15q11-13 [121–123]. Additional genetic regions of interest include 7q22-36, 16p11, 17q21, 22q11-13, and Xq27-28 [114, 123–125]. The number of autism candidate genes continues to fluctuate, highlighting the individuality of this disorder. It is likely that different genetic abnormalities alter the epigenetic landscape that in turn affects neurologically important pathways, resulting in phenotypically similar outcomes.

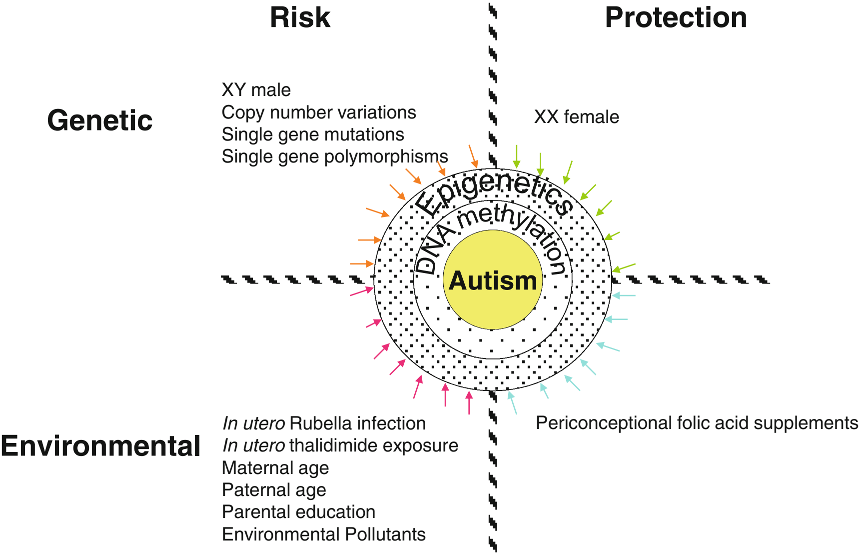


Fig. 17.2 Epigenetic mechanisms act at the interface of genes and the environment. Risk for the complex genetic disorder autism is likely influenced by multiple genetic and environmental risk and protective factors, represented by colored arrows (*orange*, genetic risk; *green*, genetic protection; *red*, environmental risk; *blue*, environmental protection). Epigenetic mechanisms such as DNA methylation (*spotted circles)* act at the interface of genetic and environmental risk and protective factors. For instance, multiple environmental risk factors correlate with global DNA hypomethylation, while folic acid supplementation promotes DNA methylation

Environmental risk factors for ASD have been identified in recent epidemiologic studies, but the understanding of environmental factors in autism is still in its infancy. A recent study of twin pairs estimates that shared environmental components play a larger role than genetic components in susceptibility to ASD [149]. Research to date indicates that a combination of genetic and environmental influences combine to affect neurodevelopment perinatally with many possible outcomes, one of which is ASD. Among the reported risk factors are increased maternal and paternal age, increased maternal education, and premature birth (less than 35 weeks gestation) [116].

Epigenetic mechanisms act at the interface of genetic and environmental risk factors for autism in considering both risk factors as well as protective factors (Fig. 17.2). Evidence for epigenetic alterations in specific gene loci has been accumulating. Deficiencies in *UBE3A* and the $GABA_A$ receptor gene *GABRB3* are common to idiopathic autism, RTT, and AS, suggesting epigenetic overlap in multiple neurodevelopmental disorders [126]. As mentioned above, reduced expression of MeCP2 is observed in 79% of autism brain samples and correlates with increased *MECP2* promoter methylation in males [109], The three 15q11-13 $GABA_A$ receptor genes, *GABRB3*, *GABRA5*, and *GABRG3,* while biallelically expressed in control

brains, show epigenetic dysregulation leading to monoallelic and reduced expression in a subset of autism cortex samples [127]. The gene encoding oxytocin receptor (*OXTR*) was found to have both genetic and DNA methylation alterations in peripheral blood of individuals with autism [128]. In addition, a recent genome-wide methylation profiling of lymphoblastoid cells identified increased methylation of the circadian clock regulator *RORA* and the repair pathway gene *BCL-2* [129]. These studies have increased the evidence for DNA methylation abnormalities in idiopathic forms of autism, but raise many questions about which environmental factors may contribute.

Recently, folate metabolism has been linked to ASD. Genetic mutations in this pathway are expected to have epigenetic effects that are highly relevant to autism epidemiology. Folate metabolism is intertwined with epigenetics in that it produces S-adenosyl-methionine (SAM), which serves as a universal methyl-group donor for DNA, RNA, and histone methylation. SAM is converted to S-adenosyl-homocysteine (SAH) during methylation, and the balance between SAM and SAH is important for methyltransferase activity. Methylene tetrahydrafolate reductase (MTHFR) is an enzyme important for SAH metabolism and is required to maintain the proper SAM:SAH ratio. Dietary intake of B vitamins provides some of the components necessary for folate metabolism, and deficiencies have been linked to neural tube defects [130].

Evidence has emerged that implicates polymorphisms in *MTHFR* (specifically C677T) in some cases of idiopathic autism [131–134]. This polymorphism is associated with a 2.79 fold increase in autism risk. Additional polymorphisms have been found in *MTHFR* at A1298C, and are suggested to increase the risk for autism 8-fold when found in combination with the C677T allele [133]. These polymorphisms reduce enzymatic activity in a dose-dependent manner [135].

These studies of mutations in genes important in folate metabolism were spurred by observations that plasma from autistic patients had decreased levels of SAM and methionine and increased levels of SAH, in addition to altered levels of other metabolites important in folate and one-carbon metabolism. The two-fold decrease in SAM:SAH observed in autistic cases suggests an imbalance in the folate pathway which could affect DNA methylation [136]. SAH levels were significantly elevated in mothers of autistic children, resulting in a decreased SAM:SAH ratio. This increase in SAH was linked to a 7-fold increase in the likelihood of having a child with autism. Additionally, the maternal samples with the highest levels of SAH were shown to have global hypomethylation. It has been theorized that maternal increases in SAH can affect the methylation status of the developing fetus by altering the SAM:SAH ratio *in utero* [137]. A combination of factors most likely influences the prenatal methylation status of children with autism, including the mother's genotype and dietary intake.

A recent work has disputed the importance of polymorphisms in *MTHFR* and instead posits that maternal polymorphisms in *RFC1*, which codes for a carrier protein involved in folate metabolism, are a major contributor to autism risk. Mothers homozygous for *RFC1* A80G were 46% more likely to have a child with autism. Interestingly, the increased autism risk was independent of the child's genotype [138].

This observation, combined with hypomethylation in the mothers of autistic children and the children themselves, supports the idea of environmental epigenetic alterations *in utero* leading to neurodevelopmental disorders. Polymorphisms in genes whose products are important in folate and one carbon metabolism may increase the individual's need for folate and other B vitamins. If these needs are not met, pre- and post-natal abnormalities in neurodevelopment may result.

Of interest is the idea that maternal genetics and environmental exposures can have an effect on the epigenetic neurodevelopment of the offspring. For example, genetic mutations in the mother that lead to increased SAH levels can influence the epigenetic state of the fetus including DNA and histone methylation. This can lead to epigenetic effects which persist after birth. Alternatively, if the genetic mutations is present in the fetus, *in utero* neurodevelopment progresses normally dependent on high maternal dietary intake of folate and input of intermediaries from the maternal folate pathway. Postnatal neurodevelopment, however, does not follow a normal trajectory due to abnormal folate and one carbon metabolism in the offspring, which essentially leads to a build up of post-natally acquired epigenetic abnormalities and possibly leads to neurodevelopmental disorders [139]. A recent epidemiology study has shown a protective effect of prenatal vitamin supplementation taking prior to conception and in the first trimester of pregnancy [140], supporting a potential role of DNA methylation in the protection.

In addition to the nutritional environment of the developing fetus, environmental toxins are expected to contribute to autism risk [141]. Environmental toxins could act through mutations and DNA repair pathways [142] or through epigenetic pathways such as DNA methylation, but an exploration of potential human exposures and autism risk is still in its infancy. Two possible environmental exposures from epidemiological studies of autism include organochloride pesticides from proximity to agricultural fields and air pollution from residential proximity to freeways [143, 144]. Interestingly, human air pollution exposure has been shown to negatively correlate with global levels of DNA methylation at *LINE-1* and *Alu* repeats [145, 146].

In addition to intersecting with genetic pathways, environmental exposures are expected to intersect with nutritional pathways. A unifying mechanism for a combination of diet and environmental exposures has been proposed [147]. Because the one-carbon metabolism and the glutathione (GSH) synthesis pathways are biochemically linked, enhanced need for GSH, an important antioxidant, to conjugate chemicals acts as an added "demand" on the availability of methyl-donors for DNA methylation. Deficiencies in methylation, GSH, and oxidative stress pathways have been implicated in autism [136].

17.4 Conclusions

It is becoming increasingly clear that environmental and genetic factors significantly impact the epigenome. The epigenome is responsible for modulating gene expression in all tissue types, and is especially important in neuronal cells.

The epigenome serves to interpret environmental signals and adjust the neuronal environment and transcriptional activity accordingly. Alterations in DNA methylation, whether large scale or site specific, can significantly affect transcription of neuronally important genes. Alterations in histone modifications can affect the spatial arrangement of chromatin, restricting or allowing access to transcriptional machinery. Understanding the improper progression of epigenetic changes that occurs in neurodevelopmental disorders can shed light on both genetic and epigenetic causes, and hopefully lead to effective treatments.

Acknowledgements The authors thank Dr. Michael Gonzales and Weston Powell for helpful review of the manuscript. This work was supported in part by NIH grants 1R01ES015171, 1R01HD048799, and 2R01HD041462 (JML); and by NIEHS training grant T32 ES007058-33 (RW).

References

1. Yen RW, Vertino PM, Nelkin BD, Yu JJ, el-Deiry W, Cumaraswamy A et al (1992) Isolation and characterization of the cDNA encoding human DNA methyltransferase. Nucleic Acids Res 20:2287–2291
2. Goto K, Numata M, Komura JI, Ono T, Bestor TH, Kondo H (1994) Expression of DNA methyltransferase gene in mature and immature neurons as well as proliferating cells in mice. Differentiation 56:39–44
3. Feng J, Chang H, Li E, Fan G (2005) Dynamic expression of de novo DNA methyltransferases Dnmt3a and Dnmt3b in the central nervous system. J Neurosci Res 79:734–746
4. Argentaro A, Yang JC, Chapman L, Kowalczyk MS, Gibbons RJ, Higgs DR et al (2007) Structural consequences of disease-causing mutations in the ATRX-DNMT3-DNMT3L (ADD) domain of the chromatin-associated protein ATRX. Proc Natl Acad Sci USA 104:11939–11944
5. Miller CA, Sweatt JD (2007) Covalent modification of DNA regulates memory formation. Neuron 53:857–869
6. LaPlant Q, Vialou V, Covington HE 3rd, Dumitriu D, Feng J, Warren BL et al (2010) Dnmt3a regulates emotional behavior and spine plasticity in the nucleus accumbens. Nat Neurosci 13:1137–1143
7. Wilkinson LS, Davies W, Isles AR (2007) Genomic imprinting effects on brain development and function. Nat Rev Neurosci 8:832–843
8. Thatcher K, LaSalle JM (2006) Dynamic changes in histone H3 lysine 9 acetylation localization patterns during neuronal maturation require MeCP2. Epigenetics 1:24–31
9. Chateauvieux S, Morceau F, Dicato M, Diederich M (2010) Molecular and therapeutic potential and toxicity of valproic acid. J Biomed Biotechnol 2010:1–19
10. Keverne EB, Fundele R, Narasimha M, Barton SC, Surani MA (1996) Genomic imprinting and the differential roles of parental genomes in brain development. Brain Res Dev Brain Res 92:91–100
11. Schanen NC (2006) Epigenetics of autism spectrum disorders. Hum Mol Genet 15(Spec No 2):R138–R150
12. Lalande M, Calciano MA (2007) Molecular epigenetics of angelman syndrome. Cell Mol Life Sci 64:947–960
13. Hogart A, Leung KN, Wang NJ, Wu DJ, Driscoll J, Vallero RO et al (2009) Chromosome 15q11-13 duplication syndrome brain reveals epigenetic alterations in gene expression not predicted from copy number. J Med Genet 46:86–93

14. Yashiro K, Riday TT, Condon KH, Roberts AC, Bernardo DR, Prakash R et al (2009) Ube3a is required for experience-dependent maturation of the neocortex. Nat Neurosci 12:777–783
15. Sato M, Stryker MP (2010) Genomic imprinting of experience-dependent cortical plasticity by the ubiquitin ligase gene Ube3a. Proc Natl Acad Sci USA 107:5611–5616
16. Greer PL, Hanayama R, Bloodgood BL, Mardinly AR, Lipton DM, Flavell SW et al (2010) The angelman syndrome protein Ube3A regulates synapse development by ubiquitinating arc. Cell 140:704–716
17. Robinson WP, Bottani A, Xie YG, Balakrishman J, Binkert F, Machler M et al (1991) Molecular, cytogenetic, and clinical investigations of prader-willi syndrome patients. Am J Hum Genet 49:1219–1234
18. Robinson WP, Dutly F, Nicholls RD, Bernasconi F, Penaherrera M, Michaelis RC et al (1998) The mechanisms involved in formation of deletions and duplications of 15q11-q13. J Med Genet 35:130–136
19. Buiting K, Saitoh S, Gross S, Dittrich B, Schwartz S, Nicholls RD et al (1995) Inherited microdeletions in the angelman and prader-willi syndromes define an imprinting centre on human chromosome 15. Nat Genet 9:395–400
20. Sutcliffe JS, Nakao M, Christian S, Orstavik KH, Tommerup N, Ledbetter DH et al (1994) Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. Nat Genet 8:52–58
21. Dittrich B, Buiting K, Korn B, Rickard S, Buxton J, Saitoh S et al (1996) Imprint switching on human chromosome 15 may involve alternative transcripts of the SNRPN gene. Nat Genet 14:163–170
22. Glenn CC, Nicholls RD, Robinson WP, Saitoh S, Niikawa N, Schinzel A et al (1993) Modification of 15q11-q13 DNA methylation imprints in unique angelman and prader-willi patients. Hum Mol Genet 2:1377–1382
23. Runte M, Huttenhofer A, Gross S, Kiefmann M, Horsthemke B, Buiting K (2001) The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. Hum Mol Genet 10:2687–2700
24. Leung KN, Vallero RO, DuBose AJ, Resnick JL, LaSalle JM (2009) Imprinting regulates mammalian snoRNA-encoding chromatin decondensation and neuronal nucleolar size. Hum Mol Genet 18:4227–4238
25. Hogart A et al (2010) The comorbidity of autism with the genomic disorders of chromosome 15q11.2-q13. Neurobiol Dis 38(2):181–191
26. Robinson WP, Kuchinka BD, Bernasconi F, Petersen MB, Schulze A, Brondum-Nielsen K et al (1998) Maternal meiosis I non-disjunction of chromosome 15: dependence of the maternal age effect on level of recombination. Hum Mol Genet 7:1011–1019
27. Cook EH Jr, Lindgren V, Leventhal BL, Courchesne R, Lincoln A, Shulman C et al (1997) Autism or atypical autism in maternally but not paternally derived proximal 15q duplication. Am J Hum Genet 60:928–934
28. Veltman MW, Thompson RJ, Craig EE, Dennis NR, Roberts SE, Moore V et al (2005) A paternally inherited duplication in the prader-willi/angelman syndrome critical region: a case and family study. J Autism Dev Disord 35:117–127
29. Mohandas TK, Park JP, Spellman RA, Filiano JJ, Mamourian AC, Hawk AB et al (1999) Paternally derived de novo interstitial duplication of proximal 15q in a patient with developmental delay. Am J Med Genet 82:294–300
30. Kao DI, Aldridge GM, Weiler IJ, Greenough WT (2010) Altered mRNA transport, docking, and protein translation in neurons lacking fragile X mental retardation protein. Proc Natl Acad Sci USA 107:15601–15606
31. Brown V, Small K, Lakkis L, Feng Y, Gunter C, Wilkinson KD et al (1998) Purified recombinant Fmrp exhibits selective RNA binding as an intrinsic property of the fragile X mental retardation protein. J Biol Chem 273:15521–15527
32. Hagerman RJ, Berry-Kravis E, Kaufmann WE, Ono MY, Tartaglia N, Lachiewicz A et al (2009) Advances in the treatment of fragile X syndrome. Pediatrics 123:378–390

33. Chonchaiya W, Schneider A, Hagerman RJ (2009) Fragile X: a family of disorders. Adv Pediatr 56:165–186
34. Hagerman PJ (2008) The fragile X prevalence paradox. J Med Genet 45:498–499
35. Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, Pizzuti A et al (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell 65:905–914
36. Fu YH, Kuhl DP, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S et al (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67:1047–1058
37. Pieretti M, Zhang FP, Fu YH, Warren ST, Oostra BA, Caskey CT et al (1991) Absence of expression of the FMR-1 gene in fragile X syndrome. Cell 66:817–822
38. Sutcliffe JS, Nelson DL, Zhang F, Pieretti M, Caskey CT, Saxe D et al (1992) DNA methylation represses FMR-1 transcription in fragile X syndrome. Hum Mol Genet 1:397–400
39. Malter HE, Iber JC, Willemsen R, de Graaff E, Tarleton JC, Leisti J et al (1997) Characterization of the full fragile X syndrome mutation in fetal gametes. Nat Genet 15:165–169
40. Kumari D, Usdin K (2010) The distribution of repressive histone modifications on silenced FMR1 alleles provides clues to the mechanism of gene silencing in fragile X syndrome. Hum Mol Genet 19:4634–4642
41. Hagerman RJ, Leehey M, Heinrichs W, Tassone F, Wilson R, Hills J et al (2001) Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. Neurology 57:127–130
42. Jacquemont S, Hagerman RJ, Leehey MA, Hall DA, Levine RA, Brunberg JA et al (2004) Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population. JAMA 291:460–469
43. Snow K, Doud LK, Hagerman R, Pergolizzi RG, Erster SH, Thibodeau SN (1993) Analysis of a CGG sequence at the FMR-1 locus in fragile X families and in the general population. Am J Hum Genet 53:1217–1228
44. Berry-Kravis E, Potanos K, Weinberg D, Zhou L, Goetz CG (2005) Fragile X-associated tremor/ataxia syndrome in sisters related to X-inactivation. Ann Neurol 57:144–147
45. Tassone F, Hagerman RJ, Taylor AK, Gane LW, Godfrey TE, Hagerman PJ (2000) Elevated levels of FMR1 mRNA in carrier males: a new mechanism of involvement in the fragile-X syndrome. Am J Hum Genet 66:6–15
46. Greco CM, Hagerman RJ, Tassone F, Chudley AE, Del Bigio MR, Jacquemont S et al (2002) Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. Brain 125:1760–1771
47. Arocena DG, Iwahashi CK, Won N, Beilina A, Ludwig AL, Tassone F et al (2005) Induction of inclusion formation and disruption of lamin A/C structure by premutation CGG-repeat RNA in human cultured neural cells. Hum Mol Genet 14:3661–3671
48. Tassone F, Adams J, Berry-Kravis EM, Cohen SS, Brusco A, Leehey MA et al (2007) CGG repeat length correlates with age of onset of motor signs of the fragile X-associated tremor/ataxia syndrome (FXTAS). Am J Med Genet B Neuropsychiatr Genet 144B:566–569
49. Willemsen R, Hoogeveen-Westerveld M, Reis S, Holstege J, Severijnen LA, Nieuwenhuizen IM et al (2003) The FMR1 CGG repeat mouse displays ubiquitin-positive intranuclear neuronal inclusions; implications for the cerebellar tremor/ataxia syndrome. Hum Mol Genet 12:949–959
50. Hagleitner MM, Lankester A, Maraschio P, Hulten M, Fryns JP, Schuetz C et al (2008) Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome). J Med Genet 45:93–99
51. Ehrlich M, Sanchez C, Shao C, Nishiyama R, Kehrl J, Kuick R et al (2008) ICF, an immunodeficiency syndrome: DNA methyltransferase 3B involvement, chromosome anomalies, and gene dysregulation. Autoimmunity 41:253–271
52. Sumner AT, Mitchell AR, Ellis PM (1998) A FISH study of chromosome fusion in the ICF syndrome: involvement of paracentric heterochromatin but not of the centromeres themselves. J Med Genet 35:833–835

53. Brown DC, Grace E, Sumner AT, Edmunds AT, Ellis PM (1995) ICF syndrome (immunodeficiency, centromeric instability and facial anomalies): investigation of heterochromatin abnormalities and review of clinical outcome. Hum Genet 96:411–416
54. Jefferson A, Colella S, Moralli D, Wilson N, Yusuf M, Gimelli G et al (2010) Altered intranuclear organisation of heterochromatin and genes in ICF syndrome. PLoS One 5:e11364
55. Jin B, Tao Q, Peng J, Soo HM, Wu W, Ying J et al (2008) DNA methyltransferase 3B (DNMT3B) mutations in ICF syndrome lead to altered epigenetic modifications and aberrant expression of genes regulating development, neurogenesis and immune function. Hum Mol Genet 17:690–709
56. Yehezkel S, Segev Y, Viegas-Pequignot E, Skorecki K, Selig S (2008) Hypomethylation of subtelomeric regions in ICF syndrome is associated with abnormally short telomeres and enhanced transcription from telomeric regions. Hum Mol Genet 17:2776–2789
57. Schoeftner S, Blasco MA (2008) Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. Nat Cell Biol 10:228–236
58. Deng Z, Campbell AE, Lieberman PM (2010) TERRA, CpG methylation and telomere heterochromatin: lessons from ICF syndrome cells. Cell Cycle 9:69–74
59. Deng Z, Norseen J, Wiedmer A, Riethman H, Lieberman PM (2009) TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. Mol Cell 35:403–413
60. Perrini B, Piacentini L, Fanti L, Altieri F, Chichiarelli S, Berloco M et al (2004) HP1 Controls telomere capping, telomere elongation, and telomere silencing by two different mechanisms in Drosophila. Mol Cell 15:467–476
61. Gecz J, Shoubridge C, Corbett M (2009) The genetic landscape of intellectual disability arising from chromosome X. Trends Genet 25:308–316
62. El-Hattab AW et al (2011) Microduplication of Xp11.23p11.3 with effects on cognition, behavior, and craniofacial development. Clin Genet 79(6):531–538
63. Tarpey PS, Smith R, Pleasance E, Whibley A, Edkins S, Hardy C et al (2009) A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. Nat Genet 41:535–543
64. Gibbons RJ, Wada T, Fisher CA, Malik N, Mitson MJ, Steensma DP et al (2008) Mutations in the chromatin-associated protein ATRX. Hum Mutat 29:796–802
65. Gibbons RJ, Picketts DJ, Villard L, Higgs DR (1995) Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). Cell 80:837–845
66. Cohn DM, Pagon RA, Hudgins L, Schwartz CE, Stevenson RE, Friez MJ (2009) Partial ATRX gene duplication causes ATR-X syndrome. Am J Med Genet A 149A:2317–2320
67. Gibbons R (2006) Alpha thalassaemia-mental retardation, X linked. Orphanet J Rare Dis 1:15
68. Berube NG, Mangelsdorf M, Jagla M, Vanderluit J, Garrick D, Gibbons RJ et al (2005) The chromatin-remodeling protein ATRX is critical for neuronal survival during corticogenesis. J Clin Invest 115:258–267
69. McDowell TL, Gibbons RJ, Sutherland H, O'Rourke DM, Bickmore WA, Pombo A et al (1999) Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. Proc Natl Acad Sci USA 96:13983–13988
70. Nan X, Hou J, Maclean A, Nasir J, Lafuente MJ, Shu X et al (2007) Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. Proc Natl Acad Sci USA 104:2709–2714
71. Baumann C, De La Fuente R (2009) ATRX marks the inactive X chromosome (Xi) in somatic cells and during imprinted X chromosome inactivation in trophoblast stem cells. Chromosoma 118:209–222
72. Gibbons RJ, McDowell TL, Raman S, O'Rourke DM, Garrick D, Ayyub H et al (2000) Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. Nat Genet 24:368–371

73. Kernohan KD, Jiang Y, Tremblay DC, Bonvissuto AC, Eubanks JH, Mann MR et al (2010) ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain. Dev Cell 18:191–202
74. Gonzales ML, LaSalle JM (2010) The role of MeCP2 in brain development and neurodevelopmental disorders. Curr Psychiatry Rep 12:127–134
75. Ng HH, Bird A (1999) DNA methylation and chromatin modification. Curr Opin Genet Dev 9:158–163
76. Shahbazian MD, Antalffy B, Armstrong DL, Zoghbi HY (2002) Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. Hum Mol Genet 11:115–124
77. Balmer D, Goldstine J, Rao YM, LaSalle JM (2003) Elevated methyl-CpG-binding protein 2 expression is acquired during postnatal human brain development and is correlated with alternative polyadenylation. J Mol Med 81:61–68
78. Braunschweig D, Simcox T, Samaco RC, LaSalle JM (2004) X-Chromosome inactivation ratios affect wild-type MeCP2 expression within mosaic Rett syndrome and Mecp2−/+ mouse brain. Hum Mol Genet 13:1275–1286
79. Nan X, Tate P, Li E, Bird A (1996) DNA methylation specifies chromosomal localization of MeCP2. Mol Cell Biol 16:414–421
80. Skene PJ, Illingworth RS, Webb S, Kerr AR, James KD, Turner DJ et al (2010) Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. Mol Cell 37:457–468
81. Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N et al (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet 19:187–191
82. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN et al (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393:386–389
83. Georgel PT, Horowitz-Scherer RA, Adkins N, Woodcock CL, Wade PA, Hansen JC (2003) Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. J Biol Chem 278:32181–32188
84. Horike S, Cai S, Miyano M, Cheng JF, Kohwi-Shigematsu T (2005) Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nat Genet 37:31–40
85. Kumar A, Kamboj S, Malone BM, Kudo S, Twiss JL, Czymmek KJ et al (2008) Analysis of protein domains and Rett syndrome mutations indicate that multiple regions influence chromatin-binding dynamics of the chromatin-associated protein MECP2 in vivo. J Cell Sci 121:1128–1137
86. Young JI, Hong EP, Castle JC, Crespo-Barreto J, Bowman AB, Rose MF et al (2005) Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. Proc Natl Acad Sci USA 102:17551–17558
87. Yasui DH, Peddada S, Bieda MC, Vallero RO, Hogart A, Nagarajan RP et al (2007) Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. Proc Natl Acad Sci USA 104:19416–19421
88. Chahrour M, Jung SY, Shaw C, Zhou X, Wong ST, Qin J et al (2008) MeCP2, a key contributor to neurological disease, activates and represses transcription. Science 320:1224–1229
89. Chao HT, Zoghbi HY, Rosenmund C (2007) MeCP2 controls excitatory synaptic strength by regulating glutamatergic synapse number. Neuron 56:58–65
90. Maezawa I, Swanberg S, Harvey D, LaSalle JM, Jin LW (2009) Rett syndrome astrocytes are abnormal and spread MeCP2 deficiency through gap junctions. J Neurosci 29:5051–5061
91. Ballas N, Lioy DT, Grunseich C, Mandel G (2009) Non-cell autonomous influence of MeCP2-deficient glia on neuronal dendritic morphology. Nat Neurosci 12:311–317
92. Traynor J, Agarwal P, Lazzeroni L, Francke U (2002) Gene expression patterns vary in clonal cell cultures from Rett syndrome females with eight different MECP2 mutations. BMC Med Genet 3:12

93. Tudor M, Akbarian S, Chen RZ, Jaenisch R (2002) Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. Proc Natl Acad Sci USA 99:15536–15541
94. Peddada S, Yasui DH, LaSalle JM (2006) Inhibitors of differentiation (ID1, ID2, ID3 and ID4) genes are neuronal targets of MeCP2 that are elevated in Rett syndrome. Hum Mol Genet 15:2003–2014
95. Deng V, Matagne V, Banine F, Frerking M, Ohliger P, Budden S et al (2007) FXYD1 is an MeCP2 target gene overexpressed in the brains of Rett syndrome patients and Mecp2-null mice. Hum Mol Genet 16:640–650
96. Urdinguio RG, Lopez-Serra L, Lopez-Nieva P, Alaminos M, Diaz-Uriarte R, Fernandez AF et al (2008) Mecp2-null mice provide new neuronal targets for Rett syndrome. PLoS One 3:e3669
97. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY (1999) Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl- CpG-binding protein 2. Nat Genet 23:185–188
98. Neul JL, Fang P, Barrish J, Lane J, Caeg EB, Smith EO et al (2008) Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. Neurology 70:1313–1321
99. Lin C, Franco B, Rosner MR (2005) CDKL5/Stk9 kinase inactivation is associated with neuronal developmental disorders. Hum Mol Genet 14:3775–3786
100. Mari F, Azimonti S, Bertani I, Bolognese F, Colombo E, Caselli R et al (2005) CDKL5 belongs to the same molecular pathway of MeCP2 and it is responsible for the early-onset seizure variant of Rett syndrome. Hum Mol Genet 14:1935–1946
101. Schanen C (2001) Rethinking the fate of males with mutations in the gene that causes Rett syndrome. Brain Dev 23(Suppl 1):S144–S146
102. Clayton-Smith J, Watson P, Ramsden S, Black GC (2000) Somatic mutation in *MECP2* as a non-fatal neurodevelopmental disorder in males. Lancet 356:830–832
103. Hoffbuhr K, Devaney JM, LaFleur B, Sirianni N, Scacheri C, Giron J et al (2001) MeCP2 mutations in children with and without the phenotype of Rett syndrome. Neurology 56:1486–1495
104. Hammer S, Dorrani N, Dragich J, Kudo S, Schanen C (2002) The phenotypic consequences of MECP2 mutations extend beyond Rett syndrome. Ment Retard Dev Disabil Res Rev 8:94–98
105. Van Esch H, Bauters M, Ignatius J, Jansen M, Raynaud M, Hollanders K et al (2005) Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. Am J Hum Genet 77:442–453
106. Lugtenberg D, Kleefstra T, Oudakker AR, Nillesen WM, Yntema HG, Tzschach A et al (2009) Structural variation in Xq28: MECP2 duplications in 1% of patients with unexplained XLMR and in 2% of male patients with severe encephalopathy. Eur J Hum Genet 17:444–453
107. Campos M Jr, Churchman SM, Santos-Reboucas CB, Ponchel F, Pimentel MM (2010) High frequency of nonrecurrent MECP2 duplications among Brazilian males with mental retardation. J Mol Neurosci 41:105–109
108. Ramocki MB, Peters SU, Tavyev YJ, Zhang F, Carvalho CM, Schaaf CP et al (2009) Autism and other neuropsychiatric symptoms are prevalent in individuals with MeCP2 duplication syndrome. Ann Neurol 66:771–782
109. Nagarajan RP, Hogart AR, Gwye Y, Martin MR, Lasalle JM (2006) Reduced MeCP2 expression is frequent in autism frontal cortex and correlates with aberrant MECP2 promoter methylation. Epigenetics 1:172–182
110. Nagarajan RP, Patzel KA, Martin M, Yasui DH, Swanberg SE, Hertz-Picciotto I et al (2008) Mecp2 Promoter methylation and X chromosome inactivation in autism. Autism Res 1:169–178
111. Samaco RC, Nagarajan RP, Braunschweig D, LaSalle JM (2004) Multiple pathways regulate MeCP2 expression in normal brain development and exhibit defects in autism-spectrum disorders. Hum Mol Genet 13:629–639

112. Patterson SY, Smith V, Jelen M (2010) Behavioural intervention practices for stereotypic and repetitive behaviour in individuals with autism spectrum disorder: a systematic review. Dev Med Child Neurol 52:318–327
113. Jones JR, Skinner C, Friez MJ, Schwartz CE, Stevenson RE (2008) Hypothesis: dysregulation of methylation of brain-expressed genes on the X chromosome and autism spectrum disorders. Am J Med Genet A 146A:2213–2220
114. Bill BR, Geschwind DH (2009) Genetic advances in autism: heterogeneity and convergence on shared pathways. Curr Opin Genet Dev 19:271–278
115. Holt R, Barnby G, Maestrini E, Bacchelli E, Brocklebank D, Sousa I et al (2010) Linkage and candidate gene studies of autism spectrum disorders in European populations. Eur J Hum Genet 18:1013–1019
116. Yirmiya N, Charman T (2010) The prodrome of autism: early behavioral and biological signs, regression, peri- and post-natal development and genetics. J Child Psychol Psychiatry 51:432–458
117. Lasalle JM, Yasui DH (2009) Evolving role of MeCP2 in Rett syndrome and autism. Epigenomics 1:119–130
118. Toro R, Konyukh M, Delorme R, Leblond C, Chaste P, Fauchereau F et al (2010) Key role for gene dosage and synaptic homeostasis in autism spectrum disorders. Trends Genet 26:363–372
119. Landrigan PJ (2010) What causes autism? Exploring the environmental contribution. Curr Opin Pediatr 22:219–225
120. Rubenstein JL (2010) Three hypotheses for developmental defects that may underlie some forms of autism spectrum disorder. Curr Opin Neurol 23:118–123
121. Delahanty RJ, Kang JQ, Brune CW, Kistner EO, Courchesne E, Cox NJ et al (2009) Maternal transmission of a rare GABRB3 signal peptide variant is associated with autism. Mol Psychiatry
122. Castermans D, Volders K, Crepel A, Backx L, De Vos R, Freson K et al (2010) SCAMP5, NBEA and AMISYN: three candidate genes for autism involved in secretion of large dense-core vesicles. Hum Mol Genet 19:1368–1378
123. Freitag CM, Staal W, Klauck SM, Duketis E, Waltes R (2010) Genetics of autistic disorders: review and clinical implications. Eur Child Adolesc Psychiatry 19:169–178
124. Abrahams BS, Geschwind DH (2010) Connecting genes to brain in the autism spectrum disorders. Arch Neurol 67:395–399
125. Gong X, Bacchelli E, Blasi F, Toma C, Betancur C, Chaste P et al (2008) Analysis of X chromosome inactivation in autism spectrum disorders. Am J Med Genet B Neuropsychiatr Genet 147B:830–835
126. Samaco RC, Hogart A, LaSalle JM (2005) Epigenetic overlap in autism-spectrum neurodevelopmental disorders: MECP2 deficiency causes reduced expression of UBE3A and GABRB3. Hum Mol Genet 14:483–492
127. Hogart A, Nagarajan RP, Patzel KA, Yasui DH, Lasalle JM (2007) 15q11-13 GABAA receptor genes are normally biallelically expressed in brain yet are subject to epigenetic dysregulation in autism-spectrum disorders. Hum Mol Genet 16:691–703
128. Gregory SG, Connelly JJ, Towers AJ, Johnson J, Biscocho D, Markunas CA et al (2009) Genomic and epigenetic evidence for oxytocin receptor deficiency in autism. BMC Med 7:62
129. Nguyen A, Rauch TA, Pfeifer GP, Hu VW (2010) Global methylation profiling of lymphoblastoid cell lines reveals epigenetic contributions to autism spectrum disorders and a novel autism candidate gene, RORA, whose protein product is reduced in autistic brain. FASEB J 24:3036–3051
130. Rush D (1994) Periconceptional folate and neural tube defect. Am J Clin Nutr 59:511S–515S, discussion 515S-516S
131. Adams M, Lucock M, Stuart J, Fardell S, Baker K, Ng X (2007) Preliminary evidence for involvement of the folate gene polymorphism 19 bp deletion-DHFR in occurrence of autism. Neurosci Lett 422:24–29

132. Pasca SP, Dronca E, Kaucsar T, Craciun EC, Endreffy E, Ferencz BK et al (2009) One carbon metabolism disturbances and the C677T MTHFR gene polymorphism in children with autism spectrum disorders. J Cell Mol Med 13:4229–4238
133. Mohammad NS, Jain JM, Chintakindi KP, Singh RP, Naik U, Akella RR (2009) Aberrations in folate metabolic pathway and altered susceptibility to autism. Psychiatr Genet 19:171–176
134. Goin-Kochel RP, Porter AE, Peters SU, Shinawi M, Sahoo T, Beaudet AL (2009) The MTHFR 677 C–>T polymorphism and behaviors in children with autism: exploratory genotype-phenotype correlations. Autism Res 2:98–108
135. Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, Matthews RG et al (1995) A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. Nat Genet 10:111–113
136. James SJ, Cutler P, Melnyk S, Jernigan S, Janak L, Gaylor DW et al (2004) Metabolic biomarkers of increased oxidative stress and impaired methylation capacity in children with autism. Am J Clin Nutr 80:1611–1617
137. James SJ, Melnyk S, Jernigan S, Hubanks A, Rose S, Gaylor DW (2008) Abnormal transmethylation/transsulfuration metabolism and DNA hypomethylation among parents of children with autism. J Autism Dev Disord 38:1966–1975
138. James SJ, Melnyk S, Jernigan S, Pavliv O, Trusty T, Lehman S et al (2010) A functional polymorphism in the reduced folate carrier gene and DNA hypomethylation in mothers of children with autism. Am J Med Genet B Neuropsychiatr Genet 153B:1209–1220
139. Rogers EJ (2008) Has enhanced folate status during pregnancy altered natural selection and possibly autism prevalence? A closer look at a possible link. Med Hypotheses. doi:10.1016/j.mehy.2008.04.013
140. Schmidt RJ, Hansen RL, Hartiala J, Allayee H, Schmidt LC, Tancredi DJ et al (2011) Prenatal vitamins, one-carbon metabolism gene variants, and risk for autism. Epidemiology 22:476–485
141. Herbert MR (2010) Contributions of the environment and environmentally vulnerable physiology to autism spectrum disorders. Curr Opin Neurol 23:103–110
142. Kinney DK, Barch DH, Chayka B, Napoleon S, Munir KM (2010) Environmental risk factors for autism: do they help cause de novo genetic mutations that contribute to the disorder? Med Hypotheses 74:102–106
143. Roberts EM, English PB, Grether JK, Windham GC, Somberg L, Wolff C (2007) Maternal residence near agricultural pesticide applications and autism spectrum disorders among children in the California Central Valley. Environ Health Perspect 115:1482–1489
144. Windham GC, Zhang L, Gunier R, Croen LA, Grether JK (2006) Autism spectrum disorders in relation to distribution of hazardous air pollutants in the San Francisco bay area. Environ Health Perspect 114:1438–1444
145. Baccarelli A, Wright RO, Bollati V, Tarantini L, Litonjua AA, Suh HH et al (2009) Rapid DNA methylation changes after exposure to traffic particles. Am J Respir Crit Care Med 179:572–578
146. Tarantini L, Bonzini M, Apostoli P, Pegoraro V, Bollati V, Marinelli B et al (2009) Effects of particulate matter on genomic DNA methylation content and iNOS promoter methylation. Environ Health Perspect 117:217–222
147. Lee DH, Jacobs DR Jr, Porta M (2009) Hypothesis: a unifying mechanism for nutrition and chemicals as lifelong modulators of DNA hypomethylation. Environ Health Perspect 117:1799–1802
148. de Greef JC et al (2011) Mutations in ZBTB24 are associated with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. Am J Hum Genet 88(6):796–804
149. Hallmayer J et al (2011) Genetic heritability and shared environmental factors among twin pairs with autism. Arch Gen Psychiatry, Accepted Article

Chapter 18
Epigenetic Epidemiology of Psychiatric Disorders

Bart P.F. Rutten and Jim van Os

Contents

18.1 Introduction 346
18.2 Epigenetic Mechanisms in Development and Aging of the Brain 346
18.3 Psychiatric Phenotypes in Disorders of Genetic Imprinting 348
18.4 Twin Discordance 348
18.5 Parent-of-Origin Effect and Paternal Age 350
18.6 Sex-Differences in Psychiatric Disorders 351
18.7 Enduring Influences of Environmental Exposures 352
18.7.1 Pregnancy and Birth Complications 352
18.7.2 Rearing Environment and Childhood Abuse 354
18.7.3 Drugs of Abuse 355
18.7.4 Minority Group Position and Social Defeat 356
18.7.5 Synergism in Environmental Exposures 357
18.8 Variants in Epigenetic-Relevant Genes and Psychiatric Disorders 358
18.8.1 MTHFR 358
18.8.2 MeCP2 358
18.8.3 Histone Genes 359
18.9 Aberrant Epigenetic Profiles in Psychiatric Disorders 359
18.9.1 Depression, PTSD, and Suicide 359
18.9.2 Schizophrenia and Bipolar Disorder 359
18.9.3 Alzheimer's Disease 360
18.10 Summary and Perspectives 361
References 364

B.P.F. Rutten (✉)
Department of Psychiatry and Psychology, School for Mental Health and Neuroscience, European Graduate School of Neuroscience (EURON), South Limburg Mental Health Research and Teaching Network (SEARCH), Maastricht University Medical Centre, PO BOX 616 (VIJV) 6200, Maastricht, MD, The Netherlands
e-mail: b.rutten@maastrichtuniversity.nl

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_18,
© Springer Science+Business Media B.V. 2012

Abstract Exciting developments in the field of epigenetics have generated great interest within psychiatric epidemiology to focus on direct and indirect evidence for epigenetic involvement in behavior, mental health, and complex psychiatric disorders. Epidemiologic evidence on epigenetics in psychiatry, however, is currently very sparsely available. With the aim to address the current status of the literature on evidence indicative of involvement of epigenetic mechanisms in psychiatric disorders, we describe a clear role for epigenetic mechanisms in development and aging of the brain, with experiences and environmental exposures particularly during early life having considerable impact on the development of functional abilities of the brain. Besides the psychiatric consequences of classical syndromes of genetic imprinting in humans, findings of twin discordance, parent-of-origin effects, paternal age effects, and sex differences in psychiatric disorders suggest epigenetic involvement in the etiology of psychiatric disorders. The evidence is further strengthened by observations of endurable effects of various environmental exposures during life on risk of psychiatric disorders, and preliminary epigenetic studies showing differential epigenetic profiles in patients with several psychiatric disorders. Findings of these first (and preliminary) epigenetic studies should be interpreted with caution because of small samples sizes, lack of replication, limitations in the etiologic validity of psychiatric diagnoses, and in accessibility of the regions and cell types of the brain at "appropriate" periods during life. Despite the sparse availability, the current evidence for epigenetic involvement in (particularly early) brain development, mental health, and psychiatric disorders appears very promising, and may be used in bringing together inherited and acquired risk factors into a neurodevelopmental etiological model of psychiatric disorders with epigenetics as a plausible key mediating mechanism. Given the dynamic nature of epigenetic regulation of gene expression and the potential reversibility of epigenetic modifications, future well-designed multidisciplinary and translational studies will be of key importance in order to identify new targets for prevention and therapeutic strategies.

