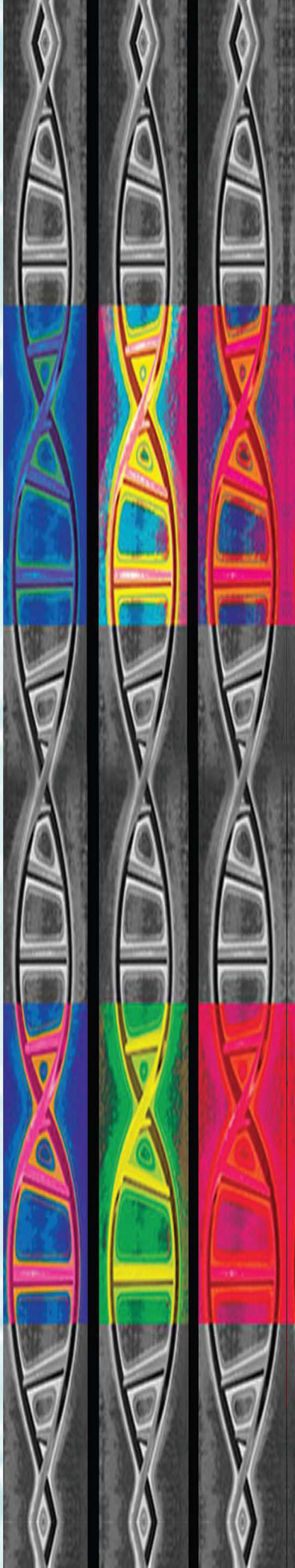
科爱传播
KE'AI COMMUNICATIONS
·导读版·

Handbook of Epigenetics

The New Molecular and Medical Genetics

表观遗传学手册

新分子遗传学与医学遗传学



Trygve Tollefsbol

ELSEVIER
原版引进

科学出版社

Handbook of Epigenetics

The New Molecular and Medical Genetics

表观遗传学手册

新分子遗传学与医学遗传学

Edited by

Trygve Tollefsbol

Department of Biology

University of Alabama at Birmingham

Birmingham , AL , 35294-1170 , USA

科 学 出 版 社

北 京

图字：01-2011-1552

ISBN：978-0-12-375709-8

图书在版编目(CIP)数据

表观遗传学手册：新分子遗传学与医学遗传学＝Handbook of Epigenetics：The New Molecular and Medical Genetics：英文/(美)托勒夫斯波(Tollefsbol，T.)主编.—北京：科学出版社，2011

ISBN 978-7-03-031742-1

Ⅰ.①表… Ⅱ.①托… Ⅲ.①分子遗传学-手册-英文②医学遗传学-手册-英文 Ⅳ.①Q75-62②R394-62

中国版本图书馆 CIP 数据核字(2011)第 121212 号

责任编辑：李小汀/责任印制：钱玉芬/封面设计：耕者设计工作室

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

双青印刷厂 印刷
科学出版社发行 各地新华书店经销
*
2011 年 7 月第 一 版 开本：787×1092 1/16
2011 年 7 月第一次印刷 印张：42 1/4
印数：1—1 800 字数：1 000 000

定价：168.00 元
(如有印装质量问题，我社负责调换)

导　读

“表观遗传学”的研究对象是一类特殊的**可继承**的生物性状的变化。不同于经典的遗传性状突变，表观遗传性状的变化并非由 DNA 突变造成。而且表观遗传性状的变化既具有可继承性，又有一定程度的可逆性。

在后基因组时代，表观遗传学是生命科学领域的一个前沿和热点。在过去的十余年中，表观遗传学研究突飞猛进，许多经典的表观遗传学现象的分子机制得到了诠释。但是，与近年来表观遗传学领域的迅猛发展以及研究论文的指数式上升所不相称的是表观遗传学领域书籍的相对贫乏。这无疑不利于青年学生和其他领域的科学家对表观遗传学产生兴趣，并加深了解。科学出版社将美国冷泉港实验室出版社 2007 年出版的《表观遗传学》先后影印及翻译出版，部分缓解了国内读者对表观遗传学领域书籍的渴望。然而，四年多来表观遗传学领域又出现了众多突破性研究成果，而且此前出版的《表观遗传学》更侧重生物学机制的阐述，较少涉及到与人类疾病直接相关的医学表观遗传学。因此，科学出版社决定引进 Elsevier 出版集团 2011 年出版的《表观遗传学手册：新分子遗传学与医学遗传学》，并以导读版的形式出版。这无疑是恰逢其时。

《表观遗传学手册：新分子遗传学与医学遗传学》由美国阿拉巴马大学 Trygve Tollefsbol 教授主编，各章节均由活跃在表观遗传学各个领域的科学家撰写。Trygve Tollefsbol 教授是研究表观遗传因素在衰老及癌症发生过程中的作用和机理的科学家，该书在介绍表观遗传学分子机理的同时，用大量的篇幅介绍了表观遗传学相关疾病，以及正在兴起的表观遗传学疗法。

由于此前冷泉港实验室出版社出版的《表观遗传学》用大量的篇幅介绍了表观遗传现象的相关分子机制。因此，Trygve Tollefsbol 教授在主编《表观遗传学手册：新分子遗传学与医学遗传学》时仅用两个章节简单地介绍了表观遗传现象的相关分子机制，但却涵盖了前书未及探讨的表观遗传学领域的两大最新进展：组蛋白去甲基化和 DNA 胞嘧啶的 5-羟甲基化。从而，有助于读者们了解表观遗传学研究的最新前沿。

《表观遗传学手册：新分子遗传学与医学遗传学》的最大特色是强调了表观遗传学机制在发育过程、正常生理状态及病理状态下的作用。这使得更多的读者能够认识到，从经典表观遗传学现象中发现的分子机制，并非仅仅局限于一些特殊的生命现象。相反，这些分子机制还在更多的生命进程中被生物体所利用，并参与正常的发育过程和生理功能。因此，表观遗传因子的异常会导致发育失常并诱发疾病。本书用多个章节介绍了表观遗传学疾病和新兴的表观遗传疗法，因此对于研究型临床医生和药物设计人员，也不失为一本有价值的参考书。

此外，本书系统性地介绍了表观遗传学领域的相关研究技术，为将要涉足表观遗传学研究的青年学生和其他领域的科学家提供了充分的入门介绍。

在介绍本书之余，笔者还希望能藉此机会向读者们介绍一下自己对表观遗传调控体系的认识。表观遗传现象，除了朊粒等极少数个例，归根到底是基因的转录水平调控。

基因的转录水平调控在绝大多数情况下取决于识别特定 DNA 序列的转录因子，例如能够将体细胞诱导成全能性细胞（iPS）的四个 Yamanaka 因子。然而，许多情况下转录因子并不足以改变基因的表达状态，因为它们的靶基因处于“封闭”的异染色质状态，阻碍了转录因子的结合。而染色质结构“封闭”或“开放”状态的形成、维持和转换，正是表观遗传学研究的核心问题。

对于刚刚对表观遗传学产生兴趣的青年学子而言，显而易见的问题是，“我们为什么需要有一个表观遗传调控体系？这一体系与经典的遗传体系的共同点、差异性和互补性又是什么?”显然，本学科的所有科学家都没有这些问题的最终答案。但一个大致的轮廓却可以被勾勒出来。表观遗传调控体系存在的基本意义毋庸置疑是为了实现多细胞复杂生物中不同细胞间功能的分化，换言之，是为了使得“拥有同一个基因组的不同体细胞能拥有不同的表观基因组，从而分别表达其特有的转录组”。因此，经典遗传体系与表观遗传体系的最大区别就在于前者是刚性的，不会发生可逆的变化；而后者则具有一定的可塑性，可以对内在或外在环境的信号作出响应，通过对表观基因组的改变实现转录组的变化。这就使得多细胞生物中不同细胞间功能的分化成为可能。而类似于经典遗传体系的是，表观遗传体系也具有一定的可继承性，从而使细胞在扩增时能够维持其特有的表观基因组和相应的转录组。

