Anna K. Naumova
Celia M.T. Greenwood *Editors*

Epigenetics and Complex Traits

Springer

Epigenetics and Complex Traits

Anna K. Naumova • Celia M.T. Greenwood
Editors

Epigenetics and Complex Traits

Springer

Editors
Anna K. Naumova
Departments of Obstetrics
and Gynecology and Human Genetics
McGill University
Montreal, QC, Canada

Celia M.T. Greenwood
Lady Davis Institute for Medical Research
Centre for Clinical Epidemiology
Jewish General Hospital
Montreal, QC, Canada

ISBN 978-1-4614-8077-8 ISBN 978-1-4614-8078-5 (eBook)
DOI 10.1007/978-1-4614-8078-5
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013946543

Printed on acid-free paper

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

Preface

Influence of Epigenetic Phenomena on Gene Expression and Inheritance of Phenotypes

One of the many definitions of an epigenetic mark is a heritable feature that does not change the DNA sequence but determines when, where, and to what extent a gene will be expressed. Hence, epigenetics is a science that studies DNA packaging and regulation of its expression. Although often introduced as a new science, epigenetics dates back to the discovery of the roles of chromatin and DNA methylation in controlling gene expression in the 60s and 70s of the last century. Despite the intimate relationship between DNA and epigenetic factors, mainstream studies of genetic traits in humans and animal models have largely ignored the existence of epigenetic factors during the past decades, while the epigenetics community, although part of both the genetics and developmental biology fields, was digging deeper and deeper into the molecular mechanisms of epigenetic phenomena but seldom tackling problems of complex genetic traits in mammals. One of the reasons for the dichotomy is the very complexity of complex traits where small effects from multiple loci define the phenotype, whereas traditional molecular biology research required focusing on one selected target at a time. Another reason was the lack of methodologies capable of analyzing large amounts of epigenetic information in large cohorts of patients and controls. Nevertheless, during the last two decades, in-depth analysis of inheritance patterns combined with molecular approaches in a number of animal models, such as agouti viable yellow mice and callipyge sheep, has provided remarkable examples of how the interplay between genetic and epigenetic factors can generate complex traits.

Rapid technological improvements are now making it possible to measure epigenetic signals at many genomic locations in an unprecedented way and conduct prior-hypothesis-free epigenetic studies. Global initiatives such as the International Human Epigenome Consortium are underway to obtain high-resolution maps of histone modifications, DNA methylation, and transcription start sites and to compare epigenome signals and the resulting transcriptional regulation in a wide variety

of tissues and different cell types. However, even hypothesis-free data analyses require knowledge of epigenetic paradigms to make informed decisions when interpreting these massive data sets.

In this book, we have focused on the relationship between epigenetics and complex traits, since this field can be daunting for those wishing to do research. The biology is complex, and the ramifications of epigenetic regulation are widespread. Epigenetic states may contribute to the penetrance of genetic polymorphisms or mutations and thereby modify inheritance patterns. This may result in apparently non-Mendelian inheritance of genetic traits. Epigenetic changes in an individual may affect several different generations, depending on when these changes occur and in which cells. Genetic factors will influence epigenetic factors, and possibly their transmission. Effects may vary depending on sex, and also on the sex of an implicated parent. Concepts that applied in genetics, such as heritability, or the proportion of variance explained by genetics, can now be expanded to explicitly consider the epigenetic contributions. Furthermore, of course, different loci may demonstrate different associations with all these factors. Design of experiments and analysis of experimental data must reflect this complexity and be carefully approached.

Therefore, this book presents 14 detailed and distinct views on the interplay between complex traits and epigenetics. The chapters are grouped into three sections: (1) Fundamental aspects of the biology in epigenetics, with focus on the period in mammalian development that is pivotal for genetic transmission, i.e., gametogenesis and early embryonic development, insight into how the epigenetic marks are established, maintained, and transmitted and their influence on gene expression; (2) The known impact of epigenetic factors on several different complex traits and diseases of interest for human genetics; and (3) Approaches to experimental design and statistical analysis in this context.

Our hope is that the two communities of basic researchers and analysts will find mutual enrichment through this combination of material. An overview of available analytic methods and their underlying assumptions could inform experimental design choices. Similarly, improved understanding of the biology could lead to better choices for analysis, and an appreciation for the many factors that may need to be considered. Ultimately, this marriage of topics could lead to improved study designs, rich and complete analytic frameworks, new approaches to analysis, and guidelines for interpretation.

Of course, this book includes only a small overview of the available knowledge and approaches, yet we anticipate that this will be a helpful first reference for researchers entering the field, and will stimulate future developments. We thank Springer for making this endeavor possible.

Montreal, QC, Canada Anna K. Naumova
Montreal, QC, Canada Celia M.T. Greenwood

Acknowledgements

We sincerely thank Jason Apostolopoulos for all his assistance in putting this book together. Celia Greenwood is grateful to the Pharmaprix Weekend to End Women's Cancers (2010) for their support.