Abbreviations

5-mC	5-methyl cytidine
AD	Alzheimer's disease
ADHD	Attention deficit hyperactivity disorder
AKT1	v-akt murine thymoma viral oncogene homolog 1

J. van Os
Department of Psychiatry and Psychology, School for Mental Health and Neuroscience, European Graduate School of Neuroscience (EURON), South Limburg Mental Health Research and Teaching Network (SEARCH), Maastricht University Medical Centre, PO BOX 616 (VIJV)6200, Maastricht, MD, The Netherlands

Department of Psychosis Studies, Institute of Psychiatry, King's College London, King's Health Partners, London, UK

APOE	Apolipoprotein E
APP	Amyloid precursor protein
BDNF	Brain derived neurotrophic factor
CB1	Cannabis-1
CDH1	Cadherin 1
Cdk5	Cyclin-dependent kinase 5
COMT	Catechol-O-methyltransferase
DC-MZ	Dichorionic monozygotic
DNMT	DNA methyl transferase
DRD2	Dopamine D2 receptor
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders 4th edition
DZ	Dizygotic
GABA	Gamma-aminobutyric acid
GAD	Glutamic-acid decarboxylase
H4K12	Histone H4 lysine 12
H4K16	Histone H4 lysine 16
HIST1H2AG	Histone cluster 1, H2ag
HIST1H2AH	Histone cluster 1, H2ah
HIST1H2BJ	Histone cluster 1, H2bj
HIST1H2BK	Histone cluster 1, H2bk
HIST1H4I	Histone cluster 1, H4i
HTERT	Telomerase reverse transcriptase
MAPT	Microtubule-associated protein tau
MC-MZ	Monochorionic monozygotic
MeCP2	Methyl CpG binding protein 2
MS	Multiple sclerosis
MTHFR	Methylenetetrahydrofolate reductase
MZ	Monozygotic
NPAS3	Neuronal PAS domain protein 3
NR3C1	Glucocorticoid receptor
OCM	One-carbon metabolism
PBCs	Pregnancy and birth complications
PCR	Polymerase chain reaction
PPIEL	Peptidylprolyl isomerae E-like
PSEN1	Presenilin 1
PTSD	Post-traumatic stress disorder
RELN	Reelin
SIRT3	Sirtuin 3
SMARCA5	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5
SMS	Spermine synthase
SNP	Single nucleotide polymorphism
THC	Δ^9-tetrahydrocannabinol
TRKB	Neurotrophic tyrosine kinase, receptor, type 2
UBE3A	Ubiquitin protein ligase E3A
WHO	World Health Organization

18.1 Introduction

Studies on the environmental and genetic epidemiology of psychiatric diseases have taken important steps forward in estimating heritability rates and identifying associations between a range of environmental and genetic factors and psychiatric phenotypes. Various environmental and genetic factors interact in complex manners throughout an individual's life to contribute to psychiatric disorders. Recent exciting developments in the field of epigenetics suggest that epigenetic mechanisms may mediate sustainable effects of environmental exposures. These ideas have generated great interest within many research disciplines, including psychiatric epidemiology. The "seductive allure of behavioral epigenetics" [1] has prompted psychiatric epidemiologists to focus on direct and indirect evidence for epigenetic involvement in mental health and normal behavior as well as in complex psychiatric disorders, in an attempt to elucidate the role of epigenetic mechanisms and possibly identify new strategies for prevention and treatment of psychiatric disorders [2]. Without attempting to provide a complete overview, this chapter addresses the current status of the literature on evidence indicative of involvement of epigenetic mechanisms in psychiatric disorders. The chapter starts with a summary of the evidence for a role of epigenetic mechanisms in development and aging of the brain and its functional abilities. Next, we exemplify that aberrant epigenetic mechanisms are linked to neuropsychiatric phenotypes by briefly describing psychiatric consequences of classical syndromes of genetic imprinting in humans. Thereafter, we summarize general epidemiologic findings that are indicative of epigenetic involvement in psychiatric disorders and review the more direct epidemiologic evidence (i.e., differential epigenetic profiles) for epigenetic involvement in the most prevalent and severe psychiatric illnesses. We will end the chapter by discussing current research challenges and future perspectives on epigenetic involvement in psychiatric disorders.

18.2 Epigenetic Mechanisms in Development and Aging of the Brain

The brain is the most complex organ of the human body and enables the individual to respond with appropriate behavior to a diverse range of internal and external stimuli, a process which requires dynamic adaptations in the molecular and cellular processes that represent the pathway from sensory perception to behavioral responses. Psychiatric disorders are characterized by mental symptoms and complaints that are accompanied with dysfunction of the human individual in interaction with the social world. The very complex molecular and cellular changes that underlie the development of functional abilities of the brain come together with sensitive time windows at which experiences and environmental exposures co-determine the developmental outcome. Pioneering work in the last decade has uncovered epigenetic regulation of gene transcription as fundamental for normal

development and functioning of the organism, especially in relation to appropriate responses to stimuli [3]. The various epigenetic mechanisms encompass DNA methylation, histone modifications, genetic imprinting, X-inactivation, and non-coding RNAs (described in detail elsewhere) which give rise to tissue- and cell-type specific profiles of gene expression and epigenetic marks during development. Besides robust changes in reprogramming of genomic methylation patterns in germline cells as well as in preimplantation embryos [4], it has become clear that dynamic changes occur in the patterns of DNA methylation, histone alterations, and expression of microRNAs throughout life and especially during development. Brain development involves cellular processes such as cellular proliferation, cellular differentiation, and maturation [5], but also myelination [6] and synaptic plasticity, and accumulating evidence indicates that these processes depend on appropriate epigenetic regulation [7]. Recent experimental animal studies have, for example, established that the functional abilities of memory formation, learning, motivation, and reward are all linked to epigenetic regulation of gene expression [8–10]. Genetic manipulation of Dnmt1 and Dnmt3a in mice has shown that long-term plasticity (which underlies learning and memory) in the mouse hippocampus depends critically on these major DNA methyl transferase enzymes [11]. Experience-driven developmental changes impact at different biological levels, such as membrane depolarization, calcium influx, and induction of transcription factors [12], and recent studies have discovered that sustainable effects of developmental exposures and experiences are mediated (and reflected) by epigenetic alterations [13], as discussed in more detail below. Evidence that humans with mutations in gene encoding methyl CpG-binding protein 2 (MeCP2) frequently show markedly decreased cognitive performances is furthermore in line with this notion. Epigenetic changes are proposed to have crucial impact on early life programming and thereby on neurodevelopmental disorders [14], but also affect age-related changes in the brain. Recent work reports that epigenetic markings are subject to change with advancing age [15, 16]. While earlier work suggested that aging was associated with a global loss of DNA methylation, more recent work indicates that age-related changes are also CpG-island dependent [15–17]. Investigating non-pathologic human tissues at 1,413 autosomal CpG loci associated with 773 genes, Christensen et al. observed highly significant CpG island–dependent correlations between age and methylation: loci in CpG islands gained methylation with age, loci not in CpG islands lost methylation with age [17]. Another large scale study using human brain tissue identified CpG loci, primarily CpG islands, with consistent positive correlation between DNA methylation, and chronological age [18]. In addition, recent findings on an aging mouse cohort kept under controlled environmental conditions throughout live showed that the level of the major *de novo* methylation enzyme DNA methyltransferase 3a (Dnmt3a), which catalyzes DNA methylation, increased with age in the hippocampus [19], and correlated with age-related increase in levels of 5-methyl cytidine (5-mC). The same study showed that caloric restriction, which increases lifespan and prevents age-related alterations and pathology in various animal species [20–23], was able to prevent these age-related changes in hippocampal levels of Dnmt3a [19] and 5-mC [251]. The speculation of a causal involvement of

epigenetic mechanisms in age-related decline of functional abilities of the brain [24] is in line with (1) findings that age-related memory disturbances in mice are associated with altered chromatin plasticity (in particular with dysregulation of H4K12ac) in the hippocampus [25], and (2) the link between sirtuins, i.e., molecules which deacetylase H4K16ac in a Nicotinamide adenine dinucleotide-dependent manner, and life span [26]. Thus, epigenetic mechanisms appear to be fundamentally involved in the neurobiological processes that govern development [14] and aging of the brain [16], and may be involved in the formation of an individuals' functional abilities and personality characteristics, as well as in neurodevelopmental and neurodegenerative trajectories of psychopathology [1, 27].

18.3 Psychiatric Phenotypes in Disorders of Genetic Imprinting

The crucial role of genetic imprinting in brain function is exemplified by the clinical pictures of two syndromes that are caused by aberrations in genetic imprinting: Angelman syndrome and Prader-Willi syndrome [28]. Angelman syndrome results from loss of expression of the maternal gene *UBE3A* at chromosome 15, by either deletion or other genetic abnormalities [29]. The maternal gene is normally expressed while the paternal gene is normally silenced. The characteristic clinical picture of Angelman syndrome comprises neurodevelopmental disabilities, motor abnormalities, seizures, and speech deficits, resembling the clinical picture in patients with deficiencies in 5,10-methylenetetrahydrofolate reductase (*MTHFR*) or in patients with mutations in methyl-CpG binding protein 2 (*MeCP2*) [30]. Prader-Willi syndrome results from the loss of paternal expression of genes (on the same region on chromosome 15 as Angelman syndrome) [28]. Prader-Willi syndrome is clinically characterized by the psychopathological features of mental retardation, obsessive-compulsive symptoms, eating problems (hyperphagia), hypersomnia, and neurodevelopmental delay of motor skills [31]. Thus, aberrations of genetic imprinting in the same region on chromosome 15 can cause two syndromes with overt neuropsychiatric illness phenotypes. The epigenetic changes underlying these syndromes are discussed in more detail in Chap. 9.

18.4 Twin Discordance

Proband-wise concordance rates for psychiatric disorders in classical twin studies are generally spanning a wide range, in the order of 20–80%, while many psychiatric disorders also have clear links to aberrant neurodevelopment, e.g., autism spectrum disorders, ADHD, and schizophrenia, that all show substantial heritabilities. The standard assumptions that greater disease concordance rates in monozygotic (MZ) versus dizygotic (DZ) twins indicate genetic contribution, and concordance rates

well below 100% in MZ twins indicate an environmental component, have been challenged by evidence indicating that (1) gene-environment interactions may contribute to both the "genetic" and "environmental" components [32], and (2) possible effects of stochastic factors exist on biological processes, such as regulation of gene expression, throughout life [2, 33].

While most MZ twins show phenotypic similarity and MZ twin pairs are expected to be concordant for congenital malformations, chromosomal abnormalities, and Mendelian disorders, phenotypic discordance in MZ twins does occur frequently and may arise from various sources such as (1) chromosomal and monogenetic variations, (2) differences in environmental exposures in the intrauterine (differential timing of the twinning process, a differential number of cells allocated to each twin, and differential placental vascularization), perinatal (such as hypoxia) and postnatal environment, as well as (3) epigenetic differences [34]. An indication of epigenetic mediation of environmental exposures during life is provided by whole-genome and locus-specific methylation analyses of DNA from lymphocytes of MZ twins (15 male twin pairs and 25 female twin pairs) showing that approximately one-third of these MZ twins harbored significant epigenetic differences in DNA methylation and histone modification that were more distinct in MZ twins who were older, had different lifestyles, and had spent more of their lives apart [35].

Recent methylation microarray analyses of DNA in lymphocytes and buccal mucosa of 114 MZ twins and 80 dizygotic (DZ) twins confirmed the presence of substantial differences in DNA methylation profiles in MZ twins, while also showing epigenetic metastability of 6,000 unique genomic regions in MZ twins, and greater epigenetic similarity in MZ co-twins than in DZ co-twins [36].

Differences in methylation profiles in buccal mucosa of MZ twins have furthermore been reported for CpG sites in a number of specific candidate genes for psychiatric disorders such as the dopamine D2 receptor (*DRD2*) gene [33] and the catechol-O-methyltransferase (*COMT*) gene [37]. In addition, differential DNA methylation profiles in buccal mucosa have been observed in MZ twins discordant for bipolar disorder (BD) [38] and the behavioral phenotype of risk-taking behavior [39]. A recent methylation microarray analysis of MZ twins discordant for BD observed increased methylation in the affected twins upstream of the spermine synthase gene (SMS) and lower methylation upstream of the peptidylprolyl isomerase E-like gene (PPIEL) [38]. However, very detailed analyses on MZ twin pairs discordant for multiple sclerosis (MS) [40] indicated that the puzzle of explaining MZ twin discordance is far from being solved. Analyses of the genome sequences of an MZ twin pair discordant for MS, and messenger RNA transcriptome and epigenome sequences of lymphocytes from three MZ twin pairs discordant for MS failed to detect reproducible differences in approximately 3.6 million single nucleotide polymorphisms (SNPs) and approximately 0.2 million insertion-deletion polymorphisms between the co-twins [40]. The authors furthermore did not observe reproducible differences between siblings of the three twin pairs in HLA haplotypes, MS-susceptibility SNPs, copy number variations, mRNA and genomic SNP, and insertion-deletion genotypes, or the expression of approximately 19,000 genes. Although the detailed analyses did detect 2–176 differences in the methylation of

approximately two million CpG dinucleotides between siblings of the three twin pairs, it is unlikely that these methylation differences can fully explain disease discordance [40].

Thus, preliminary evidence indicates epigenetic involvement in MZ twin discordance, although the complete puzzle of explaining MZ twin discordance in the various (neuro)psychiatric disorders is far from being solved. Future twin research will be of crucial relevance for the formulation and testing of new etiological models of complex psychiatric traits and disorders integrating the contributions of genetic components, environmental exposures, gene-environment interactions, stochastic factors, and epigenetics [2, 35, 41–43].

18.5 Parent-of-Origin Effect and Paternal Age

Parent-of-origin effects reflect a differential proportion of paternal or maternal disease-causing transmission to offspring. Thus, the risk-increasing effects of alleles for certain complex diseases may depend on their parent of origin. It has been proposed that parent-of-origin effects are either based on mutagenesis, causing de novo spontaneous mutations which would then propagate and accumulate in successive generations of sperm-producing cells, or on genomic imprinting, an epigenetic mechanism in which differential epigenetic modification of genes occurs depending on their parental origin [44]. Parent-of-origin effects have, for example, been described in attention deficit hyperactivity disorder (ADHD), autism, psychosis, and late onset cases of Alzheimer's disease (AD) [45–50]. It has furthermore been described that paternal and maternal ages (at the time point of conception) are risk factors for various psychiatric disorders with neurodevelopmental origins such as schizophrenia and autism. Paternal age effects have been proposed to be mediated by epigenetic mechanisms [51]. In autism, a 10-year increase in paternal age (independent of maternal age) is associated with a 22% increased risk for autism while a 10-year increase in maternal age (independent of paternal) age has been associated with a 38% increased risk for developing the disorder [52], which is in line with other findings reporting associations between parental ages and autism [53–55]. A recent meta-analysis on the association between paternal age and schizophrenia reported that both advanced paternal age (>/=30) and younger paternal age (<25) may increase the risk of schizophrenia, with further analyses indicating that younger paternal age may be associated with particularly an increased risk in male but not female offspring [56]. Given these effects of paternal (and maternal) age on risk of several psychiatric disorders, one could speculate that paternal and maternal age influence general functional abilities of the offspring's brain rather than inducing abnormalities specific for a certain disorder. An interesting recent study on 33,437 children drawn from the US Collaborative Perinatal Project reported that offspring of older fathers were impaired in tests of neurocognitive ability (at an age of 7 years) while advanced maternal age was associated with better neurocognitive abilities [57]. The pattern of childhood behavioral problems in the same population

was also differentially affected by advanced paternal age as compared with advanced maternal age; while advanced paternal age was associated with externalizing behaviors in the offspring but not with internalizing behavioral outcomes, advanced maternal age was significantly protective against externalizing behavioral outcomes, but associated with an increased risk of internalizing behavioral problems [58]. Evidence that, at least for some disorders, both high and low paternal or maternal age were associated with increased risk of psychopathology, e.g., in autism [53] and schizophrenia [56], is in line with the concept of the existence of an optimal (epigenetic-regulated) window of paternal and maternal ages at conception. Thus, parent-of-origin effects and effects of paternal (and possibly also maternal age) on risk of psychopathology in offspring point to a possible role for genomic imprinting in the etiology of psychiatric disorders.

18.6 Sex-Differences in Psychiatric Disorders

The onset, course, and phenomenology of psychiatric disorders often show sexual differentiation [59, 60]. Female sex has consistently been associated with increased risk of depressive disorders [61] and anxiety disorders such as panic disorder, post-traumatic stress disorder (PTSD), and social phobia [62], while male sex is associated with ADHD [63], autism spectrum disorders [64, 65], and substance abuse disorders [66]. Sexual differentiation is furthermore apparent in age of onset, phenomenology, and developmental trajectories toward psychopathology [67, 68]. Various epigenetic mechanisms have been proposed to underlie sexual differentiation in psychiatry. Besides differential exposures to environmental risk factors with sustainable effects on the individual's phenotype, sexual differentiation in many features of psychiatric disorders may be the result of sex hormone-induced differences in the epigenetic status of key genes impacting during sensitive time periods during development [69]. For example, methylation of the promoter of the estrogen receptor-alpha gene has been reported to increase during development [70].

Another epigenetic mechanism giving rise to sexual differentiation is X-chromosome inactivation. X-inactivation is the irreversible epigenetic process that involves silencing of one of the two X-chromosomes in female individuals during early development [71]. X-inactivation is estimated to take place just before the morula stage when the human embryo contains less than 16 cells [72, 73]. Previous studies have shown that the timing of X-inactivation is associated with the timing of chorionic twinning [72]: dichorionic monozygotic (DC-MZ) twinning is estimated to take place at, or before, the morula stage (so, just at or before X-inactivation) and the monochorionic monozygotic (MC-MZ) twinning event occurs after the morula stage (and consequently after X-inactivation) [72]. Consistent with the notion that individuals who are members of DC-MZ twins are separately subject to X-inactivation, patterns of X-inactivation are more identical in MC-MZ twins than in DC-MZ twins [72, 74]. Thus, chorionicity may be used as a proxy-measure of X-inactivation. While it has been proposed that X-inactivation may affect functional

abilities and risk of psychiatric disorders particularly in females [75], only a very limited number of studies in psychiatry have directly measured variations in X-inactivation, although various epidemiological studies have used proxy-measures of X-inactivation. For example dichorionicity in twins and female sex have been considered as a proxy measure of X-inactivation. In contrast to previous twin studies using sex and chorion type in isolation as proxy-measures for X-inactivation [75], a recent study by Peerbooms et al. combined information on sex and chorion type in an improved proxy measure and found no association between this proxy measure for X inactivation and variations in intelligence and behavioral problems in children with a mean age of 10 years, thus indicating no major role for X-inactivation in variations of these traits at this age [76]. However, a recent study by Rosa et al. using more direct measures of X-inactivation by comparing methylation at CpG islands of X-linked housekeeping genes on the active and inactive X-chromosomes in blood and buccal epithelial cells (as an index of X-chromosome inactivation) found evidence that was suggestive for a role of X-inactivation in BD. The index of X-inactivation showed a borderline statistically significant difference for discordant BD twin pairs as compared to controls [77]. Nevertheless, the role of X-inactivation in behavioral traits and psychiatric disorders has only received very little research interest thus far and requires further exploration.

18.7 Enduring Influences of Environmental Exposures

A wealth of epidemiologic evidence indicates that environmental exposures influence susceptibility to disease in later life [14, 42]. A consistent line of evidence has connected exposures to nutritional factors, childhood trauma, minority group position, drugs of abuse, and pre- and perinatal complications to an increased risk of psychiatric disorders such as depression and schizophrenia later in life, especially in genetically vulnerable people [32, 41, 68, 78, 79]. It has been argued previously that the contributions of genetic and environmental factors are unlikely to be independent and that models taking into account gene-environmental interactions may be much more plausible [80–82]. While epigenetic mechanisms have been proposed as a prime candidate for mediating some of these environmental effects, primarily on the basis of indirect evidence and findings in experimental animal studies [41, 42], direct evidence in humans remains very sparse. Based upon the currently available evidence from both animal and human research, we focus here on examples of environmental factors associated with psychiatric disorders, for which evidence suggests epigenetic involvement.

18.7.1 Pregnancy and Birth Complications

Population-based studies have suggested an association between psychiatric disorders and a wide variety of prenatal exposures, including maternal stress [83–85], maternal nutritional deficiency [86, 87], prenatal iron deficiency [88], maternal serum lead levels [89], maternal serum homocysteine levels [90], rhesus incompatibility [91],

low and high neonatal vitamin D [92], prenatal toxoplasmosis [93, 94], specific viral infections [95], bacterial infections [96], maternal pyelonephritis [97], maternal hypertension [98], and maternal use of analgesics [99] and diuretics [98] during pregnancy. As many of these exposures have been found to associate (or induce) epigenetic alterations in cell culture or animal studies, epigenetic mechanisms have been proposed to mediate at least some of these effects [41, 100–103].

There are, however, few true (i.e., corresponding in trimester timing, exposure definition, subgroup-only effects) replications, and credible non-replications, on the associations between pregnancy and birth complications (PBCs) and psychiatric disorders [104–109], such that no definitive conclusions about association (nor mediation) can be drawn at this stage. It is important to realize that while causal relationships and underlying mechanisms can be examined easier in animal research, inferences of animal findings to the human situation are limited [110, 111], which is especially relevant for the domain of psychiatric disorders. Nevertheless, it is attractive to hypothesize that the wide variety of exposures may reflect a single underlying epigenetic mechanism associated with subtle developmental perturbations that increase risk for psychotic outcomes in interaction with genetic risk variants [97]. A recent study suggested that individuals from the Dutch Hunger winter who were exposed to famine prenatally showed hypomethylation at the imprinted *IGF2* gene when compared to their unexposed same-sex siblings six decades after the period of severe famine [112]; an association that was specific for peri-conceptional exposure to famine, thus suggesting crucial relevance of epigenetic mechanisms in very early mammalian development [112]. As famine may also evoke psychological stress (or be associated with other co-variates), these effects may not be solely attributable to impaired nutritional uptake during pregnancy. Several nutritional factors such as deficiencies of vitamin A or D, protein content, essential fatty acids, and folate have been proposed to mediate the increased risk of psychiatric disorders in offspring subjected to famine during fetal development [113]. A large birth cohort study with data on schizophrenia in adulthood indicated that elevated homocysteine levels in the third trimester of pregnancy are associated with an increased risk of schizophrenia in the offspring [90]. A study in which birth interval was used as a proxy of folate levels during pregnancy (as postpartum restoration to normal maternal folate values may take up to 1 year after pregnancy) further suggested an association between folate and schizophrenia [114]. It cannot be excluded that reported findings on the wide range of prenatal exposures in part represent the effects of publication bias, multiple testing in small samples, or genetic confounding occasioned by maternal behaviors associated with genetic risk for broadly defined psychiatric disorders (given that most psychiatric disorders show a familial relationship with schizophrenia [115]). On the other hand, inconsistencies across studies and lack of reproducibility may also be due to differential confounding.

Eight prospective, population-based studies on PBCs and schizophrenia were available for meta-analysis [116]. This meta-analysis reported a wide range of PBCs that were associated with a diagnosis of schizophrenia including (1) complications of pregnancy (bleeding, diabetes, rhesus variables reflecting rhesus incompatibility, rhesus-negative mother, rhesus antibodies, pre-eclampsia); (2) abnormal fetal growth and development (low birth weight, congenital malformations, reduced head

circumference); and (3) complications of delivery (uterine atony, asphyxia, emergency Caesarean section). Pooled estimates of effect sizes were generally around 2. These studies suggest an etiological signal associated with PBCs, confirmed by later studies on this topic [117].To establish a more robust link between an exposure and a complex outcome usually require replication of association across a range of studies using a variety of different designs and approaches [118]. Prospective population-based studies in Denmark and Sweden of groups with high rates of PBCs, such as twins, have been inconclusive [119–121], and studies focusing on subclinical psychotic symptoms and PBCs similarly are not consistent [122–124]. Familial risk may be associated with greater sensitivity for the risk-increasing effect of PBCs on schizophrenia and on brain structural alterations associated with schizophrenia [125, 126], which may be linked to effects of hypoxia [127, 128], that may be epigenetically mediated [129].

18.7.2 Rearing Environment and Childhood Abuse

Adoption studies provide evidence for an association between rearing environment, childhood stress, and an increased risk of psychopathology at a later age. For example, adoptees with a positive family history for psychosis who had been brought up in dysfunctional adoptive family environment displayed an increased risk of psychiatric disorders [130–133]. Conversely, a positive rearing environment may decrease the risk of psychotic disorder later in life, as it has been shown that high-risk children with positive parental relationships have a lower risk for developing schizophrenia [134]. Studies on the effects of an adverse early psychosocial environment on later psychiatric disorder have been re-invigorated by recent prospective epidemiological findings showing that victims of bullying at ages 8 and/or 10 years have a 2-fold risk of psychotic symptoms at age 12 with an even more elevated risk when victimization is chronic or severe [135], and recent surveys in large population samples indicate that childhood adversities (especially when associated with maladaptive family functioning) account for approximately 30% of all psychiatric disorders [136], and are associated with earlier onset of psychiatric disorders, persistence of mood and anxiety disorders [137, 138] and with suicidal behavior [139]. No studies have yet examined the link between childhood stress, epigenetic changes, and the onset of psychiatric disorders in humans. Recent experimental animal research, however, has suggested that the psychosocial environment and stress can mediate changes in gene expression during key developmental periods through epigenetic mechanism with long lasting effects on behavior. Meaney and colleagues observed that postnatal maternal care in rats was apparently associated with epigenetic modification of a transcription factor binding site in the promoter region of the glucocorticoid receptor gene (*NR3C1*) which in turn directly alters gene expression and behavioral phenotypes in the offspring which persist into adult life although experiments lacked proper reproducibility studies with updated techniques [140]. In subsequent experiments methyl supplementation during the same early postnatal period seemed

to reverse the epigenetic modification induced by maternal care, with related gene expression changes and behavioral phenotypes in adult offspring [141]. Subsequent transcriptomic studies by the same group identified over 900 genes in the rodent hippocampus that are stably regulated by maternal care [142], suggesting an even more widespread effect of the early social environment on gene expression through the life course. However, these studies still await robust replication. A recent study in humans suggested epigenetic differences in a homologous *NR3C1* promoter region, comparing DNA methylation in postmortem hippocampus samples obtained from suicide victims with a history of childhood abuse to that seen in samples from either suicide victims with no childhood abuse or controls. In line with the animal findings, abused suicide victims were reported to have increased methylation of the *NR3C1* promoter with concomitant changes in mRNA [143]. These preliminary findings however await replication.

18.7.3 Drugs of Abuse

In recent years, evidence from epidemiological studies and meta-analyses has established that cannabis and other drugs of abuse can be considered risk factors for later psychotic symptoms or psychotic disorders [144–146]. The age (or developmental stage) at which individuals start using cannabis influences this association [147]. Further evidence suggests that gene-environment interactions are likely implicated in the association between cannabis and psychosis [145, 148–152]. For example, the psychotomimetic effect of cannabis is much greater in siblings of patients with a psychotic disorder, who are at increased genetic risk to develop psychotic disorder than in controls [153]. Cannabis use increased the risk for developing psychotic symptoms and schizophreniform disorder only in carriers with the valine158 allele in the gene encoding catechol-*O*-methyltransferase (*COMT*) [148], while the risk-increasing effects of cannabis on psychosis expression were moderated by genetic variation in *AKT1* in another study [150]. The primary psychoactive component of cannabis is Δ^9-tetrahydrocannabinol (THC), which is thought to exert its effects through cannabis-1 (CB1) receptor-mediated signaling [154]. Administration of THC or cannabis elicits long-term molecular and cellular changes in the brains of mice and humans [155–158], as well as on electrophysiological and biochemical measures of neuronal signaling in brain structures such as nucleus accumbens and hippocampus [159] with differential effects by duration of exposure and timing during development [159–162]. The abuse of psychostimulants, such as cocaine and amphetamine, has consistently been associated with major psychotic disorders and substance use disorders. Many drugs of abuse may act via the common mechanism of sensitization and sensitization may play a key role in psychosis. In humans and animals, repeated exposure to psychostimulants induces a sensitized state [163], i.e., an enhanced response to subsequent low dose challenges [164]. Sensitized animals (to different drugs) share a number of neurobiological changes such as long-term alterations of mesolimbic and prefrontal dopaminergic neurotransmission

while also expressing behavioral abnormalities resembling positive symptoms and enduring cognitive deficits such as seen in schizophrenia [163]. In addition, acute exposure to drugs of abuse provokes transient increases in cFos and other members of the Fos family of transcription factors in the striatum [165]. The majority of these *Fos* factors become desensitized during chronic abuse, with the exception of ΔFOSB, which itself stimulates the expression of other genes, such as *Cdk5*. [166, 167] ΔFOSB accumulates in the nucleus accumbens and dorsal striatum – two brain regions important for reward and locomotor drug responses – and remains detectable for several weeks after cessation of drug administration, suggesting a role in the onset of addiction [167]. In mouse models of cocaine response, the upregulation of cFos and ΔFOSB upon acute administration is accompanied by rapid and transient H4 hyperacetylation at the *cFos* and *FosB* gene promoters [168]. Chronic exposure, however, does not result in H4 hyperacetylation but is associated with H3 hyperacetylation of the *FosB* promoter, with no effect on the *cFos* promoter. H3 hyperacetylation of two genes associated with chronic, but not acute, drug use – *Cdk5* and *BDNF* (brain derived neutrophic factor) – was also detected after chronic cocaine administration, and persisted even 1 week after treatment ceased [168].

These findings implicate epigenetic mechanisms in the development of addiction, and particularly in the immediate and sustained behavioral effects of cocaine use [168]. Human studies found enduring changes in gene expression in subjects exposed to psychostimulants. For example, prolonged exposure to amphetamine is associated with long-term reductions of dopamine transporter density in the brain as measured by in vivo imaging studies [169]. Exposure to cannabis, cocaine, and phencyclidine have long lasting effects on intracerebral gene expression, and recent evidence suggests that in fact these drugs may all affect common neurobiological pathways. In a recent post mortem study using human brain tissue samples, exposures to cannabis, cocaine, and phencyclidine were associated with many of these transcriptional changes in the brain [170]. Hierarchical clustering of these transcripts indicated that genes from the calmodulin signaling cluster were predominantly affected in the brains of abusers, which may be of particular importance given that enhanced dopamine release due to sensitization depends on calmodulin signaling [171, 172]. Thus, chronic exposure to drugs may induce a sensitized state with enduring intracerebral changes in gene expression (particularly in dopaminergic neurotransmission), with evidence – at least from animal studies – suggesting a crucial, mediating role for the epigenetic machinery.

18.7.4 Minority Group Position and Social Defeat

Meta-analytic work shows consistency for the association between psychotic syndrome and minority group position across a wide range of approaches, endpoints, settings and cultural group definitions [173, 174], and after adjustment for a range of confounders. The possibility of cultural bias in diagnosis and selective migration has been examined but not found to have a major impact on the association. The association with minority group position is observed in both first and second

generation migrants [173, 174], as well as in minority groups without recent migration [175], indicating that pre-migration factors or migration itself are unlikely to mediate effects. Studies in four different countries have shown that the effect of minority ethnic group on psychotic syndrome depends on the ethnic density of the area the person is living in: the greater the proportion of the own ethnic group in the area, the lower the risk for psychotic disorder [176, 177]. These findings suggest that it is not ethnic group per se that increases risk, but rather the degree to which one occupies a minority position, or stands out in relation to the wider social environment. Additional research suggests that effects associated with minority group position may be mediated by chronic social adversity and discrimination [178], resulting in social marginalization or a state of social "defeat" (chronic experience of an inferior position or social exclusion [179]). Although epigenetic consequences of chronic social "defeat" has not been shown in humans yet, animal research has shown that chronic social defeat stress (by daily experience of defeat by a bigger and more aggressive mouse for 10 consecutive days) induces gene expression changes, particularly of *BDNF*, via a range of epigenetic mechanisms including histone tail modifications and DNA methylation. For example, chronic exposure to social defeat stress in mice significantly downregulated mRNA levels of histone deacetylase-5 in the nucleus accumbens [180]. Chronic defeat stress in mice also induced enduring downregulation of BDNF transcripts and increased histone methylation [181]. Interestingly, chronic treatment with the antidepressant imipramine can reverse downregulation of BDNF transcripts while increasing histone acetylation at the corresponding promoters [181]. Thus, animal research suggests that social stress may be an epigenetically mediated, environmental factor that underlies proxy risk factors such as ethnicity and migration in psychiatric disorders.

18.7.5 Synergism in Environmental Exposures

It has been proposed that psychiatric disorders arise slowly from subclinical symptoms that become abnormally persistent when synergistically combined with various environmental exposures during development that may impact on behavioral and neurotransmitter sensitization [78, 182, 183]. Although various experimental animal studies have indicated synergistic effects of environmental exposures [184–186], data from epidemiologic studies remain sparse. For example, heritable risk for depression and the combined effects of several environmental exposures (over the course of development measured as lower birth weight for gestational age, childhood adversity, and negative life events in adulthood) have recently been reported to have a synergistic impact on the psychiatric phenotype of stress-sensitivity (that is proposed to underlie depressive symptomatology) [78]. While it remains to be established whether epigenetic alterations mediate these effects, future prospective epidemiologic studies with (epi) genetically sensitive designs may further focus on synergistic effects of environmental exposures during development with inherited sensitivity and epigenetic profiles.

18.8 Variants in Epigenetic-Relevant Genes and Psychiatric Disorders

Genetic studies have indicated that variations in genes with crucial relevance for epigenetic mechanisms may increase the risk for developing psychiatric disorders.

18.8.1 MTHFR

The gene encoding 5,10-methylenetetrahydrofolate reductase (MTHFR) is an epigenetically relevant gene as it encodes a crucial enzyme involved in the folate-mediated one-carbon metabolism (OCM), which is essential for purine and thymidylate biosynthesis, methylation of DNA and amino acids, and necessary for reactions forming neurotransmitters [187]. Dysfunction of one-carbon metabolism has been linked to a range of disorders including neural tube defects [188, 189], autism [190], leukemia [191, 192], dementia [193, 194], colorectal cancer [195, 196], cardiovascular disease [197], and congenital abnormalities [198, 199]. Genetic studies on associations between genetic variants in *MTHFR* and various major psychiatric disorders such as schizophrenia, BD, and unipolar depressive disorder have yielded largely inconclusive and often mixed results [200–206]. Given the essential role of *MTHFR* in brain function and neurodevelopment [207, 208], and the fact that family and twin studies have established considerable shared genetic variance between psychiatric disorders [209–212], a recent meta-analysis (on a total of more than 29,000 subjects) tested whether genetic variation in *MTHFR* contributes to the shared genetic vulnerability of schizophrenia, BD, and unipolar depressive disorder [213]. This meta-analysis showed that *MTHFR* C677T was significantly associated with the combined group of schizophrenia, BD, and unipolar depressive disorder (odds ratio = 1.26 for TT versus CC genotype carriers; confidence interval 1.09–1.46) without evidence of modifying effects of psychiatric diagnosis, sex, ethnic group, or year of publication, thus providing evidence for shared genetic vulnerability for SZ, BPD, and unipolar depressive disorder mediated by *MTHFR* 677TT genotype [213].

18.8.2 MeCP2

The gene encoding for MeCP2 is another gene crucially involved in epigenetics. MeCP2 protein selectively binds CpG dinucleotides in the genome and mediates transcriptional repression through interaction with histone deacetylase [214]. Besides being the major cause of Rett syndrome (a rare but fulminant neurodevelopmental disorder in young girls) [215], mutations in this X-linked gene have been found to be associated with a broad array of other neurodevelopmental disorders in

males and females, including X-linked mental retardation [216, 217], severe neonatal encephalopathy, Angelman's syndrome, and autism [214]. Thus, these findings show that variants in the key epigenetic regulator gene MeCP2 impact crucially on brain development and thereby on risk of psychiatric disorders.

18.8.3 Histone Genes

A recent large meta-analysis suggested that common variants located at chromosome 6p22.1 in a cluster of the histone genes HIST1H2BJ, HIST1H2AG, HIST1H2BK, HIST1H4I, and HIST1H2AH are associated with an increased risk of schizophrenia [218]. Although the observed associations may be connected to one or several genes, intergenic elements, or longer haplotypes that include susceptibility alleles in many genes, the strongest association was observed for a region in and near a cluster of histone protein genes, making variants in histone coding genes prime candidates for further analysis, at least in schizophrenia.

18.9 Aberrant Epigenetic Profiles in Psychiatric Disorders

18.9.1 Depression, PTSD, and Suicide

Methylation microarray analyses of whole blood-derived, bisulfite-converted DNA indicated differential methylation profiles in 33 individuals who reported a lifetime history of depression, as compared to 67 non-depressed adults, suggested patterns of increased methylation in genes relevant for brain development and tryptophan metabolism and patterns of decreased methylation in other biological processes [219]. In PTSD, a similar microarray approach indicated differential methylation profiles in 23 PTSD-affected as compared to 77 PTSD-non-affected individuals; genes implicated in immune function were uniquely unmethylated in the affected individuals [220]. In post mortem obtained brain tissue, H3K27me differed in the *TRKB* promoter locus in suicide completers ($n = 20$) as compared to control subjects ($n = 10$) in the orbital frontal cortex but not in cerebellum [221].

18.9.2 Schizophrenia and Bipolar Disorder

Early studies primarily focused on differential epigenetic marks in specific candidate genes in post mortem brain tissue. Such studies reported DNA methylation differences associated with schizophrenia in catechol-*O*-methyltransferase (*COMT*) [222] and reelin (*RELN*) using methylation-specific PCR [223], although studies using

full quantitative methylation profiling methods did not confirm these findings [224–226]. Post mortem analyses of cortical GABA-ergic neurons in schizophrenia have shown increased levels of Dnmt1, which was associated with altered expression of reelin or GAD67 [223, 227]. As discussed earlier, MZ twins discordant for BD had differential DNA methylation profiles in regions upstream of the spermine synthase gene (*SMS*) and upstream of the peptidylprolyl isomerase E-like gene (*PPIEL*) [228]. While DNA methylation upstream of *SMS* did not affect expression of the gene, a strong inverse correlation between *PPIEL* gene expression and DNA methylation was observed. Methylation microarray analysis of frontal cortex tissue from patients with schizophrenia and BD revealed DNA methylation differences of numerous loci, including several involved in biological processes such as glutamatergic and GABA-ergic neurotransmission, brain development, and other processes functionally linked to disease etiology [225].

18.9.3 Alzheimer's Disease

Decreased global levels of DNA methylation were observed in the entorhinal cortex of AD patients using immunohistochemical analyses [229]. Several human post mortem brain studies analyzing the methylation status of promoter regions of candidate genes of AD have yielded mixed and inconsistent findings. Analysis of the *APP* promoter from temporal cortex failed to show a difference in methylation status of the promoter region between control and Alzheimer's disease patients [230]. DNA methylation was decreased in the promoter region of the tau protein gene in the parietal cortex, but its transcription was downregulated [231]. In a postmortem analysis of the frontal cortex and hippocampus, no significant differences between Alzheimer's disease patients and age-matched controls were found in the methylation patterns of the promoters of *MAPT*, *PSEN1*, and *APP* [232]. Another study demonstrated increased methylation within the promoter regions of *APOE* and *MTHFR* in Alzheimer's disease patients when compared to controls, in both post mortem prefrontal cortex tissue and peripheral lymphocytes [233]. No differences in methylation patterns of DNA from blood lymphocytes of AD patients versus controls were observed in promoter regions of the *SIRT3* (sirtuin 3), *SMARCA5* (SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a 5), and *CDH1* (cadherin 1) [234]. The methylation status of *HTERT* (telomerase reverse transcriptase) in DNA of blood lymphocytes, however, differed between Alzheimer patients and age-matched controls [234]. Epigenetic analyses in post mortem brain samples furthermore showed that the epigenetic distance from the norm (the median methylation of the control individuals) increased with age, and was higher in people with AD than in healthy controls [233]. The general age-dependent epigenetic drift was also found in lymphocytes of AD patients, although the differences were smaller when compared to brain tissue [233]. Thus, clear and consistent evidence for epigenetic involvement in AD is currently lacking while adequately powered epigenetic analyses are warranted.

18.10 Summary and Perspectives

It is easy to speculate about the role of epigenetic processes in mediating susceptibility to psychiatric disorders, but investigating these modifications at the combined epidemiologic and molecular level is far from straightforward. The first generation studies on epigenetic differences in psychiatric disorders are characterized by a cross-sectional design investigating DNA methylation differences of candidate genes in small numbers of patients ($n = 1–50$) and age- and sex-matched "super-controls" (i.e., healthy controls without any psychopathology). DNA was derived from a range of tissue types (whole blood, blood lymphocytes, or homogenates of various brain regions) and processed with variable laboratory procedures, i.e., bisulfite conversion, methylated DNA immunoprecipitation, and microarray techniques. As these studies harbor various limitations, such as risk of false-positive findings (e.g., due to small sample size and the comparison with supernormal controls), lack of prospective investigations, selection bias, treatment effects, disease effects, lack of primary tissue of interest (the brain), and lack of replication, results from these first generation studies should be interpreted with caution [235]. Nonetheless, the findings do suggest possible differential epigenetic profiles in psychiatric disorders and generate hypotheses for future studies.