过刚易折，过柔则弱。上善若水，刚柔相济。这或许就是表观遗传体系的写照了。

朱冰
北京生命科学研究所
2011 年 6 月于静思湖畔

前　言

表观遗传学被许多人认为是“新遗传学”，因为许多生命进程并非由基因突变所控制，而是被可逆却又可继承的表观遗传现象所控制。这些表观遗传现象涵盖 DNA 甲基化、组蛋白修饰乃至朊粒等。表观遗传进程发生在众多的物种中并且控制大量的生命功能，例如组织/器官再生、X-染色体失活、干细胞分化、基因组印迹和衰老。表观遗传异常会导致各种疾病，包括癌症和免疫系统、内分泌系统、神经系统疾病。上述疾病中的一部分已经有了相应的临床干预手段，更多的表观遗传疗法很可能将在近期出现。

《表观遗传学手册：新分子遗传学与医学遗传学》全面系统地介绍了表观遗传学，并且概括了这一迷人领域的近期研究进展。本书通过阐述表观遗传学的进化、正常生命活动和病理条件下的表观遗传机理以及表观遗传学在研究和治疗中的应用，相信会吸引学生和科研及医药工业研究人员的兴趣。

（朱　冰　译）

目　录

（朱　冰　译）

CONTENTS

CONTRIBUTORS

Masahiko Abematsu
Laboratory of Molecular Neuroscience
Graduate School of Biological Sciences
Nara Institute of Science and Technology
Nara 630-0101, Japan

David W. Anderson
Department of Pathology
St. Luke's Hospital and Health Network
801 Ostrum Street
Bethlehem, PA 18015, USA

Roberta H. Andronikos
Murdoch Childrens Research Institute
Flemington Road
Parkville, Victoria 3052, Australia

Adriana Arita
Nelson Institute of Environmental Medicine
New York University School of Medicine
57 Old Forge Road, New York, NY 10987, USA

Zoya Avramova
School of Biological Sciences, UNL
E-300 Beadle Center
P.O. Box 88066
Lincoln, NE 68588-0666, USA

Autumn J. Bernal
Department of Radiation Oncology
Box 3433
Duke University Medical Center
Durham, NC 27710, USA

Javier Campión
Department of Nutrition and Food Sciences,
Physiology and Toxicology
University of Navarra, 31080 Pamplona
Navarra, Spain

Vince Castranova
PPRB/NIOSH
Basic Pharmaceutical Sciences
West Virginia University
1095 Willowdale Road
Morgantown, WV 26505, USA

Frances A. Champagne
Department of Psychology
Columbia University
Room 406, Schermerhorn Hall
1190 Amsterdam Avenue
New York, NY 10027, USA

Wendy Chao
Department of Ophthalmology
Schepens Eye Research Institute
Harvard Medical School
20 Staniford Street
Boston, MA 02114, USA

Fei Chen
PPRB/NIOSH
Basic Pharmaceutical Sciences
West Virginia University
1095 Willowdale Road
Morgantown, WV 26505, USA

Xiaowei Sylvia Chen
Department of Biochemistry
University of Otago
Dunedin 9054, New Zealand

Xiaodong Cheng
Department of Biochemistry
Emory University School of Medicine
1510 Clifton Road
Atlanta, Georgia 30322, USA

Lesley Collins
Massey Genome Service
Massey University
Palmerston North, New Zealand

Max Costa
Nelson Institute of Environmental Medicine
New York University School of Medicine
57 Old Forge Road, New York, NY 10987, USA

James P. Curley
Department of Psychology
Columbia University
Room 406, Schermerhorn Hall
1190 Amsterdam Avenue
New York, NY 10027, USA

Walter Doerfler
Institute for Virology
Erlangen University Medical School
Schlossgarten 4
D-91054 Erlangen, Germany

Tamar Dvash
Department of Human Genetics
The Eli and Edythe Broad Center for Regenerative
Medicine and Stem Cell Research

David Geffen School of Medicine
University of California Los Angeles
695 Charles E. Young Drive South
Los Angeles, CA 90095, USA

Thomas Eggermann
Institute of Human Genetics
University Hospital
RWTH Aachen, Pauwelsstrasse 30, DE-52074
Aachen, Germany

Guoping Fan
Department of Human Genetics
The Eli and Edythe Broad Center for Regenerative Medicine and Stem Cell Research
David Geffen School of Medicine
University of California Los Angeles
695 Charles E. Young Drive South
Los Angeles, CA 90095, USA

Tamara B. Franklin
Brain Research Institute
Medical Faculty of the University of Zürich and Department of Biology
Swiss Federal Institute of Technology
Winterthurerstrasse 190
CH-8057 Zürich, Switzerland

Steffen Gay
Center of Experimental Rheumatology
University Hospital and Zurich Center of Integrative Human Physiology (ZIHP)
University Hospital Zürich,
8091, Zürich, Switzerland

Eric Gilson
Laboratory of Biology and Pathology of Genomes CNRS UMR6267/INSERM U998/UNS
28 Avenue Valombrose, Faculte de Medecine
06107 Nice Cedex 2, France

Johannes Gräff
Brain Research Institute
Medical Faculty of the University of Zürich and Department of Biology
Swiss Federal Institute of Technology
Winterthurerstrasse 190
CH-8057 Zürich, Switzerland

Hideharu Hashimoto
Department of Biochemistry
Emory University School of Medicine
1510 Clifton Road
Atlanta, GA 30322, USA

Naoko Hattori
Carcinogenesis Division
National Cancer Center Research Institute
1-1 Tsukiji 5-chome
Chuo-ku, Tokyo 104-0045, Japan

Zdenko Herceg
Epigenetics Group
International Agency for Research on Cancer (IARC)
150 cours Albert Thomas
69372 Lyon cedex 08, France

Norbert Hochstein
QIAGEN
Qiagen Strasse 1
40724, Hilden, Germany

John R. Horton
Department of Biochemistry
Emory University School of Medicine
1510 Clifton Road
Atlanta, GA 30322, USA

Simon H. House
22 Stanley Street
Southsea, Portsmouth, Hants
PO5 2DS, UK

Karolina Janitz
The Ramaciotti Centre for Gene Function Analysis
The University of New South Wales, Sydney
NSW 2052, Australia

Michal Janitz
School of Biotechnology and Biomolecular Sciences
The University of New South Wales, Sydney
NSW 2052, Australia

Randy L. Jirtle
Department of Radiation Oncology
Box 3433
Duke University Medical Center
Durham, NC 27710, USA

Ji-Hoon E. Joo
Murdoch Children's Research Institute
Flemington Road
Parkville, Victoria 3052, Australia

Berry Juliandi
Laboratory of Molecular Neuroscience
Graduate School of Biological Sciences
Nara Institute of Science and Technology
Nara 630-0101, Japan

Astrid Jungel
Center of Experimental Rheumatology
University Hospital and Zurich Center of Integrative Human Physiology (ZIHP)
University Hospital Zürich,
8091, Zürich, Switzerland

Hong Kan
PPRB/NIOSH
Basic Pharmaceutical Sciences
West Virginia University
1095 Willowdale Road
Morgantown, WV 26505, USA

Tae Hoon Kim
Yale University School of Medicine
333 Cedar Street SHM I-142B
New Haven, CT 06520-8005, USA

Yoon Jung Kim
Yale University School of Medicine
333 Cedar Street SHM I-142B
New Haven, CT 06520-8005, USA

Hervé Lalucque
Institut de Genetique et Microbiologie
UMR CNRS 8621
Univ Paris Sud 11
91405 Orsay cedex, France

Sheila C.S. Lima
Epigenetics Group
International Agency for Research on Cancer (IARC)
150 Cours Albert Thomas
Rhone-Alpes, 69008 Lyon, France

Sally A. Litherland
Metabolic Signaling and Disease Program
Diabetes & Obesity Center
Sanford-Burnham Medical Research Center
at Lake Nona
6400 Sanger Road
Orlando, FL 32827, USA

John C. Lucchesi
Department of Biology
Emory University
1510 Clifton Road, Atlanta, GA 30322, USA

Hanna Maciejewska-Rodrigues
Center of Experimental Rheumatology
University Hospital and Zurich Center of
Integrative Human Physiology (ZIHP)
University Hospital Zürich,
8091, Zürich, Switzerland