Contents

Contributors

Philippe Arnaud Genetic Reproduction and Development, Clermont Université-INSERM, Clermont-Ferrand, France

Christopher G. Bell Medical Genomics, UCL Cancer Institute, University College London, London, UK

Jordana T. Bell Department of Twin Research and Genetic Epidemiology, King's College, London, UK

Carolyn J. Brown Department of Medical Genetics, Molecular Epigenetics Group, University of British Columbia, Vancouver, BC, Canada

Carole Charlier Unit of Animal Genomics, University of Liège, Liège, Belgium

Huijun Cheng Unit of Animal Genomics, University of Liège, Liège, Belgium

Freddy B. Christiansen Department of Bioscience and Bioinformatics Research Center, University of Aarhus, Aarhus, Denmark

Noelle Cockett Department of Animal, Dairy, and Veterinary Sciences, Utah State University, Logan, UT, USA

Allison M. Cotton Department of Medical Genetics, Molecular Epigenetics Group, University of British Columbia, Vancouver, BC, Canada

Michelle M. Creek Department of Biostatistics, UCLA Fielding School of Public Health, Los Angeles, CA, USA

Marcus W. Feldman Department of Biology, Stanford University, Stanford, CA, USA

Robert E. Furrow Department of Biology, Stanford University, Stanford, CA, USA

Michel Georges Unit of Animal Genomics, University of Liège, Liège, Belgium

Alexander A. Gimelbrant Department of Genetics, Harvard Medical School, Boston, MA, USA

Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA

Tracy Hadfield-Shay Department of Animal, Dairy, and Veterinary Sciences, Utah State University, Logan, UT, USA

Lam Opal Huang Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, QC, Canada

Claire Infante-Rivard Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, QC, Canada

Mohammad Saharul Islam Medical Microbiology and Immunology, Genome Center, MIND Institute, University of California, Davis, CA, USA

Aurélie Labbe Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, QC, Canada

Douglas Mental Health University Institute, Montréal, QC, Canada

Benoit Labonté McGill Group for Suicide Studies, Department of Psychiatry, Douglas Mental Health Institute, Montreal, QC, Canada

Janine M. LaSalle Medical Microbiology and Immunology, Genome Center, MIND Institute, University of California, Davis, Davis, CA, USA

Stéphanie Maupetit-Méhouas Genetic Reproduction and Development, Clermont Université-INSERM, Clermont-Ferrand, France

Anna K. Naumova Departments of Obstetrics and Gynecology and Human Genetics, McGill University, Montreal, QC, Canada

The Research Institute of the McGill University Health Centre, McGill University, Montreal, QC, Canada

David Nury Genetic Reproduction and Development, Clermont Université-INSERM, Clermont-Ferrand, France

Maria S. Peñaherrera Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada

Child & Family Research Institute, Vancouver, BC, Canada

Samantha B. Peeters Department of Medical Genetics, Molecular Epigenetics Group, University of British Columbia, Vancouver, BC, Canada

Wendy P. Robinson Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada

Child & Family Research Institute, Vancouver, BC, Canada

Ionel Sandovici Metabolic Research Laboratories, Department of Obstetrics and Gynaecology, University of Cambridge, Cambridge, UK

Centre for Trophoblast Research, University of Cambridge, Cambridge, UK

Virginia Savova Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA

Department of Genetics, Harvard Medical School, Boston, MA, USA

Janet S. Sinsheimer Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA

Haruko Takeda Unit of Animal Genomics, University of Liège, Liège, Belgium

Amanda Ewart Toland Department of Internal Medicine, Division of Human Genetics, Columbus, OH, USA

Gustavo Turecki McGill Group for Suicide Studies, Department of Psychiatry, Douglas Mental Health Institute, Montreal, QC, Canada

Xu Xuewen Unit of Animal Genomics, University of Liège, Liège, Belgium

Abbreviations

3C	Chromosome conformation capture
5C	Chromosome conformation capture carbon copy
5caC	5-Carboxylcytosine
5fC	5-Formylcytosine
5hmC	5-Hydroxymethylcytosine
5hmU	5 Hydroxymethyluracil
5mC	5-Methylcytosine
aCGH	Array comparative genomic hybridization
ACTH	Adreno corticotropic hormone
ADHD	Attention deficit hyperactivity disorder
AdoMet	S-adenosyl-l-methionine
AEBP2	AE binding protein 2
AGRE	Autism Genetic Resource Exchange
AHEAD	Alliance for the Human Epigenome and Disease
AID or AICDA	Activation-induced cytidine deaminase
AMER1	APC membrane recruitment protein 1
AML	Acute myeloid leukemia
APOBEC1	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 1
AR	Androgen receptor
ART	Assisted reproductive technologies
AS	Angelman syndrome
ASD	Autism spectrum disorder
ASE	Allele-Specific Expression
ASHM	Allele-Specific Histone Modifications
ASM	Allele-Specific Methylation
AVP	Vasopressin
$Axin^{Fu}$	Axin fused allele
BDNF	Brain-Derived Neurotrophic Factor
BEGAIN	Brain-enriched guanylate kinase-associated protein