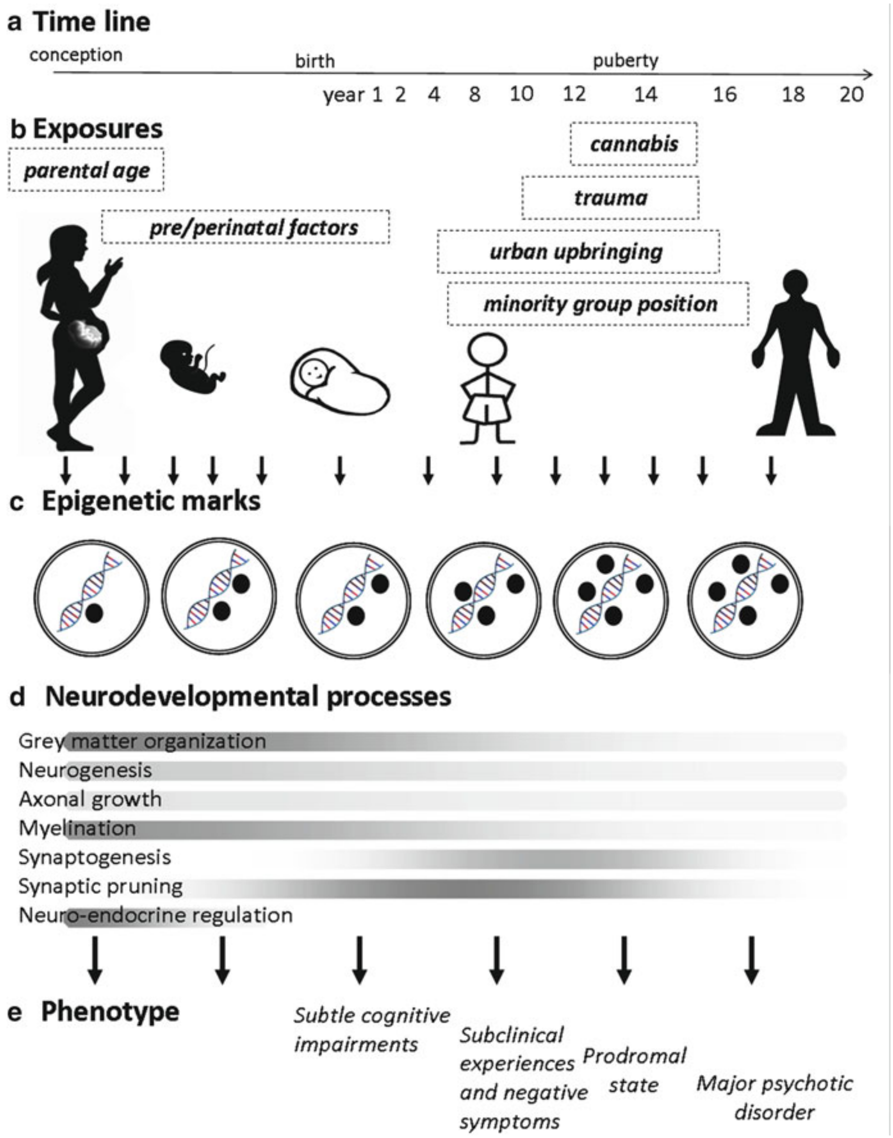
Difficulties in establishing etiological classifications of psychiatric disorders and phenotypes have defined intermediate phenotypes, dimensions of psychopathology, or sub-classes of conventional disease categories which may be more proximal to the actual neurobiological causal factors, and thus more suitable for epigenetic epidemiologic research.

Analyses of DSM-IV classifications of diagnoses in the WHO mental health surveys have established that the temporary presence of any distinct psychiatric disorder predicts subsequent onset of other psychiatric disorders, and that, therefore, "comorbidity" of psychiatric disorders is the rule rather than the exception [236]. In line with this, a diagnosis of schizophrenia has been found to be highly predictive of virtually all other Axis I and Axis II psychiatric disorders in the same person [237]. These findings suggest that dimensional approaches may be useful. Moreover, cognitive alterations, psychosis (hallucinations and delusions), and affective dysregulation that are observed in psychotic disorders (schizophrenia, BD) are also prevalent in common mental disorders (anxiety disorders, depression) [238, 239]. The relative non-specificity of psychiatric disorders extends to familial aggregation and to risk and protective factors. A family history of schizophrenia is associated with the strongest relative risk; however almost any psychiatric disorder in first-degree relatives is associated with an increased risk for schizophrenia. Similarly, nearly 30% of schizophrenia in the population can be attributed to psychiatric family history in general, compared to 6% that is attributable to a family history of schizophrenia specifically [115]. Many psychiatric syndromes/disorders share common risk and protective factors. Common genetic factors, such as polymorphisms in *MTHFR* [213] and *NPAS3* [240], as well as various environmental exposures such as childhood adversity [136] have substantial effects that transcend traditional clinical

diagnostic boundaries. These data suggest that shared genetic, epigenetic, and environmental factors contributed to neurodevelopmental alterations resulting in broad dimensions of mental ill-health [241]. While integrative investigations on genetic data, epigenetic data, and environmental data thus seem pertinent, analyzing these effects in integrative statistical models remains a tremendous challenge [242, 243].

The first epigenetic studies in psychiatric disorders have increased awareness of numerous challenges at the technical level as well. The limited accessibility to high-quality human brain tissue from well-phenotyped patients, in combination with the cell-type-specific and temporal-specific nature of epigenetic programming, poses great challenges for epigenetic studies. Whether more accessible tissue sources (like lymphocytes) for epigenetic profiling will reflect similar pertinent epigenetic changes as the respective tissue of interest (i.e., brain) remains to be explored. There is increasing evidence from other disorders that many epimutations are not limited to the affected tissue or cell type, but can also be detected in other tissues. Furthermore, there is great need for studies that establish the "normal" epigenome, especially for tissues such as brain, and patterns across tissue types and across development. Translational studies that combine findings from (a) carefully conducted epidemiologic studies including longitudinal twin studies [244], (b) molecular biology studies on prospectively collected, easily accessible human tissue (such as blood lymphocytes, buccal mucosa, or germline cells) as well as post mortem brain tissue [245], and (c) experimental animal studies on the effects of environmental exposures during sensitive time periods of development [68, 246] will allow for further exploration on the role of epigenetics in human development, physiology, and pathophysiology. Figure 18.1 illustrates a neurodevelopmental view on the role of epigenetics in the

Fig. 18.1 Neurodevelopmental model of epigenetic involvement in the etiology of major psychotic disorders. A scheme on a neurodevelopmental model of the etiology of psychiatric disorders illustrating an individual's age (Fig. 18.1a), the approximate timing of risk-increasing exposures (Fig. 18.1b), epigenetic marks that can be inherited or acquired over life, as a possible consequences of exposures (Fig. 18.1c), timing of neurodevelopmental processes (Fig. 18.1d), and the temporal sequence of the expression of psychiatric phenotypes (Fig. 18.1e). Exposures depicted in 1b may thus induce epigenetic alterations on the molecular level that may impact on crucial neurobiological processes during sensitive periods of neurodevelopment, and may synergistically give rise to aberrations in mental health, phenotypically expressed by quantitative (and/or qualitative) alterations in psychological functions that in interaction with the individual's social world may result into psychiatric disorder. *Black circles* in 1c represent epigenetic modifications to DNA (e.g., methylated cytosines or histone modifications). The row represents a nucleus of a post-mitotic cell with dynamic changes to DNA modifications which may accumulate within the individual throughout development as a consequence of the synergistical combination of multiple exposures. These dynamic changes in DNA methylation may be associated with aberrations in neurdevelopmental processes such as, for example, myelination (Fig. 18.1d), and the appearance of subclinical psychotic symptoms (Fig.18.1e) that can become abnormally persistent and ultimately lead to the onset of a major psychotic disorder (Fig. 18.1e). While this figure illustrates that development of major psychotic disorders is associated with accumulation of DNA modifications of a gene, scenarios are equally possible where the gene is associated with the removal of epigenetic modifications



etiology of psychiatric disorders. Current scientific approaches can take advantage of recent developments in genome-wide analyses of genetic variations, gene expression, and epigenetic marks that allow for explorative, hypothesis-free investigations.

The exploration of epigenetic mediation of environmental exposures may benefit from testing for gene-environment interactions by investigating genetic variants in epigenome-relevant genes for their interaction with environmental exposures in

large, well-characterized samples of subjects with psychiatric disorders [32]. Furthermore, Mendelian randomization designs [247], longitudinal studies including patients with "at risk" mental states for psychiatric disorders (such as in psychosis [248] and dementia [249]), and the use of twin studies discordant for psychiatric phenotypes and/or environmental exposures [244] may yield important insights in the role of epigenetic alterations in the onset and/or course of psychiatric disorders. Although rodent studies do not allow to make direct inferences about the human situation, animal studies will be very informative in elucidating the role of epigenetics in brain function and dysfunction, as both genetic and environmental factors can be (relatively) well controlled and brain tissue easily obtained [250].

In conclusion, the field of epigenetic research in psychiatric disorders is in a very early phase, but promising preliminary findings encourage carefully designed studies combining epidemiology, clinical science, and molecular genetics.

References

1. Epigenetics MG (2010) The seductive allure of behavioral epigenetics. Science 329:24–27
2. Petronis A (2010) Epigenetics as a unifying principle in the aetiology of complex traits and diseases. Nature 465:721–727
3. Feinberg AP (2007) Phenotypic plasticity and the epigenetics of human disease. Nature 447:433–440
4. Weaver JR, Susiarjo M, Bartolomei MS (2009) Imprinting and epigenetic changes in the early embryo. Mamm Genome 20:532–543
5. Hsieh J, Eisch AJ (2010) Epigenetics, hippocampal neurogenesis, and neuropsychiatric disorders: unraveling the genome to understand the mind. Neurobiol Dis 39:73–84
6. Yu Y, Casaccia P, Lu QR (2010) Shaping the oligodendrocyte identity by epigenetic control. Epigenetics 5:124–128
7. Hsieh J, Gage FH (2005) Chromatin remodeling in neural development and plasticity. Curr Opin Cell Biol 17:664–671
8. Levenson JM, Sweatt JD (2006) Epigenetic mechanisms: a common theme in vertebrate and invertebrate memory formation. Cell Mol Life Sci 63:1009–1016
9. Renthal W, Nestler EJ (2008) Epigenetic mechanisms in drug addiction. Trends Mol Med 14:341–350
10. Day JJ, Sweatt JD (2010) DNA methylation and memory formation. Nat Neurosci 13:1319–1323
11. Feng J, Zhou Y, Campbell SL, Le T, Li E, Sweatt JD et al (2010) Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. Nat Neurosci 13:423–430
12. Flavell SW, Greenberg ME (2008) Signaling mechanisms linking neuronal activity to gene expression and plasticity of the nervous system. Annu Rev Neurosci 31:563–590
13. MacDonald JL, Roskams AJ (2009) Epigenetic regulation of nervous system development by DNA methylation and histone deacetylation. Prog Neurobiol 88:170–183
14. Bale TL, Baram TZ, Brown AS, Goldstein JM, Insel TR, McCarthy MM et al (2010) Early life programming and neurodevelopmental disorders. Biol Psychiatry 68:314–319
15. Fraga MF, Esteller M (2007) Epigenetics and aging: the targets and the marks. Trends Genet 23:413–418
16. Chouliaras L, Rutten BP, Kenis G, Peerbooms O, Visser PJ, Verhey F et al (2010) Epigenetic regulation in the pathophysiology of Alzheimer's disease. Prog Neurobiol 90:498–510

17. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL et al (2009) Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet 5:e1000602
18. Hernandez DG, Nalls MA, Gibbs JR, Arepalli S, van der Brug M, Chong S et al (2011) Distinct DNA methylation changes highly correlated with chronological age in the human brain. Hum Mol Genet 20(6):1164–1172
19. Chouliaras L, van den Hove DL, Kenis G, Cruz JD, Lemmens MA, van Os J et al (2011) Caloric restriction attenuates age-related changes of DNA methyltransferase 3a in mouse hippocampus. Brain Behav Immun 25(4):616–623
20. Colman RJ, Anderson RM, Johnson SC, Kastman EK, Kosmatka KJ, Beasley TM et al (2009) Caloric restriction delays disease onset and mortality in rhesus monkeys. Science 325:201–204
21. Mattson MP, Duan W, Chan SL, Cheng A, Haughey N, Gary DS et al (2002) Neuroprotective and neurorestorative signal transduction mechanisms in brain aging: modification by genes, diet and behavior. Neurobiol Aging 23:695–705
22. Sohal RS, Weindruch R (1996) Oxidative stress, caloric restriction, and aging. Science 273:59–63
23. Rutten BP, Brasnjevic I, Steinbusch HW, Schmitz C (2010) Caloric restriction and aging but not overexpression of SOD1 affect hippocampal volumes in mice. Mech Ageing Dev 131:574–579
24. Liu L, van Groen T, Kadish I, Tollefsbol TO (2009) DNA methylation impacts on learning and memory in aging. Neurobiol Aging 30:549–560
25. Peleg S, Sananbenesi F, Zovoilis A, Burkhardt S, Bahari-Javan S, Agis-Balboa RC et al (2010) Altered histone acetylation is associated with age-dependent memory impairment in mice. Science 328:753–756
26. Longo VD, Kennedy BK (2006) Sirtuins in aging and age-related disease. Cell 126: 257–268
27. Mehler MF (2008) Epigenetic principles and mechanisms underlying nervous system functions in health and disease. Prog Neurobiol 86:305–341
28. Khan NL, Wood NW (1999) Prader-Willi and Angelman syndromes: update on genetic mechanisms and diagnostic complexities. Curr Opin Neurol 12:149–154
29. Lalande M, Calciano MA (2007) Molecular epigenetics of Angelman syndrome. Cell Mol Life Sci 64:947–960
30. Christodoulou J, Weaving LS (2003) MECP2 and beyond: phenotype-genotype correlations in Rett syndrome. J Child Neurol 18:669–674
31. Dimitropoulos A, Schultz RT (2007) Autistic-like symptomatology in Prader-Willi syndrome: a review of recent findings. Curr Psychiatry Rep 9:159–164
32. van Os J, Rutten BP, Poulton R (2008) Gene-environment interactions in schizophrenia: review of epidemiological findings and future directions. Schizophr Bull 34:1066–1082
33. Petronis A, Gottesman II, Kan P, Kennedy JL, Basile VS, Paterson AD et al (2003) Monozygotic twins exhibit numerous epigenetic differences: clues to twin discordance? Schizophr Bull 29:169–178
34. Zwijnenburg PJ, Meijers-Heijboer H, Boomsma DI (2010) Identical but not the same: the value of discordant monozygotic twins in genetic research. Am J Med Genet B Neuropsychiatr Genet 153B:1134–1149
35. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML et al (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 102:10604–10609
36. Kaminsky ZA, Tang T, Wang SC, Ptak C, Oh GH, Wong AH et al (2009) DNA methylation profiles in monozygotic and dizygotic twins. Nat Genet 41:240–245
37. Mill J, Dempster E, Caspi A, Williams B, Moffitt T, Craig I (2006) Evidence for monozygotic twin (MZ) discordance in methylation level at two CpG sites in the promoter region of the catechol-O-methyltransferase (COMT) gene. Am J Med Genet B Neuropsychiatr Genet 141:421–425

38. Kuratomi G, Iwamoto K, Bundo M, Kusumi I, Kato N, Iwata N et al (2008) Aberrant DNA methylation associated with bipolar disorder identified from discordant monozygotic twins. Mol Psychiatry 13:429–441
39. Kaminsky Z, Petronis A, Wang SC, Levine B, Ghaffar O, Floden D et al (2008) Epigenetics of personality traits: an illustrative study of identical twins discordant for risk-taking behavior. Twin Res Hum Genet 11:1–11
40. Baranzini SE, Mudge J, van Velkinburgh JC, Khankhanian P, Khrebtukova I, Miller NA et al (2010) Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. Nature 464:1351–1356
41. Rutten BP, Mill J (2009) Epigenetic mediation of environmental influences in major psychotic disorders. Schizophr Bull 35:1045–1056
42. Jirtle RL, Skinner MK (2007) Environmental epigenomics and disease susceptibility. Nat Rev Genet 8:253–262
43. Haque FN, Gottesman II, Wong AH (2009) Not really identical: epigenetic differences in monozygotic twins and implications for twin studies in psychiatry. Am J Med Genet C Semin Med Genet 151C:136–141
44. Malaspina D, Perrin M, Kleinhaus KR, Opler M, Harlap S (2008) Growth and schizophrenia: aetiology, epidemiology and epigenetics. Novartis Found Symp 289:196–203, discussion 203–197, 238–140
45. Delahanty RJ, Kang JQ, Brune CW, Kistner EO, Courchesne E, Cox NJ et al (2011) Maternal transmission of a rare GABRB3 signal peptide variant is associated with autism. Mol Psychiatry 16(1):86–96
46. Anney RJ, Hawi Z, Sheehan K, Mulligan A, Pinto C, Brookes KJ et al (2008) Parent of origin effects in attention/deficit hyperactivity disorder (ADHD): analysis of data from the international multicenter ADHD genetics (IMAGE) program. Am J Med Genet B Neuropsychiatr Genet 147B:1495–1500
47. De Luca V, Likhodi O, Kennedy JL, Wong AH (2007) Parent-of-origin effect and genomic imprinting of the HTR2A receptor gene T102C polymorphism in psychosis. Psychiatry Res 151:243–248
48. De Luca V, Likhodi O, Kennedy JL, Wong AH (2007) Differential expression and parent-of-origin effect of the 5-HT2A receptor gene C102T polymorphism: analysis of suicidality in schizophrenia and bipolar disorder. Am J Med Genet B Neuropsychiatr Genet 144B:370–374
49. Goos LM, Ezzatian P, Schachar R (2007) Parent-of-origin effects in attention-deficit hyperactivity disorder. Psychiatry Res 149:1–9
50. Bassett SS, Avramopoulos D, Perry RT, Wiener H, Watson B Jr, Go RC et al (2006) Further evidence of a maternal parent-of-origin effect on chromosome 10 in late-onset Alzheimer's disease. Am J Med Genet B Neuropsychiatr Genet 141B:537–540
51. Perrin MC, Brown AS, Malaspina D (2007) Aberrant epigenetic regulation could explain the relationship of paternal age to schizophrenia. Schizophr Bull 33:1270–1273
52. Grether JK, Anderson MC, Croen LA, Smith D, Windham GC (2009) Risk of autism and increasing maternal and paternal age in a large north American population. Am J Epidemiol 170:1118–1126
53. Reichenberg A, Gross R, Sandin S, Susser ES (2010) Advancing paternal and maternal age are both important for autism risk. Am J Public Health 100:772–773, author reply 773
54. Croen LA, Najjar DV, Fireman B, Grether JK (2007) Maternal and paternal age and risk of autism spectrum disorders. Arch Pediatr Adolesc Med 161:334–340
55. Reichenberg A, Gross R, Weiser M, Bresnahan M, Silverman J, Harlap S et al (2006) Advancing paternal age and autism. Arch Gen Psychiatry 63:1026–1032
56. Miller B, Messias E, Miettunen J, Alaräisänen A, Järvelin MR, Koponen H et al (2010) Meta-analysis of paternal age and schizophrenia risk in male versus female offspring. Schizophr Bull 2011 Sep 37(5):1039–1047
57. Saha S, Barnett AG, Foldi C, Burne TH, Eyles DW, Buka SL et al (2009) Advanced paternal age is associated with impaired neurocognitive outcomes during infancy and childhood. PLoS Med 6:e40

58. Saha S, Barnett AG, Buka SL, McGrath JJ (2009) Maternal age and paternal age are associated with distinct childhood behavioural outcomes in a general population birth cohort. Schizophr Res 115:130–135
59. Sherazi R, McKeon P, McDonough M, Daly I, Kennedy N (2006) What's new? The clinical epidemiology of bipolar I disorder. Harv Rev Psychiatry 14:273–284
60. Bijl RV, De Graaf R, Ravelli A, Smit F, Vollebergh WA (2002) Gender and age-specific first incidence of DSM-III-R psychiatric disorders in the general population. Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). Soc Psychiatry Psychiatr Epidemiol 37:372–379
61. Bijl RV, Ravelli A, van Zessen G (1998) Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Soc Psychiatry Psychiatr Epidemiol 33:587–595
62. Wittchen HU, Stein MB, Kessler RC (1999) Social fears and social phobia in a community sample of adolescents and young adults: prevalence, risk factors and co-morbidity. Psychol Med 29:309–323
63. Polanczyk G, Rohde LA (2007) Epidemiology of attention-deficit/hyperactivity disorder across the lifespan. Curr Opin Psychiatry 20:386–392
64. Fombonne E (2009) Epidemiology of pervasive developmental disorders. Pediatr Res 65:591–598
65. Fombonne E (1999) The epidemiology of autism: a review. Psychol Med 29:769–786
66. Nelson CB, Wittchen HU (1998) DSM-IV alcohol disorders in a general population sample of adolescents and young adults. Addiction 93:1065–1077
67. van Os J, Kapur S (2009) Schizophrenia. Lancet 374:635–645
68. van Os J, Kenis G, Rutten BP (2010) The environment and schizophrenia. Nature 468:203–212
69. McCarthy MM, Auger AP, Bale TL, De Vries GJ, Dunn GA, Forger NG et al (2009) The epigenetics of sex differences in the brain. J Neurosci 29:12815–12823
70. Schwarz JM, Nugent BM, McCarthy MM (2010) Developmental and hormone-induced epigenetic changes to estrogen and progesterone receptor genes in brain are dynamic across the life span. Endocrinology 151:4871–4881
71. Waggoner D (2007) Mechanisms of disease: epigenesis. Semin Pediatr Neurol 14:7–14
72. Monteiro J, Derom C, Vlietinck R, Kohn N, Lesser M, Gregersen PK (1998) Commitment to X inactivation precedes the twinning event in monochorionic MZ twins. Am J Hum Genet 63:339–346
73. Manning N (2008) The influence of twinning on cardiac development. Early Hum Dev 84:173–179
74. Hall LL, Lawrence JB (2003) The cell biology of a novel chromosomal RNA: chromosome painting by XIST/Xist RNA initiates a remodeling cascade. Semin Cell Dev Biol 14:369–378
75. Loat CS, Asbury K, Galsworthy MJ, Plomin R, Craig IW (2004) X inactivation as a source of behavioural differences in monozygotic female twins. Twin Res 7:54–61
76. Peerbooms OL, Wichers M, Jacobs N, Kenis G, Derom C, Vlietinck R et al (2010) No major role for X-inactivation in variations of intelligence and behavioral problems at middle childhood. Am J Med Genet B Neuropsychiatr Genet 153B:1311–1317
77. Rosa A, Picchioni MM, Kalidindi S, Loat CS, Knight J, Toulopoulou T et al (2008) Differential methylation of the X-chromosome is a possible source of discordance for bipolar disorder female monozygotic twins. Am J Med Genet B Neuropsychiatr Genet 147B:459–462
78. Wichers M, Schrijvers D, Geschwind N, Jacobs N, Myin-Germeys I, Thiery E et al (2009) Mechanisms of gene-environment interactions in depression: evidence that genes potentiate multiple sources of adversity. Psychol Med 39:1077–1086
79. Wichers MC, Purcell S, Danckaerts M, Derom C, Derom R, Vlietinck R et al (2002) Prenatal life and post-natal psychopathology: evidence for negative gene-birth weight interaction. Psychol Med 32:1165–1174
80. Oh G, Petronis A (2008) Environmental studies of schizophrenia through the prism of epigenetics. Schizophr Bull 34:1122–1129

81. Crow TJ (2007) How and why genetic linkage has not solved the problem of psychosis: review and hypothesis. Am J Psychiatry 164:13–21
82. Wong AH, Gottesman II, Petronis A (2005) Phenotypic differences in genetically identical organisms: the epigenetic perspective. Hum Mol Genet 14(Spec No 1):R11–R18
83. Khashan AS, Abel KM, McNamee R, Pedersen MG, Webb RT, Baker PN et al (2008) Higher risk of offspring schizophrenia following antenatal maternal exposure to severe adverse life events. Arch Gen Psychiatry 65:146–152
84. Huttunen MO, Niskanen P (1978) Prenatal loss of father and psychiatric disorders. Arch Gen Psychiatry 35:429–431
85. Van Os J, Selten J-P (1998) Prenatal exposure to maternal stress and subsequent schizophrenia: the May 1940 invasion of the Netherlands. Br J Psychiatry 172:324–326
86. Susser E, Neugebauer R, Hoek HW, Brown AS, Lin S, Labovitz D et al (1996) Schizophrenia after prenatal famine. Further evidence. Arch Gen Psychiatry 53:25–31
87. Xu MQ, Sun WS, Liu BX, Feng GY, Yu L, Yang L et al (2009) Prenatal malnutrition and adult schizophrenia: further evidence from the 1959–1961 Chinese famine. Schizophr Bull 35:568–576
88. Insel BJ, Schaefer CA, McKeague IW, Susser ES, Brown AS (2008) Maternal iron deficiency and the risk of schizophrenia in offspring. Arch Gen Psychiatry 65:1136–1144
89. Opler MG, Buka SL, Groeger J, McKeague I, Wei C, Factor-Litvak P et al (2008) Prenatal exposure to lead, delta-aminolevulinic acid, and schizophrenia: further evidence. Environ Health Perspect 116:1586–1590
90. Brown AS, Bottiglieri T, Schaefer CA, Quesenberry CP Jr, Liu L, Bresnahan M et al (2007) Elevated prenatal homocysteine levels as a risk factor for schizophrenia. Arch Gen Psychiatry 64:31–39
91. Hollister JM, Laing P, Mednick SA (1996) Rhesus incompatibility as a risk factor for schizophrenia in male adults. Arch Gen Psychiatry 53:19–24
92. McGrath JJ, Eyles DW, Pedersen CB, Anderson C, Ko P, Burne TH et al (2010) Neonatal vitamin D status and risk of schizophrenia: a population-based case-control study. Arch Gen Psychiatry 67(9):889–894
93. Mortensen PB, Norgaard-Pedersen B, Waltoft BL, Sorensen TL, Hougaard D, Torrey EF et al (2007) Toxoplasma gondii as a risk factor for early-onset schizophrenia: analysis of filter paper blood samples obtained at birth. Biol Psychiatry 61:688–693
94. Brown AS, Schaefer CA, Quesenberry CP Jr, Liu L, Babulas VP, Susser ES (2005) Maternal exposure to toxoplasmosis and risk of schizophrenia in adult offspring. Am J Psychiatry 162:767–773
95. Brown AS, Derkits EJ (2010) Prenatal infection and schizophrenia: a review of epidemiologic and translational studies. Am J Psychiatry 167:261–280
96. Sorensen HJ, Mortensen EL, Reinisch JM, Mednick SA (2009) Association between prenatal exposure to bacterial infection and risk of schizophrenia. Schizophr Bull 35:631–637
97. Clarke MC, Tanskanen A, Huttunen M, Whittaker JC, Cannon M (2009) Evidence for an interaction between familial liability and prenatal exposure to infection in the causation of schizophrenia. Am J Psychiatry 166:1025–1030
98. Sorensen HJ, Mortensen EL, Reinisch JM, Mednick SA (2003) Do hypertension and diuretic treatment in pregnancy increase the risk of schizophrenia in offspring? Am J Psychiatry 160:464–468
99. Sorensen HJ, Mortensen EL, Reinisch JM, Mednick SA (2004) Association between prenatal exposure to analgesics and risk of schizophrenia. Br J Psychiatry 185:366–371
100. Oberlander TF, Weinberg J, Papsdorf M, Grunau R, Misri S, Devlin AM (2008) Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress responses. Epigenetics 3:97–106
101. Zheng S, Rollet M, Pan YX (2011) Maternal protein restriction during pregnancy induces CCAAT/enhancer-binding protein (C/EBPbeta) expression through the regulation of histone modification at its promoter region in female offspring rat skeletal muscle. Epigenetics 6(2):161–170

102. Thornburg KL, Shannon J, Thuillier P, Turker MS (2010) In utero life and epigenetic predisposition for disease. Adv Genet 71:57–78
103. Waterland RA, Michels KB (2007) Epigenetic epidemiology of the developmental origins hypothesis. Annu Rev Nutr 27:363–388
104. Selten JP, Frissen A, Lensvelt-Mulders G, Morgan VA (2010) Schizophrenia and 1957 pandemic of influenza: meta-analysis. Schizophr Bull 36:219–228
105. Selten JP, Cantor-Graae E, Nahon D, Levav I, Aleman A, Kahn RS (2003) No relationship between risk of schizophrenia and prenatal exposure to stress during the Six-Day War or Yom Kippur War in Israel. Schizophr Res 63:131–135
106. McGrath J, Eyles D, Mowry B, Yolken R, Buka S (2003) Low maternal vitamin D as a risk factor for schizophrenia: a pilot study using banked sera. Schizophr Res 63:73–78
107. Cannon TD, Rosso IM, Bearden CE, Sanchez LE, Hadley T (1999) A prospective cohort study of neurodevelopmental processes in the genesis and epigenesis of schizophrenia. Dev Psychopathol 11:467–485
108. Buka SL, Tsuang MT, Torrey EF, Klebanoff MA, Bernstein D, Yolken RH (2001) Maternal infections and subsequent psychosis among offspring. Arch Gen Psychiatry 58:1032–1037
109. Brown AS, Schaefer CA, Quesenberry CP Jr, Shen L, Susser ES (2006) No evidence of relation between maternal exposure to herpes simplex virus type 2 and risk of schizophrenia? Am J Psychiatry 163:2178–2180
110. Jaaro-Peled H, Ayhan Y, Pletnikov MV, Sawa A (2010) Review of pathological hallmarks of schizophrenia: comparison of genetic models with patients and nongenetic models. Schizophr Bull 36:301–313
111. Ayhan Y, Sawa A, Ross CA, Pletnikov MV (2009) Animal models of gene-environment interactions in schizophrenia. Behav Brain Res 204:274–281
112. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES et al (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci USA 105:17046–17049
113. Brown AS, Susser ES (2008) Prenatal nutritional deficiency and risk of adult schizophrenia. Schizophr Bull 34:1054–1063
114. Smits L, Pedersen C, Mortensen P, van Os J (2004) Association between short birth intervals and schizophrenia in the offspring. Schizophr Res 70:49–56
115. Mortensen PB, Pedersen MG, Pedersen CB (2010) Psychiatric family history and schizophrenia risk in Denmark: which mental disorders are relevant? Psychol Med 40:201–210
116. Cannon M, Jones PB, Murray RM (2002) Obstetric complications and schizophrenia: historical and meta-analytic review. Am J Psychiatry 159:1080–1092
117. Byrne M, Agerbo E, Bennedsen B, Eaton WW, Mortensen PB (2007) Obstetric conditions and risk of first admission with schizophrenia: a Danish national register based study. Schizophr Res 97:51–59
118. Rutter M (2009) Understanding and testing risk mechanisms for mental disorders. J Child Psychol Psychiatry 50:44–52
119. Klaning U (1999) Greater occurrence of schizophrenia in dizygotic but not monozygotic twins. Register-based study. Br J Psychiatry 175:407–409
120. Klaning U, Pedersen CB, Mortensen PB, Kyvik KO, Skytthe A (2002) A possible association between the genetic predisposition for dizygotic twinning and schizophrenia. Schizophr Res 58:31–35
121. Kendler KS, Pedersen NL, Farahmand BY, Persson PG (1996) The treated incidence of psychotic and affective illness in twins compared with population expectation: a study in the Swedish Twin and Psychiatric Registries. Psychol Med 26:1135–1144
122. Spauwen J, Krabbendam L, Lieb R, Wittchen HU, van Os J (2004) Early maternal stress and health behaviours and offspring expression of psychosis in adolescence. Acta Psychiatr Scand 110:356–364
123. Bartels-Velthuis AA, Jenner JA, van de Willige G, van Os J, Wiersma D (2010) Prevalence and correlates of auditory vocal hallucinations in middle childhood. Br J Psychiatry 196:41–46

124. Zammit S, Odd D, Horwood J, Thompson A, Thomas K, Menezes P et al (2009) Investigating whether adverse prenatal and perinatal events are associated with non-clinical psychotic symptoms at age 12 years in the ALSPAC birth cohort. Psychol Med 39(9):1457–1467
125. Mittal VA, Ellman LM, Cannon TD (2008) Gene-environment interaction and covariation in schizophrenia: the role of obstetric complications. Schizophr Bull 34:1083–1094
126. Cannon TD (1997) On the nature and mechanisms of obstetric influences in schizophrenia: a review and synthesis of epidemiologic studies. Int Rev Psychiatry 9:387–397
127. Schmidt-Kastner R, van Os J, Steinbusch HWM, Schmitz C (2006) Gene regulation by hypoxia and the neurodevelopmental origin of schizophrenia. Schizophr Res 84:253–271
128. Nicodemus KK, Marenco S, Batten AJ, Vakkalanka R, Egan MF, Straub RE et al (2008) Serious obstetric complications interact with hypoxia-regulated/vascular-expression genes to influence schizophrenia risk. Mol Psychiatry 13(9):873–877
129. Watson JA, Watson CJ, McCann A, Baugh J (2010) Epigenetics, the epicenter of the hypoxic response. Epigenetics 5:293–296
130. Tienari P, Wynne LC, Moring J, Lahti I, Naarala M, Sorri A et al (1994) The finnish adoptive family study of schizophrenia. Implications for family research [see comments]. Br J Psychiatry Suppl 20–26
131. Wahlberg KE, Wynne LC, Oja H, Keskitalo P, Pykalainen L, Lahti I et al (1997) Gene-environment interaction in vulnerability to schizophrenia: findings from the Finnish Adoptive Family Study of Schizophrenia. Am J Psychiatry 154:355–362
132. Tienari P, Wynne LC, Sorri A, Lahti I, Laksy K, Moring J et al (2004) Genotype-environment interaction in schizophrenia-spectrum disorder. Long-term follow-up study of Finnish adoptees. Br J Psychiatry 184:216–222
133. Wahlberg KE, Wynne LC, Hakko H, Laksy K, Moring J, Miettunen J et al (2004) Interaction of genetic risk and adoptive parent communication deviance: longitudinal prediction of adoptee psychiatric disorders. Psychol Med 34:1531–1541
134. Carter JW, Parnas J, Cannon TD, Schulsinger F, Mednick SA (1999) MMPI variables predictive of schizophrenia in the Copenhagen High-Risk Project: a 25-year follow-up. Acta Psychiatr Scand 99:432–440
135. Schreier A, Wolke D, Thomas K, Horwood J, Hollis C, Gunnell D et al (2009) Prospective study of peer victimization in childhood and psychotic symptoms in a nonclinical population at age 12 years. Arch Gen Psychiatry 66:527–536
136. Kessler RC, McLaughlin KA, Green JG, Gruber MJ, Sampson NA, Zaslavsky AM et al (2010) Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. Br J Psychiatry 197:378–385
137. McLaughlin KA, Green JG, Gruber MJ, Sampson NA, Zaslavsky AM, Kessler RC (2010) Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication II: associations with persistence of DSM-IV disorders. Arch Gen Psychiatry 67:124–132
138. Green JG, McLaughlin KA, Berglund PA, Gruber MJ, Sampson NA, Zaslavsky AM et al (2010) Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication I: associations with first onset of DSM-IV disorders. Arch Gen Psychiatry 67:113–123
139. Bruffaerts R, Demyttenaere K, Borges G, Haro JM, Chiu WT, Hwang I et al (2010) Childhood adversities as risk factors for onset and persistence of suicidal behaviour. Br J Psychiatry 197:20–27
140. Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR et al (2004) Epigenetic programming by maternal behavior. Nat Neurosci 7:847–854
141. Weaver IC, Champagne FA, Brown SE, Dymov S, Sharma S, Meaney MJ et al (2005) Reversal of maternal programming of stress responses in adult offspring through methyl supplementation: altering epigenetic marking later in life. J Neurosci 25:11045–11054
142. Weaver IC, Meaney MJ, Szyf M (2006) Maternal care effects on the hippocampal transcriptome and anxiety-mediated behaviors in the offspring that are reversible in adulthood. Proc Natl Acad Sci USA 103:3480–3485

143. McGowan PO, Sasaki A, D'Alessio AC, Dymov S, Labonte B, Szyf M et al (2009) Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. Nat Neurosci 12:342–348
144. Arseneault L, Cannon M, Witton J, Murray RM (2004) Causal association between cannabis and psychosis: examination of the evidence. Br J Psychiatry 184:110–117
145. Henquet C, Di Forti M, Morrison P, Kuepper R, Murray RM (2008) Gene-environment interplay between cannabis and psychosis. Schizophr Bull 34:1111–1121
146. Murray RM, Morrison PD, Henquet C, Di Forti M (2007) Cannabis, the mind and society: the hash realities. Nat Rev Neurosci 8:885–895
147. Houston JE, Murphy J, Adamson G, Stringer M, Shevlin M (2008) Childhood sexual abuse, early cannabis use, and psychosis: testing an interaction model based on the National Comorbidity Survey. Schizophr Bull 34:580–585
148. Caspi A, Moffitt TE, Cannon M, McClay J, Murray R, Harrington H et al (2005) Moderation of the effect of adolescent-onset cannabis use on adult psychosis by a functional polymorphism in the catechol-O-methyltransferase gene: longitudinal evidence of a gene X environment interaction. Biol Psychiatry 57:1117–1127
149. Henquet C, Rosa A, Delespaul P, Papiol S, Fananas L, van Os J et al (2009) COMT ValMet moderation of cannabis-induced psychosis: a momentary assessment study of 'switching on' hallucinations in the flow of daily life. Acta Psychiatr Scand 119:156–160
150. van Winkel R (2011) Family-based analysis of genetic variation underlying psychosis-inducing effects of cannabis: sibling analysis and proband follow-up. Arch Gen Psychiatry 68(2):148–157
151. De Hert M, Wampers M, Jendricko T, Franic T, Vidovic D, De Vriendt N et al (2011) Effects of cannabis use on age at onset in schizophrenia and bipolar disorder. Schizophr Res 126(1–3):270–276
152. Jacobs N, Rijsdijk F, Derom C, Vlietinck R, Delespaul P, van Os J et al (2006) Genes making one feel blue in the flow of daily life: a momentary assessment study of gene-stress interaction. Psychosom Med 68:201–206
153. G.R.O.U.P (2010) Evidence that familial liability for psychosis is expressed as differential sensitivity to cannabis: an analysis of patient-sibling and sibling-control pairs. Arch Gen Psychiatry 68(2):138–147
154. Chevaleyre V, Takahashi KA, Castillo PE (2006) Endocannabinoid-mediated synaptic plasticity in the CNS. Annu Rev Neurosci 29:37–76
155. Villares J (2007) Chronic use of marijuana decreases cannabinoid receptor binding and mRNA expression in the human brain. Neuroscience 145:323–334
156. Ellgren M, Spano SM, Hurd YL (2007) Adolescent cannabis exposure alters opiate intake and opioid limbic neuronal populations in adult rats. Neuropsychopharmacology 32:607–615
157. Casu MA, Pisu C, Sanna A, Tambaro S, Spada GP, Mongeau R et al (2005) Effect of delta9-tetrahydrocannabinol on phosphorylated CREB in rat cerebellum: an immunohistochemical study. Brain Res 1048:41–47
158. Fernandez-Ruiz J, Gomez M, Hernandez M, de Miguel R, Ramos JA (2004) Cannabinoids and gene expression during brain development. Neurotox Res 6:389–401
159. Mato S, Chevaleyre V, Robbe D, Pazos A, Castillo PE, Manzoni OJ (2004) A single in-vivo exposure to delta 9THC blocks endocannabinoid-mediated synaptic plasticity. Nat Neurosci 7:585–586
160. Scallet AC (1991) Neurotoxicology of cannabis and THC: a review of chronic exposure studies in animals. Pharmacol Biochem Behav 40:671–676
161. Heath RG, Fitzjarrell AT, Fontana CJ, Garey RE (1980) Cannabis sativa: effects on brain function and ultrastructure in rhesus monkeys. Biol Psychiatry 15:657–690
162. Hoffman AF, Oz M, Caulder T, Lupica CR (2003) Functional tolerance and blockade of long-term depression at synapses in the nucleus accumbens after chronic cannabinoid exposure. J Neurosci 23:4815–4820

163. Featherstone RE, Kapur S, Fletcher PJ (2007) The amphetamine-induced sensitized state as a model of schizophrenia. Prog Neuropsychopharmacol Biol Psychiatry 31:1556–1571
164. Castner SA, Williams GV (2007) From vice to virtue: insights from sensitization in the nonhuman primate. Prog Neuropsychopharmacol Biol Psychiatry 31:1572–1592
165. Uslaner J, Badiani A, Norton CS, Day HEW, Watson SJ, Akil H et al (2001) Amphetamine and cocaine induce different patterns of c-fos mRNA expression in the striatum and subthalamic nucleus depending on environmental context. Eur J Neurosci 13:1977–1983
166. Bibb JA, Chen J, Taylor JR, Svenningsson P, Nishi A, Snyder GL et al (2001) Effects of chronic exposure to cocaine are regulated by the neuronal protein Cdk5. Nature 410:376–380
167. Nestler EJ, Barrot M, Self DW (2001) Delta FosB: a sustained molecular switch for addiction. Proc Natl Acad Sci 98:11042
168. Kumar A, Choi KH, Renthal W, Tsankova NM, Theobald DEH, Truong HT et al (2005) Chromatin remodeling is a key mechanism underlying cocaine-induced plasticity in striatum. Neuron 48:303–314
169. Sekine Y, Iyo M, Ouchi Y, Matsunaga T, Tsukada H, Okada H et al (2001) Methamphetamine-related psychiatric symptoms and reduced brain dopamine transporters studied with PET. Am J Psychiatry 158:1206–1214
170. Lehrmann E, Colantuoni C, Deep-Soboslay A, Becker KG, Lowe R, Huestis MA et al (2006) Transcriptional changes common to human cocaine, cannabis and phencyclidine abuse. PLoS One 1:e114
171. Greenstein R, Novak G, Seeman P (2007) Amphetamine sensitization elevates CaMKIIbeta mRNA. Synapse 61:827–834
172. Iwata SI, Hewlett GH, Ferrell ST, Kantor L, Gnegy ME (1997) Enhanced dopamine release and phosphorylation of synapsin I and neuromodulin in striatal synaptosomes after repeated amphetamine. J Pharmacol Exp Ther 283:1445–1452
173. Cantor-Graae E, Selten JP (2005) Schizophrenia and migration: a meta-analysis and review. Am J Psychiatry 162:12–24
174. Bourque F, van der Ven E, Malla A (2010) A meta-analysis of the risk for psychotic disorders among first- and second-generation immigrants. Psychol Med 41(5):1–14
175. Bresnahan M, Begg MD, Brown A, Schaefer C, Sohler N, Insel B et al (2007) Race and risk of schizophrenia in a US birth cohort: another example of health disparity? Int J Epidemiol 36:751–758
176. Veling W, Susser E, van Os J, Mackenbach JP, Selten JP, Hoek HW (2008) Ethnic density of neighborhoods and incidence of psychotic disorders among immigrants. Am J Psychiatry 165:66–73
177. Boydell J, Van Os J, McKenzie K, Allardyce J, Goel R, McCreadie RG et al (2001) Incidence of schizophrenia in ethnic minorities in London: ecological study into interactions with environment. BMJ 323:1336
178. Morgan C, Charalambides M, Hutchinson G, Murray RM (2010) Migration, ethnicity, and psychosis: toward a sociodevelopmental model. Schizophr Bull, Medline EPub. doi: sbq051 [pii] 10.1093/schbul/sbq051
179. Selten JP, Cantor-Graae E (2005) Social defeat: risk factor for schizophrenia? Br J Psychiatry 187:101–102
180. Renthal W, Maze I, Krishnan V, Covington HE 3rd, Xiao G, Kumar A et al (2007) Histone deacetylase 5 epigenetically controls behavioral adaptations to chronic emotional stimuli. Neuron 56:517–529
181. Tsankova NM, Berton O, Renthal W, Kumar A, Neve RL, Nestler EJ (2006) Sustained hippocampal chromatin regulation in a mouse model of depression and antidepressant action. Nat Neurosci 9:519–525
182. Cougnard A, Marcelis M, Myin-Germeys I, De Graaf R, Vollebergh W, Krabbendam L et al (2007) Does normal developmental expression of psychosis combine with environmental risk to cause persistence of psychosis? A psychosis proneness-persistence model. Psychol Med 37:513–527

