Progress in Drug Research 67
Series Editors: Paul L. Herrling · Alex Matter

Susan M. Gasser
En Li
Editors

Epigenetics and Disease

Pharmaceutical Opportunities

Springer

Progress in Drug Research

Volume 67

Founding Editor of the Series
Ernst Jucker

For further volumes:
http://www.springer.com/series/4857

Susan M. Gasser • En Li
Editors

Epigenetics and Disease

Pharmaceutical Opportunities

Springer

Editors
Dr. Susan M. Gasser
Friedrich Miescher Institute for
Biomedical Research
Part of the Novartis Research Foundation
Maulbeerstrasse 66
4058 Basel
Switzerland
susan.gasser@fmi.ch

Dr. En Li
China Novartis Institutes for BioMedical
Research Co., Ltd.
Lane 898 Halei Road
201203 Shanghai
Zhangjiang Hi-Tech Park
Pudong New Area
People's Republic of China
en.li@novartis.com

ISBN 978-3-7643-8988-8 e-ISBN 978-3-7643-8989-5
DOI 10.1007/978-3-7643-8989-5

Library of Congress Control Number: 2010937861

Cover design: deblik Berlin, Germany

Printed on acid-free paper

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

Preface

The field of epigenetics seeks to explain how cell-type specific controls over gene expression are maintained during self renewal and differentiation and are altered by environmental events. The field covers the biochemistry of genome organization, its expression, inheritance, as well as controls over translation, message stability, and the relationship of all these with external signals. Epigenetics is now at an exciting stage, one comparable to the field of genetics before the elucidation of the structure of DNA and the genetic code. We know most of the players, but are still unclear as to how they work together to maintain gene expression states faithfully. As our understanding of epigenetic inheritance expands, the field intersects increasingly with fields focused on human health and disease. The topics most directly addressed are control over stem cell status and modulation of cell differentiation, which are at the heart of degenerative disease and cancer. Finally, from the study of chromatin and miRNA modifications, we can expect to identify targets, biomarkers, and diagnostic tools relevant for biomedical application.

What new medical opportunities are opened by the field of epigenetics? We note that pharmaceutical developments have in the past avoided targets that control mammalian gene expression, because controls were thought to arise from combinatorial protein–DNA and protein–protein interactions that are difficult to interfere with. Advances in epigenetics have identified enzymes that modify histones, DNA, and other proteins, that collectively control the compaction and organization of chromatin domains, to regulate gene expression. They influence events ranging from transcription, splicing, and mRNA stability to translation. The definition of these molecular modifications and the enzymatic machinery that controls them render gene regulation "targetable" in ways that were not possible in the past. Given that these epigenetic targets guide gene expression both during development and in adult tissues, they become of particular biomedical relevance to a broad range of diseases, including developmental disorders, aging, cancer and tissue degeneration. They are expected to impact the discovery and development of novel treatments.

The field is still exploring what are reasonable goals for an "epigenetics" program in relation to novel medications or developments that promote human health. Below we list a few of the reasonable targets and goals that justify epigenetic studies within the medical world.

1. *Relevant and "druggable" enzymes*. These include new targets such as histone deacetylases, histone acetyltransferases, histone methyltransferases, DNA methyltransferases, demethlyating enzymes that act on either DNA or protein, ubiquitinylating E3 ligases, deubiquination enzymes, Sumoylating and Neddylating enzymes, helicases, nucleosome remodelers, as well as subdomains of histone modifying enzymes, such as SET domains. Moreover, there are enzymes that signal and control these enzymes, sending signals to chromatin and the translational apparatus. Many of these have been shown to be reasonable and effective targets in cancer cells, for they are upregulated in and essential to tumor cell growth.
2. *Diseases arising from loss or alteration of epigenetic marks*. Examples include Rett syndrome, Prader–Willi/Angelman Syndrome, Fragile X Syndrome, Beckwith–Wiedermann, ATRX, Hutchison–Gilford progeria, schizophrenia and several leukemias. These diseases are useful as models for proof of concept for drugs targeting more general defects. Mouse models of such diseases are particularly useful for understanding human disease and age-related degenerative phenotypes.
3. *Diagnostics and biomarkers*. This arises from genome-wide profiling of modifications, independent of transcription, but correlated with a disease state or response to stress or oncogenic transformation. By monitoring DNA methylation on promoters, or histone modifications generally, we are able to predict the differentiation state of cells, be it in degenerative disease or cancer. This provides a powerful read-out for toxicity, for changes in cellular state, as well as for patient stratification in clinical trials (see below).
4. *Patient stratification*. Epigenomic profiling is a means to select patients and help identify tumor type, prior to clinical trials. The same read-outs are useful to test drug toxicity on both normal and diseased tissues, or as diagnostics of response spectra.
5. *Regenerative medicine*. Cell differentiation correlates precisely with epigenetic changes on the genome-wide level. These can be monitored with high throughput sequencing and Chromatin-IP-sequencing techniques. Small molecules that alter the differentiation state and potential of cells are being discovered. The restoration of a differentiated or pluripotent status to otherwise normal cells may help treat both degenerative disease and cancer. The goal of reprogramming cell fate is within reach, as is intervention to prevent aberrant responses that might alter gene expression profiles in a heritable manner.

We note that particularly cancer and neurological disorders can be traced to misregulation of epigenetic marks. Examples are as shown in Table 1 (adapted with permission from Rodenhiser, D and Mann, M. (2006) CMAJ 174(3), pp 341–348).

Table 1 Associations between epigenetic modifications and human diseases and conditions

Disease/condition	Gene	Biological process	Disease/condition	Gene	Biological process
Cancer			Neurologic		
Bladder	Multiple genes	Hypermethylation	Schizophrenia	*RELN*	Hypermethylation
Brain (glioma)	*RASSF1A*	Hypermethylation	Bipolar disorder	*11p?*	Unknown
Brain (glioblast)	*MGMT*	Hypermethylation	Memory formation	Multiple genes	Hypo-, hypermethylation
Breast	*BRCA1*	Hypermethylation	Lupus	Retroviral DNA	Hypomethylation
Breast	Multiple genes	Hypermethylation	Cardiovascular		
Cervix	*P16*	Hypermethylation	Atherosclerosis	Multiple genes	Hypo-, hypermethylation
Colon	Multiple genes	Hypermethylation	Homocysteinemia	Multiple genes	Hypomethylation
Colorectal	L1 repeats	Hypomethylation	Vascular endothelium	*eNOS*	Hypomethylation
Esophagus	*CDH1*	Hypermethylation	Imprinting and pediatric syndromes		
Head/neck	*p16*, *MGMT*	Hypermethylation	PWS or AS	15q11-q13	Imprinting
Kidney	*TIMP-3*	Hypermethylation	BWS	11p15	Imprinting
Leukemia	p15	Hypermethylation	SRS	Chromosome 7	Imprinting
Liver	Multiple genes	Hypermethylation	UPD14	14q23-q32	Imprinting
Lung	*p16*, *p73*	Hypermethylation	PHP, AHO, MAS	20q13.2	Imprinting
Lymphoma	*DAPK*	Hypermethlation	Rett syndrome	*MECP2*	Mutation
Myeloma	*DAPK*	Hypermethylation	ICF syndrome	*DNMT3B*	Mutation
Ovary	*BRCA1*	Hypermethylation	ATRX	*ATRX*	Chromation structure
Ovary	*Sat2*	Hypomethylation	FraX	Triplet repeat	Silencing
Pancreas	*APC*	Hypermethylation	FSHD	3.3 kb repeat	Chromatin structure
Pancreas	Multiple genes	Hypomethylation	Reproductive		
Prostate	*BRCA2*	Hypermethylation	Ovarian teratoma	No paternal genome	Imprinting
		Rhabdomyosarcoma	*PAX3*		Hypermethylation
CHM	No maternal genome	Imprinting			
Stomach	*Cyclin D2*	Hypomethylation	BiCHM	Maternal genome	Imprinting
Thymus	*POMC*	Hypomethylation	Aging	Chromatin	Hypo-, hypermethylation
Urothelial	Satellite DNA	Hypomethylation			
Uterus	*hMLH1*	Hypermethylation			

Note: *PWS* Prader–Willi syndrome; *AS* Angelman syndrome; *BWS* Beckwith–Weidemann syndrome; *SRS* Silver–Russell syndrome; *UPD14* uniparental disomy 14; *PHP* pseudohypoparathyroidism; *AHO* Albright hereditary osteodystrophy; *MAS* McCune–Albright syndrome; *ICF* immunodeficiency, centromeric instability, and facial anomalies; *ATRX* a-thalassemia/mental retardation syndrome, X-linked; *FraX* Fragile X syndrome; *FSHD* facioscapulohumeral muscular dystrophy, *CHM* complete hydatidiform mole, *BiCHM* familial biparental CHM

The current book aims to explore novel ideas about diagnostics, treatments, and the power of exploiting regenerative medicine for diseases that have long been inaccessible to medicine. We do not aim at being comprehensive, but rather

forward looking and innovative as we explore the current questions facing biomedical epigenetic research. The goal of this edition is to help define the field of epigenetics in relation to human disease, in order to benefit the medical world, the pharmaceutical industry, and the academic research scientist. Indeed, there are only a few cases in which small molecule inhibitors of epigenetic enzymes have made it to clinical trials. These are the start of a dynamic interaction of epigenetics and drug discovery and will be presented here.

We thank the contributors to this volume for their readiness to submit their visions of the field at short notice. We thank our collaborators who have helped review and discuss the many aspects of epigenetics and disease, and SG wishes to thank her assistant, Nicole Jascur for exceptional support in this project.

Basel, Switzerland — Susan M. Gasser
Shanghai, People's Republic of China — En Li
July 2010

Contents

DNA Methylation and Cancer

Phillippa C. Taberlay and Peter A. Jones

Abstract DNA methylation acts in concert with other epigenetic mechanisms to regulate normal gene expression and facilitate chromatin organization within cells. Aberrant DNA methylation patterns are acquired during carcinogenic transformation; such events are often accompanied by alterations in chromatin structure at gene regulatory regions. The expression pattern of any given gene is achieved by interacting epigenetic mechanisms. First, the insertion of nucleosomes at transcriptional start sites prevents the binding of the transcriptional machinery and additional cofactors that initiate gene expression. Second, nucleosomes anchor all of the DNMT3A and DNMT3B methyltransferase proteins in the cell, which suggests a role for histone octamers in the establishment of DNA methylation patterns. During carcinogenesis, epigenetic switching and 5-methylcytosine reprogramming result in the aberrant hypermethylation of CpG islands, reducing epigenetic plasticity of critical developmental and tumor suppressor genes, rendering them unresponsive to normal stimuli. Here, we will discuss the importance of both established and novel molecular concepts that may underlie the role of DNA methylation in cancer.

1 Overview

The eukaryotic genome is complex and has evolved to enable large amounts of DNA to be contained within the boundary of the nucleus. The structural organization of DNA into chromatin involves several orders of compaction and creates an environment that is generally repressive for gene transcription. However, chromatin is a highly dynamic structure that must be modified to accommodate the transcriptional

P.C. Taberlay and P.A. Jones
Department of Urology, Biochemistry and Molecular Biology, USC/Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, California 90033, USA
e-mail: pjones@med.usc.edu

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_1,

machinery when gene expression is required, to facilitate DNA repair mechanisms, or to allow DNA replication [1]. Epigenetic regulation of these processes is typically driven in a cell type-dependent manner during and following differentiation from totipotency. It has also been established that epigenetic mechanisms, such as DNA methylation, govern many aspects of embryonic growth from conception and are necessary for the survival of mammals. Since several enzymatic systems coordinate epigenetic modifications, a high level of combinatorial control must be maintained to ensure the correct chromatin conformation and identity of each cell. To this end, it is now apparent that alterations to normal epigenetic processes deregulate biological signaling pathways, contributing to carcinogenesis and disease. Here we will discuss mechanisms that may be involved in establishing aberrant DNA methylation patterns in carcinogenesis.

2 Mechanisms of Silencing by DNA Methylation

The biology of DNA methylation events in cancer is currently the best characterized epigenetic aberration in disease [2]. DNA methylation is a relatively stable modification that occurs in the context of CpG dinucleotides in mammalian cells. The presence of CpG sites within the genome is irregular, with some regions containing a high frequency of CpG dinucleotides (CpG islands) in contrast to areas where this dinucleotide is underrepresented. The distribution of CpG sites throughout the genome has implications for cellular gene expression profiles. First, CpG rich regions are often situated in promoters that are proximal to the transcription start sites of many genes while the remainder of the genome is relatively CpG poor, including sites of viral integration as well as intergenic and intronic regions [3]. Second, not all CpG sites in the genome are methylated. CpG islands are resistant to de novo methylation in normal cells [4, 5], while CpG poor regions are predisposed to this process [6].

Distinct methylation patterns are established during embryonic development and are mitotically heritable through many cellular divisions. The faithful maintenance of normal DNA methylation patterns is disrupted in cancer, where CpG islands become susceptible to methyltransferase activity and CpG poor regions undergo hypomethylation during transformation. Consistent with this, the overall level of genomic 5-methylcytosine is decreased in cancer cells [7, 8]. Hypomethylation of bulk cellular DNA might result in genomic and chromosomal instability [9, 10] and is perhaps suggestive of a global switch mechanism that directs changes in chromatin structure concomitant with aberrations in DNA methylation patterns. The change in DNA methylation patterns is considered to be common in most cancers [11], with significant effects on gene expression patterns, cellular growth, and selective advantage. These changes can be the result of silencing of tumor suppressor genes and alterations to associated downstream pathways [2, 12], such as repression of the p53 tumor suppressor pathway [13]. It is important to emphasize

that epigenetic mechanisms act in concert to coordinate normal gene regulatory processes and that cellular deregulation in disease involves many systems.

DNA methylation is a mediator of long-term silencing [6] and contributes to the regulatory mechanisms of tissue-specific gene expression in normal cells. The covalent addition of a methyl group to DNA can influence gene transcription [14] by varying the binding of transcription factors [15, 16] or through the recruitment of methyl-binding proteins [17, 18] and chromatin modifiers such as histone deacetylases [14, 19]. These studies provided some of the first evidence that there was interplay between different epigenetic modifications. Despite this evidence, the exact mechanism for gene silencing by DNA methylation is still debated. Foremost, while covalent histone modifications are sufficient to repress gene expression, they are considered to be a less stable and reversible process [3, 20]. Therefore, it is likely that there are additional determinants that specify the establishment of permanently silenced and hypermethylated CpG islands in cancer.

CpG islands remain unmethylated in normal cells and are frequently sites of DNase hypersensitivity on a global scale [21–23]. DNase hypersensitivity has been used as a marker for genomic regions that are free of nucleosomes, suggesting that the extent of nucleosome occupancy may be correlated with gene silencing and expression. Given these data, it has been hypothesized for many years that inactive genes exhibit a closed, compact chromatin structure in contrast to active gene promoters that are less condensed to allow for the binding of transcriptional machinery. Technological advances have confirmed and extended these data, and it is now clear that the role of nucleosome positioning in gene silencing through DNA methylation is critical for gene control.

DNA accessibility is a requirement for transcription [1, 24] and can be considered to be either a constitutive state [25, 26] or one that is generated following extensive chromatin remodeling, as demonstrated for the *PHO5* promoter in yeast [27]. Genome-wide screens in several organisms have shown that the regions upstream of many transcriptional start sites are devoid of nucleosomes [28–32], indicative of the specific gene expression patterns in these eukaryotes. In support of a model whereby nucleosomes are central to gene control, it has been shown that a nucleosome depleted yeast *PHO5* promoter is maintained through DNA replication [33]. The inheritance of a nucleosome depleted state was shown to be independent of coactivator complexes, and transcription of *PHO5* was not required to maintain the nucleosome depleted region [33]. Despite such striking associations, the significance of these findings was not directly correlated with events of gene silencing during carcinogenesis until recently [24].

Extensive analyses of the *MLH1* promoter, which is frequently hypermethylated in cancers, reveal that the formation of a nucleosome depleted region is required for gene expression [24]. The precise positioning of nucleosomes can be determined at individual promoters by using a high-resolution single-molecule assay called methyltransferase-based single-promoter analysis (M-SPA) [34]. The M-SPA assay has also been utilized to confirm the requirement for a nucleosome depleted region at the *GRP78* [35] and *BRCA1* [24] promoters. By extension, a nucleosome depleted region is likely to be characteristic of expressing genes containing a CpG

island promoter. A nucleosome is inserted immediately upstream of the transcriptional start site of an inactive *MLH1* promoter, which becomes permanently silenced by DNA methylation in cancer cell lines [24]. These data suggest that changes in nucleosome occupancy contribute to the epigenetic silencing of CpG islands during transformation (Fig. 1). The mechanisms that then ensure that CpG islands remain permanently silenced are unclear, but it is feasible that this process involves protein complexes that facilitate the addition and removal of other epigenetic marks.

3 DNA Methylation, Covalent Histone Modifications, and Histone Variants

3.1 Histone Variants

Beyond the physical positioning of nucleosomes, the composition and posttranslational modification of these core particles must also be considered. Histone variants, including H2A.Z and H3.3 (Fig. 2a), have altered amino acid sequences compared with the canonical histone proteins [36] and have been shown to have profound effects on gene expression [32, 37, 38] as well as being associated with distinct chromosomal regions [39, 40]. Importantly, H2A.Z is enriched at transcriptional start sites of both active and inactive genes [41], suggesting that H2A.Z also has roles that are independent of transcription. One such function may be to maintain genes in a poised state [42] and to prevent the permanent silencing of these loci by DNA methylation in cancer [40, 43]. Alternatively, H2A.Z may contribute to the over-expression of oncogenes or cell cycle regulators during transformation. It has recently been demonstrated that the over-expression of H2A.Z is linked to the progression of estrogen-responsive breast cancers [44]. In this study, c-MYC was shown to bind to the *H2A.Z* promoter in response to estrogen, increasing H2A.Z protein expression [44]. This observation correlated with altered proliferation properties of MCF7 cells [44]. Despite this, a conclusive mechanistic link between H2A.Z and cancer progression remains to be established.

DNA methylation and H2A.Z are mutually exclusive epigenetic marks in plants [40]. Altered DNA methylation patterns are mirrored by changes in H2A.Z localization and vice versa [40], suggesting a high level of interaction between the mechanisms underlying these two epigenetic modifications. Specifically, genomic regions that exhibit a loss of DNA methylation become enriched for H2A.Z [40], which is proposed to be a direct effect of DNA hypomethylation events rather than changes in the levels of transcription [40]. The insertion of H2A.Z into nucleosomes is reliant on the Snf-2-related CREB-binding protein activator (SRCAP) chromatin remodeling complex in humans [45, 46]. A mutation in plants of the equivalent complex, *PIE1*, results in genome-wide DNA hypermethylation in *Arabidopsis thaliana* [40]. While the distribution of DNA methylation patterns

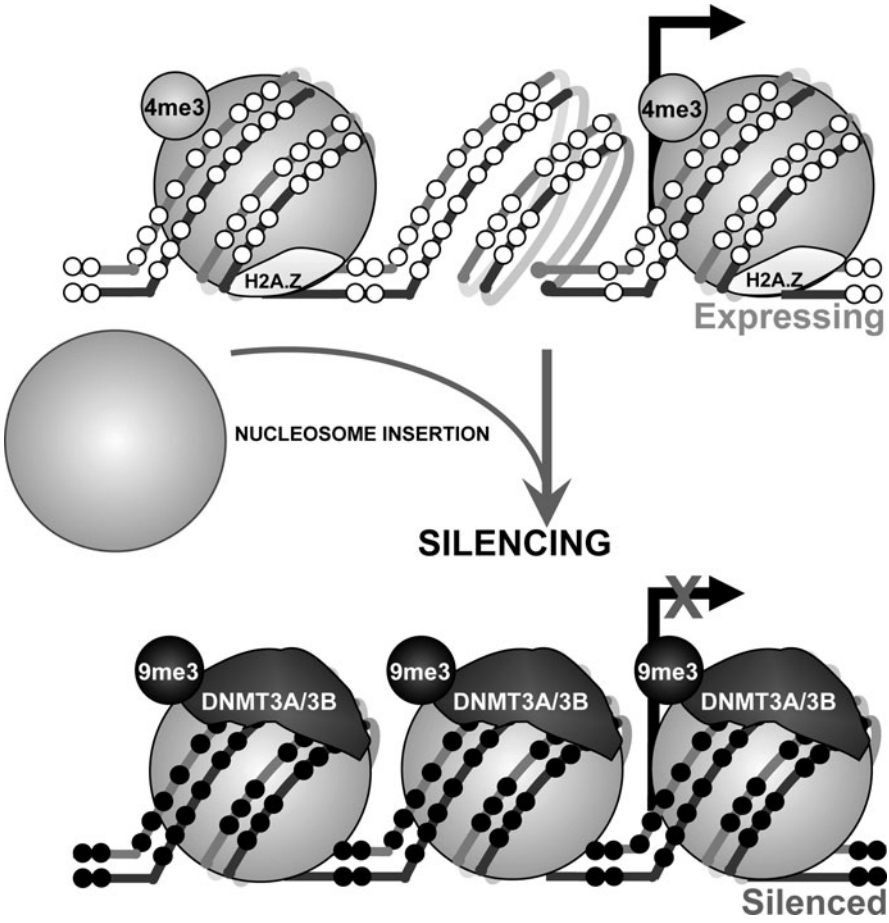


Fig. 1 Nucleosomes contribute to the epigenetic silencing of genes in concert with DNA methylation in cancer cells. In normal cells (above), active promoters are depleted of nucleosomes immediately upstream of the transcriptional start site. Nucleosome-depleted regions are flanked by nucleosomes that are enriched for active marks and are permissive for transcription, such as H3K4me3. In addition, these nucleosomes contain histone variants shown to correlate with transcription, such as H2A.Z. During the silencing process (below), a nucleosome is inserted into the nucleosome depleted region, physically interfering with the process of gene expression. DNA methylation and the acquisition of repressive histone marks, such as H3K9me3, permanently silence genes in cancer cells. DNMT3A and DNMT3B are anchored to nucleosomes associated with methylated DNA. Removal of DNA methylation leads to the eviction nucleosomes from reactivated loci after treatment of cancer cells with DNMT inhibitors such as 5-Aza-CdR (not shown; [24]) agent.

Small white circle unmethylated CpG site; *small black circle* methylated CpG site; *large circle* nucleosome; *X* silenced transcriptional start site; *4* trimethylation of histone H3 at lysine 4 (H3K4me3); *9* trimethylation of histone H3 at lysine 9 (H3K9me3); *DNMT* DNA methyltransferase; *5-Aza-CdR* 5-Aza-2′-deoxycytidine demethylating agent

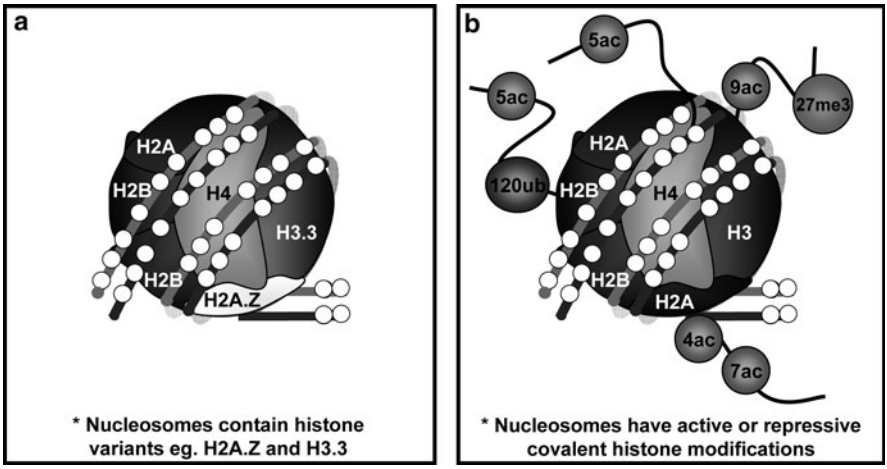


Fig. 2 Histone composition and posttranslational histone modifications correlate with transcriptional competence and nucleosome stability. (**a**) Nucleosomes consist of an octamer of core histone proteins and may contain variants such as H2A.Z and H3.3, which are often localized to the transcriptional start sites of active genes. H2A.Z is anticorrelated with DNA methylation. (**b**) Histone tail residues can be posttranslationally modified (*right*). An array of active and repressive histone modifications determine the expression status of various genes. The modifications illustrated in the figure depict some of the possible modifications that can occur on each histone tail. The exact combination of modifications present in the active, repressed, and silenced states is not known.
Small white circle unmethylated CpG site; *large circle* nucleosome; *H2A* Histone 2A; *H2B* Histone 2B; *H3.3* Histone variant 3.3; *H2A.Z* Histone variant H2A.Z; *H4* Histone 4; *4ac* acetylation of lysine 4 on histone 2A; *7ac* acetylation of lysine 7 on Histone 2A; *5ac* Acetylation of lysine 5 on Histone 2B or 4; *120ub* ubiquitylation of lysine 120 on Histone 2B; *9ac* acetylation of lysine 9 on Histone 3; *27* trimethylation of histone H3 at lysine 27 (H3K27me3)

remained similar, an increase in the total level of DNA methylation was observed [40]. Increased DNA methylation correlated with loss of H2A.Z from these loci [40], further suggesting that nucleosome composition contributes to the DNA methylation process.

3.2 Posttranslational Histone Modifications

In addition to composition of the nucleosome, histone proteins may be posttranslationally modified, acquiring active or repressive marks (Fig. 2b). A combination of up to 17 modifications, including previously characterized marks such as the trimethylation of lysine 4 on Histone 3 (H3K4me3) [47], has been proposed to correlate with highly expressed genes [48]. The H3K4me3 modification marks the 5′ regions of genes [37, 49] and is anticorrelated with DNA methylation [50]. The

trimethylation of Histone 3 at lysine 27 (H3K27me3) is mediated by the Polycomb Repressive Complex 2 (PRC2) and is associated with gene repression [51, 52]. Interestingly, H3K27me3 enrichment is also anticorrelated with DNA methylation [53, 54]. It is possible that the presence of PRC2 prevents the binding of DNA methyltransferase enzymes (DNMTs), similar to the mechanism by which DNMT 3-Like (DNMT3L) is inhibited by H3K4me3 in germ cells [55]. Perhaps H3K27me3 colocalizes with another histone modification or histone variant and prevents the aberrant hypermethylation of PRC2 target genes in normal cells.

After cellular transformation there are several characteristic marks and protein complexes that accompany DNA methylation. Accessory proteins, such as heterochromatin protein 1 (HP1) [17, 56], are proposed to contribute to the permanent silencing of DNA methylated genes since they are also associated with heterochromatic regions of the genome [57, 58]. The enrichment of H3K9me2 [59] and H3K9me3 [57, 58] is associated with genes silenced by DNA methylation in cancer. Interestingly, H3K9me2 can create a binding site for HP1 [60, 61] and is removed from promoters that have been demethylated and reactivated in cancer cells [59].

Taken together, these data indicate that DNA methylation patterning can be dependent on nucleosome placement and composition, particularly near the 5′-regions of genes that may encompass CpG islands. DNA methylation can result in, or be the result of, the exclusion of histone variant containing nucleosomes near transcriptional start sites. Therefore, a lack of epigenetic marks that specify transcriptional competence, or alternatively a permissive but repressed state, may result in a feedback mechanism that progressively silences CpG island promoters by DNA methylation in cancer.

4 Epigenetic Switching in the Cancer Genome

CpG island promoters often become hypermethylated during cancer progression, while the remainder of the genome exhibits a reduction in DNA methylation [7, 8]. Genes containing CpG islands that were PRC2 targets in embryonic stem cells appear to be predisposed for hypermethylation [62–66]. These genes are critical for development and are normally repressed by PRC2 following differentiation from pluripotency.

DNA methylation and PRC2 occupancy at gene promoters are typically considered to be mutually exclusive events [53, 54]. This may be explained partially from the results of embryonic stem cell studies, which suggest that PRC2 and DNA methylation have the potential to regulate different subsets of genes [67]. However, it is also evident that PRC2 occupied gene promoters are preferentially methylated during the transformation of somatic cells [53, 62, 63, 65, 66], suggesting that there are genes that may be regulated by both PRC2 and DNA methylation.

Genome-wide comparisons of normal prostate epithelium with a prostate cancer cell line revealed that there are three distinct subsets of genes that are silenced or

repressed by epigenetic processes as a consequence of oncogenesis [53] (Fig. 3). First, there are genes that are expressed in normal tissue, but are silenced by DNA methylation in cancer due to a process termed 5-methylcytosine reprogramming. A second group of genes are expressed in normal prostate epithelium and become repressed by PRC2 in cancer [53, 68]. This type of event is termed PRC reprogramming. Finally, there are those developmentally important genes that are repressed by PRC2 in normal cells, but are hypermethylated in cancer [53]. This type of change is termed "epigenetic switching," but does not result in gene expression changes. Epigenetic switching may reduce the capacity of PRC2 regulated promoters to respond to signals that may otherwise reactivate these genes.

The mechanisms responsible for epigenetic switching are not yet clear. One hypothesis suggests that DNMTs are actively recruited to PRC2 occupied gene promoters during the carcinogenic process [69, 70]. The concept of active DNMT recruitment contradicts evidence that DNA methylation and PRC2 occupancy are typically mutually exclusive epigenetic marks [53, 54]. A model whereby DNMTs

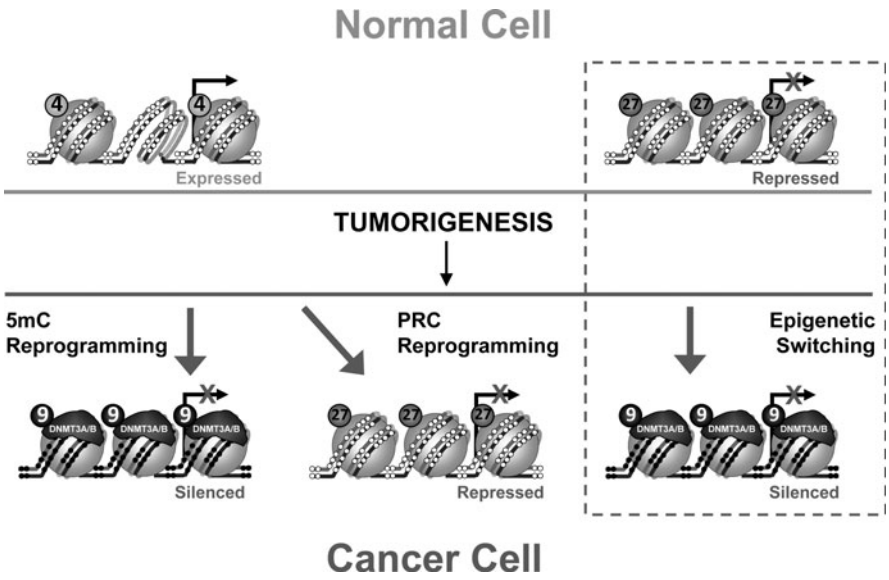


Fig. 3 A model of epigenetic switching of PRC2 target genes in cancer cells. In normal cells, genes contained within CpG islands are usually unmethylated and can be active (*left*) or repressed by PRC2, which mediates the addition of a trimethyl group to lysine 27 of histone 3 through EZH2 (27me3; *right*). Following cellular transformation, an active gene may undergo 5-methylcytosine reprogramming or PRC2 reprogramming. In addition, genes regulated by PRC2 in embryonic stem cells appear to be predisposed to become DNA hypermethylated in cancer and can undergo epigenetic switching upon transformation (*dashed box*).
Small white circle unmethylated CpG site; *small black circle* methylated CpG site; *large circle* nucleosome; *4* trimethylation of histone H3 at lysine 4 (H3K4me3); *9* trimethylation of histone H3 at lysine 9 (H3K9me3); *27* trimethylation of histone H3 at lysine 27 (H3K27me3); *X* silenced transcriptional start site; *DNMT* DNA methyltransferase; *EZH2* Enhancer of Zeste 2

are actively recruited to gene regulatory regions is therefore insufficient to explain why the majority of embryonic PCR2 targets become hypermethylated in cancer.

Alternatively, the process of epigenetic switching could encompass a series of passive events that gradually result in the hypermethylation of CpG islands in cancer. Could a change in histone composition, such as the removal of variants associated with transcriptional activation from gene promoters, mediate some of the alterations that we observe in the cancer epigenome? Ultimately, DNA methylation may be a cause or consequence of changes in nucleosome occupancy, both of which occur during transcriptional silencing and the "locking" of the genome by acquisition of specific histone methylations (such as H3K9me2). Several unique events are likely to be intricately involved in epigenetic switching, resulting in reduced epigenetic plasticity and the silencing of tumor suppressor genes in cancer. The dissection of such events may be possible using newly available high-resolution genome-wide DNA methylation assays [71].

5 Mechanisms of DNA Methylation Inheritance

DNA methylation is a mitotically heritable epigenetic modification. The classical model for the maintenance of DNA methylation patterns has recently been reviewed [72]. A revised model has now been proposed, which encompasses new data and addresses unexplained and previously inconsistent observations regarding DNA methylation inheritance [72]. There are two key characteristics of DNA methylation that partially explain the mechanisms through which it could be inherited. First, patterns of DNA methylation exist, and second, these patterns are distinguishable in somatic cell types [73, 74]. These observations, together with accumulating data, gradually led to the conclusion that mammalian cells had two classes of enzymes, de novo and maintenance methyltransferases, that use DNA as a substrate for methylation.

The DNMT family has now been well characterized. DNMT1, DNMT3A, and DNMT3B are the only enzymes to be shown to have indispensable roles in DNA methylation in somatic cells [75, 76]. DNMT3A/3B are required for de novo methyltransferase activity [76], both having an equal preference for hemi-methylated and unmethylated DNA *in vitro* [77]. In normal somatic cells, the expression of DNMT3A/3B is reduced compared with embryonic stem cells. Complete abrogation of DNMT3A results in viable litters; however, these mice die approximately 4 weeks after birth [76]. DNMT3B$^{-/-}$ mice are not viable and do not survive embryogenesis [76]. DNMT3L (DNMT3-Like) is a regulatory protein that acts to enhance the activity of DNMT3A/3B. DNMT3L is expressed primarily in gametogenesis to establish parental origin methylation patterns [78, 79]. DNMT3L mice are viable [80], suggestive of its complementary role in establishing DNA methylation patterns. In contrast to DNMT3A/3B, DNMT1 is the "maintenance methyltransferase," though it also exhibits de novo methyltransferase activity [81]. DNMT1 is preferentially targeted to hemi-methylated DNA [82] and determines

the steady-state level of methylation [83]. DNMT1 is localized to the replication fork during DNA synthesis through interactions with PCNA, a p21-dependent nuclear antigen involved in DNA replication and repair [84] and UHRF1, a DNA repair protein that also has a role in cell cycle progression [85, 86]. Once tethered to the replication fork, the patterns of methylation that were established by DNMT3A/3B are faithfully copied to the newly synthesized daughter strand of DNA by DNMT1 [87]. Like the de novo methyltransferases, abolishment of DNMT1 results in embryonic lethality [88]. It is interesting to note that DNMT1 is also required for cancer cell survival [89, 90], suggesting that this protein has essential roles in both early development and in cancer.

6 Nucleosomes and DNA Methylation Patterns in Cancer

What is the significance of the nucleosome in maintaining accurate DNA methylation patterns and how are these altered in cancer? The classical model of DNA methylation events involves DNMT3A/3B as de novo methyltransferases, together with DNMT1 as the maintenance enzyme. However, several experimental observations do not completely fit with this model of DNA methylation maintainance [72]. Of importance to our understanding of how DNA methylation patterns are inherited is the structure of chromatin, which was largely ignored in the classical model of DNA methylation maintenance [72]. As discussed previously, nucleosomes are absent from the transcriptional start sites of active genes whose promoters are located in CpG islands (Fig. 1). The physical placement of the nucleosome is sufficient to impede transcriptional initiation and correlates strongly with DNA methylation of CpG islands [24] (Fig. 1). Several lines of evidence suggest that chromatin structures contribute to DNA methylation establishment and maintenance. DNMT3 proteins have been shown to associate with heterochromatin [91, 92], which has a higher density of nucleosomes than euchromatic regions. Indeed, DNMT3A has been shown to generate a structure with DNMT3L that could physically encompass a nucleosome [55], and DNMT3A/3B are known to interact strongly with nucleosomes within methylated CpG islands and repeat sequences [93]. However, it is excluded from this interaction by the active H3K4me3 mark. These data strongly suggest that the physical anchoring of DNMT3A/3B to nucleosomes is necessary for the maintenance of DNA methylation patterns (Fig. 1).

It has been suggested that a component of PRC2 recruits DNMT3A to gene regulatory regions, yet this occurs without subsequent de novo methylation [69], indicating that this process may not be a widespread mechanism for the mistargeting of DNMTs and the resultant methylation of CpG island loci in cancer.

The expression of variant DNMT3 isoforms is increased during oncogenesis, resulting in more random patterns of DNA methylation. Altered expression of wildtype DNMT proteins, together with delta DNMT3 isoforms (ΔDNMT3) and catalytically inactive forms of DNMT3, is expressed in a tissue-specific manner [94] and is linked to several types of cancer [48, 95–97]). The ΔDNMT3 isoforms

are a family of seven transcriptional variants that are generated by alternative splicing [98]. Truncated and inactive DNMT3 isoforms can compete with wildtype proteins to alter DNA methylation patterns [99] in a promoter-dependent manner [100].

Increasing the expression of altered DNMT3 isoforms is one way that aberrant DNA methylation patterns could be established in a cancer cell. Alternatively, both wildtype and ΔDNMT3 isoforms could be mistargeted to chromatin during cellular transformation. Chromatin structure is integral to the mechanisms underlying the process of DNA methylation, and the anchoring of wildtype DNMT3A/3B proteins to nucleosomes is dependent on the N-terminal regulatory region of the proteins [93]. This might ensure that unbound DNMT3A/3B does not cause aberrant hypermethylation. Some ΔDNMT3 variants are not anchored to nucleosomes because they lack the N-terminal domain [93]. Therefore, it is possible that the unbound mutant DNMT3 proteins contribute to the establishment of aberrant DNA methylation pattern at CpG islands, irrespective of nucleosome occupancy. Another possibility is that the insertion of the nucleosome into the previously depleted region upstream of transcriptional start sites (Fig. 1) redirects wildtype DNMT3A/3B to CpG islands during oncogenesis. These two possibilities are not likely to be mutually exclusive.

7 Epigenetic Regulation of miRNAs

Small, noncoding microRNAs (miRNAs) have the potential to regulate gene expression by modulating mRNA stability or translation [101–103]. The activity of miRNAs occurs in a cell type-dependent manner [104], contributing to crucial cellular processes such as proliferation and differentiation [105, 106], and it is not surprising then that miRNA expression is altered in cancer cells [107, 108]. The genome-wide profiling of miRNAs has now been conducted for several cancers [109], revealing that most of these ~22 nt noncoding RNA molecules function as tumor suppressors in somatic cells. It has been noted that miRNAs may also be used as biomarkers for various types of cancers [110], predicting outcome and treatment options for patients. This has been correlated with variable miRNA expression within cancer subtypes at different stages of the disease [109].

miRNAs have the potential to act as epigenetic modifiers. A mechanism for the widespread effects of miRNAs in carcinogenesis may therefore be partly explained by their abilities to regulate the translation of DNMTs through epigenetic modification. For example, a reduction in DNMT3A/3B mRNA is observed concomitant with decreased expression of the miRNA-290 cluster [111]. Similarly, overexpression of the miRNA-29 family, which occurs frequently in lung cancer, correlates with reduced transcription of DNMT3A/3B [112]. Aberrant miRNA expression or activity may therefore specify a mechanism by which DNMTs are deregulated in cancer.

In a similar manner, miRNAs have been shown to alter other epigenetic pathways that are disrupted in cancer, such as PRC2 [113, 114]. As mentioned previously, many genes that were PRC2 targets in embryonic stem cells become hypermethylated in cancer due to epigenetic switching [53]. A second group of genes are not de novo target genes of PRC2, yet became repressed in prostate cancer cell lines due to PRC2 reprogramming [53]. Indeed, Enhancer of Zeste 2 (EZH2), the catalytic component of PRC2, is overexpressed in prostate [115], breast [116], and bladder cancers [117]. The overexpression of EZH2 has now been linked to reduced expression of miRNA-101 in cancer cells [113, 114]. Exemplifying the role of miRNAs in regulating epigenetic pathways, it was demonstrated that the restoration of miRNA-101 expression resulted in the reexpression of genes that had been repressed by EZH2 in cancer cells [113]. Further, reduced H3K27me3 levels are observed at known target gene promoters, *FAM84* and *DDIT4* [113] as well as *RUNX3* and *WNT1* [114] after EZH2 knockdown in cancer cells. These data indicate that complex networks are involved in epigenetic switching, 5-methylcytosine reprogramming, and PRC reprogramming. Moreover, it is clear from these results that the deregulation of key epigenetic pathways can be mediated by miRNAs.

In addition to acting as epigenetic modifiers, miRNAs themselves can be epigenetically regulated [118]. miRNA-127, located in a CpG island, is silenced by DNA methylation in tumors [118]. A widespread reduction in histone acetylation of miRNA-127 was also detected in these tumors [118], suggesting that both DNA methylation and posttranslational histone modifications play a role in epigenetically regulating miRNAs. The epigenetic deregulation of miRNAs in additional cancers has now been described, such as the hypermethylation of miRNA-34b/c in colorectal cancer [119]. Thus, the hypermethylation of CpG islands can contribute to carcinogenesis not only by silencing tumor suppressor genes, but also by silencing miRNAs.

8 DNA Methylation at CpG Poor Regions

The focus in the field thus far has been on hypermethylation of CpG island promoters in cancers. As stated previously, CpG islands constitute approximately 2% of the genome, while the remainder is CpG poor. Some promoters with intermediate CpG dinucleotide content exhibit tissue-specific patterns of expression, which can be associated with the DNA methylation status of these regulatory regions [120]. There are also examples whereby a single methylation site can influence gene activity at non-CpG islands. For example, repression of the CpG poor Interferon-β promoter has been correlated with the methylation of one CpG site, due to inhibitory effects on transcription factor binding [16].

CpG poor regions are generally methylated in normal tissue [121]. During the initiation and progression of cancer, the DNA methylation landscape is altered such that the genome-wide hypomethylation of CpG poor regions accompanies the hypermethylation events at CpG islands [2]. Such profound changes in DNA

methylation have now been linked to increased chromosomal rearrangements [122], leading to genomic instability and changes in nucleosome positioning throughout bulk chromatin. These events have been correlated with an increase in the incidence of tumors [123] and constitute potential gain of function epigenetic effects by increasing the access of transcription factors and coactivator complexes to DNA [124]. Of particular interest is the resultant reactivation of proto-oncogenes, such as c-Myc [125]. Activation of signaling pathways is also commonly reported, including Wnt/β-catenin [126], several components of which are linked to many types of cancers [127]. Effects on cellular behavior due to the activation of oncogenes and their associated signaling pathways include disruptions to the cell cycle, DNA repair mechanisms, and altered gene expression profiles, to name a few.

The genome-wide hypomethylation of a cancer cell has additional implications. Since retrotransposons and repetitive elements constitute much of the genome, demethylation of these can result in an increased frequency of chromosomal rearrangements and insertional mutagenesis [128], in addition to chromatin instability. Sense and antisense transcripts may also occur, which directly interfere with transcription of proximal genes [124].

Similar to hypermethylation of tumor suppressor genes, the hypomethylation of CpG poor gene promoters and CpG rich repetitive sequences, such as Alu or LINE-1 elements [129], can be correlated with patient prognosis. For example, the long-term survival of patients with ovarian cancer can be linked to hypomethylation of the MAL gene [130]. Interestingly, this is also indicative of resistance to current standard of care treatment with platinum reagents [130]. It could be speculated from these data that the hypomethylation of genes may determine cellular responsiveness to drugs, including epigenetic therapies.

9 Epigenetic Therapy

DNA methylation is integral to the epigenetic silencing process, and aberrations in this process occur at high frequency across a range of cancer types. DNA methylation is an attractive target from a therapeutic standpoint, particularly because CpG island promoters are seldom regulated by DNA methylation in normal cells. Several potent nucleoside analogs have been derived and have been the subject of thorough investigation, including two that have now been approved for clinical use by the Food and Drug Administration, 5-Azacytidine (5-Aza-CR) and 5-Aza-2′-deoxycytidine (5-Aza-CdR) [131]. Of the remainder, 5-Fluro-2′-deoxycytidine (FCdR) is currently undergoing clinical trials [132]. Zebularine is a highly stable inhibitor of DNA methylation [133] and is effective when administered orally. Treatment can result in the reactivation of *p16* in bladder cancer xenografts in nude mice [134] and long-term enteral administration of zebularine to cancer prone mice prevented intestinal tumors, having minimal side effects [133].

Interference of the DNA methylation process in cancer cells can promote gene reactivation, together with cellular differentiation [135], similar to observations

first made using immortalized mouse embryonic stem cells [136]. While cancer cells respond to treatment with DNMT inhibitors such as 5-Aza-CdR, it has been demonstrated that normal cells are more resistant to the effects of these compounds [134, 137, 138]. This observation is critical for the therapeutic potential of these agents because it suggests that cancer cells can be selectively targeted.

Clinical trials have demonstrated that 5-Aza-CR and 5-Aza-CdR are effective in the treatment of myelodysplastic syndrome and myelogenous leukemias. However, the administered dosage is critical and must be within a low, narrow range in order for epigenetic effects to be apparent because the drugs become cytotoxic at high concentrations [136, 139]. It was this characteristic that prompted low-dose clinical trials in elderly patients with myelodysplastic syndrome [140]. A low-dose regimen of 5-Aza-CR has approximately 50% efficacy, increasing long-term survival, remission rates, and a general improvement in quality of life [141]. These studies were the first of many demonstrating that an epigenetic agent could delay disease progression [142], although it has been noted that 5-Azanucleosides are particularly unstable and must be given as part of a chronic treatment program to avoid reversal of DNA methylation patterns.

In addition to their instability, nucleoside analogs, such as 5-Aza-CdR, may not be ideal candidates for epigenetic therapy since they become incorporated into DNA and are thus cytotoxic. It has been speculated that the use of nucleoside analogs may result in the formation of secondary tumors that arise due to chromosomal instability after drastic changes in DNA methylation [123, 143]. These hypotheses were drawn from studies in genetically manipulated mice that developed a higher number of tumors following genome-wide DNA hypomethylation [123, 143]. However, the mice in question were born with genetic aberrations in DNMT1, exhibiting phenotypic changes at birth [123, 143]. Moreover, the studies mentioned here did not specifically address problems arising from use of DNMT inhibitors in these animal models [123, 143]. To address this, an investigation into the cytotoxic actions of nucleoside inhibitors was undertaken in leukemic patients [144]. Secondary tumors were not detected in any patient undertaking a nucleoside analog treatment program, though it was not ruled out that they could possibly occur. Alternative DNMT inhibitors are also being investigated and developed [145]. SGI-1027, RG108, and MG98 are small molecule inhibitors that are proposed to decrease DNA methylation by impairing binding sites for cofactors, inhibiting the catalytic activity of DNMTs or acting much like miRNA sequences that inhibit mRNA translation [146, 147].

For the most part, it is assumed that DNA methylation itself specifically results in gene silencing. In agreement with this tenet, many genes are reexpressed in cancer cell lines after treatment with 5-Aza-CdR, including *MLH1* in the RKO colorectal carcinoma cell line [24]. However, it must be emphasized that the reactivation of *MLH1* is dependent on the eviction of a nucleosome from the promoter. Using the high-resolution MSPA technique, it was noted that approximately 50% of individual DNA molecules became demethylated after 5-Aza-CdR treatment [24]. Of the demethylated molecules, 50% exhibited a nucleosome depleted region immediately upstream of the transcriptional start site [24]. The

remaining promoters, still occupied by a nucleosome in the promoter region, are likely to represent the molecules that are demethylated but transcriptionally repressed [24]. These data illustrate the importance of the nucleosome in the process of transcriptional reactivation after demethylation and suggest that the nucleosome may serve as a mechanism for the epigenetic inheritance of DNA methylation patterns in cancer.

10 Conclusions and Future Directions

The importance of DNA methylation in cancer establishment is clear, though the specific order of events that result in transformation remains to be conclusively established. The advent of sophisticated epigenetic studies has allowed us to integrate our knowledge and study the interplay between histone modifications and nucleosome positioning together with DNA methylation. All of these processes are integral to normal cellular stability, and the initiation of cancer is likely to involve disruptions to each simultaneously. We are now in a position to fully understand the roles of aberrant DNA methylation in cancer establishment and progression. There are exciting new links between stem cell behavior and cancer cells, and perhaps thorough investigation of developmental epigenetics will direct us to the key events that are deregulated in carcinogenesis. Increased information about these mechanisms will allow us to yield novel and specific epigenetic therapies that will ensure lower toxicity and better patient outcomes.

Acknowledgments We thank Dr. Theresa Kelly, Dr. Fernanda Molognoni, Daniel de Carvalho, and Shikhar Sharma for reviewing the manuscript and their helpful suggestions.