BER	Base excision repair
BiS-seq	Bisulfite 2nd generation sequencing
BLIMP1	B-lymphocyte induced maturation protein 1
BMI	Body Mass Index
BMI	Body mass composition
BMP	Bone morphogenetic protein
CBX5	Chromobox homolog 5
CCNE1	Cycline E1
CDC25A	Cell division cycle 25A
CDX1	Caudal type homeobox 1
CDX2	Caudal type homeobox 2
CGI	CpG Island
CHD1	Chromodomain helicase DNA binding protein 1
ChIA-PET	Chromatin interaction analysis with paired-end tag sequencing
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation-sequencing
CLPG	Callipyge locus
CMCM	Case-mother, Control-mother
CNV	Copy number variation
CPA	Child physical abuse
CpG	Cytosine-phosphate-guanine
CPT	Case-parent trio
CRF	Corticotropin-releasing factor
CRH	Corticotropin-releasing hormone
CSA	Child sexual abuse
CTCF	CCCTC-binding factor [zinc finger protein]
D3	DIO3, type 3 deiodinase
DD	Developmental delay
DEX	Dexamethasone
DGS	DiGeorge syndrome
DHS	DNase I Hypersensitivity Sites
DKK1	Dickkopf 1 homolog (*Xenopus laevis*)
DLK1	Delta-like homologue 1; also known as preadipocyte factor-1 (PREF1) or fetal antigen (FA1)
DMD	Duchenne muscular dystrophy
DMR	Differentially Methylated Region
DNMT	DNA methyltransferase
DNMT1	DNA methyltransferase 1
DNMT1o	Oocyte-specific form of the DNA cytosine methyltransferase 1
DNMT3A	DNA methyltransferase 3A
DNMT3L	DNA methyltransferase 3-like protein
Dscam	Down Syndrome Cell Adhesion Molecule

DSL	Delta-Serrate-LAG-2 domain
DSL	Disease susceptibility locus
DZ	Dizygotic
EED	Embryonic ectoderm development
EFNB1	Ephrin-B1
EHMT2	Euchromatic histone-lysine N-methyltransferase 2
ELF5	E74-like factor 5 [ets domain transcription factor]
EMFG	Extended Maternal-Fetal Genotype
EMSA	Electrophoretic mobility shift assay (EMSA)
ENCODE project	Encyclopedia of DNA Elements project
EOMES	Eomesodermin
eQTL	Expression quantitative trait loci
ERV1	Class I endogenous retrovirus 1
ES	Embryonic stem cells
EWAS	Epigenome Wide Association Study
EZH2	Enhancer of zeste homolog 2 (Drosophila)
F XIII	Factor XIII
FGF	Fibroblast growth factor
FISH	Fluorescence in situ hybridization
FMR1	Fragile X mental retardation 1
FOXA	Forkhead box A
G6PD	Glucose-6-phosphate dehydrogenase
GAD1	Glutamate decarboxylase 1 [brain, 67 kDa]
GADD45a	Growth-arrest and DNA-damage-inducible protein 45 α
GATA	GATA binding protein
GATA4	GATA binding protein 4
GATA6	GATA binding protein 6
GC	Germ cell
gDMR	Germline differentially methylated region
GR	Glucocorticoid receptor
GSK3B	Glycogen synthase kinase 3 beta
GTL2	Gene trap locus 2
GWAS	Genome-wide association study
HDAC	Histone deacetylase
HDN	Hemolytic disease of the newborn
HIRA	Histone cell cycle regulation defective homolog A (*S. cerevisiae*)
HLA-DQB1	Major histocompatibility complex, class II, DQ beta 1
HLA-DRB1	Major histocompatibility complex, class II, DR beta 1
HNF4A	Hepatocyte nuclear factor 4, alpha
HOXD11	Homeobox D11
HOXD12	Homeobox D12

HPA	Hypothalamic–pituitary–adrenal axis
HSM	Haplotype-Specific Methylation
HUMARA	Human androgen receptor
IAP	Intracisternal A particle
IBD	Identically by descent
ICM	Inner cell mass
ICR	Imprinting control region
IG DMR	Intergenic differentially methylated region
iPS	Induced pluripotent stem cells
iQTL	Imprinted QTL
JARID2	Jumonji, AT rich interactive domain 2—a member of the Jumonji family of lysine demethylases)
JMJD3/KDM6B	Lysine [K]-specific demethylase 6B
KDM1B	Lysine (K)-specific demethylase 1B
KDM5C	Lysine (K)-specific demethylase 5C
KLF2	Kruppel-like factor 2
KRAB	Kruppel-associated box
L1	LINE element, long interspersed repetitive element 1
LCR	Low copy repeat
LD	Linkage disequilibrium
LG	Licking and grooming
LIF/STAT3	Leukemia inhibitory factor/signal transducer and activator of transcription 3
lincRNAs	Large intergenic noncoding RNAs
LINE1	Long interspersed repeat element 1
LIS1	Lissencephaly-1 gene
LMR	Low Methylation Region
lncRNA	Long non-coding RNA
lncRNA	Long noncoding RNA
LOI	Loss of imprinting
LRT	Long-range transgenerational
LRT-M	Long range transgenerational effects on the maternal side
LRT-P	Long range transgenerational effects on the paternal side
LSH	Lymphoid-specific helicase
LTR	Long terminal repeat
MAE	Monoallelic expression
MAOA	Monoamine oxidase A
Mat-gDMR	Maternal-germline diffentially methylated region
MBD3	Methyl-CpG binding domain protein 3
MBD4	Methyl CpG binding domain protein 4
MBD-seq	Methylated DNA binding domain sequencing
MDLS	Miller-Dieker syndrome
MDR	Methylation Determining Region