183. Dominguez MD, Wichers M, Lieb R, Wittchen HU, van Os J (2011) Evidence that onset of clinical psychosis is an outcome of progressively more persistent subclinical psychotic experiences: an 8-year cohort study. Schizophr Bull 37(1):84–93
184. Nikulina EM, Covington HE 3rd, Ganschow L, Hammer RP Jr, Miczek KA (2004) Long-term behavioral and neuronal cross-sensitization to amphetamine induced by repeated brief social defeat stress: Fos in the ventral tegmental area and amygdala. Neuroscience 123:857–865
185. Kabbaj M, Isgor C, Watson SJ, Akil H (2002) Stress during adolescence alters behavioral sensitization to amphetamine. Neuroscience 113:395–400
186. Von Frijtag JC, Reijmers LG, Van der Harst JE, Leus IE, Van den Bos R, Spruijt BM (2000) Defeat followed by individual housing results in long-term impaired reward- and cognition-related behaviours in rats. Behav Brain Res 117:137–146
187. Sugden C (2006) One-carbon metabolism in psychiatric illness. Nutr Res Rev 19:117–136
188. van der Put NM, van Straaten HW, Trijbels FJ, Blom HJ (2001) Folate, homocysteine and neural tube defects: an overview. Exp Biol Med (Maywood) 226:243–270
189. Zhang HY, Luo GA, Liang QL, Wang Y, Yang HH, Wang YM et al (2008) Neural tube defects and disturbed maternal folate- and homocysteine-mediated one-carbon metabolism. Exp Neurol 212:515–521
190. Pasca SP, Dronca E, Kaucsar T, Craciun EC, Endreffy E, Ferencz BK et al (2009) One carbon metabolism disturbances and the C677T MTHFR gene polymorphism in children with autism spectrum disorders. J Cell Mol Med 13:4229–4238
191. de Jonge R, Tissing WJ, Hooijberg JH, Jansen G, Kaspers GJ, Lindemans J et al (2009) Polymorphisms in folate-related genes and risk of pediatric acute lymphoblastic leukemia. Blood 113:2284–2289
192. Wiemels JL, Smith RN, Taylor GM, Eden OB, Alexander FE, Greaves MF (2001) Methylenetetrahydrofolate reductase (MTHFR) polymorphisms and risk of molecularly defined subtypes of childhood acute leukemia. Proc Natl Acad Sci USA 98:4004–4009
193. Kronenberg G, Colla M, Endres M (2009) Folic acid, neurodegenerative and neuropsychiatric disease. Curr Mol Med 9:315–323
194. Kim JM, Kim SW, Shin IS, Yang SJ, Park WY, Kim SJ et al (2008) Folate, vitamin b(12), and homocysteine as risk factors for cognitive decline in the elderly. Psychiatry Investig 5:36–40
195. Kim YI (1999) Folate and carcinogenesis: evidence, mechanisms, and implications. J Nutr Biochem 10:66–88
196. Levine AJ, Figueiredo JC, Lee W, Conti DV, Kennedy K, Duggan DJ et al (2010) A candidate gene study of folate-associated one carbon metabolism genes and colorectal cancer risk. Cancer Epidemiol Biomarkers Prev 19:1812–1821
197. Smulders YM, Stehouwer CD (2005) Folate metabolism and cardiovascular disease. Semin Vasc Med 5:87–97
198. Carmichael SL, Yang W, Correa A, Olney RS, Shaw GM (2009) Hypospadias and intake of nutrients related to one-carbon metabolism. J Urol 181:315–321, discussion 321
199. Wani NA, Hamid A, Kaur J (2008) Folate status in various pathophysiological conditions. IUBMB Life 60:834–842
200. Betcheva ET, Mushiroda T, Takahashi A, Kubo M, Karachanak SK, Zaharieva IT et al (2009) Case-control association study of 59 candidate genes reveals the DRD2 SNP rs6277 (C957T) as the only susceptibility factor for schizophrenia in the Bulgarian population. J Hum Genet 54:98–107
201. Feng LG, Song ZW, Xin F, Hu J (2009) Association of plasma homocysteine and methylenetetrahydrofolate reductase C677T gene variant with schizophrenia: a Chinese Han population-based case-control study. Psychiatry Res 168:205–208
202. Gaysina D, Cohen S, Craddock N, Farmer A, Hoda F, Korszun A et al (2008) No association with the 5,10-methylenetetrahydrofolate reductase gene and major depressive disorder: results of the depression case control (DeCC) study and a meta-analysis. Am J Med Genet B Neuropsychiatr Genet 147B:699–706

203. Gilbody S, Lewis S, Lightfoot T (2007) Methylenetetrahydrofolate reductase (MTHFR) genetic polymorphisms and psychiatric disorders: a HuGE review. Am J Epidemiol 165:1–13
204. Pan CC, McQuoid DR, Taylor WD, Payne ME, Ashley-Koch A, Steffens DC (2009) Association analysis of the COMT/MTHFR genes and geriatric depression: an MRI study of the putamen. Int J Geriatr Psychiatry 24:847–855
205. Yuan YG, Zhang ZJ, Li JJ (2008) Plasma homocysteine but not MTHFR gene polymorphism is associated with geriatric depression in the Chinese population. Acta Neuropsychiatr 20(5):251–255
206. Yu L, Li T, Robertson Z, Dean J, Gu NF, Feng GY et al (2004) No association between polymorphisms of methylenetetrahydrofolate reductase gene and schizophrenia in both Chinese and Scottish populations. Mol Psychiatry 9:1063–1065
207. del Rio GC, Torres-Sanchez L, Chen J, Schnaas L, Hernandez C, Osorio E et al (2009) Maternal MTHFR 677 C>T genotype and dietary intake of folate and vitamin B(12): their impact on child neurodevelopment. Nutr Neurosci 12:13–20
208. Ueland PM, Hustad S, Schneede J, Refsum H, Vollset SE (2001) Biological and clinical implications of the MTHFR C677T polymorphism. Trends Pharmacol Sci 22:195–201
209. McGuffin P, Rijsdijk F, Andrew M, Sham P, Katz R, Cardno A (2003) The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. Arch Gen Psychiatry 60:497–502
210. Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF et al (2009) Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. Lancet 373:234–239
211. Van Snellenberg JX, de Candia T (2009) Meta-analytic evidence for familial coaggregation of schizophrenia and bipolar disorder. Arch Gen Psychiatry 66:748–755
212. Cardno AG, Rijsdijk FV, Sham PC, Murray RM, McGuffin P (2002) A twin study of genetic relationships between psychotic symptoms. Am J Psychiatry 159:539–545
213. Peerbooms OL, van Os J, Drukker M, Kenis G, Hoogveld L, de Hert M et al (2010) Meta-analysis of MTHFR gene variants in schizophrenia, bipolar disorder and unipolar depressive disorder: evidence for a common genetic vulnerability? Brain Behav Immun Dec 24 [Epub ahead of print]
214. Gonzales ML, LaSalle JM (2010) The role of MeCP2 in brain development and neurodevelopmental disorders. Curr Psychiatry Rep 12:127–134
215. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet 23:185–188
216. Klauck SM, Lindsay S, Beyer KS, Splitt M, Burn J, Poustka A (2002) A mutation hot spot for nonspecific X-linked mental retardation in the MECP2 gene causes the PPM-X syndrome. Am J Hum Genet 70:1034–1037
217. Couvert P, Bienvenu T, Aquaviva C, Poirier K, Moraine C, Gendrot C et al (2001) MECP2 is highly mutated in X-linked mental retardation. Hum Mol Genet 10:941–946
218. Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I et al (2009) Common variants on chromosome 6p22.1 are associated with schizophrenia. Nature 460:753–757
219. Uddin M, Koenen KC, Aiello AE, Wildman DE, de Los SR, Galea S (2011) Epigenetic and inflammatory marker profiles associated with depression in a community-based epidemiologic sample. Psychol Med 41(5):1–11
220. Uddin M, Aiello AE, Wildman DE, Koenen KC, Pawelec G, de Los Santos R et al (2010) Epigenetic and immune function profiles associated with posttraumatic stress disorder. Proc Natl Acad Sci USA 107:9470–9475
221. Ernst C, Chen ES, Turecki G (2009) Histone methylation and decreased expression of TrkB. T1 in orbital frontal cortex of suicide completers. Mol Psychiatry 14:830–832
222. Abdolmaleky HM, Cheng KH, Faraone SV, Wilcox M, Glatt SJ, Gao F et al (2006) Hypomethylation of MB-COMT promoter is a major risk factor for schizophrenia and bipolar disorder. Hum Mol Genet 15:3132–3145

223. Grayson DR, Jia X, Chen Y, Sharma RP, Mitchell CP, Guidotti A et al (2005) Reelin promoter hypermethylation in schizophrenia. Proc Natl Acad Sci USA 102:9341–9346
224. Dempster EL, Mill J, Craig IW, Collier DA (2006) The quantification of COMT mRNA in post mortem cerebellum tissue: diagnosis, genotype, methylation and expression. BMC Med Genet 7:10
225. Mill J, Tang T, Kaminsky Z, Khare T, Yazdanpanah S, Bouchard L et al (2008) Epigenomic profiling reveals DNA methylation changes associated with major psychosis. Am J Hum Genet 82:696–711
226. Tochigi M, Iwamoto K, Bundo M, Komori A, Sasaki T, Kato N et al (2008) Methylation status of the reelin promoter region in the brain of schizophrenic patients. Biol Psychiatry 63:530–533
227. Veldic M, Guidotti A, Maloku E, Davis JM, Costa E (2005) In psychosis, cortical interneurons overexpress DNA-methyltransferase 1. Proc Natl Acad Sci USA 102:2152–2157
228. Kuratomi G, Iwamoto K, Bundo M, Kusumi I, Kato N, Iwata N et al (2007) Aberrant DNA methylation associated with bipolar disorder identified from discordant monozygotic twins. Mol Psychiatry 13:429–441
229. Mastroeni D, Grover A, Delvaux E, Whiteside C, Coleman PD, Rogers J (2008) Epigenetic changes in Alzheimer's disease: decrements in DNA methylation. Neurobiol Aging 12:2025–2037
230. Yoshikai S, Sasaki H, Doh-ura K, Furuya H, Sakaki Y (1990) Genomic organization of the human amyloid beta-protein precursor gene. Gene 87:257–263
231. Tohgi H, Utsugisawa K, Nagane Y, Yoshimura M, Ukitsu M, Genda Y (1999) The methylation status of cytosines in a tau gene promoter region alters with age to downregulate transcriptional activity in human cerebral cortex. Neurosci Lett 275:89–92
232. Barrachina M, Ferrer I (2009) DNA methylation of Alzheimer disease and tauopathy-related genes in postmortem brain. J Neuropathol Exp Neurol 68(8):880–891
233. Wang SC, Oelze B, Schumacher A (2008) Age-specific epigenetic drift in late-onset Alzheimer's disease. PLoS One 3:e2698
234. Silva PN, Gigek CO, Leal MF, Bertolucci PH, de Labio RW, Payao SL et al (2008) Promoter methylation analysis of SIRT3, SMARCA5, HTERT and CDH1 genes in aging and Alzheimer's disease. J Alzheimers Dis 13:173–176
235. Michels KB (2010) The promises and challenges of epigenetic epidemiology. Exp Gerontol 45:297–301
236. Kessler RC, Ormel J, Petukhova M, McLaughlin KA, Green JG, Russo LJ et al (2011) Development of lifetime comorbidity in the world health organization world mental health surveys. Arch Gen Psychiatry 68:90–100
237. McMillan KA, Enns MW, Cox BJ, Sareen J (2009) Comorbidity of axis I and II mental disorders with schizophrenia and psychotic disorders: findings from the national epidemiologic survey on alcohol and related conditions. Can J Psychiatry 54:477–486
238. Hanssen M, Peeters F, Krabbendam L, Radstake S, Verdoux H, Van Os J (2003) How psychotic are individuals with non-psychotic disorders? Soc Psychiatry Psychiatr Epidemiol 38:149–154
239. Weiser M, Reichenberg A, Rabinowitz J, Knobler HY, Lubin G, Yazvitzky R et al (2004) Cognitive performance of male adolescents is lower than controls across psychiatric disorders: a population-based study. Acta Psychiatr Scand 110:471–475
240. Huang J, Perlis RH, Lee PH, Rush AJ, Fava M, Sachs GS et al (2010) Cross-disorder genomewide analysis of schizophrenia, bipolar disorder, and depression. Am J Psychiatry 167:1254–1263
241. Argyropoulos SV, Landau S, Kalidindi S, Toulopoulou T, Castle DJ, Murray RM et al (2008) Twins discordant for schizophrenia: psychopathology of the non-schizophrenic co-twins. Acta Psychiatr Scand 118:214–219
242. Bjornsson HT, Fallin MD, Feinberg AP (2004) An integrated epigenetic and genetic approach to common human disease. Trends Genet 20:350–358

243. Foley DL, Craig JM, Morley R, Olsson CA, Dwyer T, Smith K et al (2009) Prospects for epigenetic epidemiology. Am J Epidemiol 169:389–400
244. Wong CC, Caspi A, Williams B, Craig IW, Houts R, Ambler A et al (2010) A longitudinal study of epigenetic variation in twins. Epigenetics 5(6):516–526
245. Pidsley R, Mill J (2011) Epigenetic studies of psychosis: current findings, methodological approaches, and implications for postmortem research. Biol Psychiatry 69:146–156
246. Andersen SL, Teicher MH (2008) Stress, sensitive periods and maturational events in adolescent depression. Trends Neurosci 31:183–191
247. Davey Smith G, Ebrahim S, Lewis S, Hansell AL, Palmer LJ, Burton PR (2005) Genetic epidemiology and public health: hope, hype, and future prospects. Lancet 366:1484–1498
248. McGorry PD, Nelson B, Amminger GP, Bechdolf A, Francey SM, Berger G et al (2009) Intervention in individuals at ultra-high risk for psychosis: a review and future directions. J Clin Psychiatry 70:1206–1212
249. Albert MS, Blacker D (2006) Mild cognitive impairment and dementia. Annu Rev Clin Psychol 2:379–388
250. Nestler EJ, Hyman SE (2010) Animal models of neuropsychiatric disorders. Nat Neurosci 13:1161–1169
251. Chouliaras L, van den Hove DL, Kenis G, Keitel S, Hof PR, van Os J, Steinbusch HW, Schmitz C, Rutten BP (2011) Neurobiol Aging. 2011 Jul 14. [Epub ahead of print]

Chapter 19
Epigenetic Epidemiology of Type 1 Diabetes

Amanda J. MacFarlane

Contents

19.1 Overview of Type 1 Diabetes ... 378
19.2 Epigenetic Regulation of Gene Expression and Its Potential Involvement in Autoimmune Diabetes ... 380
19.2.1 Autoimmunity ... 381
19.2.2 β-Cell ... 384
19.3 The Diabetic State, Homocysteine Metabolism and Cellular Methylation Potential ... 385
19.3.1 Glucose and Insulin Regulate Homocysteine Metabolism and Cellular Methylation Potential ... 385
19.3.2 Evidence That the Diabetic State Modifies Gene Expression ... 387
19.4 Strategies to Study the Epigenetic Epidemiology of Type 1 Diabetes ... 388
19.4.1 Prospective Studies of Genetically At-Risk Children ... 389
19.4.2 Monozygotic Twin Studies ... 389
19.4.3 Comparative Studies Examining the Effect of *In Utero* and Neonatal Exposures ... 391
19.4.4 Complications of Type 1 Diabetes ... 393
19.5 Conclusion ... 393
References ... 394

Abstract Type 1 diabetes (T1D) results from the immune-mediated destruction of insulin-secreting β-cells that reside in the pancreatic Islets of Langerhans. Genetic susceptibility is necessary but not sufficient for the development of autoimmune diabetes, indicating a key role for risk modification by environmental factors. Epigenetic mechanisms could mediate the effect of specific environmental factors

A.J. MacFarlane (✉)
Department of Biology, Faculty of Science, Carleton University, Ottawa, ON, Canada

Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

Nutrition Research Division, Food Directorate, Health Canada, 251 Sir Frederick Banting Drive, Ottawa, ON, K1A 0K9, Canada
e-mail: Amanda.macfarlane@hc-sc.gc.ca

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_19,
© Springer Science+Business Media B.V. 2012

and could therefore explain, at least in part, non-genetic susceptibility to Type 1 diabetes. Exposure to variable nutrition or infection in early life, including *in utero* experiences can modify T1D susceptibility, which may occur via epigenetic means. T1D is characterized by autoimmune destruction of pancreatic insulin-secreting β-cells and their aberrant development and function of β-cells; these pathogenic mechanisms are also subject to epigenetic regulation. Furthermore, the diabetic state, characterized by fluctuations in insulin and glucose, can influence genomic methylation profiles and gene expression, which can potentially impact susceptibility to co-morbidities of T1D. This chapter identifies potential epigenetic mechanisms that may modify T1D risk and describes study designs that can be implemented to determine the role played by epigenetics in disease pathogenesis.

Abbreviations

AdoMet	S-adenosylmethionine
AdoHcy	S-adenosylhomocysteine
BMI	Body mass index
CBS	Cystathionine β-synthase
ChIP	Chromatin immunoprecipitation
DAISY	Diabetes Autoimmunity Study in the Young
DIPP	Diabetes Prediction and Prevention
DNMT	DNA methyltransferase
HLA	Human leukocyte antigen
IFN	Interferon
IL	Interleukin
LFA	Lymphocyte function-associated antigen
MHC	Major histocompatibility complex
MS	Methionine synthase
NF-κB	Nuclear factor-κB
Seq	Sequencing
SNP	Single nucleotide polymorphism
TCR	T cell receptor
TEDDY	The Environmental Determinants of Diabetes in the Young
T_H	T helper
T1D	Type 1 diabetes
T_{reg}	T regulatory

19.1 Overview of Type 1 Diabetes

Type 1 diabetes is a complex autoimmune disease involving the interaction of numerous genes and environmental factors (Fig. 19.1). It results from the immune-mediated destruction of the insulin-secreting β-cells that reside in the pancreatic

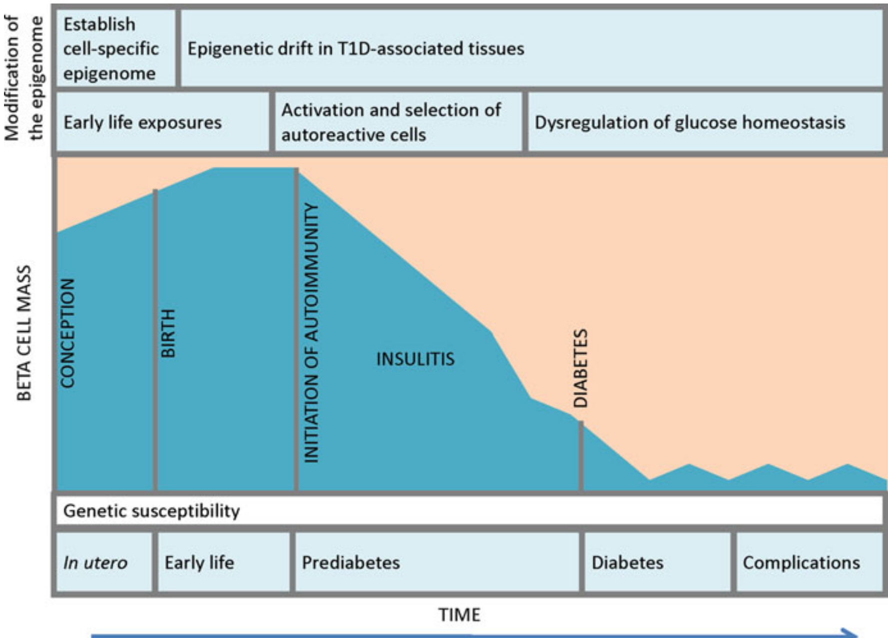


Fig. 19.1 Pathogenesis of Type 1 diabetes and possible role of epigenetic mechanisms. Type 1 diabetes (T1D) is a complex autoimmune disease involving the interplay of numerous genes and environmental factors, which may be regulated by epigenetic mechanisms. Genetically susceptible individuals are prone to the loss of self tolerance resulting in initiation, selection and progression of autoimmunity and insulitis, which culminates in the reduction of β cell mass, loss of glucose homeostasis and the requirement for exogenous insulin. Epigenetic modification of disease susceptibility may be dependent on timing as well as on the tissues involved. There are critical windows in development during which epigenetic marks are erased and re-established, making the epigenome of the developing embryo particularly vulnerable to changes in the intrauterine environment. Altered establishment of the epigenome may result in aberrant pancreas development and function making it more vulnerable to autoimmune attack. Early life experiences, such as the introduction of dietary components and exposure to infectious agents, mould the maturing immune system, a process regulated and perpetuated by epigenetic mechanisms. The epigenome of immune or β cells may also impact the timing of disease onset and disease severity. Finally, fluctuations in circulating glucose and insulin concentrations due to the diabetic state may modify homocysteine metabolism and cellular methylation potential in various tissues resulting in altered susceptibility to the development of T1D-associated complications (Modified from [10])

Islets of Langerhans (reviewed in [1]). Ultimately, the loss of β-cell mass becomes critical and glucose homeostasis is lost, resulting in lifetime dependence of the individual on exogenous insulin. Diabetes is associated with increased risk for a number of co-morbidities including heart disease, kidney failure, blindness, and limb amputation, which culminate in a decreased life expectancy.

A number of linkage and genome-wide association studies have identified the major risk loci for Type 1 diabetes. The DR3, DR4 and DQ2 susceptibility loci

are located within the major histocompatibility complex (MHC, also referred to as the human leukocyte antigen, HLA) class II region and can account for more than ~40% of increased risk [1, 2]. Overall, more than 40 risk loci have been associated with T1D, the majority of which are common polymorphisms (minor allele frequency of >10% in Western populations) in non-MHC genes that display odds ratios ≤1.3 [3, 4].

While underlying genetic susceptibility is necessary, it is not sufficient for the development of autoimmune diabetes leaving environmental factors as key contributors to disease risk. The incidence of T1D has increased more than 2–3 fold over the past half century, a pace that has not been paralleled by an increase in the frequency of the major risk alleles [3, 4]. In fact, the prevalence of the major MHC class II susceptibility loci appears to be decreasing among diabetic patients suggesting that the rise in incidence is due to the impact of environmental factors on individuals with lower-risk genotypes [5, 6]. Perhaps the most compelling indication that environment plays a significant role in diabetes development is the lack of 100% concordance among monozygotic (MZ) twins. The concordance rate among MZ twins is estimated to be between 25% and 65% [7–9]. A number of environmental factors have been implicated in the pathogenesis of autoimmune diabetes, including exposures to various dietary components and infectious agents [10, 11]. However, like the genetic side of the susceptibility coin, no single environmental agent has been found to explain all cases of the disease. As in many complex traits, autoimmune diabetes likely develops as a result of a variety of intricate interactions among many genetic loci and exposure to common environmental factors.

19.2 Epigenetic Regulation of Gene Expression and Its Potential Involvement in Autoimmune Diabetes

Epigenetics may explain, at least in part, non-genetic susceptibility to Type 1 diabetes (reviewed in [12]). Epigenetic mechanisms could act as the conduit by which some environmental factors induce variable phenotypes from identical genotypes. Of note, epigenetic modifications can occur at the level of the cell, tissue or whole organism and have the potential to be stable and mitotically heritable [13–15].

Epigenetic regulation of gene expression occurs by many mechanisms including DNA methylation, post-translational modifications of histones and the activation of microRNAs [16]. The cytosine in CpG dinucleotides in the genome can be methylated to form 5-methylcytosine. Multiple CpG dinucleotides, so called CpG islands, are often observed in the regulatory regions of genes and their methylation status influences gene expression. Heavily methylated CpG islands are often associated with gene silencing [17]. Second, post-translational covalent histone modifications, including acetylation, methylation, phosphorylation and ubiquitination also regulate gene silencing or expression within their associated genomic region(s) [18]. For example, lysine acetylation of histones H3 and H4 is commonly associated with an open

chromatin conformation that allows for active gene expression, such as the methylation of H3 lys4 (H3K4) and lys36 (H3K36). However, H3 and H4 hypoacetylation, di- and trimethylation of H3 lys9 (H3K9), trimethylation of H3 lys27 (H3K27) and methylation of H4 lys20 (H4K20) are associated with gene silencing [18]. A third mechanism involves the expression of short non-coding RNAs, the expression of which can lead to translational silencing through the specific binding and eventual degradation of transcribed RNA [19].

Despite general acceptance that environment plays a significant role in T1D [10, 11], with the exception of a few studies, the involvement of epigenetic mechanisms in T1D pathogenesis has gone essentially unstudied. Epigenetic mechanisms could play a role in the initiation or progression of autoimmunity, or alter the target tissue in such a way as to increase the likelihood that it will be targeted by an autoimmune attack. Here I propose potential epigenetic mechanisms that could be involved in T1D pathogenesis, identify a number of candidate epigenetically-regulated genes and suggest study designs that could be implemented to clarify the T1D-susceptibility epigenome.

19.2.1 Autoimmunity

T1D results from a T cell-mediated autoimmune attack on pancreatic β-cells. CD4+ T helper (T_H) cells, major protagonists involved in β-cell autoimmunity, originate from naïve $CD4^+$ T cells that differentiate into one of three main lineages, T_H1, T_H2 or T_H17, depending on extrinsic signals [20–22]. The transformation from naive to mature cells is a mitotically heritable phenotype that is passed on to all daughter cells and it is subject to complex epigenetic control (reviewed in [23–25]). A number of epigenetically-regulated genes associated with the immune response are presented in Table 19.1.

The differentiation of naïve T cells into T_H cell lineages is driven by the expression of specific cytokines. The treatment of thymoma cells with 5-azacytidine, which induces DNA hypomethylation, resulted in clones that constitutively produced interleukin (IL)-2 [26], suggesting that DNA methylation patterns were important for regulated gene expression. 5-Azacytidine treatment of stimulated naïve T_H cells was also associated with increased interferon (IFN)-γ and IL-4 expression [27], the main effector cytokines of T_H1 and T_H2 cells, respectively. The IFN-γ and IL-4 loci undergo extensive DNA and histone methylation changes that correlate with their respective gene expression during T_H1 and T_H2 cell differentiation [28–36]. Also, genes encoding T cell lineage specific transcription factors are targeted by particular epigenetic modifications [34]. Therefore, the specificity and balance of cytokine expression in differentiating naïve T_H cells is controlled at least in part at the level of the epigenome.

Interestingly, modification of methylation profiles in $CD4^+$ T cells can induce autoreactivity. Normally $CD4^+$ T cells do not react to self antigen; however, when

Table 19.1 Epigenetically-regulated genes involved in the immune response or β-cell development and function

Function/Tissue	Gene product	Epigenetic mechanism	Evidence	References
Immunity	IFN-γ	DNA methylation	IFN-gamma promoter undergoes differential methylation during in vitro T cell differentiation: hypermethylated in Th2 cells, hypomethylated in Th1 cells	[27, 35, 36]
	IFN-γ	Histone modification	Histone modifications characteristic of relaxed, accessible chromatin, observed at *Ifng* locus in T_H1 cells, and selective T-bet binding and activity, a key transcription factor associated with T_H1 pro-inflammatory responses	[34]
	IL-4	DNA methylation	Hypermethylated 5′ region of *Il4* locus becomes demethylated when naive $CD4^+$ T cells are stimulated to undergo T cell differentiation, demethylation associated with enhanced gene expression	[27, 28, 30]
	IL-4	Histone modification	H3K27me3 histone modification at *Il4* locus associated with gene suppression	[34]
	Foxp3	DNA methylation	Transcriptionally active region in *Foxp3* locus demethylated at CpG motifs in $Foxp3^+CD25^+CD4^+$ T_{reg} cells	[41, 44–46]
	Foxp3	Histone modification	H3 acetylation and trimethylation patterns indicative of open euchromatin observed at *Foxp3* locus region, associated with T_{reg} differentiation	[44]
	LFA-1	DNA methylation	Hypomethylation and upregulation of *Itgal* gene expression results in antigen-independent co-stimulation promoting T cell activation	[37–40]
	Il-17	Histone modification	Loss of the histone methyltransferase G9a during T helper cell differentiation results in loss of H3K9me2 repressive marks at the *Il17a* locus and increased expression	[116]
β-cell development and function	Insulin	Histone modification	Proximal promoter of *Ins1/2* gene displays H3 hyperacetylation and hypermethylation at H3K4, both of which are associated with euchromatin status and active gene transcription in pancreatic β-cells. Dimethylation status at H3K4 in the regulatory region of *Ins1/2* maintained by a transcriptional protein complex involving Pdx1, a key transcription factor involved in β-cell development, and the methyltransferase Set7/9, promoted continuous activation of *Ins1/2* and other genes involved in glucose-stimulated insulin secretion.	[54–58]
	PPARGC1A	DNA methylation	PPARGC1A is a transcriptional regulator associated with insulin secretion from islets. Its promoter region was hypemethylated in Type 2 diabetes patients and associated with reduced gene expression and insulin secretion	[117]
	miR-375	miRNA	Pancreatic islet specific expression inhibits insulin secretion in mouse β-cells	[59]
	miR-9	miRNA	Increased expression related to defective glucose-stimulated insulin secretion	[60]
	INK4a/Arf	Histone modification	Increased expression is associated with reduced β-cell regeneration, repression of expression in β-cells is regulated by Ezh2, a histone methyltransferase, and is dependent on H3K27me3 in the gene region	[118]

treated with 5-azacytidine, they lose the requirement for specific antigen stimulation and become activated in response to self MHC class II molecules [37]. The mechanism for this loss of self tolerance has been ascribed to the demethylation and activation of the *ITGAL* gene [38, 39]. The *ITGAL* gene product, lymphocyte function-associated antigen (LFA)-1, stabilizes T cell receptor-MHC interactions and provides co-stimulatory signals for the activation of T cell responses [40]. Its dysregulated overexpression is hypothesized to stabilize lower affinity TCR-MHC interactions and provide co-stimulation, which lead to autoimmunity similar to that observed in systemic lupus erythematosus. The data clearly indicate that changes to the methylation status of immune-activating genes can lead to the promotion of an autoimmune response by allowing indiscriminate interactions between T cells and antigen presenting cells.

The maturation and function of T regulatory (T_{reg}) cells, which are required for the maintenance of self tolerance [41], is also regulated by epigenetic mechanisms [42]. The transcription factor Foxp3 drives T_{reg} cell differentiation and its loss in mice results in the failure to generate $CD25^{+}CD4^{+}$ T_{reg} cells and the development of severe autoimmunity [43]. A conserved and transcriptionally active region within the *Foxp3* locus demonstrated epigenetic patterns associated with open and active euchromatin. Specifically, CpG motifs were demethylated and histone H3 was acetylated at the *Foxp3* locus, which correlated with gene expression in naïve $Foxp3^{+}CD25^{+}CD4^{+}$ T_{reg} cells, but not $CD25^{-}CD4^{+}$ T cells [41, 44–46]. Normally T_{reg} cells promote immune tolerance but this function is lost during autoimmunity. Therefore the dysregulation of T_{reg} cells has been suggested to play a role in T1D pathogenesis [42].

Other genes involved in the inflammatory and/or autoimmune process are also epigenetically regulated. These include the major diabetes susceptibility gene CTLA4, TGF-β, p38 mitogen-activated protein kinase, Toll-like receptors, IL-8, macrophage chemoattractant protein-1, and IL-6 [47–49]. One could speculate that the dysregulation of any one or combination of these epigenetically-regulated genes could predispose an individual to mount an autoimmune response or promote the uncontrolled progression of an autoimmune attack. However, this has not been empirically tested.

Only a few studies have examined epigenetically-regulated gene expression in the context of Type 1 diabetes. Abnormal global methylation of $CD4^{+}$ T cells has been observed in patients with latent autoimmune diabetes in adults [119] compared with healthy control subjects. Global methylation was increased compared to control subjects. Concomitantly, Foxp3 expression was decreased due to hypermethylation of its promoter region. Miao et al. [50] observed that lymphocytes from T1D patients demonstrated distinct histone modification patterns in comparison to control subjects. However, an important caveat to these studies is the earlier findings that human THP-1 monocytes cultured in high glucose conditions also revealed alterations in histone lysine methylation of various T1D-related loci [51]. Because blood glucose levels vary considerably in patients with T1D, it is difficult to differentiate whether such epigenetic changes are related to diabetes susceptibility or result from variations in blood glucose in overt diabetes [50].

In fact, cellular methylation potential is significantly influenced by glucose and insulin concentrations, which will be discussed in Sect. 19.3, highlighting the importance of obtaining tissue samples in the pre-diabetic period. Nonetheless, the data suggest that altered epigenetic regulation of gene expression could potentially influence the autoimmune process or impact diabetes-associated complications.

19.2.2 β-Cell

Type 1 diabetes is associated with progressive dysfunction of insulin-secreting β-cells, resulting in part from the dysregulation of epigenetically-regulated gene expression that is required for appropriate β-cell development and function [52, 53] (Table 19.1). For example, insulin is a major autoantigen in T1D and its gene expression is dependent on specific patterns of DNA methylation, histone acetylation and methylation. The expression of insulin mRNA is negatively correlated with DNA methylation of specific CpG islands in its promoter regions in patients with Type 2 diabetes [120]. Specifically, the proximal promoter of the *Ins1/2* gene in pancreatic β-cells displays H3 hyperacetylation and hypermethylation at H3K4, which are associated with euchromatin status and active gene transcription [54, 55]. Dimethylation status at H3K4 in the regulatory region of *Ins1/2* was shown to be maintained by a transcriptional protein complex involving Pdx1, a key transcription factor involved in β-cell development [56], and the methyltransferase Set7/9, which promoted the continuous activation of *Ins1/2* and other genes involved in glucose-stimulated insulin secretion [57, 58]. Also, two microRNAs, miR-375 and -9, have been reported to regulate glucose-stimulated insulin secretion [59, 60]. Aberrant insulin expression due to altered epigenetic regulation of *Ins1/2* could promote or antagonize an autoimmune response.

Specific genomic regions undergo a sequence of histone modifications during pancreatic organogenesis, with specific events occurring over the course of β-cell development. In zebrafish, the loss of DNMT1, which is required for the maintenance of DNA methylation, was associated with increased beta-cell regeneration, implicating DNA methylation as an important regulator during pancreatic development [121]. Also, in the absence of HDAC5 or HDAC9, two class IIa histone deacetylases, is associated with increased numbers of insulin-producing beta-cells [122]. In β-cells from mice, H3K27 trimethylation was increased among regulators of inappropriate developmental fate, allowing for the specific repression of genes that would be detrimental to the development, function and survival of β-cells [61]. This marker of gene repression was also associated with decreased mRNA expression of these factors. Interestingly the majority of the H3K27me3 marks were acquired *de novo* throughout developmental stages, from the multipotent progenitor cell stage through to terminal differentiation. Therefore, changes to histone modifications associated with the insulin gene and other developmental regulators required for β-cell development could modify the maturation or function

of β-cells resulting in predisposition for autoimmune diabetes. Additionally, aberrant gene expression could also result in the targeting of β-cells by the autoimmune response.

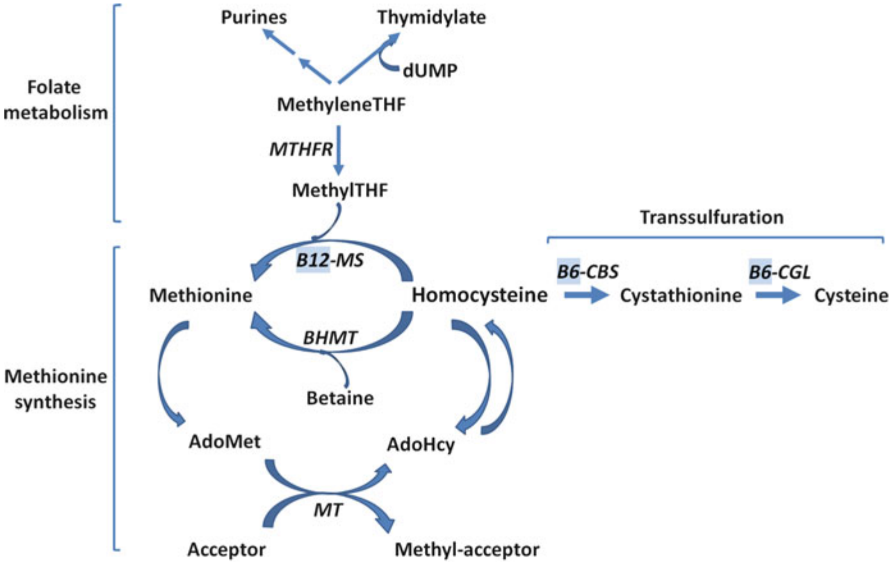
19.3 The Diabetic State, Homocysteine Metabolism and Cellular Methylation Potential

The establishment and maintenance of methylation patterns of CpG dinucleotides in DNA and histones depend on the cellular methylation potential. Homocysteine and methionine metabolism are interdependent and regulate the availability of methyl groups for cellular methylation reactions (Fig. 19.2) [62]. Homocysteine can be remethylated to form methionine, which in turn can be converted to S-adenosylmethionine (AdoMet). AdoMet is the major methyl group donor in cellular transmethylation reactions including the methylation of DNA, histones and other molecules [62]. AdoMet-dependent transmethylation reactions, mediated by methyltransferases, result in the production of S-adenosylhomocysteine (AdoHcy), which can then be converted to homocysteine, completing the methionine cycle. Alternatively, homocysteine can be catabolised by the transsulfuration pathway through its irreversible conversion to cystathionine.

Methyl group and homocysteine metabolism depend on the availability of a number of factors including folate, vitamin B6, vitamin B12, choline, betaine and methionine. Folate, the biologically active form of folic acid, mediates one-carbon metabolism (Fig. 19.2) [62]. Folate carries and activates one-carbon units for their use in the *de novo* synthesis of purines, thymidylate and methionine [63]. Methionine synthase (MS), a vitamin B12-dependent enzyme, transfers the methyl group of 5-methyltetrahydrofolate to homocysteine in its remethylation to form methionine [63]. Folate-mediated one-carbon metabolism is also dependent on vitamin B6, niacin, riboflavin, zinc, cobalt and vitamin A [62]. Folate-independent remethylation of homocysteine utilizes methyl groups donated by betaine, which can be taken up from the diet or derived from choline (Fig. 19.2) [64, 65]. The irreversible transsulfuration of homocysteine by cystathionine β-synthase (CBS) to cystathionine is vitamin B6-dependent (Fig. 19.2). Therefore, nutrient deficiencies or single nucleotide polymorphisms (SNPs) in enzymes involved in various aspects of methionine and homocysteine metabolism can influence cellular methylation potential.

19.3.1 Glucose and Insulin Regulate Homocysteine Metabolism and Cellular Methylation Potential

In the context of studying the epigenetic epidemiology of T1D, it is important to note that the diabetic physiological state can modify cellular methylation potential by altering methyl group metabolism. Specifically, it has been clearly demonstrated



↑ Glucose ∝ ↓ Homocysteine	Insulin treatment
-↑ flux through transsulfuration -↑ flux through remethylation -Potential altered methylation capacity	-normalizes homocysteine -normalizes transsulfuration

Fig. 19.2 Homocysteine and methionine metabolism. Cellular methylation potential is dependent on the production of S-adenosylmethionine (*AdoMet*), the major methyl donor in the cell. Homocysteine can be remethylated to form methionine. A methyl group can be transferred by the B12-dependent enzyme methionine synthase (*MS*) from 5-methyltetrahydrofolate (*THF*), the product of methyleneTHF reductase and the folate metabolic pathway. In the liver and kidney, homocysteine can also be remethylated by betaine:homocysteine methyltransferase (*BHMT*) using betaine as the methyl donor. Betaine can be derived directly from the diet or from choline. Methionine is converted to AdoMet, which is used by methyltransferases (*MT*) to transfer methyl groups to a variety of acceptor molecules, including DNA and histones. The reaction produces S-adenosylhomocysteine (*AdoHcy*) that can be converted to homocysteine. Alternatively, homocysteine can be irreversibly metabolized to cysteine by the B6-dependent enzymes cystathionine-β-synthase (*CBS*) and cystathionine-γ-lyase (*CGL*) in the transsulfuration pathway. Vitamin co-factors B6 and B12 are highlighted. Circulating homocysteine in T1D patients is lower than in non-diabetic control subjects as a result of increased flux through the transsulfuration pathway, and also potentially by increased flux through remethylation. Insulin treatment normalizes homocysteine by decreasing its metabolism by transsulfuration

that fluctuations in glucose and insulin concentrations, as observed in diabetes, perturbs homocysteine and methionine homeostasis.

Glucose regulates homocysteine metabolism. The concentration of circulating homocysteine is inversely correlated with blood glucose concentrations in diabetic

patients without renal dysfunction [66–68] and rodent models of diabetes [69]. The activities of enzymes involved in homocysteine transsulfuration, including CBS and cystathionine γ-lyase, are also increased in the liver of rats with uncontrolled diabetes [69] and CBS activity positively correlates with glucose concentration in cultured hepatocytes [70]. Chiang et al. suggested that homocysteine clearance under high glucose conditions in cultured hepatocytes could also be attributed to increased homocysteine remethylation flux in a methionine-dependent fashion [71]. High glucose levels were associated with increased cellular methylation potential, as indicated by the AdoMet to S-adenosylhomocysteine (AdoHcy) ratio, increased AdoMet synthase activity and DNA methyltransferase activity and global methylation potential in cultured hepatocytes [71]. The cumulative data suggest that the observed diabetes-related reduction in circulating homocysteine is due to its catabolism by the transsulfuration pathway and potentially by increased rates of its remethylation. The overall result is that the diabetic state impacts cellular methylation potential.

Conversely, insulin normalizes circulating homocysteine in T1D. Insulin treatment normalized the rates of homocysteine remethylation, methionine synthesis and homocysteine transsulfuration in T1D patients [71–73]. Similarly, insulin treatment of rats with streptozotocin-induced diabetes resulted in a normalization of circulating homocysteine [69, 74]. Insulin treatment was associated with decreased activities of transsulfuration enzymes which resulted in a dose-dependent increase in plasma homocysteine [75]. In addition, the flux through the transsulfuration pathway was reduced in hepatocytes cultured with insulin [71]. Insulin also induced folate-dependent homocysteine remethylation under methionine-restricted and normoglycemic conditions in hepatic cells [71]. Together the data suggest that insulin substitution normalizes homocysteine concentrations by reducing the capacity to eliminate it by transsulfuration.