Frédérique Magdinier
Laboratoire de Génétique Médicale et
Génomique Fonctionnelle
INSERM, UMR S_910/Université de la
Méditerannée
Faculté de Médecine de la Timone
27 Bd Jean Moulin 13385
Marseille cedex05, France

Fabienne Malagnac
Institut de Genetique et Microbiologie
UMR CNRS 8621
Univ Paris Sud 11
91405 Orsay cedex, France

Isabelle M. Mansuy
Brain Research Institute
Medical Faculty of the University of Zürich
and Department of Biology
Swiss Federal Institute of Technology
Winterthurerstrasse 190, CH-8057 Zürich
Switzerland

J. Alfredo Martinez
Department of Nutrition and Food Sciences,
Physiology and Toxicology
University of Navarra, 31080 Pamplona
Navarra, Spain

Rahia Mashoodh
Department of Psychology
Columbia University
Room 406, Schermerhorn Hall
1190 Amsterdam Avenue
New York, NY 10027, USA

Fermin I. Milagro
Department of Nutrition and Food Sciences,
Physiology and Toxicology
University of Navarra, 31080 Pamplona
Navarra, Spain

Ashleigh Miller
Department of Medicine
Division of Hematology and Oncology
University of California-San Francisco
San Francisco, CA 94143, USA

Pamela Munster
Department of Medicine
Division of Hematology and Oncology
University of California-San Francisco
San Francisco, CA 94143, USA

Susan K. Murphy
Department Gynecologic Oncology
Box 91092
Duke University Medical Center
Durham, NC 27710, USA

Rabih Murr
Epigenetics Group
International Agency for Research on
Cancer (IARC)
150 cours Albert Thomas
69372 Lyon cedex 08, France

Kinichi Nakashima
Laboratory of Molecular Neuroscience
Graduate School of Biological Sciences
Nara Institute of Science and Technology
Nara 630-0101, Japan

Anja Naumann
Institute for Virology
Erlangen University Medical School
Schlossgarten 4
D-91054 Erlangen, Germany

Alexandre Ottaviani
Laboratoire de Biologie Moléculaire de la Cellule
CNRS UMR5239, Ecole Normale Supérieure de Lyon
UCBL1, IFR128, 46 allée d'Italie
69364 Lyon Cedex 07, France

Jacob Peedicayil
Department of Pharmacology and Clinical Pharmacology
Christian Medical College
Vellore, India

Gerd P. Pfeifer
Division of Biology
Beckman Research Institute of the City of Hope
1500 East Duarte Road
Duarte, CA 91010, USA

Luis Felipe Ribeiro Pinto
Divisão de Genética — Coordenação de Pesquisa
Instituto Nacional de Câncer
Rua André Cavalcanti, 37, 40 Andar
CEP: 20231-050
Riode Janiero, RJ, Brazil

Vincenzo Pirrotta
Department of Molecular Biology and Biochemistry
Rutgers University
604 Allison Road
Piscataway, NJ 08854, USA

Pier Lorenzo Puri
Istituto Dulbecco Telethon at Istituto Di Ricovero e Cura a Carattere Scientifico
Santa Lucia Fondazione and European Brain Research Institute
64 Via del Fosso di Fiorano
00143 Rome, Italy
The Sanford/Burnham Institute for Medical Research
10901 North Torrey Pines Road
La Jolla, CA 92037-1062, USA

Tibor A. Rauch
Section of Molecular Medicine
Department of Orthopedic Surgery
Rush University Medical Center
Chicago, IL 60612, USA

Lee B. Riley
Department of Oncology,
St. Luke's Hospital and Health Network,
801 Ostrum Street,
Bethlehem, PA 18015, USA

Eric D. Roth
Department of Neurobiology
1825 University Blvd, SHEL 1070
Birmingham, AL 35294, USA

Tania L. Roth
Department of Neurobiology
1825 University Blvd, SHEL 1070
Birmingham, AL 35294, USA

Richard Saffery
Murdoch Childrens Research Institute
Flemington Road
Parkville, Victoria 3052, Australia

Vittorio Sartorelli
Laboratory of Muscle Stem Cells and Gene Regulation
National Institute of Arthritis, Musculoskeletal and Skin Diseases
National Institutes of Health
Bethesda, MD, USA

Caroline Schluth-Bolard
Laboratoire de Biologie Moléculaire de la Cellule
CNRS UMR5239, Ecole Normale Supérieure de Lyon
UCBL1, IFR128, 46 allée d'Italie
69364 Lyon Cedex 07, France

Barbara Schönfeld
Allan Wilson Centre for Molecular Ecology and Evolution
Massey University
Palmerston North, New Zealand

Axel Schumacher
Centre for Addiction and Mental Health,
Toronto, ON M5T 1R8, Canada

Philippe Silar
Institut de Genetique et Microbiologie
UMR CNRS 8621
Univ Paris Sud 11
91405 Orsay cedex, France

Kjetil Søreide
Department of Surgery
Stavanger University Hospital
Armaver Hansenvei 20
POB 8100, N-4068
Stavanger, Norway

J. David Sweatt
Department of Neurobiology
1825 University Blvd, SHEL 1070
Birmingham, AL 35294, USA

Scott Thomas
Department of Medicine
Division of Hematology and Oncology
University of California-San Francisco
San Francisco, CA 94143, USA

Kenneth T. Thurn
Department of Medicine
Division of Hematology and Oncology
University of California-San Francisco
San Francisco, CA 94143, USA

Trygve O. Tollefsbol
Department of Biology
175 Campbell Hall
1300 University Boulevard
University of Alabama at Birmingham
Birmingham, AL 35294-1170, USA

Toshikazu Ushijima
Carcinogenesis Division
National Cancer Center Research Institute
1-1 Tsukiji 5-chome
Chuo-ku, Tokyo 104-0045, Japan

David A. Wacker
Yale University School of Medicine
333 Cedar Street SHM I-142B
New Haven, CT 06520, USA

Stefanie Weber
Institute for Virology
Erlangen University Medical School
Schlossgarten 4
D-91054 Erlangen, Germany

Xing Zhang
Department of Biochemistry
Emory University School of Medicine
1510 Clifton Road
Atlanta, GA 30322, USA

PREFACE

Epigenetics is considered by many to be the "new genetics" because many biological processes are controlled not through gene mutations, but rather through reversible and heritable epigenetic phenomena ranging from DNA methylation to histone modifications to prions. Epigenetic processes occur in diverse organisms and control a vast array of biological functions, such as tissue/organ regeneration, X-chromosome inactivation, stem cell differentiation, genomic imprinting, and aging. Epigenetic aberrations underlie many diseases, including cancer and disorders of the immune, endocrine, and nervous systems; clinical intervention is already in place for some of these disorders and many novel epigenetic therapies are likely on the horizon.

Handbook of Epigenetics: The New Molecular and Medical Genetics is the first comprehensive analysis of epigenetics, and summarizes recent advances in this intriguing field of study. This book will interest students and researchers in both academics and industry by illuminating the evolution of epigenetics, the epigenetic basis of normal and pathological processes, and the practical applications of epigenetics in research and therapeutics.

CHAPTER 1

Epigenetics: The New Science of Genetics

Trygve O. Tollefsbol[1,2,3,4,5]
[1]Department of Biology, University of Alabama at Birmingham, AL 35294
[2]Center for Aging, University of Alabama at Birmingham, AL 35294
[3]Comprehensive Cancer Center, University of Alabama at Birmingham, AL 35294
[4]Nutrition Obesity Research Center, University of Alabama at Birmingham, AL 35294
[5]Comprehensive Diabetes Center, University of Alabama at Birmingham, AL 35294, USA

INTRODUCTION

The term epigenetics was first introduced in 1942 by Conrad Waddington and was defined as the causal interactions between genes and their products that allow for phenotypic expression [1]. This term has now been somewhat redefined and although there are many variants of the definition of this term today, a consensus definition is that epigenetics is the collective heritable changes in phenotype due to processes that arise independent of primary DNA sequence. This heritability of epigenetic information was for many years thought to be limited to cellular divisions. However, it is now apparent that epigenetic processes can be transferred in organisms from one generation to another [2–3]. This phenomenon was first described in plants [4] and has been expanded to include yeast, *Drosophila*, mouse and, possibly, humans [5–7].