References

1. Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. Cell 128:707–719
2. Jones PA, Baylin SB (2002) The fundamental role of epigenetic events in cancer. Nat Rev Genet 3:415–428
3. Miranda TB, Jones PA (2007) DNA methylation: the nuts and bolts of repression. J Cell Physiol 213:384–390
4. Bestor TH, Gundersen G, Kolsto AB, Prydz H (1992) CpG islands in mammalian gene promoters are inherently resistant to de novo methylation. Genet Anal Tech Appl 9:48–53
5. Nguyen C, Liang G, Nguyen TT, Tsao-Wei D, Groshen S, Lubbert M, Zhou JH, Benedict WF, Jones PA (2001) Susceptibility of nonpromoter CpG islands to de novo methylation in normal and neoplastic cells. J Natl Cancer Inst 93:1465–1472
6. Bird A (2002) DNA methylation patterns and epigenetic memory. Genes Dev 16:6–21
7. Feinberg AP, Vogelstein B (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature 301:89–92

8. Riggs AD, Jones PA (1983) 5-methylcytosine, gene regulation, and cancer. Adv Cancer Res 40:1–30
9. Almeida A, Kokalj-Vokac N, Lefrancois D, Viegas-Pequignot E, Jeanpierre M, Dutrillaux B, Malfoy B (1993) Hypomethylation of classical satellite DNA and chromosome instability in lymphoblastoid cell lines. Hum Genet 91:538–546
10. 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
11. Jones PA, Baylin SB (2007) The epigenomics of cancer. Cell 128:683–692
12. Herman JG, Baylin SB (2003) Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med 349:2042–2054
13. Harris SL, Levine AJ (2005) The p53 pathway: positive and negative feedback loops. Oncogene 24:2899–2908
14. Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet 19:187–191
15. Jaenisch R, Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 33(Suppl):245–254
16. Jones B, Chen J (2006) Inhibition of IFN-gamma transcription by site-specific methylation during T helper cell development. EMBO J 25:2443–2452
17. Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T (2003) The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. J Biol Chem 278:4035–4040
18. Nguyen CT, Gonzales FA, Jones PA (2001) Altered chromatin structure associated with methylation-induced gene silencing in cancer cells: correlation of accessibility, methylation, MeCP2 binding and acetylation. Nucleic Acids Res 29:4598–4606
19. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393:386–389
20. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 119:941–953
21. Mucha M, Lisowska K, Goc A, Filipski J (2000) Nuclease-hypersensitive chromatin formed by a CpG island in human DNA cloned as an artificial chromosome in yeast. J Biol Chem 275:1275–1278
22. Pfeifer GP, Riggs AD (1991) Chromatin differences between active and inactive X chromosomes revealed by genomic footprinting of permeabilized cells using DNase I and ligation-mediated PCR. Genes Dev 5:1102–1113
23. Tazi J, Bird A (1990) Alternative chromatin structure at CpG islands. Cell 60:909–920
24. Lin JC, Jeong S, Liang G, Takai D, Fatemi M, Tsai YC, Egger G, Gal-Yam EN, Jones PA (2007) Role of nucleosomal occupancy in the epigenetic silencing of the MLH1 CpG island. Cancer Cell 12:432–444
25. Sabo PJ, Humbert R, Hawrylycz M, Wallace JC, Dorschner MO, McArthur M, Stamatoyannopoulos JA (2004) Genome-wide identification of DNaseI hypersensitive sites using active chromatin sequence libraries. Proc Natl Acad Sci USA 101:4537–4542
26. Sasaki H, Jones PA, Chaillet JR, Ferguson-Smith AC, Barton SC, Reik W, Surani MA (1992) Parental imprinting: potentially active chromatin of the repressed maternal allele of the mouse insulin-like growth factor II (Igf2) gene. Genes Dev 6:1843–1856
27. Boeger H, Griesenbeck J, Strattan JS, Kornberg RD (2004) Removal of promoter nucleosomes by disassembly rather than sliding *in vivo*. Mol Cell 14:667–673
28. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA et al (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39:311–318
29. Lee CK, Shibata Y, Rao B, Strahl BD, Lieb JD (2004) Evidence for nucleosome depletion at active regulatory regions genome-wide. Nat Genet 36:900–905

30. Mito Y, Henikoff JG, Henikoff S (2005) Genome-scale profiling of histone H3.3 replacement patterns. Nat Genet 37:1090–1097
31. Ozsolak F, Song JS, Liu XS, Fisher DE (2007) High-throughput mapping of the chromatin structure of human promoters. Nat Biotechnol 25:244–248
32. Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K (2008) Dynamic regulation of nucleosome positioning in the human genome. Cell 132:887–898
33. Ohsawa R, Adkins M, Tyler JK (2009) Epigenetic inheritance of an inducibly nucleosome-depleted promoter and its associated transcriptional state in the apparent absence of transcriptional activators. Epigenetics Chromatin 2:11
34. Fatemi M, Pao MM, Jeong S, Gal-Yam EN, Egger G, Weisenberger DJ, Jones PA (2005) Footprinting of mammalian promoters: use of a CpG DNA methyltransferase revealing nucleosome positions at a single molecule level. Nucleic Acids Res 33:e176
35. Gal-Yam EN, Jeong S, Tanay A, Egger G, Lee AS, Jones PA (2006) Constitutive nucleosome depletion and ordered factor assembly at the GRP78 promoter revealed by single molecule footprinting. PLoS Genet 2:e160
36. Malik HS, Henikoff S (2003) Phylogenomics of the nucleosome. Nat Struct Biol 10:882–891
37. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. Cell 129:823–837
38. Deal RB, Topp CN, McKinney EC, Meagher RB (2007) Repression of flowering in Arabidopsis requires activation of FLOWERING LOCUS C expression by the histone variant H2A.Z. Plant Cell 19:74–83
39. Sarcinella E, Zuzarte PC, Lau PN, Draker R, Cheung P (2007) Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. Mol Cell Biol 27:6457–6468
40. Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S (2008) Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. Nature 456:125–129
41. Raisner RM, Hartley PD, Meneghini MD, Bao MZ, Liu CL, Schreiber SL, Rando OJ, Madhani HD (2005) Histone variant H2A.Z marks the 5′ ends of both active and inactive genes in euchromatin. Cell 123:233–248
42. Brickner DG, Cajigas I, Fondufe-Mittendorf Y, Ahmed S, Lee PC, Widom J, Brickner JH (2007) H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. PLoS Biol 5:e81
43. Meneghini MD, Wu M, Madhani HD (2003) Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. Cell 112:725–736
44. Hua S, Kallen CB, Dhar R, Baquero MT, Mason CE, Russell BA, Shah PK, Liu J, Khramtsov A, Tretiakova MS et al (2008) Genomic analysis of estrogen cascade reveals histone variant H2A.Z associated with breast cancer progression. Mol Syst Biol 4:188
45. Ruhl DD, Jin J, Cai Y, Swanson S, Florens L, Washburn MP, Conaway RC, Conaway JW, Chrivia JC (2006) Purification of a human SRCAP complex that remodels chromatin by incorporating the histone variant H2A.Z into nucleosomes. Biochemistry 45:5671–5677
46. Wong MM, Cox LK, Chrivia JC (2007) The chromatin remodeling protein, SRCAP, is critical for deposition of the histone variant H2A.Z at promoters. J Biol Chem 282: 26132–26139
47. Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T (2002) Active genes are tri-methylated at K4 of histone H3. Nature 419:407–411
48. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ et al (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 40:897–903
49. Liang G, Lin JC, Wei V, Yoo C, Cheng JC, Nguyen CT, Weisenberger DJ, Egger G, Takai D, Gonzales FA et al (2004) Distinct localization of histone H3 acetylation and H3-K4

methylation to the transcription start sites in the human genome. Proc Natl Acad Sci USA 101:7357–7362
50. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 39:457–466
51. Kirmizis A, Bartley SM, Kuzmichev A, Margueron R, Reinberg D, Green R, Farnham PJ (2004) Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. Genes Dev 18:1592–1605
52. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev 16:2893–2905
53. Gal-Yam EN, Egger G, Iniguez L, Holster H, Einarsson S, Zhang X, Lin JC, Liang G, Jones PA, Tanay A (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
54. Lindroth AM, Park YJ, McLean CM, Dokshin GA, Persson JM, Herman H, Pasini D, Miro X, Donohoe ME, Lee JT et al (2008) Antagonism between DNA and H3K27 methylation at the imprinted Rasgrf1 locus. PLoS Genet 4:e1000145
55. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD et al (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 448:714–717
56. Smallwood A, Esteve PO, Pradhan S, Carey M (2007) Functional cooperation between HP1 and DNMT1 mediates gene silencing. Genes Dev 21:1169–1178
57. Kouzarides T (2002) Histone methylation in transcriptional control. Curr Opin Genet Dev 12:198–209
58. Lachner M, O'Sullivan RJ, Jenuwein T (2003) An epigenetic road map for histone lysine methylation. J Cell Sci 116:2117–2124
59. Nguyen CT, Weisenberger DJ, Velicescu M, Gonzales FA, Lin JC, Liang G, Jones PA (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
60. Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410:120–124
61. Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410:116–120
62. Hahn MA, Hahn T, Lee DH, Esworthy RS, Kim BW, Riggs AD, Chu FF, Pfeifer GP (2008) Methylation of polycomb target genes in intestinal cancer is mediated by inflammation. Cancer Res 68:10280–10289
63. Jin B, Yao B, Li JL, Fields CR, Delmas AL, Liu C, Robertson KD (2009) DNMT1 and DNMT3B modulate distinct polycomb-mediated histone modifications in colon cancer. Cancer Res 69:7412–7421
64. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, Mohammad HP, Chen W, Daniel VC, Yu W 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
65. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, Eden E, Yakhini Z, Ben-Shushan E, Reubinoff BE 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
66. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, Weisenberger DJ, Campan M, Young J, Jacobs I et al (2007) Epigenetic stem cell signature in cancer. Nat Genet 39:157–158
67. Fouse SD, Shen Y, Pellegrini M, Cole S, Meissner A, Van Neste L, Jaenisch R, Fan G (2008) Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. Cell Stem Cell 2:160–169

68. Kondo Y, Shen L, Cheng AS, Ahmed S, Boumber Y, Charo C, Yamochi T, Urano T, Furukawa K, Kwabi-Addo B et al (2008) Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. Nat Genet 40:741–750
69. Rush M, Appanah R, Lee S, Lam LL, Goyal P, Lorincz MC (2009) Targeting of EZH2 to a defined genomic site is sufficient for recruitment of Dnmt3a but not de novo DNA methylation. Epigenetics 4:404–414
70. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM et al (2006) The Polycomb group protein EZH2 directly controls DNA methylation. Nature 439:871–874
71. Hodges E, Smith AD, Kendall J, Xuan Z, Ravi K, Rooks M, Zhang MQ, Ye K, Bhattacharjee A, Brizuela L et al (2009) High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. Genome Res 19:1593–1605
72. Jones PA, Liang G (2009) Rethinking how DNA methylation patterns are maintained. Nat Rev Genet 10:805–811
73. Holliday R, Pugh JE (1975) DNA modification mechanisms and gene activity during development. Science 187:226–232
74. Riggs AD (1975) X inactivation, differentiation, and DNA methylation. Cytogenet Cell Genet 14:9–25
75. Lyko F, Ramsahoye BH, Kashevsky H, Tudor M, Mastrangelo MA, Orr-Weaver TL, Jaenisch R (1999) Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in Drosophila. Nat Genet 23:363–366
76. Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99:247–257
77. Okano M, Xie S, Li E (1998) Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet 19:219–220
78. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH (2001) Dnmt3L and the establishment of maternal genomic imprints. Science 294:2536–2539
79. 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
80. Webster KE, O'Bryan MK, Fletcher S, Crewther PE, Aapola U, Craig J, Harrison DK, Aung H, Phutikanit N, Lyle R et al (2005) Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. Proc Natl Acad Sci USA 102:4068–4073
81. Pradhan S, Bacolla A, Wells RD, Roberts RJ (1999) Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. J Biol Chem 274:33002–33010
82. Bestor TH, Ingram VM (1983) Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. Proc Natl Acad Sci USA 80:5559–5563
83. Biniszkiewicz D, Gribnau J, Ramsahoye B, Gaudet F, Eggan K, Humpherys D, Mastrangelo MA, Jun Z, Walter J, Jaenisch R (2002) Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. Mol Cell Biol 22:2124–2135
84. Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF (1997) Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science 277:1996–2000
85. Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science 317:1760–1764
86. Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K et al (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 450:908–912
87. Bestor T, Laudano A, Mattaliano R, Ingram V (1988) Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. J Mol Biol 203:971–983

88. Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69:915–926
89. Chen T, Hevi S, Gay F, Tsujimoto N, He T, Zhang B, Ueda Y, Li E (2007) Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. Nat Genet 39:391–396
90. Egger G, Jeong S, Escobar SG, Cortez CC, Li TW, Saito Y, Yoo CB, Jones PA, Liang G (2006) Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. Proc Natl Acad Sci USA 103:14080–14085
91. Bachman KE, Rountree MR, Baylin SB (2001) Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. J Biol Chem 276:32282–32287
92. Chen T, Ueda Y, Dodge JE, Wang Z, Li E (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol Cell Biol 23:5594–5605
93. Jeong S, Liang G, Sharma S, Lin JC, Choi SH, Han H, Yoo CB, Egger G, Yang AS, Jones PA (2009) Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. Mol Cell Biol 29:5366–5376
94. Robertson KD, Uzvolgyi E, Liang G, Talmadge C, Sumegi J, Gonzales FA, Jones PA (1999) The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. Nucleic Acids Res 27:2291–2298
95. Liu J, Liang X, Zhu J, Wei L, Hou Y, Chen DY, Sun QY (2008) Aberrant DNA methylation in 5′ regions of DNA methyltransferase genes in aborted bovine clones. J Genet Genomics 35:559–568
96. Ostler KR, Davis EM, Payne SL, Gosalia BB, Exposito-Cespedes J, Le Beau MM, Godley LA (2007) Cancer cells express aberrant DNMT3B transcripts encoding truncated proteins. Oncogene 26:5553–5563
97. Saito Y, Kanai Y, Sakamoto M, Saito H, Ishii H, Hirohashi S (2002) Overexpression of a splice variant of DNA methyltransferase 3b, DNMT3b4, associated with DNA hypomethylation on pericentromeric satellite regions during human hepatocarcinogenesis. Proc Natl Acad Sci USA 99:10060–10065
98. Wang L, Wang J, Sun S, Rodriguez M, Yue P, Jang SJ, Mao L (2006) A novel DNMT3B subfamily, DeltaDNMT3B, is the predominant form of DNMT3B in non-small cell lung cancer. Int J Oncol 29:201–207
99. Wang J, Walsh G, Liu DD, Lee JJ, Mao L (2006) Expression of Delta DNMT3B variants and its association with promoter methylation of p16 and RASSF1A in primary non-small cell lung cancer. Cancer Res 66:8361–8366
100. Wang J, Bhutani M, Pathak AK, Lang W, Ren H, Jelinek J, He R, Shen L, Issa JP, Mao L (2007) Delta DNMT3B variants regulate DNA methylation in a promoter-specific manner. Cancer Res 67:10647–10652
101. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136: 215–233
102. Filipowicz W, Bhattacharyya SN, Sonenberg N (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet 9:102–114
103. Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev 20:515–524
104. Farh KK, Grimson A, Jan C, Lewis BP, Johnston WK, Lim LP, Burge CB, Bartel DP (2005) The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. Science 310:1817–1821
105. Chen CZ, Li L, Lodish HF, Bartel DP (2004) MicroRNAs modulate hematopoietic lineage differentiation. Science 303:83–86

106. Chen JF, Mandel EM, Thomson JM, Wu Q, Callis TE, Hammond SM, Conlon FL, Wang DZ (2006) The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genet 38:228–233
107. Esquela-Kerscher A, Slack FJ (2006) Oncomirs – microRNAs with a role in cancer. Nat Rev Cancer 6:259–269
108. Hermeking H (2007) p53 enters the microRNA world. Cancer Cell 12:414–418
109. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA et al (2005) MicroRNA expression profiles classify human cancers. Nature 435:834–838
110. Lee YS, Dutta A (2009) MicroRNAs in cancer. Annu Rev Pathol 4:199–227
111. Benetti R, Gonzalo S, Jaco I, Munoz P, Gonzalez S, Schoeftner S, Murchison E, Andl T, Chen T, Klatt P et al (2008) A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. Nat Struct Mol Biol 15:268–279
112. Fabbri M, Garzon R, Cimmino A, Liu Z, Zanesi N, Callegari E, Liu S, Alder H, Costinean S, Fernandez-Cymering C et al (2007) MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. Proc Natl Acad Sci USA 104:15805–15810
113. Friedman JM, Liang G, Liu CC, Wolff EM, Tsai YC, Ye W, Zhou X, Jones PA (2009) The putative tumor suppressor microRNA-101 modulates the cancer epigenome by repressing the polycomb group protein EZH2. Cancer Res 69:2623–2629
114. Varambally S, Cao Q, Mani RS, Shankar S, Wang X, Ateeq B, Laxman B, Cao X, Jing X, Ramnarayanan K et al (2008) Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. Science 322:1695–1699
115. Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, Ghosh D, Pienta KJ, Sewalt RG, Otte AP et al (2002) The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature 419:624–629
116. Kleer CG, Cao Q, Varambally S, Shen R, Ota I, Tomlins SA, Ghosh D, Sewalt RG, Otte AP, Hayes DF et al (2003) EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. Proc Natl Acad Sci USA 100:11606–11611
117. Weikert S, Christoph F, Kollermann J, Muller M, Schrader M, Miller K, Krause H (2005) Expression levels of the EZH2 polycomb transcriptional repressor correlate with aggressiveness and invasive potential of bladder carcinomas. Int J Mol Med 16:349–353
118. Saito Y, Jones PA (2006) Epigenetic activation of tumor suppressor microRNAs in human cancer cells. Cell Cycle 5:2220–2222
119. Toyota M, Suzuki H, Sasaki Y, Maruyama R, Imai K, Shinomura Y, Tokino T (2008) Epigenetic silencing of microRNA-34b/c and B-cell translocation gene 4 is associated with CpG island methylation in colorectal cancer. Cancer Res 68:4123–4132
120. Futscher BW, Oshiro MM, Wozniak RJ, Holtan N, Hanigan CL, Duan H, Domann FE (2002) Role for DNA methylation in the control of cell type specific maspin expression. Nat Genet 31:175–179
121. Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 9:465–476
122. Howard G, Eiges R, Gaudet F, Jaenisch R, Eden A (2008) Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. Oncogene 27:404–408
123. Eden A, Gaudet F, Waghmare A, Jaenisch R (2003) Chromosomal instability and tumors promoted by DNA hypomethylation. Science 300:455
124. Wilson AS, Power BE, Molloy PL (2007) DNA hypomethylation and human diseases. Biochim Biophys Acta 1775:138–162
125. Herold S, Herkert B, Eilers M (2009) Facilitating replication under stress: an oncogenic function of MYC? Nat Rev Cancer 9:441–444

126. Ying Y, Tao Q (2009) Epigenetic disruption of the WNT/beta-catenin signaling pathway in human cancers. Epigenetics 4:307–312
127. Polakis P (2000) Wnt signaling and cancer. Genes Dev 14:1837–1851
128. Walsh CP, Chaillet JR, Bestor TH (1998) Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet 20:116–117
129. Cho NY, Kim BH, Choi M, Yoo EJ, Moon KC, Cho YM, Kim D, Kang GH (2007) Hypermethylation of CpG island loci and hypomethylation of LINE-1 and Alu repeats in prostate adenocarcinoma and their relationship to clinicopathological features. J Pathol 211:269–277
130. Lee PS, Teaberry VS, Bland AE, Huang Z, Whitaker RS, Baba T, Fujii S, Secord AA, Berchuck A, Murphy SK (2010) Elevated MAL expression is accompanied by promoter hypomethylation and platinum resistance in epithelial ovarian cancer. Int J Cancer 126(6): 1378–1389
131. Oki Y, Issa JP (2007) Treatment options in advanced myelodysplastic syndrome, with emphasis on epigenetic therapy. Int J Hematol 86:306–314
132. Beumer JH, Parise RA, Newman EM, Doroshow JH, Synold TW, Lenz HJ, Egorin MJ (2008) Concentrations of the DNA methyltransferase inhibitor 5-fluoro-2′-deoxycytidine (FdCyd) and its cytotoxic metabolites in plasma of patients treated with FdCyd and tetrahydrouridine (THU). Cancer Chemother Pharmacol 62:363–368
133. Yoo CB, Chuang JC, Byun HM, Egger G, Yang AS, Dubeau L, Long T, Laird PW, Marquez VE, Jones PA (2008) Long-term epigenetic therapy with oral zebularine has minimal side effects and prevents intestinal tumors in mice. Cancer Prev Res (Phila Pa) 1:233–240
134. Cheng JC, Weisenberger DJ, Gonzales FA, Liang G, Xu GL, Hu YG, Marquez VE, Jones PA (2004) Continuous zebularine treatment effectively sustains demethylation in human bladder cancer cells. Mol Cell Biol 24:1270–1278
135. Yoo CB, Jones PA (2006) Epigenetic therapy of cancer: past, present and future. Nat Rev Drug Discov 5:37–50
136. Jones PA, Taylor SM (1980) Cellular differentiation, cytidine analogs and DNA methylation. Cell 20:85–93
137. Issa JP, Kantarjian HM (2009) Targeting DNA methylation. Clin Cancer Res 15:3938–3946
138. Liang G, Gonzales FA, Jones PA, Orntoft TF, Thykjaer T (2002) Analysis of gene induction in human fibroblasts and bladder cancer cells exposed to the methylation inhibitor 5-aza-2′-deoxycytidine. Cancer Res 62:961–966
139. Momparler RL (2005) Epigenetic therapy of cancer with 5-aza-2′-deoxycytidine (decitabine). Semin Oncol 32:443–451
140. Wijermans P, Lubbert M, Verhoef G, Bosly A, Ravoet C, Andre M, Ferrant A (2000) Low-dose 5-aza-2′-deoxycytidine, a DNA hypomethylating agent, for the treatment of high-risk myelodysplastic syndrome: a multicenter phase II study in elderly patients. J Clin Oncol 18:956–962
141. Silverman LR, Holland JF, Weinberg RS, Alter BP, Davis RB, Ellison RR, Demakos EP, Cornell CJ Jr, Carey RW, Schiffer C et al (1993) Effects of treatment with 5-azacytidine on the *in vivo* and *in vitro* hematopoiesis in patients with myelodysplastic syndromes. Leukemia 7(Suppl 1):21–29
142. Silverman LR, Demakos EP, Peterson BL, Kornblith AB, Holland JC, Odchimar-Reissig R, Stone RM, Nelson D, Powell BL, DeCastro CM 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
143. 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
144. Yang AS, Estecio MR, Garcia-Manero G, Kantarjian HM, Issa JP (2003) Comment on "Chromosomal instability and tumors promoted by DNA hypomethylation" and "Induction of tumors in nice by genomic hypomethylation". Science 302:1153, author reply 1153

145. Sharma S, Kelly TK, Jones PA (2010) Epigenetics in cancer. Carcinogenesis 31:27–36
146. Cortez CC, Jones PA (2008) Chromatin, cancer and drug therapies. Mutat Res 647:44–51
147. Datta J, Ghoshal K, Denny WA, Gamage SA, Brooke DG, Phiasivongsa P, Redkar S, Jacob ST (2009) A new class of quinoline-based DNA hypomethylating agents reactivates tumor suppressor genes by blocking DNA methyltransferase 1 activity and inducing its degradation. Cancer Res 69:4277–4285

Genome-Wide Epigenetic Modifications in Cancer

Yoon Jung Park, Rainer Claus, Dieter Weichenhan, and Christoph Plass

Abstract Epigenetic alterations in cancer include changes in DNA methylation and associated histone modifications that influence the chromatin states and impact gene expression patterns. Due to recent technological advantages, the scientific community is now obtaining a better picture of the genome-wide epigenetic changes that occur in a cancer genome. These epigenetic alterations are associated with chromosomal instability and changes in transcriptional control which influence the overall gene expression differences seen in many human malignancies. In this review, we will briefly summarize our current knowledge of the epigenetic patterns and mechanisms of gene regulation in healthy tissues and relate this to what is known for cancer genomes. Our focus will be on DNA methylation. We will review the current standing of technologies that have been developed over recent years. This field is experiencing a revolution in the strategies used to measure epigenetic alterations, which includes the incorporation of next generation sequencing tools. We also will review strategies that utilize epigenetic information for translational purposes, with a special emphasis on the potential use of DNA methylation marks for early disease detection and prognosis. The review will close with an outlook on challenges that this field is facing.

1 Epigenetic Modifications in Healthy Tissue

1.1 Epigenetic Patterns in the Human Genome

The development of an organism relies on gene expression patterns that are regulated in a spatial and time-dependent manner. This is accomplished, in part,

Y.J. Park, R. Claus, D. Weichenhan, and C. Plass
Division of Epigenomics and Cancer Risk Factors, German Cancer Research Center, Im Neuenheimer Feld 280, 69120, Heidelberg, Germany
e-mail: c.plass@dkfz.de

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_2,

by altering the DNA and associated proteins without changing the DNA sequence; these alterations are collectively termed "epigenetic modifications." Heritable information is carried by chemical modifications of both DNA and chromatin-associated proteins within the genomes of eukaryotes. Epigenetic modifications have profound influences on gene expression by modulating chromatin structure and DNA accessibility. Together, these epigenetic modifications on a genome-wide scale are referred to as the "epigenome". Examination of the epigenome suggests that epigenetic phenomena contribute to global, cell-specific patterns of gene expression. The epigenome is highly dynamic depending on the tissue type and developmental stage within an organism. The biological and biomedical significance of epigenetic modifications becomes particularly evident in mouse models where alterations of genes responsible for the addition or removal of epigenetic modifications lead to embryonic lethality (e.g., [1–4]).

Epigenetic modifications mainly fall into two categories, DNA methylation and histone modifications. In mammals, DNA can be modified by addition of a methyl group to cytosine residues, typically in a CpG context. This is mediated by DNA methyltransferases, such as the de novo DNA methyltransferases DNMT3a and DNMT3b, and/or DNMT1, which reestablishes DNA methylation at newly synthesized sequences after replication (reviewed in [5]). Conversely, it is not fully understood how DNA methylation is removed from the mammalian genome, although several proteins, including DNA methyl binding domain-containing proteins, nucleotide excision repair factors, and DNA methyltransferases, have been suggested to be part of the demethylation process (reviewed in [6]). Intriguingly, hydroxmethylcytosine was recently reported as an additional modification to cytosine residues in mouse brain and ES cells [7, 8]. Data indicated that it was generated from 5-methylcytosine by TET1 activity, which may be a key step for the demethylation process of 5-methylcytosine.

Unlike plants or invertebrates, most of which have mosaic methylation patterns, mammalian genomes have "ubiquitous" DNA methylation patterns (reviewed in [9]). CpG dinucleotides are predominantly methylated, except for those in CpG islands (CGIs). CGIs are defined as GC-rich regions that retain a higher than expected frequency of CpG dinucleotides. They frequently colocalize with promoter regions of genes (reviewed in [10]). Unmethylated CGIs account for only 1–2% of the whole genome. A small yet significant portion of them are heavily methylated and are involved in directing tissue-specific patterns of gene expression, genomic imprinting, and X chromosome inactivation.

In addition to cytosine residues of DNA, N-terminal tails of histone proteins are subject to modifications, including methylation, acetylation, phosphorylation, ribosylation, and ubiquitination. Genome-wide analysis of histone modifications revealed that acetylation around transcriptional start sites is correlated with transcriptional activation and chromatin accessibility [11]. The impact of lysine (K) methylation on gene expression varies depending on which Lys residue is modified and how many methyl groups it carries. For example, trimethylated K9 of histone H3 and K20 of histone H4 are enriched in constitutive heterochromatin, while trimethylated K27 and K9 of histone H3 are enriched in the inactive X chromosome [12]. These histone

modifications are silencing marks of gene expression. In contrast, trimethylation at H3K4 and H3K36 is associated with active chromatin [13]. High levels of monomethylation and low trimethylation at H3K4 are frequently observed in enhancer elements; such modifications have become a useful tool to identify potential enhancers together with cell type-specific expression patterns [14]. The discovery of interdependent relationships between DNA methylation and specific histone modifications has just begun. DNA methylation can provide a template to maintain certain histone modifications after replication. Conversely, specific histone methylation can guide the establishment of DNA methylation patterns (reviewed in [15]).

1.2 DNA Methylation and Patterns of Gene Expression

The conventional view ascribes a repressive role for DNA methylation on gene expression. Promoters of transcriptionally active genes are usually unmethylated, but become silenced once targeted by DNA methylation. DNA methylation-associated gene silencing is mediated either by recruiting methyl binding domain proteins and repressive factors such as histone deacetylases or by blocking recruitment of transcriptional factors (reviewed in [16, 17]). In contrast to DNA methylation in promoters, methylation within gene bodies is observed in transcriptionally active genes. It is not known, however, whether methylation reflects the default state in the genome or is targeted in a sequence- or locus-specific manner (reviewed in [18, 19]). DNA methylation-dependent transcriptional activation has been shown at some imprinted loci by preventing interactions with enhancer blocking factors [20, 21]. Whether or not this mechanism directly applies to the function of genome-wide gene body methylation remains unclear.

Genome-wide DNA methylation analysis in various human tissues has revealed the presence of tissue-specific differentially methylated regions [19, 22], which may play a role in cellular memory and tissue-specific genome function. Tissue-specific DNA methylation patterns are less frequently observed in the middle of CpG islands, instead differences have been suggested to occur in sequences of intermediate CpG density up to 2kb away, referred to as CpG island shores [23]. Distinct epigenetic patterns are also observed in genome-wide maps of ES cells vs. differentiated cells in mice [24, 25]. A bivalent chromatin state consisting of H3K27me3 and H3K4me3 at the same genetic location was observed in embryonic stem (ES) cells [24]. This state keeps genes in ES cells silenced but poised for either continued silencing or activation upon ES cell differentiation. Stable repression of genes associated with the pluripotent state also requires DNA methylation, which complements other regulatory mechanisms such as histone modifications or recruitment of transcription factors [26]. Recent findings in induced pluripotent stem cells emphasize the importance of proper patterning of DNA methylation, in conjunction with the presence of specific transcription factors, to define pluripotency [27].

1.3 DNA Methylation in Genomic Imprinting and X Inactivation

In addition to transcriptional regulation during development, epigenetic regulation is required for proper patterns of genomic imprinting and X chromosome inactivation during germ cell development and early embryogenesis (reviewed in [5, 28]). Imprinted genes are a subset of genes that show allele-specific expression defined by the allele's parent-of-origin and are controlled by epigenetic mechanisms. Allele-specific DNA methylation at imprinting control regions (ICRs) has long been known to be the main force behind imprinted gene expression [29]. DNA methylation is established during gametogenesis in a sex-dependent manner, resulting in allelic differences after fertilization. DNA methylation in ICRs contributes to allele-specific expression patterns either by inhibiting transcription of regulatory noncoding RNAs or by blocking DNA binding factors such as CTCF [30]. The parentally imprinted methylation is completely erased and is reestablished during primordial germ cell development. Epigenetic regulators of histone modifications such as G9a and the PRC2 complex have also been shown to be involved in maintenance of placental imprinting patterns [31, 32]. These imprinted modifications are faithfully maintained throughout development. Disruptions of the imprinting patterns are associated with human diseases, most notably cancer (reviewed in [33]).

Compared with the single active X chromosome in male somatic cells, inactivation of one of the X chromosomes in female mammalian somatic cells is necessary for dosage compensation of X chromosomal gene expression. X chromosome inactivation is regulated by imprinted and random patterns in extraembryonic and embryonic tissues, respectively. Similar to imprinted genes, epigenetic marks initiate and maintain inactivation of the X chromosome and undergo dynamic reprogramming during germ cell development and early embryogenesis. Embryonic random X inactivation utilizes DNA methylation to prevent *Tsix* expression, a noncoding RNA that blocks *Xist* expression, and allows for *Xist* transcription. There is, however, no consensus whether differential methylation in the control regions of *Xist*, *Tsix*, and *Xite* leads to imprinted X inactivation during early embryogenesis (reviewed in [34]). In addition to the primary role in random inactivation, DNA methylation also provides additional levels of repression for long-term inactivation across the whole X chromosome, together with the spreading of trimethylated H3K27 [35].

1.4 DNA Methylation and Genome Stability

In the human genome, DNA methylation resides predominantly in repetitive genomic regions, which include satellite DNA and retrotransposons such as LINEs, SINEs, and LTRs. Satellite DNA, mainly located in centromeric and telomeric regions, consists of

tandem repeats and forms heterochromatin characterized by DNA methylation and trimethylation of H3K9. About 40–50% of the mammalian genome consists of transposable elements. Their transcriptional repression by DNA methylation is likely to function as a host defense mechanism to maintain genomic integrity and stability [36]. DNA methylation in retrotransposons undergoes dynamic reprogramming during early embryogenesis and germ line development. Some repetitive elements, such as intracisternal A particles (IAPs), are resistant to demethylation in both the primordial germ cells (PGCs) and the zygote, possibly a critical necessity for genome stability [37]. Accordingly, establishment and maintenance of DNA methylation in transposable elements after reprogramming appear to be essential to prevent transpositions. Moreover, mutations in *Dnmt3L*, a DNA methyltransferase family member expressed during germ line development, cause loss of methylation and de-repression of transcripts in transposable elements, leading to meiotic defects and male infertility [38, 39]). Argonaute family proteins, MILI and MIWI1/2, are also indispensible for maintenance of DNA methylation in retrotransposons during germ cell development [40]. Together with genome instability, the loss of DNA methylation in retrotransposons is a hallmark of various cancers.

Altogether, epigenetic modifications are key regulators of tissue-specific gene expression, genomic imprinting, X inactivation, and repression of retrotransposons during development. Understanding the epigenetic patterns in normal tissue and how they are disrupted in cancer will provide insights how to tackle pathogenesis.

2 Genome-Wide Epigenetic Alterations in Cancer

2.1 Hypermethylation of Candidate Genes

Epigenetic alterations in cancer have initially been investigated in candidate gene approaches. Following the hypothesis that epigenetic alterations have the ability to silence gene transcription, tumor suppressor genes were tested for epigenetic changes in their promoter regions. These tests were mainly building on the identification of aberrant DNA methylation as a marker for epigenetic alterations. The base 5-methylcytosine is a stable mark of the DNA sequence, and hence, can readily be traced even in archived materials. DNA methylation assays were developed that allowed rapid testing for changes in DNA methylation in comparisons of normal and tumor tissue DNA. PCR-based assays following conversion of the DNA by sodium bisulfite treatment [41] proved sensitive, quantitative, and scalable to high throughput applications (reviewed in [42]). Initial results quickly demonstrated that numerous GC-rich promoter regions in virtually every human malignancy are targets for epigenetic alterations and gene silencing. In these studies, known tumor suppressor genes, such as *MLH1* in colon cancer [43, 44], *BRCA1* in breast cancer [45, 46], *DAPK1* in chronic lymphocytic leukemia (CLL) [47], or $p16^{INK4a}$ in head and neck cancer [48, 49] and lung cancer [50], were found to be epigenetically repressed.

2.2 DNA Hypomethylation of Cancer Genomes

Early studies, in which the 5-methylcytosine content of cancer genomes was measured, indicated that global levels in a tumor genome are reduced as compared with those in normal tissues [51]. It quickly became clear that loss of 5-methylcytosine occurs in sequences spread throughout the genome that are usually methylated in normal cells. These sequences included centromeric repeats and alpha satellite sequences located in centromeric regions, but also interspersed repetitive elements such as LINE1 sequences. The consequence of hypomethylation at these repeat sequences is genomic instability caused by an opening of the chromatin and subsequent chromosomal breakage. This may explain the numerous chromosomal aberrations found as one of the hallmarks in cancer genomes. Numerically, hypomethylation events exceed the number of hypermethylation events by far if one considers the abundance of methylated CpG dinucleotides in GC-rich repetitive sequences (>80% of the total genomic CpG content) relative to the number of unmethylated CpG dinucleotides in CpG island sequences (1–2%).

2.3 Epigenetic Changes in Imprinted Regions

Sequences that display both hypo- and hypermethylation include the ICRs. Both the gain and loss of methylation in these differentially methylated regions result in loss of genomic imprinting and dysregulation of genes controlled by protein complexes that detect the unmethylated ICR but not the methylated. Examples include loss of imprinting due to the hypermethylation of the ICR of the IGF2/H19 locus, resulting in overexpression of growth-activating *IGF2* [52, 53].

2.4 Lessons Learned from Genome-Wide Approaches

Candidate gene approaches are not sufficient to evaluate the amount of epigenetic alterations in a cancer genome (Fig. 1). However, for a long time, assays to evaluate the DNA methylation status of all 28×10^6 CpG dinucleotides simultaneously in a human genome were not available (see also below, Methylome analysis). Thus, most first generation scanning assays focused on a representation of the genome rather than attempting to cover the entire genome. Assays that focused specifically on CpG island sequences included Differential Methylation Hybridization (DMH) [54] and Restriction Landmark Genomic Scanning (RLGS) [55]. Both assays were designed to allow the calculation of overall CpG island methylation frequency. DMH was initially developed to identify methylated sequences in a cancer genome in a screen of CpG island clones (or later arrayed oligonucleotides) representing CpG island sequences. In this assay, DNAs were divided into two pools after *Mse*I

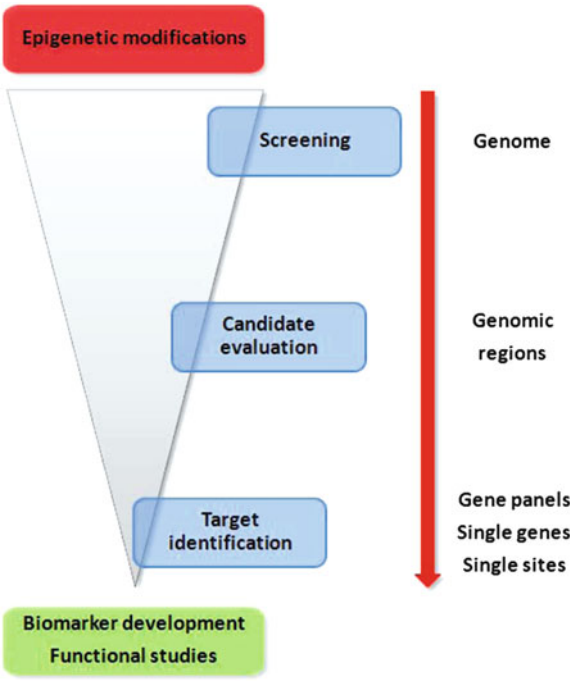


Fig. 1 Schematic outline of steps in an epigenetic screen (see text for more information)

restriction digest and linker-ligation. The control pool was amplified without further treatment whereas the test pool was digested with the methylation-sensitive restriction enzyme *Bst*UI prior to amplification. Subsequent to this procedure, both pools were labeled with different dyes and cohybridized to arrayed CpG island sequences. RLGS, on the other hand, was built on the restriction digest of genomic DNA with a methylation-sensitive restriction enzyme that preferentially cuts in CpG island sequences (e.g., *Not*I or *Asc*I). Restriction ends were radioactively labeled. Subsequent to a second restriction digest with a more frequently cutting restriction enzyme, the DNAs were separated in a tube-like agarose gel followed by in-gel digestion with a third restriction enzyme and final separation in an acrylamide gel. The gels were dried and exposed to an X-ray film which displayed up to 2000 RLGS fragments that represented unmethylated *Not*I or *Asc*I restriction sites.

While DMH and RLGS allow for the evaluation of overall levels of CpG island hypermethylation, other assays were designed specifically to identify hyper- or hypomethylated sequences from cancer genomes without providing data on the overall frequency of CpG island methylation in a tumor genome. For example, methylated CpG island amplification (MCA) was used as screening tool for the identification of novel methylated sequences in colon cancer [56]: MCA identified cancer-specific methylation events and a panel of sequences that characterizes the CpG island-methylator phenotype (CIMP). The CIMP is present in the majority of

sporadic colorectal cancers displaying microsatellite instability and being most frequently associated with *MLH1* hypermethylation. A surprise resulting from the data of the first genome scans performed by RLGS was the number of aberrantly methylated CpG islands. Leukemias, for example, demonstrated mean levels of CpG island methylation of 4.8% (chronic lymphocytic leukemia, CLL) and 1.9% (in acute myeloid leukemia, AML) [57, 58]. Similar numbers were found in solid tumors with mean levels of CpG island methylation of 5.3% in lung cancer [59], 1% in primary head and neck cancer [60], and 4.6% in ovarian cancer [61]. These numbers exceed by far the estimated number of tumor suppressor genes. The frequency of CpG island methylation raises questions: which are the initial epigenetic silencing events? And what events are accumulating during tumorigenesis, perhaps due to a loss of DNA repair or accelerated growth? Lacking in this context is also the information on DNA methylation changes in regions outside of CpG islands. These regions might exhibit changes that are of relevance in tumorigenesis. There is hope that novel genome-wide scans will provide this information in a comprehensive manner.

An additional surprising finding in these studies was that DNA methylation events are tumor-type specific [62]. Tumors and tumor subtypes display specific patterns of aberrant CpG island methylation indicating specific, yet, unknown mechanisms that lead to the silencing of specific groups of genes within a tumor. Here, either direct targeting of genes by oncogene-encoded proteins, onco-fusion proteins, or a selection process are discussed as possible mechanisms (see [63] for a detailed discussion).

2.5 Genetic vs. Epigenetic Alterations

While candidate gene approaches detected epigenetic effects in genes that had previously been identified as target genes for genetic (mutational) events, the epigenetic screens identified novel genes and gene families that were predominantly or even exclusively silenced by epigenetic mechanisms. The importance of these types of genes in normal development and their silencing in tumorigenesis is under investigation in many laboratories. One example is *DAPK1*, a gene frequently silenced in many tumor types by epigenetic alterations. However, there are no reports on genetic mutations in the coding region of *DAPK1* [64]. An additional example is *CTNNA1*, a gene silenced by both epigenetic and genetic mechanisms in myelodysplastic syndrome (MDS) and AML cases with chromosome 5q deletions [65]. The possibility of concordant genetic and epigenetic events in the inactivation of tumor suppressor genes is now being used to identify novel tumor suppressor genes in regions of chromosomal loss where searches for mutated genes have failed to pinpoint candidate cancer genes [66]. The underlying assumption for this approach is that the two hits, postulated by Knudsen's "Two-Hit Hypothesis" [67], can be a combination of genetic and epigenetic events.

It is now becoming clear that DNA methylation changes are closely linked to alterations of other epigenetic modifications, especially histone tail modifications. An intriguing observation was the finding that many of the epigenetic target genes are targets for the polycomb repressor complex and are marked by the repressive histone tail modification H3K27me3, which is mediated by EZH2 a member of the polycomb group complex [68–70]. Furthermore, cancer cells show a loss of monoacetylated and trimethylated forms of histone H4K16 and K20 residues of histone H4. These changes occur predominantly in hypomethylated, repetitive DNA sequences of the cancer genome [71].

3 Second Generation Methodologies for Epigenome-Wide Scans

Microarray and novel sequencing techniques have facilitated the comprehensive analyses of whole transcriptomes and complex genomic sequences. These techniques also paved the way for the genome-wide scanning of DNA methylation states (methylome profiling). Methylome profiling covers the whole genome, yet historically concentrates on the methylation states in CpG islands because of their frequent overlap with/or their close vicinity to promoter sequences (see above). Here, we will concentrate on a few examples of recent technical achievements in whole genome profiling (second generation methylome profiling) rather than discuss approaches targeting single candidate genes. Basically, methylome profiling can be separated into two successive processes: sequence enrichment for potentially methylated CpG sites and sequence-based analysis (Table 1).

3.1 Enrichment for Methylated CpG Sites

The rationale for sequence enrichment is the reduction of genome complexity and sequence load in later analysis. Three technical alternatives are currently in use:

Table 1 Novel methods for methylome profiling

Method/Acronym	Enrichment	Analysis	References
HELP, MIAMI, RRBS	Restriction enzyme digestion with methylation-sensitive and -resistant isoschizomers	Microarray hybridization, NGS	[72, 73, 75, 76]
meDIP, MIRA, MCIp	Protein affinity purification	Microarray hybridization, NGS	[77–79, 83]
Sequence capture	Hybridization of bisulfite-treated DNA to oligonucleotides	NGS	[80–82]

1. Cutting with methylation-sensitive and methylation-resistant restriction isoschizomers followed by linker-mediated PCR amplification (developed in the late 1990th, still often employed in combination with modern analysis technology)
2. Affinity purification using antibodies or recombinant proteins with high affinity for methylated DNA
3. Sequence capture by hybridization to complementary oligonucleotides

A variety of enzymatic options are available to distinguish between methylated and nonmethylated states at CpG sites. In two approaches, abbreviated MIAMI [72] and HELP [73], genomic DNA is cut with the methylation-sensitive restriction enzyme *Hpa*II and, as internal control, the methylation-insensitive *Msp*I; both enzymes recognize the sequence CCGG. This tetranucleotide occurs ~2.3 million times in the human genome, ~22% of them residing in CpG islands [74]. After cutting genomic DNA to completion with either enzyme, linkers for final PCR amplification are ligated to the 5′ G-overhangs of size fractionated (e.g., 200–2,000 bp) restriction fragments. In the later analysis, methylated CpG sites are recognized by their absence in the *Hpa*II- and their presence in the *Msp*I-treated sample. Recent technical improvement of the HELP approach considerably extended the representation of the addressed sequence regions [75].

Bisulfite treatment enables to discriminate between different methylation states by the conversion of unmethylated cytosine to uracil while methylated cytosine remains unconverted. In *Msp*I reduced representation bisulfite sequencing (RRBS, [76]), genomic DNA is cut only with *Msp*I, subsequently ligated to linkers containing only methylated but no unmethylated cytosines, and then bisulfite treated. PCR amplification then leads to a change from cytosine to thymine for every unmethylated cytosine while methylated cytosines are preserved as cytosines.

Affinity purification of randomly fragmented DNA (200–1,000 bp) employs antibodies or proteins with high affinity for methylated CpGs [77–79]. The currently most widespread approach is meDIP [77] using antibodies against single-stranded methylated DNA. Alternatives apply recombinant human proteins MBD2 (MCIp, [78]) or complex MBD2/MBD3L1 (MIRA, [79]), which bind with high affinity to double-stranded methylated DNA. After binding, the methylated DNA fraction is eluted from the antibodies/proteins by a high-salt buffer or a gradient of buffers with increasing salt concentrations. The gradient discriminates between states of low, intermediate, and high methylation. For microarray analysis, eluted DNA can be directly labeled in a linear amplification reaction. The labeling products reflect more reliably the relative abundance of enriched fragments than products of exponential PCR amplification. Another advantage of the affinity compared with the enzyme-based enrichment approaches is their independence of specific recognition sites, allowing, at least theoretically, examination of all potentially methylated sequence stretches. However, fragments with high methylation density are favored compared with those with low or moderate methylation density. Moreover, since affinity purification offers no direct proof for the presence of methylated CpGs, identified candidate genes need validation by a confirmatory method.

Sequence capture of bisulfite-treated DNA employs oligonucleotide capture probes, which are complementary to specific target sequences. In two of three presented strategies, padlock probes were used. These probes are usually ~100 bases long and anneal via their end sequences to a target sequence, thereby forming a padlock- or horseshoe-like structure. In the first strategy, padlock probes were designed to target nonrepetitive sequences covering ten bases with a 5′ CpG which are flanked by at least 20 bases on each side free of CpGs [80]. After annealing to the genomic sequence, the probes prime DNA synthesis of the targeted ten bases. Subsequent ligation leads to the formation of single-stranded DNA circles. Using primers derived from the common backbone of the probes, all synthesis products can be amplified in a single PCR. In the second strategy, padlock probes were designed to target longer sequence stretches of up to 225 bases. Moreover, capturing arms were allowed to contain CpGs. Consequently, multiple probes were designed considering all possible sequence combinations after bisulfite treatment [81]. The third strategy employed 60-mer probes immobilized on a microarray [82]. Similar to the second strategy, these probes were also allowed to contain CpGs (up to 15). However, only a binary probe design with respect to possible methylation states was applied: fully unmethylated or fully methylated, referring to reports that efficient hybridization tolerates polymorphic sites and even up to six distributed mismatches. Similar to enzyme-based enrichment, the presented sequence capture methods addressed specific target sequences rather than enabled to profile the whole methylome. All three capture methods are bioinformatically demanding and, therefore, require special expertise. However, coverage of substantial parts of the methylome in all three studies leads to the expectation that comprehensive capture addressing the unique sequences of the human methylome may be feasible in the near future.

3.2 Methylome Analysis

Methylome analysis after sequence enrichment is either performed by hybridization on high-density oligonucleotide tiling microarrays or by next generation sequencing (NGS). Different commercial tiling microarray platforms are currently in use, offering the flexibility of custom-designed arrays or standard arrays covering, for example, promoter or CpG island sequences of the human genome. Tiling probes are usually 45–60 bps in length and cover the regions of interest like CpG islands in close spacing or even with overlaps. DNA samples are labeled with a fluorescent dye such as Cy3 or Cy5 and cohybridized with a control sample, labeled with the complementary dye, to the microarray. Control samples, like DNA from healthy tissue, are usually enriched in the same way as the test sample or they may consist of the nonenriched test DNA. After hybridization, scanning of the array generates an image file displaying the different signal intensities of the two DNA samples on the oligonucleotide probes. Different types of feature extraction software evaluate the image file and provide both signal intensity ratios and a set of quality control values. Additional corrections by normalizing unequal distributions of bulk fluorescence

intensity values may be necessary prior to final statistical data evaluation. Since microarray data provides no direct proof for the methylation state of CpG sites, validation by other methods involving bisulfite DNA treatment is mandatory. Recently, algorithms have been proposed that allow correlations of quantitative array data with DNA methylation levels [83, 84].

NGS offers an attractive alternative to microarray analysis and has already been combined with the enrichment strategies described above. The power of NGS became particularly evident when applied on enriched bisulfite treated DNA samples. In three sequence capture studies [80–82], the Illumina Genome Analyzer was used which can provide hundreds of millions of short read sequences (~30–35 bases) in a single run. Prior to sequencing, a library of short (~100–300 bp) DNA fragments has to be prepared by ligation-mediated PCR. Use of different linkers for PCR priming allows sample multiplexing.

The huge output of a single run requires enormous data storage capacity and powerful software for proper quality check-up and final mapping of sequences on the human reference genome. Bisulfite treatment of DNA leads to the reduction of sequence complexity and to ambiguities at CpG sites, necessitating novel mapping algorithms. Repetitive sequences, such as the transposon-like repeats which are highly methylated in the human genome [74], are excluded in silico from further evaluation using masking algorithms (e.g., Repeatmasker) because they cannot be mapped back to the reference genome.

All capture studies reported encouraging experimental performance with respect to both specificity (discovery of false positives) and sensitivity (likelihood for detecting rare positives). In one study, ~3 million reads were mapped to ~7,700 of ~10,000 targets (~77%) [80], while in another study ~5.5 million reads could be mapped to 10,364 of 10,582 targets (98% sensitivity) [81]. In the latter study, the abundance of different captured fragments ranged from 1 to ~10,000-fold. This variation could be traced back to a combination of parameters including the size of the target sequence and the GC-content. Knowing these parameters and their influence should enable a more reliable sequence representation in NGS methylome profiling projects.

4 Epigenetic Biomarkers

4.1 Epigenomic Profiles as Markers for Cancer Tissues

Biomarkers are biological parameters that can be objectively measured and evaluated as indicators of biological processes. Over the last years, biomarkers have gained an enormous impact on diagnosis and treatment of cancer and cancer-related diseases. They can be used as diagnostic tools and as prognostic factors that predict the outcome of individual patients in terms of a specific clinical endpoint. In addition, they may serve as predictive factors for the effect of a specific treatment and as surrogate

endpoints that replace a clinical endpoint of interest. Recently, the term biomarker has become a synonym for "molecular biomarker," which can be measured by molecular techniques in biological samples. Molecular biomarkers include changes in nucleic acid sequences such as mutations or polymorphisms and gene expression alterations, peptides, proteins, lipid metabolites, and other small molecules.

Human malignancies can be characterized by distinct epigenomic profiles as markers of the malignant cell clone (see above). Increased DNA methylation at CpG islands is prevalent in basically every human cancer, and different types of cancer can be reliably distinguished by their unique DNA hypermethylation pattern [62]. The list of the affected loci (genes) is rapidly growing and exerting its impact on clinical decisions (Table 2). Hypermethylation at distinct genomic loci has several properties that predispose for use as an attractive potential biomarker in disease. First of all, DNA hypermethylation of many distinct genes is characteristic for neoplastic cells. It is found to a significantly lower extent in healthy individuals. Early onset of DNA methylation changes is evident in the pathogenesis of many cancer types. This makes hypermethylation signatures an attractive tool in early detection and screening, particularly in patients who exhibit increased risk. Second, 5-methylcytosine is a chemically stable covalent mark that can be reliably detected in a variety of tissue sources. In contrast to RNA-based signatures, DNA methylation patterns are less prone to storage- or handling-dependent variations which could confound measurements and consecutive interpretations. Analyses can be performed on fresh tissues, archived frozen material, or paraffin-embedded tissues. Samples can be long-term stored for intra- and interindividual references. Detection is possible not only in tumor tissue, but also on tumor-derived DNA, which may be present in body fluids (such as peripheral blood or serum). In addition, the detection of epimutations (epigenetic modifications that can be passed down from parents to offspring) which may be present in unaffected tissues becomes a promising approach for a biomarker based on DNA methylation [85]. However, in assessment, evaluation, and interpretation of DNA methylation analyses of clinical samples, it is important to consider potential cytological heterogeneity as possible confounder. Third, the technology of DNA methylation measurement has greatly improved over the last years (see above). Sensitivity, reproducibility, and applicability for clinical settings have significantly improved. The ability of quantitative measurements at single CpG dinucleotide resolution enables tight correlations with clinical endpoints, precise identification, and separation of subgroups revealing previously unidentified differences. This gains particular importance considering several reports that indicate superior prognostic or predictive significance of single CpG dinucleotides within CpG-rich areas.

A major challenge in utilizing DNA hypermethylation events for sensitive and specific diagnostic and prognostic markers is the selection of candidate genes (schematically depicted in Fig. 1). For many entities, candidate gene selection using, for example, differential regulation or supposed function ("informed best-guess") led to identification of successful DNA methylation markers. However, this approach relies on restricted observations and assumptions and might not consider independent potential markers. With the development of genome-wide tools and

Table 2 Selection of prominent examples for established and potential epigenetic biomarkers

Methylated genes	Specimen	Lesion type	(Potential) Clinical use	Reference
GSTPI	Tumor tissue	Prostate cancer	Diagnosis and risk assessment	[110]
	Urine	Prostate cancer	Screening, detection	[111]
	Serum	Prostate cancer	Screening, detection	[112]
hMLH1	Tumor tissue	Colon cancer, gastric cancer, endometrial cancer	Classification and risk assessment	[113]
	Endometrial tissue	Atypical endometrial hyperplasia (endometrial cancer)	Screening and risk assessment	[89]
	Tumor tissue	Various tumors	Response to treatment	[103]
	Tumor tissue	Diffuse large B cell lymphoma (DLBCL)	Prognosis	[114]
	Serum	Ovarian cancer	Chemosensitivity	[102]
$p16^{INK4a}$	Esophageal tissue	Barrett's esophagus (esophageal carcinoma)	Screening and risk assessment	[88]
	Serum	Esophageal cancer	Screening, detection	[115]
	Cervical cytologic specimens	Cervical neoplasia	Screening and risk assessment	[116]
	Tumor tissue, lymph nodes	Non small cell lung cancer (NSCLC)	Prognosis (early recurrence)	[117]
	Sputum	Lung cancer	Screening, detection, risk assessment	[118]
	Serum	Hepatocellular carcinoma	Screening and detection	[119]
MGMT	Colon mucosa	Colorectal adenoma (colorectal cancer)	Screening and risk assessment	[87]
	Tumor tissue	Glioma/Glioblastoma	Treatment response in Glioma	[100, 120]
SFRP1	Pancreatic excretion	Pancreatitis vs. pancreas carcinoma	Screening and risk assessment	[121]
	Tumor tissue	Breast cancer	Prognosis	[97]
e-cadherin	Tumor tissue	Bladder cancer	Prognosis	Reviewed in [122]
	Tumor tissue	Breast cancer	Prognosis, progression	[123]
	Tumor tissue	Skin cancer	Prognosis, staging	[124]
APC	Tumor tissue	Prostate cancer	Prognosis, marker for progression	[90, 125]
	Colon (tumor) tissue	Colorectal cancer	Detection	
	Breast tissue	Breast cancer	Detection	
P15	Blood/myeloblasts	AML	Prognosis, treatment response	[105]

the tremendous effort of initiatives trying to decipher nonmalignant and cancer methylomes, the exciting perspective of highly sensitive and specific DNA methylation signatures for various cancers and their respective subtypes becomes increasingly available. This might be of particular interest for distinction of tumors and nontumor diseases with similar behavior, e.g., chronic inflammations. Moreover, precise discrimination between tumor subtypes by DNA methylation signatures might help to strengthen diagnostic efforts and to conduct improved disease- and stage-specific therapy decisions. In a large number of cases, diagnosis is based on biopsies from undifferentiated metastatic tissue that makes it difficult for conventional histological and immunocytochemical approaches to determine the tumor origin (e.g., "CUP syndrome"). Since particular therapeutic options may vary greatly between different tumor types, defined methylation signatures could significantly contribute to the management of such cases.

4.2 Epigenetic Markers for (Early) Detection of Cancer Cells and Screening

Evidence from various mouse models (e.g., *TCL1*-transgenic mouse model for CLL pathogenesis) shows that DNA methylation events occur early during pathogenesis at particular genomic loci. They can already be detected in premalignant lesions [86]. Thus, DNA hypermethylation of distinct loci/genes can frequently be assessed even before the histological onset of the disease. In humans, this could convincingly be shown for CpG island hypermethylation in $p16^{INK4a}$, $p14^{ARF}$, and O6-methylguanine-DNA methyltransferase (*MGMT*) in colorectal adenomas [87], $p16^{INK4a}$, *RUNX3*, and *HPP1* in Barrett's esophagus [88], *MLH1* in atypical endometrial hyperplasia [89], *GSTPI* in prostate cancer [90], $p16^{INK4a}$ and/or *MGMT* in squamous cell lung carcinoma [91], and many more. These data strongly underline that the analysis of the DNA methylation signature plays an important role in screening and early detection of different malignancies. Particularly, in predisposed patients with either infectious or inflammatory conditions, DNA methylation signatures may be useful as markers of increased cancer risk. Individuals with family history of cancer can notably benefit from successful early cancer detection by such innovative sensitive screenings. This was impressively affirmed by studies in colorectal cancer where DNA methylation patterns of familiar cases exhibited striking similarity to those of sporadic cases [92]. With respect to practical feasibility, the presence of DNA in body fluids offers easy, noninvasive accessibility to material for numerous cancer types. A large case-control study recently demonstrated that hypomethylation measured in peripheral blood lymphocytes was strongly associated with increased risk of bladder cancer [93]. Indeed, hypomethylation of L1 LINE sequence elements in bladder cancer tissue had already been reported several years ago [94].