19.3.2 Evidence That the Diabetic State Modifies Gene Expression

Fluctuations in insulin and glucose regulate homocsyteine metabolism and impact cellular methylation potential. Therefore, the diabetic physiological state can potentially impact cellular methylation reactions and methylation-dependent epigenetically-regulated gene expression, a concept for which evidence is accumulating. High glucose concentrations have been associated with dysregulation of methyltransferase activity and expression, global DNA methylation and gene expression [71]. Uncontrolled diabetes in streptozotocin-treated rats was associated with altered homocysteine metabolism, increased expression of hepatic DNA methyltransferase (DNMT) 1 and global hypomethylation of genomic DNA [76]. El-Osta et al. found that transient exposure of aortic endothelial cells to high glucose resulted in stable activating epigenetic changes in the promoter of the pro-inflammatory

nuclear factor κB (NF-κB) subunit p65 and long-lasting activation of its expression [77, 78]. Specifically, transient high glucose conditions resulted in persistent mobilization of the histone methyltransferase Set7 and monomethylation of H3K4 at the proximal promoter of p65 [77]. Also, transient glucose exposure was associated with recruitment of the histone demethylase LSD1 and the loss of suppressive H3K9m2 and H3K9m3 methylation marks at the p65 promoter [78]. Furthermore, cardiomyocytes cultured in high glucose demonstrate decreased IGF-1 receptor expression and were more susceptible to apoptosis. The IGF-1 receptor expression was associated with decreased association of acetylated H4 with the IGF-1 receptor promoter region and consequent reduced promoter activity [123]. In addition to diabetes-dependent effects on cellular methylation potential, new evidence has linked glucose derived acetyl-CoA, the substrate of histone acetylases, to modifications of chromatin structure and global gene expression [79, 80]. Many of these studies focused on the consequences of the diabetic state in tissues related to complications of diabetes rather than those directly involved in disease pathogenesis, leaving the role played by epigenetic mechanisms on diabetes initiation or progression unclear.

19.4 Strategies to Study the Epigenetic Epidemiology of Type 1 Diabetes

A number of strategies at the population level have been employed to successfully identify genetic risk factors associated with Type 1 diabetes (reviewed in [81]). Linkage studies using pairs of affected relatives, usually siblings, allow for the identification of genomic regions associated with risk, but the results tend to be limited to loci that have a large effect. Genome-wide association studies using case–control or family-based designs are able to identify loci with a small or moderate impact on susceptibility. The addition of high-throughput sequencing and SNP genotyping technologies have aided significantly in the identification of candidate risk genes.

Similar population sampling strategies in combination with genome-wide high-throughput screening methods can be implemented in the search for epigenetically regulated susceptibility loci. The "methylome", or genome-wide methylation status of an individual can be queried using array-based or genome-wide methods (reviewed in [82, 83]). Array-based methods, including SNP arrays and high density promoter and CpG island arrays, have been used extensively but are limited in terms of genome coverage and their restriction to specific gene promoters and CpG islands. Alternatively, high-throughput sequencing methods allow for genome-wide analysis of the methylome but are currently limited by their high cost. Genomic DNA associated with specific histone modifications can also be determined using chromatin immunoprecipitation (ChIP) assays followed by microarray (ChIP-on-chip) or direct sequencing of bound DNA (ChIP-Seq). Use of these methods, or a combination

thereof, will allow for the determination of the diabetes-risk epigenome. The key to any study will be the population from which the samples are drawn, the tissues that are sampled and how the data are interpreted.

19.4.1 Prospective Studies of Genetically At-Risk Children

The diabetic state is characterized by fluctuations in insulin and glucose, which can influence cellular methylation potential, global methylation profiles and gene expression. As such, epigenetic changes observed in overt diabetic patients could be related to the diabetic state rather than diabetes risk, making the interpretation of differential epigenetic modifications difficult. Therefore, it will be important to study genetically at-risk, normoglycemic individuals before the initiation of the autoimmune process to elucidate the pre-existing risk epigenome.

A number of prospective cohort studies of genetically at-risk children have already been implemented to determine environmental risk factors for Type 1 diabetes. These include the BABYDIAB study in Germany [84], the Diabetes Autoimmunity Study in the Young (DAISY) in the USA [85], the Finnish Diabetes Prediction and Prevention (DIPP) study [86] and The Environmental Determinants of Diabetes in the Young (TEDDY) study in the USA and Europe [87]. These studies have examined the relation among genetics, *in utero* exposures, diet and infectious agents, to name but a few, in the development of autoimmunity, as indicated by the presence of autoantibodies and progression to overt diabetes. It is easy to conceive that these studies or ones modeled after them could also identify diabetes-associated epigenetic modifications and their relation with various environmental factors. Ideally, sampling would occur in the prediabetic state, as well as during the progression to disease. This would allow for the epigenetic features of diabetes risk to be differentiated from those of the diabetic state, as well as the determination of key epigenetic events that occur during disease pathogenesis. Also, the identification of epigenetic changes due to active disease will be important as they could play a causal role in the development of diabetes complications, including cardiovascular disease. Furthermore, it will be imperative to collect detailed information on exposures to various environmental factors to begin to understand the association of specific factors with changes to the epigenome.

19.4.2 Monozygotic Twin Studies

As previously mentioned, the strongest evidence that environmental factors modify T1D is the lack of 100% concordance in twin pairs. Studies in MZ twins suggest that while genetics play a significant role in the development of autoimmune T1D, it is not sufficient to explain 35–75% of cases [7–9]. Some differences in disease concordance may be explained by environment-dependent differences in the epigenome

between twin pairs. Indeed, DNA methylation content and histone modifications can vary significantly in genetically identical twins. Significant differences in 5-methylcytosine content and H3 and H4 acetylation patterns of genomic DNA were observed in up to 35% of monozygotic twin pairs ranging in age from 3 to 74 years [88] .The epigenetic differences also correlated with age and time spent together, with younger twin pairs being more similar than older twin pairs, and those who had spent more time together being more similar than those who had been separated. Of note, the epigenetic patterns correlated with changes in gene expression [88]. Interestingly, even newborn MZ twins demonstrated variable methylation profiles at differentially methylated regions, such as the *IGF2/H19* locus, indicating the importance of the intrauterine environment in the establishment of the epigenome [89]. Another study demonstrated that the epigenetic profile of an individual can change over time, the concept of epigenetic drift, as demonstrated by changes in global DNA methylation. In two samples taken 11–16 years apart, more than 25% of individuals showed a 10% or greater difference in global DNA methylation [90]. Furthermore, Waterland et al. have identified candidate metastable epigenetic loci, the methylation profile of which significantly correlate among tissues within an individual but demonstrated increased inter-individual variation [91]. Of note, MZ twin pairs demonstrated discordant methylation patterns at some of the identified metastable loci, indicating the stochastic nature of the establishment of methylation patterns at these loci [91]. While these early studies relied on small samples sizes and, in some cases, less sophisticated analyses of epigenetic variation, such as global DNA methylation, together they support the concept that the epigenotype of genetically identical individuals can diverge in a time- and environment-dependent manner. In terms of diabetes discordance in monozygotic twins, this epigenetic drift could explain differences in time to disease onset, disease severity and predisposition to secondary complications.

Twin studies will provide a unique opportunity to elucidate the role of epigenetics in autoimmune diabetes, as they have for other autoimmune diseases such as systemic lupus erythematosus [92]. The pairwise examination of discordant twins will eliminate the confounding effect of genetic polymorphisms on the epigenome, which can influence methylation patterns. For example, a SNP in the SUV39H2 histone methyltransferase was shown to be associated with risk for diabetic nephropathy [124]. However, there exist some limitations. Twin studies could be confounded by significant shared environment and experiences, including those *in utero*. The tissues selected for study will also be important. As indicated previously, many genes involved in the autoimmune process are regulated by epigenetic mechanisms, which could be independent of disease susceptibility. By choosing tissues that are not involved in the immune process and by comparing multiple tissues within and between twin pairs it should be possible to minimize the impact of the autoimmune process. However, the diabetic state can also impact on cellular methylation potential and methylation-dependent epigenetic mechanisms, which could be misinterpreted as being associated with disease risk. It will be necessary to include MZ twins in a prospective cohort of genetically at-risk individuals to allow for their assessment in a prediabetic, normoglycemic state.

19.4.3 Comparative Studies Examining the Effect of In Utero and Neonatal Exposures

The epigenome is likely the most plastic and sensitive to environmental-mediated changes during embryonic development. Upon fertilization, gamete-related epigenetic marks are erased and replaced with embryonic marks [93]. Later in fetal development, developing gametes undergo epigenetic reprogramming in which parental marks are replaced by stem cell marks [93]. It is during these periods of flux that the epigenome will be the most susceptible to modifications by nutritional and other environmental factors. In addition, since epigenetic mechanisms regulate many of the cytokines involved in immune responses, including tolerance to dietary antigens and immunity to infections, exposures to environmental factors in the neonatal period can impact the establishment and maturation of the immune system, which will have long term consequences for future immune responses.

19.4.3.1 Gestational Diabetes

Gestational diabetes is diagnosed when an abnormal oral glucose tolerance is observed for the first time during pregnancy (reviewed in [94]). The incidence of gestational diabetes is increasing in parallel with the obesity epidemic, however, 10% of cases are of autoimmune origin. Consequences for offspring affected by gestational diabetes include increased risk for obesity, metabolic syndrome and Type 2 diabetes. However, evidence is mounting that Type 1 diabetes could also be associated with this *in utero* experience [95, 96]. Increased maternal body mass index (BMI) before and excessive weight gain during pregnancy, risk factors for the development of gestational diabetes, are associated with increased risk of islet autoimmunity in genetically at-risk children [97]. Also, children born to mothers with gestational diabetes are more likely to have increased birth weight, faster weight gain and higher BMI, which have been associated with increased risk for islet autoimmunity, progression to T1D and earlier age at disease onset in genetically at-risk children [98–101].

Of note, diabetic pregnancies have a 2–5-fold increased risk of being complicated by a congenital anomaly [102, 103]. Animal studies show that folic acid treatment can prevent gestational diabetes-associated anomalies [104], indicating that diabetes-related altered folate metabolism impacts the developing embryo. As such, altered folate metabolism during pregnancy as a result of gestational diabetes could also impact the cellular methylation potential and epigenome of the offspring. Ideally, prospective samples will be taken from children exposed to gestational diabetes *in utero* beginning at birth to determine the effect of the intrauterine environment on the epigenetic underpinnings of disease risk, as well as the relation with timing and severity of disease development. Information such as maternal BMI, diet, supplement use and other lifestyle factors from the periconceptional period and pregnancy should be collected to allow for associations to be made between *in utero* experiences and epigenetic-dependent risk in the offspring.

19.4.3.2 Maternal and Neonatal Diet

Early diet exposures, including exposures *in utero*, have been associated with the development of Type 1 diabetes in both humans and animals [11]. Whether or not the various dietary manipulations affect epigenetic-dependent disease susceptibility remains to be determined. However, the Agouti mouse model has been exploited to demonstrate an elegant proof of principle that maternal diet during gestation modifies the epigenome [105]. The Agouti mouse has an IAP retrotransposon insertion that contains a cryptic promoter upstream of the *agouti* gene. Hypomethylation of the cryptic promoter results in ectopic expression of the gene and yellow coat color. Conversely, hypermethylation inhibits promoter activity and results in brown coat color. Therefore, coat color acts as a phenotypic proxy for the methylation status of the promoter. Maternal nutrition, specifically the intake of methyl donors during pregnancy, was shown to modify coat color as a result of altered methylation of the IAP transposable element in the offspring [105]. In addition to these early animal studies, the methylation of candidate metastable epialleles in humans was associated with maternal nutritional status during early pregnancy [91].

In relation to T1D, maternal diet during gestation has been associated with the progression of autoimmunity and the development of disease [106]. One example is maternal intake of vitamin D during pregnancy, which has been associated with decreased risk of and increased time to autoantibody development in the offspring [107, 108]. Also, the introduction of diet components and the timing of their introduction in the neonatal period are important in determining the risk for diabetes in genetically at-risk infants. For instance, the timing of exposure to specific cereals has been associated with increased risk for autoantibody development in at-risk children [109, 110]. Other early diet exposures linked to diabetes risk have included duration of breast-feeding and introduction and timing of introduction of cow's milk formula [111].

In addition, heavier babies at birth, a potential indication of over-nutrition during gestation, have a small but significant increased risk for developing T1D [101]. Conversely, animal models demonstrate that maternal under-nutrition during pregnancy is associated with reduced risk for diabetes and abnormal development of the pancreas in offspring. Under-nutrition due to caloric restriction *in utero* reduces the incidence of diabetes in NOD mice [112]. A low protein diet fed during pregnancy and lactation was associated with decreased islet mass and decreased autoimmunity in NOD mice, and decreased islet mass and islet vascularity in rats [113, 114]. The smaller islet size and impaired insulin release observed in the offspring of low protein fed dams could be attributed to changes in the balance of beta cell replication and apoptosis in fetal and neonatal life [115]. It is possible that the changes observed in offspring are due to nutrition-dependent epigenetic modifications in the developing embryo but this relationship requires empirical determination. This can be achieved with the careful selection of study subjects, such as mother's that are themselves at genetic-risk for T1D, as well as detailed information about food intake, nutritional status and supplement use during pregnancy.

19.4.4 Complications of Type 1 Diabetes

Chronic diabetes is associated with a number of complications including diabetic nephropathy, neuropathy, and cardiovascular disease, which diminish quality of life and life expectancy. As previously noted, the diabetic state defined by high glucose and low insulin can modify the cellular methylation potential by altering the flux of homocysteine remethylation and transsulfuration, thereby regulating the availability of AdoMet for cellular methylation reactions (Sect. 19.3, Fig. 19.2). The careful choice of control subjects will be critical for defining the impact of the diabetic state on epigenetic drivers of diabetes complications. It will be necessary to use control subjects with T1D but without the specific co-morbidity in addition to healthy control subjects to differentiate diabetes/dysregulated glucose homeostasis-dependent from diabetes complication-dependent epigenetic signatures. Other critical factors to consider will be age at diagnosis, length of disease and disease severity. A recent study performed a genome-wide CpG island methylation analysis to identify differentially methylated sites associated with diabetic nephropathy. The methylation status of 19 CpG sites were shown to correlate with time to nephropathy, including one CpG site in close proximity to a gene that had previously been associated with nephropathy [125]. The cases and controls were selcted to control for T1D duration, age at diagnosis, age and sex.

19.5 Conclusion

Based on our understanding of disease mechanisms and the growing body of evidence for the involvement of epigenetics in autoimmune and other diseases, it is plausible that epigenetic mechanisms underlie at least a proportion of the risk for Type 1 diabetes. However, very few studies have empirically tested the relationship. A number of epigenetically-regulated genes associated with T1D have been identified, representing a starting point for a candidate gene approach to determine their role in disease pathogenesis. Altered epigenetic regulation could affect T1D susceptibility in a number of ways and could be dependent on timing as well as on the tissues involved. The intrauterine environment, especially during periods of epigenetic flux such as post-fertilization, may be the most important environment to affect the T1D-susceptibility epigenome. Epigenetic mechanisms could act as T1D-susceptibility modifiers during other periods such as pancreas development, immune system maturation, the initiation and progression of the autoimmune process and even once overt diabetes has been established, through the impact of epigenetically regulated genes on the development of T1D-related co-morbidities. As such, studies must be thoughtfully designed to address numerous caveats, including the timing of recruitment and sampling, the collection of data relevant to environmental exposures including diet and infections, and the tissues sampled. Ideally, data will be derived from prospective cohort studies of genetically at-risk individuals using

powerful genome-wide high-throughput technologies. Only then will it be possible to begin to unravel the T1D risk epigenome and the environmental factors that influence it.

Acknowledgments The author would like to thank Alfred Aziz, Fraser Scott and Alex Wong for helpful discussions. This work is funded by Health Canada.

References

1. Bluestone JA, Herold K, Eisenbarth G (2010) Genetics, pathogenesis and clinical interventions in type 1 diabetes. Nature 464:1293–1300
2. Ziegler AG, Nepom GT (2010) Prediction and pathogenesis in type 1 diabetes. Immunity 32:468–478
3. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA et al (2009) Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 41:703–707
4. Patterson CC, Dahlquist GG, Gyurus E, Green A, Soltesz G (2009) Incidence trends for childhood type 1 diabetes in Europe during 1989–2003 and predicted new cases 2005–20: a multicentre prospective registration study. Lancet 373:2027–2033
5. Fourlanos S, Varney MD, Tait BD, Morahan G, Honeyman MC, Colman PG et al (2008) The rising incidence of type 1 diabetes is accounted for by cases with lower-risk human leukocyte antigen genotypes. Diabetes Care 31:1546–1549
6. Gillespie KM, Bain SC, Barnett AH, Bingley PJ, Christie MR, Gill GV et al (2004) The rising incidence of childhood type 1 diabetes and reduced contribution of high-risk HLA haplotypes. Lancet 364:1699–1700
7. Redondo MJ, Jeffrey J, Fain PR, Eisenbarth GS, Orban T (2008) Concordance for islet autoimmunity among monozygotic twins. N Engl J Med 359:2849–2850
8. Redondo MJ, Rewers M, Yu L, Garg S, Pilcher CC, Elliott RB et al (1999) Genetic determination of islet cell autoimmunity in monozygotic twin, dizygotic twin, and non-twin siblings of patients with type 1 diabetes: prospective twin study. BMJ 318:698–702
9. Hyttinen V, Kaprio J, Kinnunen L, Koskenvuo M, Tuomilehto J (2003) Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs: a nationwide follow-up study. Diabetes 52:1052–1055
10. Knip M, Veijola R, Virtanen SM, Hyoty H, Vaarala O, Akerblom HK (2005) Environmental triggers and determinants of type 1 diabetes. Diabetes 54(Suppl 2):S125–S136
11. Lefebvre DE, Powell KL, Strom A, Scott FW (2006) Dietary proteins as environmental modifiers of type 1 diabetes mellitus. Annu Rev Nutr 26:175–202
12. MacFarlane AJ, Strom A, Scott FW (2009) Epigenetics: deciphering how environmental factors may modify autoimmune type 1 diabetes. Mamm Genome 20:624–632
13. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A et al (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454:766–770
14. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z et al (2007) High-resolution profiling of histone methylations in the human genome. Cell 129:823–837
15. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S et al (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 40:897–903
16. Wolffe AP, Matzke MA (1999) Epigenetics: regulation through repression. Science 286:481–486
17. Bird A (2002) DNA methylation patterns and epigenetic memory. Genes Dev 16:6–21
18. Bernstein BE, Meissner A, Lander ES (2007) The mammalian epigenome. Cell 128:669–681
19. Bernstein E, Allis CD (2005) RNA meets chromatin. Genes Dev 19:1635–1655

20. Harrington LE, Hatton RD, Mangan PR, Turner H, Murphy TL, Murphy KM et al (2005) Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. Nat Immunol 6:1123–1132
21. Mosmann TR, Cherwinski H, Bond MW, Giedlin MA, Coffman RL (1986) Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. J Immunol 136:2348–2357
22. Park H, Li Z, Yang XO, Chang SH, Nurieva R, Wang YH et al (2005) A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. Nat Immunol 6:1133–1141
23. Wilson CB, Rowell E, Sekimata M (2009) Epigenetic control of T-helper-cell differentiation. Nat Rev Immunol 9:91–105
24. Sawalha AH (2008) Epigenetics and T-cell immunity. Autoimmunity 41:245–252
25. Aune TM, Collins PL, Chang S (2009) Epigenetics and T helper 1 differentiation. Immunology 126:299–305
26. Ballas ZK (1984) The use of 5-azacytidine to establish constitutive interleukin 2-producing clones of the EL4 thymoma. J Immunol 133:7–9
27. Bird JJ, Brown DR, Mullen AC, Moskowitz NH, Mahowald MA, Sider JR et al (1998) Helper T cell differentiation is controlled by the cell cycle. Immunity 9:229–237
28. Lee DU, Agarwal S, Rao A (2002) Th2 lineage commitment and efficient IL-4 production involves extended demethylation of the IL-4 gene. Immunity 16:649–660
29. Santangelo S, Cousins DJ, Winkelmann NE, Staynov DZ (2002) DNA methylation changes at human Th2 cytokine genes coincide with DNase I hypersensitive site formation during CD4(+) T cell differentiation. J Immunol 169:1893–1903
30. Agarwal S, Rao A (1998) Modulation of chromatin structure regulates cytokine gene expression during T cell differentiation. Immunity 9:765–775
31. Melvin AJ, McGurn ME, Bort SJ, Gibson C, Lewis DB (1995) Hypomethylation of the interferon-gamma gene correlates with its expression by primary T-lineage cells. Eur J Immunol 25:426–430
32. Jones B, Chen J (2006) Inhibition of IFN-gamma transcription by site-specific methylation during T helper cell development. EMBO J 25:2443–2452
33. Makar KW, Wilson CB (2004) DNA methylation is a nonredundant repressor of the Th2 effector program. J Immunol 173:4402–4406
34. Wei G, Wei L, Zhu J, Zang C, Hu-Li J, Yao Z et al (2009) Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4+ T cells. Immunity 30:155–167
35. Yano S, Ghosh P, Kusaba H, Buchholz M, Longo DL (2003) Effect of promoter methylation on the regulation of IFN-gamma gene during in vitro differentiation of human peripheral blood T cells into a Th2 population. J Immunol 171:2510–2516
36. Fitzpatrick DR, Shirley KM, McDonald LE, Bielefeldt-Ohmann H, Kay GF, Kelso A (1998) Distinct methylation of the interferon gamma (IFN-gamma) and interleukin 3 (IL-3) genes in newly activated primary CD8+ T lymphocytes: regional IFN-gamma promoter demethylation and mRNA expression are heritable in CD44(high)CD8+ T cells. J Exp Med 188:103–117
37. Richardson B (1986) Effect of an inhibitor of DNA methylation on T cells. II. 5-Azacytidine induces self-reactivity in antigen-specific T4+ cells. Hum Immunol 17:456–470
38. Richardson BC, Strahler JR, Pivirotto TS, Quddus J, Bayliss GE, Gross LA et al (1992) Phenotypic and functional similarities between 5-azacytidine-treated T cells and a T cell subset in patients with active systemic lupus erythematosus. Arthritis Rheum 35:647–662
39. Lu Q, Ray D, Gutsch D, Richardson B (2002) Effect of DNA methylation and chromatin structure on ITGAL expression. Blood 99:4503–4508
40. Wulfing C, Sumen C, Sjaastad MD, Wu LC, Dustin ML, Davis MM (2002) Costimulation and endogenous MHC ligands contribute to T cell recognition. Nat Immunol 3:42–47
41. Lal G, Zhang N, van der Touw W, Ding Y, Ju W, Bottinger EP et al (2009) Epigenetic regulation of Foxp3 expression in regulatory T cells by DNA methylation. J Immunol 182:259–273

42. Sgouroudis E, Piccirillo CA (2009) Control of type 1 diabetes by CD4+Foxp3+ regulatory T cells: lessons from mouse models and implications for human disease. Diabetes Metab Res Rev 25:208–218
43. Fontenot JD, Gavin MA, Rudensky AY (2003) Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. Nat Immunol 4:330–336
44. Floess S, Freyer J, Siewert C, Baron U, Olek S, Polansky J et al (2007) Epigenetic control of the foxp3 locus in regulatory T cells. PLoS Biol 5:e38
45. Kim HP, Leonard WJ (2007) CREB/ATF-dependent T cell receptor-induced FoxP3 gene expression: a role for DNA methylation. J Exp Med 204:1543–1551
46. Janson PC, Winerdal ME, Marits P, Thorn M, Ohlsson R, Winqvist O (2008) FOXP3 promoter demethylation reveals the committed Treg population in humans. PLoS One 3:e1612
47. Saccani S, Pantano S, Natoli G (2001) Two waves of nuclear factor kappaB recruitment to target promoters. J Exp Med 193:1351–1359
48. Saccani S, Pantano S, Natoli G (2002) p38-Dependent marking of inflammatory genes for increased NF-kappa B recruitment. Nat Immunol 3:69–75
49. Saccani S, Natoli G (2002) Dynamic changes in histone H3 Lys 9 methylation occurring at tightly regulated inducible inflammatory genes. Genes Dev 16:2219–2224
50. Miao F, Smith DD, Zhang L, Min A, Feng W, Natarajan R (2008) Lymphocytes from patients with type 1 diabetes display a distinct profile of chromatin histone H3 lysine 9 dimethylation: an epigenetic study in diabetes. Diabetes 57:3189–3198
51. Miao F, Wu X, Zhang L, Yuan YC, Riggs AD, Natarajan R (2007) Genome-wide analysis of histone lysine methylation variations caused by diabetic conditions in human monocytes. J Biol Chem 282:13854–13863
52. Matveyenko AV, Butler PC (2008) Relationship between beta-cell mass and diabetes onset. Diabetes Obes Metab 10(Suppl 4):23–31
53. Eizirik DL, Colli ML, Ortis F (2009) The role of inflammation in insulitis and beta-cell loss in type 1 diabetes. Nat Rev Endocrinol 5:219–226
54. Chakrabarti SK, Francis J, Ziesmann SM, Garmey JC, Mirmira RG (2003) Covalent histone modifications underlie the developmental regulation of insulin gene transcription in pancreatic beta cells. J Biol Chem 278:23617–23623
55. Mutskov V, Raaka BM, Felsenfeld G, Gershengorn MC (2007) The human insulin gene displays transcriptionally active epigenetic marks in islet-derived mesenchymal precursor cells in the absence of insulin expression. Stem Cells 25:3223–3233
56. Babu DA, Deering TG, Mirmira RG (2007) A feat of metabolic proportions: Pdx1 orchestrates islet development and function in the maintenance of glucose homeostasis. Mol Genet Metab 92:43–55
57. Francis J, Chakrabarti SK, Garmey JC, Mirmira RG (2005) Pdx-1 links histone H3-Lys-4 methylation to RNA polymerase II elongation during activation of insulin transcription. J Biol Chem 280:36244–36253
58. Deering TG, Ogihara T, Trace AP, Maier B, Mirmira RG (2009) Methyltransferase Set7/9 maintains transcription and euchromatin structure at islet-enriched genes. Diabetes 58:185–193
59. Poy MN, Eliasson L, Krutzfeldt J, Kuwajima S, Ma X, Macdonald PE et al (2004) A pancreatic islet-specific microRNA regulates insulin secretion. Nature 432:226–230
60. Plaisance V, Abderrahmani A, Perret-Menoud V, Jacquemin P, Lemaigre F, Regazzi R (2006) MicroRNA-9 controls the expression of Granuphilin/Slp4 and the secretory response of insulin-producing cells. J Biol Chem 281:26932–26942
61. van Arensbergen J, Garcia-Hurtado J, Moran I, Maestro MA, Xu X, Van de Casteele M et al (2010) Derepression of Polycomb targets during pancreatic organogenesis allows insulin-producing beta-cells to adopt a neural gene activity program. Genome Res 20:722–732
62. Fox JT, Stover PJ (2008) Folate-mediated one-carbon metabolism. Vitam Horm 79:1–44
63. Wagner C (1995) Biochemical role of folate in cellular metabolism. In: Bailey LB (ed) Folate in health and disease. Marcell Dekker, Inc., New York, pp 23–42

64. Finkelstein JD, Martin JJ, Harris BJ, Kyle WE (1982) Regulation of the betaine content of rat liver. Arch Biochem Biophys 218:169–173
65. Zeisel SH (2006) Choline: critical role during fetal development and dietary requirements in adults. Annu Rev Nutr 26:229–250
66. Heilman K, Zilmer M, Zilmer K, Kool P, Tillmann V (2009) Elevated plasma adiponectin and decreased plasma homocysteine and asymmetric dimethylarginine in children with type 1 diabetes. Scand J Clin Lab Invest 69:85–91
67. Robillon JF, Canivet B, Candito M, Sadoul JL, Jullien D, Morand P et al (1994) Type 1 diabetes mellitus and homocyst(e)ine. Diabete Metab 20:494–496
68. Mazza A, Bossone E, Mazza F, Distante A (2005) Reduced serum homocysteine levels in type 2 diabetes. Nutr Metab Cardiovasc Dis 15:118–124
69. Jacobs RL, House JD, Brosnan ME, Brosnan JT (1998) Effects of streptozotocin-induced diabetes and of insulin treatment on homocysteine metabolism in the rat. Diabetes 47:1967–1970
70. Dicker-Brown A, Fonseca VA, Fink LM, Kern PA (2001) The effect of glucose and insulin on the activity of methylene tetrahydrofolate reductase and cystathionine-beta-synthase: studies in hepatocytes. Atherosclerosis 158:297–301
71. Chiang EP, Wang YC, Chen WW, Tang FY (2009) Effects of insulin and glucose on cellular metabolic fluxes in homocysteine transsulfuration, remethylation, S-adenosylmethionine synthesis, and global deoxyribonucleic acid methylation. J Clin Endocrinol Metab 94:1017–1025
72. Abu-Lebdeh HS, Barazzoni R, Meek SE, Bigelow ML, Persson XM, Nair KS (2006) Effects of insulin deprivation and treatment on homocysteine metabolism in people with type 1 diabetes. J Clin Endocrinol Metab 91:3344–3348
73. Fonseca V, Dicker-Brown A, Ranganathan S, Song W, Barnard RJ, Fink L et al (2000) Effects of a high-fat-sucrose diet on enzymes in homocysteine metabolism in the rat. Metabolism 49:736–741
74. Ratnam S, Maclean KN, Jacobs RL, Brosnan ME, Kraus JP, Brosnan JT (2002) Hormonal regulation of cystathionine beta-synthase expression in liver. J Biol Chem 277:42912–42918
75. Gursu MF, Baydas G, Cikim G, Canatan H (2002) Insulin increases homocysteine levels in a dose-dependent manner in diabetic rats. Arch Med Res 33:305–307
76. Williams KT, Garrow TA, Schalinske KL (2008) Type I diabetes leads to tissue-specific DNA hypomethylation in male rats. J Nutr 138:2064–2069
77. El-Osta A, Brasacchio D, Yao D, Pocai A, Jones PL, Roeder RG et al (2008) Transient high glucose causes persistent epigenetic changes and altered gene expression during subsequent normoglycemia. J Exp Med 205:2409–2417
78. Brasacchio D, Okabe J, Tikellis C, Balcerczyk A, George P, Baker EK et al (2009) Hyperglycemia induces a dynamic cooperativity of histone methylase and demethylase enzymes associated with gene-activating epigenetic marks that coexist on the lysine tail. Diabetes 58:1229–1236
79. Rathmell JC, Newgard CB (2009) Biochemistry, a glucose-to-gene link. Science 324:1021–1022
80. Wellen KE, Hatzivassiliou G, Sachdeva UM, Bui TV, Cross JR, Thompson CB (2009) ATP-citrate lyase links cellular metabolism to histone acetylation. Science 324:1076–1080
81. Concannon P, Rich SS, Nepom GT (2009) Genetics of type 1A diabetes. N Engl J Med 360:1646–1654
82. Feinberg AP (2010) Genome-scale approaches to the epigenetics of common human disease. Virchows Arch 456:13–21
83. Hurd PJ, Nelson CJ (2009) Advantages of next-generation sequencing versus the microarray in epigenetic research. Brief Funct Genomic Proteomic 8:174–183
84. Ziegler AG, Hummel M, Schenker M, Bonifacio E (1999) Autoantibody appearance and risk for development of childhood diabetes in offspring of parents with type 1 diabetes: the 2-year analysis of the German BABYDIAB Study. Diabetes 48:460–468

85. Barker JM, Goehrig SH, Barriga K, Hoffman M, Slover R, Eisenbarth GS et al (2004) Clinical characteristics of children diagnosed with type 1 diabetes through intensive screening and follow-up. Diabetes Care 27:1399–1404
86. Kimpimaki T, Kupila A, Hamalainen AM, Kukko M, Kulmala P, Savola K et al (2001) The first signs of beta-cell autoimmunity appear in infancy in genetically susceptible children from the general population: the Finnish type 1 diabetes prediction and prevention study. J Clin Endocrinol Metab 86:4782–4788
87. TEDDY Study Group (2008) The environmental determinants of diabetes in the young (TEDDY) study. Ann N Y Acad Sci 1150:1–13
88. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML et al (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 102:10604–10609
89. Ollikainen M, Smith KR, Joo EJ, Ng HK, Andronikos R, Novakovic B et al (2010) DNA methylation analysis of multiple tissues from newborn twins reveals both genetic and intra-uterine components to variation in the human neonatal epigenome. Hum Mol Genet 19:4176–4188
90. Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H et al (2008) Intra-individual change over time in DNA methylation with familial clustering. JAMA 299:2877–2883
91. Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M et al (2010) Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. PLoS Genet 6:e1001252
92. Javierre BM, Fernandez AF, Richter J, Al-Shahrour F, Martin-Subero JI, Rodriguez-Ubreva J et al (2010) Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. Genome Res 20:170–179
93. Morgan HD, Santos F, Green K, Dean W, Reik W (2005) Epigenetic reprogramming in mammals. Hum Mol Genet 14(Spec No 1):R47–R58
94. Kaaja R, Ronnemaa T (2008) Gestational diabetes: pathogenesis and consequences to mother and offspring. Rev Diabet Stud 5:194–202
95. Pirkola J, Vaarasmaki M, Leinonen E, Bloigu A, Veijola R, Tossavainen P et al (2008) Maternal type 1 and gestational diabetes: postnatal differences in insulin secretion in offspring at preschool age. Pediatr Diabetes 9:583–589
96. Ziegler AG, Hillebrand B, Rabl W, Mayrhofer M, Hummel M, Mollenhauer U et al (1993) On the appearance of islet associated autoimmunity in offspring of diabetic mothers: a prospective study from birth. Diabetologia 36:402–408
97. Rasmussen T, Stene LC, Samuelsen SO, Cinek O, Wetlesen T, Torjesen PA et al (2009) Maternal BMI before pregnancy, maternal weight gain during pregnancy, and risk of persistent positivity for multiple diabetes-associated autoantibodies in children with the high-risk HLA genotype: the MIDIA study. Diabetes Care 32:1904–1906
98. Couper JJ, Beresford S, Hirte C, Baghurst PA, Pollard A, Tait BD et al (2009) Weight gain in early life predicts risk of islet autoimmunity in children with a first-degree relative with type 1 diabetes. Diabetes Care 32:94–99
99. EURODIAB Substudy 2 Study Group (2002) Rapid early growth is associated with increased risk of childhood type 1 diabetes in various European populations. Diabetes Care 25:1755–1760
100. Dabelea D, D'Agostino RB Jr, Mayer-Davis EJ, Pettitt DJ, Imperatore G, Dolan LM et al (2006) Testing the accelerator hypothesis: body size, beta-cell function, and age at onset of type 1 (autoimmune) diabetes. Diabetes Care 29:290–294
101. Cardwell CR, Stene LC, Joner G, Davis EA, Cinek O, Rosenbauer J et al (2010) Birthweight and the risk of childhood-onset type 1 diabetes: a meta-analysis of observational studies using individual patient data. Diabetologia 53:641–651
102. Aberg A, Westbom L, Kallen B (2001) Congenital malformations among infants whose mothers had gestational diabetes or preexisting diabetes. Early Hum Dev 61:85–95

103. Martinez-Frias ML (1994) Epidemiological analysis of outcomes of pregnancy in diabetic mothers: identification of the most characteristic and most frequent congenital anomalies. Am J Med Genet 51:108–113
104. Wentzel P, Gareskog M, Eriksson UJ (2005) Folic acid supplementation diminishes diabetes- and glucose-induced dysmorphogenesis in rat embryos in vivo and in vitro. Diabetes 54:546–553
105. Waterland RA, Jirtle RL (2003) Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol 23:5293–5300
106. Lamb MM, Myers MA, Barriga K, Zimmet PZ, Rewers M, Norris JM (2008) Maternal diet during pregnancy and islet autoimmunity in offspring. Pediatr Diabetes 9:135–141
107. Fronczak CM, Baron AE, Chase HP, Ross C, Brady HL, Hoffman M et al (2003) In utero dietary exposures and risk of islet autoimmunity in children. Diabetes Care 26:3237–3242
108. Brekke HK, Ludvigsson J (2007) Vitamin D supplementation and diabetes-related autoimmunity in the ABIS study. Pediatr Diabetes 8:11–14
109. Norris JM, Barriga K, Klingensmith G, Hoffman M, Eisenbarth GS, Erlich HA et al (2003) Timing of initial cereal exposure in infancy and risk of islet autoimmunity. JAMA 290:1713–1720
110. Ziegler AG, Schmid S, Huber D, Hummel M, Bonifacio E (2003) Early infant feeding and risk of developing type 1 diabetes-associated autoantibodies. JAMA 290:1721–1728
111. Hypponen E, Kenward MG, Virtanen SM, Piitulainen A, Virta-Autio P, Tuomilehto J et al (1999) Infant feeding, early weight gain, and risk of type 1 diabetes. Childhood diabetes in Finland (DiMe) study group. Diabetes Care 22:1961–1965
112. Oge A, Isganaitis E, Jimenez-Chillaron J, Reamer C, Faucette R, Barry K et al (2007) In utero undernutrition reduces diabetes incidence in non-obese diabetic mice. Diabetologia 50:1099–1108
113. Chamson-Reig A, Arany EJ, Summers K, Hill DJ (2009) A low protein diet in early life delays the onset of diabetes in the non-obese diabetic mouse. J Endocrinol 201:231–239
114. Boujendar S, Arany E, Hill D, Remacle C, Reusens B (2003) Taurine supplementation of a low protein diet fed to rat dams normalizes the vascularization of the fetal endocrine pancreas. J Nutr 133:2820–2825
115. Petrik J, Reusens B, Arany E, Remacle C, Coelho C, Hoet JJ et al (1999) A low protein diet alters the balance of islet cell replication and apoptosis in the fetal and neonatal rat and is associated with a reduced pancreatic expression of insulin-like growth factor-II. Endocrinology 140:4861–4873
116. Lehnertz B, Northrop JP, Antignano F, Burrows K, Hadidi S, Mullaly SC et al (2010) Activating and inhibitory functions for the histone lysine methyltransferase G9a in T helper cell differentiation and function. J Exp Med 207:915–922
117. Ling C, Del Guerra S, Lupi R, Ronn T, Granhall C, Luthman H et al (2008) Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion. Diabetologia 51:615–622
118. Chen H, Gu X, Su IH, Bottino R, Contreras JL, Tarakhovsky A et al (2009) Polycomb protein Ezh2 regulates pancreatic beta-cell Ink4a/Arf expression and regeneration in diabetes mellitus. Genes Dev 23:975–985
119. Li Y, Zhao M, Hou C, Liang G, Yang L, Tan Y, Wang Z, Yin H, Zhou Z, Lu Q (2011) Abnormal DNA methylation in CD4(+) T cells from patients with latent autoimmune diabetes in adults. Diabetes Res Clin Pract 94:242–248 doi:10.1016/j.diabres.2011.07.027
120. Yang BT, Dayeh TA, Kirkpatrick CL, Taneera J, Kumar R, Groop L, Wollheim CB, Nitert MD, Ling C (2011) Insulin promoter DNA methylation correlates negatively with insulin gene expression and positively with HbA1c levels in human pancreatic islets. Diabetologia 54(2):360–367
121. Anderson RM, Bosch JA, Goll MG, Hesselson D, Dong PD, Shin D, Chi NC, Shin CH, Schlegel A, Halpern M, Stainier DY (2009) Loss of Dnmt1 catalytic activity reveals multiple roles for DNA methylation during pancreas development and regeneration. Dev Biol 334(1):213–23

122. Lenoir O, Flosseau K, Ma FX, Blondeau B, Mai A, Bassel-Duby R, Ravassard P, Olson EN, Haumaitre C, Scharfmann R (2011) Specific control of pancreatic endocrine {beta}- and {delta}-cell mass by class IIa histone deacetylases HDAC4, HDAC5, and HDAC9. Diabetes 60:2861–2871 doi:10.2337/db11-0440
123. Yu XY, Geng YJ, Liang JL, Lin QX, Lin SG, Zhang S, Li Y (2010) High levels of glucose induce apoptosis in cardiomyocyte via epigenetic regulation of the insulin-like growth factor receptor. Exp Cell Res. 316(17):2903–2909
124. Syreeni A, El-Osta A, Forsblom C, Sandholm N, Parkkonen M, Tarnow L, Parving HH, McKnight AJ, Maxwell AP, Cooper ME, Groop PH; on the behalf of the FinnDiane Study Group (2011) Genetic examination of SETD7 and SUV39H1/H2 methyltransferases and the risk of diabetes complications in patients with Type 1 diabetes. Diabetes 60:3073–3080 doi:10.2337/db11-0073
125. Bell CG, Teschendorff AE, Rakyan VK, Maxwell AP, Beck S, Savage DA (2010) Genome-wide DNA methylation analysis for diabetic nephropathy in type 1 diabetes mellitus. BMC Med Genomics 3:33

Chapter 20
Epigenetic Epidemiology of Obesity, Type 2 Diabetes, and Metabolic Disorders

Reinhard Stöger

Contents

20.1 Introduction 402
20.2 Complementing Genetic Approaches to Understand Metabolic Syndrome: Challenges of a Young Discipline 403
20.3 Muscle: Genetic and Epigenetic Analysis of COX7A1 404
20.4 Pancreatic Tissue 407
20.4.1 Specific Epigenetic Signatures and Genetic Variation 407
20.4.2 Epigenetic Nomenclature and the Characterization of Epialleles 408
20.4.3 The Histone Methylation Profile of Pancreatic Islets 408
20.4.4 Epigenetic Signatures of the Insulin Locus 409
20.5 Brain Tissue 409
20.6 Peripheral Blood Lymphocytes 410
20.6.1 Example and Concept of PBL DNA for Epigenetic Studies 410
20.6.2 Establishment of Epigenetic Patterns Reflecting T2D and Obesity Risk 410
20.6.3 Results from the Dutch Hunger Winter Cohort 411
20.6.4 Results from the AGES Cohort 412
20.7 Imprinted Genes: A Contributing Factor in the General Population? 415
20.7.1 Rare Imprinting Disorders Associated with a Metabolic Phenotype 416
20.7.2 The Imprinting Cluster on Chromosome 11p15.5 and IGF2 416
20.8 Conclusions 417
References 418

Abstract Knowledge and inclusion of epigenetic variants as potential risk factors for obesity and type 2 diabetes (T2D) will improve diagnosis and mechanistic understanding of these complex metabolic disorders. A number of practical and conceptual issues still have to be worked out before epidemiological studies can fully integrate epigenomic approaches. Altered patterns of histone modifications,

R. Stöger (✉)
University of Nottingham, School of Biosciences, Nottingham, UK
e-mail: reinhard.stoger@nottingham.ac.uk

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_20,

© Springer Science+Business Media B.V. 2012

DNA methylation, and microRNA levels are implicated in obesity and T2D. Recent studies are discussed, thereby illustrating strengths and challenges to accurately identify and measure epigenetic differences that covary with metabolic disorders. Unlike classical, DNA sequence-based approaches, the choice of cell or tissue type as source material may confound the results of such epigenomic studies. It is becoming clear that various types of epigenetic signatures exist; some are very stable and do not change over many years within an individual, whereas others are liable to change and perhaps more influenced by age and environmental and physiological conditions. A standardized classification system to accurately describe and characterize the properties of newly identified epigenetic signatures will enable the generation of reproducible epigenetic data.