THE BASICS OF DNA METHYLATION AND HISTONE MODIFICATIONS

In most eukaryotes DNA methylation, the most studied of epigenetic processes, consists of transfer of a methyl moiety from S-adenosylmethionine (SAM) to the 5-position of cytosines in certain CpG dinucleotides. This important transfer reaction is catalyzed by the DNA methyltransferases (DNMTs). The three major DNMTs are DNMT1, 3A and 3B and DNMT1 catalyzes what is referred to as maintenance methylation that occurs during each cellular replication as the DNA is duplicated. The other major DNMTs, 3A and 3B, are known more for their relatively higher *de novo* methylation activity where new 5-methylcytosines are introduced in the genome at sites that were not previously methylated. The most significant aspect of DNA methylation, which can also influence such processes as X chromosome inactivation and cellular differentiation, is its effects on gene expression. In general, the more methylated a gene regulatory region, the more likely it is that the gene activity will become down-regulated and vice versa although there are some notable exceptions to this dogma [8]. Chapter 2 of this book reviews the mechanisms of DNA methylation, methyl-CpG recognition and demethylation in mammals. Recent advances have highlighted important roles of UHRF1 and DNMT3L that are required for maintenance and *de novo* methylation,

Handbook of Epigenetics: The New Molecular and Medical Genetics. DOI: 10.1016/B978-0-12-375709-8.00001-0

respectively, and the potential inclusion of 5-hydroxymethylcytosine with 5-methylcytosine in expressing the impact of DNA methylation on the genome.

Chromatin changes are another central epigenetic process that have an impact not only on gene expression, but also many other biological processes. Posttranslational modifications of histones such as acetylation and methylation occur in a site-specific manner that influences the binding and activities of other proteins that influence gene regulation. The histone acetyltransferases (HATs) catalyze histone acetylation and the histone deacetylases (HDACs) result in removal of acetyl groups from key histones that comprise the chromatin. These modifications can occur at numerous sties in the histones and are most common in the amino terminal regions of these proteins as reviewed in Chapter 3. In general, increased histone acetylation is associated with greater gene activity and vice versa. By contrast, methylation of histones has variable effects on gene activity where lysine 4 (K4) methylation of histone H3 is often associated with increasing gene activity whereas methylation of lysine 9 (K9) of histone H3 may lead to transcriptional repression. There is also considerable crosstalk between DNA methylation and histone modifications [9] such that cytosine methylation may increase the likelihood of H3-K9 methylation and H3-K9 methylation may promote cytosine methylation.

ADDITIONAL EPIGENETIC PROCESSES

Among the most exciting advances in epigenetics have been the discoveries that many other processes besides DNA methylation and histone modifications impact the epigenetic behavior of cells. For instance, non-coding RNA (Chapter 4) including both short and long forms, often share protein and RNA components with the RNA interference (RNAi) pathway and they may also influence more traditional aspects of epigenetics such as DNA methylation and chromatin marking. These effects appear to be widespread and occur in organisms ranging from protists to humans. Prions are fascinating in that they can influence epigenetic processes independent of DNA and chromatin. In Chapter 5 it is shown that structural heredity also is important in epigenetic expression where alternative states of macromolecular complexes or regulatory networks can have a major effect on phenotypic expression independent of changes in DNA sequences. The prion proteins are able to switch their structure in an autocatalytic manner that can not only influence epigenetic expression, but also lead to human disease. The position of a gene in a given chromosome can also greatly influence its expression (Chapter 6). Upon rearrangement, a gene may be relocated to a heterochromatic region of the genome leading to gene silencing and many other gene position effects have been described, some of which may also lead to various human diseases. Polycomb mechanisms are another relatively new aspect of epigenetics that control all of the major cellular differentiation pathways and are also involved in cell fate. Polycomb repression is very dynamic and can be easily reversed by activators and they also raise the threshold of the signals or activators required for transcriptional activation which places these fascinating proteins within the realm of epigenetic processes (Chapter 7). Therefore, although DNA methylation and histone modifications are mainstays of epigenetics, recent advances have greatly expanded the epigenetic world to include many other processes such as non-coding RNA, prions, chromosome position effects and Polycomb mechanisms.

EPIGENETIC TECHNOLOGY

Many of the advances in epigenetics that have driven this field for the past two decades can be traced back to the technological breakthroughs that have made the many discoveries possible. We now have a wealth of information about key gene-specific epigenetic changes that occur in a myriad of biological processes. In Chapter 8, gene-specific techniques for determining DNA methylation are reviewed. These methods include bisulfite sequencing,

methylation-specific PCR (MSP) and quantitative MSP. These techniques can be applied not only to mechanisms of epigenetic gene control, but to diagnostic processes as well. In addition, there have been important breakthroughs in analyses of the methylome at high resolution. Microarray platforms and high-throughput sequencing have made possible new techniques to analyze genome-wide features of epigenetics that are based on uses of methylation-sensitive restriction enzymes, sodium bisulfite conversion and affinity capture with antibodies or proteins that select methylated DNA sequences. Techniques such as restriction landmark genomic sequencing (RLGS), methylation-sensitive restriction fingerprinting, methylation-specific digital karyotyping, targeted and whole genome bisulfite sequencing, methylated DNA immunoprecipitation (MeDIP) and the methylated-CpG island recovery assay are reviewed in Chapter 9. Mechanisms for lysine 9 methylation of histone H3 are reviewed in Chapter 10 and chromatin immunoprecipitation (ChIP) and chromosome conformation capture (3C) are covered in Chapter 11. The 3C-based method allows analyses of the spatial proximity of distant functional genomic sites to render a three dimensional view of the genome within the nucleus itself. Since there has been much information derived from epigenomic approaches, methods to analyze data from ChIP-on-chip and ChIP-seq, for example, are becoming increasingly important and are delineated in Chapter 12. There is no question that developments in the tools for assessing epigenetic information have been and will continue to be important factors in advancing epigenetics.

MODEL ORGANISMS OF EPIGENETICS

Epigenetic processes are widespread and much of our extant knowledge about epigenetics has been derived from model systems, both typical and unique. The ease of manipulation of eukaryotic microbes has facilitated discoveries in the molecular mechanisms of basic epigenetic processes (Chapter 13). In these cases epigenetics may play a key role in genomic protection from invasive DNA elements and in identifying the importance of gene silencing mechanisms in evolution. *Drosophila* is a mainstay model in biology in general and the epigenetics field is not an exception in this regard. For example, Chapter 14 offers a number of examples of transgenerational inheritance in *Drosophila* and this model system also shows promise in unraveling the evolutionary aspects of epigenetics. Probably the most useful model system in epigenetics to date is the mouse model (Chapter 15). Randy Jirtle and colleagues review numerous different mouse models that are important in many epigenetic processes such as transgenerational epigenetics and imprinting and these models have potential in illuminating human diseases such as diabetes, neurological disorders and cancer. Plant models (Chapter 16) are of great importance in epigenetics due in part to their plasticity and their ability to silence transposable elements. RNAi silencing in plants has been at the forefront of epigenetics and plant models will likely lead the way in several other epigenetic processes in the future. Thus, model development, like the advances in techniques, have made many of the most exciting discoveries in epigenetics possible for a number of years.

METABOLISM AND EPIGENETICS

Epigenetics is intricately linked to changes in the metabolism of organisms and these two processes cannot be fully understood separately. S-adenosylmethionine (SAM) is a universal methyl donor and drives many epigenetic processes (Chapter 17) and the importance of SAM in epigenetic mechanisms is vast. Metabolic functions can also influence the chromatin which is a major mediator of epigenetic processes (Chapter 18). It is now apparent that various environmental influences and metabolic compounds can regulate the many enzymes that modify histones in mammals. Thus, metabolic processes impact DNA methylation and chromatin remodeling, the two major epigenetic mediators, and it is likely that this relatively new field will continue to advance in an exponential manner.