MECAP-seq	Methylated DNA capture by affinity purification sequencing
MECP2	Methyl CpG binding protein 2
Me-DIP	Methylated DNA immunoprecipitation
MeDIP-seq	Methylation Dependent Immunoprecipitation 2nd generation sequencing
MEK	MAP kinase/ERK kinase]
me-QTLs	Methylation quantitative trait loci
methOR	Methylation odds ratios
MFG	Maternal fetal genotype
MGMT	O^6-methylguanine-DNA methyltransferase
MHC	Major histocompatibility region
Mirg	Micro-RNA containing gene (cluster of ~50 miRNAs expressed from the maternal allele)
miRNA	Micro RNA
MLL/Trithorax complex	Myeloid/lymphoid or mixed-lineage leukemia [trithorax homolog, Drosophila]
MLL3/KMT2C	Lysine (K)-specific methyltransferase 2C
MLL4/KMT2D	Lysine (K)-specific methyltransferase 2D
MNase	Micrococcal nuclease
M-PCR	Methylation-specific PCR
MRFs	Myogenic regulatory bHLH-containing factors
MSUC	Meiotic silencing of unsynapsed chromatin
MT	Mouse transcript
MTHFR	Methylenetetrahydrofolate reductase (NAD(P)H)
MVH	Mouse vasa homolog
MVP	Methylation Variable Position
MZ	Monozygotic twins
NANOG	Nanog homeobox
NAP-1	Nucleosome assembly protein-1
ncRNA	Noncoding RNA
ND	Neurodevelopmental disorders
NER	Nucleotide excision repair
Nes	Nestin
NF1	Neurofibromatosis type 1
NGF	Nerve growth factor
NIMA	Non-inherited maternal antigen
NIPA	Non-inherited paternal antigen
NIPBL	Nipped-B homolog [Drosophila]
NP95 (or UHRF1)	Nuclear protein of 95 kDa (or ubiquitin-like with PHD and ring finger domains 1)
NT3/4	Neurotrophin 3 and 4
NuRD	Nucleosome-remodeling

OATL1	Ornithine aminotransferase-like 1
OCT4	Octamer-binding transcription factor 4, also known as POU5F1 – POU domain, class 5, transcription factor 1
Om	Ovum mutant
PAI-1	Plasminogen activator inhibitor-1
PAR	Pseudoautosomal region
pasRNA	Promoter-associated small RNA
PAT	Parental Asymmetry Test
Pat-gDMR	Paternal-germline differentially methylated region
PcG	Polycomb group
PCL2	Polycomb-like 2 protein
PCR	Polymerase chain reaction
PCSK1N	Proprotein convertase subtilisin/kexin type 1 inhibitor
PFG	Paternal fetal genotype
PGC	Primordial germ cell
PGK1	Phosphoglycerate kinase 1
PGL/PCC	Paraganlioma/pheochromocytoma
PHF6	PHD finger protein 6
piRNA	PIWI-interacting RNA
POE	Parent of origin effects
POF	Premature ovarian failure
PO-LRT	Parent-of-Origin Likelihood Ratio Test
POMC	Pro opiomelanocortin
PoO	Parent of origin
PPB	Pleuropulmonary blastoma
PRC1	Repressive complex 1
PRC2	Polycomb repressive complex 2
PRDM	PR-domain-zinc-finger protein
PRMT5	Protein arginine N-methyltransferase 5
PTSD	Post traumatic stress disorder
PVN	Paraventricular nucleus
PWS	Prader-Willi syndrome
qTCAs	Transcriptional clonality assays
QTL	Quantitative trait locus
RASGRF1	RAS protein-specific guanine nucleotide-releasing factor 1
RFLP	Restriction fragment length polymorphism
RM	Recurrent miscarriage
RNA pol II	RNA polymerase II
RNAP II	RNA polymerase II
RNA-seq	RNA-sequencing
RNF2	Ring finger protein 2, also known as RING1B
RRBS	Reduced representation bisulfite sequencing
RRBS-seq	Reduced Representation BiSulfite 2nd generation sequencing