Abbreviations

AS	Angelman syndrome
BMI	Body mass index
CNV	Copy number variation
DMR	Differentially methylated region
GWAS	Genome-wide association study
LOI	Loss of imprinting
PBL	Peripheral blood lymphocyte
PWS	Prader-Willi syndrome
SNP	Single nucleotide polymorphism
T2D	Type 2 diabetes
VMR	Variably methylated region

20.1 Introduction

This chapter discusses challenges and issues related to the budding field of epigenetic epidemiology of T2D and non-syndromic obesity. Recurrent themes are the choice of tissues and cells, the difficulty to efficiently measure epigenetic marks with accuracy in many samples, and the necessity to collect additional information such as environmental variables, age, and gender – all of which can influence the nature and stability of epigenetic marks. Since almost any change in cellular physiology is likely to result in some change in epigenetic marks, there will always be an association with metabolic disease. Distinguishing between association and causality of epigenetic marks will become a main issue for this research field. In this context, the small number of available studies are examined that highlight strengths and problems in determining the role of epigenetic variation in T2D, obesity, and related metabolic disorders.

20.2 Complementing Genetic Approaches to Understand Metabolic Syndrome: Challenges of a Young Discipline

Epigenetic research into T2D and obesity is still a very young field. The role of epigenetic mechanisms in the etiology of these disorders and related metabolic abnormalities such as dyslipidemia, hypertension, and hyperglycemia is not well understood. Nevertheless, interest is growing rapidly and epigenetic concepts are now being integrated in models to understand complex disease phenotypes [1].

This comes at a time when we realize that variations in DNA sequence reveal much less about disease predisposition than was widely expected. The genetic contribution to T2D and obesity – long presumed to be significant with heritability estimates reaching 70% in twin studies [2–4] – may be lower in reality. Genome-wide association studies (GWAS) have extracted a number of single nucleotide polymorphisms (SNPs) that are consistently associated with disease risk within the general population. Their contributions are small and currently can explain less than 10% of the presumed heritability of metabolic conditions [5]. As SNPs fail to sufficiently enlighten the heritable basis of non-syndromic T2D and obesity, attention has also turned to larger structural differences among human genomes. Quantitative variation of genomic regions, including insertions and deletions (indels) and common copy number variations (CNVs), were thought to add to disease risk. Yet, recent studies indicate that these types of genomic differences also fail to explain susceptibility to metabolic syndrome-related diseases [5, 6].

Studies on the role of epigenetics in T2D and obesity are lacking. At present, there is a disparity between the large number of hypothesis-based review articles and the few reports documenting solid, experimental evidence of an epigenetic role in T2D and obesity in human populations. Although for the remaining part of this chapter epigenetics is discussed as likely contributing factor, it is possible that epigenetics plays no causal role in the etiology of complex metabolic disorders.

Several factors account for the paucity of epigenetic data. Because it is a young research field, resources are only starting to be committed to this new line of investigation. Genomic loci carrying epigenetic signatures must be identified, which are unmistakably associated with phenotypes of metabolic syndrome. A number of candidate genes for such studies now exist, but further characterization is necessary before they can be confirmed as suitable biomarkers. It has been argued that large networks of non-imprinted and imprinted genes may acquire epigenetic modifications during early stages of life, resulting in a disease predisposition [7, 8]. Microarray-expression studies suggest that transcription levels of hundreds or thousands of genes are changed in animals and humans with symptoms of T2D and obesity [9, 10]. Whether epigenetic modifications underlie and maintain these altered gene expression states over time has not been established. If, as we expect, many genomic loci become modified and acquire epigenetic marks associated with T2D or obesity as a result of adverse lifestyle factors such as poor diet, lack of exercise [11], we must further determine whether they are causative and disease-promoting.

Existing methods enable epigenetic profiling of the entire human genome (Chap. 4). Many of these techniques were developed by probing DNA, histones, or chromatin from cell lines – homogenous populations of cells – and the resulting datasets have moderate levels of “background noise.” The complexity and heterogeneous mix of cell types typical for human tissues is a complicating factor that adds to the challenges of epigenetic epidemiology. What is the currently available evidence in humans on the role of epigenetic variation in metabolic disorders and what can we learn?

20.3 Muscle: Genetic and Epigenetic Analysis of COX7A1

A small cohort study explored the relation between genetic and non-genetic factors and their potential roles in contributing to altered gene expression in insulin resistance and type 2 diabetes [12]. Insulin resistance in skeletal muscles is an early characteristic in diabetics and insulin-insensitivity may be linked with impaired mitochondrial function, which is another symptomatic abnormality. Nuclear-encoded genes contributing to mitochondrial function are often down-regulated in insulin-resistant muscle [13–15]. One of these genes is *COX7A1*, which encodes a component of the mitochondrial respiratory chain. A compelling feature of the work by Ling and colleagues is that they examined *COX7A1* in twins by sampling the *vastus lateralis* muscle [12]. This muscle is prone to become insulin insensitive and is pathologically affected in obese and diabetic individuals. Variables contributing to a reduction of *COX7A1* expression levels in muscle were considered and included (a) SNPs, representing variations in genotype; (b) age, representing a non-genetic factor; and (c) DNA methylation, representing epigenetic variation.

Genotype: Although one SNP (rs753420) appeared to influence basal *COX7A1* expression, the effect of this particular genotype was only detectable in muscle of young twins ($n=110$/age 28.0 ± 1.9), which were a subset of the study participants [12]. Overall, *COX7A1* genotypes did not reveal any associations with T2D-risk in the sampled group ($n=4{,}296$).

Age: One of the key findings was that *COX7A1* expression levels decrease significantly with age. This non-genetic variable must be considered when conducting and interpreting gene-expression data related to metabolic traits – age appears to have a strong influence on certain genomic loci (see also Chap. 10)

DNA methylation: To measure DNA methylation levels of the *COX7A1* promoter, samples from a small subset of individuals were analyzed, which included ten young twins (age 28.0 ± 1.9) and ten older twins (age 62.4 ± 2.0). PCR products derived from sodium bisulfite converted DNA were cloned and subsequently sequenced. The promoter region contains 22 CpG dinucleotides; these are potential methylation sites. The sequencing data from each individual was pooled and used to determine the proportion of *COX7A1* methylation by dividing the number of methylated CpGs

by the total number of CpGs in the data set compared to the young twins, DNA methylation levels were higher in older twins. While the *COX7A1* DNA methylation levels were higher, they were also found to be more diverse among individuals of the elderly twin-group. Thus, DNA methylation patterns diverge with age and as a result of environmental exposures, a phenomenon that has to be considered in the design of population-based, epigenetic studies [16] (see also Chap. 10).

DNA methylation levels surrounding a promoter are generally assumed to provide an adequate proxy for gene activity, yet Ling and colleagues did not observe a significant correlation between DNA methylation levels and *COX7A1* expression [12]. Measurements of this epigenetic modification were performed on DNA samples derived from skeletal muscle – relevant to disease – and the same tissue samples were used to determine *COX7A1* mRNA levels. Does the lack of correlation between DNA methylation and gene expression indicate a lack of functionality? Is it worth to assess DNA methylation levels at individual loci, or will such studies consume disproportionate amounts of resources without providing meaningful information, as so often is the case for SNP-studies? Is there a possible explanation for the mismatch between the *COX7A1* methylation profile and expression levels?

For each individual, methylation patterns of ten DNA molecules (epialleles) from ten cells were analyzed. While this approach – subcloning and sequencing of PCR products derived from bisulfite-treated genomic DNA – yields the most accurate information of cytosine methylation patterns of individual DNA molecules, it also bears several disadvantages, particularly if only a limited number of sequences are being used to measure DNA methylation levels. Biased PCR amplification of a subset of template DNA molecules, or biases during the subcloning procedure, can often lead to a quantitative misrepresentation. Failing to reflect the true distribution of the many possible types of epialleles, even subtle biases in sampling regularly contribute to the distortion of data sets. Redundant sampling of the same DNA molecule – and thus the same DNA methylation pattern – is particularly frequent when only limited amounts of template DNA are available. The muscle biopsies studied by Rönn and colleagues were likely composed of different cell types making the comparison, on the basis of about ten cells per individual, more challenging. The above-mentioned technique to analyze DNA methylation patterns is unsurpassed at the level of individual DNA molecules – every potential cytosine methylation event within the targeted region can be detected. This is a suitable method if the DNA methylation state of a genomic locus is examined for the first time and is considered as a potential epigenetic biomarker for T2D and obesity.

To establish robust methylation profiles for loci that have a broad range of DNA methylation densities, a large number of epialleles must be examined. Exceptions would be genes and promoters that normally remain methylation-free in all human tissues and cell types; here, any trace of cytosine methylation could indicate a disease predisposition. Such epigenetic lesions exist in tumor samples (see Chap. 12). Identification of this type of clear-cut, unambiguous epigenetic alteration associated with metabolic disease risk would greatly advance population-based screens.

Given an unbiased sampling, working with large numbers of epialleles also reduces the problem of cell type diversity in tissue specimen. The more epialleles we analyze per tissue sample, the better the representation will be of the various cell types present in a specimen. Thus, even if the epialleles are derived from a tissue composed of heterogeneous cell types, an overall representative methylation profile will emerge from a locus that carries diagnostic value. That is, the epigenetic profile will likely differ between healthy and diabetic/obese individuals, but the differences might be very subtle. Unlike genetic testing, the results will have to be interpreted in "shades of grey" rather than "black and white." Technically, small research groups

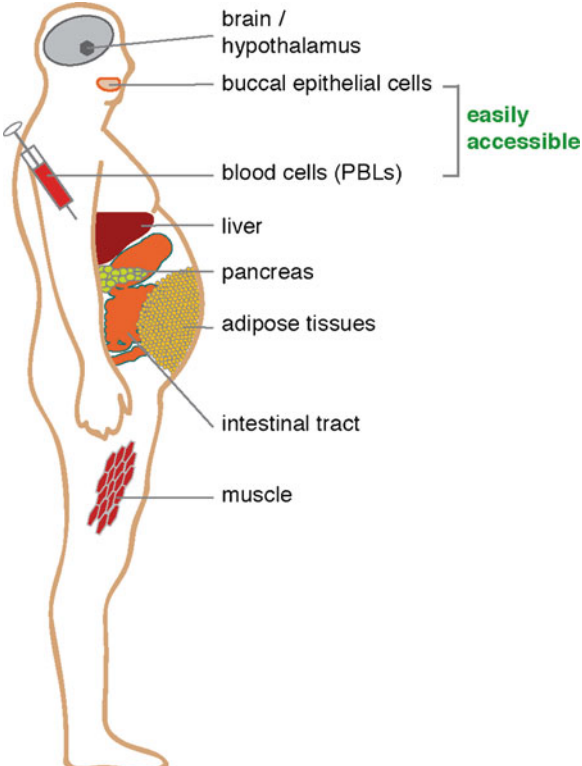


Fig. 20.1 Energy homeostasis is a fundamental physiological process that requires interaction and regulation of multiple organ systems, tissues, and cell types of the body. The central nervous system – the hypothalamus in particular – integrates signals from peripheral tissues and controls food intake, satiety, and energy expenditure. Most of the tissues relevant for our understanding of T2D and obesity (muscle, intestinal tract, adipose tissues, pancreas, liver, and brain) cannot be easily sampled and are not suitable for large, epigenetic population studies. In contrast, peripheral blood cells and buccal epithelial cells can be collected without difficulty and are frequently used as a DNA source for genetic screens; whether these cell types carry epigenetic biomarkers that identify disease risk for either T2D or obesity is insufficiently studied

can carry out this technique. Nevertheless, it is labor and time intensive, and as a result costly and not well suited for epidemiological studies.

In order to advance the research field it will be important to establish epigenetic profiles for genomic loci in tissues, such as muscle, which play a pathophysiological role. Because of the obvious difficulties of obtaining these tissues, other research groups cannot easily validate such data. Future studies will benefit from comparative epigenetic testing, by matching data from, for example, muscle, with data derived from peripheral blood cells from one and the same person. In this way, it will be possible to build databases that allow us to determine whether the epigenetic state of a given genomic locus is similar in disease-affected tissues and simple-to-collect tissues. In an ideal setting, the epigenotype in the affected tissue will be mirrored by the epigenotype in peripheral blood lymphocytes (see also Fig. 20.1).

20.4 Pancreatic Tissue

Suboptimal insulin secretion and changes in the number of beta-cells – one of the endocrine cell types in pancreatic islets – contribute to the etiology of T2D [17]. Altered gene expression programs in pancreatic islets could be an underlying factor for disease risk. Several research groups have recently started to investigate the role of chromatin, epigenetic modifications, and genetic variants in pancreatic islets.

20.4.1 Specific Epigenetic Signatures and Genetic Variation

Gaulton and colleagues established a map of open chromatin sites in human pancreatic islets and compared the results with chromatin maps of non-islet cell lines [18]. Their chromatin-profiling experiments demonstrated that pancreatic islets have a distinct, tissue-specific distribution of open chromatin sites. Thus, a particular epigenetic information – relevant for our understanding of T2D and obesity – can be confined to difficult-to-collect tissues.

The authors discovered that one of the islet-specific open chromatin sites overlaps with the genomic position of a T2D-associated SNP (rs7903146) [18]. This SNP maps to intron 3 of the gene encoding transcription factor 7-like 2 (*TCF7L2*). The T risk-allele was found to be in a more “open chromatin” structure, indicating reduced nucleosome occupancy at this locus in comparison with the non-risk C-allele. Such a “relaxed” chromatin state correlates well with increased *TCF7L2* expression of the T risk-allele, which had been observed earlier [19, 20]. Carriers of the *TCF7L2* risk-allele are found to have impaired insulin secretion [21, 22].

Establishment of cell type specific chromatin profiles and their correlation with disease-associated SNPs is a promising approach to unravel the molecular and epigenetic mechanisms that lead to the phenotypes associated with the metabolic syndrome [23].

20.4.2 Epigenetic Nomenclature and the Characterization of Epialleles

The correlation between the T risk-allele and an open chromatin state suggests that epigenetic variation at the *TCF7L2* locus may either be "obligate" or "facilitated." These terms have been proposed by Richards and allow classification of epigenetic variants [24, 25]. An obligate epigenetic variant directly corresponds with a given genotype; in this case, the T-risk allele would always be associated with open chromatin. The facilitated epigenetic variant is somewhat dependent on the genotype, but would not necessarily have to be linked with the DNA sequence variation; in this case, the T risk-allele would often be associated with open chromatin. It is likely that many associations between epigenotype and T2D/obesity-risk phenotype exist that are completely independent of the genotype [7]. Using the nomenclature proposed by Richards such genotype-independent epigenetic variants would be labeled as "pure" [24].

20.4.3 The Histone Methylation Profile of Pancreatic Islets

Human pancreatic islets were also subject of a recent epigenetic analysis by Kaestner's group [26]. A genome-wide map of histone modifications was established by chromatin immuno-precipitation and massively parallel sequencing technology (ChIP-seq). The relative distribution of two marks was recorded: methylation of histone 3 lysine 4 (H3K4me) is usually associated with active genomic loci; in contrast, methylation of histone 3 lysine 27 (H3K27me) usually correlates with genomic repression. Unlike many other epigenomic studies that sample cell lines – and therefore obtain fairly clear-cut results – this study looked at the mixture of cells present in the pancreatic islets. Assuming that these pancreatic cells were obtained from individuals with no apparent metabolic dysfunction, the results provide a glimpse of the normal histone-modification profile present in cells enriched from this endocrine tissue. For example, pancreatic and duodenal homeobox 1 (*PDX1*) and *MAFA* are genes involved in differentiation and regulation of beta-cell function and despite relative low expression levels, their promoters are highly occupied with H3K4 trimethylated nucleosomes. Unexpectedly, only low levels of H3K4 trimethylation were detected around the insulin (*INS*) and glucagon (*GCG*) promoters, leading the authors to speculate that alternative epigenetic modifications and regulatory mechanisms promote high-level transcription of both genes in vivo [26].

The genome-wide map of H3K4me3 marks in pancreatic islets created by Kaestner and colleagues corroborate the tissue-specific distribution of this histone modification [26]. The pancreas dataset was compared with a high-resolution histone modification profile derived from CD14+ T-cells described by Barski et al. [27]. Islet cells have comparatively high H3K4me3 levels in a small subset of genes (< 100) and the associated gene promoter sequences have a low content of CpG dinucleotides. This small list of genes with a pancreatic-specific histone modification profile could prove very useful for comparative epigenetic studies and might indicate where differences in chromatin structure and gene expression will be detected

in the diseased state. As pancreatic islets play a significant role in the etiology of diabetes and carry their own distinct epigenetic profile, it becomes clear that this tissue could provide important insights, but unfortunately it is an impractical choice for large, population-based studies (Fig. 20.1)

20.4.4 Epigenetic Signatures of the Insulin Locus

The epigenetic makeup of the insulin (*INS*) locus has been characterized in more detail by Mutskov and Felsenfeld, who also used freshly isolated human pancreatic islets as study material [28]. *INS* maps to a region of human chromosome 11p15.5, within an ~ 80 kb gene-rich cluster including tyrosine hydroxylase (*TH*), insulin-like growth factor 2 antisense (*IGF2AS*), and insulin-like growth factor 2 (*IGF2*). High levels of histone acetylation and H3K4 dimethylation – both epigenetic marks associated with gene activity – were present throughout the genomic region, indicating that this gene cluster lies within an unexpectedly large and open chromatin domain. Overall the epigenetic profile was characteristic for all tested pancreatic islet cells. However, the authors noted some variation among the samples without further exploring the cause of these presumably subtle, epigenetic differences [28]. Variations in the purity and quality of the isolated pancreatic islets may explain such differences, exemplifying one of the practical difficulties that must be overcome in order to establish "normal" base-line data. It will be of interest to create similar epigenetic profiles in pancreatic islets from obese or diabetic individuals and to determine whether or not the large, open *INS* chromatin domain at 11p15.5 is significantly altered in the diseased state.

Tissue-specific, intergenic transcription of non-coding RNA was observed within the 80 kb *INS* gene cluster in pancreatic islets, but was absent in HeLa cells and human brain tissue [28]. Islet-specific intergenic transcription correlates in a linear fashion with *INS* expression levels. In addition, two sites within the *INS* locus are devoid of DNA methylation, perhaps demarcating transcriptional start sites for non-coding, intergenic islet transcripts. At this stage we do not know if and how DNA methylation levels change at 11p15.5 in individuals with metabolic syndrome, nor is it clear whether alterations of intergenic RNAs around *INS* locus correlate with – or even contribute to – diminished metabolic health. A role of non-coding RNAs must be considered, as microRNAs are known to modulate insulin promoter activity and insulin mRNA levels in the adult mouse pancreatic islets [29].

20.5 Brain Tissue

The brain, and the hypothalamus in particular, plays a key role in modulating appetite and energy expenditure (Fig. 20.1). This region of the brain integrates information from peripheral organs, including the intestinal tract and adipose tissues and connects

via neuronal pathways with different brain regions and the endocrine system [30–32]. Within the hypothalamus the suprachiasmatic nucleus (SCN) orchestrates peripheral, circadian body clocks, which in turn regulate crucial metabolic processes. Recent studies provide evidence for a "circadian epigenome," linking daily metabolic rhythms with oscillating chromatin changes and histone modifications (reviewed in [33]). Because this tissue cannot be easily sampled, animal studies may have to suffice in the identification of T2D or obese-specific epigenetic patterns in the hypothalamus.

20.6 Peripheral Blood Lymphocytes

Peripheral blood lymphocytes (PBLs) are routinely collected and are an economic and feasible source for molecular epidemiology studies. A pertinent question then is whether this source of DNA can be used to probe epigenetic differences that accurately reflect a given complex phenotype.

20.6.1 Example and Concept of PBL DNA for Epigenetic Studies

For certain conditions, such as Fragile-X syndrome, DNA methylation patterns derived from PBLs are indeed good diagnostic indicators. The degree of cytosine methylation mosaicism at a single locus – the *FMR1* promoter – in PBL DNA was shown to correlate remarkably well with the degree of cognitive impairment of a fragile X individual [34, 35]. The symptoms of fragile X syndrome can range from mild learning impairments to severe intellectual disabilities [36] and although blood cells do not contribute to the fragile-X phenotype they are suitable for epigenetic profiling of this predominantly neurological disorder because aberrant DNA methylation at the *FMR1* locus is likely established in the germline or during early stages of embryonic development [37, 38]. As cells divide and differentiate, *FMR1* methylation is maintained in neurons but also in the multiple distinct cell types present in peripheral blood. Proportions of the different cell types in a population of PBLs can vary dramatically even between samples taken from a single individual at different time points; an infection, for example, significantly alters the cell type ratio in PBLs. Thus, if an epigenetic marker is not derived from the germline, or undifferentiated embryonic cells, but is particular to one blood cell type, proportional changes in the PBL pool of cells can result in apparent – but misleading – changes of the epigenotype.

20.6.2 Establishment of Epigenetic Patterns Reflecting T2D and Obesity Risk

Epidemiologic studies suggest that at least some epigenetic patterns, indicative of a predisposition for metabolic disorders, could be established during early stages in

life, propagated and maintained in cells of an adult person [39–43]. That is, a population of embryonic and fetal cells may acquire epigenetic marks that are passed on to developing tissues that contribute to metabolic disease. These include the pancreas, muscle, adipose, and brain, but also to cell types not directly associated with T2D and obesity, such as hematopoietic stem cells, the precursors of mature blood cells. Questions then arise about the causes, timing and stability of varying epigenetic states. Early life events indeed seem to have lasting effects, influencing metabolic disease risk in adulthood; this topic is covered in detail in Chaps. 6 and 7. Briefly, there is clear evidence that environmental factors – nutrition in particular – experienced during pregnancy and perinatal development can predispose to the metabolic syndrome. Studies of the "Dutch hunger winter" cohort are a classic example. A transportation blockade imposed by the Nazis caused a severe famine during the winter of 1944/1945 in the western region of the Netherlands. Epidemiological studies revealed that individuals who were exposed in utero to this severe food shortage are significantly more likely to be diabetic and overweight [44, 45].

20.6.3 Results from the Dutch Hunger Winter Cohort

Heijmans and colleagues recently reported their first results of epigenetic profiling of individuals of the Dutch hunger winter cohort, using DNA derived from PBLs [46, 47]. Genes chosen for their methylation measurements have been implicated in the etiology of obesity, T2D or related metabolic dysfunctions. More than 60 years after the Dutch famine, DNA methylation levels were detected that appear to differ between individuals who were exposed to the famine during pregnancy and their same-sex siblings who were either born before, or conceived after the hunger winter [46, 47]. In exposed individuals, seven loci were found to have different DNA methylation levels; they include the imprinted genes *IGF2, GNASAS, MEG3*, and *INSIGF* and non-imprinted genes *IL10, ABCA1, LEP*. The results are remarkable and imply that hematopoietic stems cells of adults retain epigenetic memories of environmental events that occurred in early life. The results further suggest that it may be possible to use PBLs for epigenetic epidemiology studies of the intrauterine origin of T2D and obesity although these cell types are not directly implicated in contributing to the complex, metabolic disease phenotypes (although it could be argued that infiltration of marcophages and pro-inflammatory T-lymphocytes in adipose tissue is an important feature in obese and insulin-resistant patients [48–52]).

DNA methylation levels in the Dutch hunger winter cohort were measured by a method based on mass spectrometry (Epityper, Sequenom) [46, 47]. While the quantitative nature and reproducibility of this approach is generally considered to be accurate, the results provide only average cytosine methylation estimates for all CpGs combined of a given genomic region, and can be influenced by genetic polymorphism. Importantly, this approach does not reveal the nature of methylation patterns on individual DNA strands; this is relevant for our understanding of the processes that shape the DNA methylation profile at a given locus.

Leptin (*LEP*), encoding an adipocyte hormone regulating energy homeostasis, was one of the genes found to have higher methylation levels in exposed individuals of the Dutch hunger winter cohort study [47]. Cytosine methylation patterns vary considerably at the human *LEP* locus in adipose tissues, where *LEP* is highly expressed [52, 53]. Pronounced epigenetic variation – mosaicism of cytosine methylation – of the *LEP* promoter has also been detected in DNA isolated from PBLs [52]. Examples of such methylation mosaicism among *LEP* epialleles from PBLs are shown in Fig. 20.2; they illustrate the complex nature of epigenetic patterns that may be present in many autosomal, non-imprinted genomic loci. Based on the broad range of methylation densities in healthy individuals, it may be a challenging task to differentiate between normal methylation levels and those that predispose to disease.

We can assume that *LEP* methylation densities reported for the individuals by Heijmans et al. reflect the heterogeneous mixture of epialleles present within each of the sampled PBL-DNAs. As Fig. 20.2 depicts, some individuals have a larger proportion of more densely methylated epialleles and will therefore register to have higher *LEP* methylation in PBL-DNA. If *LEP* methylation densities also reflect an early-life event – as the results by Heijmans and colleagues imply – it must be assumed that DNA methylation patterns do not change significantly at the *LEP* locus as cells divide and differentiate from a pool of hematopoietic stem cells into the short-lived PBLs. This would suggest that epigenetic marks established at the *LEP* promoter are very stable in cells of the hematopoietic system and the locus would have the potential to be used as an epigenetic biomarker for metabolic disease risk set in utero. Detailed studies and longitudinal methylation profiles need to be established that trace the *LEP* DNA methylation in order to determine the epigenetic stability of this genomic locus.

There is evidence that DNA methylation levels are influenced by gender. PBL-derived DNA of males has methylation levels that are reportedly somewhat elevated in *Alu* and *Line-1* elements, as well as in differentially methylated regions (DMRs) of imprinted loci, including *PEG3, NESP55*, and *H19* [54]. Sex-dependent variations were also observed in DNA samples of the Dutch hunger winter study. In particular, exposure to famine around the time of conception influences the methylation levels of *LEP, INSIGF*, and *GNASAS* in males [47]. Such findings add another layer of complexity and any epigenetic biomarker considered must first be scrutinized for gender-specific differences if epidemiologic studies are to succeed.

20.6.4 Results from the AGES Cohort

A recent study by Feinberg and colleagues also suggests the existence of epigenetic marks in lymphocyte DNA that correlate with the body mass index (BMI) of an individual [55]. An approach termed "CHARM" was used to measure DNA methylation levels across the genome. CHARM is a comprehensive high-throughput

Hematopoietic stem cells

- Are *Leptin* methylation patterns established in early life?
- Do individual stem cells maintain a particular pattern throughout their life span?
- Do these methylation patterns change significantly during differentiation?
- Are they influenced by environment, age, gender, circadian rhythms, etc ?
- Do these patterns reflect DNA methylation levels in other tissues of a person?
- Do these methylation patterns modulate *Leptin* expression levels?

Leptin methylation in mature, peripheral blood lymphocytes

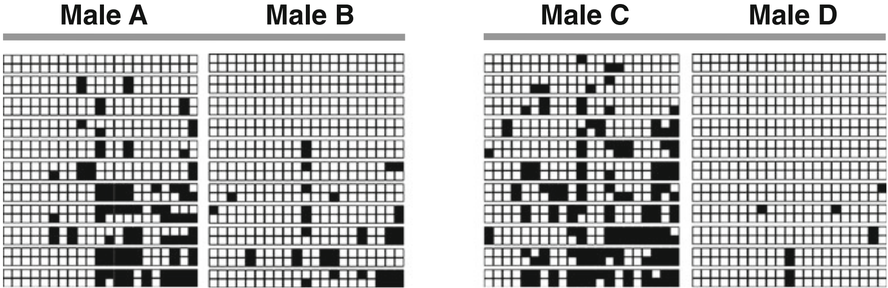


Fig. 20.2 An example of DNA methylation mosaicism. Marked variation in methylation densities is characteristic for the Leptin (*LEP*) promoter in cells of an individual, as well as between individuals. DNA samples from peripheral blood lymphocytes (PBLs) were analyzed in two age-matched pairs of adult males: males A & B, and males C & D, respectively. The DNA samples were received anonymized and it was not possible to determine whether the *LEP* methylation status of an individual correlates with body mass index, health status, or environmental variables. Further epigenetic characterization of the *LEP* promoter will be necessary before we know if this genomic locus can be used as a reliable biomarker for epigenetic epidemiological studies. It is not clear if and how different methylation densities of the promoter influence transcription levels of *LEP* mRNA. The data were obtained using hairpin-bisulfite PCR, a method that provides information on the methylation status of CpG sites on both strands of individual DNA molecules. For each male, a column is depicted with methylation patterns from 11 double-stranded DNA molecules. The two rows of boxes within each thin rectangle stand for CpG sites on the two strands of individual DNA molecules; unmethylated CpG sites are shown in *white*, while methylated CpG sites are shown in *black*

array-based relative methylation analysis platform, combining tiling arrays and statistical procedures [56]. With a hypothesis-neutral analysis of around four million CpG sites (approximately one seventh of all CpGs of the human genome), this study currently represents the largest screen to identify genomic loci whose epigenetic make-up relate to BMI [55].

The authors had access to samples of lymphocyte DNA from participants of the Age, gene/environment, susceptibility (AGES) cohort. The AGES cohort consists of Icelandic individuals partaking in an ongoing longitudinal study, which originally began in 1967 [57]. The phenotypes of these individuals have been documented in detail, including body composition and metabolism. Two DNA samples from each individual, collected about 11 years apart, were included in the methylation analysis. At the time of the first of these two DNA sample collections, individuals were between 58 and 85 years of age. Overall, Feinberg and colleagues analyzed data from 74 individuals [55].

Methylation levels of more than 200 genomic loci were found to vary considerably among the tested individuals. These variably methylated regions (VMRs) could be further categorized into three groups. Site-specific "dynamic" VMRs (41 out of 227) seem to undergo large changes in methylation levels over time within an individual (intra-individual variation). About a third of the VMRs (67 out of 227) are "ambiguous" with respect to intra-individual methylation change over time. The third group of VMRs (119 out of 227) was categorized as "stable"; the methylation levels of these particular genomic regions did not change in an 11-year time span. Indeed, in analogy to a person's SNP profile, it was possible to identify individuals on the basis of their unique methylation profiles of stable VMRs [55]. It is not clear why the methylation levels of these VMRs are stable but vary among individuals. Using Richards' aforementioned classification system [24], future studies will likely clarify whether these stable VMRs are "obligate" (entirely related to an individual's genotype), "facilitated" (somewhat influenced by genotype), or "pure" (entirely uninfluenced by genotype) epigenetic variants.

Detection of stable VMRs allowed Feinberg and colleagues to explore the relationship between epigenotype and BMI [55]. Following stringent statistical procedures four stable VMRs located near genes were found to correlate over time with an individual's BMI. The four genes are *PM20D1, MMP9, PRKG1*, and *RFC5*, respectively. For example, the level of methylation at *PM20D1* was elevated by about 20% in obese participants (≥30 BMI), in comparison with normal-weight, individuals (BMI ≤25). *PM20D1* encodes an unstudied putative carboxypeptidase and its role in energy metabolism is not known. The authors point out that the remaining three genes with stable VMRs may be implicated in metabolic syndrome [55]. That is, the matrix metalloproteinase 9 (*MMP9*) is upregulated in preadipocytes of obese individuals [58]. Overactivity of *MMP9* has also been observed in cultured adipocytes exposed to macrophage-conditioned medium [59] and in atherosclerotic plaques from subjects with T2D [60]. The cyclic GMP-dependent protein kinase-1 gene (*PKRG1*) is linked with food-related behaviors in genetic model organisms [61]; in humans, however, genetic variations associated with obesity have not been found in *PKRG1* [62]. Deregulation of *RFC5,* encoding a

component of the DNA replication and repair clamp loader machinery, may be involved in diabetes-associated DNA damage as the authors suggest [55]. These four genomic regions were not detected previously by large, genome-wide association studies as potential risk loci for common obesity. This may indicate that SNPs do not significantly influence DNA methylation levels of the associated genes. In other words, genetic variation does not influence the variance of methylation of the four BMI-related VMRs; these loci appear to harbor "pure" epigenetic variants.

CHARM is clearly an effective approach to establish an individual's gene-specific DNA methylation profile on a genome scale. In principle, it should be possible to scale-up the procedure permitting large, population-based epigenomic screens of complex phenotypes such as obesity. It remains to be shown if other research groups will adopt the CHARM approach, or if rapid advancement of technologies will supersede this procedure.

The methylation data were obtained from blood DNA. As discussed earlier, we currently do not know if epigenomic signatures from lymphocyte DNA are a suitable proxy for other tissues of an individual. It is promising that four loci were identified whose DNA methylation levels showed covariation with BMI. Will these VMRs turn out to be accurate epigenetic biomarkers for diagnosis of obesity? The study by Feinberg et al. – as all work referred to in this chapter – illustrates that it is still early days for the research field of epigenetic epidemiology of obesity. Confidence will grow in epigenetic biomarkers if independent measurements lead to similar results.

20.7 Imprinted Genes: A Contributing Factor in the General Population?

Imprinted genes carry parental-specific epigenetic information. Epigenetic variation and instability in this class of genes was observed soon after the first imprinted genes had been discovered [63]. Perturbations of genomic imprints have mainly been analyzed in the context of tumor biology, where loss of imprinting (LOI) is a common event in a variety of cancers [64]. Whether or not degeneration of parental imprints commonly occurs in metabolic diseases remains to be determined [8, 65, 66]. A possible role for imprinted genes in obesity and T2D can be inferred from to kinship theory of genomic imprinting. This hypothesis predicts that imprinted genes are generally involved in managing resources, ranging from regulating cell growth to feeding behaviors [67–69].

A recent case in point is the maternally expressed, imprinted gene *KLF14*, which encodes a Krüppel-like transcription factor [70]. *KLF14* operates as a master trans regulator, modulating the expression of a gene network in adipose tissue [70]. *KLF14* and many of its target genes have been associated with T2D and related metabolic phenotypes in large population-based studies [71].

20.7.1 Rare Imprinting Disorders Associated with a Metabolic Phenotype

Certainly some rare human imprinting syndromes—caused by epigenetic or genetic abnormalities—feature metabolic imbalances as part of the phenotypic spectrum. For example, perturbations of parent-of-origin specific gene expression on human chromosome 15q11-13 can cause Prader-Willi syndrome (PWS) or Angelman syndrome (AS), respectively. PWS individuals become obese during adulthood as they have hypothalamic abnormalities and a defective satiety-sensing system promoting uncontrollable hyperphagia [72].

A cluster of imprinted genes map to 15q11-13, and although all the genes and transcripts have been identified, it is still not clear how many of them contribute to the clinical manifestations of PWS. Some genes such as *UBE3A and ATP10A* show tissue-specific imprinting; in neuronal tissues the paternally inherited allele is mostly inactive, while bi-allelic expression is observed in all other tissue types [73, 74]. Hypothalamic tissue would be the choice for epigenetic studies of the imprinted gene cluster on 15q11-13, but use of this tissue is obviously not feasible in any epidemiologic approach to metabolic disorders.

The imprinted *GNAS* locus on 20q13.2-13.3 has been linked with the development of an early-onset obesity phenotype in a subgroup of individuals with Albright hereditary osteodystrophy (AHO). Mainly patients with mutations in their active, maternally inherited G-protein alpha-subunit gene ($G_s a$) allele become obese. Metabolic phenotypes observed in a range of different *Gnas* knock-out mouse models support the finding that perturbation of imprinting at this locus affects glucose and energy metabolism [75, 76].

20.7.2 The Imprinting Cluster on Chromosome 11p15.5 and *IGF2*

Another large genomic region harboring at least 12 imprinted genes lies on human chromosome 11p15.5 and includes the aforementioned *INS* chromatin domain (Sect. 20.4.4). Prominent imprinted genes in this cluster are *IGF2, CDKN1C, KCNQ1*, and the non-coding transcripts *H19* and *KCNQ1OT1*. Both genetic and epigenetic defects in this locus are associated with a number of human diseases such as the Beckwith-Wiedemann syndrome and transient neonatal diabetes mellitus. Reduced expression in the pancreas of *CDKN1C*, a cyclin-dependent kinase inhibitor, has been observed in hyperinsulinism of infancy and is associated with an increased proliferation of pancreatic beta-cells [77]. SNPs associated with T2D susceptibility were recently identified in the imprinted region on 11p15 [5, 78]; some of them confer disease risk when inherited from the father. Interestingly, alleles carrying the risk variant of one of these SNPs (rs2334499) have decreased methylation levels in

PBL-DNA at five CpG sites located within a DMR of the 11p15 imprinting cluster [79]. If confirmed, this SNP would provide another example of an "obligate" or "facilitated" epigenotype, where a genetic variation modulates the epigenotype of a genomic region associated with T2D risk in the general population.

IGF2 is an imprinted gene of interest within the 11p15 imprinting cluster. Epigenetic lesions are thought to cause bi-alleleic expression of *IGF2*, a gene that is normally transcribed only from the paternal allele. Within the general population bi-allelic *IGF2* expression has been observed at a frequency of about 10% [80, 81 #45]. Loss of imprinting and other, as yet unknown, events causing elevated *IGF2* levels can promote cell growth and division. In culture, growth-arrested pancreatic beta-cells resume cell proliferation upon treatment with IGF2 at physiological levels [82, 83]. Similarly, transgenic mice with beta-cell specific *IGF2* overexpression have pancreatic islet hyperplasia and subsequently develop insulin resistance as a consequence of prolonged hyperinsulinemia [84]. Is there a relationship between the *IGF2* imprinting status and pancreatic beta-cell mass? The number of beta-cells is thought to vary considerably among individuals [85]. Again, the difficulty to collect these endocrine micro-organs in humans makes it impractical to explore if epigenetic dysegulation of *IGF2* is associated with changes in beta-cell mass in metabolic disorders.

20.8 Conclusions

Integration of epigenetics as a component of population-based T2D and obesity studies will undoubtedly accelerate our understanding and shape our views of these public health problems in the coming years. Much expectation was placed in the power of genetics. We must now look to epigenetics to fulfill the promises of mechanistic insight, accurate diagnosis, prognosis, and potential therapies. As PBLs will be a preferred source for epidemiological studies, genomic loci from these cells will have to be rigorously characterized. In addition to classic epigenetic systems, non-coding RNAs have been identified as regulators of adipose tissue and insulin sensitivity [86]. MicroRNAs are likely to become a relevant class of biomarkers alongside DNA and histone modifications. The dynamic nature of epigenetic marks, and the many variables that can influence them, pose considerable challenges for this research field. The emerging data may not provide easily identifiable markers, and will probably require the adjustment or development of new statistical tools. Collection and use of additional data will bring the results into focus. Cumulative environmental exposures, age, gender, and even circadian rhythms are factors that must be included when embarking on epigenetic epidemiology studies. The task is complex – but so is the metabolic syndrome.

Acknowledgments I am grateful to the editor, Karin Michels, as well as Catherine Suter and David Martin for comments and suggestions on the manuscript.