In reviewing the identification of new epigenetically altered candidate genes and the rapid development of technologies over the last few years, the challenging goal

is now to set up panels for reliable early detection markers especially in high risk patients. At the same time, early detection using methylation signatures opens up novel treatment options that include epigenetic therapeutic strategies. In addition, sensitive epigenetic screening panels may not only allow effective identification of early cancer stages but may also be used as instruments for disease monitoring and detection of relapses. Continued sequential analyses of DNA methylation signatures can be indicative for recurring disease at stages where clinical symptoms are still absent and conventional diagnostic tools do not offer sufficient sensitivity. First promising results come from a small prospective study demonstrating methylation of a gene panel in saliva for the early detection of relapses in head and neck squamous cell carcinomas [95].

4.3 DNA Methylation Profiles as Marker for Risk Assessment, Tumor Progression, and Prognosis

Usually, detection of cancer is directly followed by assessments of stage and risk of the malignant disease. As DNA methylation signatures fulfill requirements for prognostic biomarkers (the baseline value of the biomarker, or changes in the biomarker over time, should be correlated with the clinical endpoint in untreated or in treated patients), they can be used to supplement conventional staging. Several studies have demonstrated that DNA hypermethylation of distinct genes can be correlated with clinical parameters or even substitute for them. This implies that DNA methylation signatures can significantly contribute to risk stratification in malignant diseases and may furthermore define new prognostic subgroups. Striking examples come from colorectal and lung cancer where methylation of $p16^{INK4a}$ is accompanied by particularly poor prognosis [96]. In breast cancer, *SFRP1* promoter hypermethylation is associated with unfavorable prognosis and poor overall survival in patients in early stages of the disease [97]. Some studies extend the prognostic ability of epigenetic biomarkers from clinical endpoints like overall survival to distinct properties of the disease course. This has been demonstrated for the increased metastatic potential in cervical cancer assessed by hypermethylation in *MYOD1* CpG island [98]. In CLL, methylation of single CpGs separates the disease into major prognostic subgroups in addition to established prognostic parameters like IgVH mutation status or ZAP70 protein expression [99]. Precise and significant prognosis estimates are not only highly informative about the course of the disease but they allow more risk-adapted treatment decisions.

4.4 Predicting Therapy Response by Epigenomic Profiles

Treatment decisions in oncology are based on risk-adapted procedures. However, the efficiency of a particular therapy and the sensitivity of an individual cancer are difficult to predict. The exciting possibility of treatment response prediction may be

the most challenging and promising task for potential biomarkers. DNA methylation signatures have been demonstrated to serve as predictive markers, as their baseline value or their changes over time have been correlated with the effect of treatment. The most prominent example is *MGMT* hypermethylation in glioblastoma patients. Expression of *MGMT* leads to reduced toxicity of alkylating agents such as temozolomide due to rapid reversal of DNA adduct formation. DNA methylation of *MGMT* (and its methylation-associated silencing) is the best independent predictor for treatment response in glioblastoma [100]. Further evidence for predictive DNA methylation signatures comes from hypermethylation of *hMLH1*, a DNA mismatch repair gene. *hMLH1* hypermethylation frequently occurs in various tumors and is associated with increased resistance to chemotherapeutics like cisplatin [101]. In ovarian cancer, acquired *hMLH1* methylation in peripheral blood predicts for adverse response to chemotherapy [102]. Interestingly, in human tumor xenograft models, demethylation of the *hMLH1* promoter resulted in sensitization to cisplatin [103].

4.5 Monitoring Epigenetic Therapies

Despite some advances in our understanding of the mechanisms underlying epigenetic therapies, we still do not understand completely how these drugs work. Epigenetic therapies have tremendously evolved over the last few years. The increasing number of studies that report the significance of epigenetic alterations in cancerogenesis and distinct tumor cell properties build a strong rationale for the use of epigenetically modifying drugs in cancer treatment. In addition, frequent reports of epigenetic biomarkers and identification of hypermethylated genes sets that contribute to alteration of chemosensitivity support the rationale for reversal of DNA methylation patterns as an effective therapeutic approach. The DNA methyltransferase (DNMT) inhibitors 5-aza-2′-deoxycytidine/decitabine (Dacogen) and 5-azacytidine (Vidaza) have recently been approved by the U.S. Food and Drug Administration (FDA) for treatment of MDS. Therapeutic principles have been attributed to reversal of hypermethylation and reactivation of tumor suppressor genes (e.g., $p15^{INK4b}$). However, recent studies reported on particular DNA repair mechanisms and others to be involved in response to therapy. Several studies have used methylation signatures of single genes or gene combinations to monitor therapeutic effects of demethylating agents. Clinical trials in AML and MDS have shown decrease of DNA methylation at genome-wide levels (assessed by L1 LINE methylation) and at the $p15^{INK4b}$ promoter upon therapy [104, 105]. Quantitative measurements of DNA methylation allow precise monitoring of even minor demethylation effects, which are possibly indicative for effective therapy. Fetal hemoglobin (HbF) has been reported to be reactivated in patients upon treatment with demethylating agents [106]. It appears reasonable to further investigate HbF as an in vivo marker for the application, efficacy, and the assessment of treatment response. Markers for epigenetic therapy might be particularly important when

considering an extended use of demethylating agents in solid tumors in the future, where early treatment response is often more difficult to assess than in patients with hematologic neoplasia.

Taken together, DNA methylation changes at single sites, single genes, or panels of genes can serve as promising potent biomarkers that facilitate detection and clinical management of various cancer types.

5 Future Directions in Cancer Epigenetics

5.1 Understanding the Underlying Mechanisms of Epigenetic Regulation

Despite significant advances in the field, there are many open questions and challenges that remain in the understanding of the methylome and utilizing this information in basic and translational research. We still do not understand the underlying mechanisms that lead to epigenetic alterations. Major efforts should be made to precisely characterize epigenomic patterns during development and to understand aberrant epigenetic processes, with the ultimate goal to perturb them in cancer development. Mouse models for cancer may be helpful tools that recapitulate epigenetic defects [86, 107, 108]. These mice might allow us to dissect the cascade of events that leads to a global genome-wide epigenetic defect and to develop strategies that might help to prevent these alterations and subsequently cancer development.

5.2 Third Generation DNA Methylome Profiling

Although major advances have been made just recently in the development of novel scanning protocols, it can already be foreseen that further improvements will lead to a more detailed characterization of epigenetic patterns. Still ongoing improvements in probe density and design as well as relative simplicity in both performance and evaluation will assure that microarray technology will retain a share in methylome profiling studies during the next few years. In comparison, NGS is still reserved to only a few laboratories able to cope with the relatively high equipment costs and technological and bioinformatical requirements. It can be envisaged, however, that overall costs for both equipment and consumables will considerably drop soon, making NGS available to a much broader scientific community. In-depth methylome profiling of tumor samples, particularly in early stages, is often complicated or even prohibited by limited amounts of DNA, necessitating considerable improvement in target-sequence enrichment and assay sensitivity. This has been already addressed by some of the cited studies (e.g., [75]), and a novel development of

ultra-sensitive quantum-dot technology in profiling single candidate genes in large patient sample sets has just been published [109]. Both technical and financial reasons make it presently more likely that only a limited set of candidate genes or sequences rather than the whole methylome will be analyzed in the clinical routine. It remains to be seen whether such a candidate approach will be sufficient for proper molecular diagnostics, staging, and prognosis of a malignant disease.

5.3 Integration of Epigenetic Markers in Clinical Settings

Epigenetic alterations, in particular DNA methylation, have become promising new tools for cancer screening and diagnosis, risk assessment and prognosis estimation, as well as therapeutic management. However, more systematic evaluation in large, well-characterized patient cohorts is strongly desirable. As an important part of translational investigations, putative epigenetic biomarkers should be frequently incorporated in clinical trials enabling prospective sampling and more comprehensive evaluations. Especially, the development of predictive markers (foremost markers for prediction of therapy responses) is an important and urgent request for many cancer entities. Therapy response predictors (as described above for *MGMT* in glioblastoma) might lead to more patient-, risk-, and disease stage-adapted therapeutic strategies that directly translate into clinical benefit for the patient.

It is obvious that epigenetics markers have started to move from bench to bedside. They are entering the clinical field and becoming essential factors in the clinical management of cancer patients.

Acknowledgments The authors would like to thank Christopher Oakes for critical reading of the review. Work in the Division is funded in part by NIH grants CA101956 and DE013123 (C.P.). R.C. is supported by a fellowship of the Deutsche Forschungsgemeinschaft and YJ. P. holds a Roman Herzog Stipend of the Alexander von Humboldt-Stiftung.

References

1. Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69(6):915–926
2. Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99(3):247–257
3. Tachibana M, Sugimoto K, Nozaki M et al (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev 16(14):1779–1791
4. O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T (2001) The polycomb-group gene Ezh2 is required for early mouse development. Mol Cell Biol 21 (13):4330–4336

5. Sasaki H, Matsui Y (2008) Epigenetic events in mammalian germ-cell development: reprogramming and beyond. Nat Rev Genet 9(2):129–140
6. Ooi SK, Bestor TH (2008) The colorful history of active DNA demethylation. Cell 133 (7):1145–1148
7. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324(5929):929–930
8. Tahiliani M, Koh KP, Shen Y et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324(5929):930–935
9. Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 9(6):465–476
10. Illingworth RS, Bird AP (2009) CpG islands – 'a rough guide'. FEBS Lett 583(11): 1713–1720
11. Kurdistani SK, Tavazoie S, Grunstein M (2004) Mapping global histone acetylation patterns to gene expression. Cell 117(6):721–733
12. Schotta G, Lachner M, Sarma K et al (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev 18(11):1251–1262
13. Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. Cell 128 (4):707–719
14. Heintzman ND, Hon GC, Hawkins RD et al (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 459(7243):108–112
15. Cedar H, Bergman Y (2009) Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet 10(5):295–304
16. Gibbons RJ, McDowell TL, Raman S et al (2000) Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. Nat Genet 24 (4):368–371
17. Knoepfler PS, Eisenman RN (1999) Sin meets NuRD and other tails of repression. Cell 99 (5):447–450
18. Eckhardt F, Lewin J, Cortese R et al (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet 38(12):1378–1385
19. Rakyan VK, Down TA, Thorne NP et al (2008) An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). Genome Res 18(9):1518–1529
20. Yoon B, Herman H, Hu B et al (2005) Rasgrf1 imprinting is regulated by a CTCF-dependent methylation-sensitive enhancer blocker. Mol Cell Biol 25(24):11184–11190
21. Bell AC, Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene [see comments]. Nature 405(6785):482–485
22. Illingworth R, Kerr A, Desousa D et al (2008) A novel CpG island set identifies tissue-specific methylation at developmental gene loci. PLoS Biol 6(1):e22
23. Irizarry RA, Ladd-Acosta C, Wen B et al (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 41(2):178–186
24. Bernstein BE, Mikkelsen TS, Xie X et al (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125(2):315–326
25. Mikkelsen TS, Ku M, Jaffe DB et al (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448(7153):553–560
26. Fouse SD, Shen Y, Pellegrini M et al (2008) Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. Cell Stem Cell 2(2):160–169
27. Surani MA, Durcova-Hills G, Hajkova P, Hayashi K, Tee WW (2008) Germ line, stem cells, and epigenetic reprogramming. Cold Spring Harb Symp Quant Biol 73:9–15
28. 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(8):526–537

29. Plass C, Soloway PD (2002) DNA methylation, imprinting and cancer. Eur J Hum Genet 10 (1):6–16
30. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus [see comments]. Nature 405(6785):486–489
31. Umlauf D, Goto Y, Cao R et al (2004) Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. Nat Genet 36(12):1296–1300
32. Wagschal A, Sutherland HG, Woodfine K et al (2008) G9a histone methyltransferase contributes to imprinting in the mouse placenta. Mol Cell Biol 28(3):1104–1113
33. Jelinic P, Shaw P (2007) Loss of imprinting and cancer. J Pathol 211(3):261–268
34. Payer B, Lee JT (2008) X chromosome dosage compensation: how mammals keep the balance. Annu Rev Genet 42:733–772
35. Lee TI, Jenner RG, Boyer LA et al (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. Cell 125(2):301–313
36. Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. Annu Rev Biochem 74:481–514
37. Lees-Murdock DJ, De Felici M, Walsh CP (2003) Methylation dynamics of repetitive DNA elements in the mouse germ cell lineage. Genomics 82(2):230–237
38. Bourc'his D, Bestor TH (2004) Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature 431(7004):96–99
39. Webster KE, O'Bryan MK, Fletcher S et al (2005) Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. Proc Natl Acad Sci USA 102(11):4068–4073
40. Kuramochi-Miyagawa S, Watanabe T, Gotoh K et al (2008) DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev 22(7):908–917
41. Clark SJ, Harrison J, Paul CL, Frommer M (1994) High sensitivity mapping of methylated cytosines. Nucleic Acids Res 22(15):2990–2997
42. Brena RM, Huang TH, Plass C (2006) Quantitative assessment of DNA methylation: potential applications for disease diagnosis, classification, and prognosis in clinical settings. J Mol Med 84(5):365–377
43. Kane MF, Loda M, Gaida GM et al (1997) Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Res 57(5):808–811
44. Herman JG, Umar A, Polyak K et al (1998) Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci USA 95 (12):6870–6875
45. Dobrovic A, Simpfendorfer D (1997) Methylation of the BRCA1 gene in sporadic breast cancer. Cancer Res 57(16):3347–3350
46. Rice JC, Massey-Brown KS, Futscher BW (1998) Aberrant methylation of the BRCA1 CpG island promoter is associated with decreased BRCA1 mRNA in sporadic breast cancer cells. Oncogene 17(14):1807–1812
47. Raval A, Tanner SM, Byrd JC et al (2007) Downregulation of death-associated protein kinase 1 (DAPK1) in chronic lymphocytic leukemia. Cell 129(5):879–890
48. Merlo A, Herman JG, Mao L et al (1995) 5′ CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers [see comments]. Nat Med 1(7):686–692
49. Bartkova J, Lukas J, Muller H, Strauss M, Gusterson B, Bartek J (1995) Abnormal patterns of D-type cyclin expression and G1 regulation in human head and neck cancer. Cancer Res 55 (4):949–956
50. Otterson GA, Khleif SN, Chen W, Coxon AB, Kaye FJ (1995) CDKN2 gene silencing in lung cancer by DNA hypermethylation and kinetics of p16INK4 protein induction by 5-aza 2′deoxycytidine. Oncogene 11(6):1211–1216

51. Diala ES, Cheah MS, Rowitch D, Hoffman RM (1983) Extent of DNA methylation in human tumor cells. J Natl Cancer Inst 71(4):755–764
52. Rainier S, Johnson LA, Dobry CJ, Ping AJ, Grundy PE, Feinberg AP (1993) Relaxation of imprinted genes in human cancer. Nature 362(6422):747–749
53. Ogawa O, Becroft DM, Morison IM et al (1993) Constitutional relaxation of insulin-like growth factor II gene imprinting associated with Wilms' tumour and gigantism. Nat Genet 5 (4):408–412
54. Huang TH, Perry MR, Laux DE (1999) Methylation profiling of CpG islands in human breast cancer cells. Hum Mol Genet 8(3):459–470
55. Smiraglia DJ, Plass C (2002) The study of aberrant methylation in cancer via restriction landmark genomic scanning. Oncogene 21(35):5414–5426
56. Toyota M, Ho C, Ahuja N et al (1999) Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. Cancer Res 59(10):2307–2312
57. Rush LJ, Dai Z, Smiraglia DJ et al (2001) Novel methylation targets in de novo acute myeloid leukemia with prevalence of chromosome 11 loci. Blood 97(10):3226–3233
58. Raval A, Rush LJ, Funchain P et al (2002) Aberrant DNA methylation in chronic lymphocytic leukemia: a role in pathogenesis? Blood 100(11):379a
59. Dai Z, Lakshmanan RR, Zhu WG et al (2001) Global methylation profiling of lung cancer identifies novel methylated genes. Neoplasia 3(4):314–323
60. Smiraglia DJ, Smith LT, Lang JC et al (2003) Differential targets of CpG island hypermethylation in primary and metastatic head and neck squamous cell carcinoma (HNSCC). J Med Genet 40(1):25–33
61. Wei SH, Chen CM, Strathdee G et al (2002) Methylation microarray analysis of late-stage ovarian carcinomas distinguishes progression-free survival in patients and identifies candidate epigenetic markers. Clin Cancer Res 8:2246–2252
62. Costello JF, Fruhwald MC, Smiraglia DJ et al (2000) Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. Nature Genet 24(2):132–138
63. Plass C, Smiraglia DJ (2006) Genome-wide analysis of DNA methylation changes in human malignancies. Curr Top Microbiol Immunol 310:179–198
64. Bialik S, Kimchi A (2006) The death-associated protein kinases: structure, function, and beyond. Annu Rev Biochem 75:189–210
65. Liu TX, Becker MW, Jelinek J et al (2007) Chromosome 5q deletion and epigenetic suppression of the gene encoding alpha-catenin (CTNNA1) in myeloid cell transformation. Nat Med 13(1):78–83
66. Smith LT, Lin M, Brena RM et al (2006) Epigenetic regulation of the tumor suppressor gene TCF21 on 6q23-q24 in lung and head and neck cancer. Proc Natl Acad Sci USA 103 (4):982–987
67. Knudson AG Jr (1971) Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci USA 68(4):820–823
68. Ohm JE, McGarvey KM, Yu X et al (2007) A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 39(2):237–242
69. Widschwendter M, Fiegl H, Egle D et al (2007) Epigenetic stem cell signature in cancer. Nat Genet 39(2):157–158
70. Schlesinger Y, Straussman R, Keshet I et al (2007) Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet 39 (2):232–236
71. Fraga MF, Ballestar E, Villar-Garea A et al (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet 37(4):391–400
72. Hatada I, Fukasawa M, Kimura M et al (2006) Genome-wide profiling of promoter methylation in human. Oncogene 25(21):3059–3064

73. Khulan B, Thompson RF, Ye K et al (2006) Comparative isoschizomer profiling of cytosine methylation: the HELP assay. Genome Res 16(8):1046–1055
74. Fazzari MJ, Greally JM (2004) Epigenomics: beyond CpG islands. Nat Rev Genet 5 (6):446–455
75. Oda M, Glass JL, Thompson RF et al (2009) High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers. Nucleic Acids Res 37(12):3829–3839
76. Meissner A, Mikkelsen TS, Gu H et al (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454(7205):766–770
77. Weber M, Davies JJ, Wittig D et al (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 37(8):853–862
78. Gebhard C, Schwarzfischer L, Pham TH et al (2006) Genome-wide profiling of CpG methylation identifies novel targets of aberrant hypermethylation in myeloid leukemia. Cancer Res 66(12):6118–6128
79. Rauch T, Li H, Wu X, Pfeifer GP (2006) MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells. Cancer Res 66(16):7939–7947
80. Ball MP, Li JB, Gao Y et al (2009) Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 27(4):361–368
81. Deng J, Shoemaker R, Xie B et al (2009) Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. Nat Biotechnol 27(4):353–360
82. Hodges E, Smith A, Kendall J et al (2009) High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. Genome Res 19 (9):1593–1605
83. Down TA, Rakyan VK, Turner DJ et al (2008) A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nat Biotechnol 26(7):779–785
84. Pelizzola M, Koga Y, Urban AE et al (2008) MEDME: an experimental and analytical methodology for the estimation of DNA methylation levels based on microarray derived MeDIP-enrichment. Genome Res 18(10):1652–1659
85. Chan TL, Yuen ST, Kong CK et al (2006) Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. Nat Genet 38(10):1178–1183
86. Chen SS, Raval A, Johnson A et al (2009) Epigenetic changes during disease progression in a murine model of human chronic lymphocytic leukemia. Proc Natl Acad Sci USA 106(32): 13433–13438
87. Judson H, Stewart A, Leslie A et al (2006) Relationship between point gene mutation, chromosomal abnormality, and tumour suppressor gene methylation status in colorectal adenomas. J Pathol 210(3):344–350
88. Schulmann K, Sterian A, Berki A et al (2005) Inactivation of p16, RUNX3, and HPP1 occurs early in Barrett's-associated neoplastic progression and predicts progression risk. Oncogene 24(25):4138–4148
89. Banno K, Yanokura M, Susumu N et al (2006) Relationship of the aberrant DNA hypermethylation of cancer-related genes with carcinogenesis of endometrial cancer. Oncol Rep 16(6):1189–1196
90. Jeronimo C, Usadel H, Henrique R et al (2001) Quantitation of GSTP1 methylation in non-neoplastic prostatic tissue and organ-confined prostate adenocarcinoma. J Natl Cancer Inst 93(22):1747–1752
91. Palmisano WA, Divine KK, Saccomanno G et al (2000) Predicting lung cancer by detecting aberrant promoter methylation in sputum. Cancer Res 60(21):5954–5958
92. Esteller M, Fraga MF, Guo M et al (2001) DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. Hum Mol Genet 10(26):3001–3007

93. Moore LE, Pfeiffer RM, Poscablo C et al (2008) Genomic DNA hypomethylation as a biomarker for bladder cancer susceptibility in the Spanish Bladder Cancer Study: a case-control study. Lancet Oncol 9(4):359–366
94. Jurgens B, Schmitz-Drager BJ, Schulz WA (1996) Hypomethylation of L1 LINE sequences prevailing in human urothelial carcinoma. Cancer Res 56(24):5698–5703
95. Righini CA, de Fraipont F, Timsit JF et al (2007) Tumor-specific methylation in saliva: a promising biomarker for early detection of head and neck cancer recurrence. Clin Cancer Res 13(4):1179–1185
96. Esteller M, Gonzalez S, Risques RA et al (2001) K-ras and p16 aberrations confer poor prognosis in human colorectal cancer. J Clin Oncol 19(2):299–304
97. Veeck J, Niederacher D, An H et al (2006) Aberrant methylation of the Wnt antagonist SFRP1 in breast cancer is associated with unfavourable prognosis. Oncogene 25 (24):3479–3488
98. Widschwendter A, Muller HM, Fiegl H et al (2004) DNA methylation in serum and tumors of cervical cancer patients. Clin Cancer Res 10(2):565–571
99. Corcoran M, Parker A, Orchard J et al (2005) ZAP-70 methylation status is associated with ZAP-70 expression status in chronic lymphocytic leukemia. Haematologica 90 (8):1078–1088
100. Hegi ME, Diserens AC, Gorlia T et al (2005) MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med 352(10):997–1003
101. Strathdee G, MacKean MJ, Illand M, Brown R (1999) A role for methylation of the hMLH1 promoter in loss of hMLH1 expression and drug resistance in ovarian cancer. Oncogene 18:2335–2341
102. Gifford G, Paul J, Vasey PA, Kaye SB, Brown R (2004) The acquisition of hMLH1 methylation in plasma DNA after chemotherapy predicts poor survival for ovarian cancer patients. Clin Cancer Res 10(13):4420–4426
103. Plumb JA, Strathdee G, Sludden J, Kaye SB, Brown R (2000) Reversal of drug resistance in human tumor xenografts of 2′-deoxy-5-azacytidine-induced demethylation of the hMLH1 gene promoter. Cancer Res 60:6039–6044
104. Daskalakis M, Nguyen TT, Nguyen C et al (2002) Demethylation of a hypermethylated P15/INK4B gene in patients with myelodysplastic syndrome by 5-Aza-2′-deoxycytidine (decitabine) treatment. Blood 100(8):2957–2964
105. Garcia-Manero G, Kantarjian HM, Sanchez-Gonzalez B et al (2006) Phase I/II study of the combination of 5-aza-2′-deoxycytidine with valproic acid in patients with leukemia. Blood 108(10):3271–3279
106. Saunthararajah Y, Hillery CA, Lavelle D et al (2003) Effects of 5-aza-2′-deoxycytidine on fetal hemoglobin levels, red cell adhesion, and hematopoietic differentiation in patients with sickle cell disease. Blood 102(12):3865–3870
107. Laird PW, Jackson-Grusby L, Fazeli A et al (1995) Suppression of intestinal neoplasia by DNA hypomethylation. Cell 81(2):197–205
108. Opavsky R, Wang SH, Trikha P et al (2007) CpG island methylation in a mouse model of lymphoma is driven by the genetic configuration of tumor cells. PLoS Genet 3(9):1757–1769
109. Bailey VJ, Easwaran H, Zhang Y et al (2009) MS-qFRET: a quantum dot-based method for analysis of DNA methylation. Genome Res 19(8):1455–1461
110. Nakayama M, Bennett CJ, Hicks JL et al (2003) Hypermethylation of the human glutathione S-transferase-pi gene (GSTP1) CpG island is present in a subset of proliferative inflammatory atrophy lesions but not in normal or hyperplastic epithelium of the prostate: a detailed study using laser-capture microdissection. Am J Pathol 163(3):923–933
111. Goessl C, Krause H, Muller M et al (2000) Fluorescent methylation-specific polymerase chain reaction for DNA-based detection of prostate cancer in bodily fluids. Cancer Res 60 (21):5941–5945

112. Goessl C, Muller M, Miller K (2000) Methylation-specific PCR (MSP) for detection of tumour DNA in the blood plasma and serum of patients with prostate cancer. Prostate Cancer Prostatic Dis 3(S1):S17
113. Ogino S, Kawasaki T, Kirkner GJ, Kraft P, Loda M, Fuchs CS (2007) Evaluation of markers for CpG island methylator phenotype (CIMP) in colorectal cancer by a large population-based sample. J Mol Diagn 9(3):305–314
114. Esteller M, Gaidano G, Goodman SN et al (2002) Hypermethylation of the DNA repair gene O(6)-methylguanine DNA methyltransferase and survival of patients with diffuse large B-cell lymphoma. J Natl Cancer Inst 94(1):26–32
115. Hibi K, Taguchi M, Nakayama H et al (2001) Molecular detection of p16 promoter methylation in the serum of patients with esophageal squamous cell carcinoma. Clin Cancer Res 7(10):3135–3138
116. Feng Q, Balasubramanian A, Hawes SE et al (2005) Detection of hypermethylated genes in women with and without cervical neoplasia. J Natl Cancer Inst 97(4):273–282
117. Brock MV, Hooker CM, Ota-Machida E et al (2008) DNA methylation markers and early recurrence in stage I lung cancer. N Engl J Med 358(11):1118–1128
118. Belinsky SA, Palmisano WA, Gilliland FD et al (2002) Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. Cancer Res 62 (8):2370–2377
119. Wong IH, Lo YM, Zhang J et al (1999) Detection of aberrant p16 methylation in the plasma and serum of liver cancer patients. Cancer Res 59(1):71–73
120. Esteller M, Garcia-Foncillas J, Andion E et al (2000) Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. N Engl J Med 343 (19):1350–1354
121. Watanabe H, Okada G, Ohtsubo K et al (2006) Aberrant methylation of secreted apoptosis-related protein 2 (SARP2) in pure pancreatic juice in diagnosis of pancreatic neoplasms. Pancreas 32(4):382–389
122. Friedrich MG, Chandrasoma S, Siegmund KD et al (2005) Prognostic relevance of methylation markers in patients with non-muscle invasive bladder carcinoma. Eur J Cancer 41 (17):2769–2778
123. Lombaerts M, van Wezel T, Philippo K et al (2006) E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. Br J Cancer 94(5):661–671
124. Chiles MC, Ai L, Zuo C, Fan CY, Smoller BR (2003) E-cadherin promoter hypermethylation in preneoplastic and neoplastic skin lesions. Mod Pathol 16(10):1014–1018
125. Richiardi L, Fiano V, Vizzini L et al (2009) Promoter methylation in APC, RUNX3, and GSTP1 and mortality in prostate cancer patients. J Clin Oncol 27(19):3161–3168

DNA Repair and the Control of DNA Methylation

Primo Schär and Olivier Fritsch

Abstract The successful establishment and stable maintenance of cell identity are critical for organismal development and tissue homeostasis. Cell identity is provided by epigenetic mechanisms that facilitate a selective readout of the genome. Operating at the level of chromatin, they establish defined gene expression programs during cell differentiation. Among the epigenetic modifications in mammalian chromatin, the 5′-methylation of cytosine in CpG dinucleotides is unique in that it affects the DNA rather than histones and the biochemistry of the DNA methylating enzymes offers a mechanistic explanation for stable inheritance. Yet, DNA methylation states appear to be more dynamic and their maintenance more complex than existing models predict. Also, methylation patterns are by far not always faithfully inherited, as best exemplified by human cancers. Often, these show widespread hypo- or hypermethylation across their genomes, reflecting an underlying epigenetic instability that may have contributed to carcinogenesis. The phenotype of unstable methylation in cancer illustrates the importance of quality control in the DNA methylation system and implies the existence of proof-reading mechanisms that enforce fidelity to DNA methylation in healthy tissue. Fidelity seems particularly important in islands of unmethylated CpG-rich sequences where an accurate maintenance of un- or differentially methylated states is critical for stable expression of nearby genes. Methylation proof-reading in such sequences requires a system capable of recognition and active demethylation of erroneously methylated CpGs. Active demethylation of 5-methylcytosine has been known to occur for long, but the underlying mechanisms have remained enigmatic and controversial. However, recent progress in this direction substantiates a role of DNA repair in such processes. This review will address general aspects of cytosine methylation stability in mammalian DNA and explore a putative role of DNA repair in methylation control.

P. Schär and O. Fritsch
Institute of Biochemistry and Genetics, Department of Biomedicine, University of Basel, Mattenstrasse 28, CH-4058 Basel, Switzerland
e-mail: primo.schaer@unibas.ch

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_3,

1 Dynamic Stability of the "DNA Methylation Code"

DNA Methylation: The methylation of the 5′-position of cytosine (5-mC) in DNA, together with a variety of histone modifications, plays a central role in directing stable cell type-specific gene expression and suppressing transcriptional noise as well as transposon activity [1]. In mammals, this epigenetic DNA modification mainly concerns cytosine in the symmetrical CpG dinucleotide sequence, affecting approximately 60–90% of such sites. Isolated DNA islands with a higher than average content of CpG dinucleotides, so-called CpG islands (CGIs), are often present at the transcription start site of mammalian genes. Generally, such promoter CGIs are protected from cytosine methylation, although differential methylation can occur in a cell- or tissue-dependent manner (e.g., [2, 3]). CGI methylation correlates with the transcriptional activity of nearby genes, and aberrant hypermethylation is a frequent cause of improper gene silencing in human cancers [4]. DNA methylation patterns in mammalian genomes are thought to be established during early development and cell differentiation by the action of the *de novo* methyltransferases DNMT3A or DNMT3B (reviewed in [5, 6]). Assisted by DNMT3L, a methyltransferase-like protein, these enzymes interact with un- or hemimethylated CpGs to catalyze the methyl-transfer to cytosine, using S-Adenosyl L-methionine (SAM) as a methyl donor (reviewed in [7]). It is essential for embryonic development that this de novo methylation system discriminates between CpG sites that must be modified and sites that must not [8], but exactly how this is achieved is unknown. It seems increasingly clear though that the establishment of methylation patterns is linked with gene regulatory processes and may be instructed by histone marks (e.g., [9–12]). Histone modifications have been associated in various ways with DNA methylation and vice versa, but the hierarchical relationship between the two types of epigenetic modifications is complex and remains to be resolved. Both systems seem to be interlocked in a feedback loop that reinforces the maintenance of chromatin states (reviewed in [13]).

Once established, the mitotic inheritance of methylation patterns is assured by a maintenance system that copies the methylation signature from parental to daughter DNA strands during DNA replication and/or repair. Key to this system is the so-called maintenance methyltransferase, DNMT1, which shows a higher specificity for hemimethylated than for unmethylated DNA (reviewed in [7]). From a biochemical point of view, the establishment and the faithful inheritance of DNA methylation could thus be explained by a simple model based on the sequential action of DNMTs with de novo and maintenance methylation properties. However, the effect of DNMT defects on methylation patterns in mammalian cells indicates that the system is more complex and involves specific targeting of and cooperation between DNMTs (discussed in [6]). A newly emerging concept is therefore that, depending on the CpG density and chromatin state of a genomic region, CpG methylation may be maintained by both, a templated process involving DNMT1 and a stochastic process involving *de novo* methylation by DNMT3 enzymes.

In x this scenario, DNMT1 may be targeted to newly synthesized, hemimethylated DNA by interactions with PCNA [14, 15] and/or the SRA protein UHRF1/NP95 [16], whereas DNMT3 may associate more selectively with certain genomic regions through interactions with transcriptional silencers, such as EZH2 [10], or G9a [17] and HP1 [18].

1.1 Fidelity of the Methylation System

The fidelity of the methylation system is difficult to assess because of an unknown contribution of natural epigenetic plasticity in cell populations derived from tissue or grown in culture. A system involving DNMT1 and DNMT3, however, would predict that fidelity can be provided at two distinct levels. Faithful copying at hemimethylated CpGs would preserve sequence-specific patterns of methylation, whereas de novo methylation instructed by chromatin modifications would be suited to maintain methylation states across genomic regions. Indeed, data available on methylation fidelity indicate that the methylation status of individual CpGs is regulated dynamically and, to some extent, stochastically by both maintenance and de novo methylation systems in a cell type-, tissue- or chromosomal location-dependent manner [19–22]. As regards errors of maintenance methylation, an assessment of methylation instability in clonal populations of human cells indicated that maintenance errors occur with a higher rate in unmethylated than in methylated regions, and that unmethylated CGIs associated with gene promoters are protected more effectively from de novo methylation than unmethylated CGIs outside promoter regions [22]. These observations suggest that methylation errors are mainly due to de novo methylation in unmethylated DNA and that gene promoter-specific protection mechanism(s) may operate to protect CGIs from erroneous de novo methylation.

1.2 Dynamics of Methylation States

To what extent CpG methylation is a dynamic, i.e., reversible, epigenetic modification has been a matter of debate. In this regard, the widespread erasure of genomic DNA methylation observed under two developmental conditions in mouse, i.e., in differentiating primordial germ cells (PGCs) and in fertilized oocytes, may represent a paradigm. In PGCs, methylation clearance includes the CGIs of imprinted genes and therefore serves to reset maternal or paternal imprinting marks in the germ line. In fertilized oocytes, imprinted gene loci are unaffected but demethylation elsewhere is necessary for the reactivation of pluripotency genes that are silenced in germ cells but needed in the early stages of development. Genome wide demethylation in both situations also assures that a newly emerging organism develops without carry over of epigenetic traits from the previous generation.

Both active and passive mechanisms have been proposed to explain the kinetics of methylation clearance under these conditions, but the molecular pathways have remained obscure (reviewed in [23]).

Locally restricted changes in CpG methylation have also been observed, mostly upon nuclear- or hormone-receptor activation of gene promoters (e.g., [24–28, 29]), but here as well, the underlying mechanisms are just about to reveal themselves. Recently, two studies reported strand- and site-specific methylation–demethylation to occur in two estrogen responsive gene promoters (pS2/TFF1, Wisp-2) within minutes of exposure of human breast cancer cells to estradiol [30, 31]. With similar dynamics, cyclic methylation–demethylation was also observed at four estrogen-independent gene promoters following reactivation of doxorubicin-inhibited transcription [31]. Moreover, rapid methylation and demethylation was shown to accompany vitamin D receptor and parathyroid hormone-mediated repression and derepression of the CYP27B1 gene, respectively [32]. Remarkably, whereas the methylation activity was associated with DNMT1 and DNMT3 in all these cases, the demethylation steps were linked with an active process involving DNA repair (Fig. 1). Reproducibility provided the discovery of dynamic methylation–demethylation processes may lead to a paradigm-shift in how the establishment and the maintenance of CpG methylation have to be conceived. CpG methylation levels and patterns even within individual cells may turn out to reflect a steady-state much more than a stable condition.

2 Manifestation and Origin of DNA Methylation Instability

CpG methylation is established during early development and cell differentiation not only to direct cell lineage-specific gene expression and provide cell identity (reviewed in [33]), but also to suppress transposon- and recombinational activities. Hence, the faithful maintenance of methylation patterns is critical for tissue homeostasis and chromosomal stability (reviewed in [23, 34]). Aberrations in DNA methylation can be observed but they are usually associated with biological malfunctions in human diseases, such as cancer or imprinting disorders [35]. In many ways, cancers reflect a condition of inaccurate establishment or loss of cell identity [36–38] and, thus, illustrate best features and functional consequences of epigenetic instability, much like they do for genetic instability.

Aberrant CpG Methylation and Cancer: Alterations in the distribution and levels of 5-mC can be found in most, if not all cancers. Aberrations consist of both increases and decreases in the density of CpG methylation in distinct genomic regions [34, 39]. They are not only found in cancers but also in benign precursor lesions, suggesting that methylation errors arise early in tumorigenesis [40, 41]. Once occurred, such errors will be stabilized by the maintenance methylation system and eventually give rise to clonally inheritable changes in chromatin and gene expression states. This way, hypermethylation of promoter CGIs in cancers

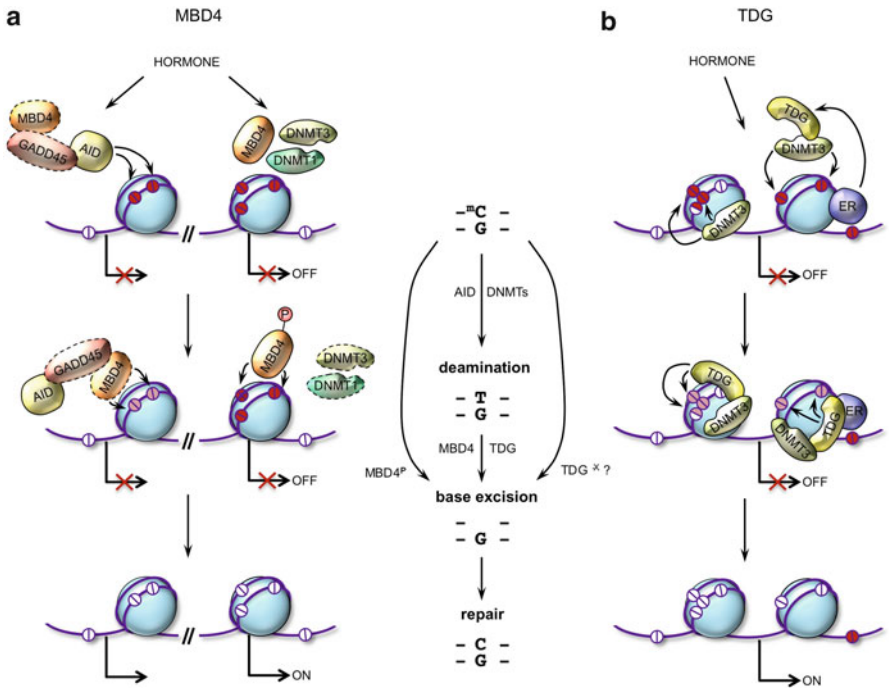


Fig. 1 DNA repair-mediated active DNA demethylation. Known processes of active cytosine demethylation involve either the MBD4 (**a**) or TDG (**b**) DNA glycosylases and BER as indicated in the *center*. (**a**) Active demethylation involving AID (*left*) is initiated by enzymatic deamination of 5-mC to T, generating a G·T mismatch. T is then excised by DNA glycosylases and repaired by BER to restore unmethylated G·C. This appears to be happening during global demethylation in zebrafish embryos, involving coupling of deamination (AID) and base excision (MDB4) through GADD45 [75]. GADD45 has also been associated with DNA repair-mediated demethylation in mammals [78, 79], although there are discrepant observations [80]. AID has also been implicated in global active demethylation in primordial germ cells [76] and during somatic cell reprogramming towards pluripotency [77]. During de-repression of the CYP27B1 gene by parathyroid hormone (*right*) [32], a complex containing DNMT1, DNMT3B and MBD4 binds to its promoter, which then undergoes MBD4-dependent active demethylation. The hormonal treatment induces phosphorylation of MBD4, then able to efficiently incise 5-mC. (**b**) Estradiol (E2) induced activation of estrogen-receptor target genes (TTF1/pS2, ERα) was shown to be accompanied by cyclic methylation–demethylation of promoter CpGs. Demethylation was proposed to result from primary deamination of 5-mC by DNMT3A/B [30, 31], both interacting physically with TDG [68, 69, 93]. The resulting G·T mismatch would then be processed by TDG (or MBD4) and repaired to unmethylated G·C establishing a transcriptionally active state. Alternatively, TDG may gain the ability to efficiently excise 5-mC through covalent modifications as MBD4 (*center*). *Dashed-line* outlines indicate uncertainty about the presence of the factor. DNA methylation status: *white*, unmethylated C; *red*, 5-mC; *pink*, deaminated 5-mC (T)

was shown to stably silence factors that control critical aspects of carcinogenesis, such as genomic stability (e.g., MLH1, MGMT, BRCA1), cell cycle control (e.g., Rb, p16), apoptosis (e.g., Caspase 8), and tumor cell invasion capacity (e.g., E-cadherin),

just to mention a few (reviewed in [42]). Hypomethylation, i.e., a decrease in DNA methylation density mostly in CpG-poor regions, seems even more prevalent in cancers. In promoter sequences, hypomethylation can lead to aberrant (over)expression of genes and therefore directly affect gene function [34]. DNA hypomethylation, however, is more often observed in repetitive DNA as well as in the vicinity of telomeres. There, the widespread loss of 5-mC may affect chromatin structure in a way that favors gross genomic instability [43] and/or interferes with telomere homeostasis and chromosome segregation (e.g., [44, 45]).

Biology of Aberrant Methylation: Little is known about the biological processes underlying the instability of DNA methylation in somatic tissue. Whereas in humans some promoter associated CGIs appear to undergo methylation as a function of age, others were found to do so exclusively in neoplastic tissue [46]. Whether these seemingly age- and cancer-specific processes reflect distinct pathways or a common underlying epigenetic instability that hits CGIs at random remains to be clarified. The finding of concurrent de novo methylation of multiple gene loci in a subset of colorectal cancers suggested that cells can acquire a "CpG island methylator phenotype" (CIMP), i.e., a condition characterized by an increased rate of CpG methylation errors [47–49]. Although the underlying molecular defects of CGI hypermethylation or CIMP are still obscure, it seems clear that in one way or another the integrity of the maintenance methylation system must be affected. This may concern the DNMTs themselves as in the case of the human ICF syndrome, which is caused by germline mutations in DNMT3B [50], or in mouse models with defects in DNMT genes (e.g., [8, 51]). Likewise, expression, RNA splicing, or targeting of DNMTs can be lost or imbalanced and the metabolic pathways providing the methyl donor SAM can be deregulated, all of which has been observed in cancer cells (e.g., [52–54]). Interestingly, however, a significant number of gene promoters known to be silenced by hypermethylation in cancers were shown to be controlled by the polycomb system in embryonic stem and tissue progenitor cells [36, 55, 56]. Polycomb repressive complexes (PRCs) establish long-term silencing of developmental genes, assuring stem cell maintenance and proper lineage commitment [57]. Key to polycomb-mediated silencing is the trimethylation of K27 in histone 3 (H3K27me3) catalyzed by the PRC2 component EZH1/2. Addition of this repressive mark on top of the activating mark H3K4me2/3 establishes a bivalent chromatin state that keeps the promoter silent but poised for reactivation, i.e., free of CpG methylation. Thus considered, it is remarkable that EZH2 also recruits DNMTs to PRC controlled promoters [10]. Although it is not clear whether this is a peculiarity of cancer cells only [56], the observation indicates that aberrant CGI methylation may reflect an inability of cells to correctly interpret histone modifications, i.e., to control DNMTs that happen to be recruited to CGI that are not supposed to be methylated.

Irrespective of the underlying mechanisms, cancer cells document that the fidelity of DNA methylation can be highly variable in both directions, toward gain and loss of methylation, and this implies a need for fidelity enforcing pathways in normal cells.

3 Enforcing Stability to DNA Methylation

DNA methylation turns out to be considerably more dynamic than originally thought and, yet, methylation patterns are relatively stably inherited. Faithful inheritance cannot be explained solely by the biochemistry of the DNMTs and is likely to require quality control mechanisms that enforce fidelity to their action.

Methylation Control: DNA replication needs proof-reading because DNA polymerases make errors. This ensures faithful inheritance of the "genetic code." The stable maintenance of the "epigenetic DNA code" is likely to involve some form of proof-reading as well, although the level of fidelity required may vary across different parts of the genome. In methylated regions, where chromatin is compact and little gene regulation occurs, the maintenance of a generally methylated state may be sufficient and patterns may be less important. Maintenance errors in such sequences would result in the occasional loss of a methylated site, which may be compensated for by ongoing de novo methylation elsewhere (Fig. 2a). In other regions, however, such as in promoter associated CGIs where methylation patterns confer biological function, the maintenance of un- or differentially methylated states at individual CpGs may be more critical. It seems unlikely that such patterns can be faithfully inherited purely on the basis of the biochemical properties of the DNMTs and their targeting factors, i.e., the selectivity of DNMT1/UHRF1 for hemimethylated CpGs [16, 58], or the site-specific recruitment and/or positioning of DNMT3 by histone modifications (e.g., [59]). As for errors of methyltransferases in unmethylated sequences, they would generate *de novo* methylated CpGs, and the restoration of the originally unmethylated state would require a system capable of actively demethylating DNA (Fig. 2b).

Active DNA Demethylation: Active demethylation has been known to occur for long, but the enzymatic processes involved in have remained enigmatic. Various mechanisms have been proposed, ranging from a direct enzymatic removal of the methyl group from 5-mC to the excision and replacement of the entire nucleotide. Direct removal of the methyl group would indeed represent a most straightforward mechanism, but evidence for such an activity of the MBD2 protein [60] could not be corroborated. Issues of reproducibility but also the lack of a plausible chemistry for such an enzymatic reaction have cast doubts over the concept.

The main body of current evidence points at a more complex scenario involving DNA excision repair. This appears most safely established in plants, where genetic and biochemical data consistently support a function of the DNA glycosylases ROS1 and DEMETER in the base excision repair (BER) of 5-mC [61, 62]. In vertebrates, the evidence is less clear, although two DNA glycosylases, TDG and MBD4, have been implicated in similar processes. DNA glycosylases are small DNA repair proteins capable of detecting and excising chemically modified DNA bases [63], thereby inducing a BER process that leads to the replacement of the modified with an unmodified base [64]. Thus, DNA glycosylases seem optimally suited to recognize erroneously methylated CpGs and to trigger their replacement with an unmethylated cytosine. Both TDG and MBD4 were originally identified as

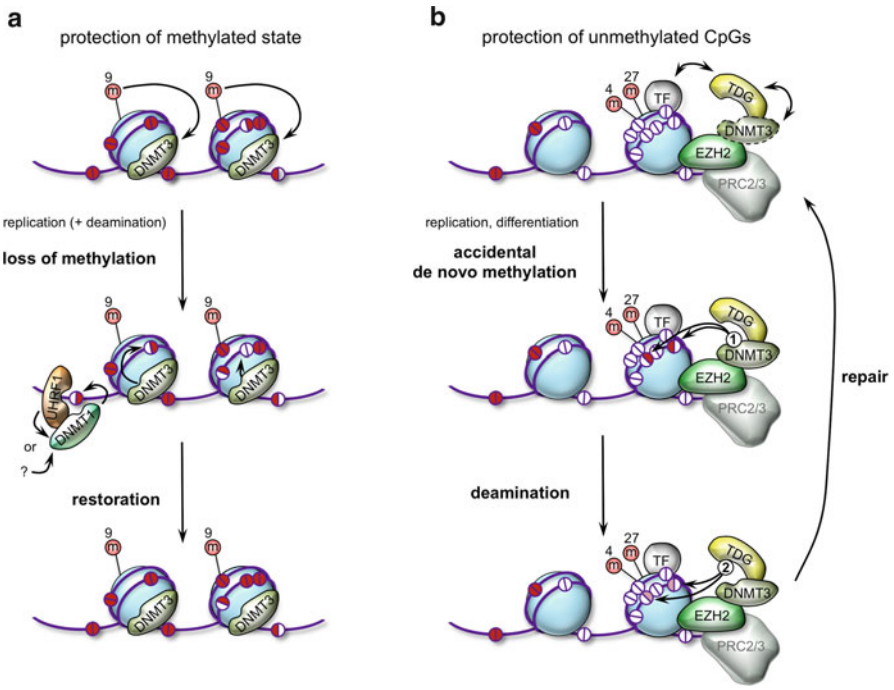


Fig. 2 Protection against DNA methylation instability. DNA methylation instability is a concern in both methylated (**a**) and unmethylated (**b**) regions. (**a**) Methylated regions need to maintain their methylated state across rounds of replication and spontaneous deamination events. Although the bulk of methylation is maintained by DNMT1 during replication, errors may occur, leaving many CpGs in a hemimethylated state. Hemimethylated CpGs may also arise upon hydroxylation of 5-mC to 5-hmC [82, 83], which will only poorly be recognized by DNMT1 [96]. Consequently, observations that DNMT3A/B are recruited to methylated DNA [56, 97] provide a context for the maintenance of methylation levels in such regions, rather than exact patterns, through ongoing de novo methylation. DNMT1 may contribute to proof-reading in this context as well, most likely guided by UHRF1 [58, 98, 99]. (**b**) Unmethylated CpGs in CGIs must be protected against unwanted methylation. We consider here promoter CGIs of polycomb target genes bound by EZH2 in a PRC2/3 context. Such targets, initially in a bivalent silent state (H3K4 and H3K27 methylated) as in ES cells, need to remain free of DNA methylation, but are prone to acquire stochastic DNMT3-mediated CpG methylation. CGIs in bivalent chromatin are bound by TDG as long as they are unmethylated and an activating histone mark (H3K4me) is present, and the loss of TDG is associated with increased aberrant de novo methylation at such sites (P. Schär, unpublished observations). We therefore propose that TDG, in cooperation with a 5-mC deaminase activity (DNMT, AID/Apobec), removes such erroneous methylation through BER, thereby maintaining an unmethylated state. The glycosylase may be targeted to such sites through its physical interactions with transcription factors (TF) or DNMTs

mismatch-specific DNA glycosylases removing thymine and uracil from G·T and G·U mismatches, respectively [65–67]. Remarkably, both were also shown to complex with DNMTs in cells [32, 68, 69] and to excise 5-mC from DNA, albeit with very poor efficiency [32, 70, 71] (Fig. 1). Recent work, showing that MDB4

gains robust 5-mC glycosylase activity upon phosphorylation by protein-kinase-C [32] (Fig. 1a, right), however, indicated that the lack of appropriate posttranslational modifications may have impeded the reproducible measurement of such activity.

Nonetheless, another and perhaps more plausible scenario for an engagement of G·T glycosylases in DNA demethylation is through prior deamination of 5-mC to T, which generates a G·T substrate in DNA. For TDG, such a pathway has been proposed to explain its contribution to cyclic methylation/demethylation following estrogen stimulated gene activation (Fig. 1b) [30]. In this case, the conversion of 5-mC to T was associated with the DNMT activity itself, which, under conditions of limiting supply of the methyl-donor SAM, appeared to turn into a deaminase, as previously reported for bacterial DNA methyltransferases [72, 73]. A similar pathway was proposed to explain active DNA demethylation by MBD4 in zebrafish embryos. There, however, cytidine deaminases of the AID/Apobec-type [74] were implicated in the deamination step and the "growth arrest and DNA-damage-inducible protein 45alpha" (Gadd45α) as a factor coordinating deamination with MBD4-mediated base excision (Fig. 1a, left) [75]. Consistently, AID was also shown to contribute to genome-wide erasure of DNA methylation in mouse PGCs [76] and to rapid promoter demethylation during reprogramming toward pluripotency of somatic cell nuclei [77]. Gadd45α, on the other hand, had previously been associated with a DNA repair process mediating demethylation of 5-mC in *Xenopus laevis* oocytes and mammalian cells [78, 79], although the circumstances under which this pathway is active remain to be clarified [80, 81].

Finally, the recent discovery of 5-hydroxymethylcytosine (5-hmC) as a prominent sixth base in the DNA of mammalian cells [82] suggested yet another potential pathway for active demethylation. 2-oxoglutarate- and Fe(II)-dependent hydroxylases like TET1, a member of the TET oncogene family, may convert 5-mC into 5-hmC [83] to generate a substrate for an as yet unspecified 5-hmC DNA glycosylase [84]. As described for the deamination pathways (Fig. 1), this glycosylase may then excise the modified base and initiate BER to insert an unmethylated C. Thus, although more firmly established in plants, the concept of DNA repair-mediated cytosine demethylation finds increasing support in experimental evidence also for vertebrate systems (Table 1).