rRNA	Ribosomal RNA
RTL1	Retrotransposon-like 1
RTT	Rett syndrome
SALL4	Sal-like protein 4
SAM	S-adenosylmethionione
SAT1	Spermidine/spermine N1-acetyltransferase
Satb2	SATB homeobox 2
SETDB1	SET domain, bifurcated 1
SGA	Small-for-gestational-age
SINE	Short interspersed repeat element
siRNA	Small interfering RNA
SKI	v-ski sarcoma viral oncogene homolog (SKI)
Smarca5	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5
SMS	Smith-Magenis syndrome
SMS	Spermine synthase
SMUG1	Single-strand-selective monofunctional uracil DNA glycosylase 1
sno-RNA	Small-nucleolar RNA
SNP	Single nucleotide polymorphism
SOX17	SRY [sex determining region Y]-box 17
Sox2	SRY-box containing gene 2
SOX2	SRY [sex determining region Y]-box 2
SOX7	SRY [sex determining region Y]-box 7
SRY	Sex-determining region Y
Ssm 1	Strain-specific modifier of transgene methylation 1
STAG2	Stromal antigen 2
STR	Short tandem repeat
STS	*STS* (steroid sulfatase (microsomal), *isozyme S*)
SUZ12	Suppressor of zeste 12 homolog [Drosophila]
T2D	Type 2 diabetes
TAT	Transmission Asymmetry Test
TCF7L2	Transcription factor 7-like 2 [T-cell specific, HMG-box]
TDG	Thymine DNA glycosylase
TDRD	Tudor domain
TDT	Transmission disequilibrium test
TE	Trophectoderm
TEAD4	TEA domain family member 4
TET	Ten-eleven-translocation
TET3	Ten-eleven translocation
Tex19.1	Testis expressed gene 19.1
TFBS	Transcription Factor Binding Sites
TGF	Transforming growth factor

TIMP1	TIMP metallopeptidase inhibitor 1
TIP60/ KAT5–P400	Lysine acetyltransferase 5/ E1A binding protein p400
tiRNA	Transcription initiation RNA
TRD	Transmission ratio distortion
TRIM28	Tripartite motif containing 28
TrkB	Tropomyosin-Related Kinase B
TS	Trophoblast stem cells
TSG	Tumor suppressor gene
TSSs	Transcription start sites
UBE3A-AS	Antisense transcript of *UBE3A*
UBF	Upstream binding factor
UCE	Upstream control element
UPD	Uniparental disomy
UTX/KDM6A	Lysine [K]-specific demethylase 6A
VCFS	Velo-cardiofacial syndrome
WBS	Williams-Beuren syndrome
WGAS	Whole Genome sequencing Association Study
WGBS	Whole genome bisulfite sequencing
XCI	X-chromosome inactivation
XEN	Extraembryonic endoderm stem cells
XIST	X-inactive specific transcript
XIST/*Xist*	Inactive X specific transcripts
ZDHHC15	DHHC-type containing 15
ZFP57	Zinc finger protein 57
ZNF274	Zinc finger protein 274

Part I
Epigenetic Phenomena in the Germ Line and Early Embryonic Development and Their Effects on the Inheritance of Genetic Traits

Chapter 1
Epigenetic Reprogramming in the Mammalian Germline

Stéphanie Maupetit-Méhouas, David Nury, and Philippe Arnaud

Abstract Epigenetic modifications are crucial for maintaining and faithfully transmitting the identity of each cell type during cell division. During mammalian germ cell development, the acquisition of the ability to form a totipotent zygote is associated with extensive epigenetic reprogramming that affects all major developmental processes, including genomic imprinting, X-inactivation, retroelement silencing and gene expression. The existing epigenetic patterns are first erased during primordial germ cell development, followed by acquisition of a germline-specific epigenetic signature that can be eventually transmitted to and interpreted by the progeny. A better characterisation of the underlying mechanisms is relevant for both fundamental and clinical research dealing with epigenetic inheritance, epigenetic control of mammalian development and regenerative medicine. In this review we present and discuss recent advances on the nature, mechanisms and consequences of resetting the epigenetic pattern during primordial germ cell formation and (re) acquiring a new set of epigenetic marks at later stages of germline development.

1.1 Introduction

During somatic development of higher organisms, pluripotent cells progressively reduce their differentiation potential and become committed to a particular cell fate with specific gene expression and functional profiles. This tightly regulated process requires the concerted action of specific factors and is accompanied or caused by dynamic chromatin changes that influence gene expression patterns and phenotype.

S. Maupetit-Méhouas • D. Nury • P. Arnaud (✉)
GReD (Genetic Reproduction and Development), CNRS UMR6293-Clermont Université-INSERM U1103, 28 Place Henri Dunant BP38, 63001 Clermont-Ferrand Cedex, France
e-mail: stephanie.maupetit_mehouas@udamail.fr; david.nury@u-clermont1.fr; philippe.arnaud@udamail.fr

A.K. Naumova and C.M.T. Greenwood (eds.), *Epigenetics and Complex Traits*,
DOI 10.1007/978-1-4614-8078-5_1,