References

1. Flintoft L (2010) Complex disease: adding epigenetics to the mix. Nat Rev Genet 11:94–95
2. Stunkard AJ, Harris JR, Pedersen NL, McClearn GE (1990) The body-mass index of twins who have been reared apart. N Engl J Med 322:1483–1487
3. Turula M, Kaprio J, Rissanen A, Koskenvuo M (1990) Body weight in the Finnish Twin Cohort. Diabetes Res Clin Pract 10(Suppl 1):S33–S36
4. Wardle J, Carnell S, Haworth CM, Plomin R (2008) Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. Am J Clin Nutr 87:398–404
5. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP et al (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 42:579–589
6. Consortium WTCC (2010) Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature 464:713–720
7. Stöger R (2008) The thrifty epigenotype: an acquired and heritable predisposition for obesity and diabetes? Bioessays 30:156–166
8. Stöger R (2008) Epigenetics and obesity. Pharmacogenomics 9:1851–1860
9. Koza RA, Nikonova L, Hogan J, Rim JS, Mendoza T, Faulk C et al (2006) Changes in gene expression foreshadow diet-induced obesity in genetically identical mice. PLoS Genet 2:e81
10. Chen Y, Zhu J, Lum PY, Yang X, Pinto S, MacNeil DJ et al (2008) Variations in DNA elucidate molecular networks that cause disease. Nature 452:429–435
11. Ling C, Groop L (2009) Epigenetics: a molecular link between environmental factors and type 2 diabetes. Diabetes 58:2718–2725
12. Rönn T, Poulsen P, Hansson O, Holmkvist J, Almgren P, Nilsson P et al (2008) Age influences DNA methylation and gene expression of COX7A1 in human skeletal muscle. Diabetologia 51:1159–1168
13. Sreekumar R, Halvatsiotis P, Schimke JC, Nair KS (2002) Gene expression profile in skeletal muscle of type 2 diabetes and the effect of insulin treatment. Diabetes 51:1913–1920
14. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J et al (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 34:267–273
15. Patti ME, Butte AJ, Crunkhorn S, Cusi K, Berria R, Kashyap S et al (2003) Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: potential role of PGC1 and NRF1. Proc Natl Acad Sci USA 100:8466–8471
16. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML et al (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 102:10604–10609
17. Rhodes CJ (2005) Type 2 diabetes-a matter of beta-cell life and death? Science 307:380–384
18. Gaulton KJ, Nammo T, Pasquali L, Simon JM, Giresi PG, Fogarty MP et al (2010) A map of open chromatin in human pancreatic islets. Nat Genet 42:255–259
19. Lyssenko V, Lupi R, Marchetti P, Del Guerra S, Orho-Melander M, Almgren P et al (2007) Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. J Clin Invest 117:2155–2163
20. Lee SH, Demeterco C, Geron I, Abrahamsson A, Levine F, Itkin-Ansari P (2008) Islet specific Wnt activation in human type II diabetes. Exp Diabetes Res 2008:728763
21. Ingelsson E, Langenberg C, Hivert MF, Prokopenko I, Lyssenko V, Dupuis J et al (2010) Detailed physiologic characterization reveals diverse mechanisms for novel genetic Loci regulating glucose and insulin metabolism in humans. Diabetes 59:1266–1275
22. Liu Z, Habener JF (2010) Wnt signaling in pancreatic islets. Adv Exp Med Biol 654:391–419
23. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB et al (2011) Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 473:43–49

24. Richards EJ (2006) Inherited epigenetic variation–revisiting soft inheritance. Nat Rev Genet 7:395–401
25. Richards EJ (2008) Population epigenetics. Curr Opin Genet Dev 18:221–226
26. Bhandare R, Schug J, Le Lay J, Fox A, Smirnova O, Liu C et al (2010) Genome-wide analysis of histone modifications in human pancreatic islets. Genome Res 20:428–433
27. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z et al (2007) High-resolution profiling of histone methylations in the human genome. Cell 129:823–837
28. Mutskov V, Felsenfeld G (2009) The human insulin gene is part of a large open chromatin domain specific for human islets. Proc Natl Acad Sci USA 106:17419–17424
29. Melkman-Zehavi T, Oren R, Kredo-Russo S, Shapira T, Mandelbaum AD, Rivkin N, Nir T, Lennox KA, Behlke MA, Dor Y, Hornstein E (2011) miRNAs control insulin content in pancreatic beta-cells via downregulation of transcriptional repressors. EMBO J 30:835–845
30. Schwartz MW, Porte D Jr (2005) Diabetes, obesity, and the brain. Science 307:375–379
31. Simpson KA, Martin NM, Bloom SR (2008) Hypothalamic regulation of appetite. Expert Rev Endocrinol Metab 3:577–592
32. Vucetic Z, Reyes MT (2010) Central dopaminergic circuitry controlling food intake and reward: implications for the regulation of obesity. Wiley Interdiscip Rev Syst Biol Med 2:577–593
33. Bellet MM, Sassone-Corsi P (2010) Mammalian circadian clock and metabolism – the epigenetic link. J Cell Sci 123:3837–3848
34. McConkie-Rosell A, Lachiewicz AM, Spiridigliozzi GA, Tarleton J, Schoenwald S, Phelan MC et al (1993) Evidence that methylation of the FMR-I locus is responsible for variable phenotypic expression of the fragile X syndrome. Am J Hum Genet 53:800–809
35. Stöger R, Kajimura TM, Brown WT, Laird CD (1997) Epigenetic variation illustrated by DNA methylation patterns of the fragile-X gene FMR1. Hum Mol Genet 6:1791–1801
36. Terracciano A, Chiurazzi P, Neri G (2005) Fragile X syndrome. Am J Med Genet C Semin Med Genet 137C:32–37
37. Laird CD (1987) Proposed mechanism of inheritance and expression of the human fragile-X syndrome of mental retardation. Genetics 117:587–599
38. Willemsen R, Bontekoe CJ, Severijnen LA, Oostra BA (2002) Timing of the absence of FMR1 expression in full mutation chorionic villi. Hum Genet 110:601–605
39. Waterland RA, Garza C (1999) Potential mechanisms of metabolic imprinting that lead to chronic disease. Am J Clin Nutr 69:179–197
40. Hales CN, Ozanne SE (2003) For debate: Fetal and early postnatal growth restriction lead to diabetes, the metabolic syndrome and renal failure. Diabetologia 46:1013–1019
41. Simmons R (2005) Developmental origins of adult metabolic disease: concepts and controversies. Trends Endocrinol Metab 16:390–394
42. Gallou-Kabani C, Junien C (2005) Nutritional epigenomics of metabolic syndrome: new perspective against the epidemic. Diabetes 54:1899–1906
43. Holness MJ, Sugden MC (2006) Epigenetic regulation of metabolism in children born small for gestational age. Curr Opin Clin Nutr Metab Care 9:482–488
44. Stein Z (1975) Famine and human development: the Dutch hunger winter of 1944–1945. Oxford University Press, New York
45. Stein AD, Kahn HS, Rundle A, Zybert PA, van der Pal-de BK, Lumey LH (2007) Anthropometric measures in middle age after exposure to famine during gestation: evidence from the Dutch famine. Am J Clin Nutr 85:869–876
46. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES et al (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci USA 105:17046–17049
47. Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD et al (2009) DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Hum Mol Genet 18:4046–4053
48. Wellen KE, Hotamisligil GS (2003) Obesity-induced inflammatory changes in adipose tissue. J Clin Invest 112:1785–1788

49. Xu H, Barnes GT, Yang Q, Tan G, Yang D, Chou CJ et al (2003) Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance. J Clin Invest 112:1821–1830
50. Weisberg SP, McCann D, Desai M, Rosenbaum M, Leibel RL, Ferrante AW Jr (2003) Obesity is associated with macrophage accumulation in adipose tissue. J Clin Invest 112:1796–1808
51. Kintscher U, Hartge M, Hess K, Foryst-Ludwig A, Clemenz M, Wabitsch M et al (2008) T-lymphocyte infiltration in visceral adipose tissue: a primary event in adipose tissue inflammation and the development of obesity-mediated insulin resistance. Arterioscler Thromb Vasc Biol 28:1304–1310
52. Stöger R (2006) In vivo methylation patterns of the Leptin promoter in human and mouse. Epigenetics 1:155–162
53. Noer A, Sorensen AL, Boquest AC, Collas P (2006) Stable CpG hypomethylation of adipogenic promoters in freshly isolated, cultured, and differentiated mesenchymal stem cells from adipose tissue. Mol Biol Cell 17:3543–3556
54. El-Maarri O, Becker T, Junen J, Manzoor SS, Diaz-Lacava A, Schwaab R et al (2007) Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. Hum Genet 122:505–514
55. Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T et al (2010) Personalized epigenomic signatures that are stable over time and covary with body mass index. Sci Transl Med 2:49–67
56. Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddeloh JA et al (2008) Comprehensive high-throughput arrays for relative methylation (CHARM). Genome Res 18:780–790
57. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G et al (2007) Age, gene/environment susceptibility-Reykjavik study: multidisciplinary applied phenomics. Am J Epidemiol 165:1076–1087
58. Nair S, Lee YH, Rousseau E, Cam M, Tataranni PA, Baier LJ et al (2005) Increased expression of inflammation-related genes in cultured preadipocytes/stromal vascular cells from obese compared with non-obese Pima Indians. Diabetologia 48:1784–1788
59. O'Hara A, Lim FL, Mazzatti DJ, Trayhurn P (2009) Microarray analysis identifies matrix metalloproteinases (MMPs) as key genes whose expression is up-regulated in human adipocytes by macrophage-conditioned medium. Pflugers Arch 458:1103–1114
60. Cardellini M, Menghini R, Martelli E, Casagrande V, Marino A, Rizza S et al (2009) TIMP3 is reduced in atherosclerotic plaques from subjects with type 2 diabetes and increased by SirT1. Diabetes 58:2396–2401
61. Kent CF, Daskalchuk T, Cook L, Sokolowski MB, Greenspan RJ (2009) The Drosophila foraging gene mediates adult plasticity and gene-environment interactions in behaviour, metabolites, and gene expression in response to food deprivation. PLoS Genet 5:e1000609
62. Zakharkin SO, Belay AT, Fernandez JR, De Luca V, Kennedy JL, Sokolowski MB et al (2005) Lack of association between polymorphism of the human cyclic GMP-dependent protein kinase gene and obesity. Int J Obes (Lond) 29:872–874
63. Ogawa O, Eccles MR, Szeto J, McNoe LA, Yun K, Maw MA et al (1993) Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. Nature 362:749–751
64. Feinberg AP (2007) Phenotypic plasticity and the epigenetics of human disease. Nature 447:433–440
65. Constancia M, Kelsey G, Reik W (2004) Resourceful imprinting. Nature 432:53–57
66. Charalambous M, da Rocha ST, Ferguson-Smith AC (2007) Genomic imprinting, growth control and the allocation of nutritional resources: consequences for postnatal life. Curr Opin Endocrinol Diabetes Obes 14:3–12
67. Moore T, Haig D (1991) Genomic imprinting in mammalian development: a parental tug-of-war. Trends Genet 7:45–49
68. Haig D (2000) The kinship theory of genomic imprinting. Annu Rev Ecol Syst 31:9–32

69. Haig D (2004) Genomic imprinting and kinship: how good is the evidence? Annu Rev Genet 38:553–585
70. Parker-Katiraee L, Carson AR, Yamada T, Arnaud P, Feil R, Abu-Amero SN, Moore GE, Kaneda M, Perry GH, Stone AC, Lee C, Meguro-Horike M et al (2007) Identification of the imprinted KLF14 transcription factor undergoing human-specific accelerated evolution. PLoS Genet 3:e65
71. Small KS, Hedman AK, Grundberg E, Nica AC, Thorleifsson G, Kong A, Thorsteindottir U, Shin SY, Richards HB, Soranzo N, Ahmadi KR, Lindgren CM et al (2011) Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. Nat Genet 43:561–564
72. Shapira NA, Lessig MC, He AG, James GA, Driscoll DJ, Liu Y (2005) Satiety dysfunction in Prader-Willi syndrome demonstrated by fMRI. J Neurol Neurosurg Psychiatry 76:260–262
73. Herzing LB, Kim SJ, Cook EH Jr, Ledbetter DH (2001) The human aminophospholipid-transporting ATPase gene ATP10C maps adjacent to UBE3A and exhibits similar imprinted expression. Am J Hum Genet 68:1501–1505
74. Hogart A, Patzel KA, LaSalle JM (2008) Gender influences monoallelic expression of ATP10A in human brain. Hum Genet 124:235–242
75. Kelsey G (2009) Epigenetics and imprinted genes: insights from the imprinted Gnas locus. Horm Res 71(Suppl 2):22–29
76. Weinstein LS, Xie T, Qasem A, Wang J, Chen M (2010) The role of GNAS and other imprinted genes in the development of obesity. Int J Obes (Lond) 34:6–17
77. Kassem SA, Ariel I, Thornton PS, Hussain K, Smith V, Lindley KJ et al (2001) p57(KIP2) expression in normal islet cells and in hyperinsulinism of infancy. Diabetes 50:2763–2769
78. Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, Furuta H et al (2008) Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. Nat Genet 40:1092–1097
79. Kong A, Steinthorsdottir V, Masson G, Thorleifsson G, Sulem P, Besenbacher S et al (2009) Parental origin of sequence variants associated with complex diseases. Nature 462:868–874
80. Sakatani T, Wei M, Katoh M, Okita C, Wada D, Mitsuya K et al (2001) Epigenetic heterogeneity at imprinted loci in normal populations. Biochem Biophys Res Commun 283:1124–1130
81. Cui H, Onyango P, Brandenburg S, Wu Y, Hsieh CL, Feinberg AP (2002) Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. Cancer Res 62:6442–6446
82. Petrik J, Pell JM, Arany E, McDonald TJ, Dean WL, Reik W et al (1999) Overexpression of insulin-like growth factor-II in transgenic mice is associated with pancreatic islet cell hyperplasia. Endocrinology 140:2353–2363
83. Milo-Landesman D, Efrat S (2002) Growth factor-dependent proliferation of the pancreatic beta-cell line betaTC-tet: an assay for beta-cell mitogenic factors. Int J Exp Diabetes Res 3:69–74
84. Devedjian JC, George M, Casellas A, Pujol A, Visa J, Pelegrin M et al (2000) Transgenic mice overexpressing insulin-like growth factor-II in beta cells develop type 2 diabetes. J Clin Invest 105:731–740
85. Bouwens L, Rooman I (2005) Regulation of pancreatic beta-cell mass. Physiol Rev 85:1255–1270
86. Trajkovski M, Hausser J, Soutschek J, Bhat B, Akin A, Zavolan M, Heim MH, Stoffel M (2011) MicroRNAs 103 and 107 regulate insulin sensitivity. Nature 474:649–653

Chapter 21
Epigenetic Epidemiology of Atherosclerosis

Einari Aavik, Mikko P. Turunen, and Seppo Ylä-Herttuala

Contents

21.1 Introduction 425
21.2 Etiology of Atherosclerosis 426
21.3 Hypercholesterolemia and Potential Dietary Effects: An Epigenetic View 426
21.4 The Role of Homocysteinemia in Atherosclerosis 427
21.5 DNA Methylation and Atherosclerosis 428
  21.5.1 Hypermethylation and Atherosclerosis 429
  21.5.2 Hypomethylation and Atherosclerosis? 431
21.6 Histone Code and Atherosclerosis 432
21.7 Small RNAs and Atherosclerosis 432
21.8 Conclusions and Future Perspectives 433
References 434

E. Aavik
Department of Biotechnology and Molecular Medicine, A. I. Virtanen Institute,
University of Eastern Finland, Neulaniementie 2, Kuopio, Finland
e-mail: einari.aavik@uef.fi

M.P. Turunen
Ark Therapeutics LTD, Microkatu 1S, Kuopio, Finland
e-mail: mikko.turunen@uef.fi

S. Ylä-Herttuala (✉)
Department of Biotechnology and Molecular Medicine, A. I. Virtanen Institute,
University of Eastern Finland, Neulaniementie 2, Kuopio, Finland

Gene Therapy Unit, Kuopio University Hospital, Kuopio, Finland
e-mail: Seppo.Ylaherttuala@uef.fi

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2_21,

© Springer Science+Business Media B.V. 2012

Abstract The multitude of atherosclerotic risk factors can be divided into three main categories – genetical, environmental, and habitual issues. This classification can be simplified even further to genetical and environmental issues only as human habits manifest as environmental factors at cellular level. The link between large artery cellular responses to changed environment and epigenetic modifications of the chromosomal DNA has been established only recently. This review discusses the general epigenetic mechanisms operational in atherosclerotic tissue.

Abbreviations

ApoE	Apolipoprotein E
CHD	Coronary heart disease
CpG	Cytosine-guanine dinucleotide
CTCF	CCCTC-binding factor
ECM	Extracellular matrix
ERα	Estrogen receptor alpha
ERβ	Estrogen receptor beta
FGF2	Fibroblast growth factor 2
HDAC	Histone deacetylase
HAT	Histone acetyl transferase
Hcy	Homocysteine
HIF1α	Hypoxia inducible factor 1α
IGF2	Insulin-like growth factor 2
KAT2B	Lysine acetyltransferase 2B
LDL	Low density lipoprotein
LDLR	LDL receptor
MBD	Methyl-cytosine binding protein
NOS	Nitric oxide synthase
POL2	RNA polymerase II
PBL	Peripheral blood lymphocytes
RISC	RNA induced silencing
SAH	S-Adenosylhomocysteine
SAM	S-adenosylmethionine
SiRNA	small interfering RNA
SMC	Smooth muscle cell
SNP	Single nucleotide polymorphism
SRF	Serum response factor
TGS	Transcriptional gene silencing

21.1 Introduction

Atherosclerosis is a complex disease, and most of the complexity lies in the interplay of factors contributing to the disease etiology. Male gender, genetic variability, and some environmental and habitual issues have been recognized as the most important risk factors for atherosclerosis. The impact of genetic variability on atherosclerosis, with only a few exceptions, has been very difficult to assess as the contribution of the identified single nucleotide polymorphism (SNP) is usually only marginally above normal variability [1]. The chromosomal region 9p21 appears the strongest polymorphic site that has been related to atherosclerosis, but the statistical significance of the detected variability still means only a small increase in the risk for disease development [2]. On the other hand, research done on experimental animals has shown that there are strong non-genetic factors that can be linked to the increased risk for atherosclerosis in the progeny [3]. These hereditary events not explainable by genetics have attracted attention long before any possible mechanisms for that had been proposed and the widely used term epigenetics was invented by Conrad Waddington already in 1942. Waddington defined epigenetics as something that is needed in addition to genes "to bring the phenotype into being" [4].

Epigenetic mechanisms include several levels of regulation which contribute to the activity of gene expression. The principal mechanisms of epigenetic regulation in mammals are DNA methylation and modifications of histone tails which alter chromatin structure [5]. Importantly, these changes in chromosomal structure are not permanent and they leave the primary genetic code (= nucleotide sequence) unchanged. Thus, epigenetic regulation necessitates thinking that a particular phenotype is not entirely defined by the DNA primary sequence. Moreover, the most striking piece of information from epigenetic studies is supporting the assumption that some acquired traits can be passed to progeny, both in humans [6] and in experimental animals although direct evidence in humans is still lacking [3].

Genetic variability and epigenetic mechanisms of gene regulation are intricately combined in the phenomenon of sexual dimorphism [7]. The impact of sexually dimorphic genes on complex diseases can only be analyzed after sequencing of possibly thousands of genomes with the ability to collect data separately from maternal and paternal genomes. Human haploid genome lengths can vary by up to 3%, when all insertions and deletions are included. Recent advancements in DNA sequencing technology have increased the speed of sequence data collection by eight orders of magnitude during the last 25 years, but the biggest challenges are still in the computational efficiency of simultaneous comparisons of thousands of genomes and relating findings to phenotypic information [8].

Most studies relating epigenetics to human pathologies have focused on cancer. Studies about the involvement of epigenetic mechanisms in cardiovascular diseases have been disappointingly sparse. The first study demonstrating a link between hypercholesterolemia and chromatin structure was published already in 1980 [9]. Acridine orange intercalated to unwound DNA regions devoid of histones and made it possible to distinguish between actively transcribed and/or replicating euchromatin

and densely packed heterochromatin. Lehmann and coworkers could detect more acridine orange binding nuclei in aortic wall of hypercholesterolemic rabbit [9]. Already decades ago researchers noticed that commonly recognized risk factors (genetic background, diet, and lifestyle) do not explain all cases of cardiovascular morbidity [10]. Studies in the area of epigenetics may significantly contribute to our understanding of the pathogenesis of cardiovascular diseases and suggest new therapeutic possibilities for atherosclerosis-related diseases.

21.2 Etiology of Atherosclerosis

Atherosclerosis is a chronic disease of large and medium sized arteries which is characterized by accumulation of cholesterol in the arterial wall together with proliferation of arterial smooth muscle cells (SMC) and accumulation of extracellular matrix components which lead to occlusion of blood vessels, myocardial infarction, stroke, aneurysms, and peripheral vascular disease [11–14]. Chronic inflammatory response with infiltration of monocyte–macrophages and T-cells and endothelial dysfunction are also prominent features of atherogenesis [15, 16].

Atherosclerosis develops over several decades of life. Its incidence reached epidemics in the Western world in the 1960s and 1970s whereafter the incidence declined, thanks to effective blood lipid lowering, treatment of hypertension, dietary changes, and the decline of smoking in many countries. However, in other parts of the world, such as in the Far East and in Third World countries, its incidence is rapidly increasing. It is obvious that unfavorable changes in dietary habits and smoking contribute to the changes via increases in blood lipid levels and uptake of cholesterol in the arterial wall. However, it has been speculated that epigenetic changes might also contribute to atherogenesis since it involves polyclonal proliferation of SMCs and some dietary components, recognized as risk factors of atherosclerosis, may affect methylation machinery in the arterial cells [17, 18] (Fig. 21.1).

21.3 Hypercholesterolemia and Potential Dietary Effects: An Epigenetic View

Hypercholesterolemia is central to the initial stages of atherosclerosis in large arteries [19]. Genetic studies have provided support for the existence of inherited patterns of cardiovascular diseases and maternal hypercholesterolemia appears to be associated with a higher incidence and a faster progression of atherosclerosis in humans [6], rabbits [20], and mice [21, 22]. However, there is at least one report where the link between maternal hypercholesterolemia and enhanced atherosclerosis in mouse progeny was not demonstrated [23].

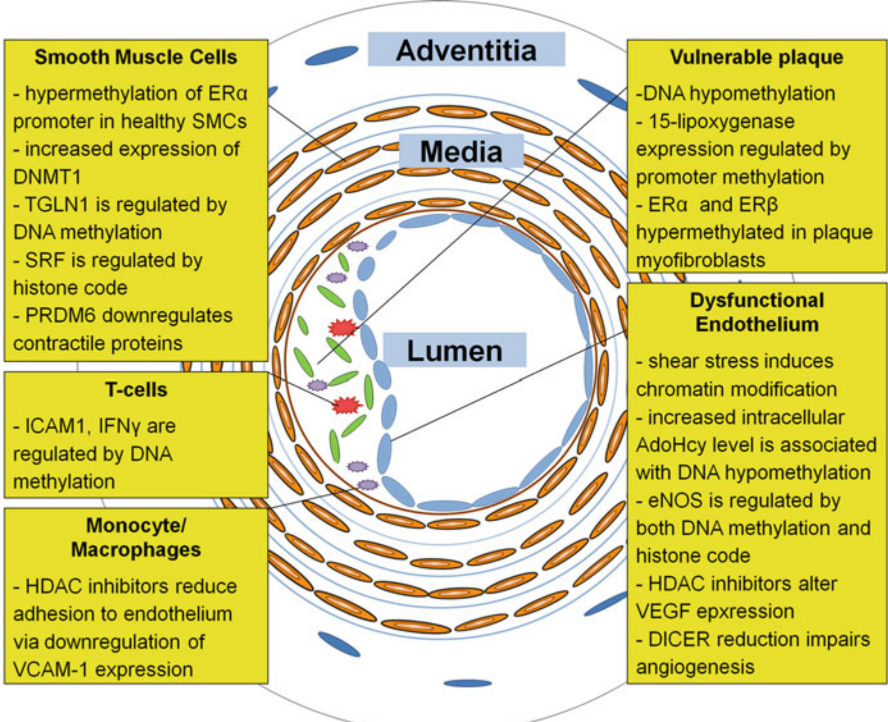


Fig. 21.1 Summary of epigenetic effects in atherosclerotic vascular wall

Dietary modification can have a marked effect on epigenetic modifications of the genome. A diet deficient for methionine and folate impairs the biosynthesis of necessary methyl donor S-adenosylmethionine (SAM), and this, in turn, can lead to altered imprinting of IGF2 [24]. While the risk of colorectal cancer is linked to dietary folate deficiency and some variants of methylenetetrahydrofolate reductase that direct folate to remethylation of homocysteine (Hcy) to methionine [25], high folic acid supplementation has also been linked to a promotion of underlying colorectal cancer [26–28]. Epigenetic mechanisms may connect dietary and other environmental exposures with the risk of cardiovascular disease.

21.4 The Role of Homocysteinemia in Atherosclerosis

About 10% of human population may be affected by hyperhomocysteinemia and elevated plasma total Hcy is associated with increased risk for ischemic heart disease and stroke [29, 30]. Hcy is metabolized to methionine after activation to SAM,

which is known to act as a methyl donor. Underlying pathogenic mechanisms that lead to endothelial dysfunction and atherosclerosis have been thoroughly studied but remain poorly understood. Recent evidence suggests that the pathologic role of hyperhomocysteinemia in vascular diseases might be mediated via S-adenosylhomocysteine (SAH) accumulation and DNA methylation [31].

Castro et al. have observed that patients with vascular disease have increased plasma Hcy and SAH concentrations and disturbed global DNA methylation status [31]. High blood Hcy levels (> 75 μmol/l) correlate with DNA hypomethylation and atherosclerosis [32] and can lead to a 35% reduction in DNA methylation status of peripheral blood lymphocytes [31]. High Hcy concentration competes with SAM for binding sites on DNA methylases and may lead to passive loss of methylation in replicating DNA [33]. Aberrant DNA methylation secondary to environmental and nutritional factors has been documented in atherogenesis, both in humans and in animal models [34, 35].

Sharma et al. [36] reviewed literature of 135 genes which either modulate the blood level of Hcy or are regulated by the elevated level of Hcy. Mapping of these genes to their respective pathways revealed that an elevated level of Hcy may lead to atherosclerosis either by affecting lipid metabolism and transport or via oxidative stress and endoplasmic reticulum stress. Elevated level of Hcy also decreases the bioavailability of nitric oxide [36].

21.5 DNA Methylation and Atherosclerosis

DNA methylation is an epigenetic mechanism maintaining chromosomal DNA structure and regulating gene expression (e.g., X chromosome inactivation, imprinting, controlling transposable genetic element mobility and activity of developmentally regulated genes). DNA methylation pattern is replicated in embryogenesis and in tissue repair at each mitotic cycle [37]. Methyl-cytosine binding (MBD) proteins effectively prevent gene transcription from methylated promoters [5] and induce stable blockade of transcription by recruitment of transcriptional co-repressors with histone methylase and deacetylase activities [38, 39].

In healthy human somatic cells about 1% of all cytosines appear to be methylated [40], but there is some variation between tissues and even 3–4% of all cytosines have been reported to be methylated in some tissues [41]. This corresponds to 70–90% of all cytosines in CpG dinucleotides [42]. CpG dinucleotides present in the 5′ regulatory areas are mostly unmethylated, whereas CpG islands in intronic and repetitive DNA elements are usually methylated [42]. Regulatory DNA regions rich in CpG dinucleotides are known as CpG islands with lengths varying from a few hundred to several thousand base pairs [37]. Some components of the DNA methylation machinery are shared with the enzyme complex involved in histone methylation (see below; [39, 43]).

It is now generally accepted that DNA methylation patterns are replicated during embryonal development and that changes introduced into DNA methylation pattern

in early life can be passed to the progeny at least in rodents [3]. There is also some evidence that the epigenetic programming in early life can be reversed by pharmacological therapy later in adulthood [6, 44–46].

Age-related changes in global methylation status of DNA appear to be genetically determined. Bjornsson et al. have shown that some individuals gain and others loose cytosine methylation in their DNA over time and the paired data on intra-individual changes could be clustered family-wise [47]. This result may explain why the averaged levels of DNA methylation in age groups did not change [48].

SMCs make up most of the cellular mass in arterial wall, and consequently epigenetic changes detected in vascular wall are mainly a characteristic of SMCs. Transgelin is the earliest marker of SMC differentiation and is regulated also by DNA methylation [49]. Another example of epigenetic regulation is the estrogen receptor (ER) α promoter, which is hypermethylated in proliferating arterial SMCs, but not in endothelial cells [50]. Endothelial nitric oxide synthase (eNOS) is hypomethylated in endothelium, but heavily methylated in other cell types [51], including SMC lines [52]. Inducible NOS (iNOS) is expressed in atherosclerotic plaque neointima [53, 54], but repressed by methylation in most other tissues [51].

21.5.1 Hypermethylation and Atherosclerosis

Hypermethylation of promoter CpG islands is characteristic to transcriptional silencing. Gene repression by DNA methylation is achieved either by prevention of the binding of transcription factors to their cognate sites by cytosine methylation or by occupying the promoter region by one of the five methyl-cytosine recognizing proteins MBD1, MBD2, MBD4, MeCP2, or Kaiso [5]. In atherosclerosis the binding of certain transcription factors, such as HIF1α [55], myc [56], and CTCF [57] have been shown to be affected by DNA methylation. CTCF target sites on paternal chromosomal DNA are methylated and CTCF regulates maternal genes like IGF2 [58]. YY1 is another eukaryotic insulator regulating paternal genes (e.g., XIST, GNAS, and PEG3, which are all methylated on maternal chromosome). Kaiso recognizes also unmethylated TNGCAGGA and interacts with CTCF [59].

Transcriptional silencing by hypermethylation is poorly studied in other diseases than cancer. There is some existing controversy about the existence of DNA hypermethylation in peripheral blood lymphocytes (PBL) in coronary heart disease (CHD) patients. Two recent articles [60, 61] have reported an increase in PBL methylation status of CHD patients, whereas some earlier reports have documented opposite findings [32, 62].

The complex nature of atherosclerosis can be seen in the dualistic effects of epigenetic regulation of specific genes (Table 21.1). For example, while hyperhomocysteinemia has the potential to reduce DNA methylation in atherosclerotic patient PBL [31], some promoters like iNOS [63] in cultured human monocytes and FGF2 [64] in cultured human coronary artery endothelial cells are still hypermethylated. Increased methylation of ERα promoter has been detected in atheromatous

Table 21.1 Genes related to atherosclerotic diseases which are at least partly regulated by epigenetic mechanisms

Gene	Function	Target	Mechanism	Reference
eNOS	Blood pressure control	Endothelium	DNA methylation and histone code	[52, 104]
iNOS	Blood pressure control	Inflammation, macrophages	DNA methylation and histone code	[51]
Fads2	Fatty acid desaturase	Contributes to pathology of Hcy	DNA methylation	[105]
c-*fos*	Transcription factor	Shear stress	Histone code	[106]
Estrogen receptor α	Transcription factor	Coronary artery	DNA methylation	[50]
Estrogen receptor β	Transcription factor	Coronary artery Mammary artery Carotid artery Femoral artery Saphenous vein	DNA methylation	[66]
P66Shc	Docking protein in cell signaling	End-stage renal disease	DNA methylation	[107]
15-LO (ALOX15)	Lipid peroxidation	Pathogenesis of atherosclerosis	DNA methylation	[108]
EC-SOD	Oxidative stress	Pathogenesis of atherosclerosis	DNA methylation	[17]
H19/Igf2	H19 – tumor suppressor IGF2 – fetal growth factor	Regulated by hyperhomo-cysteinemia	Imprinting, normally maternally expressed	[105, 109]
MMP-2	Matrix degradation	Extracellular matrix	DNA methylation	[110]
MMP-7	Matrix degradation	Extracellular matrix	DNA methylation	[110]
MMP-9	Matrix degradation	Extracellular matrix	DNA methylation	[110]
TIMP-3	Inhibitor of ECM degradation	Extracellular matrix	DNA methylation	[111]
IFN-γ	Cytokine	Inflammatory response	DNA methylation	[112]
PDGF-A	Growth factor	Cell proliferation	DNA methylation	[113]
ICAM-1	Adhesion	Inflammatory reactions	DNA methylation	[114]
p53	Tumor suppressor	Apoptosis	DNA methylation	[115]

proliferating SMCs [50, 65]. Also, ERβ promoter shows a high level of methylation in cells from the plaque area compared to non-plaque regions [66]. Although it is apparent that ERα and ERβ play a role in actively proliferating atheromatous SMCs, the interpretation of the exact role of ER in vivo remains unclear. There is a steady age-dependent increase in ER promoter methylation, which may reach 99% methylation level in the arterial SMC isolated from elderly cadaveric patients [65].

Age-related hypermethylation of promoters is not limited to ERα [67, 68] and hypermethylated promoters of immunomodulatory genes and tumor suppressors [68] may have profound effects also on the pathogenesis of atherosclerosis.

Inflammation promotes global hypermethylation of peripheral blood lymphocyte DNA, and inflammation in dialysis patients correlates with increased cardiovascular mortality [69].

21.5.2 Hypomethylation and Atherosclerosis?

Loss of DNA methylation can be either passive (DNA replication-dependent) or active (DNA-replication independent) [42]. It is not known which of these processes predominates in atherosclerosis, although active replication of SMCs in neointima seems to favor the idea of passive hypomethylation.

Ageing and atherosclerosis occur simultaneously, and it is somewhat difficult to distinguish between these two processes as causes for global hypomethylation. Global hypomethylation of genomic DNA has been considered as a characteristic of aging [70], but recent data show that hypomethylation is characteristic to only 20–30% of the population [47]. Global hypomethylation is common in some other pathological conditions like tumor growth [71], hyperhomocysteinemia [32], and schizophrenia [72]. Global hypomethylation is also characteristic of areas of SMC proliferation in advanced human atherosclerotic plaques where a 9% decrease in cytosine methylation has been detected compared to normal artery [41]. In mice atherosclerosis and aging can be studied separately, since in ApoE knock-out mice DNA hypomethylation appears to precede any histological signs of atherosclerosis [73]. Similar hypomethylation is present in atheromas of ApoE knockout mice [73] and in neointimal thickenings of New Zealand White rabbit aortas [17, 41]. A significant decrease in DNA methylation status has also been detected in ApoE knockout mouse aortas already as early as 4 weeks of age, and considerable advancement in demethylation was detectable at 6 months [73]. Whereas hypomethylation of eNOS and iNOS promoter areas are characteristic of healthy human vascular endothelial cells, reduction in the methylation status of extracellular superoxide dismutase [17] and 15-lipoxygenase [41] promoters are characteristic of human atherosclerotic lesions.

Transcriptional regulator specifically recognizing unmethylated DNA is CXXC1 [74] which acts in concert with histone 3 lysine 4 methylases MLL, MLL2, MLL3, and hSET1 [75]. Hypomethylation of DNA repeated sequences can lead to activation of transposable elements which, in turn, may cause insertional mutagenesis and chromosomal instability by DNA recombination events as shown in human myeloid leukemia [76]. However, hypomethylation in atherosclerotic lesions is a somewhat paradoxical finding since DNMT1 mRNA has been shown to be upregulated in atheroma tissue [41]. Atherosclerosis is also associated with humoral and adaptive immunological responses, including autoantibody formation against oxidized LDL [77]. In this context it is interesting to note that hypomethylating agents can promote autoimmunity in mice [78].

21.6 Histone Code and Atherosclerosis

Inflammatory processes in atherosclerosis are largely mediated by transcription factor NFkB activation which includes histone code modulation. Histone acetylation of KAT2B has a costimulatory effect on NFkB-dependent transcription of inflammatory genes in PBL from diabetic humans [79, 80], whereas histone deacetylase activation suppresses inflammation [81]. In addition to histone modifications, both KAT2B and p300 can acetylate NFkB proteins directly and thereby increase its binding affinity to DNA and thus its transactivating activity [82, 83]. Trichostatin A, a specific inhibitor of histone deacetylase, significantly increased the formation of fatty streak lesions and macrophage infiltration in LDLR knock-out mice [84]. Dysregulation of epigenetic histone modifications may be a major underlying mechanism for metabolic memory and sustained proinflammatory phenotype in SMCs in diabetic vascular disease [43, 85].

Histone lysine demethylase KDM1 (LSD1), a component of the CoREST2CtBP co-repressor complex, was discovered in 2004 and since then a number of histone demethylases have been identified and shown to play important roles in the regulation of gene expression, as well as in cellular differentiation and development [86]. Recently, Reddy et al. showed that SMCs derived from a mouse model of type 2 diabetes displayed enhanced inflammatory gene expression and proatherogenic responses [85]. These results demonstrated functional roles for LSD1 and H3K4 methylation in SMCs and inflammation. Dysregulation of their actions may be a major mechanism for chronic vascular inflammation, and metabolic memory associated with diabetic vascular complications.

21.7 Small RNAs and Atherosclerosis

Small RNA molecules have been shown to regulate gene transcription by interacting with the promoter region and modifying the histone code [87–90]. A process known as transcriptional gene silencing (TGS) involves promoter-targeted small interfering RNA (siRNA) and leads to silent state epigenetic profile containing dimethylated histone H3 lysine 9 (H3K9me2) and trimethylated histone H3 lysine 27 (H3K27me3) [91]. This process was first recognized in plants but has since been identified also in mammals [89, 92]. Interestingly, besides transcriptional gene silencing, the small RNAs have been reported to induce gene activation [88, 90]. This is associated with a loss of H3K9me2 at the targeted promoter sequences and seems to require a member of a RISC complex, Argonaute-2 protein. TGS has earlier been noticed to require Argonaute-1 and -2 [93].

The chromatin modifications induced by siRNAs have been shown to require RNA polymerase II (POL2) and the transcription of the homologous target sequence in *S. pombe* [94]. Truncation of the C-terminal domain of POL2 did not cause any deficiencies in transcription or in production of siRNAs but it impaired

the RNAi-directed chromatin modifications. Thus, POL2 via its C-terminal domain might stabilize the interactions between the components of RNAi machinery and the transcript (the RNA/RNA model). It has been also proposed that during transcription the passage of POL2 might act by opening the chromatin and providing access for siRNAs to the complementary DNA sequence, allowing the siRNA-DNA interactions to take place (the RNA/DNA model). One suggested mechanism of action for promoter targeted small RNA mediated gene activation and silencing is that the antisense strand of the small RNA binds to a complementary non-coding promoter-associated antisense RNA [95–97].

21.8 Conclusions and Future Perspectives

Clonal proliferation of lesion SMCs has been documented in human atherosclerotic lesions [98]. It is likely that some SMCs acquire properties that enable them to proliferate more rapidly and thus increase their progeny in the lesion area. No data is available about possible association of epigenetic modifications with this clonal proliferation process. However, SMC-specific gene expression is dependent on regulatory elements controlled by epigenetic mechanisms altering chromatin structure. The binding of serum response factor (SRF) to SMC-specific promoters is associated with modification of adjacent histones by methylation and acetylation [99, 100]. Likewise, myocardin-induced SMC-specific transcription is modulated by the acetylation of histones flanking SRF-binding sites [101]. A recently discovered histone methylase PRDM6 specifically downregulates contractile protein expression in SMCs and aids switching to the rapidly proliferating synthetic phenotype [102]. Because of the rapid fluctuation in the incidence of cardiovascular diseases within various countries, epigenetic effects are likely involved since human genetic code does not change that rapidly through classical mutations. Similarly, epigenetic mechanisms could explain, at least partly, marked increases with age of several common diseases, including cardiovascular diseases.

Potential links of direct dietary effects on vascular cells and their gene expression via epigenetic mechanisms cannot be excluded. This is an area that clearly requires further study. Several drugs are currently in development that target epigenetic changes in malignant cells and processes like autoimmunity. Testing these compounds for potential effects on atherosclerosis might prove fruitful and provide direct evidence whether these processes are involved in the pathogenesis of cardiovascular diseases.

Double stranded RNAs have recently been shown to influence epigenetic properties of several cell types [89, 91, 92, 103]. Indeed, microRNAs expressed from genes and intergenic areas constitute a large pool of potential effectors and modifiers of gene expression, some effects of which are likely to involve epigenetic changes at DNA methylation and histone level [90]. The role of microRNAs in fine tuning gene expression and cellular genetic memory are areas where further studies are clearly warranted.

In summary, epigenetic changes have been detected in atherosclerotic lesions. These changes may contribute to the pathogenesis of vascular lesions and modify properties and gene expression in lesion cells. Treatments directed toward modification of epigenetic changes in vascular cells may provide new possibilities for the treatment of cardiovascular diseases.

Acknowledgments This work was supported by Finnish Academy, Leducq Foundation, CliniGene EU grant, and Ark Therapeutics Ltd.