FUNCTIONS OF EPIGENETICS

The functions of epigenetics are indeed numerous and it would be next to impossible to do complete justice in one book to this ever-expanding field. However, Chapters 19–25 illustrate a few of the many different functions that epigenetics mediates. Stem cells rely in part on signals from the environment and epigenetic mechanisms such as DNA methylation, histone modification, and microRNA (miRNA) have central roles in how stem cells respond to environmental influences (Chapter 19). Regenerative medicine is dependent upon stem cells and skeletal muscle regeneration (Chapter 20) involves key changes in the epigenome that regulate gene expression in muscle progenitors through chromatin as well as microRNA epigenetic changes. It has been known that epigenetics is important in X chromosome inactivation for quite some time although advances in this area are continuing to move rapidly. It is now apparent that X chromosome inactivation is regulated not only through the genes *Tsix* and *Xist*, but also pluripotency factors that affect *Xist* expression (Chapter 21). Genomic imprinting, likewise, has been known to be epigenetic-based for many years, but discoveries in this area of epigenetics continue to move at a rapid pace. Genomic imprinting is not limited to mammals but also occurs through analogous processes in plants and invertebrates and it can occur in specific tissues or during critical developmental stages (Chapter 22). Profound new discoveries have recently occurred in the area of the epigenetics of memory processes. Recent exciting discoveries have shown that gene regulation through epigenetic mechanisms is necessary for changes in adult brain function and behavior based on life experiences (Chapter 23). Moreover, new drugs that impact epigenetic mechanisms may have future uses in treating or alleviating cognitive dysfunction. Transgenerational inheritance (Chapter 24) is also a form of memory based in part on epigenetics in that early life experiences that impact epigenetic markers can greatly influence adult health and risk for diseases. In addition, the aging process is a form of epigenetic memory and experience, in that our genes are epigenetically modified from our parents and also during our entire life spans, that can significantly impact the longevity of humans as well as our risk for the numerous age-related diseases, many of which are also epigenetically-based (Chapter 25). It is therefore apparent that epigenetics influences a number of different functions and it is highly likely that many additional functions of epigenetics will be discovered in the future.

EVOLUTIONARY EPIGENETICS

Although many think of epigenetic processes as being inherent and static to a specific organism, it is apparent that epigenetics has been a major force behind the evolutionary creation of new species. Chapter 26 reveals that epigenetic mechanisms have a major influence on mutations. The evolutionary impact of epigenetics is in full force even today with the ever-changing environment that can modulate gene expression through epigenetic processes. For example, rapid changes in diet and the modern lifestyle as well as environmental pollution are undoubtedly impacting not only the human epigenome, but also the evolution of many of the more primitive species that in turn greatly affect the environment.

EPIGENETIC EPIDEMIOLOGY

Dietary factors are highly variable not only between individuals, but also among human populations and various nonhuman species. Many studies have shown that diet has a profound effect on the epigenetic expression of the genome and therefore on the phenotype. DNA methylation is the epigenetic process that has been most often associated with diet and changes in the diet may not only induce varying epigenetic expressions, but, paradoxically, a changed diet may also transfix epigenetic changes that can then be transferred to the next generation in a stable manner (Chapter 27). Environmental agents other than diet also impact the epigenome. For example, Chapter 28 reviews the many environmental agents

that can lead to alterations in the epigenome thereby inducing toxicity or carcinogenesis. Moreover, invasion by foreign agents can influence the epigenome (Chapter 29). Viruses and bacteria, for example, play a major role in altering the epigenetic expression of the genome and these processes may lead to human diseases such as cancer. Chapter 30 by Walter Doerfler and colleagues illustrates details of the role of adenovirus type 12 (Ad12) in reshaping the hamster genome and they also provide analyses of the human *FMR1* promoter that is impacted by DNA methylation in the fragile X syndrome. Drugs also reshape the epigenome, which has opened the new field of pharmacoepigenomics. It is clear that certain populations respond differently to drugs and much of this variation may be explained by epigenetic factors (Chapter 31). Thus, epidemiological factors have great importance in epigenetics and this is influenced by diet, environmental agents, infections, drugs and likely many other factors as well.

EPIGENETICS AND HUMAN DISEASES

For the medical community, a major interest in epigenetics stems from the role of epigenetic changes in the etiology, progression and diagnosis of human diseases. Cancer has long been associated with epigenetic alterations, and DNA methylation, chromatin modifications, and RNA-dependent regulation have all been shown to affect the incidence and severity of cancer (Chapter 32). Many immune disorders such as systemic lupus erythematosus (SLE) and rheumatoid arthritis as well as autoimmune disorders such as multiple sclerosis have been associated with epigenetic aberrations (Chapter 33) and epigenetic processes have also been linked to brain disorders (Chapter 34). In the latter case, the Rett syndrome, Alzheimer's disease, Huntington's disease and even autism to name a few have been associated in at least some way with epigenetic alterations. Even schizophrenia and depression may have an epigenetic basis in their expression. System metabolic disorders may also be related to epigenetic aberrations. For example, obesity, gestational diabetes and hypertension can influence the fetal chromatin and lead to an increased incidence in adult disease later in life (Chapter 35). Since genomic imprinting is based on epigenetic mechanisms, it may come as no surprise that defects in imprinting can lead to a number of human diseases (Chapter 36). The Prader–Willi syndrome, Angelman syndrome, Silver Russell syndrome and many other imprinting disorders such as transient neonatal diabetes mellitus are due to imprinting disorders that are based on epigenetic defects. Therefore, the number of diseases impacted by epigenetic processes is large and advances in the treatment of these disorders will likely depend in part on breakthroughs in epigenetic therapy.

EPIGENETIC THERAPY

Although there are many epigenetic therapies that are in use and on the horizon, histone modifying drugs have probably received the most attention in the clinics. Chief among these are the histone deacetylase (HDAC) inhibitors. Vorinostat (Zolinza), for example, has been approved by the Food and Drug Administration for use in the treatment of patients with cutaneous T-cell lymphoma (Chapter 37). Many different HDAC inhibitors have been developed and it is likely that significant improvements will occur for HDAC inhibitors as well as many other drugs that can normalize aberrations in not only histone modifications, but also DNA methylation and perhaps some of the many other epigenetic processes that have been discovered.

CONCLUSION

Advances in understanding the basic mechanisms of DNA methylation and histone modifications have raised the field of epigenetics well beyond original expectations. This area of research has also significantly expanded horizontally with the illumination of other epigenetic processes such as non-coding RNA, prion changes and Polycomb mechanisms

and it is likely that additional epigenetic processes will be discovered in the not too distant future. A major driving force in epigenetics has been the outstanding development of new technology that has not only served to stimulate new discoveries, but has also expanded the field by allowing for novel discoveries possible only through the use of these new tools. Advances in new model organisms for understanding epigenetic processes have also greatly stimulated this field of study. We now know that epigenetics is not only intricately associated with metabolism, but also functions in stem cell behavior, X chromosome inactivation, tissue regeneration, genomic imprinting, the transfer of information through generations, neurological memory processes and even the aging of organisms. Epigenetics has also played roles in evolution and has served as a molecular driver of mutations. Moreover, the changing environment is currently reshaping the evolution of many organisms through plastic epigenetic processes. Epidemiological factors such as diet, environmental exposure, microbial infections and drugs are also influencing our daily lives through epigenetics. Diseases that have been associated with epigenetic processes range from schizophrenia to cancer and the list of these diseases is rapidly expanding. Fortunately, the field of epigenetic therapy is also expanding and the hope is that the future will see many novel treatments for the numerous diseases that are derived from epigenetic defects.

References

1. Waddington CH. The epigenotype. Endeavour 1942;1:18–20.
2. Liu L, Li Y, Tollefsbol TO. Gene-environment interactions and epigenetic basis of human diseases. Curr Issues Mol Biol 2008;10:25–36.
3. Chong S, Whitelaw E. Epigenetic germline inheritance. Curr Opin Genet Dev 2004;14:692–6.
4. Brink RA, Styles ED, Axtell JD. Paramutation: directed genetic change. Paramutation occurs in somatic cells and heritably alters the functional state of a locus. Science 1968;159:161–70.
5. Cavalli G, Paro R. Epigenetic inheritance of active chromatin after removal of the main transactivator. Science 1999;286:955–8.
6. Grewal SI, Klar AJ. Chromosomal inheritance of epigenetic states in fission yeast during mitosis and meiosis. Cell 1996;86:95–101.
7. Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. Metastable epialleles in mammals. Trends Genet 2002;18:348–51.
8. Lai SR, Phipps SM, Liu L, Andrews LG, Tollefsbol TO. Epigenetic control of telomerase and modes of telomere maintenance in aging and abnormal systems. *Front Biosci* 2005;**10**:1779–96.
9. Fuks F, Burgers WA, Brehm A, Hughes-Davies L, Kouzarides T. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat Genet* 2000;**24**:88–91.