These changes occur at the level of DNA methylation, histone tail modifications, nucleosome remodelling and regulation of higher order chromatin structures. Most (but not all) of these modifications are heritable from one cell generation to the next and are thus referred as being epigenetic. Thus, each cell type in an organism is characterised by a specific and stable epigenetic signature (epigenome) that is transmitted to the daughter cells. Once specified, the epigenome of a cell type is relatively stable. However, in mammals, there are two key developmental stages in which epigenetic patterns are profoundly modified, with erasure of the existing epigenetic marks and acquisition of a new set. This so-called epigenetic reprogramming occurs first in early embryogenesis, following fertilisation, when the epigenetic information carried by the mature gametes is removed and replaced by an embryonic/somatic signature at the peri-implantation stage. This "embryonic" reprogramming is incomplete as some genomic regions, notably the *cis*-acting regulatory sequences of imprinted gene loci (imprinting control regions, ICRs), escape this process. A more thorough epigenetic reprogramming occurs during gametogenesis and it virtually impacts all epigenetic-based developmental processes: genomic imprinting, X-inactivation, retroelement silencing and gene expression. The understanding of the underlying mechanisms is relevant for both fundamental and clinical research. It will enable to better define the role of epigenetics in the control of mammalian development and also to elucidate the mechanism of in vitro-induced reprogramming/pluripotency.

This review focuses on the germline epigenetic reprogramming and discusses recent findings on the mechanisms involved in erasing the epigenetic pattern during primordial germ cell (PGC) formation and in (re)acquiring a new set of epigenetic marks at later stages of germline development.

1.2 Temporal and Spatial Dynamics of Mouse Germ Cell Development

Among all the cell lineages of a complex organism, only germ cells can give rise to a new individual, allowing the transmission of genetic and possibly epigenetic information to the next generation. Germ cell development initiates with the specification of PGCs, which following colonisation of the embryonic gonads will develop into oocytes or spermatozoa. In mammals, most of our knowledge on the temporal and spatial dynamics of this tightly regulated process comes from the mouse model (Fig. 1.1).

Unlike other non-mammalian species, such as *D. melanogaster* and zebrafish, mouse PGCs are not predetermined at fertilisation but are specified in the post-implantation embryo. At embryonic day 4.5 (E4.5), following blastocyst implantation, there is a rapid increase in the number of inner cell mass cells, leading to the formation of the epiblast (the source of all the body cell lineages). Germ cell fate

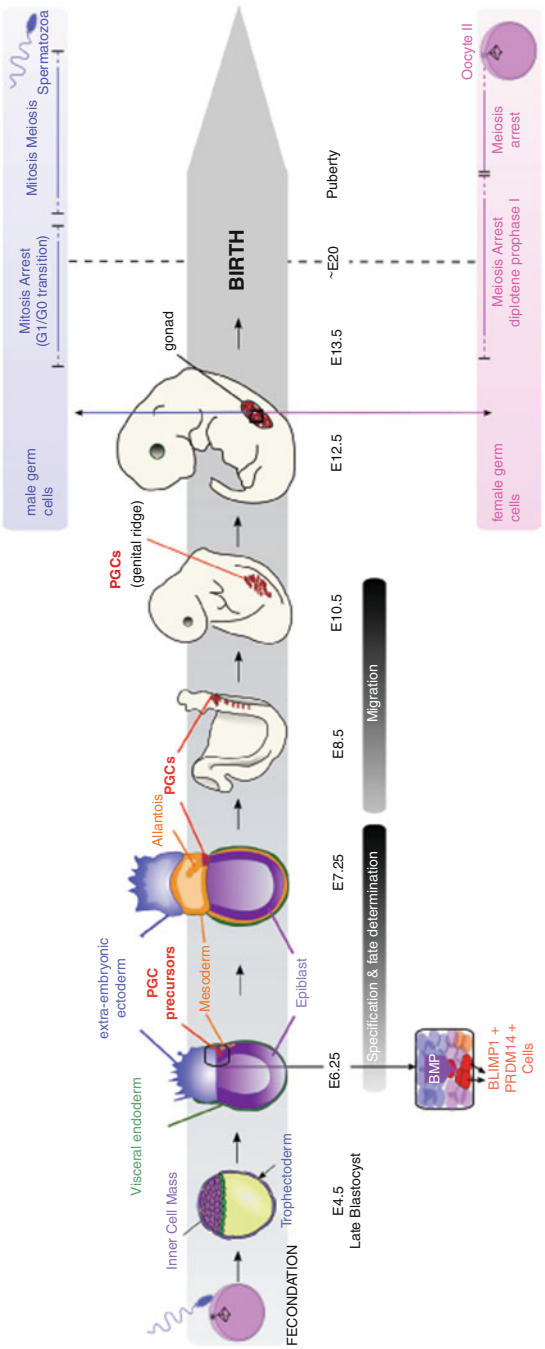

Fig. 1.1 Temporal and spatial dynamics of mouse germ cell development. Germ cell fate is induced at E6.25, and the lineage is determined around E7.25. From E7.5, primordial germ cells (PGCs) initiate migrating and then colonise the nascent genital ridge at around E10.5. At E12.5, PGCs start differentiating into male or female gametes. E4.5: Embryonic day 4.5. BMP: Bone morphogenetic protein. BLIMP1: B-lymphocyte maturation-induced protein 1. PRDM: PR domain zinc finger protein