References

1. Moore JH, Asselbergs FW, Williams SM (2010) Bioinformatics challenges for genome-wide association studies. Bioinformatics 26:445–455
2. Palomaki GE, Melillo S, Bradley LA (2010) Association between 9p21 genomic markers and heart disease: a meta-analysis. JAMA 303:648–656
3. Collas P, Noer A, Timoskainen S (2007) Programming the genome in embryonic and somatic stem cells. J Cell Mol Med 11b:602–620
4. Goldberg AD, Allis CD, Bernstein E (2007) Epigenetics: a landscape takes shape. Cell 128:635–638
5. Matouk CC, Marsden PA (2008) Epigenetic regulation of vascular endothelial gene expression. Circ Res 102:873–887
6. Napoli C, Glass CK, Witztum JL, Deutsch R, D'Armiento FP, Palinski W (1999) Influence of maternal hypercholesterolaemia during pregnancy on progression of early atherosclerotic lesions in childhood: fate of early lesions in children (FELIC) study. Lancet 354:1234–1241
7. Gabory A, Attig L, Junien C (2009) Sexual dimorphism in environmental epigenetic programming. Mol Cell Endocrinol 304:8–18
8. Venter JC (2010) Multiple personal genomes await. Nature 464:676–677
9. Lehmann R, Denes R, Nienhaus R, Steinbach T, Lusztig G, Sanatger RM (1980) Alteration of chromatin in early experimental arteriosclerosis. Artery 8:288–293
10. Rose G (1964) Familial patterns in ischaemic heart disease. Br J Prev Soc Med 18:75–80
11. Lamon BD, Hajjar DP (2008) Inflammation at the molecular interface of atherogenesis: an anthropological journey. Am J Pathol 175:1253–1264
12. Yla-Herttuala S, Palinski W, Rosenfeld ME, Parthasarathy S, Carew TE, Butler S et al (1989) Evidence for the presence of oxidatively modified low density lipoprotein in atherosclerotic lesions of rabbit and man. J Clin Invest 84:1086–1095
13. Wissler RW (1991) Update on the pathogenesis of atherosclerosis. Am J Med 91(1B):3S–9S
14. Sanchez LA, Veith FJ (1998) Diagnosis and treatment of chronic lower extremity ischemia. Vasc Med 3:291–299
15. Lusis AJ (2000) Atherosclerosis. Nature 407:233–241
16. Ross R (1999) Atherosclerosis – an inflammatory disease. N Engl J Med 14:115–126
17. Laukkanen MO, Mannermaa S, Hiltunen MO, Aittomaki S, Airenne K, Janne J et al (1999) Local hypomethylation in atherosclerosis found in rabbit ec-sod gene. Arterioscler Thromb Vasc Biol 19:2171–2178
18. Hiltunen MO, Yla-Herttuala S (2003) DNA methylation, smooth muscle cells, and atherogenesis. Arterioscler Thromb Vasc Biol 23:1750–1753
19. Hansson GK (2005) Inflammation, atherosclerosis, and coronary artery disease. N Engl J Med 352:1685–1695
20. Napoli C, Witztum JL, Calara F, de Nigris F, Palinski W (2000) Maternal hypercholesterolemia enhances atherogenesis in normocholesterolemic rabbits, which is inhibited by antioxidant

or lipid-lowering intervention during pregnancy: an experimental model of atherogenic mechanisms in human fetuses. Circ Res 10:946–952

21. Alkemade FE, Gittenberger-de Groot AC, Schiel AE, VanMunsteren JC, Hogers B, van Vliet LS et al (2007) Intrauterine exposure to maternal atherosclerotic risk factors increases the susceptibility to atherosclerosis in adult life. Arterioscler Thromb Vasc Biol 10:2228–2235
22. Napoli C, de Nigris F, Welch JS, Calara FB, Stuart RO, Glass CK et al (2002) Maternal hypercholesterolemia during pregnancy promotes early atherogenesis in LDL receptor-deficient mice and alters aortic gene expression determined by microarray. Circulation 105:1360–1367
23. Madsen C, Dagnaes-Hansen F, Moller J, Falk E (2003) Hypercholesterolemia in pregnant mice does not affect atherosclerosis in adult offspring. Atherosclerosis 68:221–228
24. Dobosy JR, Fu VX, Desotelle JA, Srinivasan R, Kenowski ML, Almassi N et al (2008) A methyl-deficient diet modifies histone methylation and alters Igf2 and H19 repression in the prostate. Prostate 68:1187–1195
25. Kono S, Chen K (2005) Genetic polymorphisms of methylenetetrahydrofolate reductase and colorectal cancer and adenoma. Cancer Sci 96:535–542
26. Pellis L, Dommels Y, Venema D, Polanen A, Lips E, Baykus H et al (2008) High folic acid increases cell turnover and lowers differentiation and iron content in human HT29 colon cancer cells. Br J Nutr 99:703–708
27. Crott JW, Liu Z, Keyes MK, Choi SW, Jang H, Moyer MP et al (2008) Moderate folate depletion modulates the expression of selected genes involved in cell cycle, intracellular signaling and folate uptake in human colonic epithelial cell lines. J Nutr Biochem 19:328–335
28. Sauer J, Mason JB, Choi SW (2009) Too much folate: a risk factor for cancer and cardiovascular disease? Curr Opin Clin Nutr Metab Care 12(1):30–36
29. Boushey CJ, Beresford SA, Omenn GS, Motulsky AG (1995) A quantitative assessment of plasma homocysteine as a risk factor for vascular disease. Probable benefits of increasing folic acid intakes. JAMA 274(13):1049–1057
30. Homocysteine Studies Collaboration (2002) Homocysteine and risk of ischemic heart disease and stroke: a meta-analysis. JAMA 288(16):2015–2022
31. Castro R, Rivera I, Struys EA, Jansen EE, Ravasco P, Camilo ME et al (2003) Increased homocysteine and S-adenosylhomocysteine concentrations and DNA hypomethylation in vascular disease. Clin Chem 49(8):1292–1296
32. Lee ME, Wang H (1999) Homocysteine and hypomethylation: a novel link to vascular disease. Trends Cardiovasc Med 9(1–2):49–54
33. James SJ, Melnyk S, Pogribna M, Pogribny IP, Caudill MA (2002) Elevation in S-adenosylhomocysteine and DNA hypomethylation: potential epigenetic mechanism for homocysteine-related pathology. J Nutr 132(8 Suppl):2361S–2366S
34. Chao CL, Kuo TL, Lee YT (2000) Effects of methionine-induced hyperhomocysteinemia on endothelium-dependent vasodilation and oxidative status in healthy adults. Circulation 101(5):485–490
35. Dayal S, Arning E, Bottiglieri T, Boger RH, Sigmund CD, Faraci FM et al (2004) Cerebral vascular dysfunction mediated by superoxide in hyperhomocysteinemic mice. Stroke 35(8):1957–1962
36. Sharma P, Senthilkumar RD, Brahmachari V, Sundaramoorthy E, Mahajan A, Sharma A et al (2006) Mining literature for a comprehensive pathway analysis: a case study for retrieval of homocysteine related genes for genetic and epigenetic studies. Lipids Health Dis 5:1
37. Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. Annu Rev Biochem 74(1):481–514
38. Ikegami K, Ohgane J, Tanaka S, Yagi S, Shiota K (2009) Interplay between DNA methylation, histone modification and chromatin remodeling in stem cells and during development. Int J Dev Biol 53(2–3):203–214
39. Fuks F, Hurd PJ, Deplus R, Kouzarides T (2003) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Res 31(9):2305–2312

40. Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA et al (1982) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res 10(8):2709–2721
41. Hiltunen MO, Turunen MP, Hakkinen TP, Rutanen J, Hedman M, Makinen K et al (2002) DNA hypomethylation and methyltransferase expression in atherosclerotic lesions. Vasc Med 7(1):5–11
42. Miranda TB, Jones PA (2007) DNA methylation: the nuts and bolts of repression. J Cell Physiol 213(2):384–390
43. Villeneuve LM, Reddy MA, Lanting LL, Wang M, Meng L, Natarajan R (2008) Epigenetic histone H3 lysine 9 methylation in metabolic memory and inflammatory phenotype of vascular smooth muscle cells in diabetes. Proc Natl Acad Sci USA 105(26):9047–9052
44. Weaver IC, Champagne FA, Brown SE, Dymov S, Sharma S, Meaney MJ et al (2005) Reversal of maternal programming of stress responses in adult offspring through methyl supplementation: altering epigenetic marking later in life. J Neurosci 25(47):11045–11054
45. Weaver IC, Meaney MJ, Szyf M (2006) Maternal care effects on the hippocampal transcriptome and anxiety-mediated behaviors in the offspring that are reversible in adulthood. Proc Natl Acad Sci USA 103(9):3480–3485
46. Bannister AJ, Kouzarides T (2005) Reversing histone methylation. Nature 436(7054): 1103–1106
47. Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H et al (2008) Intra-individual change over time in DNA methylation with familial clustering. JAMA 299(24):2877–2883
48. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M et al (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet 38(12):1378–1385
49. Yamamura H, Masuda H, Ikeda W, Tokuyama T, Takagi M, Shibata N et al (1997) Structure and expression of the human SM22alpha gene, assignment of the gene to chromosome 11, and repression of the promoter activity by cytosine DNA methylation. J Biochem 122(1):157–167
50. Post WS, Goldschmidt-Clermont PJ, Wilhide CC, Heldman AW, Sussman MS, Ouyang P et al (1999) Methylation of the estrogen receptor gene is associated with aging and atherosclerosis in the cardiovascular system. Cardiovasc Res 43(4):985–991
51. Chan GC, Fish JE, Mawji IA, Leung DD, Rachlis AC, Marsden PA (2005) Epigenetic basis for the transcriptional hyporesponsiveness of the human inducible nitric oxide synthase gene in vascular endothelial cells. J Immunol 175(6):3846–3861
52. Fish JE, Matouk CC, Rachlis A, Lin S, Tai SC, D'Abreo C et al (2005) The expression of endothelial nitric-oxide synthase is controlled by a cell-specific histone code. J Biol Chem 280(26):24824–24838
53. Luoma JS, Stralin P, Marklund SL, Hiltunen TP, Sarkioja T, Yla-Herttuala S (1998) Expression of extracellular SOD and iNOS in macrophages and smooth muscle cells in human and rabbit atherosclerotic lesions: colocalization with epitopes characteristic of oxidized LDL and peroxynitrite-modified proteins. Arterioscler Thromb Vasc Biol 18(2):157–167
54. Wilcox JN, Subramanian RR, Sundell CL, Tracey WR, Pollock JS, Harrison DG et al (1997) Expression of multiple isoforms of nitric oxide synthase in normal and atherosclerotic vessels. Arterioscler Thromb Vasc Biol 17(11):2479–2488
55. Wenger RH, Kvietikova I, Rolfs A, Camenisch G, Gassmann M (1998) Oxygen-regulated erythropoietin gene expression is dependent on a CpG methylation-free hypoxia-inducible factor-1 DNA-binding site. Eur J Biochem 253(3):771–777
56. Perini G, Diolaiti D, Porro A, Della Valle G (2005) In vivo transcriptional regulation of N-myc target genes is controlled by E-box methylation. Proc Natl Acad Sci USA 102(34):12117–12122
57. Bell AC, Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 405(6785):482–485
58. Han L, Lee DH, Szabo PE (2008) CTCF is the master organizer of domain-wide allele-specific chromatin at the H19/Igf2 imprinted region. Mol Cell Biol 28(3):1124–1135

59. Defossez PA, Kelly KF, Filion GJ, Perez-Torrado R, Magdinier F, Menoni H et al (2005) The human enhancer blocker CTC-binding factor interacts with the transcription factor kaiso. J Biol Chem 280(52):43017–43023
60. Sharma P, Kumar J, Garg G, Kumar A, Patowary A, Karthikeyan G et al (2008) Detection of altered global DNA methylation in coronary artery disease patients. DNA Cell Biol 27(7):357–365
61. Kim M, Long TI, Arakawa K, Wang R, Yu MC, Laird PW (2010) DNA methylation as a biomarker for cardiovascular disease risk. PLoS One 5(3):e9692
62. Castro R, Rivera I, Blom HJ, Jakobs C, Tavares de Almeida I (2006) Homocysteine metabolism, hyperhomocysteinaemia and vascular disease: an overview. J Inherit Metab Dis 29(1):3–20
63. Jiang Y, Zhang J, Xiong J, Cao J, Li G, Wang S (2007) Ligands of peroxisome proliferator-activated receptor inhibit homocysteine-induced DNA methylation of inducible nitric oxide synthase gene. Acta Biochim Biophys Sin (Shanghai) 39(5):366–376
64. Chang PY, Lu SC, Lee CM, Chen YJ, Dugan TA, Huang WH et al (2008) Homocysteine inhibits arterial endothelial cell growth through transcriptional downregulation of fibroblast growth factor-2 involving G protein and DNA methylation. Circ Res 102(8):933–941
65. Ying AK, Hassanain HH, Roos CM, Smiraglia DJ, Issa JJ, Michler RE et al (2000) Methylation of the estrogen receptor-alpha gene promoter is selectively increased in proliferating human aortic smooth muscle cells. Cardiovasc Res 46(1):172–179
66. Kim J, Kim JY, Song KS, Lee YH, Seo JS, Jelinek J et al (2007) Epigenetic changes in estrogen receptor beta gene in atherosclerotic cardiovascular tissues and in-vitro vascular senescence. Biochim Biophys Acta 1772(1):72–80
67. Waki T, Tamura G, Sato M, Motoyama T (2003) Age-related methylation of tumor suppressor and tumor-related genes: an analysis of autopsy samples. Oncogene 22(26):4128–4133
68. Thompson RF, Atzmon G, Gheorghe C, Liang HQ, Lowes C, Greally JM et al (2010) Tissue-specific dysregulation of DNA methylation in aging. Aging Cell 9(4):506–518
69. Stenvinkel P, Karimi M, Johansson S, Axelsson J, Suliman M, Lindholm B et al (2007) Impact of inflammation on epigenetic DNA methylation – a novel risk factor for cardiovascular disease? J Intern Med 261(5):488–499
70. Wilson VL, Jones PA (1983) DNA methylation decreases in aging but not in immortal cells. Science 220(4601):1055–1057
71. Butcher LM, Beck S (2008) Future impact of integrated high-throughput methylome analyses on human health and disease. J Genet Genomics 35(7):391–401
72. Shimabukuro M, Sasaki T, Imamura A, Tsujita T, Fuke C, Umekage T et al (2007) Global hypomethylation of peripheral leukocyte DNA in male patients with schizophrenia: a potential link between epigenetics and schizophrenia. J Psychiatr Res 41(12):1042–1046
73. Lund G, Andersson L, Lauria M, Lindholm M, Fraga MF, Villar-Garea A et al (2004) DNA methylation polymorphisms precede any histological sign of atherosclerosis in mice lacking apolipoprotein E. J Biol Chem 279(28):29147–29154
74. Voo KS, Carlone DL, Jacobsen BM, Flodin A, Skalnik DG (2000) Cloning of a mammalian transcriptional activator that binds unmethylated CpG motifs and shares a CXXC domain with DNA methyltransferase, human trithorax, and methyl-CpG binding domain protein 1. Mol Cell Biol 20(6):2108–2121
75. Ansari KI, Mishra BP, Mandal SS (2008) Human CpG binding protein interacts with MLL1, MLL2 and hSet1 and regulates hox gene expression. Biochim Biophys Acta 1779(1):66–73
76. Roman-Gomez J, Jimenez-Velasco A, Agirre X, Cervantes F, Sanchez J, Garate L et al (2005) Promoter hypomethylation of the LINE-1 retrotransposable elements activates sense/antisense transcription and marks the progression of chronic myeloid leukemia. Oncogene 24(48):7213–7223
77. Yla-Herttuala S, Palinski W, Butler SW, Picard S, Steinberg D, Witztum JL (1994) Rabbit and human atherosclerotic lesions contain IgG that recognizes epitopes of oxidized LDL. Arterioscler Thromb 14(1):32–40

78. Richardson BC (2002) Role of DNA methylation in the regulation of cell function: autoimmunity, aging and cancer. J Nutr 132(8 Suppl):2401S–2405S
79. Miao F, Gonzalo IG, Lanting L, Natarajan R (2004) In vivo chromatin remodeling events leading to inflammatory gene transcription under diabetic conditions. J Biol Chem 279(17):18091–18097
80. Vogel NL, Boeke M, Ashburner BP (2006) Spermidine/Spermine N1-acetyltransferase 2 (SSAT2) functions as a coactivator for NF-kappaB and cooperates with CBP and P/CAF to enhance NF-kappaB-dependent transcription. Biochim Biophys Acta 1759(10):470–477
81. Ashburner BP, Westerheide SD, Baldwin AS Jr (2001) The p65 (RelA) subunit of NF-kappaB interacts with the histone deacetylase (HDAC) corepressors HDAC1 and HDAC2 to negatively regulate gene expression. Mol Cell Biol 21(20):7065–7077
82. Chen LF, Greene WC (2003) Regulation of distinct biological activities of the NF-kappaB transcription factor complex by acetylation. J Mol Med 81(9):549–557
83. Liu Y, Denlinger CE, Rundall BK, Smith PW, Jones DR (2006) Suberoylanilide hydroxamic acid induces akt-mediated phosphorylation of p300, which promotes acetylation and transcriptional activation of RelA/p65. J Biol Chem 281(42):31359–31368
84. Choi JH, Nam KH, Kim J, Baek MW, Park JE, Park HY et al (2005) Trichostatin A exacerbates atherosclerosis in low density lipoprotein receptor-deficient mice. Arterioscler Thromb Vasc Biol 25(11):2404–2409
85. Reddy MA, Villeneuve LM, Wang M, Lanting L, Natarajan R (2008) Role of the lysine-specific demethylase 1 in the proinflammatory phenotype of vascular smooth muscle cells of diabetic mice. Circ Res 103(6):615–623
86. Lan F, Nottke AC, Shi Y (2008) Mechanisms involved in the regulation of histone lysine demethylases. Curr Opin Cell Biol 20(3):316–325
87. Janowski BA, Younger ST, Hardy DB, Ram R, Huffman KE, Corey DR (2007) Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. Nat Chem Biol 3(3):166–173
88. Li LC, Okino ST, Zhao H, Pookot D, Place RF, Urakami S et al (2006) Small dsRNAs induce transcriptional activation in human cells. Proc Natl Acad Sci USA 103(46):17337–17342
89. Morris KV, Chan SW, Jacobsen SE, Looney DJ (2004) Small interfering RNA-induced transcriptional gene silencing in human cells. Science 305(5688):1289–1292
90. Turunen MP, Lehtola T, Heinonen SE, Assefa GS, Korpisalo P, Girnary R et al (2009) Efficient regulation of VEGF expression by promoter-targeted lentiviral shRNAs based on epigenetic mechanism: a novel example of epigenetherapy. Circ Res 105(6):604–609
91. Weinberg MS, Villeneuve LM, Ehsani A, Amarzguioui M, Aagaard L, Chen ZX et al (2006) The antisense strand of small interfering RNAs directs histone methylation and transcriptional gene silencing in human cells. RNA 12(2):256–262
92. Mette MF, Aufsatz W, van der Winden J, Matzke MA, Matzke AJ (2000) Transcriptional silencing and promoter methylation triggered by double-stranded RNA. EMBO J 19(19):5194–5201
93. Janowski BA, Huffman KE, Schwartz JC, Ram R, Nordsell R, Shames DS et al (2006) Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. Nat Struct Mol Biol 13(9):787–792
94. Schramke V, Sheedy DM, Denli AM, Bonila C, Ekwall K, Hannon GJ et al (2005) RNA-interference-directed chromatin modification coupled to RNA polymerase II transcription. Nature 435(7046):1275–1279
95. Han J, Kim D, Morris KV (2007) Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. Proc Natl Acad Sci USA 104(30):12422–12427
96. Morris KV, Santoso S, Turner AM, Pastori C, Hawkins PG (2008) Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. PLoS Genet 4:e1000258
97. Schwartz JC, Younger ST, Nguyen NB, Hardy DB, Monia BP, Corey DR et al (2008) Antisense transcripts are targets for activating small RNAs. Nat Struct Mol Biol 15:842–848

98. Benditt EP, Benditt JM (1973) Evidence for a monoclonal origin of human atherosclerotic plaques. Proc Natl Acad Sci USA 70:1753–1756
99. Qiu P, Li L (2002) Histone acetylation and recruitment of serum responsive factor and CREB-binding protein onto SM22 promoter during SM22 gene expression. Circ Res 90:858–865
100. McDonald OG, Wamhoff BR, Hoofnagle MH, Owens GK (2006) Control of SRF binding to CArG box chromatin regulates smooth muscle gene expression in vivo. J Clin Invest 116:36–48
101. Cao D, Wang Z, Zhang CL, Oh J, Xing W, Li S et al (2005) Modulation of smooth muscle gene expression by association of histone acetyltransferases and deacetylases with myocardin. Mol Cell Biol 25:364–376
102. Davis CA, Haberland M, Arnold MA, Sutherland LB, McDonald OG, Richardson JA et al (2006) PRISM/PRDM6, a transcriptional repressor that promotes the proliferative gene program in smooth muscle cells. Mol Cell Biol 26:2626–2636
103. Wang S, Aurora AB, Johnson BA, Qi X, McAnally J, Hill JA et al (2008) The endothelial-specific microRNA miR-126 governs vascular integrity and angiogenesis. Dev Cell 15:261–271
104. Chan Y, Fish JE, D'Abreo C, Lin S, Robb GB, Teichert AM et al (2004) The cell-specific expression of endothelial nitric-oxide synthase: a role for DNA methylation. J Biol Chem 279:35087–35100
105. Devlin AM, Singh R, Wade RE, Innis SM, Bottiglieri T, Lentz SR (2007) Hypermethylation of Fads2 and altered hepatic fatty acid and phospholipid metabolism in mice with hyperhomocysteinemia. J Biol Chem 282:37082–37090
106. Hastings NE, Simmers MB, McDonald OG, Wamhoff BR, Blackman BR (2007) Atherosclerosis-prone hemodynamics differentially regulates endothelial and smooth muscle cell phenotypes and promotes pro-inflammatory priming. Am J Physiol Cell Physiol 293:C1824–C1833
107. Geisel J, Schorr H, Heine GH, Bodis M, Hubner U, Knapp JP et al (2007) Decreased p66Shc promoter methylation in patients with end-stage renal disease. Clin Chem Lab Med 45:1764–1770
108. Liu C, Xu D, Sjoberg J, Forsell P, Bjorkholm M, Claesson HE (2004) Transcriptional regulation of 15-lipoxygenase expression by promoter methylation. Exp Cell Res 29:61–67
109. Devlin AM, Bottiglieri T, Domann FE, Lentz SR (2005) Tissue-specific changes in H19 methylation and expression in mice with hyperhomocysteinemia. J Biol Chem 280:25506–25511
110. Sato N, Maehara N, Su GH, Goggins M (2003) Effects of 5-aza-2′-deoxycytidine on matrix metalloproteinase expression and pancreatic cancer cell invasiveness. J Natl Cancer Inst 95:327–330
111. Wild A, Ramaswamy A, Langer P, Celik I, Fendrich V, Chaloupka B et al (2003) Frequent methylation-associated silencing of the tissue inhibitor of metalloproteinase-3 gene in pancreatic endocrine tumors. J Clin Endocrinol Metab 88:1367–1373
112. White GP, Watt PM, Holt BJ, Holt PG (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 168:2820–2827
113. Lin XH, Guo C, Gu LJ, Deuel TF (1993) Site-specific methylation inhibits transcriptional activity of platelet-derived growth factor A-chain promoter. J Biol Chem 268:17334–17340
114. Tanaka Y, Fukudome K, Hayashi M, Takagi S, Yoshie O (1995) Induction of ICAM-1 and LFA-3 by Tax1 of human T-cell leukemia virus type 1 and mechanism of down-regulation of ICAM-1 or LFA-1 in adult-T-cell-leukemia cell lines. Int J Cancer 60:554–561
115. Schroeder M, Mass MJ (1997) CpG methylation inactivates the transcriptional activity of the promoter of the human p53 tumor suppressor gene. Biochem Biophys Res Commun 235:403–406

Index

A
Adaptive plasticity, 108
Addiction, 356
Age, 2, 16, 26, 70, 94, 111, 128, 140, 165, 185–193, 203, 232, 274, 293, 311, 324, 344, 390, 402, 429
Aging, 11, 13–18, 94, 166, 187–190, 192, 193, 198, 203–204, 209, 210, 212, 213, 230, 232, 233, 251, 313, 344, 346–348, 431
Agouti gene, 95, 110, 392
Agouti mouse, 110, 392
Air pollution, 198, 209–210, 310, 313–315, 334
Alcohol, 23, 29, 32, 200–203, 211, 231, 232, 236, 274, 294
Allergic disease, 314–315
Alzheimer's disease (AD), 24, 174, 202, 350, 360
Analysis of variance, 63, 66, 69, 166, 176
Angelman Syndrome (AS), 118, 119, 130–133, 139, 140, 142, 147–150, 211, 323, 324, 348, 359, 416
Arcsine transformation, 61, 63
Arsenic, 198, 208, 239, 313
Asbestos, 210
Assisted reproductive technology (ART), 2, 117–133, 171, 211
Association, 1, 2, 22, 26, 27, 30, 31, 42, 62–64, 66–73, 80, 93, 106, 110, 118, 122, 125, 126, 129, 130, 132, 133, 157, 168–170, 172, 175, 186, 190, 200–205, 208, 210, 211, 231–234, 236–243, 245, 246, 248–250, 274, 294, 296, 310–315, 331, 346, 350, 352–359, 379, 388, 389, 391, 402–404, 408, 415, 433
Asthma, 2, 24, 307–316
Atherosclerosis, 2, 192, 423–434
Autism, 2, 321–335
Autism spectrum disorder, 324, 325, 331, 348, 351
Autoimmune diabetes, 380–385, 390
Autoimmune disease, 292, 293, 295, 297, 378, 379, 390

B
Beckwith-Wiedemann syndrome (BWS), 108, 118, 125–130, 133, 143–147, 173, 211, 416
Benzene, 17, 207, 232, 245
Beta distribution, 61, 73
Beta-cells, 378, 381, 382, 384, 385, 407, 416, 417
Bioinformatics, 25, 31, 40, 49
Biospecimens, 22, 27, 29, 38–39, 168, 179, 296
Bipolar disorder, 176, 349, 359–360
Birth cohort, 26, 28, 29, 353
Birth complication, 352–354
Birthweight, 106, 111
Bisulfite converted DNA, 359, 404
Blood glucose, 383, 386
Body mass index (BMI), 175, 199–202, 391, 412–415
Bonferroni-adjustment, 63, 69, 71, 72
Brain, 12, 18, 22, 24, 89, 106, 107, 131, 149, 153, 188, 209, 247–248, 322–324, 326, 330–333, 346–348, 350, 354–356, 358–362, 364, 409–411

K.B. Michels (ed.), *Epigenetic Epidemiology*, DOI 10.1007/978-94-007-2495-2,
© Springer Science+Business Media B.V. 2012

C
Cancer, 1, 8, 22, 92, 106, 151, 179, 188, 198, 225–251, 271, 308, 358, 415, 425
Candidate gene studies, 59, 62, 63, 66
Cannabis, 355, 356
Carcinogenesis, 14, 18, 193, 204, 213, 227, 231, 232, 237, 240, 241, 243, 270, 275, 282
Cardiovascular disease, 1, 106, 201, 358, 389, 393, 425–427, 433, 434
Case-control study, 22, 26–28, 128, 176, 234, 236, 238, 240–242, 244, 312
CHARM. *See* Comprehensive high-throughput arrays for relative methylation (CHARM)
Childhood abuse, 354–355
Chromatin, 6, 23, 39, 59, 80, 108, 144, 189, 208, 229, 282, 290, 309, 322, 348, 381, 404, 425
Classification, 66, 67, 73, 143, 361, 408, 414
Classification and regression tree (CART), 67, 73
Classification system, 414
Cluster analysis (clustering), 62–68, 72, 73, 249, 356
Cohort, 26–29, 54, 55, 80, 111, 126–130, 132, 178, 201, 231, 234–237, 239, 242, 246, 250, 294, 310–314, 316, 347, 353, 389, 390, 393, 404, 411–415
Cohort study, 26, 27, 29, 128–130, 132, 201, 234, 236, 237, 250, 294, 311, 312, 314, 353, 404, 412
Comorbidity, 361, 393
Comprehensive high-throughput arrays for relative methylation (CHARM), 59, 412, 415
Confirmation study, 59, 70
Confounding, 22, 24, 29–32, 126, 142, 165, 179, 235, 294, 311, 353, 390
CpG. *See* Cytosine-phosphate-guanine (CpG)
CpG island (methylation), 13, 45, 65, 84, 108, 186, 204, 227, 271, 314, 325, 347, 380, 428
Cross validation, 68, 70
Cross-sectional study, 26, 27, 211, 239
Cytosine-phosphate-guanine (CpG), 7, 24, 40, 59, 84, 108, 123, 170, 186, 203, 227, 271, 309, 322, 347, 380, 404, 428

D
Data management, 40
Demethylation, 11–14, 25, 45, 81, 82, 85, 87, 89, 119–121, 187, 234, 295, 296, 313–315, 382, 383, 431
Depression, 24, 330, 352, 357, 359, 361
Developmental Origins of Health and Disease (DOHaD), 106–113
Developmental plasticity, 106, 107, 112, 113
Diabetes, 1, 20, 106, 139, 353, 377–394, 401–417, 432
Diet, 13, 30, 95, 96, 190, 199–206, 208, 231, 236, 293, 298, 310, 334, 385, 386, 389, 391–393, 403, 426, 427
Differentially methylated regions (DMRs), 85, 87, 88, 93, 109, 112, 119, 128, 130, 139, 155, 156, 166, 169, 170, 211, 390, 412
Discordance, 163, 164, 166–170, 172, 173, 176, 179, 180, 348–350, 390
Disease risk, 22–24, 26, 31, 32, 200, 201, 213, 380, 390, 391, 403, 405–407, 411, 412, 416
Dizygotic (DZ), 163–167, 169–172, 178, 179, 331, 349
DNA hydroxymethylation, 44–45
DNA methylation, 1, 8, 22, 38, 58, 80, 108, 119, 139, 164, 185–193, 199, 227, 270, 294, 309, 322, 347, 380, 404, 425
DOHaD. *See* Developmental Origins of Health and Disease (DOHaD)
Drugs of abuse, 352, 355–356

E
Early life, 106, 108, 113, 142, 166, 236, 296, 347, 379, 411, 412, 429
Effect modification, 30
Embryo development, 119, 125, 152
Embryonic stem cells, 10–12, 31, 44, 84, 89, 190, 192, 211, 281
Endocrine disruptors, 94, 96, 206–207
Environment, 8, 10, 13–18, 81, 92, 107, 109–111, 163–165, 170, 172, 180, 199, 203, 215, 231, 232, 235–237, 290, 294, 296, 299, 324, 332, 335, 349, 350, 354–355, 357, 363, 379–381, 389–391, 393, 414
Environmental exposures, 2, 12, 111, 164, 166, 190, 198–206, 211, 212, 215, 230, 236, 298, 309–314, 334, 346, 349, 350, 352–357, 361–364, 393, 405, 417, 427
Epialleles, 94, 95, 109, 110, 112, 392, 405, 406, 408, 412
Epidemiologic study, 2, 22, 25, 31, 72
Epigenetic genes, 233, 245
Epigenetic inheritance, 80, 94, 95, 111–112, 120, 121, 235–237

Epigenetic marker, 29, 174, 227–230, 245, 247, 410
Epigenetic methods, 2, 40–47
Epigenetic nomenclature, 408
Epigenetic regulation, 16, 107, 109, 110, 191, 203, 238, 297, 308–316, 328, 329, 346, 347, 380–385, 393, 425, 429
Epigenetic reprogramming, 80–89, 94, 108, 119, 199, 236, 391
Epigenetic signatures, 27, 393, 403, 407, 409
Epstein-Barr virus (EBV), 271, 272, 274, 275, 277, 293
Exploratory study, 59
Exposures, 2, 11, 24, 94, 106, 164, 190, 198, 228, 271, 293, 309, 334, 346, 379, 405, 427

F
Factor analysis (FA), 63–66
False discovery rate (FDR), 59, 63, 69, 71–73, 175, 177, 178
Family-based study, 26
Familywise error rate (FWER), 63, 69–73
Fisher's exact test, 62, 66, 171
Folate, 22, 28, 190, 199–203, 208, 214, 231, 232, 236, 237, 239, 250, 311, 333, 334, 353, 358, 385–387, 391, 427
Folic acid, 25, 110, 111, 208, 310–312, 332, 385, 391, 427
Fragile-X syndromes, 168, 325–326, 410

G
Gametes, 28, 80–87, 93, 94, 97, 111, 112, 119, 123, 125, 235, 324, 391
Gender, 204, 237, 402, 412, 417, 425
Gene environment interactions, 349, 350, 355, 363
Gene expression, 1, 6, 25, 39, 79, 121, 139, 163, 186, 198, 227, 291, 310, 325, 347, 380, 403, 425
Genetic, 2, 6, 24, 47, 59, 87, 107, 126, 138, 163, 191, 198, 231, 274, 292, 309, 322, 346, 380, 403, 425
Genetic imprinting, 138–139, 346–348
Genome-wide methylation, 32, 85, 278, 333, 388
Genomic DNA, 39–47, 172, 176, 200, 314, 387, 388, 390, 405, 431
Genomic imprinting, 2, 25, 50, 108, 112, 119, 139, 199, 211, 350, 351, 415
Genotype, 62, 107, 201, 202, 210, 232, 240, 244, 313, 325, 333, 358, 404, 408, 414
Germ cells, 80–82, 84, 85, 89, 94, 119
Gestational diabetes, 106, 391
Global DNA methylation, 26, 58, 123, 209, 387, 390, 428
Global hypomethylation, 208, 241, 242, 247, 296, 327, 333, 387, 431
Glucocorticoid receptor (GRs), 11, 316, 354

H
Heatmap, 68, 69, 73
Helicobacter pylori (H. pylori), 189, 270, 272–280, 283
Hepatitis virus, 271, 272, 274–275, 278
Heritability, 18, 163–165, 168, 170, 172, 292, 346, 403
Hierarchical clustering, 65–68, 73, 356
Histone, 1, 6, 23, 38, 58, 80, 108, 119, 165, 227, 280, 290, 309, 322, 347, 380, 404, 425
Histone code, 291, 297, 298, 322, 430, 432
Histone modifications, 1, 9, 10, 24, 25, 27, 30, 33, 38, 39, 45, 46, 53, 58, 81, 84, 86, 88–90, 93, 112, 119, 121, 124, 199, 209, 212, 227, 229, 238, 239, 243, 249, 251, 281, 282, 290, 297–300, 309, 322, 328, 335, 347, 349, 362, 380, 382–384, 388, 390, 408, 410, 417, 432
Homocysteine (Hcy), 199, 201, 202, 208, 210, 250, 353, 379, 385–388, 427, 428
Homocysteinemia, 427–428
Human disease, 2, 8, 22, 119, 204, 214, 323, 416
Human genetic imprinting, 138–139, 346, 348
Hypermethylation, 16, 23, 29, 94, 110, 126, 132, 144, 147, 148, 154, 155, 188, 189, 204–208, 212–214, 227, 228, 231, 232, 236, 237, 239–241, 243, 244, 247–249, 312, 322, 329, 382, 384, 392, 429–431
Hypertension, 106, 353, 403, 426
Hypomethylation, 15, 23, 44, 88, 121, 143, 188, 200, 227, 295, 327, 353, 381, 428

I
IGF2. *See* Insulin-like growth factor 2 gene (IGF2)
Imprinted genes, 17, 50, 80–82, 84, 85, 87, 91, 93, 94, 96, 108, 109, 111, 112, 119, 121–124, 138, 139, 144, 147, 148, 151, 153, 157, 187, 250, 295, 322–324, 329, 403, 411, 415–417

Imprinting, 2, 22, 50, 58, 85, 106, 118, 137–157, 166, 187, 199, 237, 295, 322, 346, 415, 427
Imprinting disorders, 118, 122, 125, 129, 130, 132, 133, 137–157, 211, 416
Infection, 17, 189, 244, 247, 270, 271, 273–280, 283, 293, 295, 299, 300, 311, 327, 330, 353, 391, 393, 410
Inflammation, 2, 84, 189, 190, 204, 212, 235, 274, 276–277, 283, 289–300, 310, 311, 316, 430–432
Innate immunity, 291
Insulin, 22, 107, 151, 237, 295, 312, 378, 379, 382, 384–387, 389, 392, 393, 404, 407–409, 411, 417
Insulin-like growth factor (IGF), 22, 237, 295, 312, 409
Insulin-like growth factor 2 (IGF2), 23, 26, 30, 50, 108, 110, 111, 126, 144, 147, 166, 167, 170, 174, 232, 237, 295, 312, 353, 390, 409, 411, 416–417, 427, 429, 430
Insulin-like growth factor 2 (IGF2) gene, 110, 237, 295, 312, 353
Interleukin (IL), 291, 314, 381
Intervention study, 26
Interindividual variation, 22, 25, 93, 108, 110, 112, 390
Intracytoplasmic sperm injection (ICSI), 118, 119, 121–123, 125, 126, 129–132
Intraindividual variation, 25, 32, 414
Intrauterine exposures, 111–112, 310
Intrauterine growth retardation (IUGR), 111
Intrauterine insemination (IUI), 129, 131
In utero, 94, 107, 110, 119, 163, 164, 166, 170, 199, 206, 207, 211, 333, 334, 389–392, 411, 412
In vitro culture, 119–124
In vitro fertilization (IVF), 121–126, 129, 130, 132, 153, 166, 171, 211
Islets of Langerhans, 379
IUGR. *See* Intrauterine growth retardation (IUGR)

K
K-means clustering, 65
Kernel machines, 67, 73
KLF14, 415
Kruskal-Wallis test, 61, 63, 66

L
Large offspring syndrome (LOS), 121, 211
Leave-one-out-cross-validation (LOOCV), 68
Leptin, 412, 413
Limma, 69, 73
Lineage specification, 80, 86, 88–89
Linear discriminant analysis, 67, 169
Liver fluke, 271, 272, 276
Locus-by-locus analysis, 60, 62–63, 69
Longitudinal studies, 364
Loss of imprinting (LOI), 22, 23, 26, 30, 31, 33, 50, 58, 59, 62, 108, 119, 173, 211, 237, 295, 296, 415, 417

M
Mann-Whitney test, 63, 66
Maternal age, 128, 324, 350, 351
Maternal care, 354, 355
MeCP2. *See* Methyl CpG binding protein 2 (MeCP2)
Mendelian randomization, 364
Metabolic syndrome, 391, 403–404, 407, 409, 411, 414, 417
Metals, 198, 209, 313, 414
Methyl CpG binding protein 2 (MeCP2), 328–332, 347, 348, 358–359, 429
Methylation, 1, 8, 22, 38, 58, 80, 108, 119, 139, 164, 185–193, 199, 227, 270, 290, 309, 322, 347, 379, 404, 425
Methylation fingerprint, 271, 277–284
Methylenetetrahydrofolate reductase (MTHFR), 201, 232, 244, 333, 348, 358, 360, 361
Micro-RNA (miRNA), 25, 58, 59, 61, 68, 125, 203, 205, 207–210, 212, 229–230, 233, 234, 238, 240–242, 244, 245, 250, 313, 316, 382
Microarrays, 25, 39, 45, 46, 49, 50, 53, 58–60, 62, 63, 66, 68–73, 80, 94, 97, 111, 278, 327, 349, 359–361, 388, 403
Mixture model, 63–66, 70, 73
Moderated t-test, 69
Monozygotic (MZ), 94, 95, 126, 147, 163–180, 203, 292, 331, 348–351, 360, 380, 389–390
Mosaicism, 142, 149, 167, 169, 191–193, 326, 330, 410, 412, 413
MTHFR. *See* Methylenetetrahydrofolate reductase (MTHFR)
Multiple comparisons, 59, 60, 63, 69–72, 111
Muscle, 150, 169, 404–407, 411

N
Naïve Bayes procedure, 66
Nested case-control study, 22, 26–28

Neurodevelopment, 2, 321–335, 347, 348, 350, 358, 362
Next generation sequencing, 59, 80, 89, 176, 215, 250
Normal distribution (normality), 61, 67
Normalization (of microarrays), 59, 68
Nutrition, 8, 10, 14, 17, 25, 29, 80, 95, 96, 106–110, 112, 200, 230, 232, 296, 297, 299, 300, 311, 322, 334, 352, 353, 378, 391, 392, 411, 428

O

Obesity, 2, 106, 138, 140, 142, 149, 150, 323, 391, 401–417
Observational study, 23, 214
Oncogene, 14, 227, 230, 234, 241, 244, 246, 247, 277
One carbon metabolism (OCM), 28, 95, 190, 199–203, 208, 210, 230–232, 333, 334, 358, 385

P

p-value, 59, 66, 70–72, 165, 169, 171, 178
Pancreas, 151, 250, 379, 392, 393, 406, 408, 411, 416
Papilloma virus, 244, 271, 272, 275
Parent-of-origin, 22, 48, 50, 51, 62, 108, 138, 154, 322, 350–351, 416
Particulates, 209–210, 313, 315
Paternal age, 332, 350–351
Pearson chi-square test, 62
Pearson correlation, 63, 175
Peripheral blood lymphocytes (PBL), 235, 407, 410–415, 417, 428, 429, 431, 432
Permutation tests, 66
Phenotype, 16, 31, 62–67, 70–72, 79, 85, 86, 94, 95, 107, 109, 124, 140, 143, 147, 150–153, 155, 157, 163, 164, 168, 172, 175, 179, 199, 205, 212, 213, 230, 233, 236, 248, 275, 309, 310, 312, 316, 324, 325, 329, 331, 346, 348, 349, 351, 354, 355, 357, 361, 362, 364, 380, 381, 403, 407, 408, 410, 411, 414–416, 425, 432, 433
Phenotypic discordance, 163, 167, 172, 179, 180, 349
Plasticity, 16, 89–94, 106–108, 112, 113, 149, 192, 322, 323, 347, 348
Post-traumatic stress disorder (PTSD), 351, 359
PowerAtlas, 72
Prader-Willi syndrome (PWS), 119, 130, 139, 140, 147–151, 154, 323, 324, 348, 416
Pregnancy, 94, 106, 110, 111, 205, 206, 311, 312, 334, 352–354, 391, 392, 411
Preimplantation embryos, 80, 84, 85, 87–89, 94, 211, 347
Principal components analysis (PCA), 62–66, 170
Prognosis, 23, 153, 157, 227–231, 236, 239, 240, 245, 247–249, 417
Prognostic (risk, indicator) factor, 244, 247, 248
Promoter (CpG island), 13, 14, 25, 32, 84, 92, 108, 186–189, 192, 212, 213, 236, 247, 248, 250, 271, 274, 314, 325, 384, 388, 429
Prospective study, 132, 232, 244, 250
Pseudohypoparathyroidism (PHP), 139, 141, 155–156
Pyrosequencing, 14, 31, 43, 46, 51, 59, 143, 169, 173, 175, 176, 247

Q

q-value, 71–73
Quantitative PCR (qPCR), 49, 171

R

Radiation, 207, 210, 245
Random forest, 67, 73
Randomized controlled trial, 26, 28, 32
Rearing environment, 354–355
Retrospective study, 26–28, 125, 133, 242, 311, 312
Rett syndrome (RTT), 329–332, 358
Rheumatoid arthritis (RA), 2, 172, 175, 289–300
Risk, 1, 16, 22, 106, 118, 139, 168, 193, 200, 227, 273, 290, 310, 332, 349, 379, 403, 425
RNA, 7–9, 38, 39, 43, 47–52, 54, 58, 59, 61, 85, 97, 139, 144, 151, 177, 178, 186, 209, 214, 227, 229, 230, 280, 282, 299, 310, 324–326, 328, 330, 333, 349, 381, 409, 432, 433

S

S-adenosylhomocysteine (SAH), 202, 230, 333, 334, 385–387, 428
S-adenosylmethionine (SAM), 10, 44, 72, 73, 190, 199, 201, 202, 208, 214, 230, 312, 333, 385, 386, 427, 428

Sample collection, 23, 39, 54, 250, 251, 414
Schistosoma, 271, 272, 276
Schizophrenia, 106, 173, 322, 348, 350–354, 356, 358–361, 431
Screening, 46, 143, 236, 238, 242–244, 388
Semi-supervised analysis, 70
Sensitivity, 23, 95, 107, 207, 228, 233, 236, 238, 278, 279, 354, 357, 404, 417
Sex differences, 298–299, 351–352
Significance Analysis of Microarrays (SAM), 72, 73
Silver Russell syndrome (SRS), 108, 139, 140, 143, 144, 146–147, 174
Single nucleotide polymorphism (SNP), 16, 33, 43, 50, 51, 54, 62, 210, 233, 234, 238, 240–242, 244, 247, 349, 385, 388, 390, 403–405, 407, 414–417, 425
Small RNAs, 48, 51, 84, 432–433
SNP. *See* Single nucleotide polymorphism (SNP)
SNP analysis, 233
Social defeat, 356–357
Social minority, 356–357
Spearman correlation, 63, 170, 175
Specificity, 22, 23, 29, 42, 46, 52, 54, 96, 187–189, 206, 228, 238, 243, 250, 278, 282, 308, 361, 381
Statistical Significance, 425
Stroke, 106, 426, 427
Suicide, 355, 359
Supervised analysis, 66–68, 70
Support vector machine (SVM), 67, 73
Synergism, 357

T

T helper differentiation, 309, 382
T regulatory cells, 309, 383
t-test, 61, 63, 66, 69
Temple syndrome (TS), 139, 141, 153–155
Tissue
Tobacco, 15, 111, 198, 204–206, 210, 211, 213, 232, 233, 238, 239, 313
Transgeneration inheritance, 26, 28, 81, 82, 94, 96, 97, 111–112, 206, 235–237, 326
Transient Neonatal Diabetes (TND), 139, 140, 142, 143, 151–153, 416
Transposable elements, 85, 109, 295, 296, 392, 431
Tumor
Tumor suppressor, 13, 15, 29, 132, 151, 204, 206, 210, 227, 228, 230, 232, 234, 237, 239, 241, 243, 244, 247, 249, 271, 274–278, 430, 431
Twins, 95, 126, 147, 161–180, 203, 292, 331, 349, 351, 352, 354, 360, 380, 389, 390, 404, 405
Type 1 diabetes (T1D), 377–394
Type I error, 59, 63, 70, 73

U

Uniparental disomy (UPD), 131, 139, 141, 142, 144, 146–151, 153–156, 167, 323, 324
Univariate analysis, 62–63, 70, 73
Unsupervised analysis, 60, 62–67, 72

V

Variably methylated region (VMR), 414, 415
Variation, 1, 17, 22, 25, 27, 32, 52, 54, 59, 62, 85, 93–96, 107–109, 112, 163–165, 167–171, 179, 180, 185–193, 203, 233, 235, 242, 244, 293, 309, 311, 315, 326, 331, 349, 352, 355, 358, 363, 383, 390, 402–404, 407–409, 412–415, 417, 428
Vitamin B12, 110, 190, 250, 385
Vitamin B6, 200, 236, 237, 250, 385
VMR.*See* Variably methylated region (VMR)

W

Wang syndrome (WS), 139, 141, 153, 155

X

X-inactivation, 93, 94, 168, 187, 299, 322, 326, 328–330, 347, 351, 352

Z

ZFP57.*See* Zinc-finger protein 57 (ZFP57)
Zinc-finger protein 57 (ZFP57), 88, 141, 143, 152, 153