SECTION I

Molecular Mechanisms of Epigenetics

CHAPTER 2

Mechanisms of DNA Methylation, Methyl-CpG Recognition, and Demethylation in Mammals

Xiaodong Cheng, Hideharu Hashimoto, John R. Horton, and Xing Zhang
Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA

INTRODUCTION

The control of transcription initiation in mammalian cells can be very broadly divided into three categories: intrinsic promoter strength and availability of core transcription machinery [1–3], the actions of promoter- or regulon-specific transcription factors (positive and negative) [4–6], and the control of DNA accessibility by altering chromatin structure [6–8]. This latter category, including posttranslational modifications to histones and postreplicational modification of DNA, is the focus of recent extensive studies. Nucleosomes are the fundamental building blocks of eukaryotic chromatin, and consist of ~146 base pairs of DNA wrapped twice around a histone octamer [9]. A variety of protein-modifying enzymes (including methyltransferases, MTases] is responsible for histone modification, primarily at their flexible N-termini [10–12]. Here, we summarize the most recent structural and biochemical advances in the study of mammalian DNA MTases and their associated protein factor(s), and will touch on the functional links between histone modification and that of DNA.

In mammals and other vertebrates, DNA methylation occurs at the C5 position of cytosine (5mC), mostly within CpG dinucleotides (Fig. 2.1A), with the Dnmt enzymes using a conserved mechanism [13] that has been studied best in the bacterial 5mC MTase M.HhaI [14–18]. Briefly, this mechanism involves MTase binding to the DNA, eversion of the target nucleotide so that it projects out of the double helix ("base flipping"), covalent attack of a conserved Cys nucleophile on cytosine C6, transfer of the methyl group from *S*-adenosyl-L-methionine (AdoMet) to the activated cytosine C5, and the various release steps. This methylation, together with histone modifications, plays an important role in modulating chromatin structure, thus controlling gene expression and many other chromatin-dependent processes [19]. The resulting epigenetic effects maintain the various patterns of gene expression in different cell types [20]. Epigenetic processes include genomic imprinting [21], gene silencing [22,23], X chromosome inactivation [24], reprogramming in transferred nuclei [25,26], and some elements of carcinogenesis [27].

Handbook of Epigenetics: The New Molecular and Medical Genetics. DOI: 10.1016/B978-0-12-375709-8.00002-2

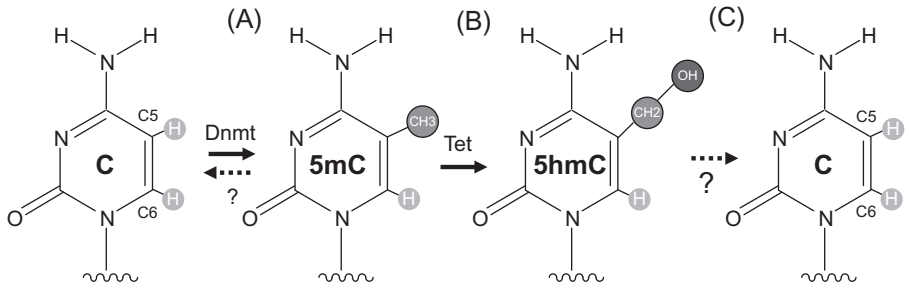


FIGURE 2.1

DNA cytosine methylation, hydroxylation, and demethylation. (A) The question mark indicates possible activity of DNA demethylases [150–155]. (B) Conversion of 5 mC to 5 hmC in mammalian DNA by the MLL fusion partner TET1 [129]. (C) It is currently unknown whether 5hmC is an end product or an intermediate in active DNA demethylation. The question mark indicates a possible MTase-assisted removal of the C5-bound hydroxymethyl group [131] (Please refer to color plate section)

DNA methylation is also associated with phenomena such as DNA repair [28], initiation of sexual dimorphism [29], progression through cell division checkpoints [30], and suppression of the huge number of transposable and retroviral elements in the mammalian genome [31–33].

MAMMALIAN DNA MTases

In mammals, Dnmts include three members, in two families that are structurally and functionally distinct (Fig. 2.2A). The Dnmt3a and Dnmt3b establish the initial CpG methylation pattern *de novo*, while Dnmt1 maintains this pattern during chromosome replication [34] and repair [35] (Fig. 2.2B). As befits a maintenance MTase, Dnmt1 has a 30–40-fold preference for hemimethylated sites [discussed in refs 36 and 37]. However, this division of labor is not absolute, as Dnmt1 activity is required for *de novo* methylation at non-CpG cytosines [38], and perhaps to an extent even in CpG islands [39,40].

The Dnmt3 family includes two active *de novo* Dnmts, Dnmt3a and Dnmt3b, and one regulatory factor, Dnmt3-Like protein (Dnmt3L) [41] (Fig. 2.2A). Dnmt3a and Dnmt3b have similar domain arrangements: both contain a variable region at the N-terminus, followed by a PWWP domain that may be involved in nonspecific DNA binding [42,43], a Cys-rich 3-Zn binding domain (comprising six CXXC motifs), and a C-terminal catalytic domain. The amino acid sequence of Dnmt3L is very similar to that of Dnmt3a and Dnmt3b in the Cys-rich 3-Zn binding domain, but it lacks the conserved residues required for DNA MTase activity in the C-terminal domain.

Dnmt3L IS A REGULATORY FACTOR FOR *DE NOVO* DNA METHYLATION

The phenotype of Dnmt3L knockout mice is indistinguishable from that of a Dnmt3a germ-cell-specific conditional knockout, with both having altered sex-specific *de novo* methylation of DNA sequences in germ cells and dispersed retrotransposons [44–47]. These results indicate that Dnmt3a and Dnmt3L are both required for the methylation of most imprinted loci in germ cells. Dnmt3L co-localizes and co-immunoprecipitates with both Dnmt3a and Dnmt3b [48], and enhances *de novo* methylation by both of these MTases [49–53]. The minimal regions required for interaction between Dnmt3L and Dnmt3a (or Dnmt3b), and for stimulated activity, are in the C-terminal domains of both proteins [50–54], as illustrated by the structure of the complex between C-terminal domains of Dnmt3a and Dnmt3L [55] (Fig. 2.3A).