DNA Repair-Mediated Methylation Proof-Reading: Given the capacity of cells to actively excise 5-mC from DNA, the question arises to what extent such mechanisms contribute to the fidelity of the methylation system. Newly emerging experimental evidence leads us to speculate that DNA repair systems may not only counteract genetic instability but also epigenetic instability by proof-reading methylation marks and correcting occasional DNMT errors (Fig. 2b). Methylation proofreading may be particularly important during DNA replication and repair where DNMT1 maintains patterns at sites of DNA synthesis, as well as in the context of gene regulation and/or cell differentiation where DNMT3s seem to induce dynamic methylation changes in gene promoters (Fig. 1). Common to both these situations is that DNMTs are likely to associate with CpG-rich sequences that are not supposed to be methylated and therefore need to be protected. Numerous interactions

Table 1 Evidence for an association between factors involved in DNA repair and DNA methylation processes

Factor	Link	Evidence	Context	Reference
TDG	DNMT3A	Physical interaction, co-immunoperecipitation	*In vitro* and 293T cells	[68]
	DNMT3A	Association *in vivo*	Vitronectin-dependent gene activation	[69]
	DNMT3B	Affinity capture, reconstituted protein complex	*In vitro*	[93]
	DNMT3A, DNMT3B	Co-occupancy at sites of active demethylation	Estrogen-dependent gene activation	[30]
	DNMT3A, DNMT3B	Co-occupancy at sites of active demethylation	Reactivation of methylated promoters	[31]
	5^{m}C-glycosylase activity	Mismatch specific T-glycosylase activity copurified with a 5th C-glycosylase activity	*In vitro* with chicken embryo extracts	[100]
	5^{m}C-glycosylase activity	Weak 5th C-glycosylase activity as compared to mismatch T	*In vitro*	[87]
	phenotypic	Knock-down results in reduced genome-wide demethylation	Differentiating mouse fibroblasts	[101]
MND4	DNMT3B	Reconstituted complex	*In vitro*	[93]
	DNMT1, DNMT3B	Co-immunoprecipitation	Hormone-dependent DNA demethylation	[32]
	5^{m}C-glycosylase activity	Weak 5mC-glycosylase activity	*In vitro*	[70, 71]
	5^{m}C-glycosylase activity	Phosphorylated MBD4 incises efficiently 5^{m}C	*In vitro*	[32]
	phenotypic	Altered genome methylation upon MBDA knock-down and AID/MBD4 overexpression	Zebrafish embryos	[75]
AID	5^{m}C-deamination	5^{m}C-deamination in ssDNA	*In vitro*	[74]
	5^{m}C-deamination	5^{m}C-deamination of injected DNA	*In vivo*, injection in zebrafish embryos	[75]
	phenotypic	Altered genome methylation upon AID/MBD4 overexpression or AID knock-down	Zebrafish embryos	[75]
	phenotypic	Required for active demethylation at OCT4 and NANOG promoters	ES cells	[77]
	phenotypic	Impaired erasure of methylation in AID deficient cells	Primordial germ cells	[76]
GADD45α	phenotypic	Knock-down and overexpression alter global methylation levels	*Xenopus leavis* oocytes, human cells	[78]
	phenotypic	Knock-down and overexpression alter rDNA methylation levels, early recruitment during demethylation	Human cells	[79]

between BER proteins and the methylation system have been reported, arguing for a functional link between the two processes (Table 1), and recent work with TDG-deficient cell lines and mice in our laboratory seems to corroborate a role for DNA repair in methylation control and provide insight into a possible pathway.

Unlike other DNA glycosylases [64], TDG turned out be essential for mouse embryonic development (P. Schär, unpublished observations). Given the physical and functional interactions of TDG with nuclear receptors, DNMTs, and histone acetyl transferases, this developmental phenotype is more likely related to a defect in gene regulation than in DNA repair [68, 69, 85, 86]. Indeed, gene-expression profiling revealed dramatic differences between TDG proficient and deficient mouse embryonic fibroblasts (MEFs). Developmental factors (e.g., HoxA and HoxD genes) were significantly enriched amongst the differentially expressed genes, most of them having CGIs in their promoters and being targets of the polycomb repressive system at a certain developmental state. We also found TDG to bind with high preference to the promoter CGIs of such genes as long as they remain unmethylated and carry the activating histone mark H3K4me2. The loss of TDG was associated with increased aberrant de novo methylation of promoter CGIs in differentiated cells (MEFs), while global DNA methylation was largely unaffected, and this correlated with an increase of repressive histone methylation (H3K27me3 and H3K9me2/3) and a loss of H3K4 methylation. Whether TDG primarily controls histone or the DNA modification states remains to be clarified. Its biochemical properties, however, and the observation that XRCC1, an essential component of the BER system, localizes to the same CGIs in a TDG-dependent manner support the hypothesis that TDG-dependent BER operates to maintain CGIs in an unmethylated state (P. Schär, unpublished observations).

How can this be envisaged mechanistically? TDG has been associated in various ways with active demethylation [30, 31, 87], and it was repeatedly shown to be targeted to gene regulatory sequences through interactions with transcription factors [85, 86, 88–90]. Gene regulatory sequences also bind DNMT3A/B, either as part of transcription factor complexes [69, 91, 92] or upon recruitment to bivalent chromatin through interaction with the PRC2 component EZH2 [10]. Hence, there seem to be conditions where de novo methyltransferases associate with gene regulatory sequences that must not be methylated, such as promoter CGIs in a bivalent chromatin state, and TDG may be present to control their activity [68, 69, 93]. Figure 2b illustrates the situation of a CGI associated with a promoter of a developmental gene. The promoter is kept silent by the polycomb repressive system but must be protected from *de novo* DNA methylation to assure reactivation at a later stage in development. In this situation, DNMT3 can bind to the PRC2 component EZH2 and accidentally methylate a CpG. TDG may then recognize the erroneous 5-mC and initiate excision repair that restores the unmethylated state. It may do so on its own, provided some form of posttranslational modification or allosteric activation stimulates its latent 5-mC glycosylase activity, as reported for MBD4. More likely, though, TDG may act in concert with a 5-mC deaminase, such as DNMT3 itself, or an AID/Apobec-type of activity as discussed before (Fig. 1). While such a mechanism for methylation proof-reading at CGIs is consistent with

the phenotype of TDG deficient cells, it remains hypothetical until important details like the targeting of TDG to CGIs and the biochemistry of 5-mC processing will be clarified.

MBD4 has enzymatic properties comparable to those of TDG [67] and could contribute in a similar way to methylation proof-reading (Fig. 1a). Unlike TDG, however, MBD4 appears to be dispensable for epigenetic programming during cell differentiation and development. MBD4 knockout mice do not show a developmental phenotype or other signs of DNA methylation instability [94]. Nevertheless, MBD4 has been associated with active demethylation in at least two situations [32, 75]. Both concern initially methylated rather than unmethylated sequences, to which MBD4 may be targeted through its methylated DNA-binding domain [95]. By contrast, we found TDG to be targeted preferentially to unmethylated CGIs, presumably through its interaction with transcription factor complexes (P. Schär, unpublished observations). This considered, it is tempting to speculate that the two G·T glycosylases serve complementary epigenetic purposes. While TDG appears to preserve the unmethylated state of CGIs, MBD4 may function in demethylation of methylated sequences, mainly in the context of promoter derepression. Hence, TDG- and MBD4-dependent pathways would operate in different genomic contexts, the first to maintain and the second to alter methylation patterns.

4 Concluding Remark

Considering the devastating consequences of DNA methylation instability for mammalian development and tissue homeostasis, it seems likely that proof-reading mechanisms exist to ensure faithful inheritance of the epigenetic DNA code. Such a function may be important mainly in regions where defined methylation patterns fulfill biological function, such as in CGIs that modulate gene promoter activity and are often unmethylated. The maintenance of unmethylated states would require a system capable of recognizing and actively demethylating erroneously methylated CpGs. Active 5-mC demethylation in mammals has been controversial but an increasingly robust body of experimental evidence implicates DNA repair in such a process. Given their ability to recognize and excise small base alterations, DNA glycosylases seem to be destined as proof-readers of methylation errors, and BER would be optimally suited to restore unmethylated CpGs. Protection from de novo methylation appears to be particularly important during the transition of epigenetic states associated with cell differentiation, which may explain the developmental defect of TDG deficient embryos. It is thus plausible that BER not only provides genetic stability but also contributes to epigenetic stability. Although speculative at this point, the predictions of such a model are clear and warrant careful investigation.

References

1. Jones PA, Takai D (2001) The role of DNA methylation in mammalian epigenetics. Science 293:1068–1070
2. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 39:457–466
3. Illingworth RS, Bird AP (2009) CpG islands – 'a rough guide'. FEBS Lett 583:1713–1720
4. Gronbaek K, Hother C, Jones PA (2007) Epigenetic changes in cancer. APMIS 115:1039–1059
5. Reik W, Dean W, Walter J (2001) Epigenetic reprogramming in mammalian development. Science 293:1089–1093
6. Jones PA, Liang G (2009) Rethinking how DNA methylation patterns are maintained. Nat Rev Genet 10:805–811
7. Jeltsch A (2006) Molecular enzymology of mammalian DNA methyltransferases. Curr Top Microbiol Immunol 301:203–225
8. Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99:247–257
9. Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y, Cedar H, Bergman Y (2006) G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat Cell Biol 8:188–194
10. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM, Bollen M, Esteller M, Di Croce L, de Launoit Y, Fuks F (2006) The Polycomb group protein EZH2 directly controls DNA methylation. Nature 439:871–874
11. Hu JL, Zhou BO, Zhang RR, Zhang KL, Zhou JQ, Xu GL (2009) The N-terminus of histone H3 is required for de novo DNA methylation in chromatin. Proc Natl Acad Sci USA 106 (52):22187–22192
12. Jelinic P, Stehle JC, Shaw P (2006) The testis-specific factor CTCFL cooperates with the protein methyltransferase PRMT7 in H19 imprinting control region methylation. PLoS Biol 4:e355
13. Klose RJ, Bird AP (2006) Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci 31:89–97
14. Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF (1997) Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science 277:1996–2000
15. Mortusewicz O, Schermelleh L, Walter J, Cardoso MC, Leonhardt H (2005) Recruitment of DNA methyltransferase I to DNA repair sites. Proc Natl Acad Sci USA 102:8905–8909
16. Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M, Koseki H (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 450:908–912
17. Dong KB, Maksakova IA, Mohn F, Leung D, Appanah R, Lee S, Yang HW, Lam LL, Mager DL, Schubeler D, Tachibana M, Shinkai Y, Lorincz MC (2008) DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. EMBO J 27:2691–2701
18. Fuks F, Hurd PJ, Deplus R, Kouzarides T (2003) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Res 31:2305–2312
19. Turker MS, Swisshelm K, Smith AC, Martin GM (1989) A partial methylation profile for a CpG site is stably maintained in mammalian tissues and cultured cell lines. J Biol Chem 264:11632–11636
20. Pfeifer GP, Steigerwald SD, Hansen RS, Gartler SM, Riggs AD (1990) Polymerase chain reaction-aided genomic sequencing of an X chromosome-linked CpG island: methylation

patterns suggest clonal inheritance, CpG site autonomy, and an explanation of activity state stability. Proc Natl Acad Sci USA 87:8252–8256
21. Riggs AD, Xiong Z, Wang L, LeBon JM (1998) Methylation dynamics, epigenetic fidelity and X chromosome structure. Novartis Found Symp 214:214–225, discussion 225-232
22. Ushijima T, Watanabe N, Okochi E, Kaneda A, Sugimura T, Miyamoto K (2003) Fidelity of the methylation pattern and its variation in the genome. Genome Res 13:868–874
23. Reik W (2007) Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 447:425–432
24. Wilks AF, Cozens PJ, Mattaj IW, Jost JP (1982) Estrogen induces a demethylation at the 5′ end region of the chicken vitellogenin gene. Proc Natl Acad Sci USA 79:4252–4555
25. Saluz HP, Jiricny J, Jost JP (1986) Genomic sequencing reveals a positive correlation between the kinetics of strand-specific DNA demethylation of the overlapping estradiol/glucocorticoid-receptor binding sites and the rate of avian vitellogenin mRNA synthesis. Proc Natl Acad Sci USA 83:7167–7171
26. Jost JP, Saluz HP, Pawlak A (1991) Estradiol down regulates the binding activity of an avian vitellogenin gene repressor (MDBP-2) and triggers a gradual demethylation of the mCpG pair of its DNA binding site. Nucleic Acids Res 19:5771–5775
27. Lucarelli M, Fuso A, Strom R, Scarpa S (2001) The dynamics of myogenin site-specific demethylation is strongly correlated with its expression and with muscle differentiation. J Biol Chem 276:7500–7506
28. Thomassin H, Flavin M, Espinas ML, Grange T (2001) Glucocorticoid-induced DNA demethylation and gene memory during development. EMBO J 20:1974–1983
29 . Kress C, Thomassin H, Grange T (2006) Active cytosine demethylation triggered by a nuclear receptor involves DNA strand breaks. Proc Natl Acad Sci USA 103:11112–11117
30. Metivier R, Gallais R, Tiffoche C, Le Peron C, Jurkowska RZ, Carmouche RP, Ibberson D, Barath P, Demay F, Reid G, Benes V, Jeltsch A, Gannon F, Salbert G (2008) Cyclical DNA methylation of a transcriptionally active promoter. Nature 452:45–50
31. Kangaspeska S, Stride B, Metivier R, Polycarpou-Schwarz M, Ibberson D, Carmouche RP, Benes V, Gannon F, Reid G (2008) Transient cyclical methylation of promoter DNA. Nature 452:112–115
32. Kim MS, Kondo T, Takada I, Youn MY, Yamamoto Y, Takahashi S, Matsumoto T, Fujiyama S, Shirode Y, Yamaoka I, Kitagawa H, Takeyama K, Shibuya H, Ohtake F, Kato S (2009) DNA demethylation in hormone-induced transcriptional derepression. Nature 461:1007–1012
33. Mohn F, Schubeler D (2009) Genetics and epigenetics: stability and plasticity during cellular differentiation. Trends Genet 25:129–136
34. Ehrlich M (2002) DNA methylation in cancer: too much, but also too little. Oncogene 21:5400–5413
35. Robertson KD (2005) DNA methylation and human disease. Nat Rev Genet 6:597–610
36. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, Weisenberger DJ, Campan M, Young J, Jacobs I, Laird PW (2007) Epigenetic stem cell signature in cancer. Nat Genet 39:157–158
37. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, Weinberg RA (2008) An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nat Genet 40:499–507
38. Gupta PB, Chaffer CL, Weinberg RA (2009) Cancer stem cells: mirage or reality? Nat Med 15:1010–1012
39. Baylin SB, Esteller M, Rountree MR, Bachman KE, Schuebel K, Herman JG (2001) Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. Hum Mol Genet 10:687–692
40. Chan AO, Rashid A (2006) CpG island methylation in precursors of gastrointestinal malignancies. Curr Mol Med 6:401–408

41. Menigatti M, Truninger K, Gebbers JO, Marbet U, Marra G, Schar P (2009) Normal colorectal mucosa exhibits sex- and segment-specific susceptibility to DNA methylation at the hMLH1 and MGMT promoters. Oncogene 28:899–909
42. Momparler RL (2003) Cancer epigenetics. Oncogene 22:6479–6483
43. Eden A, Gaudet F, Waghmare A, Jaenisch R (2003) Chromosomal instability and tumors promoted by DNA hypomethylation. Science 300:455
44. Rauch TA, Zhong X, Wu X, Wang M, Kernstine KH, Wang Z, Riggs AD, Pfeifer GP (2008) High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. Proc Natl Acad Sci USA 105:252–257
45. Vera E, Canela A, Fraga MF, Esteller M, Blasco MA (2008) Epigenetic regulation of telomeres in human cancer. Oncogene 27:6817–6833
46. Toyota M, Issa JP (1999) CpG island methylator phenotypes in aging and cancer. Semin Cancer Biol 9:349–357
47. Toyota M, Ahuja N, Suzuki H, Itoh F, Ohe-Toyota M, Imai K, Baylin SB, Issa JP (1999) Aberrant methylation in gastric cancer associated with the CpG island methylator phenotype. Cancer Res 59:5438–5442
48. Ogino S, Cantor M, Kawasaki T, Brahmandam M, Kirkner GJ, Weisenberger DJ, Campan M, Laird PW, Loda M, Fuchs CS (2006) CpG island methylator phenotype (CIMP) of colorectal cancer is best characterised by quantitative DNA methylation analysis and prospective cohort studies. Gut 55:1000–1006
49. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, Kang GH, Widschwendter M, Weener D, Buchanan D, Koh H, Simms L, Barker M, Leggett B, Levine J, Kim M, French AJ, Thibodeau SN, Jass J, Haile R, Laird PW (2006) CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 38:787–793
50. Ehrlich M (2003) The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. Clin Immunol 109:17–28
51. 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
52. Saito Y, Kanai Y, Sakamoto M, Saito H, Ishii H, Hirohashi S (2002) Overexpression of a splice variant of DNA methyltransferase 3b, DNMT3b4, associated with DNA hypomethylation on pericentromeric satellite regions during human hepatocarcinogenesis. Proc Natl Acad Sci USA 99:10060–10065
53. Gopalakrishnan S, Van Emburgh BO, Shan J, Su Z, Fields CR, Vieweg J, Hamazaki T, Schwartz PH, Terada N, Robertson KD (2009) A novel DNMT3B splice variant expressed in tumor and pluripotent cells modulates genomic DNA methylation patterns and displays altered DNA binding. Mol Cancer Res 7:1622–1634
54. Fang JY, Xiao SD (2003) Folic acid, polymorphism of methyl-group metabolism genes, and DNA methylation in relation to GI carcinogenesis. J Gastroenterol 38:821–829
55. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, Mohammad HP, Chen W, Daniel VC, Yu W, Berman DM, Jenuwein T, Pruitt K, Sharkis SJ, Watkins DN, Herman JG, Baylin SB (2007) A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 39:237–242
56. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, Eden E, Yakhini Z, Ben-Shushan E, Reubinoff BE, Bergman Y, Simon I, Cedar H (2007) Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet 39:232–236
57. Ringrose L, Paro R (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu Rev Genet 38:413–443
58. Arita K, Ariyoshi M, Tochio H, Nakamura Y, Shirakawa M (2008) Recognition of hemimethylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature 455:818–821

59. Ciccone DN, Su H, Hevi S, Gay F, Lei H, Bajko J, Xu G, Li E, Chen T (2009) KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. Nature 461:415–418
60. Bhattacharya SK, Ramchandani S, Cervoni N, Szyf M (1999) A mammalian protein with specific demethylase activity for mCpG DNA. Nature 397:579–583
61. Gong Z, Morales-Ruiz T, Ariza RR, Roldan-Arjona T, David L, Zhu JK (2002) ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. Cell 111:803–814
62. Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL (2002) DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. Cell 110:33–42
63. Schärer OD, Jiricny J (2001) Recent progress in the biology, chemistry and structural biology of DNA glycosylases. BioEssays 23:270–281
64. Robertson AB, Klungland A, Rognes T, Leiros I (2009) DNA repair in mammalian cells: base excision repair: the long and short of it. Cell Mol Life Sci 66:981–993
65. Neddermann P, Jiricny J (1993) The purification of a mismatch-specific thymine-DNA glycosylase from HeLa cells. J Biol Chem 268:21218–21224
66. Bellacosa A, Cicchillitti L, Schepis F, Riccio A, Yeung AT, Matsumoto Y, Golemis EA, Genuardi M, Neri G (1999) MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. Proc Natl Acad Sci USA 96:3969–3974
67. Hendrich B, Hardeland U, Ng HH, Jiricny J, Bird A (1999) The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. Nature 401:301–304
68. Li YQ, Zhou PZ, Zheng XD, Walsh CP, Xu GL (2007) Association of Dnmt3a and thymine DNA glycosylase links DNA methylation with base-excision repair. Nucleic Acids Res 35:390–400
69. Gallais R, Demay F, Barath P, Finot L, Jurkowska R, Le Guevel R, Gay F, Jeltsch A, Metivier R, Salbert G (2007) Dnmt 3a and 3b associate with the nuclear orphan receptor COUP-TFI during gene activation. Mol Endocrinol 21:2085–2098
70. Zhu B, Zheng Y, Hess D, Angliker H, Schwarz S, Siegmann M, Thiry S, Jost JP (2000) 5-methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. Proc Natl Acad Sci USA 97:5135–5139
71. Zhu B, Zheng Y, Angliker H, Schwarz S, Thiry S, Siegmann M, Jost JP (2000) 5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence. Nucleic Acids Res 28:4157–4165
72. Shen JC, Zingg JM, Yang AS, Schmutte C, Jones PA (1995) A mutant HpaII methyltransferase functions as a mutator enzyme. Nucleic Acids Res 23:4275–4282
73. Yebra MJ, Bhagwat AS (1995) A cytosine methyltransferase converts 5-methylcytosine in DNA to thymine. Biochemistry 34:14752–14757
74. Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK (2004) Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. J Biol Chem 279:52353–52360
75. 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
76. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE, Reik W (2010) Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature doi:10.1038/nature08829
77. Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM (2009) Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature doi:10.1038/nature08752

78. Barreto G, Schafer A, Marhold J, Stach D, Swaminathan SK, Handa V, Doderlein G, Maltry N, Wu W, Lyko F, Niehrs C (2007) Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature 445:671–675
79. Schmitz KM, Schmitt N, Hoffmann-Rohrer U, Schafer A, Grummt I, Mayer C (2009) TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. Mol Cell 33:344–353
80. Jin SG, Guo C, Pfeifer GP (2008) GADD45A does not promote DNA demethylation. PLoS Genet 4:e1000013
81. Engel N, Tront JS, Erinle T, Nguyen N, Latham KE, Sapienza C, Hoffman B, Liebermann DA (2009) Conserved DNA methylation in Gadd45a(−/−) mice. Epigenetics 4:98–99
82. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324:929–930
83. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324:930–935
84. Cannon SV, Cummings A, Teebor GW (1988) 5-Hydroxymethylcytosine DNA glycosylase activity in mammalian tissue. Biochem Biophys Res Commun 151:1173–1179
85. Um S, Harbers M, Benecke A, Pierrat B, Losson R, Chambon P (1998) Retinoic acid receptors interact physically and functionally with the T:G mismatch-specific thymine-DNA glycosylase. J Biol Chem 273:20728–20736
86. Tini M, Benecke A, Um SJ, Torchia J, Evans RM, Chambon P (2002) Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. Mol Cell 9:265–277
87. Cortazar D, Kunz C, Saito Y, Steinacher R, Schär P (2007) The enigmatic thymine DNA glycosylase. DNA Repair (Amst) 6:489–504
88. Missero C, Pirro MT, Simeone S, Pischetola M, Di Lauro R (2001) The DNA glycosylase T:G mismatch-specific thymine DNA glycosylase represses thyroid transcription factor-1-activated transcription. J Biol Chem 276:33569–33575
89. Chen D, Lucey MJ, Phoenix F, Lopez-Garcia J, Hart SM, Losson R, Buluwela L, Coombes RC, Chambon P, Schär P, Ali S (2003) T:G mismatch-specific thymine-DNA glycosylase potentiates transcription of estrogen-regulated genes through direct interaction with estrogen receptor alpha. J Biol Chem 278:38586–38592
90. Lucey MJ, Chen D, Lopez-Garcia J, Hart SM, Phoenix F, Al-Jehani R, Alao JP, White R, Kindle KB, Losson R, Chambon P, Parker MG, Schär P, Heery DM, Buluwela L, Ali S (2005) T:G mismatch-specific thymine-DNA glycosylase (TDG) as a coregulator of transcription interacts with SRC1 family members through a novel tyrosine repeat motif. Nucleic Acids Res 33:6393–6404
91. Sato N, Kondo M, Arai K (2006) The orphan nuclear receptor GCNF recruits DNA methyltransferase for Oct-3/4 silencing. Biochem Biophys Res Commun 344:845–851
92. Hervouet E, Vallette FM, Cartron PF (2009) Dnmt3/transcription factor interactions as crucial players in targeted DNA methylation. Epigenetics 4:487–499
93. Boland MJ, Christman JK (2008) Characterization of Dnmt3b:thymine-DNA glycosylase interaction and stimulation of thymine glycosylase-mediated repair by DNA methyltransferase(s) and RNA. J Mol Biol 379:492–504
94. Millar CB, Guy J, Sansom OJ, Selfridge J, MacDougall E, Hendrich B, Keightley PD, Bishop SM, Clarke AR, Bird A (2002) Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. Science 297:403–405
95. Hendrich B, Bird A (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol 18:6538–6547
96. Valinluck V, Sowers LC (2007) Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. Cancer Res 67:946–950

97. Jeong S, Liang G, Sharma S, Lin JC, Choi SH, Han H, Yoo CB, Egger G, Yang AS, Jones PA (2009) Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. Mol Cell Biol 29:5366–5376
98. Liang G, Chan MF, Tomigahara Y, Tsai YC, Gonzales FA, Li E, Laird PW, Jones PA (2002) Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. Mol Cell Biol 22:480–491
99. Schermelleh L, Haemmer A, Spada F, Rosing N, Meilinger D, Rothbauer U, Cardoso MC, Leonhardt H (2007) Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. Nucleic Acids Res 35:4301–4312
100. Jost JP, Siegmann M, Sun L, Leung R (1995) Mechanisms of DNA demethylation in chicken embryos. J Biol Chem 270:9734–9739
101. Jost JP, Oakeley EJ, Zhu B, Benjamin D, Thiry S, Siegmann M, Jost Y (2001) 5-Methylcytosine DNA glycosylase participates in the genome-wide loss of DNA methylation occurring during mouse myoblast differentiation. Nucleic Acids Res 29:4452–4461

Errors in Erasure: Links Between Histone Lysine Methylation Removal and Disease

Elizabeth M. Duncan and C. David Allis

Abstract Many studies have demonstrated that covalent histone modifications are dynamically regulated to cause both chemical and physical changes to the chromatin template. Such changes in the chromatin template lead to biologically significant consequences, including differential gene expression. Histone lysine methylation, in particular, has been shown to correlate with gene expression both positively and negatively, depending on the specific site and degree (i.e., mono-, di-, or tri-) of methylation within the histone sequence. Although genetic alterations in the proteins that establish, or "write," methyl modifications and their effect in various human pathologies have been documented, connections between the misregulation of proteins that remove, or "erase," histone methylation and disease have emerged more recently. Here we discuss three mechanisms through which histone methylation can be removed from the chromatin template. We describe how these "erasure" mechanisms are linked to pathways that are known to be misregulated in diseases, such as cancer. We further describe how errors in the removal of histone methylation can and do lead to human pathologies, both directly and indirectly.

1 Introduction

In vivo, the genome of a cell is packaged into a three-dimensional structure known as chromatin: DNA wraps around a core of histone proteins to form a chain of repeating units or "beads on a string" [1]. This association of genomic DNA with histone proteins provides a framework for "epigenetic" phenotypes, which are defined as heritable phenotypes that are not caused by differences in DNA sequence [2]. Although histone proteins were long regarded as an inert scaffold for the

E.M. Duncan and C.D. Allis
Department of Neurobiology and Anatomy, University of Utah, 20 North 1900 East, Salt Lake City, UT 84132
e-mail:alliscd@rockefeller.edu

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_4,

genome, a wealth of more recent studies shows that they are dynamically regulated, both chemically and physically, in ways that lead to functional change. These pathways of chromatin regulation establish, maintain, and propagate different patterns of gene expression during normal tissue development and differentiation [2]; it then follows that errors made in the regulation of the "epigenome" can lead to pathological situations, such as tumorigenesis.

One of the fundamental mechanisms used to regulate the accessibility of the chromatin template is the posttranslational modification of histone proteins, such as histone methylation [3–7]. This covalent modification has been mapped to both lysine and arginine residues within all histone species; predictably, most sites of methylation are found within the histone tails (as compared with the less accessible histone core region). The downstream functions of lysine methylation can vary widely depending on the sequence context in which the modification occurs and are often mediated by the binding of "effector" proteins and/or their associated protein complexes. For example, trimethylation of lysine 4 of histone H3 (H3K4me3) is enriched at the promotor regions of genomic loci that are competent for transcription, whereas H3K27me3 largely correlates with gene repression [8]. The H3K4me3-binding protein BPTF, the largest subunit of the NUcleosome Remodeling Factor (NURF) complex, and the H3K27me3-associated complex Polycomb Repressive Compex 1 (PRC1) have been shown to mediate functions of these marks, respectively [4, 9]. The fact that each of these methyl marks has a distinct relationship with transcription indicates that robust regulatory mechanisms exist to establish and maintain specific modification states, as well as those that remove or alter them specifically in response to cellular change (e.g., during differentiation and development).

Methyl marks on lysine residues in histone tails are added, or "written," by enzymes that usually contain the evolutionarily conserved SET domain (named after three lysine methyltransferases, or KMTs, that contain it: Su(var)3-9, Enhancer of Zeste, Trithorax); SET domain proteins are often chromatin associated, and many have been shown to have critical roles in development [10]. As suggested by the distinct functions of their methylated histone substrates, KMTs are very specific to the site of histone lysine methylation that they catalyze; many KMTs have been shown to methylate only a single lysine residue within the histone proteins [8]. KMTs are also known to methylate nonhistone substrates, and these modifications have been shown to have downstream biological consequences as well [11, 12]. It follows that mutation and/or misregulation of such KMTs would correlate with abnormal animal development as well as some human disease states, such as cancer; indeed, this connection has been demonstrated in numerous elegant studies, which have been reviewed previously [13].

Logically, histone methylation must also be removed, or "erased," from the chromatin template in order for it to serve as a dynamic regulator of gene expression. Here we will discuss three potential mechanisms for its erasure: (I) enzymatic demethylation, (II) replacement/exchange of the methylated histone protein, and (III) regulated proteolysis of the histone tail (Fig. 1). These three removal

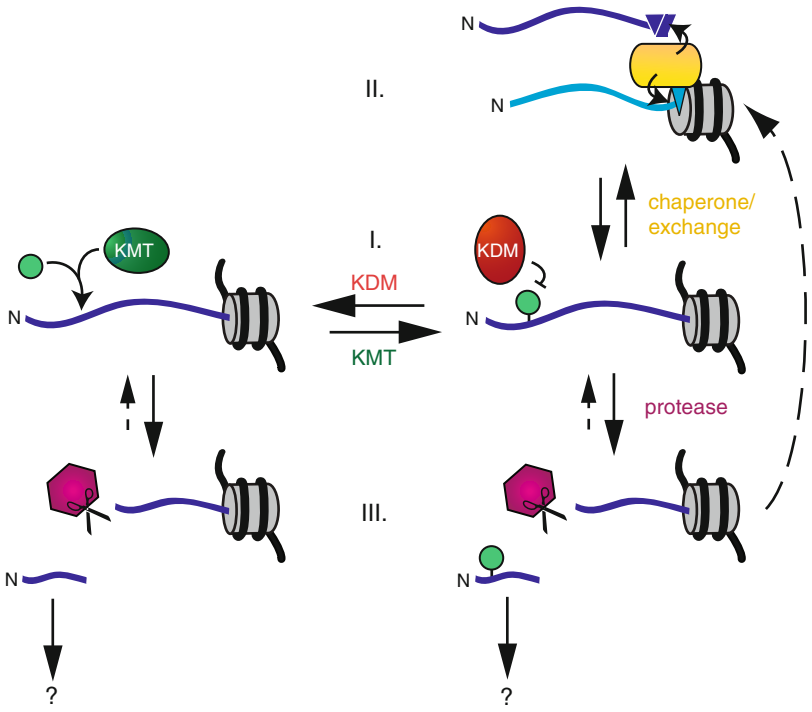


Fig. 1 Schematic representation of three mechanisms that lead to the removal of posttranslational modifications from the chromatin template. For clarity, this review focuses specifically on the removal of methylation on histone H3 (also see [14]). I. One mechanism by which demethylation is achieved is the enzymatic demethylation by a lysine (K) or arginine demethylase (not shown). Before such enzymes were shown to exist and function *in vivo*, methylation was hypothesized to be a "permanent" posttranslational modification. Lysine demethylases (KDMs) "erase" the marks that are "written" by lysine methyltransferases (KMT). II. A second mechanism by which methyl marks can be removed from the chromatin template is through the regulated replacement of the entire histone protein by a chaperone protein (or protein complex). III. A third mechanism of modification removal is through the proteolytic cleavage of the histone tail. Such proteolysis leads to the production of both a C-terminal peptide, which remains incorporated in the chromatin template until it is replaced (possibly by mechanism II, as represented by the *dashed arrow*), and an N-terminal peptide, which may have a subsequent, independent function once liberated

mechanisms have been reviewed previously [14]; however, experimental evidence for two of them was largely nonexistent at the time. Moreover, since both enzymatic demethylation and proteolysis of histone tails have been described more recently, the field is only beginning to uncover connections between these mechanisms of modification "erasure" and the pathologies of human disease. Below we will review the evidence linking these mechanisms to human disease and also propose potential models through which misregulation of histone methylation removal might lead to abnormal pathologies.

2 Lysine Demethylases Are Directly and Indirectly Associated with Both Oncogenic and Tumor-Suppressing Proteins

For many years, it was hypothesized that histone methylation was a stable, irreversible covalent modification and that its permanence made methylation a true "epigenetic" mark [15]. However, this view was invalidated by the identification of the first lysine-specific histone demethylase, LSD1 [16]. Previous studies had demonstrated that the peptidylarginine deiminase (PAD) family of enzymes is capable of converting methylated arginine residues to citrulline [17], but LSD1 was the first enzyme shown to be a bona fide lysine demethylase (KDM) in that it both removes mono or di-methyl groups from a lysine residue and concomitantly returns it to its unmodified state. Shi et al. further demonstrated that there is a functional relationship between LSD1 and H3K4 dimethylation occupancy at gene promoters, and that altering this occupancy using RNAi affects gene expression. The discovery that LSD1, an amine oxidase, demethylates histones was proceeded quickly by the identification of another family of KDMs, the Jumonji family [18–21]. The Jumonji family of KDMs differs from the amine oxidases LSD1 and AOF1 [22] in that they are able to demethylate tri-methylated histones. Together, these findings affirm that histone methylation, in all states, is a dynamic, reversible mark and that its enzymatic erasure is an important element of chromatin regulatory mechanisms.

As mentioned above, the mutation or misregulation of methylation "writers" is often detrimental to a given cell and/or organism. It follows that mutation or misregulation of methyl-histone "erasers" would also lead to improper gene expression and disordered biological outcomes. In fact, despite the relative newness of the identification of histone KDMs, there is both concrete and suggestive evidence that many of them are indeed linked to human diseases, such as cancer. Below we discuss a few notable examples.

LSD1. Interestingly, there is evidence showing that LSD1 is strongly expressed in aggressive prostate cancer cells [23]; however, its expression is decreased in breast cancer cells [24]. One explanation for this discrepancy is that LSD1 demethylates H3K9me1/2, rather than H3K4me1/2, upon interaction with the androgen receptor (AR) (Fig. 2), which then leads to the derepression of AR target genes such as the prostate-specific antigen (PSA) gene [25]. Given that prostate carcinomas have been shown to express high levels of AR and their proliferation is dependent on its expression [23], these data strongly suggest that LSD1 overexpression can lead to erroneous gene expression in AR-expressing cells. In the breast cancer model, on the other hand, LSD1 maintains its conventional specificity for H3K4me, a marker of active genes, and therefore must be down-regulated to promote gene expression. Wang et al. further showed that LSD1 is a subunit of the nucleosome remodeling and deacetylase (NuRD) complex, an association that not only combines two activities that induce gene repression (demethylation and deacetylation) but also forms a functional link between LSD1 and the chromatin template in breast cancer cells [24]. In addition, they demonstrate that this interaction

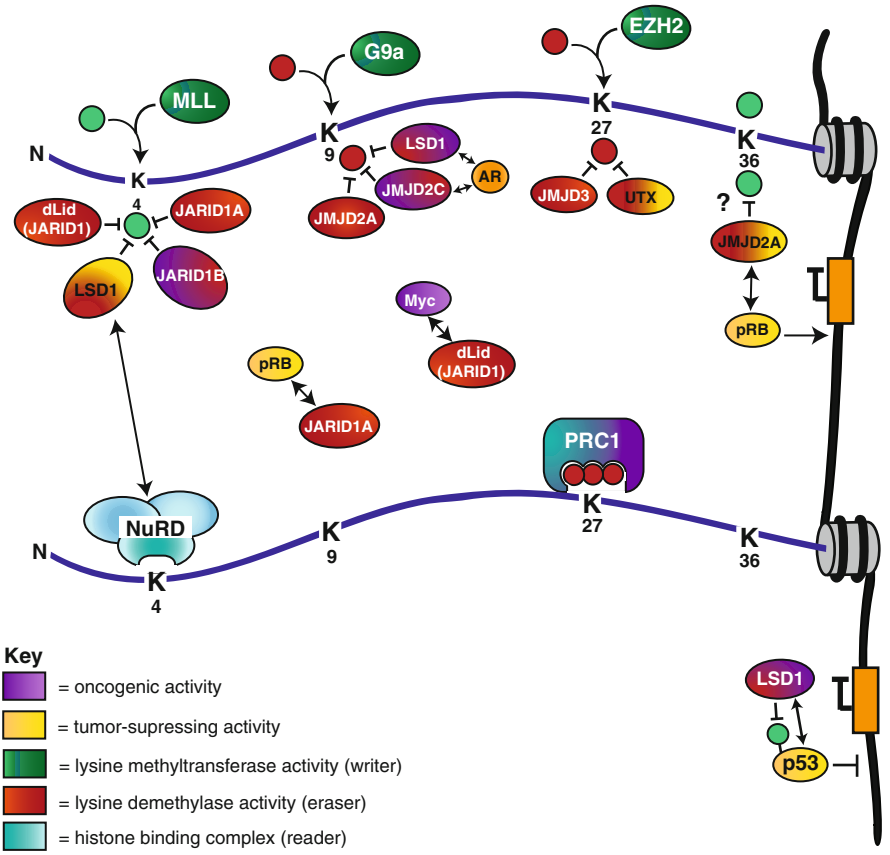


Fig. 2 Many lysine demethylases directly and/or indirectly possess oncogenic and/or tumor-suppressing activity. LSD1 demethylates H3K4 methylation, which leads to gene repression and tumor suppression via interaction with the NuRD complex (a complex that binds nonmethylated histone H3 tails). When in complex with the androgen receptor (AR), LSD1 demethylates H3K9 methylation, which leads to gene activation and cell proliferation. LSD1 has also been shown to demethylate methylated p53, which prevents it from binding DNA and leads to inactivation of p53 target genes. JMJD2C behaves similarly. JMJD2A demethylates H3K9 as well as H3K36 methylation; it is also known to interact with the tumor-suppressor pRB and enhance its tumor-suppressing gene repression. Such repression could be reinforced by simultaneous demethylation of the activating mark H3K36 by JMJD2A, although this mechanism is hypothetical. JARID1 family proteins demethylate H3K4 and have been shown to interact with both the tumor suppressor pRB and the oncoprotein Myc. JMJD3 and UTX both demethylate K27; UTX has been shown to have tumor-suppressing activity, balancing the oncogenic activity of the Polycomb Group (PcG) proteins, specifically Polycomb Repressive Complex 1 (PRC1). *Double-headed arrows* represent known physical associations between proteins (*dashed arrow* indicated association has been suggested but not published). Protein activities are shaded according to the key (*bottom left*). Select lysines are shown in the histone H3 tail (not drawn to scale), represented by the *solid blue line* (*horizontal*); DNA is represented by the *solid black line* (*vertical*). Genes encoded in the DNA sequence (*black*) are represented by *orange boxes*

is necessary for the proper silencing of downstream target genes, including genes that are linked to cancer and metastasis. In a separate study, Huang et al. also provide evidence that LSD1 has oncogenic activity in both osteosarcoma (U2OS) and breast cancer (MCF7) cell lines, although in this case such activity is mediated through the demethylation of p53 rather than histone H3 [26]. Together, these data suggest that LSD1 can have both oncogenic and tumor-suppressing activities, depending on the cellular context.

JMJD2C. The Jumonji-domain containing lysine demethylase JMJD2C (also known as Gene Amplified in Squamous cell Carcinoma 1, GASC1) is also able to demethylate histone H3K9 methylation [19]. Similar to the H3K9 demethylase complex LSD1/AR in prostate cancer, JMJD2C is overexpressed in esophageal squamous cell carcinoma, lung sarcomatoid carcinoma, and desmoplastic medulloblastoma [27–31] and appears to have oncogenic activity: inhibition of JMJD2C decreases cell proliferation and its ectopic expression leads to the delocalization of heterochromatin protein 1 (HP1) *in vivo* [19]. Unlike the LSD1/AR complex, which demethylates H3K9me1/2, JMJD2C can demethylate di- and tri-methylated H3K9 (H3K9me2/3). Notably, Wissmann et al. observe that these two demethylases interact physically and mechanistically with each other to promote the activation of AR target loci, in both normal human embryonic kidney cells (HEK-293) and prostate tumor cells (LNCaP) [32]. Supporting the idea that JMJD2C, and possibly other JMJD2 family members, has oncogenic activity, RNAi knockdown of the *C. elegans* JMJD2 orthologue (CeJMJD2) results in DNA damage and p53-dependent apoptosis in the germline [21]. In contrast, a related Jumonji-family member with H3K9 and H3K36 lysine demethylase activity, JMJD2A, has been shown to interact with the tumor-suppressor protein pRB to enhance its effect in silencing E2F responsive promotors [33]. These observations suggest that, much like LSD1, the oncogenic or tumor-supressing activity of the JMJ2D family is also dependent on the context in which it associates with the chromatin template.

JARID1 family. There are four JARID (Jumonji, AT-rich interactive domain) family proteins in mammalian cells: JARID1A, B, C, and D. Mutations in JARID1C (SMCX) associate with X-linked mental retardation, including several mutations that decrease its H3K4 demethylase activity [34–37]. JARID1D (SMCY) is also located on a sex chromosome (the male Y chromosome vs. the X for JARID1C/SMCX) and is also known to have H3K4 demethylase activity, although its *in vivo* function is not known [38].

JARID1A and B are highly similar to one another and both have been suggested to play a role in cancer [39]. The function of JARID1A (RBP2) in either normal or disease pathologies is still unclear. JARID1A knockout mice are both viable and fertile; its function is presumably at least partially redundant with another JARID family protein, although JARID1A −/− mice show derepression of a number of target loci and decreased apoptosis of hematopoetic progenitor compartments, perhaps suggesting a role in cell cycle regulation [40]. It is also possible, however, that JARID1A plays a yet uncovered role in tumorigenesis, given that it was originally identified as a pRB-binding protein [41]. Interestingly, Benevolenskaya et al. show that the interaction between JARID1A and pRB prevents JARID1A

from repressing its target genes and that this derepression promotes differentiation, suggesting the possibility that its mutation and/or misregulation could lead to hyperproliferation in certain contexts [42]. Moreover, their data show that the tumor-suppressing activity of pRB may be enhanced through its interaction with JARID1A.

JARID1B (PLU-1), on the other hand, has been more clearly linked to breast cancer [43, 44]. Prior to its identification as an H3K4 demethylase, Lu et al. showed that JARID1B is up-regulated in breast cancer cells compared with normal tissue (in which its expression is restricted to the testis) or colon cancer cell lines [43]. Yamane et al. later demonstrated that JARID1B demethylates H3K4me (di- and tri-) both *in vitro* and *in vivo* [44]. They further showed that knockdown of JARID1B leads to decreased proliferation of the breast cancer cell line MCF-7 and interferes with tumor growth in the murine breast cancer cell line 4T1, suggesting that JARID1B exhibits oncogenic activity in breast cancer cells, although it is not clear that this activity is directly related to its demethylase activity. In support of these findings, the *Drosophila* homologue of the JARID1 family, Lid, has been shown to interact both physically and genetically with the oncoprotein Myc [45]. Moreover, this interaction appears to enhance the oncogenic activity of Myc in that it leads to greater Myc-induced cell growth. Interestingly, this effect is not synergistic with the H3K4 demethylase activity of Lid since it was shown to be enzymatically inactive when bound to Myc [45]. The interaction between Lid and Myc may be conserved in mammals as JARID1A (RBP2) also interacts with Myc *in vitro* [45]. Given that the JARID1 family proteins contain the chromatin-binding PHD finger and ARID domains, there are many possible mechanisms through which Myc-bound JARID1 could be recruited to chromatin in order to regulate (or misregulate, as in over-expressing cancer cells) gene expression.

JMJD3 and UTX. Both the Jumonji domain-containing protein D3 (JMJD3) and the ubiquitously transcribed X-chromosome tetratricopeptide repeat protein (UTX) have been shown to demethylate histone H3 at K27 [46–49]. Predictably, both reverse the silencing effect of H3K27 methylation mediated by the Polycomb Group (PcG) protein EZH2. Importantly, demethylation of H3K27 by both UTX and JMJD3 leads to the derepression of *Hox* genes, which are key developmental regulators; as a result, the inhibition of UTX in zebrafish and JMJD3 in *C. elegans* leads to impaired posterior trunk and gonadal development, respectively [47, 49]. Although these H3K27 demethylases were identified only recently, two studies provide some evidence that they may play a role in human cancers. Xiang et al. show that JMJD3 is up-regulated in prostate cancer cells, particularly those that are thought to be more metastatic [50]. UTX, on the other hand, has been shown to have inactivating somatic mutations in several different types of cancers, including multiple myeloma, esophageal squamous cell carcinomas, and renal carcinomas, suggesting that it has tumor-suppressing, rather than oncogenic, activity [51]. van Haaften et al. also show that rescue with wild-type UTX in these mutated cancer cells slows their proliferation, strengthening this hypothesis. The notion that the H3K27me "eraser" UTX suppresses cellular proliferation is not unexpected; opposing "writers" and "readers" of H3K27 methylation are known to have oncogenic

activity, specifically the Polycomb Repressive Complex 1 (PRC1) proteins Bmi1 and CBX7 [52, 53]. Nevertheless, further studies will be needed to dissect exactly how these "writers," "readers," and "erasers" interact mechanistically at specific loci, such as p16INK4A-ARF, and how mutations and/or misregulation of them lead to human disease.

Other KDMs and disease. One possible universal connection between lysine demethylases and disease pathways is their role in the inflammatory response. Both lysine demethylases and their counterpart methyltransferases have been linked to the inflammatory response and its key regulator, NF-kappaB, including enzymes discussed above (LSD1 [54], JMJD3 [48]) and others (FBXL11 [55], SET7/9 [56]). Their link to the inflammatory response expands the role of lysine demethylases, such as LSD1 and JMJD3, to include diseases such as diabetes and other vascular pathologies, in addition to their roles in cancer. Moreover, these examples broaden the potential roles of all lysine demethylases, including those not mentioned here, and the pathways through which they affect disease pathologies.

3 Proteins Required for Histone Deposition Are Linked to Development and Disease

The bulk of histone deposition occurs during S-phase as DNA is replicated. However, long-standing evidence indicated that regulated histone replacement/exchange occurred outside of S-phase: studies in rat neurons [57], mammalian tissues [58], and *Tetrahymena* [59] suggested that variant isoforms of core histone proteins are incorporated into the chromatin fiber *independent* of replication (Fig. 1, mechanism II.). Using the *Drosophila* model, Ahmad and Henikoff then demonstrated that histone variant H3.3 is deposited into chromatin outside of S-phase and that it is specifically targeted to active loci, despite the fact that it differs from the canonical H3 in only four amino acids [60].

Although these studies in *Drosophila* established that regulated histone replacement/exchange exists *in vivo*, the exact mechanism by which it occurs is still unclear. A major contribution to our understanding of this process was made when Tagami et al. identified proteins responsible for the distinct H3.1 and H3.3 deposition pathways by purifying two separate chaperone complexes from mammalian cells with epitope-tagged H3.1 (i.e., canonical H3) vs. H3.3 [61]. Below we discuss one of these proteins, HIRA, and its link to human disease; we also describe the potential roles of another protein family, the CHD family, and its connections to similar disorders (see Table 1). In addition, we highlight new findings from our laboratory that implicate a third protein, alpha thalassemia/mental retardation syndrome X-linked (ATRX), in proper H3 localization.

HIRA. Prior to the discovery of its role in nucleosome assembly [62] and, more specifically, histone variant H3.3 deposition [61], human HIRA had been identified as one of the genes potentially responsible for the developmental disorder known as

Table 1 Histone replacement proteins and their mutant phenotypes

Protein	Mutation	Organism	Phenotype
H3.3	Homozygous null (both A & B genes)	*Drosophila*	Male and female sterility [67, 68]
HIRA	Homozygous null	*Drosophila*	Female sterility [65, 66]
	Homozygous null	Mouse	Embryonic lethality [64]
	Heterozygous 22q11 deletion	Human	Neurocristophathy (DiGeorge/ VCFS syndrome [60, 61])
CHD1	Homozygous null	*Drosophila*	Male and female sterility [72]
CHD5	Syntenic 1q36.3 deletion & shRNA kd	Mouse	Increased proliferation & tumorigenesis [78]
	1q36.3 deletion	Human	Multiple cancers [78–84]
CHD7	Multiple heterozygous mutations	Human	Neurocristopathy (CHARGE syndrome [85–87])
	shRNA kd	hESCs	Inability to differentiate into hNCLCs [88]
	Morpholino kd	*Xenopus*	CHARGE-like features [88]

DiGeorge syndrome [63]. DiGeorge syndrome, as well as other similar syndromes, is associated with a deletion in the long arm of human chromosome 22 (22q11) and phenotypically characterized by multiple malformations including the absence or hypoplasia of the thymus and parathyroid glands, cardiac defects, cleft palate, and other craniofacial abnormalities [64]. Human HIRA maps within the smallest critical region for these syndromes and is so named because of its homology with the *S. cerevisiae* HIR1 and HIR2 genes [63]. There are several genes within the 22q11 locus (including the Tbx1 gene that has been linked to DiGeorge-like heart defects in mice [65, 66]) and considerable variation in phenotype among those people affected by its deletion, making it difficult to pinpoint the responsible gene (or genes) causing these syndromes. Interestingly, 22q11 deletion syndromes are examples of haploinsufficiency disorders, i.e., their phenotypes are caused by deletion of only one copy of the locus. However, while Tbx1 haploinsufficiency in mice causes heart defects that are similar to those seen in patients with 22q11 haploinsufficiency [65, 66], mice heterozygous for a null mutation in the HIRA gene have not been reported to mimic the human 22q11 deletion phenotype and homozygous HIRA null mice die during early embryogenesis, before the relevant structures have developed [67]. Nevertheless, given the complexity of 22q11 syndrome phenotypes and the difficulty in proving causation by a single gene, it remains possible that HIRA haploinsufficiency contributes to these disorders.

Although homozygosity for a HIRA null mutation causes embryonic lethality in mice, similar mutations in *Drosophila* produce adult flies [68, 69]. Although this difference in viability suggests that HIRA (or perhaps its substrate, H3.3) may have different roles in *Drosophila* and mammals, it also makes *Drosophila* a good model for studying HIRA and H3.3 function. To this end, both the HIRA mutation *sésame* ($Hira^{ssm}$, a R225K point mutation) and a loss-of-function mutation generated by homologous recombination ($Hira^{HR1}$) result in female sterility, in which females produce eggs that do not hatch [68, 69]. Although this sterility is a maternal effect,

it is mediated through the failure to properly remodel the male germline: the DNA of sperm nuclei is packaged with sperm nuclear basic proteins (SNBPs), rather than the four core histones, which are then replaced with maternal histones upon fertilization. This replacement does not occur in the fertilized eggs of female *Hira* mutant flies, leading to the formation of a fatally abnormal male pronucleus in these embryos [68, 69]. Notably, normal flies appear to preferentially deposit the variant H3.3 into the chromatin of the male pronucleus, even when the canonical H3 is expressed using the regulatory sequence of the H3.3A gene and present in roughly equal amounts in the egg cytoplasm [69], suggesting that the role of HIRA in the male pronucleus is mediated though H3.3-specific deposition. These data are supported by further studies in *Drosophila* that show null mutations in both H3.3 genes (A and B) also lead to sterility, in this case in both male and female flies [70, 71]. The sterility of both male and female flies in the H3.3 mutants suggests that H3.3 plays a role in the proper formation of the germ line in male flies as well as in the remodeling of paternal chromatin in female flies. Moreover, the fact that male sterility is not seen in the HIRA mutants suggests that H3.3 can function independently of HIRA. Importantly, although mice null for HIRA and H3.3 are nonviable and unavailable, respectively, observational data suggest the roles of HIRA and H3.3 in remodeling paternal chromatin and the role of H3.3 in remodeling the male germ line may be conserved in mammals [72–74].

CHD family proteins. As discussed above, the fact that *Drosophila* mutations in both H3.3 genes produce sterility in both male and female flies while *Hira* null mutations only affect the viability of embryos via a maternal effect suggests that HIRA is not required for H3.3 deposition in the germ line of *Drosophila* males. Furthermore, Bonnefoy et al. suggest that H3.3 deposition is unaffected in *Hira* mutant flies, outside of male pronucleus remodeling, since they see comparable epitope-tagged H3.3 incorporation in both wild type and $Hira^{HR1}$ mutants [68]. Although this is an interesting and likely conclusion (see below), it is important to note that it is based on the global localization of H3.3 by immunofluorescence; it is possible that mutations in HIRA affect H3.3 deposition at specific loci, as could be determined by Chromatin ImmunoPrecipitation followed by DNA sequencing (ChIP-seq). Nevertheless, these data draw attention to the fact that there are likely HIRA-independent mechanisms for H3.3 deposition and histone H3 replacement.

One protein that is likely responsible for HIRA-independent H3 replacement, at least for some loci and developmental stages, is the motor protein CHD1. In a study that nicely builds upon the *Hira* mutant data described above, Konev et al. showed that CHD1 is also required for H3.3 deposition into chromatin in *Drosophila* [75]. Importantly, they show that mutating CHD1 causes sterility in both male and female flies [75], which phenocopies the effect of mutating both H3.3 genes [70, 71]; this effect is different from that found in *Hira* mutants, of which only the females are sterile [68, 69]. Although their data mostly focus on demonstrating the effect of the null CHD1 mutation on the male pronucleus in the eggs of mutant females, the fact that male flies are sterile as well suggests that CHD1 has an additional role in the chromatin remodeling of the male germ line. In addition, Konev et al. also suggest that H3.3 deposition is impaired at later stages of embryonic

development, since they see poor localization of epitope-tagged H3.3 with DNA in CHD1 mutants as compared with wild-type embryos at a similar developmental stage [75]. These authors conclude that this effect is independent of both the haploid state of the embryo (the result of an abnormal male pronucleus) and HIRA, since previous studies in haploid *Hira* mutant embryos did not appear to have defects in global H3.3 incorporation [68, 69]. Although their conclusion that CHD1 may play a global, HIRA-independent role in H3.3 deposition is intriguing, and perhaps likely, more quantitative studies will be needed to confirm the differences in localization of H3.3 in wild-type, CHD1 mutant and *Hira* mutant embryos.