is induced in the proximal epiblast in a dose-dependent manner by bone morphogenetic protein (BMP) signals from the extra-embryonic ectoderm at ~E6.25 (Lawson et al. 1999). This leads to the formation of a pool of PGC precursors of which only a limited number (about 6 cells), characterised by the expression of the zinc finger transcriptional regulators BLIMP1 (B-lymphocyte-induced maturation protein 1, also known as PR-domain-zinc-finger protein 1, PRDM1) and PRDM14 (PR-domain-zinc-finger protein 14), acquire a PGC fate. As a result, "fate-determined" PGCs emerge at ~E7.25 as a cluster of ~20–40 cells located at the base of the forming allantois (Ginsburg et al. 1990; Ohinata et al. 2005, 2009; Yamaji et al. 2008). From ~E7.5, PGCs migrate through the hindgut and mesentery and start colonising the nascent genital ridges (i.e., the future gonads) at ~E10.5. During this process, PGCs rapidly proliferate: from around 100 PGCs at E8.5 to ~200 at E9.5 and ~600 at E10.5. In the genital ridges, PGCs still proliferate up to E13.5 (~26 000 cells), when they stop dividing (Mochizuki and Matsui 2010; Kagiwada et al. 2012) (Fig. 1.2).

Following colonisation of the developing gonad, at E12.5, PGCs, now referred to as germ cells (GCs), start differentiating into male or female gametes. In the developing ovary, at E13.5, female GCs initiate meiosis I that will be blocked at the diplotene stage of prophase I at about the time of birth and until puberty. Following ovulation, the oocyte resumes meiosis I and halts in metaphase of meiosis II that will be completed after fertilisation (Smallwood and Kelsey 2012).

Conversely, male GCs do not initiate meiosis in the embryo and stop dividing from E13.5 (Western et al. 2008). At sexual maturity, male GCs will differentiate into spermatogonial stem cells and resume mitotic proliferation to form spermatocytes that will give rise, following meiosis, to haploid spermatids that will develop into spermatozoa.

1.3 Primordial Germ Cell Development and Reprogramming

After implantation, epiblast cells mature and prepare for gastrulation and formation of all the body cell lineages. This process is associated with major epigenetic changes, as illustrated by the genome-wide increase in DNA methylation in pre-gastrulating embryos that will be almost complete by E6.5 (Borgel et al. 2010). Thus, by ~E6.25, the prospective PGCs have accumulated several layers of epigenetic information and are already primed towards a somatic fate. Upon PGC specification, these epigenetic features will be erased through a major transcriptional and epigenetic reprogramming that might be important for the production of a totipotent zygote following fertilisation.