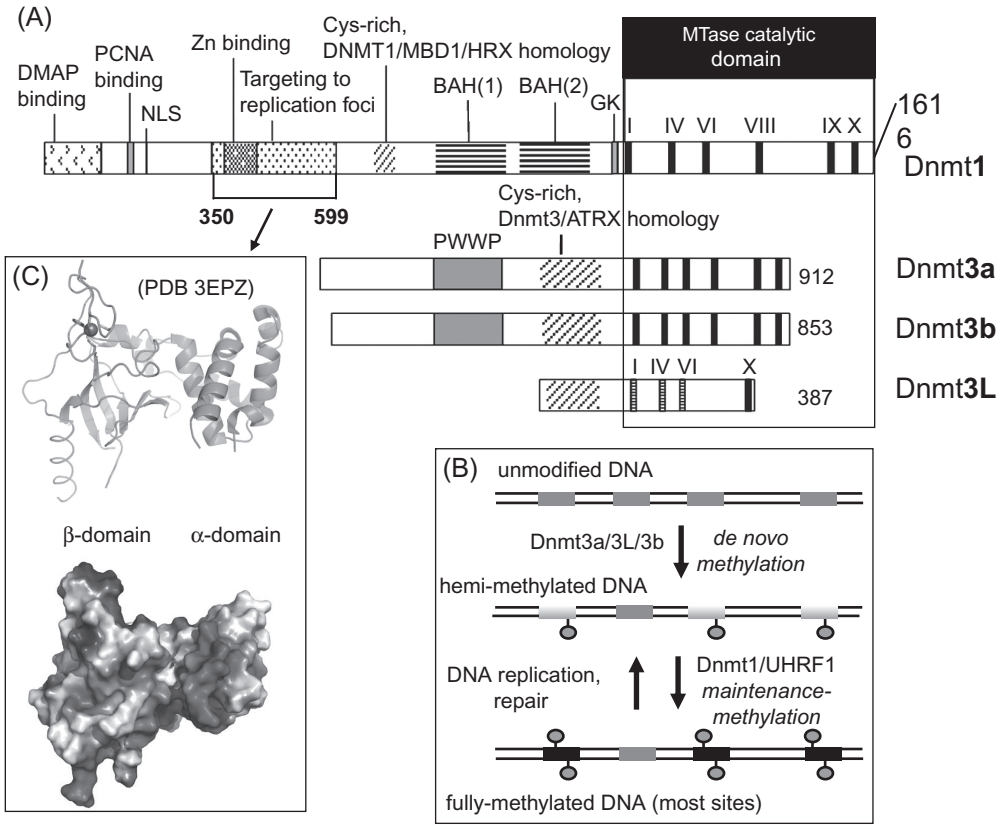


FIGURE 2.2

Schematic representation of Dnmt1 and Dnmt3. (A) Roman numerals refer to conserved motifs of DNA MTases [156]; motif IV includes the Cys nucleophile that forms a transient covalent bond to C6 of the target cytosine. (B) Maintenance vs. *de novo* methylation. The rectangular segments are substrate sequences (usually CpG), and the small ball shapes represent methyl groups on the cytosines. Following replication or repair, the duplex is methylated on one strand only. (C) The first domain structure of Dnmt1 (residues 350–599; PDB 3EPZ) [89] contains targeting sequence association with replication foci [88]. (Please refer to color plate section)

Both Dnmt3a and Dnmt3L C-terminal domains have the characteristic fold of Class I AdoMet-dependent MTases [56]. However, the methylation reaction product *S*-adenosyl-L-homocysteine (AdoHcy) was found only in Dnmt3a and not in Dnmt3L. This is consistent with Dnmt3a being the catalytic component of the complex, while Dnmt3L is inactive and unable to bind AdoMet [52,53]. The overall Dnmt3a/Dnmt3L C-terminal complex is ~16 nm long, which is greater than the diameter of a 11-nm core nucleosome (Fig. 2.3A). This complex contains two monomers of Dnmt3a and two of Dnmt3L, forming a tetramer with two 3L-3a interfaces and one 3a-3a interface (3L-3a-3a-3L). Substituting key non-catalytic residues at the Dnmt3a-3L or Dnmt3a-3a interfaces eliminates enzymatic activity, indicating that both interfaces are essential for catalysis [55].

DIMERIC Dnmt3a SUGGESTS *DE NOVO* DNA METHYLATION DEPENDS ON CpG SPACING

Among known active DNA MTases, Dnmt3a and Dnmt3b have the smallest DNA binding domain (though it is absent altogether in Dnmt3L). However, dimerization via the 3a-3a interface brings two active sites together and effectively doubles the DNA-binding surface. Superimposing the Dnmt3a structure, onto that of M.HhaI complexed with a short oligonucleotide [14], yielded a model such that the two active sites are located in the DNA major groove and dimeric Dnmt3a could methylate two CpGs

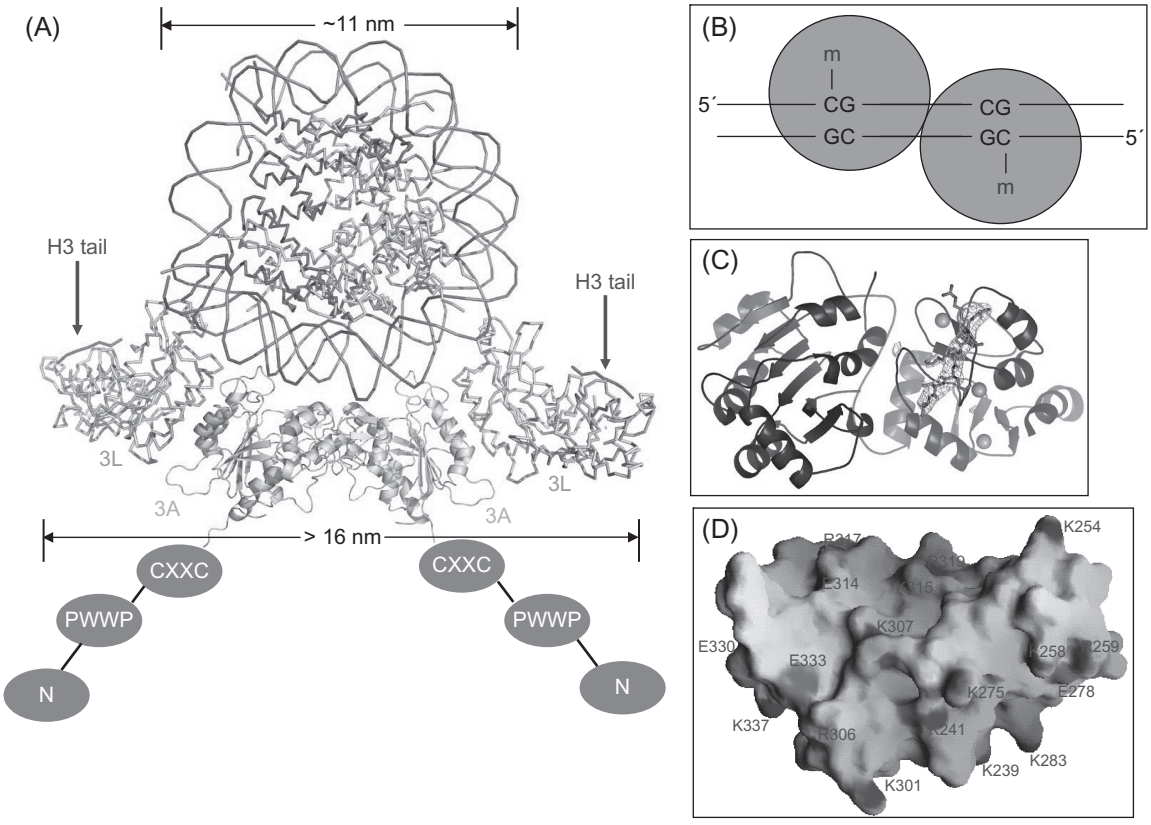


FIGURE 2.3

A model of interactions between Dnmt3a-3L tetramer and a nucleosome. (A) A nucleosome is shown, docked to a Dnmt3L-3a-3a-3L tetramer (3a-C in green; 3L full length in gray). The position of a peptide derived from the sequence of the histone H3 amino terminus (purple) is shown, taken from a co-crystal structure with this peptide bound to Dnmt3L [71]. Wrapping the tetramer around the nucleosome, the two Dnmt3L molecules could bind both histone tails from one nucleosome. The amino-proximal portion of Dnmt3a is labeled as N (for N-terminal domain), PWWP domain, and CXXC domain. By analogy to Dnmt3L, the CXXC domain of Dnmt3a might interact with histone tails from neighboring nucleosomes. (B) The Dnmt3a dimer could in theory methylate two CpGs separated by one helical turn in one binding event. (C) Structure of Dnmt3L with a bound histone H3 N-terminal tail (orange) [71]. (D) The PWWP domain structure of mouse Dnmt3b, rich in basic residues [42]. (Please refer to color plate section)

separated by one helical turn in one binding event (Fig. 2.3B). A periodicity in the activity of Dnmt3a on long DNA substrates revealed a correlation of methylated CpG sites at distances of 8–10 base pairs, and the structural model of oligomeric Dnmt3a docked to DNA may explain this pattern [55]. Similar periodicity is observed for the frequency of CpG sites in the differentially-methylated regions of 12 maternally-imprinted mouse genes [55]. These results suggest a basis for the recognition and methylation of differentially-methylated regions in imprinted genes, involving detection of both CpG spacing and nucleosome modification (see next section). Zhang et al. (2009) analyzed the methylation status of a large number of CpG sites (total of 580,427) of chromosome 21 and found that CpG DNA methylation patterns are correlated with the CpG periodicity of nine base pairs [57]. More recently, an 8–10 base-pair periodicity has also been evident for non-CpG methylation in embryonic stem cells [58]. Non-CG methylation disappeared upon induced differentiation of the embryonic stem cells, and was restored in induced pluripotent stem cells. Similarly, a 10-bp correlation of non-CpG DNA methylation by *Arabidopsis thaliana* DRM2 (which is related to mammalian Dnmt3a) has been observed [59].