Together, the data described above not only suggest that there is likely one or more HIRA-independent mechanisms of H3.3 deposition and H3 replacement, but also that CHD1, and perhaps other CHD family proteins, may play a role in such alternative mechanisms. Although the functional experiments described above were done in *Drosophila*, these studies may have revealed an interesting and important role for a family of proteins that are implicated in both development and disease.

CHD stands for chromatin helicase DNA-binding and its family members are characterized by a tandem chromodomain motif in the N-terminus and a SNF2-like ATPase domain in the structural center [76]. Several of the CHD proteins also contain DNA-binding domains and/or PHD finger motifs that, along with the helicase and chromodomains, are thought to mediate interactions with chromatin. Of the nine proteins in the CHD family (1–9), at least four of them are associated with human disease [76]. Both CHD3 and CHD4 have been linked specifically to dermatomyositis, a disease in which patients suffer from inflammation of the skin and muscle tissue [77–79]; autoantibodies to CHD3 and CHD4 (or Mi-2, as they were originally named after the *Drosophila* homologue) are detected specifically in patients with dermatomyositis, but not in patients with a related disease [77]. CHD3 and CHD4 are also both components of the NuRD complex (which interacts with the demethylase LSD1, see Sect. 1). In addition, CHD3 has been shown to physically interact with the protein Ki1/57, a marker for cancer cells in patients with Hodgkin's lymphoma [80], perhaps linking CHD3 itself to the disease.

CHD5 has also been linked to human disease and, moreover, has been shown to be a tumor suppressor [81]. It bears the characteristic tandem chromodomains and ATPase domain of the CHD family and is predicted to have two PHD finger motifs (similar to CHD3 and CHD4) as well [76, 82]. CHD5 is one of the genes that map to a region of human chromosome 1p36.3 that is commonly deleted in neuroblastoma [82]. Thompson et al. not only identified CHD5 as one of the genes located in the 1p36.3 neuroblastoma deletion, but also demonstrated that its mRNA is more highly expressed in the nervous system and that this expression is significantly decreased in neuroblastomas from patients with poor outcomes [82]. Importantly, 1p36.3 is also deleted in several other types of cancer, including both hematopoietic [83, 84] and epithelial malignancies [85–87]. To this end, Bagchi et al. have shown that CHD5 acts as a tumor suppressor *in vivo* in mice [81], strongly implicating this gene in the human cancers with the 1p36.3 deletion.

The most direct connection between a CHD family protein and human disease is that between CHD7 and CHARGE syndrome; approximately two out of three

affected patients have heterozygous mutations in CHD7 [88, 89]. CHARGE stands for Coloboma of the eye, Heart malformation, Atresia of chonae, Retardation of growth and development, Genital hypoplasia, and Ear defects and is one of the most common classifications of congenital anomalies [90]. Although the connection between CHD7 and CHARGE syndrome is well established, data on the potential mechanism by which CHD7 haploinsufficiency leads to the disease phenotype have been published only recently [91]. Although it had been hypothesized previously that CHARGE syndrome may be due to improper specification and/or development of the neural crest [92], Bajpai et al. are the first to test this theory experimentally and connect it directly to a CHD7-dependent mechanism of gene expression regulation. In a separate study, Scnetz et al. use chromatin immunoprecipitation (ChIP) to show that CHD7 maps to sites of H3K4 methylation, with the strongest correlation to H3K4 monomethylation. Separately, these authors demonstrate that its chromodomains bind directly and specifically to methylated H3K4, although they do not appear to discriminate between the mono-, di-, and trimethylated states in their assay [93]. Together these data suggest that the binding of CHD7 to H3K4 methylation may localize and/or stabilize it at particular genes. Interestingly, both CHD7 and HIRA are implicated in syndromes that may originate from the improper development of the neural crest (i.e., neurocristopathies). HIRA is expressed in the neural crest of both chick [94] and mouse embryos [95], and patients with a 22q11 deletion that includes the human HIRA gene suffer from DiGeorge syndrome, which is characterized by defects in many of the same structures as those affected in patients with CHARGE syndrome. However, further studies will be needed to determine whether there is a mechanistic connection between these two proteins.

ATRX. Recent work from our laboratory [96] has revealed that another protein is critical for histone H3 localization: ATRX. Using a genome-wide approach, Goldberg et al. show that histone H3.3 localizes to specific regions of the genome in mouse embryonic stem cells and neuronal precursor cells, including genes, regulatory regions, and telomeres. Furthermore, they demonstrate that distinct factors are responsible for H3.3 localization at these specific regions. H3.3 chaperone HIRA is required for H3.3 localization to genes and particular regulatory regions, but ATRX is required for telomeric localization of H3.3. Goldberg et al. also show that ATRX is required for the repression of telomeric RNA in mouse embryonic stem cells and that H3.3 biochemically associates with ATRX, independent of its interaction with HIRA. Interestingly, patients with mutations in ATRX (leading to the namesake disorder) show misexpression of their alpha-globin genes, which are located extremely near the telomere of human chromosome 16 [97]. It is possible that the proximity of alpha-globin genes to the telomere may explain their misregulation in ATRX patients: the inability of mutant ATRX to properly localize histone H3.3 at this locus could lead to errors in the regulation of nearby genes. Goldberg et al. are the first to document that ATRX plays a role at telomeres; further studies will dissect the exact mechanism of H3.3 deposition/replacement by ATRX and its interacting partners.

4 Mislocalization of Cathepsin Proteases Is a Marker of Cancer

Although histone cleavage as a mechanism for removing histone methylation (Fig. I, mechanism III.) had been proposed previously [14] and the cleavage of histone H3 tails had been demonstrated previously in the literature [98, 99], evidence of endogenous H3 cleavage in mammalian cells was documented only recently [100]. This study was also the first to identify a protease responsible for the observed H3 cleavage: Cathepsin L was shown to create the pattern of H3 cleavage that is generated *in vivo* during mouse embryonic stem cell (ESC) differentiation. A subsequent study in *S. cerevisiae* demonstrated that an identical cleavage site is created in the histone H3 of yeast, suggesting that this H3 proteolysis mechanism may be conserved [101]. The fact that H3 proteolysis was observed during a narrow window of ESC differentiation and that the cleavage activity of both Cathepsin L and the yeast H3 protease is affected by histone modifications on the H3 tail itself suggests that histone proteolysis is a highly regulated event that occurs at a particular stage of development, differentiation, and/or cell cycle progression.

The finding that Cathepsin L localizes to the nucleus and associates with chromatin was surprising, given that it is a well-known lysosomal protein [102]. Nevertheless, its nuclear localization was not novel: Hiwasa and Sakiyama showed that Cathepsin L (or MEP, for major excreted protein [103]) was enriched in the nuclear fraction of *ras* and *erb*B2-transformed mouse fibroblasts (NIH3T3) compared with nontransformed or v-*mos*-transformed controls [104]; Goulet et al. demonstrated that Cathepsin L localizes to the nucleus during S-phase in NIH3T3 cells [105]; Boudreau et al. not only showed that Cathepsin L is in the nuclear fraction of intestinal epithelial cells, but also that its expression increases upon differentiation [106]. Interestingly, these studies not only provide supporting evidence for the nuclear localization of Cathepsin L, but also suggest that such alternate localization may be an instigator and/or an indicator of cancer cell transformation.

The connection between Cathepsin L, as well as other cathepsin enzymes, and cancer had been established previously [107, 108]; however, those studies mainly focused on the link between the secretion of cathepsin enzymes and malignancy. Although the study by Hiwasa and Sakiyama did not determine whether the nuclear localization of Cathepsin L was a cause or an effect of the malignant transformation by *ras* or *erb*B2 [104], that by Goulet et al. demonstrated that a shorter form of nuclear Cathepsin L (lacking its N-terminal signal peptide) cleaves the CDP/Cux transcription factor and that one of the products of this proteolysis, p110, accelerates the G1/S transition of the cell cycle [105]. Goulet et al. then show that *ras* transformation of human cells increases both the production of the signal-less form of Cathepsin L and the processing of its substrate, CDP/Cux [109]. Boudreau et al. also show that inhibition of intracellular Cathepsin L, but not extracellular enzyme, impairs the polarization and differentiation of epithelial cells, and that mice with inactive Cathepsin L (*furless*) show increased intestinal neoplasia and intestinal polyps (when crossed with Apc^{Min} mice) [106]. However, although the authors

attribute these effects to intracellular Cathepsin L, it is unclear whether or not they are mediated by its activity in the nucleus or another intracellular compartment.

The studies described above suggest a link between nuclear Cathepsin L and both differentiation and transformation; however, the question remains as to whether its proteolysis of H3 correlates with or causes such events. Although there is no direct, *in vivo* evidence indicating that mammalian differentiation or transformation is mediated through cleavage of the H3 tail (and it is difficult to test since Cathepsin L has at least one other nuclear substrate, CDP/Cux, and the many copies of mammalian H3 make it impractical to create a noncleavable mutant), there are many ways in which histone proteolysis might regulate the gene expression programs responsible for these transitions (see Fig. 3): the removal of the N-terminal tail and its modifications, the creation of a new N-terminus, the regulated replacement of the cleaved histone in the chromatin fiber, and the liberation of the cleaved N-terminal tail peptide. Although these four modes of gene regulation are

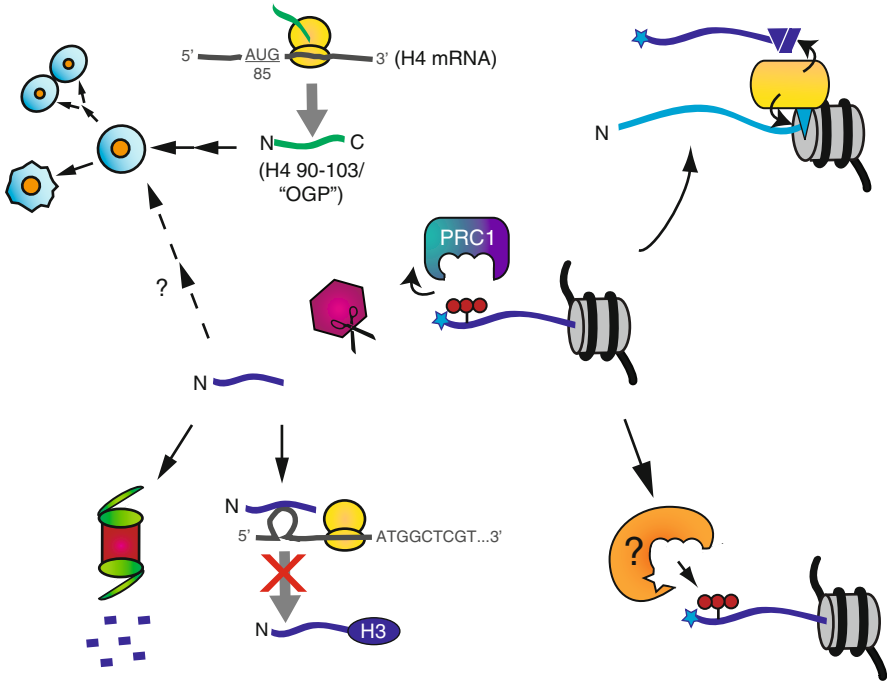


Fig. 3 Potential ways that gene expression might be regulated as a result of the N-terminal cleavage of histone H3. Loss of the N-terminal tail leads to loss of any modifications found there as well as their downstream effector proteins (e.g., the NURF complex subunit BPTF, which binds to H3K4me3) from the chromatin fiber. Loss of key N-terminal residues also leads to reduced binding by the H3K27me3-associated protein Polycomb. It is possible, however, that the new N-terminus produced by this cleavage (represented by the *star*) could recruit novel binding proteins (labeled with *question mark*). The C-terminal peptide may also be actively and specifically replaced by chaperone proteins. There are also many possible fates for the released N-terminal peptide: degradation by the proteasome, signaling to other cells, and/or inhibition of mRNA translation

somewhat speculative, we propose that they may be both important and conserved among organisms.

The removal of the N-terminal tail of H3 has been shown to regulate gene expression directly in yeast [101]; *in vitro* binding studies also indicate that the loss of the H3 tail impacts Polycomb binding [100], which could lead to significant downstream effects in the context of both normal development and cancer. Similarly, the creation of a new N-terminus, and a new context for the remaining modifications, could allow for the binding of proteins that were previously unable to engage with the H3 tail. Such proteins might even include chaperones that could purposefully replace the cleaved histone with a new molecule. Interestingly, this type of cleavage-replacement mechanism appears to exist in sea urchins during male pronucleus remodeling [110, 111]: inhibition of a sperm histone (SpH)-specific cysteine protease prevents their degradation and, subsequently, inhibits male chromatin decondensation and proper embryo development [111]. Moreover, this SpH protease is highly homologous to the Cathepsin L family of proteases and is inhibited by Cathepsin L inhibitor I [112]. The effect of SpH protease inhibition is strikingly similar to that of HIRA, CHD1, and H3.3 mutant flies (described above), although it is important to note that the sea urchin SpH protease is not specific to H3. Nevertheless, it would be interesting to test whether such a cleavage-replacement mechanism has been conserved and adapted in higher organisms, during both postfertilization sex chromatin remodeling and in more general chromatin remodeling that occurs throughout differentiation and development.

Finally, the liberation of the N-terminal tail peptide of H3 may also contribute to the regulation of gene expression following H3 proteolysis. For example, a recent study by Lee et al. suggests that an N-terminal H3 peptide, such as that released by Cathepsin L cleavage, could bind its own mRNA and therefore regulate its own translation [113]. Although the H3 peptide used in this study was synthesized *in vitro*, its effect on translation reveals a possible mechanism through which the nonchromatin-bound postcleavage peptide might influence the expression of protein within the cell. Notably, an endogenous peptide that is identical to the C-terminus of histone H4 was purified from medium conditioned with cultured regenerating rat bone marrow and identified as a factor that can induce osteogenic growth [114, 115]. Although this peptide (named OGP for osteogenic growth peptide) appears to originate from an internal, suboptimal translation start codon in the H4 gene [116, 117], the finding that (secreted) histone H4 peptide mediates signal transduction between cells and induces osteogenic growth raises the question as to whether the release of the N-terminal peptide of H3 after cleavage might also have a proliferative effect.

5 Conclusion

The removal of covalent histone modifications, such as lysine methylation, is a highly regulated process. The three mechanisms described above are influenced by a variety of factors, all of which contribute to the fine-tuning of gene expression

programs that drive differentiation and development. Mutations in one or more parts of this machinery can lead to the misregulation of gene expression and, ultimately, disease. Here, we aimed to review both established links between mechanisms of histone modification removal and disease as well as several emerging links that are more speculative. We look forward to reviewing further studies that test such hypotheses and uncover more details about how the mechanisms that erase histone modifications influence both normal and pathological cellular states.

Acknowledgments We would like to thank Gang Wang, Ping Chi, Joanna Wysocka, and Aaron Goldberg for a critical reading of the manuscript and for helpful advice in its preparation. We apologize to all those whose work we could not cite due to space constraints. C.D.A. is supported by NIH Grant GM53512, and both C.D.A. and E.M.D. are supported by a grant from the Tri-Institutional Stem Cell Initiative (funded by the Starr Foundation).

References

1. Van Holde KE (1989) Chromatin. Springer-Verlag, New York
2. Allis CD, Jenuwien T, Reinberg D (2006) Epigenetics. Cold Spring Harbor Laboratory Press, New York
3. Allfrey VG, Faulkner R, Mirsky AE (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. Proc Natl Acad Sci USA 51:786–794
4. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298:1039–1043
5. Kouzarides T (2002) Histone methylation in transcriptional control. Curr Opin Genet Dev 12:198–209
6. Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, de la Cruz CC, Otte AP, Panning B, Zhang Y (2003) Role of histone H3 lysine 27 methylation in X inactivation. Science 300:131–135
7. Strahl BD, Ohba R, Cook RG, Allis CD (1999) Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. Proc Natl Acad Sci USA 96:14967–14972
8. Kouzarides T (2007) Chromatin modifications and their function. Cell 128:693–705
9. Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P et al (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature 442:86–90
10. Dillon SC, Zhang X, Trievel RC, Cheng X (2005) The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol 6:227
11. Chuikov S, Kurash JK, Wilson JR, Xiao B, Justin N, Ivanov GS, McKinney K, Tempst P, Prives C, Gamblin SJ et al (2004) Regulation of p53 activity through lysine methylation. Nature 432:353–360
12. Su IH, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT, Tarakhovsky A (2003) Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. Nat Immunol 4:124–131
13. Schneider R, Bannister AJ, Kouzarides T (2002) Unsafe SETs: histone lysine methyltransferases and cancer. Trends Biochem Sci 27:396–402
14. Bannister AJ, Schneider R, Kouzarides T (2002) Histone methylation: dynamic or static? Cell 109:801–806

15. Jenuwein T, Allis CD (2001) Translating the histone code. Science 293:1074–1080
16. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 119:941–953
17. Wang Y, Wysocka J, Sayegh J, Lee YH, Perlin JR, Leonelli L, Sonbuchner LS, McDonald CH, Cook RG, Dou Y et al (2004) Human PAD4 regulates histone arginine methylation levels via demethylimination. Science 306:279–283
18. Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, Zhang Y (2006) Histone demethylation by a family of JmjC domain-containing proteins. Nature 439:811–816
19. Cloos PA, Christensen J, Agger K, Maiolica A, Rappsilber J, Antal T, Hansen KH, Helin K (2006) The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. Nature 442:307–311
20. Klose RJ, Yamane K, Bae Y, Zhang D, Erdjument-Bromage H, Tempst P, Wong J, Zhang Y (2006) The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. Nature 442:312–316
21. Whetstine JR, Nottke A, Lan F, Huarte M, Smolikov S, Chen Z, Spooner E, Li E, Zhang G, Colaiacovo M et al (2006) Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell 125:467–481
22. Ciccone DN, Su H, Hevi S, Gay F, Lei H, Bajko J, Xu G, Li E, Chen T (2009) KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. Nature 461:415–418
23. Kahl P, Gullotti L, Heukamp LC, Wolf S, Friedrichs N, Vorreuther R, Solleder G, Bastian PJ, Ellinger J, Metzger E et al (2006) Androgen receptor coactivators lysine-specific histone demethylase 1 and four and a half LIM domain protein 2 predict risk of prostate cancer recurrence. Cancer Res 66:11341–11347
24. Wang Y, Zhang H, Chen Y, Sun Y, Yang F, Yu W, Liang J, Sun L, Yang X, Shi L et al (2009) LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. Cell 138:660–672
25. Metzger E, Wissmann M, Yin N, Müller JM, Schneider R, Peters AHFM, Günther T, Buettner R, Schüle R (2005) LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature 437:436–439
26. Huang J, Sengupta R, Espejo AB, Lee MG, Dorsey JA, Richter M, Opravil S, Shiekhattar R, Bedford MT, Jenuwein T et al (2007) p53 is regulated by the lysine demethylase LSD1. Nature 449:105–108
27. Yang ZQ, Imoto I, Fukuda Y, Pimkhaokham A, Shimada Y, Imamura M, Sugano S, Nakamura Y, Inazawa J (2000) Identification of a novel gene, GASC1, within an amplicon at 9p23-24 frequently detected in esophageal cancer cell lines. Cancer Res 60:4735–4739
28. Ehrbrecht A, Muller U, Wolter M, Hoischen A, Koch A, Radlwimmer B, Actor B, Mincheva A, Pietsch T, Lichter P et al (2006) Comprehensive genomic analysis of desmoplastic medulloblastomas: identification of novel amplified genes and separate evaluation of the different histological components. J Pathol 208:554–563
29. Italiano A, Attias R, Aurias A, Perot G, Burel-Vandenbos F, Otto J, Venissac N, Pedeutour F (2006) Molecular cytogenetic characterization of a metastatic lung sarcomatoid carcinoma: 9p23 neocentromere and 9p23-p24 amplification including JAK2 and JMJD2C. Cancer Genet Cytogenet 167:122–130
30. Yang ZQ, Imoto I, Pimkhaokham A, Shimada Y, Sasaki K, Oka M, Inazawa J (2001) A novel amplicon at 9p23–24 in squamous cell carcinoma of the esophagus that lies proximal to GASC1 and harbors NFIB. Jpn J Cancer Res 92:423–428
31. Hu N, Wang C, Hu Y, Yang HH, Giffen C, Tang ZZ, Han XY, Goldstein AM, Emmert-Buck MR, Buetow KH et al (2005) Genome-wide association study in esophageal cancer using GeneChip mapping 10K array. Cancer Res 65:2542–2546
32. Wissmann M, Yin N, Muller JM, Greschik H, Fodor BD, Jenuwein T, Vogler C, Schneider R, Gunther T, Buettner R et al (2007) Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. Nat Cell Biol 9:347–353

33. Gray SG, Iglesias AH, Lizcano F, Villanueva R, Camelo S, Jingu H, Teh BT, Koibuchi N, Chin WW, Kokkotou E et al (2005) Functional characterization of JMJD2A, a histone deacetylase- and retinoblastoma-binding protein. J Biol Chem 280:28507–28518
34. Iwase S, Lan F, Bayliss P, de la Torre-Ubieta L, Huarte M, Qi HH, Whetstine JR, Bonni A, Roberts TM, Shi Y (2007) The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. Cell 128:1077–1088
35. Jensen LR, Amende M, Gurok U, Moser B, Gimmel V, Tzschach A, Janecke AR, Tariverdian G, Chelly J, Fryns JP et al (2005) Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. Am J Hum Genet 76:227–236
36. Tahiliani M, Mei P, Fang R, Leonor T, Rutenberg M, Shimizu F, Li J, Rao A, Shi Y (2007) The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. Nature 447:601–605
37. Tzschach A, Lenzner S, Moser B, Reinhardt R, Chelly J, Fryns JP, Kleefstra T, Raynaud M, Turner G, Ropers HH et al (2006) Novel JARID1C/SMCX mutations in patients with X-linked mental retardation. Hum Mutat 27:389
38. Lee MG, Norman J, Shilatifard A, Shiekhattar R (2007) Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein. Cell 128:877–887
39. Secombe J, Eisenman RN (2007) The function and regulation of the JARID1 family of histone H3 lysine 4 demethylases: the Myc connection. Cell Cycle 6:1324–1328
40. Klose RJ, Yan Q, Tothova Z, Yamane K, Erdjument-Bromage H, Tempst P, Gilliland DG, Zhang Y, Kaelin WG Jr (2007) The retinoblastoma binding protein RBP2 is an H3K4 demethylase. Cell 128:889–900
41. Defeo-Jones D, Huang PS, Jones RE, Haskell KM, Vuocolo GA, Hanobik MG, Huber HE, Oliff A (1991) Cloning of cDNAs for cellular proteins that bind to the retinoblastoma gene product. Nature 352:251–254
42. Benevolenskaya EV, Murray HL, Branton P, Young RA, Kaelin WG Jr (2005) Binding of pRB to the PHD protein RBP2 promotes cellular differentiation. Mol Cell 18:623–635
43. Lu PJ, Sundquist K, Baeckstrom D, Poulsom R, Hanby A, Meier-Ewert S, Jones T, Mitchell M, Pitha-Rowe P, Freemont P et al (1999) A novel gene (PLU-1) containing highly conserved putative DNA/chromatin binding motifs is specifically up-regulated in breast cancer. J Biol Chem 274:15633–15645
44. Yamane K, Tateishi K, Klose RJ, Fang J, Fabrizio LA, Erdjument-Bromage H, Taylor-Papadimitriou J, Tempst P, Zhang Y (2007) PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation. Mol Cell 25:801–812
45. Secombe J, Li L, Carlos L, Eisenman RN (2007) The Trithorax group protein Lid is a trimethyl histone H3K4 demethylase required for dMyc-induced cell growth. Genes Dev 21:537–551
46. Lee MG, Villa R, Trojer P, Norman J, Yan K-P, Reinberg D, Di Croce L, Shiekhattar R (2007) Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. Science 318:447–450
47. Agger K, Cloos PA, Christensen J, Pasini D, Rose S, Rappsilber J, Issaeva I, Canaani E, Salcini AE, Helin K (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature 449:731–734
48. De Santa F, Totaro MG, Prosperini E, Notarbartolo S, Testa G, Natoli G (2007) The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. Cell 130:1083–1094
49. Lan F, Bayliss PE, Rinn JL, Whetstine JR, Wang JK, Chen S, Iwase S, Alpatov R, Issaeva I, Canaani E et al (2007) A histone H3 lysine 27 demethylase regulates animal posterior development. Nature 449:689–694
50. Xiang Y, Zhu Z, Han G, Lin H, Xu L, Chen CD (2007) JMJD3 is a histone H3K27 demethylase. Cell Res 17:850–857

51. van Haaften G, Dalgliesh GL, Davies H, Chen L, Bignell G, Greenman C, Edkins S, Hardy C, O'Meara S, Teague J et al (2009) Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. Nat Genet 41:521–523
52. Gil J, Bernard D, Martinez D, Beach D (2004) Polycomb CBX7 has a unifying role in cellular lifespan. Nat Cell Biol 6:67–72
53. Jacobs JJ, Kieboom K, Marino S, DePinho RA, van Lohuizen M (1999) The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. Nature 397:164–168
54. 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 103:615–623
55. Lu T, Jackson MW, Wang B, Yang M, Chance MR, Miyagi M, Gudkov AV, Stark GR (2010) Regulation of NF-kappaB by NSD1/FBXL11-dependent reversible lysine methylation of p65. Proc Natl Acad Sci USA 107:46–51
56. Li Y, Reddy MA, Miao F, Shanmugam N, Yee JK, Hawkins D, Ren B, Natarajan R (2008) Role of the histone H3 lysine 4 methyltransferase, SET7/9, in the regulation of NF-kappaB-dependent inflammatory genes. Relevance to diabetes and inflammation. J Biol Chem 283:26771–26781
57. Pina B, Suau P (1987) Changes in histones H2A and H3 variant composition in differentiating and mature rat brain cortical neurons. Dev Biol 123:51–58
58. Lennox RW, Cohen LH (1988) The production of tissue-specific histone complements during development. Biochem Cell Biol 66:636–649
59. Yu L, Gorovsky MA (1997) Constitutive expression, not a particular primary sequence, is the important feature of the H3 replacement variant hv2 in Tetrahymena thermophila. Mol Cell Biol 17:6303–6310
60. Ahmad K, Henikoff S (2002) The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Mol Cell 9:1191–1200
61. Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell 116:51–61
62. Ray-Gallet D, Quivy JP, Scamps C, Martini EM, Lipinski M, Almouzni G (2002) HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. Mol Cell 9:1091–1100
63. Lamour V, Lecluse Y, Desmaze C, Spector M, Bodescot M, Aurias A, Osley MA, Lipinski M (1995) A human homolog of the S. cerevisiae HIR1 and HIR2 transcriptional repressors cloned from the DiGeorge syndrome critical region. Hum Mol Genet 4:791–799
64. Scambler PJ (2000) The 22q11 deletion syndromes. Hum Mol Genet 9:2421–2426
65. Lindsay EA, Vitelli F, Su H, Morishima M, Huynh T, Pramparo T, Jurecic V, Ogunrinu G, Sutherland HF, Scambler PJ et al (2001) Tbx1 haploinsufficieny in the DiGeorge syndrome region causes aortic arch defects in mice. Nature 410:97–101
66. Merscher S, Funke B, Epstein JA, Heyer J, Puech A, Lu MM, Xavier RJ, Demay MB, Russell RG, Factor S et al (2001) TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. Cell 104:619–629
67. Roberts C, Sutherland HF, Farmer H, Kimber W, Halford S, Carey A, Brickman JM, Wynshaw-Boris A, Scambler PJ (2002) Targeted mutagenesis of the Hira gene results in gastrulation defects and patterning abnormalities of mesoendodermal derivatives prior to early embryonic lethality. Mol Cell Biol 22:2318–2328
68. Bonnefoy E, Orsi GA, Couble P, Loppin B (2007) The essential role of Drosophila HIRA for de novo assembly of paternal chromatin at fertilization. PLoS Genet 3:1991–2006
69. Loppin B, Bonnefoy E, Anselme C, Laurencon A, Karr TL, Couble P (2005) The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. Nature 437:1386–1390

70. Hodl M, Basler K (2009) Transcription in the absence of histone H3.3. Curr Biol 19:1221–1226
71. Sakai A, Schwartz BE, Goldstein S, Ahmad K (2009) Transcriptional and developmental functions of the H3.3 histone variant in Drosophila. Curr Biol 19(21):1816–1820
72. Torres-Padilla M-E, Bannister AJ, Hurd PJ, Kouzarides T, Zernicka-Goetz M (2006) Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. Int J Dev Biol 50:455–461
73. van der Heijden GW, Derijck AA, Posfai E, Giele M, Pelczar P, Ramos L, Wansink DG, van der Vlag J, Peters AH, de Boer P (2007) Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. Nat Genet 39:251–258
74. Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F, Lee C, Almouzni G, Schneider R, Surani MA (2008) Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature 452:877–881
75. Konev AY, Tribus M, Park SY, Podhraski V, Lim CY, Emelyanov AV, Vershilova E, Pirrotta V, Kadonaga JT, Lusser A et al (2007) CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin *in vivo*. Science 317:1087–1090
76. Marfella CG, Imbalzano AN (2007) The Chd family of chromatin remodelers. Mutat Res 618:30–40
77. Ge Q, Nilasena DS, O'Brien CA, Frank MB, Targoff IN (1995) Molecular analysis of a major antigenic region of the 240-kD protein of Mi-2 autoantigen. J Clin Invest 96:1730–1737
78. Seelig HP, Moosbrugger I, Ehrfeld H, Fink T, Renz M, Genth E (1995) The major dermatomyositis-specific Mi-2 autoantigen is a presumed helicase involved in transcriptional activation. Arthritis Rheum 38:1389–1399
79. Seelig HP, Renz M, Targoff IN, Ge Q, Frank MB (1996) Two forms of the major antigenic protein of the dermatomyositis-specific Mi-2 autoantigen. Arthritis Rheum 39:1769–1771
80. Schwab U, Stein H, Gerdes J, Lemke H, Kirchner H, Schaadt M, Diehl V (1982) Production of a monoclonal antibody specific for Hodgkin and Sternberg-Reed cells of Hodgkin's disease and a subset of normal lymphoid cells. Nature 299:65–67
81. Bagchi A, Papazoglu C, Wu Y, Capurso D, Brodt M, Francis D, Bredel M, Vogel H, Mills AA (2007) CHD5 is a tumor suppressor at human 1p36. Cell 128:459–475
82. Thompson PM, Gotoh T, Kok M, White PS, Brodeur GM (2003) CHD5, a new member of the chromodomain gene family, is preferentially expressed in the nervous system. Oncogene 22:1002–1011
83. Mori N, Morosetti R, Mizoguchi H, Koeffler HP (2003) Progression of myelodysplastic syndrome: allelic loss on chromosomal arm 1p. Br J Haematol 122:226–230
84. Mori N, Morosetti R, Spira S, Lee S, Ben-Yehuda D, Schiller G, Landolfi R, Mizoguchi H, Koeffler HP (1998) Chromosome band 1p36 contains a putative tumor suppressor gene important in the evolution of chronic myelocytic leukemia. Blood 92:3405–3409
85. Bieche I, Champeme MH, Matifas F, Cropp CS, Callahan R, Lidereau R (1993) Two distinct regions involved in 1p deletion in human primary breast cancer. Cancer Res 53:1990–1994
86. Cheung TH, Lo KW, Yim SF, Poon CS, Cheung AY, Chung TK, Wong YF (2005) Clinicopathologic significance of loss of heterozygosity on chromosome 1 in cervical cancer. Gynecol Oncol 96:510–515
87. Kleer CG, Bryant BR, Giordano TJ, Sobel M, Merino MJ (2000) Genetic changes in chromosomes 1p and 17p in thyroid cancer progression. Endocr Pathol 11:137–143
88. Sanlaville D, Verloes A (2007) CHARGE syndrome: an update. Eur J Hum Genet 15: 389–399
89. Vissers LE, van Ravenswaaij CM, Admiraal R, Hurst JA, de Vries BB, Janssen IM, van der Vliet WA, Huys EH, de Jong PJ, Hamel BC et al (2004) Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. Nat Genet 36:955–957

90. Kallen K, Robert E, Mastroiacovo P, Castilla EE, Kallen B (1999) CHARGE Association in newborns: a registry-based study. Teratology 60:334–343
91. Bajpai R, Chen DA, Rada-Iglesias A, Zhang J, Xiong Y, Helms J, Chang CP, Zhao Y, Swigut T, Wysocka J (2010) CHD7 cooperates with PBAF to control multipotent neural crest formation. Nature 463:958–962
92. Siebert JR, Graham JM Jr, MacDonald C (1985) Pathologic features of the CHARGE association: support for involvement of the neural crest. Teratology 31:331–336
93. Schnetz MP, Bartels CF, Shastri K, Balasubramanian D, Zentner GE, Balaji R, Zhang X, Song L, Wang Z, Laframboise T et al (2009) Genomic distribution of CHD7 on chromatin tracks H3K4 methylation patterns. Genome Res 19:590–601
94. Roberts C, Daw SC, Halford S, Scambler PJ (1997) Cloning and developmental expression analysis of chick Hira (Chira), a candidate gene for DiGeorge syndrome. Hum Mol Genet 6:237–245
95. Wilming LG, Snoeren CA, van Rijswijk A, Grosveld F, Meijers C (1997) The murine homologue of HIRA, a DiGeorge syndrome candidate gene, is expressed in embryonic structures affected in human CATCH22 patients. Hum Mol Genet 6:247–258
96. Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, Dewell S, Law M, Guo X, Li X et al (2010) Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell 140:678–691
97. Higgs DR, Garrick D, Anguita E, De Gobbi M, Hughes J, Muers M, Vernimmen D, Lower K, Law M, Argentaro A et al (2005) Understanding alpha-globin gene regulation: aiming to improve the management of thalassemia. Ann N Y Acad Sci 1054:92–102
98. Allis CD, Bowen JK, Abraham GN, Glover CV, Gorovsky MA (1980) Proteolytic processing of histone H3 in chromatin: a physiologically regulated event in Tetrahymena micronuclei. Cell 20:55–64
99. Falk MM, Grigera PR, Bergmann IE, Zibert A, Multhaup G, Beck E (1990) Foot-and-mouth disease virus protease 3C induces specific proteolytic cleavage of host cell histone H3. J Virol 64:748–756
100. Duncan EM, Muratore-Schroeder TL, Cook RG, Garcia BA, Shabanowitz J, Hunt DF, Allis CD (2008) Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation. Cell 135:284–294
101. Santos-Rosa H, Kirmizis A, Nelson C, Bartke T, Saksouk N, Cote J, Kouzarides T (2009) Histone H3 tail clipping regulates gene expression. Nat Struct Mol Biol 16:17–22
102. Barrett AJ, Rawlings ND, Woessner JF (eds) (2003) Handbook of proteolytic enzymes. Academic Press, London
103. Troen BR, Gal S, Gottesman MM (1987) Sequence and expression of the cDNA for MEP (major excreted protein), a transformation-regulated secreted cathepsin. Biochem J 246:731–735
104. Hiwasa T, Sakiyama S (1996) Nuclear localization of procathepsin L/MEP in ras-transformed mouse fibroblasts. Cancer Lett 99:87–91
105. Goulet B, Baruch A, Moon NS, Poirier M, Sansregret LL, Erickson A, Bogyo M, Nepveu A (2004) A cathepsin L isoform that is devoid of a signal peptide localizes to the nucleus in S phase and processes the CDP/Cux transcription factor. Mol Cell 14:207–219
106. Boudreau F, Lussier CR, Mongrain S, Darsigny M, Drouin JL, Doyon G, Suh ER, Beaulieu JF, Rivard N, Perreault N (2007) Loss of cathepsin L activity promotes claudin-1 over-expression and intestinal neoplasia. FASEB J 21:3853–3865
107. Denhardt DT, Greenberg AH, Egan SE, Hamilton RT, Wright JA (1987) Cysteine proteinase cathepsin L expression correlates closely with the metastatic potential of H-ras-transformed murine fibroblasts. Oncogene 2:55–59
108. Lah TT, Kos J (1998) Cysteine proteinases in cancer progression and their clinical relevance for prognosis. Biol Chem 379:125–130
109. Goulet B, Sansregret L, Leduy L, Bogyo M, Weber E, Chauhan SS, Nepveu A (2007) Increased expression and activity of nuclear cathepsin L in cancer cells suggests a novel mechanism of cell transformation. Mol Cancer Res 5:899–907

110. Imschenetzky M, Diaz F, Montecino M, Sierra F, Puchi M (1997) Identification of a cysteine protease responsible for degradation of sperm histones during male pronucleus remodeling in sea urchins. J Cell Biochem 67:304–315
111. Concha C, Monardes A, Even Y, Morin V, Puchi M, Imschenetzky M, Geneviere AM (2005) Inhibition of cysteine protease activity disturbs DNA replication and prevents mitosis in the early mitotic cell cycles of sea urchin embryos. J Cell Physiol 204:693–703
112. Morin V, Sanchez A, Quiñones K, Huidobro JG, Iribarren C, Bustos P, Puchi M, Genevière AM, Imschenetzky M (2008) Cathepsin L inhibitor I blocks mitotic chromosomes decondensation during cleavage cell cycles of sea urchin embryos. J Cell Physiol 216:790–795
113. Lee KH, Lee NJ, Hyun S, Park YK, Yang EG, Lee J-K, Jeong S, Yu J (2009) Histone H3 N-terminal peptide binds directly to its own mRNA: a possible mode of feedback inhibition to control translation. Chembiochem 10:1313–1316
114. Bab I, Gavish H, Namdar-Attar M, Muhlrad A, Greenberg Z, Chen Y, Mansur N, Shteyer A, Chorev M (1999) Isolation of mitogenically active C-terminal truncated pentapeptide of osteogenic growth peptide from human plasma and culture medium of murine osteoblastic cells. J Pept Res 54:408–414
115. Bab I, Gazit D, Chorev M, Muhlrad A, Shteyer A, Greenberg Z, Namdar M, Kahn A (1992) Histone H4-related osteogenic growth peptide (OGP): a novel circulating stimulator of osteoblastic activity. EMBO J 11:1867–1873
116. Bab I, Smith E, Gavish H, Attar-Namdar M, Chorev M, Chen YC, Muhlrad A, Birnbaum MJ, Stein G, Frenkel B (1999) Biosynthesis of osteogenic growth peptide via alternative translational initiation at AUG85 of histone H4 mRNA. J Biol Chem 274:14474–14481
117. Smith E, Meyerrose TE, Kohler T, Namdar-Attar M, Bab N, Lahat O, Noh T, Li J, Karaman MW, Hacia JG et al (2005) Leaky ribosomal scanning in mammalian genomes: significance of histone H4 alternative translation *in vivo*. Nucleic Acids Res 33:1298–1308

Histone Modifications in Cancer Biology and Prognosis

Siavash K. Kurdistani

Abstract Cancer is a disease of genome sequence alterations as well as epigenetic changes. Epigenetics refers in part to the mechanisms by which histones affect various DNA-based processes, such as gene regulation. Histones are proteins around which the DNA wraps itself to form chromatin – the physiologically relevant form of the human genome. Histones are modified extensively by posttranslational modifications that alter chromatin structure and serve to recruit to or exclude protein complexes from DNA. Aberrations in histone modifications occur frequently in cancer including changes in their levels and distribution at gene promoters, gene coding regions, repetitive DNA sequences, and other genomic elements. Locus-specific alterations in histone modifications may have adverse effects on expression of nearby genes but so far have not been shown to have clinical utility. Cancer cells also exhibit alterations in global levels of specific histone modifications, generating an additional layer of epigenetic heterogeneity at the cellular level in tumor tissues. Unlike locus-specific changes, the cellular epigenetic heterogeneity can be used to define previously unrecognized subsets of cancer patients with distinct clinical outcomes. In general, increased prevalence of cells with lower global levels of histone modifications is prognostic of poorer clinical outcome such as increased risk of tumor recurrence and/or decreased survival probability. Prognostic utility of histone modifications has been demonstrated independently for multiple cancers including those of prostate, lung, kidney, breast, ovary, and pancreas, suggesting a fundamental association between global histone modification levels and tumor aggressiveness, regardless of cancer tissue of origin. Cellular levels of histone modifications may also predict response to certain chemotherapeutic agents, serving as predictive biomarkers that could inform clinical decisions on choice and course of therapy. The challenge before us is to understand how global levels of histone

S.K. Kurdistani
Departments of Biological Chemistry and Pathology and Laboratory Medicine, Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, David Geffen School of Medicine at UCLA, Los Angeles, California 90095, USA
e-mail: skurdistani@mednet.ucla.edu

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_5,

modifications are established and maintained and what their mechanistic links are to the cancer clinical behavior.

1 Introduction

Cancer epigenetics has become a vibrant and multilayered field of scientific inquiry, providing important insights into the basic biology of cancer as well as presenting novel clinical tools and approaches for development and application of more effective therapeutics [1, 2]. Epigenetics, in its original sense, is the study of inherited phenotypes that are not directly encoded by the DNA sequence. As it relates to chromatin, epigenetics commonly refers to DNA methylation, histone modifications, nucleosome positioning and composition, noncoding RNAs, and other chromatin elements that can regulate DNA-based processes including gene expression. However, in most cases the heritability of these elements has not been demonstrated [3, 4]. The field of cancer epigenetics has been largely focused on the roles of DNA methylation in gene regulation, especially on how promoter hypermethylation leads to repression of genes with tumor suppression functions. In the past few years, other epigenetic processes have been increasingly studied for their contributions to the cancer phenotype. For example, flurries of findings have linked microRNAs to several aspects of cancer biology including growth and metastasis [5, 6]. With the identification of most histone modifying enzymes, studies of histone modifications in cancer have also risen to prominence, spurring generation of an international task force, Alliance for the Human Epigenome and Disease (AHEAD), "to decode the human epigenome" which, among other things, aims to determine the distribution of selected histone modifications across the genome of several cell types including cancer cells [3]. Although the value of such large scale effort has been questioned [7], similar to the Human Genome Project, AHEAD hopes to provide in essence a "reference epigenome" that may be of great value in both basic and applied research.

In this section, I focus mainly on the roles of histone modifications in cancer. The major emphasis will be on the clinical utility of histone modifications, but I will also provide a brief account of how they may lead to cancer initiation.

2 Histone Modifications in Cancer Initiation

Epigenetic alterations that may precede and contribute to the onset of cancer initiation are not well understood. The majority of studies compare already cancerous cells to their normal counterparts, making it impossible to determine which epigenetic alterations are causative or upstream of cancer initiation. In the case of DNA methylation, limited studies of mouse models have suggested that aberrant promoter DNA methylation patterns can increase the risk of future genetic

mutationsby deregulating expression of genes involved in control of cellular growth [8]. In the case of histone modifications, the contributions to cancer initiation are essentially unknown because the complexity of the genetics and molecular biology of histone modifications makes it difficult to generate proper animal models. Mammalian cells contain many copies of each histone gene, making genetic manipulations of histone genes very difficult, if not impossible, given the current technologies. In addition, histone modifications, especially histone acetylation, exhibit redundant functions and can be deposited by more than one enzyme, precluding single genetic changes as a way of assigning unambiguous function.

A model system for the study of early epigenetic alterations required for cellular transformation is viral-induced oncogenesis [9, 10]. DNA tumor viruses, such as the human papilloma virus (HPV), simian virus 40 (SV40) and adenovirus, often infect host animals by infecting cells that are fully differentiated and nondividing and therefore nonconducive for viral replication. Consequently, many of these viruses have evolved proteins that are expressed immediately after infection and exploit the cellular machinery to force the host cell to reenter the cell cycle and divide so that progeny virions are produced. These viral "oncoproteins," such as HPV E6/E7, SV40 large T antigen, and adenovirus small e1a (a splice variant of large E1A), have become invaluable tools for uncovering fundamental molecular processes that regulate cell proliferation. For instance, studies of e1a-induced transformation of normal cells have helped elucidate the functions of retinoblastoma (RB) tumor suppressor, and its family members p130 and p107, in controlling the cell cycle, and the importance of p53 inactivation in tumorigenesis [9, 10]. Interestingly, e1a also interacts with a number of histone and chromatin modifiers such as p300, CBP, Gcn5, p400, some of which are essential for the ability of e1a to induce quiescent, nondividing cells to enter S phase [11]. Some of these epigenetic modifiers that interact with viral oncoproteins are mutated in nonviral, primary human cancer, suggesting that deregulated function of chromatin modifiers may be important for carcinogenesis [12, 13]. If we can understand how and when e1a or other viral oncoproteins utilize chromatin modifying enzymes in the course of cellular transformation, we could generate testable hypotheses as to how the same chromatin modifiers function to promote primary cancer.

We have recently used an adenoviral-induced cell transformation system to understand specifically the consequences of e1a interactions with p300 and its close homologue CBP, histone acetyltransferases (HAT). The p300 protein was first identified as an e1a-interacting protein and shown to be required for the oncogenic properties of e1a [9]. We discovered that e1a-p300/CBP interaction causes an ~70% reduction of histone H3 lysine 18 acetylation (H3K18ac) but has no effect on global levels of several other histone modifications [14]. The remaining H3K18ac is redistributed to the regulatory regions of genes involved in cell cycle progression, which are upregulated in the e1a-infected cells [15]. The reduction and redistribution of H3K18ac is associated with an orchestrated and precise rearrangement of multiple regulators of gene expression including p300/CBP and other proteins with epigenetic activities. This e1a-induced reprogramming results in S-phase

induction in cells that should remain in a resting state [15]. The inappropriate entry into S-phase is in fact a fundamental feature of virtually all cancers [16].

Although the exact role of H3K18ac remains to be determined, the unique effect of e1a on acetylation levels of this residue has underscored H3K18 as a distinct and important site of histone modification implicated in cell transformation. Such insight would have been hidden in other standard assays such as genome-wide mapping experiments. Therefore, the cellular transformation by viral oncoproteins could prove to be a powerful system to study the epigenetic changes that occur during an oncogenic process. This system is reproducible and amenable to manipulation and temporal measurements with proven relevance to primary human cancer. Considering the rich history of DNA tumor viruses, the oncogenic reprogramming by e1a provides a unique opportunity to understand how utilization of epigenetic modifiers leads to cellular transformation.

3 Histone Modifications as Clinical Tools

Cancer is a heterogeneous disease, resulting in different clinical outcomes even for individuals with the same affected tissue such as prostate or breast cancer. Clinical outcome may be measured as, but is not limited to, risk of tumor recurrence after removal of the primary tumor, risk of metastasis, survival probability, and/or degree of response to therapeutic agents. The ability to assess or predict the clinical behavior of cancers is critical for clinicians to determine the most appropriate treatments including type, intensity, and duration of therapies [17].

For most cancers, clinical outcome prediction is based on tumor stage, which is a measure of disease burden (e.g., tumor size) and degree of spread from the site of primary tumor. Additional predictive information may be provided by tumor grade (i.e., degree of differentiation), histological type, and patient demographics. In general, lower stage and/or grade cancers have better prognosis. However, each stage or grade category still contains large groups of patients who, although as a group behave similarly compared with patients in other stage and grade categories, still display significant differences in the course and outcome of disease. Therefore, subdividing patients into smaller groups with more cohesive clinical behavior is useful for the development of targeted and more effective therapies and eventually for personalization of patient care. This need is more pressing for patients with lower stage or grade tumors for which the choice of therapy and its intensity is not necessarily evident. In this regard, there has been much effort to discover molecular biomarkers that can stratify cancer patients with distinct clinical outcomes to expand our prognostic capabilities.

The molecular biomarkers include single nucleotide polymorphisms, chromosomal translocations, gene mutations, expression patterns of groups of genes, methylation status of specific gene promoters, or secreted proteins. In most cases, the relevant prognosticator biomarker is different for different cancers, since the genetic mutations or gene expression patterns reflect the cancer's tissue of origin.

Interestingly, while the number of potential biomarkers in the literature is steadily increasing, the number of FDA-approved biomarkers per year is not increasing but rather decreases [17]. This is because approval for clinical use requires that extensive statistical and clinical criteria are met. But the majority of biomarkers identified in initial studies eventually do not meet these criteria. Nonetheless it is clear that cancer biomarkers will play increasingly important roles in cancer treatment in the coming years.

3.1 Alterations of Histone Modifications in Cancer

Altered patterns of histone modifications are commonly observed in cancer. Promoter regions contain important regulatory sequences for transcription control of nearby genes. As a result, the vast majority of alterations in histone modifications have been identified through examination of one or more gene promoters. Deregulation of histone modifications at an individual promoter is intimately linked to misexpression of the downstream gene, which may have critical consequences for the cancer phenotype. However, to my knowledge, none of the promoter-specific changes in histone modifications has so far been related correlatively or causally to clinical outcome. This is perhaps due to the fact that changes have been mapped for only a few histone modifications at a relatively small number of loci. These are likely to be inadequate for clinical outcome predictions.

An additional level of variability in distribution patterns of histone modifications is their differences in global levels between individual cells within a given tissue. Immunohistochemical examination of histone modifications in primary tissues using site-specific antibodies has revealed dissimilar global levels of histone modifications in individual cells (Fig. 1a). There is a great deal of heterogeneity in the percentage of cells that are positively stained and in the intensity of staining (brown nuclei in Fig. 1a), not only between patients but within a single patient as well. At the global level, the cells that do not stain for histone modifications (blue nuclei in Fig. 1a) may still contain the histone modifications at few genomic loci, but their levels are below the detection limits of immunohistochemistry (IHC). The dissimilar levels of specific histone modifications in different cells generate a diversity of epigenetic patterns within cell populations that can nonetheless be readily quantified by pathologists as "percent cell staining." Other measures that take the intensity of staining into account can also be applied. Remarkably, this cellular epigenetic heterogeneity is predictive of cancer clinical behavior and may also correlate with the degree of response to certain chemotherapeutics.

3.2 Histone Modifications as Prostate Cancer Prognostic Markers

In the initial study that linked global patterns of histone modifications to clinical outcome, the global levels H3K4me2, H3K9ac, H3K18ac, H4R3me2, and H4K12ac

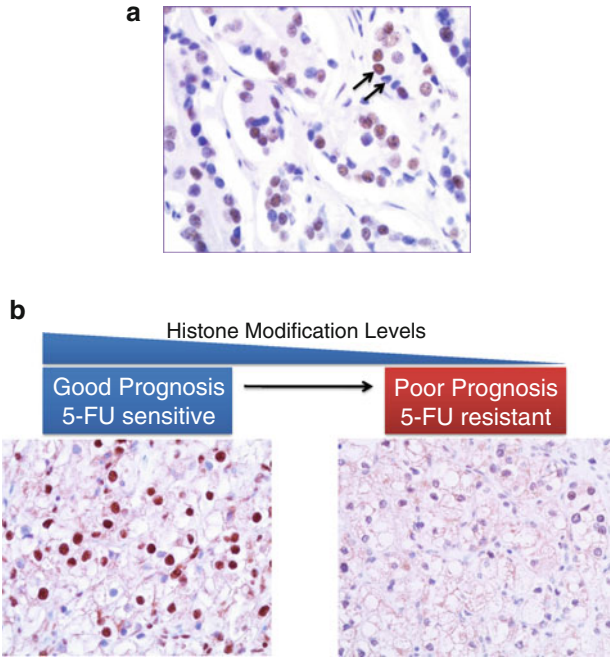


Fig. 1 Histone modifications as cancer prognostic markers. (**a**) A primary cancer tissue immunostained with anti-H3K18ac antibody shows cellular heterogeneity in global levels of H3K18ac. The arrows point to neighboring cancer cell nuclei that have high (*brown*) and low (*blue*) levels of H3K18ac. (**b**) Global levels of histone modifications predict prognosis in multiple cancers. Generally, lower levels of histone modifications are predictive of poorer prognosis. Lower levels of histone modifications also predict poorer response to 5-FU therapy in pancreatic cancer. Treatment of cancers with low levels of histone modifications and poor prognosis with histone deacetylase (HDAC) and demethylase (HDM) inhibitors may shift the global patterns of histone modifications to a more favorable prognosis. In certain cancers such as pancreas adenocarcinoma, the assessment of global histone modification levels as shown may inform the choice of therapy

were examined by IHC in primary prostate cancer tissues [18]. The choice of these modifications was based on their known biological roles in gene regulation, but also on the abilities of their respective antibodies to work on suitably fixed, paraffin-embedded tissues. The prostate cancer tissues were collected from patients who had undergone radical prostatectomy. Importantly, the tissues were associated with follow-up clinical and pathological data, enabling retrospective studies.

All five modifications showed cellular heterogeneity, with some fraction of cells (0–100%) staining positively in each tissue. While no single modification was predictive on its own, unbiased clustering of patients, based on the global histone modification patterns, defined two groups of patients with high or low levels of histone modifications with significantly different clinical outcomes. Surprisingly,

patients with lower global levels of histone modifications (i.e., decreased percent cell staining) had poorer prognosis with increased risk of tumor recurrence after removal of the primary tumor. This predictive power was independent of all clinicopathological variables tested, indicating that histone modifications provide additional prognostic information that would otherwise not be revealed by what we already know. A second independent set of patient tissues validated the result of the primary study, confirming the statistical significance of low global levels of modifications predicting poorer clinical outcome [18].

The fraction of histones that must be modified for an antibody to generate a detectable IHC signal is unclear, but certainly varies for different antibodies, since each will have different affinity for its antigen. Notwithstanding these differences, the histone modification levels (i.e., percent cell staining) correlated positively among different antibodies across all tissue specimens. Tissues with increased (or decreased) percentage of positively stained cells for a given modification were likely to have high (or low) percentages of other modifications [18]. This is somewhat surprising because the examined modifications, although they are all associated with activated gene expression, have different distributions throughout the genome and may be enriched at different sets of genes. Note that the low cellular levels of histone modifications are unlikely to represent a complete absence of these marks at the molecular level from the genome. Nonetheless on a global level, some cancers do show gain or loss of all five modifications. This suggests that the global levels of different histone modifications may be regulated through the same or related pathways (see below).