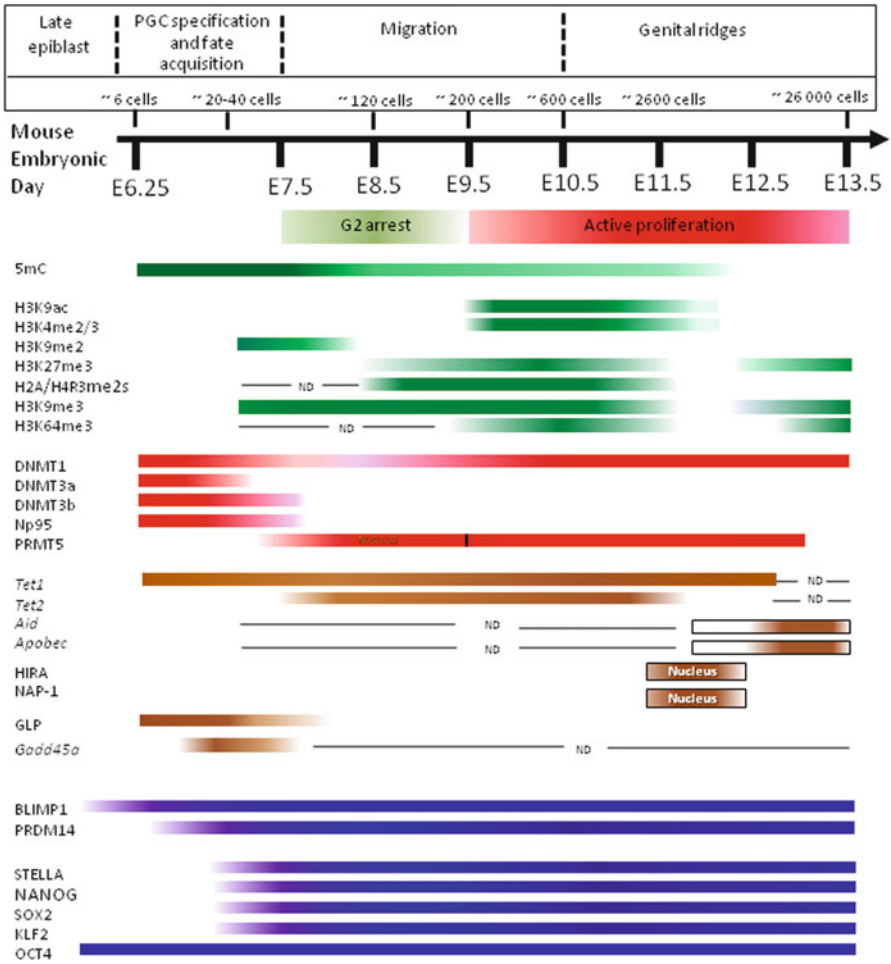

Fig. 1.2 Temporal schematic of epigenetic reprogramming during mouse primordial germ cell development. Genome-wide dynamics of DNA methylation and main histone modifications during PGC development (mainly revealed by immunochemistry analysis) are depicted. The dynamic expression of key epigenetic modifiers and pluripotency factors is also shown. Based on Kurimoto et al. (2008a, b); Ancelin et al. (2006); Seki et al. (2005, 2007); Hajkova et al. (2008, 2010); Daujat et al. (2009); and Hackett et al. (2013). PGC: Primordial germ cells, BLIMP1: B-lymphocyte-induced maturation protein 1, PRDM14: PR-domain-zinc-finger protein 14, SOX2: SRY (sex-determining region Y)-box 2, KLF2: Kruppel-like factor 2, OCT4 or POU5F1: POU class 5 homeobox 1, Dnmt: DNA (cytosine-5)-methyltransferase, NP95 or UHRF1: ubiquitin-like with PHD and ring finger domains 1, PRMT5: protein arginine methyltransferase 5, *Tet*: ten-eleven-translocation, *Aid* or *Aicda*: activation-induced cytidine deaminase, *Apobec*: apolipoprotein B mRNA editing enzyme catalytic polypeptide, HIRA: histone cell cycle regulation defective homolog A, NAP-1: nucleosome assembly protein 1, *Gadd45*a: growth arrest and DNA-damage-inducible protein 45 alpha

1.3.1 Primordial Germ Cell Specification: Reprogramming Their Transcription Pattern

PGC specification is associated with major changes in gene transcription to repress the somatic cell program and activate the germ cell-specific program, reacquire their pluripotency potential and prepare for the imminent genome-wide epigenetic reprogramming. This highly ordered process is regulated by BLIMP1 and PRDM14. At ~E6.25 these transcriptional regulators co-mark epiblast cells that will form PGCs and in the absence of either of these proteins, nascent PGC precursors fail to properly develop (Ohinata et al. 2005; Yamaji et al. 2008; Vincent et al. 2005; Kurimoto et al. 2008a, b). A single-cell microarray approach to establish the genome-wide transcription dynamics of developing PGCs and their somatic neighbours from E6.25 to E8.25 revealed that germ cell specification involves the up-regulation of nearly 500 "germ cell-specification" genes and the down-regulation of 330 "somatic program" genes (Kurimoto et al. 2008a). Among the down-regulated "somatic" genes there are many genes involved in embryonic development (e.g., *Hox* genes, *Dkk1*, *Cdx1* ...), cell cycle regulation (e.g., *Ccne1*, *Cdc25a* ...) as well as DNA methylation and histone modification, such as the de novo DNA methyltransferases DNMT3A and DNMT3B, the nuclear protein of 95 kDa (NP95, a factor essential to maintain the DNA methylation pattern during cell division) and the H3K9me2 histone methyltransferase GLP (G9a-like protein). Conversely, the "germ cell specification" category includes genes associated with germ cell development, such as *Stella* or *Fragilis*, and also the pluripotency genes *Nanog*, *Sox2* (*Sry-box2)* and *Klf2* (*Kruppel-like factor 2)* (Kurimoto et al. 2008a).

Further analysis conducted using BLIMP1-deficient PGC-like cells showed that BLIMP1 functions as a dominant repressor of the somatic program and is also involved in the reacquisition of the pluripotency potential and in the forthcoming epigenetic reprogramming. On the other hand, PRDM14 is required for *Sox2* up-regulation and *Glp* repression and is essential for the reacquisition of the pluripotency potential and for epigenetic reprogramming (Yamaji et al. 2008; Kurimoto et al. 2008b). Importantly, BLIMP1, although unnecessary to induce *Prdm14* expression, is strictly required for its maintenance (Yamaji et al. 2008).

How precisely these two proteins regulate germ cell specification remains to be established. Both BLIMP1 and PRDM14 contain a zinc-finger and histone methyltransferase SET domains, but no associated histone-modifying activity has been reported. Alternatively, they could exert their functions by recruiting effector partners to their target genes. BLIMP1 can recruit different chromatin-modifying proteins, such as histone deacetylases (HDAC) (Yu et al. 2000), G9A (Gyory et al. 2004) and the arginine methyltransferase PRMT5 (Ancelin et al. 2006). BLIMP1 and PRMT5 co-localise in the nuclei of migrating PGCs (Ancelin et al. 2006); however, it is not known whether the putative BLIMP1/PRMT5 complex is formed also during PGC specification and whether it contributes to repression of the somatic program.