Dnmt3L CONNECTS UNMETHYLATED HISTONE H3 LYSINE 4 TO *DE NOVO* DNA METHYLATION

DNA methylation and histone modifications are intricately connected with each other [60–62]. In fact, genome-scale DNA methylation profiles suggest that DNA methylation is better correlated with histone methylation patterns than with the underlying genome sequence context [62]. Specifically, DNA methylation is correlated with the absence of H3K4 methylation and the presence of H3K9 methylation. Methylation of histone H3 lysine 4 (H3K4) [63] has been suggested to protect gene promoters from *de novo* DNA methylation in somatic cells [64,65]. There have been reports of an inverse relationship between H3K4 methylation and allele-specific DNA methylation at differentially methylated regions [57,62,66–69]. More recently, AOF1 (amine-oxidase flavin-containing domain 1), a homolog of histone H3 lysine 4 demethylase (LSD1), has been shown to be required for *de novo* DNA methylation of imprinted genes in oocytes [70], suggesting that demethylation of H3K4 is critical for establishing the DNA methylation imprints during oogenesis.

The mammalian *de novo* DNA methylation Dnmt3L-Dnmt3a machinery could translate patterns of H3K4 methylation, which are not known to be themselves preserved during chromosome replication, into heritable patterns of DNA methylation that mediate transcriptional silencing of the affected sequences [71]. Dnmt3a is fully active on nucleosomal DNA *in vitro* [72]. Dnmt3a2 is a shorter isoform of Dnmt3a, predominant in embryonic stem cells and embryonal carcinoma cells and detectable in testis, ovary, thymus, and spleen, that is also required for genomic imprinting [73]. Dnmt3a2 and Dnmt3b, along with the four core histones, were identified as the main *in vivo* interaction partners of epitope-tagged Dnmt3L [71]. Peptide interaction assays showed that Dnmt3L specifically interacts with the extreme amino terminus of histone H3; this interaction was strongly inhibited by H3K4 methylation, but was insensitive to modifications at other positions [71]. Co-crystallization of Dnmt3L with the amino tail of H3 showed this tail bound to the Cys-rich 3-Zn binding domain of Dnmt3L (Fig. 2.3C), and substitution of key residues in the binding site eliminated the H3-Dnmt3L interaction. These data suggest that Dnmt3L is a probe of H3K4 methylation, and if the methylation is absent then Dnmt3L induces *de novo* DNA methylation by docking activated Dnmt3a2 to the nucleosome.

Mouse ES cells that lack the H3 lysine 9 (H3K9) MTases Suv39h1 and Suv39h2 show slight demethylation of satellite DNA [74]. G9a and GLP (G9a-like protein) – two related euchromatin-associated H3K9 methyltransferases [75] – have been implicated in DNA methylation at various loci, including imprinting center [76,77], retrotransposons and satellite repeats [78], a G9a/GLP target promoter [79], and a set of embryonic genes [80]. In filamentous fungi *Neurospora*, the H3K9 methyltransferase DIM-5 is required for DNA methylation [81–84], whereas in *Arabidopsis* the H3K9 methyltransferase KRYPTONITE is required for DNA methylation [85]. This suggests an evolutionarily-conserved silencing pathway in which H3K9 methylation correlates with DNA methylation. However, how H3K9 methylation contributes to DNA methylation is not clear, particularly in mammalian cells. G9a interacts directly with Dnmt1 during replication [86]. In addition, the G9a ankyrin repeat domain has been suggested to interact with Dnmt3a [80,87], a possible way for G9a to induce *de novo* DNA methylation [78].

A STRUCTURAL FRAGMENT OF Dnmt1

At the time of this writing (August, 2009), one domain structure is available (Fig. 2.2C) for part of the large 183 kDa Dnmt1 protein. The region (residues 350–599 of human Dnmt1) was initially identified as a novel targeting sequence association with replication foci [88]. This sequence has the properties expected of a targeting sequence in that it is not required for enzymatic activity, prevents proper targeting when deleted, and, when fused to

β-galactosidase, causes the fusion protein to associate with replication foci in a cell cycle-dependent manner. The domain structure, solved by the Structural Genomics Consortium at Toronto (PDB 3EPZ) [89], adopts a mainly β structure in the N-terminal half and a helix bundle in the C-terminal half (Fig. 2.2C).

Dnmt1 itself is subject to posttranslational modifications, including phosphorylation (Ser515 in mouse Dnmt1) [90,91] and methylation (Lys142 in human Dnmt1) [92]. Methylation of Dnmt1 at Lys142, mediated by Set7 (a protein lysine methyltransferase), resulted in its decreased stability [92]. Reciprocally, enhanced Dnmt1 methylation in the background of total deletion of LSD1 (a protein lysine demethylase) correlates with reduced Dnmt1 stability *in vivo* and progressive loss of DNA methylation [93]. Furthermore, it was hypothesized that polymers present on PARP-1 (PARylated poly(ADP-ribose) polymerase 1) interact noncovalently with Dnmt1, preventing Dnmt1 enzymatic activity. In the absence of poly(ADP-ribosyl)ation of PARP-1, Dnmt1 is free to methylate DNA; if, in contrast, high levels of PARylated PARP-1 persist, Dnmt1 will be stably inhibited, preventing DNA methylation [94].

THE SRA DOMAIN OF UHRF1 FLIPS 5-METHYLCYTOSINE OUT OF THE DNA HELIX

An accessory protein UHRF1 (ubiquitin-like, containing PHD and RING finger domains 1) targets Dnmt1 to hemimethylated replication forks (and presumably repair sites) [95–97]. The murine ortholog of this protein is also known as NP95 (nuclear protein of 95 kDa) [98–100]; the human ortholog is called ICBP90 (inverted CCAAT binding protein of 90 kDa) [101].

The crystal structure of the SET and RING associated (SRA) domain of UHRF1 in complex with DNA containing a hemimethylated CpG site was recently determined [102–104]. They reveal that the SRA domain flips the 5-methylcytosine (5mC) completely out of the DNA helix (Fig. 2.4A,B) and is positioned in a binding pocket with planar stacking contacts, and both Watson-Crick polar hydrogen bonds and van der Waals interactions specific for 5mC that distinguishes 5mC from cytosine. The structure also suggests an explanation for the preference for hemimethylated sites. In the major groove side, a backbone carbonyl oxygen is close to the C5 ring carbon of the unmethylated cytosine, forming a C=O···H-C hydrogen bond. The addition of a methyl group to C5 of the unmethylated cytosine would cause a steric clash between the methyl group and SRA.

BASE FLIPPING MECHANISM

Base flipping is a conserved mechanism that is widely used by nucleotide modifying enzymes, including DNA MTases [13,14], DNA repair enzymes [105–108], and RNA modification enzymes [109]. This mechanism, first discovered in the bacterial 5mC MTase M.HhaI [14], involves enzyme binding to the DNA and eversion of the target nucleotide so that it projects out of the double helix and into the active-site pocket. The SRA domain is the first-discovered non-enzymatic sequence-specific DNA binding protein domain that uses the base flipping mechanism in its interaction with DNA.

There is no apparent sequence or structural similarity between the SRA and the DNA MTase domain (or of DNA repair enzymes). However, the phosphodiester backbone pinching [110] due to extensive protein-phosphate contacts surrounding the flipped nucleotide, the use of two loops to approach DNA from the major and minor grooves simultaneously, and the binding of the flipped base in a concave pocket are analogous to the DNA MTases (Fig. 2.5A,B) [111]. Furthermore, enzymes use base flipping to gain access to a DNA base to perform chemistry on it, but the SRA domain probably uses base flipping to increase its protein-DNA interface and to prevent the SRA domain from linear diffusion away from the