3.3 Histone Modifications as General Prognostic Markers for Adenocarcinomas

Is the identified prognostic histone modification pattern specific to prostate cancer? In the initial study, the focus on prostate cancer was based on the availability of a tissue cohort with follow-up clinical data. Since histones and their modifications are highly conserved and ubiquitously present, the data from prostate cancer suggested that cellular histone modification patterns could be informative of prognosis in other cancers as well. This is indeed the case for H3K4me2 and H3K18ac, the two modifications that were most informative about prostate cancer. These two modifications, in a pattern similar to that identified in prostate cancer, predict two disease subtypes with significantly distinct survival probabilities in lung and kidney cancers [19]. In all cancers, the predictions are independent of relevant clinicopathological variables, indicating that histone modifications provide unique prognostic information. Generally, patients who have higher percentage of cancer cells that stain positively for H3K4me2 and H3K18ac have better prognosis than those with lower percentages (Fig. 1b). So, the increased prevalence of malignant cells with little or no detectable H3K4me2 and H3K18ac is associated with poorer

outcome in cancers with distinct tissues of origin. It is remarkable that despite the genetic and gene expression differences between these three different cancers, the epigenetic changes at the cellular level are similar. The broad applicability of histone modifications as prognostic markers in these adenocarcinomas is unique among cancer biomarkers.

H3K4me2 and H3K18ac, as well as the other three modifications examined in prostate cancer, are all associated with gene activity at the molecular level. It was conceivable that modifications that correlate with gene repression may have inverse correlations with clinical outcome, i.e., high levels, poor prognosis, as compared with the active modifications. However, this is not the case. Surprisingly, levels of H3K9me2, a modification that is associated with gene repression, revealed similar associations with clinical outcome as other modifications. Low levels of H3K9me2 were associated with poorer outcome in both prostate (increased risk of recurrence) and kidney (decreased survival) cancers [19]. Thus, loss of modifications irrespective of their effect on gene transcription is associated with an aggressive cancer phenotype. It should be noted that all aforementioned cancers are adenocarcinomas derived from glandular epithelium. The predictive patterns of histone modifications may differ among tumors of different histological derivation (see below). Nonetheless, the overall similarity in cancer-associated epigenetic patterns in different cancers suggests the existence of a common molecular process that regulates the epigenetic state of cancers.

The tumor suppressor p53 is mutated in a large fraction of all cancers, and p53 mutation is generally associated with poorer outcome. Surprisingly, in both lung and kidney cancers, the poorer prognostic patterns of histone modifications (i.e., lower global levels) correlated with wild-type p53 status, suggesting that the decrease in global levels of histone modifications is independent of misregulation of p53-mediated pathways [19]. The significance of this observation remains to be determined.

Due to inherent biases in clinical datasets (e.g., small sample size or bias in tissue collection), it is critical that findings from one dataset be replicated and validated by others with independent tissue cohorts. The prognostic power of global histone modification levels has now been validated in several laboratories, including cancers of breast [20], pancreas [21, 22], prostate [18, 19, 23], ovary [22], lung [19, 24, 25], and squamous cell carcinoma of esophagus [26]. Table 1 lists the published studies that have examined the prognostic power of histone modifications in different cancers to date.

The study by Elsheikh et al. not only confirmed the prognostic value of histone modification in breast cancer, but also discovered interesting associations with histological subtypes [20]. This analysis of >800 breast cancers and seven modifications represents the largest study to date examining the association of histone modifications with clinical parameters of cancer. The authors found three clusters of patients with high, intermediate, and low levels of histone modification, which correlated progressively with outcome: as global levels of modifications decreased, outcome worsened. Interestingly, the high modification group was enriched for luminal type breast cancer whereas the low modification group comprised basal

Table 1 Global histone modification patterns predict prognosis in multiple cancers

Study	Cancer type	Histone modifications
Barlesi et al. 2009	Lung	H2AK5ac, H2BK12ac, H3K4me2, H3K9ac, H4K8ac
Ellinger et al. 2009	Prostate	H3K4mel, me2, me3, H3K9me1, me2, me3, H3ac, H4ac
Elsheikh et al. 2009	Breast	H3K4me2, H3K9ac, H3K18ac, H4R3me2, H4K12ac, H4K16ac, H4K20me3
Manuyakorn et al. 2009[a]	Pancreas	H3K4me2, H3K9me2, H3K18ac
Park et al. 2008	Stomach	H3K9me3, H4K16ac, H3K20me3
Seligson et al. 2005	Prostate	H3K4me2, H3K9ac, H3K18ac, H4R3me2, H4K12ac
Seligson et al. 2009	Lung	H3K4me2, H3K18ac
	Kidney	H3K4me2, H3K18ac, H3K9me2
	Prostate	H3K9me2
Tzao et al. 2009	Esophagus	H3K4me2, H3K18ac, H3K27me3, H4R3me2, H4K12ac
Van Den Broeck et al. 2008	Lung	H4K5ac, H4K8ac, H4K12ac, H4K16ac, H4K20me3
Wei at al. 2008	Breast	H3K27me3
	Ovary	H3K27me3
	Pancreas	H3K27me3

[a]This study also examined the ability of global histone modification patterns to predict response to chemotherapeutics

A list of publications, cancer types and histone modifications examined to date are indicated

type and HER2 (Human Epidermal growth factor Receptor 2) positive breast cancers. The luminal and basal subtypes refer to the cellular origin of cancer cells and have different gene expression patterns; the prognosis is better for luminal vs. basal subtype. HER2 is a receptor tyrosine kinase, which when over-expressed confers a poorer prognosis in breast cancer. So, lower levels of histone modifications were associated with more aggressive subtypes of cancer. Nonetheless, the modifications showed distinct clinic-pathological associations, suggesting some specificity to their functions. For instance, H3K18ac appeared to be an independent prognosticator, identifying patients with different survival times, while low levels of H4R3me2, H3K9ac, and H4K16ac were associated with large tumor size, high levels of H4R3me2 and H3K9ac with low lymph node stage, and low levels of H4K16ac with vascular invasion. Notably, levels of H4K16ac were low in the majority of tumors, prompting the authors to suggest that "loss of H4K16ac may be an early event in the pathogenesis of invasive breast cancer [20]." Loss of H4K16ac (and H4K20me3) also occurred early in the development of a mouse model of multistage skin carcinogenesis [27]. These data indicate that while loss of histone modifications is generally associated with poorer outcome, individual modifications may still contribute differentially to cancer progression. For this reason, more detailed mapping of specific modifications may be needed for complementary prognostic information.

A limited number of studies have shown the opposite relationship between histone modifications and clinical outcome. Park et al. [28] showed that in gastric adenocarcinoma, higher levels of H3K9me3 were associated with decreased

survival. They also found that a large majority of gastric caners stain positively for H4K16ac. Tzao et al. [26] found that better survival in patients with squamous cell carcinoma of the esophagus is associated with low levels of H3K18ac or H3K27me3. If the results of Park et al. and Tzao et al. [26, 28] can be independently validated, they may indicate that the relationship of histone modification levels and clinical outcome in some cancers may be opposite that found in cancers of prostate, breast, pancreas, lung, and ovary.

3.4 Histone Modifications as Therapeutic Response Markers

Predictive prognostic markers for the aggressivity of cancer suggest implicitly that differential treatment may be beneficial for patients. However, until it is shown directly that such information can be used to deliver effective therapies with measurable benefits for patients, the prognostic information may be of limited clinical interest. Another important area in which biomarkers can be useful is prediction of response to chemotherapeutics. A recent study has provided the initial evidence that histone modifications can be useful in this area as well [21].

Manuyakorn et al. [21] examined the ability of three histone modifications, H3K4me2, H3K9me2, and H3K18ac, to predict response of patients with pancreatic cancer to 5-fluorouracil (5-FU) and gemcitabine. 5-FU is an inhibitor of thymidylate synthase, which converts uridine to thymidine for DNA synthesis; 5-FU is therefore a nucleotide synthesis inhibitor. Gemcitabine is a nucleoside analog and inhibits DNA synthesis by incorporating in the replicating DNA chain. Gemcitabine is also a ribonucleotide reductase (RNR) inhibitor, depleting cells of deoxyribonucleotides required for DNA replication. The pancreatic cancer tissues were collected as part of a phase III randomized postoperative adjuvant treatment trial comparing 5-FU to gemcitabine before and after chemoradiation (RTOG 9704). As in other cancers, histone modification levels were highly significant and independent prognostic factors in pancreatic cancer, with lower levels predicting poorer survival probability. More important, low cellular levels of histone modifications predicted worse survival outcome for patients receiving adjuvant 5-FU chemotherapy, but not for those receiving gemcitabine. These data suggest that cellular levels of histone modifications define previously unrecognized subsets of pancreatic adenocarcinoma patients with distinct response to 5-FU (Fig. 1b).

The clinical value of predicting response to therapeutics has implications for single or combinatorial therapies. For instance, 5-FU is also used in treatment of other malignancies including cancers of breast, colon, rectum, esophagus, and gallbladder. Histone modification levels could possibly serve as predictive biomarkers for adjuvant 5-FU therapy, perhaps in combination with other 5-FU predictive markers such as thymidylate synthase, which has also been associated with resistance to 5-FU chemotherapy [29]. Given that low global histone acetylation is associated with worse response to 5-FU, it is not surprising to note that histone deacetylase inhibitors (HDACi), which increase global histone acetylation levels,

function in synergy with 5-FU. HDAC inhibitors enhance 5-FU's cytotoxic and growth inhibitory effects in cancer cell lines [30, 31], raising the possibility of synergistic effects between certain chemotherapeutics and drugs that target epigenetic modifiers.

While still speculative, an analysis of histone modifications at the global level may inform the choice and regimen of various "epigenetic drugs" against HDACs or histone demethylases. Patients with low levels of histone modifications could perhaps benefit more from HDACis or require a different regimen than those with high levels of histone modification. The current HDACis in clinical trials inhibit multiple HDACs, although *in vivo* each HDAC has distinct specificities for histones and lysine residues. IHC analysis of histone modifications in primary cancer tissues may help guide the development of HDAC-specific inhibitors and their use in a more targeted approach to histone modification.

Inhibition of HDACs results in pleiotropic effects on cancer cells including growth arrest, apoptosis, and differentiation [32]. It is unclear whether these effects are linked to transcriptional reactivation of certain genes with tumor suppressor functions and/or simply shift the global patterns of histone modifications to the more favorable prognostic category (Fig. 1b). Whatever the case may be, the simplicity and robustness of IHC analysis of histone modifications should facilitate the development of standard and effective assays for patient stratification.

4 Regulatory Mechanisms of Global Histone Modification Levels

Why are there global alterations in the level of histone modifications? Histone modifications are maintained dynamically through enzymes with opposing activities, such as histone acetyltransferases-deacetylases and methyltransferases-demethylases (HMTs/HDMs). The steady-state global levels of histone modifications thus result from a balance between enzymes that add and those that remove a modification. Histone modifying enzymes commonly reside in multiprotein complexes which can be recruited by distinct transcription factors to different classes of genes where they function as transcriptional coactivators or corepressors. Thus, global modulation of histone modifications includes promoter-specific recruitment of histone-modifying enzymes by transcription factors.

However, the global decrease in histone modifications that is observed by IHC is unlikely to arise from changes at a few gene promoters. This was examined in two prostate cancer cell lines, PC3 and LNCaPs. PC3 cells have ~50% less H3K9me2 levels than LNCaPs [19]. Chromatin immunoprecipitation combined with microarrays (ChIP-chip) failed to detect consistent differences in ~17,000 promoters that could account for the differences in global H3K9me2 levels. However, analysis of histones associated with several DNA repetitive elements showed consistently lower H3K9me2 levels in PC3 vs. LNCaP cells [19]. Lower global levels of H4K16ac and

H4K29me3 in hematological malignancies could also be mapped to DNA repeat elements [27]. Considering that such repetitive elements make up the bulk of DNA in human cells, it is likely that global levels of histone modifications detected by IHC reflect their molecular levels at DNA repetitive elements. These elements also often show loss of DNA methylation in cancer cells. So, the DNA repetitive elements in cancer cells are generally "de-modified" both on DNA and their associated histones, potentially conferring a more aggressive phenotype to the cancer cell.

What mechanism(s) lead to loss of histone modifications? Obviously either loss of enzyme expression or aberrant recruitment of histone modifying enzymes to DNA repetitive elements could account for loss of modifications at these regions. Whereas mutations and changes in expression of histone modifying enzymes have been reported in cancer, it is difficult to envision a scenario in which different histone modifying enzymes are all affected similarly. Indeed, high expression of various HDACs has been correlated both with improved and poor prognosis [33]. One study that examined histone modification levels and HDAC expression in the same breast cancer tissues found that loss of histone modifications was associated surprisingly with reductions in expression of HDACs 1, 2, and 6 [34]. This counterintuitive finding suggests a more complex connection between histone modifying enzymes and histone modifications themselves. The number and diversity of histone modifying enzymes (i.e., 18 HDACs in humans) with overlapping and redundant substrate specificities pose significant challenges for understanding how these enzymes combine to generate specific histone modification patterns.

Altered allocation of acetyl coenzyme A (AcCoA) and S-adenosyl methionine (SAM), which are required by HATs and HMTs to modify histones, could also explain the loss of histone modifications in more aggressive cancers. AcCoA and SAM lie at the center of metabolic pathways required for cell growth and division. AcCoA is generated through breakdown of glucose, fatty acids, and amino acids and used for anaplerotic pathways such as the Krebs (TCA) cycle. Many anabolic pathways such as the de novo synthesis of fatty acids require AcCoA. SAM is the "one-carbon" currency of the cell and is used by a variety of methyltransferases for transfer of single methyl groups to DNA, RNA, and proteins. AcCoA and SAM reside in different pools within the cell that are not freely interchangeable, restricting their use for certain purposes. In fact, regulated transport mechanisms exist to ensure transfer of these compounds between different pools. An example is the citrate shuttle which results in the net transfer of one AcCoA molecule out of mitochondria and into the cytosol. In the case of histone acetylation, the enzymes in the metabolism of AcCoA are required for maintenance of proper levels of histone acetylation. For instance, ATP-citrate lyase (ACL) converts the citrate that is transported out of the mitochondria to AcCoA in the cytoplasm. This ACL-generated AcCoA is required for histone acetylation but not acetylation of other proteins such as p53 [35]. Similar pathways could also preferentially affect histone methylation.

Cancer cells do exhibit extensively altered metabolism, which is required to support their deregulated division [36]. They increase glucose uptake and breakdown significantly to generate high levels of ATP and intermediates for other pathways such as nucleotide synthesis [36]. Cancer cells also rely mainly on de novo synthesis of nucleotides, as opposed to the salvage pathways, for DNA replication. Because cancer cells proliferate rapidly, they have a continuous need for macromolecular biosynthesis which puts a premium on AcCoA, SAM, and other metabolites for anabolism. It is conceivable that in cancers with low levels of histone acetylation and methylation, AcCoA and SAM are partially diverted away from histones to more vital pathways. Limited availability of these compounds for HAT or HMT reactions would result in global loss of most, if not all, histone modifications. This may extend to DNA methylation as well: Global decreases in DNA methylation of cancer cells may also be tracked back to a limited availability of SAM to DNA methyltransferases. If rerouting of AcCoA and SAM to pathways critical for growth and division enhances cancer cell aggressivity, then it may be of interest to ensure that histone acetylation and methylation are maintained in a genome. This hypothesis suggests that HDACi therapy may be inappropriate for cancers with low levels of histone acetylation, because the problem may rather be a shortage of AcCoA for histone acetylation. The "limited cofactor availability" hypothesis needs to be tested and requires a better understanding of metabolic rewiring by cancer cells.

While changes in histone modifications at promoter regions may predictably alter gene expression, the consequences of global decrease in the levels of multiple histone modifications are more difficult to foresee as the altered histone modifications are associated with both transcriptional activation (e.g., H3K4me2 and H3K18ac) and repression (H3K27me3, H3K9me2). A reduction of global histone modification levels may result in increased heterochromatin formation and silencing of large regions of the genome or may provide cells a protective measure against genotoxic stress by limiting DNA exposure. Another possibility is that global reductions in histone modifications, similar to global DNA hypomethylation, lead to genomic instability. In support of this hypothesis, reduction in H3K9me2 levels by knockdown of the histone methyltransferase G9a results in chromosomal instability in cancer cell lines [37]. The suggested possibilities are not mutually exclusive and may each partially explain the observed reduced global levels of histone modifications in subsets of more clinically aggressive cancers.

5 Beyond Biomarkers: Future Outlook

Cancer patients display varied clinical behavior, ranging from indolent, slow growing tumors to highly aggressive, metastatic disease. Accumulating data from several patient cohorts indicate that global histone modification patterns are correlated with aggressiveness of cancers and can be used as prognostic markers.

Global histone modifications may also directly inform therapeutic options by predicting response to certain drugs. However, considering the critical roles of histone modifications in regulating the genetic information, it becomes tempting to speculate that the heterogeneity in histone modifications may partly cause the clinical heterogeneity that is seen within cancer patients. To determine if this is the case, several questions need to be answered. Do cancer cells with low levels of histone modifications have a clonal origin and eventually out-compete other cells in the clinically aggressive subset of cancers? Or do histone modifications change differentially in different cells in response to signals or selective pressures from the local microenvironment milieu? Are global histone modification patterns associated with specific genetic mutations and/or metabolic states? Could we find ways to reversibly induce global changes in levels of histone modifications in cancer cell lines or model systems?

Histone modifications have become the *Ulysses* in the library of epigenetics, compositionally complex but carefully structured with broad implications for human biology and disease.

Acknowledgment S.K.K. is funded by the Howard Hughes Medical Institute (early career award), the California Institute for Regenerative Medicine (new faculty award) and the American Cancer Society.

References

1. Esteller M (2008) Epigenetics in cancer. N Engl J Med 358:1148–1159
2. Gal-Yam EN, Saito Y, Egger G, Jones PA (2008) Cancer epigenetics: modifications, screening, and therapy. Annu Rev Med 59:267–280
3. American Association for Cancer Research Human Epigenome Task Force; European Union, Network of Excellence, Scientific Advisory Board (2008) Moving AHEAD with an international human epigenome project. Nature 454:711–715
4. Ptashne M (2007) On the use of the word 'epigenetic'. Curr Biol 17:R233–R236
5. Tavazoie SF, Alarcon C, Oskarsson T, Padua D, Wang Q, Bos PD, Gerald WL, Massague J (2008) Endogenous human microRNAs that suppress breast cancer metastasis. Nature 451:147–152
6. Valastyan S, Reinhardt F, Benaich N, Calogrias D, Szasz AM, Wang ZC, Brock JE, Richardson AL, Weinberg RA (2009) A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis. Cell 137:1032–1046
7. Madhani HD, Francis NJ, Kingston RE, Kornberg RD, Moazed D, Narlikar GJ, Panning B, Struhl K (2008) Epigenomics: a roadmap, but to where? Science 322:43–44
8. Sakatani T, Kaneda A, Iacobuzio-Donahue CA, Carter MG, de Boom WS, Okano H, Ko MS, Ohlsson R, Longo DL, Feinberg AP (2005) Loss of imprinting of Igf2 alters intestinal maturation and tumorigenesis in mice. Science 307:1976–1978
9. Berk AJ (2005) Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus. Oncogene 24:7673–7685
10. DeCaprio JA (2009) How the Rb tumor suppressor structure and function was revealed by the study of Adenovirus and SV40. Virology 384:274–284
11. Ferrari R, Berk AJ, Kurdistani SK (2009) Viral manipulation of the host epigenome for oncogenic transformation. Nat Rev Genet 10:290–294

12. Gayther SA, Batley SJ, Linger L, Bannister A, Thorpe K, Chin SF, Daigo Y, Russell P, Wilson A, Sowter HM et al (2000) Mutations truncating the EP300 acetylase in human cancers. Nat Genet 24:300–303
13. Giles RH, Peters DJ, Breuning MH (1998) Conjunction dysfunction: CBP/p300 in human disease. Trends Genet 14:178–183
14. Horwitz GA, Zhang K, McBrian MA, Grunstein M, Kurdistani SK, Berk AJ (2008) Adenovirus small e1a alters global patterns of histone modification. Science 321:1084–1085
15. Ferrari R, Pellegrini M, Horwitz GA, Xie W, Berk AJ, Kurdistani SK (2008) Epigenetic reprogramming by adenovirus e1a. Science 321:1086–1088
16. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100:57–70
17. Ludwig JA, Weinstein JN (2005) Biomarkers in cancer staging, prognosis and treatment selection. Nat Rev Cancer 5:845–856
18. 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
19. Seligson DB, Horvath S, McBrian MA, Mah V, Yu H, Tze S, Wang Q, Chia D, Goodglick L, Kurdistani SK (2009) Global levels of histone modifications predict prognosis in different cancers. Am J Pathol 174:1619–1628
20. Elsheikh SE, Green AR, Rakha EA, Powe DG, Ahmed RA, Collins HM, Soria D, Garibaldi JM, Paish CE, Ammar AA et al (2009) Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. Cancer Res 69:3802–3809
21. Manuyakorn A, Paulus R, Farrell J, Dawson NA, Tze S, Hines OJ, Reber H, Seligson DB, Horvath S, Kurdistani SK et al (2009) Cellular histone modification patterns predict prognosis and treatment response in resectable pancreatic adenocarcinoma: results from RTOG 9704. J Clin Oncol 28(8):1358–1365
22. Wei Y, Xia W, Zhang Z, Liu J, Wang H, Adsay NV, Albarracin C, Yu D, Abbruzzese JL, Mills GB et al (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
23. Ellinger J, Kahl P, von der Gathen J, Rogenhofer S, Heukamp LC, Gutgemann I, Walter B, Hofstadter F, Buttner R, Muller SC et al (2010) Global levels of histone modifications predict prostate cancer recurrence. Prostate 70(1):61–69
24. 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
25. Van Den Broeck A, Brambilla E, Moro-Sibilot D, Lantuejoul S, Brambilla C, Eymin B, Khochbin S, Gazzeri S (2008) Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. Clin Cancer Res 14:7237–7245
26. Tzao C, Tung HJ, Jin JS, Sun GH, Hsu HS, Chen BH, Yu CP, Lee SC (2009) Prognostic significance of global histone modifications in resected squamous cell carcinoma of the esophagus. Mod Pathol 22:252–260
27. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K et al (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet 37:391–400
28. Park YS, Jin MY, Kim YJ, Yook JH, Kim BS, Jang SJ (2008) The global histone modification pattern correlates with cancer recurrence and overall survival in gastric adenocarcinoma. Ann Surg Oncol 15:1968–1976
29. Ciaparrone M, Quirino M, Schinzari G, Zannoni G, Corsi DC, Vecchio FM, Cassano A, La Torre G, Barone C (2006) Predictive role of thymidylate synthase, dihydropyrimidine dehydrogenase and thymidine phosphorylase expression in colorectal cancer patients receiving adjuvant 5-fluorouracil. Oncology 70:366–377
30. Lee JH, Park JH, Jung Y, Kim JH, Jong HS, Kim TY, Bang YJ (2006) Histone deacetylase inhibitor enhances 5-fluorouracil cytotoxicity by down-regulating thymidylate synthase in human cancer cells. Mol Cancer Ther 5:3085–3095

31. Tumber A, Collins LS, Petersen KD, Thougaard A, Christiansen SJ, Dejligbjerg M, Jensen PB, Sehested M, Ritchie JW (2007) The histone deacetylase inhibitor PXD101 synergises with 5-fluorouracil to inhibit colon cancer cell growth *in vitro* and *in vivo*. Cancer Chemother Pharmacol 60:275–283
32. Minucci S, Pelicci PG (2006) Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. Nat Rev Cancer 6:38–51
33. Weichert W (2009) HDAC expression and clinical prognosis in human malignancies. Cancer Lett 280:168–176
34. Suzuki J, Chen YY, Scott GK, Devries S, Chin K, Benz CC, Waldman FM, Hwang ES (2009) Protein acetylation and histone deacetylase expression associated with malignant breast cancer progression. Clin Cancer Res 15:3163–3171
35. 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
36. Vander Heiden MG, Cantley LC, Thompson CB (2009) Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science 324:1029–1033
37. Kondo Y, Shen L, Ahmed S, Boumber Y, Sekido Y, Haddad BR, Issa JP (2008) Down-regulation of histone H3 lysine 9 methyltransferase G9a induces centrosome disruption and chromosome instability in cancer cells. PLoS One 3:e2037

Dynamics of Histone Lysine Methylation: Structures of Methyl Writers and Erasers

Anup K. Upadhyay and Xiaodong Cheng

Abstract In Eukarya, the packaging of DNA into chromatin provides a barrier that allows for regulation of access to the genome. Chromatin is refractory to processes acting on DNA. ATP-dependent chromatin remodeling machines and histone-modifying complexes can overcome this barrier (or strengthen it in silencing processes). Both components of chromatin (DNA and histones) are subject to postsynthetic covalent modifications, including methylation of lysines (the focus of this chapter). These lysine marks are generated by a host of histone lysine methyltransferases (writers) and can be removed by histone lysine demethylases (erasers). Importantly, epigenetic modifications impact chromatin structure directly or can be read by effector regulatory modules. Here, we summarize current knowledge on structural and functional properties of various histone lysine methyl-transfereases and demethylases, with emphasis on their importance as druggable targets.

1 Introduction

Unlike lysine acetylation, methylation of lysines does not alter the effective charge, but the hydrophobic and steric properties. The degree of lysine methylation can be mono-, di-, or tri-methylated depending on the specific functional properties of the associated methyltransferase [1–3]. These different lysine methylation marks serve as the binding site for different effector proteins with cognate recognition domains specific to different methylated lysine residues. For example, plant homeodomain (PHD) of bromodomain-PHD-transcription-factor (BPTF) binds tri- or di-methylated lysine 4 of histone H3 (H3K4me3/me2) and recruits the nucleosome

A.K. Upadhyay and X. Cheng
Department of Biochemistry, Emory University School of Medicine, 1510 Clifton Road, Atlanta, Georgia 30322, USA

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_6,

remodeling factor (NURF) complex to the target gene leading to gene activation [4, 5]. In an opposite mechanism, the chromodomain of heterochromatin protein 1 (HP1) binds tri-methylated lysine 9 of histone H3 (H3K9me3) mark, which initiates heterochromatin formation and gene silencing [6, 7].

Recent evidences have indicated that specific recognition domains, either present in a protein complex or in the same polypeptide, combinatorially recognize different histone modifications through a crosstalk mechanism leading to the propagation of active or repressive state of the chromatin. One such example includes the polycomb repressive complex 2 (PRC2) in maintaining and propagating repressive tri-methylated lysine 27 of histone H3 (H3K27me3) through allosteric interaction between EZH2 and EED subunits [8]. Similar examples also include histone lysine methylating enzymes like mammalian G9a and G9a-like protein (GLP) (for H3K9me2/me1) and yeast Clr4 (for H3K9me3), containing both a catalytic SET domain and methyl-lysine recognition module (ankyrin repeats or chromodomain) within the same polypeptide [9, 10]. Therefore, methylation of specific lysines on histones regulates the recruitment of various downstream DNA processing proteins onto the chromatin, which in turn regulate a multitude of biological processes including heterochromatin formation, X-chromosome inactivation, DNA methylation, and gene silencing [11, 12].

The extensively studied histone lysine methylation marks include lysines 4, 9, 27, 36, and 79 of histone H3 and lysine 20 of histone H4. In general, H3K4, H3K36, and H3K79 methylation have been associated with transcriptionally active euchromatin, whereas H3K9, H3K27, and H4K20 methylations are associated with transcriptional inactive heterochromatin [2, 11]. Aberrant methylation of histone lysines has been implicated in various disease etiologies including cancer and X-linked mental retardation [3, 12–15]. Therefore, a proper understanding of the structural and functional regulations of the enzymes responsible for reversible modifications of histone lysines is of immense importance in developing future therapeutics for many of these diseases. Following is a summary of our understanding on the structural properties of known enzymes responsible for catalyzing specific lysine methylation and enzymes responsible for selective removal of these methylation marks.

2 Histone Lysine (K) Methyltransferases (HKMTs)

With the exception of Dot1 [16–18], all known HKMTs contain an evolutionarily conserved SET domain comprised of 130 amino acids [19–23]. The SET domain was first identified as a shared sequence motif in three *Drosophila* proteins, suppressor of variegation [Su(var)3-9], enhancer of zeste [E(z)], and homeobox gene regulator trithorax [Trx] [24]. Mammalian homologues of *Drosophila* Su(var) 3-9 protein, SUV39H1 in human and Suv39h in mouse, were the first characterized HKMTs involved in H3K9 methylation [24]. Since then, more than 50 SET

domain-containing proteins with proven or predicted enzymatic role in carrying out lysine methylation on histone tail have been identified in human [19, 25].

With a few exceptions (e.g., Set8), the majority of the SET-containing HKMTs contain at least one additional protein module in their protein sequence. Based on the sequence homology within and around the catalytic SET domain, as well as based on other protein modules and their architectures, SET-containing HKMTs are grouped into six different subfamilies: SET1, SET2, SUV39, EZH, SMYD, and PRDM [19, 20, 25]. A number of SET-containing HKMTs, however, do not fall into the above six subfamilies, due to lacking sequences (and conservation) flanking their SET domains. Examples of such proteins include Set8/PR_Set7 (mono-methylates H4K20), SUV4-20H1 and SUV4-20H2 (di- and tri-methylates H4K20), Set7/9 (mono-methylates H3K4 and many other nonhistone substrates), as well as MLL5, SetD5 (KIAA1757), and SetD6 (FLJ21148) with currently unknown role in histone lysine methylation.

3 Structures of SET Domains

Structures of many SET domains from different subfamilies have been solved in various combinations with bound substrate peptide and methyl donor (S-adenosyl-l-methionine, AdoMet) or reaction product (S-adenosyl-l-homocysteine, AdoHcy) (Table 1). Representative structures of the SET-domain are displayed in Fig. 1. The SET domain adopts a unique structural design formed by a series of β-strands folded into three sheets surrounding a knot-like structure (Fig. 1). The knot-like structure is formed by the C-terminal segment of the SET domain, which passes through a loop formed by the preceding stretch of sequences. Formation of this

Table 1 List of HKMTs with known structures (PDB ID)

Position	HKMT	PDB ID
H3K4	MLL1	2W5Y, 2W5Z
	SET7/9 (including nonhistone substrates)	3CBO, 3CBM, 3CBP, 2F69, 1XQH, 1O9S, 1N6C, 1N6A, 1H3I, 1MUF, 1MT6
H3K9	SUV39H2	2R3A
	G9a (EHMT2)	2O8J, 3K5K
	GLP (EHMT1)	2RFI, 3FPD, 3HNA, 2IGQ,
	RIZ1 (PRDM2)	2JV0, 2QPW
	DIM-5	1PEG, 1ML9
	Clr4	1MVX, 1MVH
H3K36	SET2	3H6L
H3K79	DOT1L	1NW3, 1U2Z
H4K20	PR-SET7 (SET8)	3F9W, 3F9X, 3F9Y, 3F9Z, 2BQZ, 1ZKK
Others	PRDM10	3IHX
	PRDM12	3EP0
	PRDM1	3DAL
	SETMER	3BO5

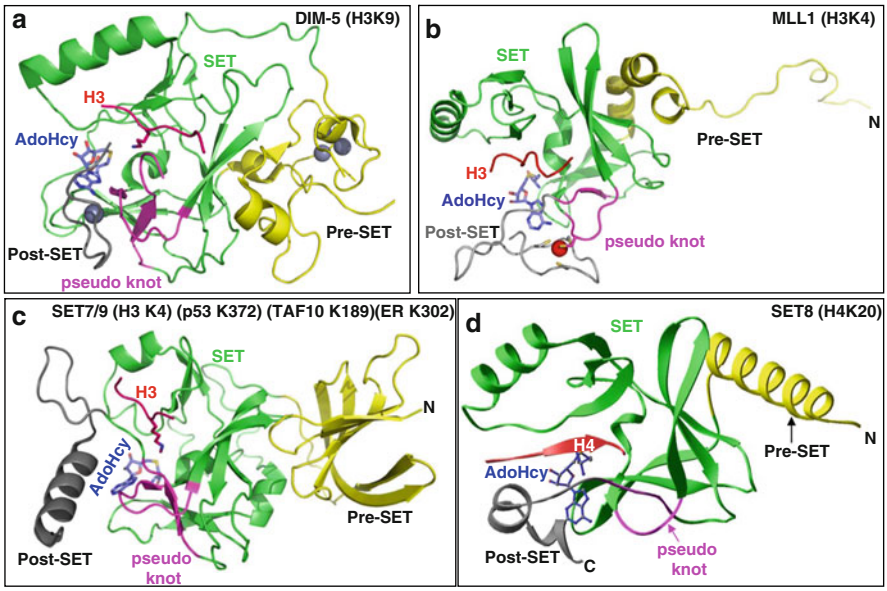


Fig. 1 Examples of SET domain structures. Ribbon diagram of (**a**) *Neurospora* DIM-5 [26], (**b**) human MLL1 [102], (**c**) human SET7/9 [31], and (**d**) human SET8 [103] (or PR-SET7 [104])

knot-like structure brings two conserved sequence motifs of the SET domain, consisting of RFINHxCxPN and ELx(F/Y)DY, in close proximity to the AdoMet-binding region and peptide-binding channel (Fig. 2a–b). Interestingly, biochemical studies performed with F/Y mutants of the conserved ELx(F/Y)DY motif in DIM-5 (F281Y), G9a (F1205Y), Set8 (Y334F), Set7/9 (Y305F), and Set1 (Y1052F) suggest that the F/Y switch regulates the product specificity (mono-, di-, or tri-methylation) of SET-containing HKMTs [26–29].

4 Structural Properties of Pre-SET and Post-SET Modules

Available crystal structures of the SUV39 subfamily (DIM-5, Clr4, GLP/EHMT1, G9a/EHMT2, and SUV39H2 – all H3K9 HKMTs) show the presence of two closely packed cysteine rich-modules in the pre-SET and post-SET (before and after the SET domain) (Fig. 1a). These two modules are important in maintaining structural stability (pre-SET) and forming part of the active site lysine channel (post-SET) [26, 30]. The pre-SET module of SUV39 subfamily contains nine conserved cysteines (Fig. 2c), which coordinate three Zn^{2+} atoms in a triangular geometry (Fig. 2d). The post-SET module of SUV39 as well as Set1 (Fig. 1b) and Set2 subfamilies contains three conserved cysteines, which along with a cysteine from the conserved RFINHxCxPN motif of the SET domain tetrahedrally coordinate one

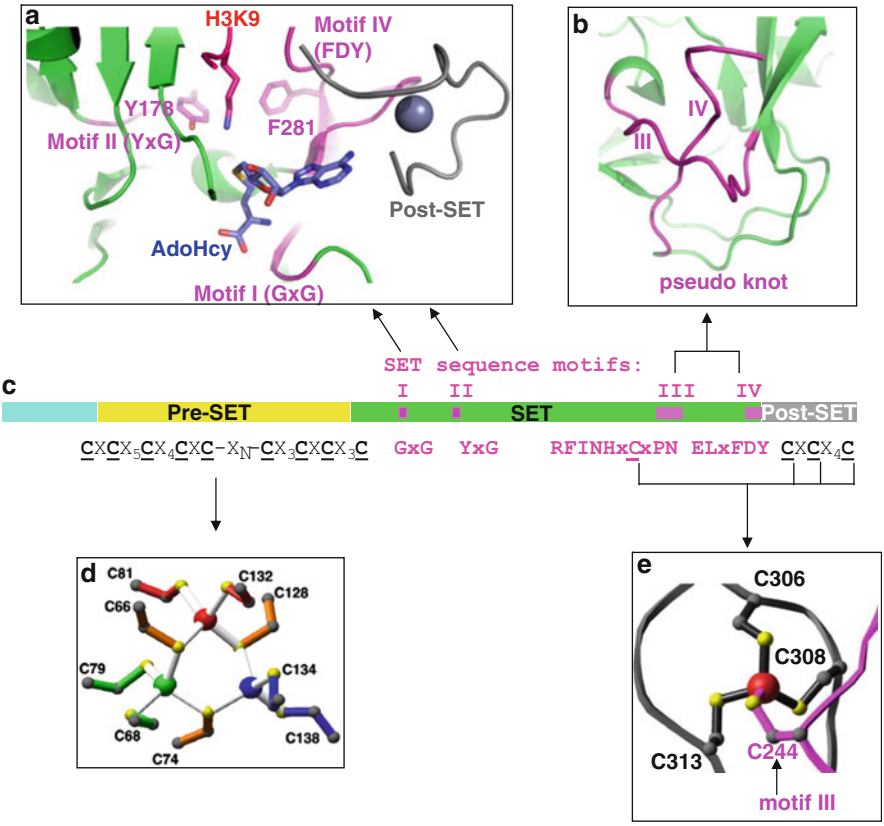


Fig. 2 Structural features of *Neurospora* DIM-5 [26, 30]. (**a**–**b**) Ribbon diagram of the pseudo knot formed by motifs III and IV. (**c**) DIM-5 contains four segments: a weakly conserved amino-terminal region, a pre-SET domain containing nine invariant cysteines, the SET region containing four signature motifs, and the post-SET domain containing three invariant cysteines. (**d**) Illustration of pre-SET Zn_3Cys_9 triangular zinc cluster and (**e**) post-SET zinc center

Zn^{2+} atom near the active site (Fig. 2e). Binding of this Zn^{2+} at the active site is essential for the activity of SUV39 subfamily and therefore is a promising site for drug targeting [26].

The pre-SET and post-SET sequences in Set7/9 do not contain any cysteine-rich region. Instead, pre-SET in Set7/9 is occupied with a β-sheet structure comprised of 12 antiparallel β-strands, while the post-SET is occupied with a small α-helix [31] (Fig. 1c). Packing of the post-SET helix into the catalytic SET domain is important to form the substrate-binding groove in Set7/9. Similar variations in the sequences flanking the SET domain have also been observed in other subfamilies of HKMTs and suggest a convergent evolution of SET-containing HKMTs. This variation may also explain the differences in substrate specificities among the SET-containing HKMTs.

Structural and biochemical studies suggested that consensus substrate recognition sequences for G9a and Set7/9 contain only two to three residues: RK (G9a) and (R/K)(S/T)K (Set7/9) [32, 33]. The short recognition sequences enable these two enzymes to methylate many nonhistone substrates, including Set7/9-mediated methylation of p53 [34], components of the TBP complex, TAF10 [35] and TAF7 [33], estrogen receptor α [36], DNA methyltransferase 1 [37], and G9a-mediated methylation of chromodomain Y-like protein (CDYL1) and widely interspaced zinc finger motifs protein (WIZ) [32], CCAAT/enhancer-binding protein-β (C/EBPβ) [38], as well as G9a auto-methylation [39]. It appears that the dynamic lysine methylation of nonhistone proteins is a rapidly developing new field [40].

5 Structure of Inhibitor Bound G9a and GLP SET Domains

Methylation of H3K9 occurs in heterochromatin, which requires trimethylation of histone H3 at lysine 9 (H3K9me3) by Suv39h [41, 42], and in euchromatin, which requires mono- and di-methylation of H3K9 (H3K9me1/me2) mostly by G9a and GLP [43, 44]. H3K9me1/me2 are the only silencing marks that are lost when tumor suppressor genes, e.g., in colorectal cancer cells [45] and in breast cancer cells [46], are reactivated following treatment with 5-aza-2′-deoxycytidine, a DNA demethylation drug [47]. Thus, the enzymes that produce H3K9me1/me2 are appealing targets for inhibition.

A small molecule, BIX-01294 (a diazepin-quinazolin-amine derivative), was originally identified as a G9a inhibitor during a chemical library screen of small molecules [48]. The compound inhibits G9a and GLP activities (IC_{50} in low μM range) [48, 49] and reduces the methylation levels of H3K9 at several G9a target genes [48, 50]. BIX-01294 was used in combination with genetic factors to improve the efficiency of generation of induced pluripotent stem cells [51–53]. This is consistent with the observation that repressive H3K9 methylation by G9a is associated with the inactivation of Oct3/4, one of the four Yamanaka genetic factors required for included pluripotency [54], during differentiation [55].

BIX-01294 was crystallized with the catalytic SET domain of GLP in the presence of AdoHcy [49]. The inhibitor is bound in the acidic substrate peptide groove at the location where the histone H3 residues N-terminal to the target lysine lie. The inhibitor resembles the bound conformation of histone H3K4 to H3R8 and is positioned by residues specific for G9a and GLP through specific interactions. Most importantly, the inhibitor-bound SET domain structure provides avenues for improving the potency of the inhibitor. One of suggested improvements is by extending the branch of O7-methoxy-CH_3 into the target lysine-binding channel, which should provide additional binding energy by increasing the surface area of binding [49]. Indeed, a recent report of chemical exploration of BIX-01294 identified a derivative (UNC0224) as a potent and selective G9a inhibitor [56]. UNC0224 contains an extended *N*-dimethylamino-propoxy arm occupying the target lysine-binding channel.

6 Histone Lysine Specific Demethylase (LSD1)

The discovery of lysine specific demethylase 1 (LSD1) [57] established that protein lysine methylation is a reversible posttranslational modification. LSD1 is a flavin-dependent amine oxidase, which demethylates H3K4me2/me1 [57], H3K9me2/me1 (in an androgen receptor-mediated pathway) [58], and p53 [59]. The closely related LSD2 demethylates H3K4me2/me1 [60] and has been linked with imprinting of the maternal genome [61]. Both LSD1 and LSD2 demethylate methyl-lysine by forming of an imine intermediate, which undergoes hydrolysis in aqueous buffer (Fig. 3a) to complete the demethylation process. Mechanistic requirements for a protonated amine in this demethylation pathway do not permit either LSD1 or LSD2 to demethylate trimethylated lysines [62].

LSD1 is found in histone modification complexes that control cell-specific gene expression [57]. Within these complexes, REST (RE1-silencing transcription factor) corepressor CoREST enables LSD1 to demethylate nucleosomes [63, 64],

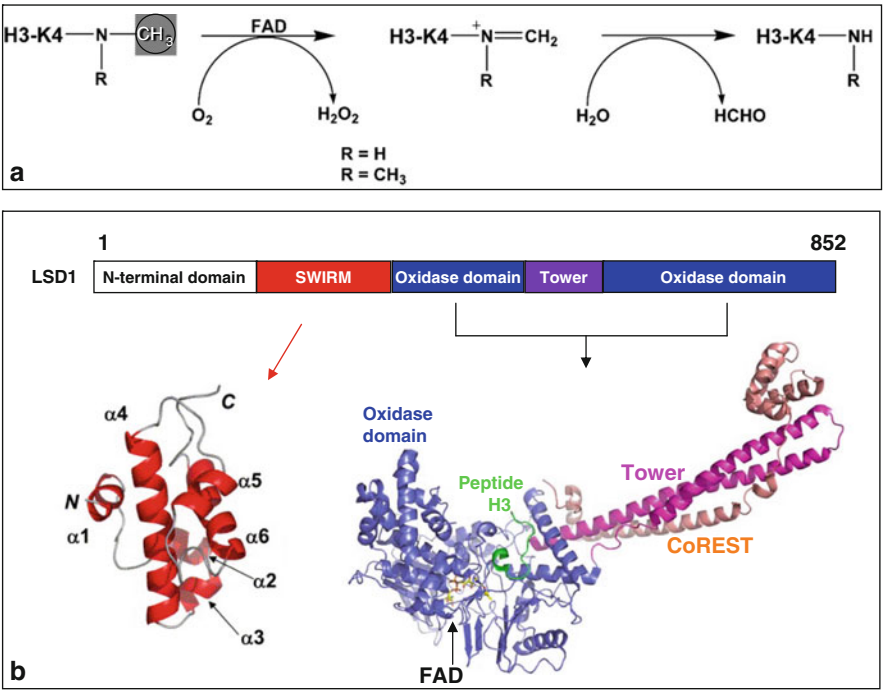


Fig. 3 Demethylation by oxidation. (**a**) Scheme of the demethylation reaction catalyzed by LSD1. (**b**) Schematic representation of human LSD1 domain organization. The oxidase domain contains an atypical insertion of the Tower domain not found in other oxidases. The solution NMR structure of the SIWRM domain of LSD1 is shown in *red* [65]. Crystal structure of LSD1 (residues 171–836 in *blue*)-CoREST (residues 308–440 in *red*) in complex with H3 peptide (residues 1–16 in *green*), and the FAD cofactor is shown as a *yellow ball-and-stick* [71]

while BHC80 (BRAF–HDAC complex) inhibits LSD1 activity [63]. The LSD1 polypeptide chain can be divided into several structural/functional regions (Fig. 3b): the N-terminal putative nuclear localization signal, followed by a SWIRM (Swi3p, Rsc8p, and Moira) domain [65] – found in several nucleosome-interacting proteins – and a monoamine oxidase domain – capable of demethylating lysines in a flavin-dependent manner [66]. Thus far, crystal structures of LSD1 alone [67, 68], LSD1 in complex with CoREST [69], and LSD1-CoREST in complex with H3 peptide [70, 71] have been determined. Using a 21-residue peptide bearing a methionine in place of target methyl-K4 – a 30-fold increase in binding affinity making the mutant peptide a strong inhibitor and an ideal candidate for structural work – Forneris et al. (2007) were able to resolve the first 16 residues of the H3 peptide, in perfect agreement with their previous biochemical data that LSD1 is active on peptide substrates longer than 16 amino acids [66]. This study is the first in which a long, structured histone tail has been visualized in histone-modifying enzymes and protein domains that recognize (decode) methyl-lysine signals. In comparison, a similar study of LSD1-histone peptide, using the approach of covalent tethering of peptide substrate to cofactor FAD, observed the first 7 residues (out of 21 residues used) of H3 peptide [70].

7 Jumonji-Containing Lysine Demethylases

In search of enzymes capable of reversing methylated lysines, Trwick et al. [72] hypothesized that Jumonji domain containing Fe^{2+}- and α-ketoglutarate-dependent dioxygenases can reverse lysine methylation via a similar mechanism as followed by bacterial AlkB family of DNA repair enzymes (Fig. 4a). This hypothesis was quickly verified with the discovery of JHDM1 as the Jumonji domain-containing histone demethylase 1 [73]. Jumonji-containing proteins are members of the cupin superfamily with functional roles in various biological processes including DNA/RNA repair through the demethylation of N-methylated nucleic acids (e.g., 3-methylcytosine, 1-methyladenine) [74, 75], hydroxylation of protein and lipid side chains [76], protein lysyl-5-hydroxylation [77], as well as recently characterized role in oxidizing 5-methylcytosine to 5-hydroxymethylcytosine [78]. Demethylation reactions catalyzed by Jumonji enzymes follow a hydroxylation pathway, which can demethylate mono-, di-, or tri-methylated lysines (Fig. 4a) [79, 80].

Currently, there are nearly 30 Jumonji-containing proteins identified in human proteome, 20 of which have known function in histone demethylation [2]. The majority of Jumonji-containing demethylases contains at least one additional structural domain in their sequence. Based on the phylogenetic relationships and domain architectures, these proteins are divided into seven subfamilies [2]. Additional structural motifs (other than the Jumonji domain) present in these proteins are thought to be important in substrate recognition or facilitating protein–protein interactions. For example, the H3K4 demethylase RBP2 contains a DNA-binding domain, the AT-rich interaction domain (ARID). ARID binds DNA sequence motif

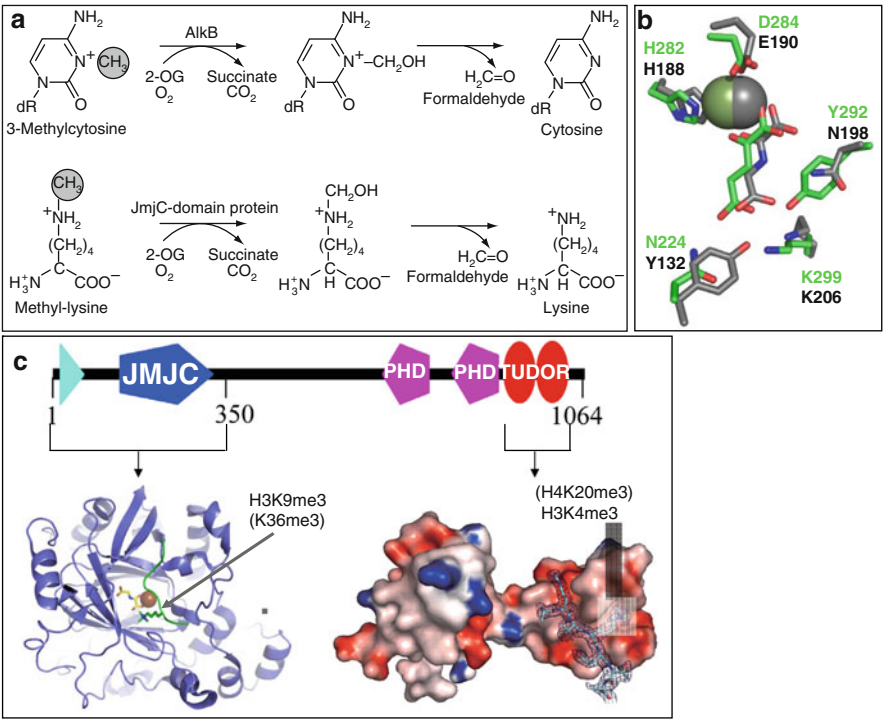


Fig. 4 Demethylation by hydroxylation. (**a**) Mechanisms of demethylation of 3-methylcytosine by AlkB (*top*) and of methyl-lysine by Jumonji-domain proteins (*bottom*). (**b**) Coordinations of Fe^{2+} (*sphere*), α-ketoglutarate in JMJD2A (in *gray*), and KIAA1718 (in *green*). (**c**) Schematic representation of JMJD2A domain organization, including the structures of the N-terminal Jumonji (*ribbons*) [85] and the C-terminal double Tudor domain (surface representation) [87]

Table 2 List of histone lysine demethylases with known structures (PDB ID)

Position	HDM	PDB ID
H3K4	LSD1	2IW5, 2HKO, 2V1D, 2UXN, 2UXX, 2DW4, 2Z3Y, 2EJR, 2Z5U
H3K9	PHF8	3K3O, 3K3N
	JMJD2A	2VD7, 2Q8C, 2Q8D, 2Q8E, 2P5B, 2PXJ, 2OQ6, 2OQ7, 2OS2, 2OT7, 2OXO, 2GP3, 2GP5
	JMJD2D	3DXT
H3K36	JHDM1A (FBXL11)	2YU1, 2YU2

(CCGCCC) and is required for RBP2 demethylase activity in cells and that DNA recognition is essential to regulate transcription [81].

Thirteen crystal structures for the Jumonji domain of JMJD2A in various configurations are currently available (Table 2) [67, 82–85]. In addition, one structure is available for JMJD2D, two for JHDM1A [86], and two for PHF8 Jumonji domain (Table 2). Like in other cupin family members, the Jumonji domain adopts the

conserved double-stranded-β-helix or jelly-roll structure formed by eight antiparallel β-strands, which harbors the Fe^{2+} (coordinated by two histones and one aspartate or glutamate) and α-ketoglutarate in a conserved coordination environment (Fig. 4b). The co-substrate α-ketoglutarate is coordinated to the Fe^{2+} center through C1-carboxylate and C2-keto group. The C5-carboxylate of α-ketoglutarate forms hydrogen-bonding interactions with Jumonji domain.

8 JMJD2A

JMJD2A contains an N-terminal Jumonji domain and C-terminal PHD and Tudor domains (Fig. 4c). The JMJD2A Jumonji domain alone is capable of demethylating tri- and di-methylated H3K9 (H3K9me3/2) and H3K36 (H3K36me3/2), though with a very low turnover rate [84]. Structural studies revealed that the JMJD2A Jumonji domain predominantly recognizes the backbone of the histone peptides (unusual for a sequence-specific enzyme), allowing the enzyme to demethylate both H3K9me3/2 and H3K36me3/2 [83–85]. On the other hand, JMJD2A Tudor domain binds two different histone sequences (H3K4me3 and H4K20me3) via radically different approaches [87, 88]. The functional connection between the methyl mark reader and eraser in JMJD2A is not clear.

9 PHF8 and KIAA1718

PHF8 and KIAA1718 belong to a small family of Jumonji proteins with three members in mice and human (PHF2, PHF8, and KIAA1718) [2]. These proteins harbor two domains in the N-terminal half (Fig. 5a): a PHD domain that binds H3K4me3 and a Jumonji domain that demethylates H3K9me2, H3K27me2, as well as H3K36me2 [89]. However, the presence of H3K4me3 on the same peptide as H3K9me2 makes the doubly methylated peptide a significantly better substrate of PHF8 [90]. In contrast, the presence of H3K4me3 has the opposite effect in that it diminishes the H3K9me2 demethylase activity of KIAA1718 with no adverse effect on its H3K27me2 activity. Differences in substrate specificity between the two enzymes are explained by a bent conformation of PHF8, allowing each of its domains to engage their respective targets, and an extended conformation of KIAA1718, which prevents its access to H3K9me2 by its Jumonji domain when its PHD domain engages H3K4me3 (Fig. 5a). This study concludes that the structural linkage between the PHD domain binding to H3K4me3 and the placement of the catalytic Jumonji domains relative to this "on" epigenetic mark determines which repressive marks are removed in both demethylases. Taken together, we suggest that the PHF8 and KIAA1718 Jumonji domains on their own are promiscuous enzymes; it is the associated PHD domains and linker – a determinant for the relative positioning of the two domains – that are mainly responsible for substrate specificity.

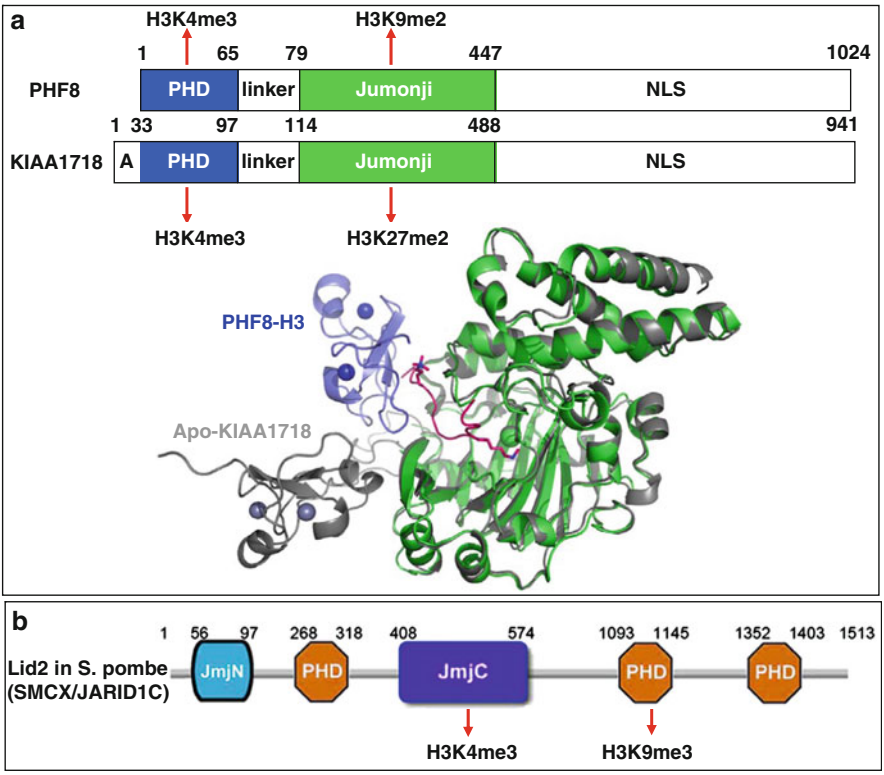


Fig. 5 Crosstalk between Jumonji and PHD within the same polypeptide. (**a**) Schematic representations of PHF8 and KIAA1718. Superimposition of PHF8 (*colored*) and KIAA1718 (*gray*) in their respective Jumonji domains indicates that the PHF8 PHD domain adopts a bent conformation toward the Jumonji domain in the presence of H3 substrate binding, whereas the PHD and Jumonji domains of KIAA1718 adopt an extended conformation in its apo-structure [90]. (**b**) Schematic representation of Lid2 in *S. pombe* [93] (SMCX/JARID1C [92])

Using domain cooperativity to enhance an enzyme's activity and its substrate specificity may be a general mechanism for Jumonji-containing protein lysine demethylases. For example, JHDM2A-mediated histone H3K9me1/2 demethylation requires a zinc finger N-terminal to the Jumonji domain for its enzymatic activity [91]. JARID Jumonji family proteins (including Lid2 in *S. pombe*) contain a Jumonji domain that demethylates H3K4me3 surrounded by several PHD domains and at least one of them binds H3K9me3 [92, 93] (Fig. 5b). Mutation or deletion of this PHD domain impairs the demethylase activity on H3K4me3 [92, 93]. We speculate that the ideal substrate for JARID family is H3 trimethylated at both K4 and K9, allowing the enzyme to remove any local activating methyl groups of H3K4me3 by the Jumonji in a repressing environment with H3K9me3 bound by the PHD (Fig. 5b). We further speculate that a similar situation might occur for JMJD2A where each of the two demethylase activities (H3K9me3/2 and H3K36me3/2) correlates with one of the methyl marks (H3K4me3 and H4K20me3) recognized by the Tudor domain (Fig. 4c).

10 Perspective

The histone code hypothesis suggests that multiple covalent histone modifications can be read combinatorially through effectors that are recruited to these marks and subsequently act on the local chromatin structure or transcriptional machinery via crosstalk among histone modifications [94–97]. Several histone-methylating enzymes contain components (domains) that both synthesize and bind a specific histone mark, such as mammalian G9a/GLP (for H3K9me1/me2) [9] and *S. pombe* Clr4 (for H3K9me3) [10]. They contain modules, within the same polypeptide, for both making (via the SET domain) and recognizing (via the ankyrin repeats or chromodomain) a given methyl mark – allowing for a mechanism of crosstalk to propagate a given methyl mark. PHF8 and KIAA1718 (Fig. 5a), and perhaps JARID/Lid2 (Fig. 5b) and JMJD2A (Fig. 4c), contain modules, within the same polypeptide, for both recognizing (via the PHD or Tudor) and removing (via the Jumonji domain) two opposing methyl marks – a mechanism of crosstalk removes an "off" methyl mark based on an existing "on" methyl mark. Understanding the function and crosstalk of individual letters (one methyl mark, two methyl marks, and so on) may allow us eventually decipher the complex language of the histone code [94, 98].

The availability of human and other model research organism genome sequences, proteomics, and transcriptomics has provided answers to a wide range of questions that in some cases we did not even previously know to ask. Global analyses of genomic DNA methylation and histone modifications [99–101] are playing a similar role, yielding powerful insights into normal development and diseases, such as cancer and diabetes. The experimental characterization of individual modifying enzymes (writers) and demodifying enzymes (erasers) of the histone code is providing a growing and convergent picture of the kinetic mechanisms, binding partners, chromatin recognition, and in some cases structures of these proteins. However, it is clear that the activities of writers, erasers, and readers of the histone code are regulated in multicomponent complexes that have yet to be fully defined and characterized.

Acknowledgments The work in the Cheng laboratory was supported by grants GM06860 and DK082678 from the National Institutes of Health (NIH). X.C. is a Georgia Research Alliance Eminent Scholar.

References

1. Martens JH, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P, Jenuwein T (2005) The profile of repeat-associated histone lysine methylation states in the mouse epigenome. Embo J 24:800–812
2. Klose RJ, Kallin EM, Zhang Y (2006) JmjC-domain-containing proteins and histone demethylation. Nat Rev Genet 7:715–727

3. Shi Y (2007) Histone lysine demethylases: emerging roles in development, physiology and disease. Nat Rev Genet 8:829–833
4. Li H, Ilin S, Wang W, Duncan EM, Wysocka J, Allis CD, Patel DJ (2006) Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature 442:91–95
5. Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, Wu C, Allis CD (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature 442:86–90
6. Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410:120–124
7. Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410:116–120
8. Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ, Voigt P, Martin SR, Taylor WR, De Marco V, Pirrotta V, Reinberg D, Gamblin SJ (2009) Role of the polycomb protein EED in the propagation of repressive histone marks. Nature 461(7265):762–767
9. Collins RE, Northrop JP, Horton JR, Lee DY, Zhang X, Stallcup MR, Cheng X (2008) The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. Nat Struct Mol Biol 15:245–250
10. Zhang K, Mosch K, Fischle W, Grewal SI (2008) Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nat Struct Mol Biol 15:381–388
11. Martin C, Zhang Y (2005) The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol 6:838–849
12. Ng SS, Yue WW, Oppermann U, Klose RJ (2009) Dynamic protein methylation in chromatin biology. Cell Mol Life Sci 66:407–422
13. Schneider R, Bannister AJ, Kouzarides T (2002) Unsafe SETs: histone lysine methyltransferases and cancer. Trends Biochem Sci 27:396–402
14. Spannhoff A, Hauser AT, Heinke R, Sippl W, Jung M (2009) The emerging therapeutic potential of histone methyltransferase and demethylase inhibitors. ChemMedChem 4:1568–1582
15. Spannhoff A, Sippl W, Jung M (2009) Cancer treatment of the future: inhibitors of histone methyltransferases. Int J Biochem Cell Biol 41:4–11
16. van Leeuwen F, Gafken PR, Gottschling DE (2002) Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell 109:745–756
17. Min J, Feng Q, Li Z, Zhang Y, Xu RM (2003) Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. Cell 112:711–723
18. Sawada K, Yang Z, Horton JR, Collins RE, Zhang X, Cheng X (2004) Structure of the conserved core of the yeast Dot1p, a nucleosomal histone H3 lysine 79 methyltransferase. J Biol Chem 279:43296–43306
19. Cheng X, Collins RE, Zhang X (2005) Structural and sequence motifs of protein (histone) methylation enzymes. Annu Rev Biophys Biomol Struct 34:267–294
20. Dillon SC, Zhang X, Trievel RC, Cheng X (2005) The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol 6:227
21. Qian C, Zhou MM (2006) SET domain protein lysine methyltransferases: structure, specificity and catalysis. Cell Mol Life Sci 63:2755–2763
22. Couture JF, Trievel RC (2006) Histone-modifying enzymes: encrypting an enigmatic epigenetic code. Curr Opin Struct Biol 16:753–760
23. Cheng X, Zhang X (2007) Structural dynamics of protein lysine methylation and demethylation. Mutat Res/Fundam Mol Mech Mutagen 618:102–115
24. Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, Jenuwein T (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406:593–599

25. Volkel P, Angrand PO (2007) The control of histone lysine methylation in epigenetic regulation. Biochimie 89:1–20
26. Zhang X, Yang Z, Khan SI, Horton JR, Tamaru H, Selker EU, Cheng X (2003) Structural basis for the product specificity of histone lysine methyltransferases. Mol Cell 12:177–185
27. Collins RE, Tachibana M, Tamaru H, Smith KM, Jia D, Zhang X, Selker EU, Shinkai Y, Cheng X (2005) In vitro and in vivo analyses of a Phe/Tyr switch controlling product specificity of histone lysine methyltransferases. J Biol Chem 280:5563–5570
28. Couture JF, Dirk LM, Brunzelle JS, Houtz RL, Trievel RC (2008) Structural origins for the product specificity of SET domain protein methyltransferases. Proc Natl Acad Sci USA 105:20659–20664
29. Takahashi YH, Lee JS, Swanson SK, Saraf A, Florens L, Washburn MP, Trievel RC, Shilatifard A (2009) Regulation of H3K4 trimethylation via Cps40 (Spp 1) of COMPASS is monoubiquitination independent: implication for a Phe/Tyr switch by the catalytic domain of Set1. Mol Cell Biol 29:3478–3486
30. Zhang X, Tamaru H, Khan SI, Horton JR, Keefe LJ, Selker EU, Cheng X (2002) Structure of the Neurospora SET domain protein DIM-5, a histone H3 lysine methyltransferase. Cell 111:117–127
31. Xiao B, Jing C, Wilson JR, Walker PA, Vasisht N, Kelly G, Howell S, Taylor IA, Blackburn GM, Gamblin SJ (2003) Structure and catalytic mechanism of the human histone methyltransferase SET7/9. Nature 421:652–656
32. Rathert P, Dhayalan A, Murakami M, Zhang X, Tamas R, Jurkowska R, Komatsu Y, Shinkai Y, Cheng X, Jeltsch A (2008) Protein lysine methyltransferase G9a acts on non-histone targets. Nat Chem Biol 4:344–346
33. Couture JF, Collazo E, Hauk G, Trievel RC (2006) Structural basis for the methylation site specificity of SET7/9. Nat Struct Mol Biol 13:140–146
34. Chuikov S, Kurash JK, Wilson JR, Xiao B, Justin N, Ivanov GS, McKinney K, Tempst P, Prives C, Gamblin SJ, Barlev NA, Reinberg D (2004) Regulation of p53 activity through lysine methylation. Nature 432:353–360
35. Kouskouti A, Scheer E, Staub A, Tora L, Talianidis I (2004) Gene-specific modulation of TAF10 function by SET9-mediated methylation. Mol Cell 14:175–182
36. Subramanian K, Jia D, Kapoor-Vazirani P, Powell DR, Collins RE, Sharma D, Peng J, Cheng X, Vertino PM (2008) Regulation of estrogen receptor alpha by the SET7 lysine methyltransferase. Mol Cell 30:336–347
37. Esteve PO, Chin HG, Benner J, Feehery GR, Samaranayake M, Horwitz GA, Jacobsen SE, Pradhan S (2009) Regulation of DNMT1 stability through SET7-mediated lysine methylation in mammalian cells. Proc Natl Acad Sci USA 106:5076–5081
38. Pless O, Kowenz-Leutz E, Knoblich M, Lausen J, Beyermann M, Walsh MJ, Leutz A (2008) G9a-mediated lysine methylation alters the function of CCAAT/enhancer-binding protein-beta. J Biol Chem 283:26357–26363
39. Sampath SC, Marazzi I, Yap KL, Sampath SC, Krutchinsky AN, Mecklenbrauker I, Viale A, Rudensky E, Zhou MM, Chait BT, Tarakhovsky A (2007) Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. Mol Cell 27:596–608
40. Huang J, Berger SL (2008) The emerging field of dynamic lysine methylation of non-histone proteins. Curr Opin Genet Dev 18:152–158
41. Jenuwein T (2006) The epigenetic magic of histone lysine methylation. Febs J 273:3121–3135
42. Grewal SI, Jia S (2007) Heterochromatin revisited. Nat Rev Genet 8:35–46
43. Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H, Shinkai Y (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev 16:1779–1791

44. Tachibana M, Ueda J, Fukuda M, Takeda N, Ohta T, Iwanari H, Sakihama T, Kodama T, Hamakubo T, Shinkai Y (2005) Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. Genes Dev 19:815–826
45. McGarvey KM, Fahrner JA, Greene E, Martens J, Jenuwein T, Baylin SB (2006) Silenced tumor suppressor genes reactivated by DNA demethylation do not return to a fully euchromatic chromatin state. Cancer Res 66:3541–3549
46. Wozniak RJ, Klimecki WT, Lau SS, Feinstein Y, Futscher BW (2007) 5-Aza-2′-deoxycytidine-mediated reductions in G9A histone methyltransferase and histone H3 K9 dimethylation levels are linked to tumor suppressor gene reactivation. Oncogene 26:77–90
47. Yoo CB, Jeong S, Egger G, Liang G, Phiasivongsa P, Tang C, Redkar S, Jones PA (2007) Delivery of 5-aza-2′-deoxycytidine to cells using oligodeoxynucleotides. Cancer Res 67:6400–6408
48. Kubicek S, O'Sullivan RJ, August EM, Hickey ER, Zhang Q, Teodoro ML, Rea S, Mechtler K, Kowalski JA, Homon CA, Kelly TA, Jenuwein T (2007) Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. Mol Cell 25:473–481
49. Chang Y, Zhang X, Horton JR, Upadhyay AK, Spannhoff A, Liu J, Snyder JP, Bedford MT, Cheng X (2009) Structural basis for G9a-like protein lysine methyltransferase inhibition by BIX-01294. Nat Struct Mol Biol 16:312–317
50. Trojer P, Zhang J, Yonezawa M, Schmidt A, Zheng H, Jenuwein T, Reinberg D (2009) Dynamic histone H1 isotype 4 methylation and demethylation by histone lysine methyltransferase G9a/KMT1C and the jumonji domain-containing JMJD2/KDM4 proteins. J Biol Chem 284:8395–8405
51. Shi Y, Do JT, Desponts C, Hahm HS, Scholer HR, Ding S (2008) A combined chemical and genetic approach for the generation of induced pluripotent stem cells. Cell Stem Cell 2:525–528
52. Xu Y, Shi Y, Ding S (2008) A chemical approach to stem-cell biology and regenerative medicine. Nature 453:338–344
53. Feng B, Ng JH, Heng JC, Ng HH (2009) Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells. Cell Stem Cell 4:301–312
54. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126:663–676
55. Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y, Cedar H, Bergman Y (2006) G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat Cell Biol 8:188–194
56. Liu F, Chen X, Allali-Hassani A, Quinn AM, Wasney GA, Dong A, Barsyte D, Kozieradzki I, Senisterra G, Chau I, Siarheyeva A, Kireev DB, Jadhav A, Herold JM, Frye SV, Arrowsmith CH, Brown PJ, Simeonov A, Vedadi M, Jin J (2009) Discovery of a 2,4-diamino-7-aminoalkoxyquinazoline as a potent and selective inhibitor of histone lysine methyltransferase G9a. J Med Chem 52(24):7950–7953
57. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 119:941–953
58. Metzger E, Wissmann M, Yin N, Muller JM, Schneider R, Peters AH, Gunther T, Buettner R, Schule R (2005) LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature 437:436–439
59. Huang J, Sengupta R, Espejo AB, Lee MG, Dorsey JA, Richter M, Opravil S, Shiekhattar R, Bedford MT, Jenuwein T, Berger SL (2007) p53 is regulated by the lysine demethylase LSD1. Nature 449:105–108
60. Karytinos A, Forneris F, Profumo A, Ciossani G, Battaglioli E, Binda C, Mattevi A (2009) A novel mammalian flavin-dependent histone demethylase. J Biol Chem 284:17775–17782

61. Ciccone DN, Su H, Hevi S, Gay F, Lei H, Bajko J, Xu G, Li E, Chen T (2009) KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. Nature 461 (7262):415–418
62. Shi Y, Whetstine JR (2007) Dynamic regulation of histone lysine methylation by demethylases. Mol Cell 25:1–14
63. Shi YJ, Matson C, Lan F, Iwase S, Baba T, Shi Y (2005) Regulation of LSD1 histone demethylase activity by its associated factors. Mol Cell 19:857–864
64. Lee MG, Wynder C, Cooch N, Shiekhattar R (2005) An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature 437:432–435
65. Tochio N, Umehara T, Koshiba S, Inoue M, Yabuki T, Aoki M, Seki E, Watanabe S, Tomo Y, Hanada M, Ikari M, Sato M, Terada T, Nagase T, Ohara O, Shirouzu M, Tanaka A, Kigawa T, Yokoyama S (2006) Solution structure of the SWIRM domain of human histone demethylase LSD1. Structure 14:457–468
66. Forneris F, Binda C, Vanoni MA, Battaglioli E, Mattevi A (2005) Human histone demethylase LSD1 reads the histone code. J Biol Chem 280:41360–41365
67. Chen Y, Yang Y, Wang F, Wan K, Yamane K, Zhang Y, Lei M (2006) Crystal structure of human histone lysine-specific demethylase 1 (LSD1). Proc Natl Acad Sci USA 103:13956–13961
68. Stavropoulos P, Blobel G, Hoelz A (2006) Crystal structure and mechanism of human lysine-specific demethylase-1. Nat Struct Mol Biol 13:626–632
69. Yang M, Gocke CB, Luo X, Borek D, Tomchick DR, Machius M, Otwinowski Z, Yu H (2006) Structural basis for CoREST-dependent demethylation of nucleosomes by the human LSD1 histone demethylase. Mol Cell 23:377–387
70. Yang M, Culhane JC, Szewczuk LM, Gocke CB, Brautigam CA, Tomchick DR, Machius M, Cole PA, Yu H (2007) Structural basis of histone demethylation by LSD1 revealed by suicide inactivation. Nat Struct Mol Biol 14:535–539
71. Forneris F, Binda C, Adamo A, Battaglioli E, Mattevi A (2007) Structural basis of LSD1-CoREST selectivity in histone H3 recognition. J Biol Chem 282:20070–20074
72. Trewick SC, McLaughlin PJ, Allshire RC (2005) Methylation: lost in hydroxylation? EMBO Rep 6:315–320
73. Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, Zhang Y (2006) Histone demethylation by a family of JmjC domain-containing proteins. Nature 439:811–816
74. Falnes PO, Johansen RF, Seeberg E (2002) AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. Nature 419:178–182
75. Trewick SC, Henshaw TF, Hausinger RP, Lindahl T, Sedgwick B (2002) Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature 419:174–178
76. Dann CE 3rd, Bruick RK (2005) Dioxygenases as O2-dependent regulators of the hypoxic response pathway. Biochem Biophys Res Commun 338:639–647
77. Webby CJ, Wolf A, Gromak N, Dreger M, Kramer H, Kessler B, Nielsen ML, Schmitz C, Butler DS, Yates JR 3rd, Delahunty CM, Hahn P, Lengeling A, Mann M, Proudfoot NJ, Schofield CJ, Bottger A (2009) Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. Science 325:90–93
78. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by the MLL fusion partner TET1. Science 324(5929):930–935
79. Hoffart LM, Barr EW, Guyer RB, Bollinger JM Jr, Krebs C (2006) Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase. Proc Natl Acad Sci USA 103:14738–14743
80. Ozer A, Bruick RK (2007) Non-heme dioxygenases: cellular sensors and regulators jelly rolled into one? Nat Chem Biol 3:144–153
81. Tu S, Teng YC, Yuan C, Wu YT, Chan MY, Cheng AN, Lin PH, Juan LJ, Tsai MD (2008) The ARID domain of the H3K4 demethylase RBP2 binds to a DNA CCGCCC motif. Nat Struct Mol Biol 15:419–421

82. Chen Z, Zang J, Whetstine J, Hong X, Davrazou F, Kutateladze TG, Simpson M, Mao Q, Pan CH, Dai S, Hagman J, Hansen K, Shi Y, Zhang G (2006) Structural insights into histone demethylation by JMJD2 family members. Cell 125:691–702
83. Chen Z, Zang J, Kappler J, Hong X, Crawford F, Wang Q, Lan F, Jiang C, Whetstine J, Dai S, Hansen K, Shi Y, Zhang G (2007) Structural basis of the recognition of a methylated histone tail by JMJD2A. Proc Natl Acad Sci USA 104:10818–10823
84. Couture JF, Collazo E, Ortiz-Tello PA, Brunzelle JS, Trievel RC (2007) Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. Nat Struct Mol Biol 14:689–695
85. Ng SS, Kavanagh KL, McDonough MA, Butler D, Pilka ES, Lienard BM, Bray JE, Savitsky P, Gileadi O, von Delft F, Rose NR, Offer J, Scheinost JC, Borowski T, Sundstrom M, Schofield CJ, Oppermann U (2007) Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. Nature 448:87–91
86. Han Z, Liu P, Gu L, Zhang Y, Li H, Chen S, Chai J (2007) Structural basis for histone demethylation by JHDM1. Frontier Sci 1:52–67
87. Huang Y, Fang J, Bedford MT, Zhang Y, Xu RM (2006) Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. Science 312:748–751
88. Lee J, Thompson JR, Botuyan MV, Mer G (2008) Distinct binding modes specify the recognition of methylated histones H3K4 and H4K20 by JMJD2A-tudor. Nat Struct Mol Biol 15:109–111
89. Loenarz C, Ge W, Coleman ML, Rose NR, Cooper CD, Klose RJ, Ratcliffe PJ, Schofield CJ (2010) PHF8, a gene associated with cleft lip/palate and mental retardation, encodes for an N {varepsilon}-dimethyl lysine demethylase. Hum Mol Genet 19:217–222
90. Horton JR, Upadhyay AK, Qi HH, Zhang X, Shi Y, Cheng X (2010) Enzymatic and structural insights for substrate specificity of a family of Jumonji histone lysine demethylases. Nat Struct Mol Biol 17:38–43
91. Yamane K, Toumazou C, Tsukada Y, Erdjument-Bromage H, Tempst P, Wong J, Zhang Y (2006) JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. Cell 125:483–495
92. Iwase S, Lan F, Bayliss P, de la Torre-Ubieta L, Huarte M, Qi HH, Whetstine JR, Bonni A, Roberts TM, Shi Y (2007) The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. Cell 128:1077–1088
93. Li F, Huarte M, Zaratiegui M, Vaughn MW, Shi Y, Martienssen R, Cande WZ (2008) Lid2 is required for coordinating H3K4 and H3K9 methylation of heterochromatin and euchromatin. Cell 135:272–283
94. Strahl BD, Allis CD (2000) The language of covalent histone modifications. Nature 403:41–45
95. Jenuwein T, Allis CD (2001) Translating the histone code. Science 293:1074–1080
96. Turner BM (2007) Defining an epigenetic code. Nat Cell Biol 9:2–6
97. Suganuma T, Workman JL (2008) Crosstalk among histone modifications. Cell 135:604–607
98. Berger SL (2007) The complex language of chromatin regulation during transcription. Nature 447:407–412
99. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454:766–770
100. Zhang Y, Rohde C, Tierling S, Jurkowski TP, Bock C, Santacruz D, Ragozin S, Reinhardt R, Groth M, Walter J, Jeltsch A (2009) DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. PLoS Genet 5:e1000438
101. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462(7271):315–322

102. Southall SM, Wong PS, Odho Z, Roe SM, Wilson JR (2009) Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. Mol Cell 33:181–191
103. Couture JF, Collazo E, Brunzelle JS, Trievel RC (2005) Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. Genes Dev 19:1455–1465
104. Xiao B, Jing C, Kelly G, Walker PA, Muskett FW, Frenkiel TA, Martin SR, Sarma K, Reinberg D, Gamblin SJ, Wilson JR (2005) Specificity and mechanism of the histone methyltransferase Pr-Set7. Genes Dev 19:1444–1454

Epigenetic Mechanisms of Mental Retardation

Anne Schaefer, Alexander Tarakhovsky and Paul Greengard

Abstract Mental retardation is a common form of cognitive impairment affecting ~3% of the population in industrialized countries. The mental retardation syndrome incorporates a highly diverse group of mental disorders characterized by the combination of cognitive impairment and defective adaptive behavior. The genetic basis of the disease is strongly supported by identification of the genetic lesions associated with impaired cognition, learning, and social adaptation in many mental retardation syndromes. Several of the impaired genes encode epigenetic regulators of gene expression. These regulators exert their function through genome-wide posttranslational modification of histones or by mediating and/or recognizing DNA methylation. In this chapter, we review the most recent advances in the field of epigenetic mechanisms of mental retardation. In particular, we focus on animal models of the human diseases and the mechanism of transcriptional deregulation associated with changes in the cell epigenome.

1 Introduction

Recent efforts to elucidate the cellular and molecular bases of brain function using the brain–computer interface (BCI) approach showed that even the simplest brain circuit generates several gigabytes of information per second [1–4]. While this information is encrypted in electric impulses produced by neurons, the foundation for neuronal activity lies in the coordinated function of neuron-expressed gene

A. Schaefer and P. Greengard
Laboratory of Molecular and Cellular Neuroscience, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA
e-mail: aschaefer@mail.rockefeller.edu

A. Tarakhovsky
Laboratory of Lymphocyte Signaling, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_7,

networks. Similar to other cells in the organism, differentiated neurons maintain stable patterns of gene expression [5, 6]. Such patterns define the neuron's identity as well as its functional specificity. The activity of numerous neuronal genes, on the other hand, fluctuates significantly in response to environmental signals [7–15]. In order to respond adequately to the rapidly changing environment, it is essential for differentiated neurons to maintain a certain degree of flexibility. Finally, the genetic nature of memory [16, 17] implies the existence of mechanisms that hard-wire external signals into genetic changes that account for retrievable patterns of neuronal activity.

It is relatively straightforward to draw a parallel between the coding capacity of the brain and the coding capacity of chromatin. A single unit of chromatin, which is represented by a single nucleosome, possesses an immense coding power embedded within four pairs of structurally distinct histone proteins and the DNA wrapped around them. The N-terminal portions of histones are subjected to numerous post-translational modifications, including methylation, acetylation, and phosphorylation [18–22]. Given the reversible nature of these modifications, it is very likely that at any given moment in time, each nucleosome posseses a unique pattern of histone modifications. This, in turn, should lead to highly dynamic, individual nucleosome-specific recruitment of the histone binding proteins that recognize individual histone modifications. The recognition of the modified histones is achieved with the help of specific protein domains such as the PHD or bromo domains, which recognize methylated or acetylated lysines within the N-terminal portions of the histone proteins, respectively [19, 23–25]. Some of the histone binding proteins carry multiple recognition domains that may enable, in a neuronal network-like fashion, the establishment of connections not only between various modifications within the individual histone proteins, but between histones within an individual nucleosome or even between individual nucleosomes [26]. Finally, proteins that bind to histones may facilitate looping of large segments of DNA, thus establishing connections between genes located on various chromosomes. All of these epigenetic processes are likely to yield a unique epigenetic signature of individual neurons. This epigenetic diversity is likely to contribute to the distinct features of individual neurons and to the enormous functional plasticity of the neuronal circuitry.

Epigenetic modifications are introduced by histone-modifying or DNA-methylating enzymes, such as acetyltransferases, methyltransferases, and kinases ("writers" by David Allis's definition) [19, 22, 23, 27]. These modifications can be removed by another class of specific enzymes ("erasers"), including deacetlyases, demethylases, and phosphatases [21, 28–35]. Therefore, one might predict that defects in the individual epigenetic "writers" or "erasers", as well as in proteins that specifically recognize the individual histone marks ("readers"), may compromise the central role of chromatin in maintenance of neuronal identity and plasticity. Such reduction in neuronal diversity and/or plasticity might lead, in turn, to a reduction in higher brain function in affected individuals.

In support of the important role of epigenetic regulation in normal brain function, about 7% of the genes known to be associated with mental retardation in humans encode various epigenetic regulators (Table 1) (reviewed in [36–44]).

Table 1 Epigenetic regulators associated with human mental retardation syndromes

Locus	Gene	Protein	Function	Disease	OMIM	Mouse model
Xp22.1	*ARX*	ARX	DB, TR	Nonsyndromic X-linked MR, XLAG syndrome, West syndrome, Partington syndrome	300419	+
Xq13	*ATRX*	ATRX	ATPase/helicase	a-Thalassemia MR syndrome X-linked, MR hypotonic facies syndrome	301040, 309580	+
Xp11.4	*BCOR*	BCOR	DB, Co-repressor	Microphthalmia, Syndromic 2; OFCD syndrome	300166	–
Xp22.2	*CDKL5*	CDKL5	Serine-threonin kinase	Infantile Spasm syndrome	300672	–
8q12	*CHD7*	CHD7	HB	CHARGE syndrome	214800	+
16p13.3	*CREBBP*	CBP	HAT, CREB binding protein	Rubinstein–Taybi syndrome	180849	+
20q11.2	*DNMT3B*	DNMT3B	DNMT	Immunodeficiency-Centromeric Instability-Facial Anomalies syndrome	242860	–
9q34	*EHMT1*	EHMT1	HMT	9q34 subtelomeric deletion syndrome	610253	+
22q13.1	*EP300*	P300	HAT	Rubinstein–Taybi syndrome	180849	+
7q11.23	*GTF2IRD1*	Gtf2ird1	TR	Williams–Beuren syndrome	194050	+
Xp11.2	*JARID1C*	JARID1C/SMCX	HDM	Nonsyndromic X-linked MR	300534	–
20p12	*MACROD2*	MACROD2	Chromatin remodeling	Kabuki syndrome	147920	–
Xq28	*MECP2*	MeCP2	Methyl DNA binding protein	Rett syndrome, Angelman syndrome	312750, 105830	+
Xq13	*Med12*	MED12	TR	Lujan–Fryns syndrome, Opitz–Kaveggia syndrome	309520, 305450	–
5q35	*NSD1*	NSD1	HMT	Sotos syndrome, Weaver syndrome	606681	+
4q16.3	*NSD2*	NSD2/MMSET	HMT	Wolf–Hirschhom syndrome	194190	+
Xq26	*PHF6*	PHF6	TR	Borjeson–Forssman–Lehmann syndrome	301900	–
Xp11.2	*PHF8*	PHF8	HDM, TR	Siderius X-linked MR with cleft palate	300560, 300263	–
Xp22.2	*RPS6KA3*	RSK2	Serine-threonin kinase	Coffin–Lowry syndrome	303600	+
Xq21	*RPS6KA6*	RSK4	Serine-threonin kinase	Nonsyndromic MR	300303	–
Xq27	*SOX3*	SOX3	DB, TR	MR with short stature	300123	–
Xq11.3	*ZNF41*	ZNF41	DB	Nonsyndromic X-linked MR	314995	–
Xp11.3	*ZNF674*	ZNF674	DB	Nonsyndromic X-linked MR	300573	–
Xq21.1	*ZNF711*	ZNF711	DB	Nonsyndromic X-linked MR	300803	–
Xp11.23	*ZNF81*	ZNF81	DB	Nonsyndromic X-linked MR	314998	–

DB DNA binding protein; *DNMT* DNA methyltransferase; *HAT* histone acetyltransferase; *HB* histone binding protein; *HDAC* histone deactelylase; *HDM* histone demethylase; *HMT* histone methyltransferase; *MR* Mental retardation; *TR* transcriptional regulator
OMIM Online Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/omim)

Many of the key symptoms of mental retardation, such as cognitive impairment and deficits in environmental adaptation, are also hallmarks of other neurological diseases such as depression and autism-spectrum disorders. Furthermore, several recent studies have linked abnormal epigenetic regulation to development of various neurological and psychiatric diseases, including Alzheimer's disease, Huntington's disease, drug addiction, depression, autism-spectrum disorders, and schizophrenia (reviewed in [17, 42, 44–51]).

In the following sections, we will discuss some compelling cases for an epigenetic basis of mental retardation. In particular, we will focus on the role of specific histone modifications and DNA methylation in regulation of gene expression. The mechanisms of dysregulation of gene expression in these diseases remain poorly understood. Their understanding will require the elucidation of the nature of genes that are targeted by the specific modifications. This task is greatly complicated by the necessity to identify mRNA and noncoding RNA expression in neuron subpopulations that, in some cases, represent a minor fraction of the extremely heterogeneous brain tissue. Therefore, in the final section, we shall discuss recent methodological advances that may facilitate the understanding of epigenetic mechanisms of regulation of gene expression in individual classes of neurons in vivo.

1.1 Mental Retardation and Suppressive Histone Lysine Methylation

Methylation of histones on individual lysine residues is associated with distinct functional outcomes. The di- or tri-methylation of histone H3 on lysine 9 (H3K9me2, H3K9me3), tri-methylation of lysine 27 (H3K27me3), as well as mono-methylation of lysine 20 of histone H4 (H4K20me1) have been linked to suppression of genes [52]. Contrary to these modifications, tri-methylation of lysine 4 of histone H3 (H3K4me3) is a hallmark of transcriptional activation [53–56]. The tri-methylation of lysine 36 of histone H3 (H3K36me3) contributes to the elongation of the mRNA transcripts [57–63], while methylation of lysine 79 of histone H3 (H3K79me) has been implicated in DNA repair [64]. These, as well as other less characterized lysine methylation marks, are introduced with the help of highly specific histone methyltransferases (HMTases) that target specific lysine residues and catalyze their mono-, di-, or tri-methylation (reviewed in [22]).

Several lines of evidence support a key role for H3K9me2 in regulation of normal brain function. H3K9me2 is catalyzed by a ubiquitously expressed multiprotein HMTase complex, the catalytic core of which is comprised of the HMTases G9a and GLP [65–71]. Genetic ablation of either of these enzymes results in loss of euchromatic H3K9me2 [65, 67–71]. It has been demonstrated that aberrant function of G9a in nonneuronal cells leads to derepression of neuronal genes outside of the nervous system [72–74]. Furthermore, deficiency in Mediator 12, one of the GLP/G9a binding partners and a possible causal factor of human mental retardation

associated with the Lujan–Fryns and the Opitz–Kaveggia syndromes (Table 1) [75], causes aberrant neuronal gene expression in non-neuronal cells [74]. Conversely, in neuronal cells, postnatal ablation of the histone methyltransferases GLP/G9a in mouse forebrain leads to derepression of non-neuronal and early neuronal progenitor genes in various brain areas [76].

Aberrant expression of non-neuronal genes in GLP/G9a deficient neurons is associated with the development of behavioral abnormalities and impaired cognition and learning [76]. Mice with a postnatal forebrain-specific ablation of GLP or G9a show a diminished exploratory behavior in response to a new environment. This behavioral abnormality is not attributable either to motor dysfunction or to increased anxiety. On the contrary, these mice display a general loss of curiosity and interest in their surroundings, which seems to leave them unaware of the potential danger of the environment [76]. Furthermore, postnatal ablation of GLP/G9a leads to severe deficits in motivation and is associated with the development of obesity in mice and humans [76–80].

The cognitive and behavioral defects, caused by ablation of GLP/G9a in postnatal mouse neurons [76], are remarkably similar to symptoms of a human mental retardation syndrome associated with deletion of the subtelomeric region of chromosome 9q34, which contains the *GLP*/*EHMT1* gene (Table 1) [80, 81]. In addition to major defects in cognition and learning, the human 9q34 syndrome is characterized by obesity, childhood hypotonia, and a gradual, age-dependent development of severe apathetic behavior, reduced motor activity, and loss of goal-directed activities [43, 77, 80–82]. The suggested causal role of the *GLP*/*EHMT1* gene alterations in the human 9q34 mental retardation syndrome has been underscored by the identification of various intragenic *GLP*/*EHMT1* mutations in patients with a mental retardation syndrome clinically indistinguishable from the 9q34 deletion syndrome [43, 80]. One of the identified mutations in the *GLP*/*EHMT1* gene leads to an amino acid substitution that is predicted to affect the conformation and hence the activity of the highly conserved, histone methyltransferase encoding SET domain of the GLP/EHMT1 protein [80]. It is not yet known whether defective expression of GLP in humans results in a dysregulation of gene expression as has been seen in mice [76]. Should this indeed be the case, then it is plausible that ectopic expression of non-neuronal genes in adult neurons might contribute to the cognitive and behavioral defects in mice and humans with impaired GLP or G9a expression or function.

Expression of proteins that are normally not expressed in adult neurons may wreak havoc on the otherwise tightly controlled signaling and transcriptional networks in these cells. Thus, the GLP/G9a deficient mouse neurons ectopically express a collection of potent signaling proteins involved in calcium and cAMP signaling (Bank1, Annexin-10, ArhGAP15, CaBP5, MuSK, Plce1), as well as cytoskeletal function (Myosin-1, Myosin-7, Myomesin-2, Titin, Tnnt2) [76]. Expression of these proteins may interfere significantly with the function of the firmly regulated intracellular signaling networks in the brain. Additionally, proteins of the laminin family like Lama2, Lama3, and Lamb3, which become upregulated in the absence of GLP/G9a, are known to control cell-to-cell communication [83] and may, therefore, interfere with neuronal connectivity in the brain. Finally, several of the upregulated

proteins, including Alpha-fetoprotein, Afamin, Carboxylesterase 7, Beta-defensin 1, Gastrokine-1, Olfactomedin-4, Serpin B1b, and Serpin B5, have been described as potent secreted regulators of various cell functions [84–89]. Production of these proteins by neurons could induce secondary changes in neighboring neurons or non-neuronal cells.

The important role of H3K9me2 in maintenance of normal cognition is further supported by the mental retardation syndrome seen in patients with defective function of the PHD finger protein 8 (PHF8) [90], a ubiquitously expressed enzyme that specifically demethylates H3K9me2 [91]. Mutations within the catalytic JmjC (Jumonji-like C terminus) domain of the PHF8 gene are associated with the development of the Siderius type X-linked mental retardation syndrome, which is characterized by mild to severe mental retardation and cleft lip and palate [90, 92, 93]. In addition to its JmjC demethylase domain, the PHF8 protein harbors two functional PHD domains that are implicated in transcriptional regulation and chromatin remodeling [90]. Further studies will be needed to explore whether PHF8 and G9a/GLP represent functionally opposing enzymatic forces that control the expression of neuron and non-neuron specific genes through regulation of H3K9me2 levels.

1.2 Mental Retardation and Activating Histone Lysine Methylation

In contrast to H3K9 di- or tri-methylation, the H3K4me3 is associated with actively transcribed genes or genes that are primed for transcription [22, 54–56]. Accordingly, demethylation of H3K4me3 is associated with decreased transcriptional activity [21, 28–30, 32–35]. The level of H3K4me3 is determined by the balance between the gene locus-specific recruitment and activity of H3K4 HMTases, such as MLL1-5, Set1, Set7/9, Ash1, and H3K4 demethylases [22]. One of the H3K4-specific demethylases that reverses H3K4 tri-methylation to di- and mono-methylated products is JARID1C [73, 94]. Mutations in the JARID1C gene have been linked to the development of nonsyndromic-X-linked mental retardation and epilepsia in humans (Table 1) [73, 94–97]. The patterns of mutations within the JARID1C gene that have been associated with XLMR suggest the loss of JARID1C-mediated control of gene expression as the potential causal mechanism for the development of the disease. Patients carrying point mutations within the exons that encode the catalytic domain of JARID1C have been identified [94–97]. These mutations reduce the activity of the enzyme. Notably, the degree of reduced H3K4me3 demethylase activity caused by the specific mutations in various patients correlated with the severity of the associated mental retardation, indicating a dose-dependent effect of H3K4me3 levels in regulation of cognitive function [73]. Additionally, the XLMR syndrome has been associated with mutations that specifically target the PHD domain of JARID1C, which is required for recruiting JARID1C to the neighboring H3K9me3 [94]. It is likely that recruitment of JARID1C to

H3K9me3 is part of the suppressive mechanism that combines H3K9me3 and, dependent on it, JARID1C-mediated demethylation of H3K4me3.

Given the ubiquitous pattern of JARID1C expression [94], mutations in this gene are likely to suppress the activity of genes in various cell types. The association shown between JARID1C, G9a, and the transcriptional repressor REST in a human kidney cell line suggests the possibility of a common pattern of genes targeted by G9a and JARID1C [73]. Furthermore, loss of JARID1C activity in this cell line leads to derepression of several REST-targeted neuronal genes [73]. Although nothing is known about the nature of JARID1C-regulated genes in mouse or human brain, one might expect that impaired JARID1C function could lead to derepression of G9a targeted genes in neuronal and non-neuronal cells.

In actively transcribed gene loci, H3K4me3 is frequently associated with methylation of lysine 36 of histone H3 [98,99]. This modification correlates with transcriptional elongation, although the exact mechanism of this process is not fully understood. There are several HMTases that methylate H3K36 in vitro, and several of these HMTases, including the Nuclear receptor Set Domain containing proteins 1 and 2 (NSD1, NSD2), have been shown to catalyze tri-methylation of H3K36 in vivo [100–102]. Mutations in the NSD1 gene are associated with mental retardation that hallmarks the human Sotos and the related Weaver syndrome [103–105], whereas mutations in the NSD2 gene are associated with the Wolf Hirschhorn mental retardation syndrome (Table 1) [102]. In addition to the Set domain that is associated with the "writing" of the histone methyl mark, NSD1 also contains several PHD finger domains that specifically "read" the methylation status of the histone lysine residues. Mutations in each of the five PHD finger domains of the NSD1 gene have been described in patients suffering from the disease [40, 104, 105]. Given the known ability of the PHD domains to bind to methylated lysines [26, 56], it is likely that NSD1 recruitment to methylated H3K4 potentiates the transcriptional efficiency of the loci that play a key role in the development and/or function of mature neurons. Early embryonic death of the NSD1 deficient mice precludes the assessment of the role of NSD1 in neuronal gene expression and brain function [100]. However, the presence of various Wolf–Hirschhorn syndrome-like midline defects in haploinsufficient NSD2 mice [102] suggests the possibility of using haploinsufficient NSD1 mice to address the mechanism of NDS1 involvement in regulation of brain function and the development of the human disease. Gene expression analysis of NSD2-deficient, embryonic heart tissue suggests that the observed pathophysiology might be associated with NSD2-mediated suppression of inappropriate gene transcripts in the mammalian heart [102]. However, the nature of NSD2-specific gene targets in the mammalian brain has not yet been addressed.

1.3 Mental Retardation and Histone Lysine Acetylation

The first report describing histone acetylation and its dynamics in developing and adult brain tissue goes back to 1970 [106]. Since then, dynamic acetylation of

individual lysine residues within the N-terminal portion of histone proteins has been identified as a pivotal epigenetic switch in the developmental and temporal control of gene expression [21, 23, 107, 108]. The essential role of histone acetylation in learning and memory was initially suggested by the demonstration of the positive role of histone H4K8 acetylation in the formation of long-term synaptic plasticity in *Aplysia* [109]. These studies showed that learning and memory may be improved through increased histone H4 acetylation. Accordingly, an overall increase in histone acetylation following sytemic in vivo administration of histone deacetylase (HDAC) inhibitors facilitates learning and memory in wild-type mice [110–113].

In agreement with the central role of histone lysine acetylation in memory and learning, reduced expression of histone acetlytransfereses (HAT) leads to severe impairment of brain function. In humans, loss of function mutations in genes encoding either HAT CBP (CREB binding protein) [114–117] or HAT EP300 [118] lead to mental retardation in patients with Rubinstein–Taybi syndrome (RSTS) (Table 1) (reviewed in [119]). Loss of HAT activity of the CBP or p300 protein seems to be sufficient to account for most of the clinical manifestation of the disease [115,120]. The key role of CBP and p300 in RSTS has been underscored by findings that show development of the RSTS-like syndrome in mice heterozygous for null mutations in the CBP or p300 genes [111,121–125]. More specifically, mice that are haplodeficient for CBP exhibit deficits in hippocampus dependent long-term memory tasks and electrophysiological learning and memory paradigms such as late phase L-LTP [111,126,127]. Similar to observations in human patients, the HAT activity of CBP proves to be critically important in regulating memory formation in mice [110,111,128]. These findings suggest that learning and memory in mice and humans depends on the balance between CBP/p300-mediated acetylation and deacetylation, suggesting that learning and memory is controlled by a specific group of HATs and, possibly, specific HDACs. The existence of HDACs that contribute selectively to learning and memory has been suggested by findings that show enhanced memory formation and synaptic plasticity in mice deficient for HDAC2, but not HDAC1 [129].

The mechanism by which the overall or locus-specific increase in histone acetylation contributes positively to memory and synaptic plasticity remains poorly understood. It is possible that long-lasting histone acetylation achieved by pharmacological or genetic HDAC inactivation establishes a surplus of positive regulators of neuronal function without affecting the expression of negative regulators. This scenario is possible only under conditions where the access of HATs to genes encoding negative regulators of memory and learning is limited by the selective lysine methylation of these loci via suppressive lysine methyltransferases. In addition, histone acetylation can favor establishment of memory and learning through direct control of epigenetic suppressors of gene expression. Thus, CBP suppresses the expression of ESET/Setdb1 [130], a H3K9-specific histone methyltransferase that can trigger transcriptional repression and heterochromatin formation [131–133]. Consistent with these observations, deficiency of CBP leads to an increase of H3K9 methylation in mammalian neurons and presumable suppression of gene expression [130].

1.4 Mental Retardation and Histone Phosphorylation

Serine 10 of histone H3 (H3S10ph) is the most thoroughly characterized site of histone phosphorylation [134,135]. Phosphorylation of H3S10 is associated with activation of gene transcription [22,136,137]. This function of H3S10ph could be due to the specific recognition of phosphorylated chromatin by the phospho-binding "reader" proteins 14-3-3, which function as transcriptional activators [138]. Concomitantly, with the recruitment of the 14-3-3 and potentially other transcriptional activators, phosphorylation of H3S10 hinders interaction of the transcriptional suppressor HP1 with the neighboring H3K9me2. The latter event facilitates derepression of gene loci associated with H3S10ph [134,139,140].

In mammals, phosphorylation of H3S10 is catalyzed by the kinases MSK1/2 and RSK2 [141–143]. The latter has been implicated in a clinically distinct mental retardation syndrome, the Coffin–Lowry syndrome (CLS) (Table 1), which is characterized by severe mental retardation in combination with skeletal and cardiac abnormalities, kyphoscoliosis, as well as auditory and visual abnormalities (reviewed in [144]). CLS is associated with mutations in the RPS6KA3 gene, which encodes the RSK2 protein [145]. Notably, deletion of another member of the RSK family, RSK4, is linked to mental retardation in patients suffering from X-linked deafness (Table 1) [146]. Mutations in two members of the RSK family, both leading to mental retardation, emphasize the important role of this kinase family in neuronal function and memory formation. In support of the important role of RSK2 in mental health, the RSK2 deficient mice display impairment of spatial working memory, delayed acquisition of a spatial reference memory task, and long-term spatial memory deficits [147,148].

Several mechanisms may account for the involvement of RSK2 in mental health. The RSK2-mediated H3S10 phosphorylation can activate genes that control cognition and memory. However, the exact gene targets of H3S10 phosphorylation in neurons are not fully characterized. In addition to its direct function on gene transcription activation via phosphorylation of histone H3, it is likely that RSK2 can regulate gene expression through binding and phosphorylation of the transcriptional activators CBP and Creb1, respectively [149–151]. Psychostimulant-induced H3S10 phosphorylation in striatal neurons leads to the transcriptional activation of *c-Fos* and *Jun* and other immediate early genes that orchestrate the dopamine-induced behavioral response in these mice [152,153]. This function of H3S10ph may contribute to the impaired brain function of patients with reduced activity in the H3S10 phosphorylating kinases.

1.5 Mental Retardation and DNA Methylation

Methylation of DNA is the most stable form of epigenetic modification that contributes to suppression of gene expression [154]. Methylated CpG nucleotides,

like modified histones, regulate gene expression either by recruiting transcriptional repressors or by hindering binding of transcriptional activators [154–159]. The DNA methylation is catalyzed by structurally and functionally distinct DNA methyltransferases (DNMTs) [27,154,160]. Recent studies from the laboratories of Nathaniel Heintz and Anjana Rao suggested the possibility of DNA demethylation via the generation of 5′hydroxymethylcytosine [161,162]. The generation of this mark is catalyzed by enzymes of the Tet protein family [161]. However, at this point it is not clear whether 5′hydroxymethylcytosine represents a novel stable form of DNA modification or the intermediate product of DNA demethylation [160]. Furthermore, nothing is known about the functional impact of this mark on regulation of gene expression.

The importance of DNA methylation for brain function is supported by the demonstration that DNMT gene expression is upregulated in the adult rat hippocampus following learning and memory paradigms and that DNMT inhibition blocks memory formation [163, 164]. Furthermore, association between several mental retardation syndromes and disabling mutation in DNMTases or proteins that recognize methylated DNA [37, 43, 51, 165] demonstrates that DNA methylation is a crucial step in regulation of cognition.

One of these mental retardation syndromes is Immunodeficiency, Centromere instability and Facial anomalies (ICF) syndrome (Table 1), which is caused by a mutation in the DNA-methyltransferase-3b (Dnmt3b) [166,167]. The most frequent symptoms of the syndrome are facial dysmorphism, mental retardation, and variable immune deficiency [168,169]. Judging from the analysis of gene expression in cell lines derived from ICF patients, mutations in DNMT3b cause ectopic expression of genes critical for immune function, development, and neurogenesis [170,171]. A fraction of the ectopically upregulated genes display low levels of DNA methylation due to the impaired function of DNMT3b [170,171]. It is likely that, similar to G9a or GLP deficient neurons, the function of neurons expressing the mutant DNMT3b is affected by erroneous expression of proteins that are normally silenced by DNA methylation.

Rett syndrome (RTT) is a postnatal, progressive disorder that affects girls in early childhood after a period of 6–18 months of normal development (Table 1) (reviewed in [165]). The onset of the disease is characterized by developmental stagnation, followed by a rapid regression and loss of all acquired skills such as speech, social interaction, and motor coordination. RTT is caused by sporadic mutations in the X-chromosome localized gene encoding the transcriptional regulator methyl-CpG-binding protein 2 (MeCP2) [172]. The analysis of the MeCP2 containing protein complexes suggested an ability of MeCP2 to coordinate the assembly of potent repressor protein complexes at the sites of methylated DNA. MeCP2 was found to be associated with corepressors such as Sin3a, the histone deacetylases HDAC1 and HDAC2, DNA methyltransferase 1 (Dnmt1), the histone methyltransferase SuvH1, corepressors c-Ski and N-Cor, the transcription factor TFIIB, the SWI/SNF-related chromatin-remodeling protein Brama, and the SWI2/SFN2 DNA helicase/ATPase [173–180]. Mutations in the latter gene are responsible for causing the alpha-thalassemia/mental retardation syndrome X-linked

(ATR-X) (Table 1) [181]. Given the general role of MeCP2 in regulation of gene expression, it is likely that MeCP2 can trigger a chain of transcriptional changes that will distort neuronal function. Therefore, it is not surprising that not only deficiency but also upregulation of MeCP2 function, due to duplication of the MeCP2 gene, can cause severe mental retardation in humans [165,182].

The causal role of MeCP2 deficiency or overexpression in mental retardation is supported by the development of a Rett sydrome-like disease in mice with germline or neuron-specific deficiency or upregulation of MeCP2 [183–187]. Ablation of MeCP2 in embryonic brain is sufficient to cause the disease in mice, and even postnatal forebrain-specific deletion leads to similar, albeit slightly less severe, abnormalities [183,187]. The critical role of MeCP2 in mature neurons is further supported by the fact that either postnatal and/or postnatal neuron-specific MeCP2 expression in MeCP2 deficient mice can partly rescue the neurological phenotypes [188–190]. These studies indicate that neuronal MeCP2 dysfunction in postmitotic neurons is partly responsible for the neurological abnormalities, and that the changes causing the neurological abnormalities in these mice remain reversible. Similar to the situation with humans, overexpression of MeCP2 in mice results in neurological abnormalities that correlate positively with levels of MeCP2 overexpression [186].

1.6 Overcoming Difficulties Associated with Neuronal Heterogeneity

Epigenetic regulators have the potential to control a large variety of different genes. Elucidation of the nature of the genes that are regulated by specific modifications of histones and DNA is essential for an understanding of the specific physiological functions of these modifications. Interestingly, the analysis of mRNA expression changes in mouse and human tissues deficient for individual histone or DNA modifying enzymes has revealed a relatively high degree of selectivity for genes controlled by distinct epigenetic regulators. These results suggest that it may be possible to identify genes and/or corresponding pathways responsible for the defective cognition, memory, and social adaptation observed in genetically engineered mice that model human mental retardation syndromes. Furthermore, a comparison of the target genes controlled by individual epigenetic regulators may help to elucidate their specific and/or overlapping functions. It is possible that mutations in different epigenetic regulators are associated with the disruption of large regulatory complexes leading to changes in a common subset of target genes.

The task of precise characterization of gene expression in the brain is greatly complicated by the morphological and functional heterogeneity of neuronal and non-neuronal subpopulations. Recent technological advances have made it possible to characterize mRNA expression patterns in small and functionally distinct cell populations [191,192]. This technology, called TRAP (Translating Ribosome

Affinity Purification), utilizes cell type-specific tagging of the large ribosomal subunit protein L10 with the enhanced green fluorescent protein (eGFP). Expression of an eGFP tagged ribosomal protein selectively in neurons of interest is followed by immunoaffinity purification and extraction of the polyribosome bound, cell type-specific mRNAs from crude brain extracts [191,192]. Currently, the neuronal cell type specificity of ribosomal tagging is achieved by using Bacterial Artificial Chromosome (BAC) transgenes to express the eGFP-tagged ribosomal protein in specific cell types. Cell type-specific expression of the tagged ribosomes is achieved by using a cell type-specific promotor, which, in most cases, reproduces faithfully the expression pattern of the endogenous genes [191,192]. However, one can envision modifications of this technology that may involve virus-mediated expression of eGFP-tagged ribosomes specifically in small groups of neurons in distinct brain regions or the generation of knock-in mice expressing the eGFP-tagged ribosomal proteins under the control of endogenous neuronal gene promoters. Regardless of the details of the tagging approach, extraction of mRNA from eGFP-labeled polyribosomes allows the identification of translated mRNAs at any time point in any given cell type of interest in response to any environmental, genetic, or pharmacological perturbation.

Unlike traditional approaches, the TRAP methodology has the advantage of combining the detection of translated mRNAs with cell type specificity. The methodology also avoids lengthy and frequently inefficient cell separation procedures, which besides severing the mRNA containing axons and dendrites also promotes RNA degradation and cell death. The use of established bacTRAP mouse lines ensures that the mRNA translational profiles can be reproducibly obtained and directly compared from the same neuronal cell population in experimental and control mice. In conclusion, the TRAP technology should largely overcome the obstacles presented by neuronal heterogeneity in the study of the genetic basis of various neurological and psychiatric disorders.

2 Prospectus

Epigenetic control of brain function is an emerging field of neuroscience. With each neuron being potentially different from all others, and with the constant flow of information through the brain, the epigenetic regulators are likely to play a central role in transmitting information about the environment to the chromatin. Epigenetic lesions associated with mental retardation offer a glimpse into the mechanisms underlying gene regulation by histone- and DNA-modifying enzymes as well as the proteins that recognize these modifications. The epigenetic targets and signal transduction mechanisms that control gene target specificity remain largely unknown. Recent studies suggest a significant role for noncoding RNAs in regulation of gene function. These long noncoding RNAs may control targeting of the epigenetic regulators to specific gene loci. In addition, it is plausible that many epigenetic regulators control the expression of miRNAs, which play a major role in

regulation of protein expression in neurons. In this respect, it is possible that some of the genes involved in mental retardation may contribute to the epigenetic control of brain function through production of specific ncRNAs or miRNAs. The development of the ultra-sensitive techniques of RNA sequencing, which could be combined with the TRAP method or similar technologies, may allow a single cell analysis of the neuronal transcriptome. This, in turn, may facilitate the identification of critical genetic and neuronal circuitry junctions that play a key role in cognition, learning, and memory.

Acknowledgments We would like to especially thank Eric Nestler for critical reading of the review and Debra Poulter for her help in the preparation of the chapter. We would like to recognize and apologize to colleagues whose work could not be acknowledged due to space limitations.

References

1. Jackson A, Moritz CT, Mavoori J, Lucas TH, Fetz EE (2006) The Neurochip BCI: towards a neural prosthesis for upper limb function. IEEE Trans Neural Syst Rehabil Eng 14:187–190
2. Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, Branner A, Chen D, Penn RD, Donoghue JP (2006) Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature 442:164–171
3. Velliste M, Perel S, Spalding MC, Whitford AS, Schwartz AB (2008) Cortical control of a prosthetic arm for self-feeding. Nature 453:1098–1101
4. Moritz CT, Perlmutter SI, Fetz EE (2008) Direct control of paralysed muscles by cortical neurons. Nature 456:639–642
5. Qiu Z, Ghosh A (2008) A brief history of neuronal gene expression: regulatory mechanisms and cellular consequences. Neuron 60:449–455
6. Ooi L, Wood IC (2008) Regulation of gene expression in the nervous system. Biochem J 414:327–341
7. Greenberg ME, Greene LA, Ziff EB (1985) Nerve growth factor and epidermal growth factor induce rapid transient changes in proto-oncogene transcription in PC12 cells. J Biol Chem 260:14101–14110
8. Greenberg ME, Ziff EB, Greene LA (1986) Stimulation of neuronal acetylcholine receptors induces rapid gene transcription. Science 234:80–83
9. Ginty DD, Bading H, Greenberg ME (1992) Trans-synaptic regulation of gene expression. Curr Opin Neurobiol 2:312–316
10. Guitart X, Thompson MA, Mirante CK, Greenberg ME, Nestler EJ (1992) Regulation of cyclic AMP response element-binding protein (CREB) phosphorylation by acute and chronic morphine in the rat locus coeruleus. J Neurochem 58:1168–1171
11. Montminy MR, Low MJ, Tapia-Arancibia L, Reichlin S, Mandel G, Goodman RH (1986) Cyclic AMP regulates somatostatin mRNA accumulation in primary diencephalic cultures and in transfected fibroblast cells. J Neurosci 6:1171–1176
12. Montminy MR, Sevarino KA, Wagner JA, Mandel G, Goodman RH (1986) Identification of a cyclic-AMP-responsive element within the rat somatostatin gene. Proc Natl Acad Sci USA 83:6682–6686
13. Montminy MR, Bilezikjian LM (1987) Binding of a nuclear protein to the cyclic-AMP response element of the somatostatin gene. Nature 328:175–178
14. Morgan JI, Cohen DR, Hempstead JL, Curran T (1987) Mapping patterns of c-fos expression in the central nervous system after seizure. Science 237:192–197

15. Sheng M, Thompson MA, Greenberg ME (1991) CREB: a Ca(2+)-regulated transcription factor phosphorylated by calmodulin-dependent kinases. Science 252:1427–1430
16. Kandel ER (2001) The molecular biology of memory storage: a dialogue between genes and synapses. Science 294:1030–1038
17. Borrelli E, Nestler EJ, Allis CD, Sassone-Corsi P (2008) Decoding the epigenetic language of neuronal plasticity. Neuron 60:961–974
18. Cheung P, Allis CD, Sassone-Corsi P (2000) Signaling to chromatin through histone modifications. Cell 103:263–271
19. Strahl BD, Allis CD (2000) The language of covalent histone modifications. Nature 403:41–45
20. Crosio C, Heitz E, Allis CD, Borrelli E, Sassone-Corsi P (2003) Chromatin remodeling and neuronal response: multiple signaling pathways induce specific histone H3 modifications and early gene expression in hippocampal neurons. J Cell Sci 116:4905–4914
21. Berger SL (2007) The complex language of chromatin regulation during transcription. Nature 447:407–412
22. Kouzarides T (2007) Chromatin modifications and their function. Cell 128:693–705
23. Jenuwein T, Allis CD (2001) Translating the histone code. Science 293:1074–1080
24. Daniel JA, Pray-Grant MG, Grant PA (2005) Effector proteins for methylated histones: an expanding family. Cell Cycle 4:919–926
25. Brenner C, Fuks F (2007) A methylation rendezvous: reader meets writers. Dev Cell 12:843–844
26. Ruthenburg AJ, Li H, Patel DJ, Allis CD (2007) Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol 8:983–994
27. Klose RJ, Bird AP (2006) Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci 31:89–97
28. Bannister AJ, Kouzarides T (2005) Reversing histone methylation. Nature 436:1103–1106
29. Klose RJ, Zhang Y (2007) Regulation of histone methylation by demethylimination and demethylation. Nat Rev Mol Cell Biol 8:307–318
30. Shi Y (2007) Histone lysine demethylases: emerging roles in development, physiology and disease. Nat Rev Genet 8:829–833
31. Shi Y, Whetstine JR (2007) Dynamic regulation of histone lysine methylation by demethylases. Mol Cell 25:1–14
32. Cloos PA, Christensen J, Agger K, Helin K (2008) Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. Genes Dev 22:1115–1140
33. Agger K, Christensen J, Cloos PA, Helin K (2008) The emerging functions of histone demethylases. Curr Opin Genet Dev 18:159–168
34. Pasini D, Bracken AP, Agger K, Christensen J, Hansen K, Cloos PA, Helin K (2008) Regulation of stem cell differentiation by histone methyltransferases and demethylases. Cold Spring Harb Symp Quant Biol 73:253–263
35. Lan F, Nottke AC, Shi Y (2008) Mechanisms involved in the regulation of histone lysine demethylases. Curr Opin Cell Biol 20:316–325
36. Inlow JK, Restifo LL (2004) Molecular and comparative genetics of mental retardation. Genetics 166:835–881
37. Chelly J, Khelfaoui M, Francis F, Cherif B, Bienvenu T (2006) Genetics and pathophysiology of mental retardation. Eur J Hum Genet 14:701–713
38. Froyen G, Bauters M, Voet T, Marynen P (2006) X-linked mental retardation and epigenetics. J Cell Mol Med 10:808–825
39. Maekawa M, Watanabe Y (2007) Epigenetics: relations to disease and laboratory findings. Curr Med Chem 14:2642–2653
40. Baker LA, Allis CD, Wang GG (2008) PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. Mutat Res 647:3–12
41. Grant ME (2008) The epigenetic origins of mental retardation. Clin Genet 73:528–530

42. Akbarian S, Huang HS (2009) Epigenetic regulation in human brain-focus on histone lysine methylation. Biol Psychiatry 65:198–203
43. Kramer JM, van Bokhoven H (2009) Genetic and epigenetic defects in mental retardation. Int J Biochem Cell Biol 41:96–107
44. Graff J, Mansuy IM (2009) Epigenetic dysregulation in cognitive disorders. Eur J Neurosci 30:1–8
45. Levenson JM, Sweatt JD (2005) Epigenetic mechanisms in memory formation. Nat Rev Neurosci 6:108–118
46. Tsankova N, Renthal W, Kumar A, Nestler EJ (2007) Epigenetic regulation in psychiatric disorders. Nat Rev Neurosci 8:355–367
47. Renthal W, Nestler EJ (2008) Epigenetic mechanisms in drug addiction. Trends Mol Med 14:341–350
48. Jiang Y, Langley B, Lubin FD, Renthal W, Wood MA, Yasui DH, Kumar A, Nestler EJ, Akbarian S, Beckel-Mitchener AC (2008) Epigenetics in the nervous system. J Neurosci 28:11753–11759
49. Costa E, Chen Y, Dong E, Grayson DR, Kundakovic M, Maloku E, Ruzicka W, Satta R, Veldic M, Zhubi A et al (2009) GABAergic promoter hypermethylation as a model to study the neurochemistry of schizophrenia vulnerability. Expert Rev Neurother 9:87–98
50. Akbarian S (2009) The molecular pathology of schizophrenia-Focus on histone and DNA modifications. Brain Res Bull [Epub ahead of print]
51. Urdinguio RG, Sanchez-Mut JV, Esteller M (2009) Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies. Lancet Neurol 8:1056–1072
52. Oda H, Okamoto I, Murphy N, Chu J, Price SM, Shen MM, Torres-Padilla ME, Heard E, Reinberg D (2009) Monomethylation of histone H4-lysine 20 is involved in chromosome structure and stability and is essential for mouse development. Mol Cell Biol 29:2278–2295
53. Cao R, Zhang Y (2004) The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr Opin Genet Dev 14:155–164
54. Martin C, Zhang Y (2005) The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol 6:838–849
55. Shilatifard A (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu Rev Biochem 75:243–269
56. Ruthenburg AJ, Allis CD, Wysocka J (2007) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. Mol Cell 25:15–30
57. Li J, Moazed D, Gygi SP (2002) Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. J Biol Chem 277:49383–49388
58. Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C et al (2003) Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol 23:4207–4218
59. Xiao T, Hall H, Kizer KO, Shibata Y, Hall MC, Borchers CH, Strahl BD (2003) Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. Genes Dev 17:654–663
60. Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol 25:3305–3316
61. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. Cell 129:823–837
62. Yuan W, Xie J, Long C, Erdjument-Bromage H, Ding X, Zheng Y, Tempst P, Chen S, Zhu B, Reinberg D (2009) Heterogeneous nuclear ribonucleoprotein L Is a subunit of human KMT3a/Set2 complex required for H3 Lys-36 trimethylation activity in vivo. J Biol Chem 284:15701–15707

63. Li Y, Trojer P, Xu CF, Cheung P, Kuo A, Drury WJ 3rd, Qiao Q, Neubert TA, Xu RM, Gozani O et al (2009) The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. J Biol Chem 284:34283–34295
64. Humpal SE, Robinson DA, Krebs JE (2009) Marks to stop the clock: histone modifications and checkpoint regulation in the DNA damage response. Biochem Cell Biol 87:243–253
65. Tachibana M, Sugimoto K, Fukushima T, Shinkai Y (2001) Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. J Biol Chem 276:25309–25317
66. Ogawa H, Ishiguro K, Gaubatz S, Livingston DM, Nakatani Y (2002) A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. Science 296:1132–1136
67. Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H et al (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev 16:1779–1791
68. Rice JC, Briggs SD, Ueberheide B, Barber CM, Shabanowitz J, Hunt DF, Shinkai Y, Allis CD (2003) Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol Cell 12:1591–1598
69. Peters AH, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y et al (2003) Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell 12:1577–1589
70. Tachibana M, Ueda J, Fukuda M, Takeda N, Ohta T, Iwanari H, Sakihama T, Kodama T, Hamakubo T, Shinkai Y (2005) Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. Genes Dev 19:815–826
71. Sampath SC, Marazzi I, Yap KL, Sampath SC, Krutchinsky AN, Mecklenbrauker I, Viale A, Rudensky E, Zhou MM, Chait BT et al (2007) Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. Mol Cell 27:596–608
72. Roopra A, Qazi R, Schoenike B, Daley TJ, Morrison JF (2004) Localized domains of G9a-mediated histone methylation are required for silencing of neuronal genes. Mol Cell 14:727–738
73. Tahiliani M, Mei P, Fang R, Leonor T, Rutenberg M, Shimizu F, Li J, Rao A, Shi Y (2007) The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. Nature 447:601–605
74. Ding N, Zhou H, Esteve PO, Chin HG, Kim S, Xu X, Joseph SM, Friez MJ, Schwartz CE, Pradhan S et al (2008) Mediator links epigenetic silencing of neuronal gene expression with x-linked mental retardation. Mol Cell 31:347–359
75. Schwartz CE, Tarpey PS, Lubs HA, Verloes A, May MM, Risheg H, Friez MJ, Futreal PA, Edkins S, Teague J et al (2007) The original Lujan syndrome family has a novel missense mutation (p.N1007S) in the MED12 gene. J Med Genet 44:472–477
76. Schaefer A, Sampath S, Intrator A, Min A, Gertler T, Surmeier DJ, Tarakhovsky A, Greengard P (2009) Control of cognition and adaptive behavior by the GLP/G9a epiegnetic suppressor complex. Neuron 64(5):678–691
77. Cormier-Daire V, Molinari F, Rio M, Raoul O, de Blois MC, Romana S, Vekemans M, Munnich A, Colleaux L (2003) Cryptic terminal deletion of chromosome 9q34: a novel cause of syndromic obesity in childhood? J Med Genet 40:300–303
78. Goldstone AP, Beales PL (2008) Genetic obesity syndromes. Front Horm Res 36:37–60
79. Fyffe SL, Neul JL, Samaco RC, Chao HT, Ben-Shachar S, Moretti P, McGill BE, Goulding EH, Sullivan E, Tecott LH et al (2008) Deletion of Mecp2 in Sim1-expressing neurons reveals a critical role for MeCP2 in feeding behavior, aggression, and the response to stress. Neuron 59:947–958
80. Kleefstra T, van Zelst-Stams WA, Nillesen WM, Cormier-Daire V, Houge G, Foulds N, van Dooren M, Willemsen MH, Pfundt R, Turner A et al (2009) Further clinical and

molecular delineation of the 9q subtelomeric deletion syndrome supports a major contribution of EHMT1 haploinsufficiency to the core phenotype. J Med Genet 46(9):598–606
81. Verhoeven WM, Kleefstra T, Egger JI (2010) Behavioral phenotype in the 9q subtelomeric deletion syndrome: a report about two adult patients. Am J Med Genet B Neuropsychiatr Genet 153B(2):536–541
82. Kleefstra T, Brunner HG, Amiel J, Oudakker AR, Nillesen WM, Magee A, Genevieve D, Cormier-Daire V, van Esch H, Fryns JP et al (2006) Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. Am J Hum Genet 79:370–377
83. Scheele S, Nystrom A, Durbeej M, Talts JF, Ekblom M, Ekblom P (2007) Laminin isoforms in development and disease. J Mol Med 85:825–836
84. Benarafa C, Cooley J, Zeng W, Bird PI, Remold-O'Donnell E (2002) Characterization of four murine homologs of the human ov-serpin monocyte neutrophil elastase inhibitor MNEI (SERPINB1). J Biol Chem 277:42028–42033
85. Oien KA, McGregor F, Butler S, Ferrier RK, Downie I, Bryce S, Burns S, Keith WN (2004) Gastrokine 1 is abundantly and specifically expressed in superficial gastric epithelium, down-regulated in gastric carcinoma, and shows high evolutionary conservation. J Pathol 203:789–797
86. Niyonsaba F, Ushio H, Nagaoka I, Okumura K, Ogawa H (2005) The human beta-defensins (−1, −2, −3, −4) and cathelicidin LL-37 induce IL-18 secretion through p38 and ERK MAPK activation in primary human keratinocytes. J Immunol 175:1776–1784
87. Bailey CM, Khalkhali-Ellis Z, Seftor EA, Hendrix MJ (2006) Biological functions of maspin. J Cell Physiol 209:617–624
88. Holmes RS, Cox LA, Vandeberg JL (2008) Mammalian carboxylesterase 5: comparative biochemistry and genomics. Comp Biochem Physiol Part D Genomics Proteomics 3: 195–204
89. Jensen EV, Jacobson HI, Walf AA, Frye CA (2010) Estrogen action: a historic perspective on the implications of considering alternative approaches. Physiol Behav 99(2):151–162
90. Laumonnier F, Holbert S, Ronce N, Faravelli F, Lenzner S, Schwartz CE, Lespinasse J, Van Esch H, Lacombe D, Goizet C et al (2005) Mutations in PHF8 are associated with X linked mental retardation and cleft lip/cleft palate. J Med Genet 42:780–786
91. Loenarz C, Ge W, Coleman ML, Rose NR, Cooper CD, Klose RJ, Ratcliffe PJ, Schofield CJ (2009) PHF8, a gene associated with cleft lip/palate and mental retardation, encodes for an N{varepsilon}-dimethyl lysine demethylase. Hum Mol Genet 19(2):217–222
92. Abidi FE, Miano MG, Murray JC, Schwartz CE (2007) A novel mutation in the PHF8 gene is associated with X-linked mental retardation with cleft lip/cleft palate. Clin Genet 72:19–22
93. Koivisto AM, Ala-Mello S, Lemmela S, Komu HA, Rautio J, Jarvela I (2007) Screening of mutations in the PHF8 gene and identification of a novel mutation in a Finnish family with XLMR and cleft lip/cleft palate. Clin Genet 72:145–149
94. Iwase S, Lan F, Bayliss P, de la Torre-Ubieta L, Huarte M, Qi HH, Whetstine JR, Bonni A, Roberts TM, Shi Y (2007) The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. Cell 128:1077–1088
95. Jensen LR, Amende M, Gurok U, Moser B, Gimmel V, Tzschach A, Janecke AR, Tariverdian G, Chelly J, Fryns JP et al (2005) Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. Am J Hum Genet 76:227–236
96. Santos C, Rodriguez-Revenga L, Madrigal I, Badenas C, Pineda M, Mila M (2006) A novel mutation in JARID1C gene associated with mental retardation. Eur J Hum Genet 14:583–586
97. Tzschach A, Lenzner S, Moser B, Reinhardt R, Chelly J, Fryns JP, Kleefstra T, Raynaud M, Turner G, Ropers HH et al (2006) Novel JARID1C/SMCX mutations in patients with X-linked mental retardation. Hum Mutat 27:389

98. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP et al (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448:553–560
99. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP et al (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458:223–227
100. Rayasam GV, Wendling O, Angrand PO, Mark M, Niederreither K, Song L, Lerouge T, Hager GL, Chambon P, Losson R (2003) NSD1 is essential for early post-implantation development and has a catalytically active SET domain. EMBO J 22:3153–3163
101. Wang GG, Cai L, Pasillas MP, Kamps MP (2007) NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. Nat Cell Biol 9:804–812
102. Nimura K, Ura K, Shiratori H, Ikawa M, Okabe M, Schwartz RJ, Kaneda Y (2009) A histone H3 lysine 36 trimethyltransferase links Nkx2-5 to Wolf–Hirschhorn syndrome. Nature 460:287–291
103. Kurotaki N, Imaizumi K, Harada N, Masuno M, Kondoh T, Nagai T, Ohashi H, Naritomi K, Tsukahara M, Makita Y et al (2002) Haploinsufficiency of NSD1 causes Sotos syndrome. Nat Genet 30:365–366
104. Douglas J, Hanks S, Temple IK, Davies S, Murray A, Upadhyaya M, Tomkins S, Hughes HE, Cole TR, Rahman N (2003) NSD1 mutations are the major cause of Sotos syndrome and occur in some cases of Weaver syndrome but are rare in other overgrowth phenotypes. Am J Hum Genet 72:132–143
105. Tatton-Brown K, Douglas J, Coleman K, Baujat G, Cole TR, Das S, Horn D, Hughes HE, Temple IK, Faravelli F et al (2005) Genotype-phenotype associations in Sotos syndrome: an analysis of 266 individuals with NSD1 aberrations. Am J Hum Genet 77:193–204
106. Bondy SC, Roberts S, Morelos BS (1970) Histone-acetylating enzyme of brain. Biochem J 119:665–672
107. Wolffe AP (1996) Histone deacetylase: a regulator of transcription. Science 272:371–372
108. Grunstein M (1997) Histone acetylation in chromatin structure and transcription. Nature 389:349–352
109. Guan Z, Giustetto M, Lomvardas S, Kim JH, Miniaci MC, Schwartz JH, Thanos D, Kandel ER (2002) Integration of long-term-memory-related synaptic plasticity involves bidirectional regulation of gene expression and chromatin structure. Cell 111:483–493
110. Levenson JM, O'Riordan KJ, Brown KD, Trinh MA, Molfese DL, Sweatt JD (2004) Regulation of histone acetylation during memory formation in the hippocampus. J Biol Chem 279:40545–40559
111. Alarcon JM, Malleret G, Touzani K, Vronskaya S, Ishii S, Kandel ER, Barco A (2004) Chromatin acetylation, memory, and LTP are impaired in CBP+/− mice: a model for the cognitive deficit in Rubinstein–Taybi syndrome and its amelioration. Neuron 42:947–959
112. Yeh SH, Lin CH, Gean PW (2004) Acetylation of nuclear factor-kappaB in rat amygdala improves long-term but not short-term retention of fear memory. Mol Pharmacol 65:1286–1292
113. Vecsey CG, Hawk JD, Lattal KM, Stein JM, Fabian SA, Attner MA, Cabrera SM, McDonough CB, Brindle PK, Abel T et al (2007) Histone deacetylase inhibitors enhance memory and synaptic plasticity via CREB:CBP-dependent transcriptional activation. J Neurosci 27:6128–6140
114. Petrij F, Giles RH, Dauwerse HG, Saris JJ, Hennekam RC, Masuno M, Tommerup N, van Ommen GJ, Goodman RH, Peters DJ et al (1995) Rubinstein–Taybi syndrome caused by mutations in the transcriptional co-activator CBP. Nature 376:348–351
115. Kalkhoven E, Roelfsema JH, Teunissen H, den Boer A, Ariyurek Y, Zantema A, Breuning MH, Hennekam RC, Peters DJ (2003) Loss of CBP acetyltransferase activity by PHD finger mutations in Rubinstein–Taybi syndrome. Hum Mol Genet 12:441–450
116. Coupry I, Monnet L, Attia AA, Taine L, Lacombe D, Arveiler B (2004) Analysis of CBP (CREBBP) gene deletions in Rubinstein–Taybi syndrome patients using real-time quantitative PCR. Hum Mutat 23:278–284

117. Coupry I, Roudaut C, Stef M, Delrue MA, Marche M, Burgelin I, Taine L, Cruaud C, Lacombe D, Arveiler B (2002) Molecular analysis of the CBP gene in 60 patients with Rubinstein-Taybi syndrome. J Med Genet 39:415–421
118. Roelfsema JH, White SJ, Ariyurek Y, Bartholdi D, Niedrist D, Papadia F, Bacino CA, den Dunnen JT, van Ommen GJ, Breuning MH et al (2005) Genetic heterogeneity in Rubinstein–Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. Am J Hum Genet 76:572–580
119. Hennekam RC (2006) Rubinstein–Taybi syndrome. Eur J Hum Genet 14:981–985
120. Murata T, Kurokawa R, Krones A, Tatsumi K, Ishii M, Taki T, Masuno M, Ohashi H, Yanagisawa M, Rosenfeld MG et al (2001) Defect of histone acetyltransferase activity of the nuclear transcriptional coactivator CBP in Rubinstein–Taybi syndrome. Hum Mol Genet 10:1071–1076
121. Tanaka Y, Naruse I, Maekawa T, Masuya H, Shiroishi T, Ishii S (1997) Abnormal skeletal patterning in embryos lacking a single Cbp allele: a partial similarity with Rubinstein–Taybi syndrome. Proc Natl Acad Sci USA 94:10215–10220
122. Yao TP, Oh SP, Fuchs M, Zhou ND, Ch'ng LE, Newsome D, Bronson RT, Li E, Livingston DM, Eckner R (1998) Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. Cell 93:361–372
123. Oike Y, Hata A, Mamiya T, Kaname T, Noda Y, Suzuki M, Yasue H, Nabeshima T, Araki K, Yamamura K (1999) Truncated CBP protein leads to classical Rubinstein–Taybi syndrome phenotypes in mice: implications for a dominant-negative mechanism. Hum Mol Genet 8:387–396
124. Kung AL, Rebel VI, Bronson RT, Ch'ng LE, Sieff CA, Livingston DM, Yao TP (2000) Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. Genes Dev 14:272–277
125. Zhang Z, Hofmann C, Casanova E, Schutz G, Lutz B (2004) Generation of a conditional allele of the CBP gene in mouse. Genesis 40:82–89
126. Wood MA, Kaplan MP, Park A, Blanchard EJ, Oliveira AM, Lombardi TL, Abel T (2005) Transgenic mice expressing a truncated form of CREB-binding protein (CBP) exhibit deficits in hippocampal synaptic plasticity and memory storage. Learn Mem 12:111–119
127. Wood MA, Attner MA, Oliveira AM, Brindle PK, Abel T (2006) A transcription factor-binding domain of the coactivator CBP is essential for long-term memory and the expression of specific target genes. Learn Mem 13:609–617
128. Korzus E, Rosenfeld MG, Mayford M (2004) CBP histone acetyltransferase activity is a critical component of memory consolidation. Neuron 42:961–972
129. Guan JS, Haggarty SJ, Giacometti E, Dannenberg JH, Joseph N, Gao J, Nieland TJ, Zhou Y, Wang X, Mazitschek R et al (2009) HDAC2 negatively regulates memory formation and synaptic plasticity. Nature 459:55–60
130. Lee J, Hagerty S, Cormier KA, Kim J, Kung AL, Ferrante RJ, Ryu H (2008) Monoallele deletion of CBP leads to pericentromeric heterochromatin condensation through ESET expression and histone H3 (K9) methylation. Hum Mol Genet 17:1774–1782
131. Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ 3rd (2002) SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev 16:919–932
132. Yang L, Xia L, Wu DY, Wang H, Chansky HA, Schubach WH, Hickstein DD, Zhang Y (2002) Molecular cloning of ESET, a novel histone H3-specific methyltransferase that interacts with ERG transcription factor. Oncogene 21:148–152
133. Wang H, An W, Cao R, Xia L, Erdjument-Bromage H, Chatton B, Tempst P, Roeder RG, Zhang Y (2003) mAM facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause transcriptional repression. Mol Cell 12:475–487
134. Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H, Allis CD (2005) Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. Nature 438:1116–1122

135. Johansen KM, Johansen J (2006) Regulation of chromatin structure by histone H3S10 phosphorylation. Chromosome Res 14:393–404
136. Zippo A, De Robertis A, Serafini R, Oliviero S (2007) PIM1-dependent phosphorylation of histone H3 at serine 10 is required for MYC-dependent transcriptional activation and oncogenic transformation. Nat Cell Biol 9:932–944
137. Ivaldi MS, Karam CS, Corces VG (2007) Phosphorylation of histone H3 at Ser10 facilitates RNA polymerase II release from promoter-proximal pausing in Drosophila. Genes Dev 21:2818–2831
138. Kagan A, McDonald TV (2005) Dynamic control of hERG/I(Kr) by PKA-mediated interactions with 14-3-3. Novartis Found Symp 266:75–89, discussion 89-99
139. Mateescu B, England P, Halgand F, Yaniv M, Muchardt C (2004) Tethering of HP1 proteins to chromatin is relieved by phosphoacetylation of histone H3. EMBO Rep 5:490–496
140. Hirota T, Lipp JJ, Toh BH, Peters JM (2005) Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. Nature 438:1176–1180
141. Sassone-Corsi P, Mizzen CA, Cheung P, Crosio C, Monaco L, Jacquot S, Hanauer A, Allis CD (1999) Requirement of Rsk-2 for epidermal growth factor-activated phosphorylation of histone H3. Science 285:886–891
142. Soloaga A, Thomson S, Wiggin GR, Rampersaud N, Dyson MH, Hazzalin CA, Mahadevan LC, Arthur JS (2003) MSK2 and MSK1 mediate the mitogen- and stress-induced phosphorylation of histone H3 and HMG-14. EMBO J 22:2788–2797
143. Chwang WB, O'Riordan KJ, Levenson JM, Sweatt JD (2006) ERK/MAPK regulates hippocampal histone phosphorylation following contextual fear conditioning. Learn Mem 13: 322–328
144. Hanauer A, Young ID (2002) Coffin-Lowry syndrome: clinical and molecular features. J Med Genet 39:705–713
145. Trivier E, De Cesare D, Jacquot S, Pannetier S, Zackai E, Young I, Mandel JL, Sassone-Corsi P, Hanauer A (1996) Mutations in the kinase Rsk-2 associated with Coffin–Lowry syndrome. Nature 384:567–570
146. Yntema HG, van den Helm B, Kissing J, van Duijnhoven G, Poppelaars F, Chelly J, Moraine C, Fryns JP, Hamel BC, Heilbronner H et al (1999) A novel ribosomal S6-kinase (RSK4; RPS6KA6) is commonly deleted in patients with complex X-linked mental retardation. Genomics 62:332–343
147. Dufresne SD, Bjorbaek C, El-Haschimi K, Zhao Y, Aschenbach WG, Moller DE, Goodyear LJ (2001) Altered extracellular signal-regulated kinase signaling and glycogen metabolism in skeletal muscle from p90 ribosomal S6 kinase 2 knockout mice. Mol Cell Biol 21:81–87
148. Yang X, Matsuda K, Bialek P, Jacquot S, Masuoka HC, Schinke T, Li L, Brancorsini S, Sassone-Corsi P, Townes TM et al (2004) ATF4 is a substrate of RSK2 and an essential regulator of osteoblast biology; implication for Coffin–Lowry Syndrome. Cell 117:387–398
149. Xing J, Ginty DD, Greenberg ME (1996) Coupling of the RAS-MAPK pathway to gene activation by RSK2, a growth factor-regulated CREB kinase. Science 273:959–963
150. De Cesare D, Jacquot S, Hanauer A, Sassone-Corsi P (1998) Rsk-2 activity is necessary for epidermal growth factor-induced phosphorylation of CREB protein and transcription of c-fos gene. Proc Natl Acad Sci USA 95:12202–12207
151. Merienne K, Pannetier S, Harel-Bellan A, Sassone-Corsi P (2001) Mitogen-regulated RSK2-CBP interaction controls their kinase and acetylase activities. Mol Cell Biol 21:7089–7096
152. Brami-Cherrier K, Valjent E, Herve D, Darragh J, Corvol JC, Pages C, Arthur SJ, Girault JA, Caboche J (2005) Parsing molecular and behavioral effects of cocaine in mitogen- and stress-activated protein kinase-1-deficient mice. J Neurosci 25:11444–11454
153. Stipanovich A, Valjent E, Matamales M, Nishi A, Ahn JH, Maroteaux M, Bertran-Gonzalez J, Brami-Cherrier K, Enslen H, Corbille AG et al (2008) A phosphatase cascade by which rewarding stimuli control nucleosomal response. Nature 453:879–884
154. Bird A (2002) DNA methylation patterns and epigenetic memory. Genes Dev 16:6–21

155. Prendergast GC, Ziff EB (1991) Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. Science 251:186–189
156. Wolffe AP, Matzke MA (1999) Epigenetics: regulation through repression. Science 286:481–486
157. Jorgensen HF, Bird A (2002) MeCP2 and other methyl-CpG binding proteins. Ment Retard Dev Disabil Res Rev 8:87–93
158. Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T (2003) The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. J Biol Chem 278:4035–4040
159. Miranda TB, Jones PA (2007) DNA methylation: the nuts and bolts of repression. J Cell Physiol 213:384–390
160. Ooi SK, Bestor TH (2008) The colorful history of active DNA demethylation. Cell 133:1145–1148
161. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324:930–935
162. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324:929–930
163. Miller CA, Sweatt JD (2007) Covalent modification of DNA regulates memory formation. Neuron 53:857–869
164. Miller CA, Campbell SL, Sweatt JD (2008) DNA methylation and histone acetylation work in concert to regulate memory formation and synaptic plasticity. Neurobiol Learn Mem 89:599–603
165. Chahrour M, Zoghbi HY (2007) The story of Rett syndrome: from clinic to neurobiology. Neuron 56:422–437
166. Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99:247–257
167. Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CM, Gartler SM (1999) The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. Proc Natl Acad Sci USA 96:14412–14417
168. Gimelli G, Varone P, Pezzolo A, Lerone M, Pistoia V (1993) ICF syndrome with variable expression in sibs. J Med Genet 30:429–432
169. Hagleitner MM, Lankester A, Maraschio P, Hulten M, Fryns JP, Schuetz C, Gimelli G, Davies EG, Gennery A, Belohradsky BH et al (2008) Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome). J Med Genet 45:93–99
170. Ehrlich M, Sanchez C, Shao C, Nishiyama R, Kehrl J, Kuick R, Kubota T, Hanash SM (2008) ICF, an immunodeficiency syndrome: DNA methyltransferase 3B involvement, chromosome anomalies, and gene dysregulation. Autoimmunity 41:253–271
171. Jin B, Tao Q, Peng J, Soo HM, Wu W, Ying J, Fields CR, Delmas AL, Liu X, Qiu 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
172. 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
173. Nan X, Campoy FJ, Bird A (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell 88:471–481
174. Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet 19:187–191
175. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393:386–389

176. Kaludov NK, Wolffe AP (2000) MeCP2 driven transcriptional repression in vitro: selectivity for methylated DNA, action at a distance and contacts with the basal transcription machinery. Nucleic Acids Res 28:1921–1928
177. Kokura K, Kaul SC, Wadhwa R, Nomura T, Khan MM, Shinagawa T, Yasukawa T, Colmenares C, Ishii S (2001) The Ski protein family is required for MeCP2-mediated transcriptional repression. J Biol Chem 276:34115–34121
178. Kimura H, Shiota K (2003) Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1. J Biol Chem 278:4806–4812
179. Harikrishnan KN, Chow MZ, Baker EK, Pal S, Bassal S, Brasacchio D, Wang L, Craig JM, Jones PL, Sif S et al (2005) Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. Nat Genet 37:254–264
180. Nan X, Hou J, Maclean A, Nasir J, Lafuente MJ, Shu X, Kriaucionis S, Bird A (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
181. 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
182. Van Esch H, Bauters M, Ignatius J, Jansen M, Raynaud M, Hollanders K, Lugtenberg D, Bienvenu T, Jensen LR, Gecz J 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
183. Chen RZ, Akbarian S, Tudor M, Jaenisch R (2001) Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. Nat Genet 27:327–331
184. Guy J, Hendrich B, Holmes M, Martin JE, Bird A (2001) A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat Genet 27:322–326
185. Shahbazian M, Young J, Yuva-Paylor L, Spencer C, Antalffy B, Noebels J, Armstrong D, Paylor R, Zoghbi H (2002) Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. Neuron 35:243–254
186. Collins AL, Levenson JM, Vilaythong AP, Richman R, Armstrong DL, Noebels JL, David Sweatt J, Zoghbi HY (2004) Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. Hum Mol Genet 13:2679–2689
187. Gemelli T, Berton O, Nelson ED, Perrotti LI, Jaenisch R, Monteggia LM (2006) Postnatal loss of methyl-CpG binding protein 2 in the forebrain is sufficient to mediate behavioral aspects of Rett syndrome in mice. Biol Psychiatry 59:468–476
188. Luikenhuis S, Giacometti E, Beard CF, Jaenisch R (2004) Expression of MeCP2 in postmitotic neurons rescues Rett syndrome in mice. Proc Natl Acad Sci USA 101:6033–6038
189. Guy J, Gan J, Selfridge J, Cobb S, Bird A (2007) Reversal of neurological defects in a mouse model of Rett syndrome. Science 315:1143–1147
190. Giacometti E, Luikenhuis S, Beard C, Jaenisch R (2007) Partial rescue of MeCP2 deficiency by postnatal activation of MeCP2. Proc Natl Acad Sci USA 104:1931–1936
191. Heiman M, Schaefer A, Gong S, Peterson JD, Day M, Ramsey KE, Suarez-Farinas M, Schwarz C, Stephan DA, Surmeier DJ et al (2008) A translational profiling approach for the molecular characterization of CNS cell types. Cell 135:738–748
192. Doyle JP, Dougherty JD, Heiman M, Schmidt EF, Stevens TR, Ma G, Bupp S, Shrestha P, Shah RD, Doughty ML et al (2008) Application of a translational profiling approach for the comparative analysis of CNS cell types. Cell 135:749–762

Histone and DNA Modifications in Mental Retardation

Shigeki Iwase and Yang Shi

Abstract Mental retardation (MR), which affects 1–3% of the total population, refers to a pathological condition whereby the affected individuals suffer from cognitive impairment, which is diagnosed by a low intelligence quotient (IQ) (<70). Over the years, human genetic studies identified a plethora of candidate genes causing MR, but mechanisms by which these candidates regulate cognitive function remain poorly understood. While the functions of MR genes range from cell signaling and gene expression to synaptic plasticity, there is growing evidence supporting a critical role for epigenetic and chromatin regulatory proteins in MR. Excitingly, recent molecular and genetic studies suggest the possibility of improving cognitive functions via modulation of epigenetic regulators, highlighting a potentially new avenue for therapeutic intervention. In this review, we discuss recent studies on epigenetic regulation in MR and explore the concept of epigenetic therapy for MR.

1 Introduction

Mental retardation (MR) is a developmental disability characterized by "significant limitation both in intellectual functioning and in adaptive behavior as expressed in conceptual, social, and practical adaptive skills with onset before the age of 18

S. Iwase
Department of Pathology, Harvard Medical School, 77 Ave Louis Pasteur, Boston, MA 02115, USA
Y. Shi
Department of Pathology, Harvard Medical School, 77 Ave Louis Pasteur, Boston, MA 02115, USA
Division of Newborn Medicine, Department of Medicine, Children's Hospital Boston, 300 Longwood Avenue, Boston, MA 02115, USA
e-mail: yang_shi@hms.harvard.edu

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_8,

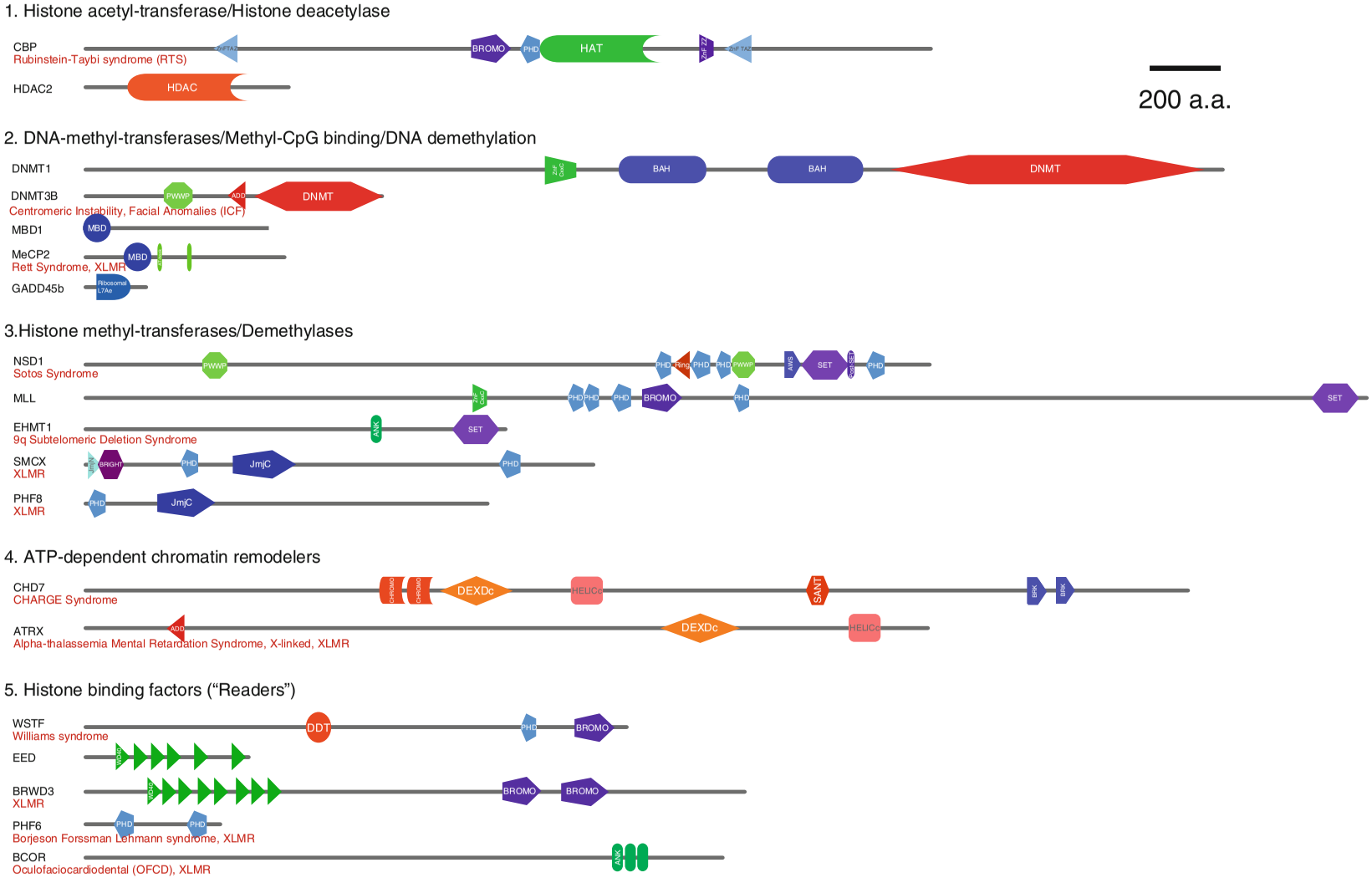
years" [1, 2]. MR is one of the main symptoms for referrals seeking pediatric, neurological, and genetics services and represents 5–10% of health care spending in some developed countries [2]. Causative genetic mutations include both autosomal and X-linked genes (X-linked mental retardation or XLMR). Defects of X-linked genes have long been considered to be important causes of MR, based on the clinical understanding that MR is significantly more common in males than in females [3]. To date, mutations in 82 X-linked genes have been reported to give rise to MR [4]. Among the 82 XLMR genes, 20 of them (24%) encode synaptic proteins, which may play a direct role in synaptogenesis and/or synaptic plasticity, suggesting that abnormal synaptic function leads to deficits in cognitive functions [4]. Importantly, 27 XLMR genes (33%) encode either known or predicted nuclear proteins [4], suggesting an important role for nuclear events including epigenetic regulation in MR.

Epigenetics refers to stable and heritable gene expression changes that do not involve alterations in DNA sequences. Chromatin regulation represents one of the important epigenetic regulatory mechanisms. While the role of DNA methylation in mammalian epigenetic regulation has been well established, the roles of the various covalent modifications on histones are less clear. However, histone methylation has been linked to DNA methylation and therefore is likely to play an epigenetic role, even if indirectly [5–7]. Importantly, recent studies have shown that a number of MR genes encode proteins and enzymes that regulate DNA CpG methylation and histone methylation [5], supporting the emerging relationship of epigenetic regulation and MR (Fig. 1).

In this review, we discuss the roles and mechanisms of action of MR genes that regulate chromatin via covalent modifications on DNA and histones, as well as noncovalent modifications of nucleosomes through ATP-dependent nucleosome remodelers. Section 2 focuses on factors involved in regulation of the core of the epigenetic mechanism, i.e., DNA methylation, including DNA methyl transferases and methyl CpG binding proteins. In Sect. 3, we discuss how covalent modifications on histones, including acetylation and methylation, contribute to cognitive functions. Sect. 4 is devoted to discussions of the involvement of ATP-dependent remodeling proteins play a role in MR, possibly by regulating the higher-order structures of chromatin in neurons. Utilizing the conceptual framework of epigenetic regulation in MR laid out in the preceding sections, we explore potential strategies to ameliorate MR in Sect. 5.

2 DNA CpG Methylation in MR

In mammalian cells, CpG pairs can be covalently modified by the addition of a methyl group to the C5 position of the cytosine ring by a number of DNA methyltransferases [6–8]. Methyl CpGs in general coincide with repressive genomic loci, including the centromeric/pericentrometic heterochromatin, the inactive X-chromosome, the silenced alleles of the imprinted genes, transposable elements as well as silenced promoters. Interestingly, the level of methylated cytidine is


1. Histone acetyl-transferase/Histone deacetylase
CBP
Rubinstein-Taybi syndrome (RTS)
HDAC2
ZnF TAZ
BROMO
PHD
HAT
ZnF ZZ
ZnF TAZ
HDAC
200 a.a.
2. DNA-methyl-transferases/Methyl-CpG binding/DNA demethylation
DNMT1
DNMT3B
Centromeric Instability, Facial Anomalies (ICF)
MBD1
MeCP2
Rett Syndrome, XLMR
GADD45b
ZnF CxxC
BAH
BAH
DNMT
PWWP
ADD
DNMT
MBD
MBD
Ribosomal L7Ae
3.Histone methyl-transferases/Demethylases
NSD1
Sotos Syndrome
MLL
EHMT1
9q Subtelomeric Deletion Syndrome
SMCX
XLMR
PHF8
XLMR
PWWP
PHD
Ring
PHD
PHD
PWWP
AWS
SET
SRI/SSH
PHD
ZnF CxxC
PHD
PHD
PHD
BROMO
PHD
SET
ANK
SET
BRIGHT
PHD
JmjC
PHD
PHD
JmjC
4. ATP-dependent chromatin remodelers
CHD7
CHARGE Syndrome
ATRX
Alpha-thalassemia Mental Retardation Syndrome, X-linked, XLMR
CHROMO
CHROMO
DEXDc
HELICc
SANT
BRK
BRK
ADD
DEXDc
HELICc
5. Histone binding factors (“Readers”)
WSTF
Williams syndrome
EED
BRWD3
XLMR
PHF6
Borjeson Forssman Lehmann syndrome, XLMR
BCOR
Oculofaciocardiodental (OFCD), XLMR
DDT
PHD
BROMO
WD40
BROMO
BROMO
PHD
PHD
ANK

higher in the brain than in other tissues, suggesting its potential importance in neuronal function [9–11]. The methylated CpG is recognized by a subset of proteins via a conserved sequence motif termed MBD (Methyl CpG Binding Domain, Fig. 1) [12, 13]. Importantly, five methyl-CpG-related factors have been implicated in MR and CNS function (Fig. 1). One of the best-understood MR genes is the methyl-CpG binder, MeCP2, whose mutations are considered to be the primary cause of the most common female MR, Rett syndrome [14], which is discussed below.

2.1 The Roles of DNA Methyltransferases in MR

Notably, two human congenital diseases accompanied with MR are caused by mutations of DNA methyltransferases. The de novo DNA methyltransferase

Fig. 1 Schematic representation of domain architecture of 19 MR-implicated nuclear factors. The associated neurological disorders are denoted in *red*. *XLMR* X-linked Mental Retardation. (1) Histone acetyltransferase/Histone deacetylase. CBP is responsible for Rubinstein–Taybi syndrome (RTS) [66]. CBP possesses a HAT domain (*green tube*) which governs histone acetylation. CBP also carries two types of zinc fingers: ZnF TAZ (*purple trapezoid*) and ZnF ZZ (*light blue triangle*), a bromodomain (*purple pentagon*), and a PHD finger (*light blue pentagon*). HDAC2 has an HDAC domain responsible for enzymatic activity. (2) DNA-methyl-transferases/Methyl-CpG binding/DNA demethylation. DNMT1 and DNMT3b carry the DNMT domain (*red hexagon*) catalyzing DNA methylation. MBD1 and MeCP2 carry a MBD domain (*blue oval*), which binds methylated CpG. DNMT3B and MeCP2 are causative for ICF syndrome [15] and Rett syndrome [14], respectively. DNMT1 also carries a CxxC-type ZnF (*green trapezoid*) and 2 BAH domain (*blue tubes*). DNMT3B harbors a PWWP domain (*green octagon*) and an ADD domain (*red triangle*). AT-hook in MeCP2 is represented by a *thin green oval*. GADD45b has a conserved, Ribosomal L7 Ae motif (*truncated blue oval*). (3) Histone methyl-transferases/demethylases. NSD1 and EHMT1 are implicated in the Sotos syndrome [87] or 9q subtelomeric deletion syndrome [97], respectively. SMCX [159] and PHF8 [160] are both responsible for XLMR and catalyze histone demethylation via their JmjC domain (*deep blue pentagon*). NSD1, MLL, and EHMT1 are all histone methyl transferases catalyzing methylation reactions via their SET domain (*purple hexagon*). All five proteins carry PHD fingers (*light blue pentagon*). NSD1 has two PWWPs (*green octagon*), one ring finger (*red triangle*), and an AWS domain (*truncated blue hexagon*). Ankyrin repeats in EHMT1 are denoted by *green oval*. The potential DNA binding motif, BRIGHT in SMCX, is represented by *purple hexagon*. (4) ATP-dependent chromatin remodelers. Both CHD7 and ATRX carry ATP-dependent remodeling motifs, DEXDc (*orange diamond*) and HELICc (*pink rectangle*). CHD7 mutations give rise to CHARGE syndrome [155], and ATRX is mutated in Alpha-thalassemia mental retardation syndrome, X-linked [152]. CHD7 carries a double chromodomain (*skewed orange rectangles*), a potential histone binding domain SANT (*orange hexagon*), and two BRK motifs (*truncated blue hexagon*). ATRX has a potential histone binding motif ADD in its N-terminus (*red triangle*). (5) Histone binding factors ("readers"). WSTF is deleted in the Williams syndrome [117]. Three conserved motifs are found, DDT (*orange oval*), PHD finger (*light blue pentagon*), and bromodomain (*purple pentagon*). EED and BRWD3 have potential histone binding repeats WD40 (*green triangle*). BRWD3 is implicated in XLMR [128]. PHF6 and BCOR are causative for the Börjeson–Forssman–Lehmann syndrome [161] and Oculofaciocardiodental syndrome [130], respectively. Ankyrin repeats are found in the BCOR C-terminus (*green oval*)

DNMT3B has been found to be mutated in immunodeficiency, centromeric instability, and facial anomalies (ICF) syndrome [15]. ICF syndrome individuals show immunodeficiency, including the absence or a severe reduction of the immunoglobulin and a reduced number of T cells, making these individuals prone to infection and death before adulthood [16]. Centromeric instability is believed to be attributable to CpG demethylation of the satellite DNA adjacent to the centromeric regions of some chromosomes [16]. Most individuals also exhibit growth and MR and facial dysmorphism [16]. Global gene expression analysis using patient lymphocytes identified misregulation of a number of neurogenesis/neural function-related genes, implying that DNMT3B might serve as a transcriptional regulator in neurons via its de novo methylation activity [17]. Since a majority of the studies on DNMT3B have been focused on its mechanism of CpG methylation and pathogenesis of immunodeficiency, the function of DNMT3B in the nervous system remains to be fully elucidated.

In contrast, the maintenance DNA methyltransferase, DNMT1, has been implicated both in chromosomal integrity during neurogeneration and in neuronal survival [18]. Conditional depletion of DNMT1 in the CNS precursor cells results in a reduced survival of postnatal differentiating/differentiated neurons, without detectable proliferation defects in the precursor cells. DNMT1-deficient precursors showed hypomethylation at the centromeric repeat sequences and retroviral DNA Intra-cisternal A particles (IAPs) [18]. Since a conditional knockout of the Dnmt1 gene in postmitotic neurons had no impact on neuronal survival, it seems likely that hypomethylation of genomic DNA in the precursors may somehow have affected neuronal survival after the neurons exit from the cell cycle [18]. Given that maintenance methylation by DNMT1 needs de novo methylation by DNMT3A/3B to begin with, and that DNMT3B deficiency causes centromeric abnormality and the ICF syndrome, it is conceivable that DNMT3B and DNMT1 may collaborate to maintain centromere integrity in the dividing neuroprecursors (Fig. 2a). Thus, one of the important roles of CpG methylation is protecting genomic integrity in dividing neuroprecursors, and a compromise in such a mechanism impacts the survival of postmitotic neurons.

2.2 Reversibility of CpG Methylation

The removal of methyl groups from DNA can be accomplished by either a passive or active process. Passive demethylation occurs during DNA replication. However, increasing evidence points to the importance of active DNA demethylation in several cellular processes during development, such as rapid demethylation of the male pronucleus in the mammalian zygote. Several possible mechanisms have been proposed to account for active demethylation, which include excision of the methylated nucleotides followed by DNA repair (reviewed in [19]). Recent studies also identified enzymes that convert methyl C to hydroxymethyl C [20, 21], revealing yet another potential mechanism for the regulation of DNA methylation

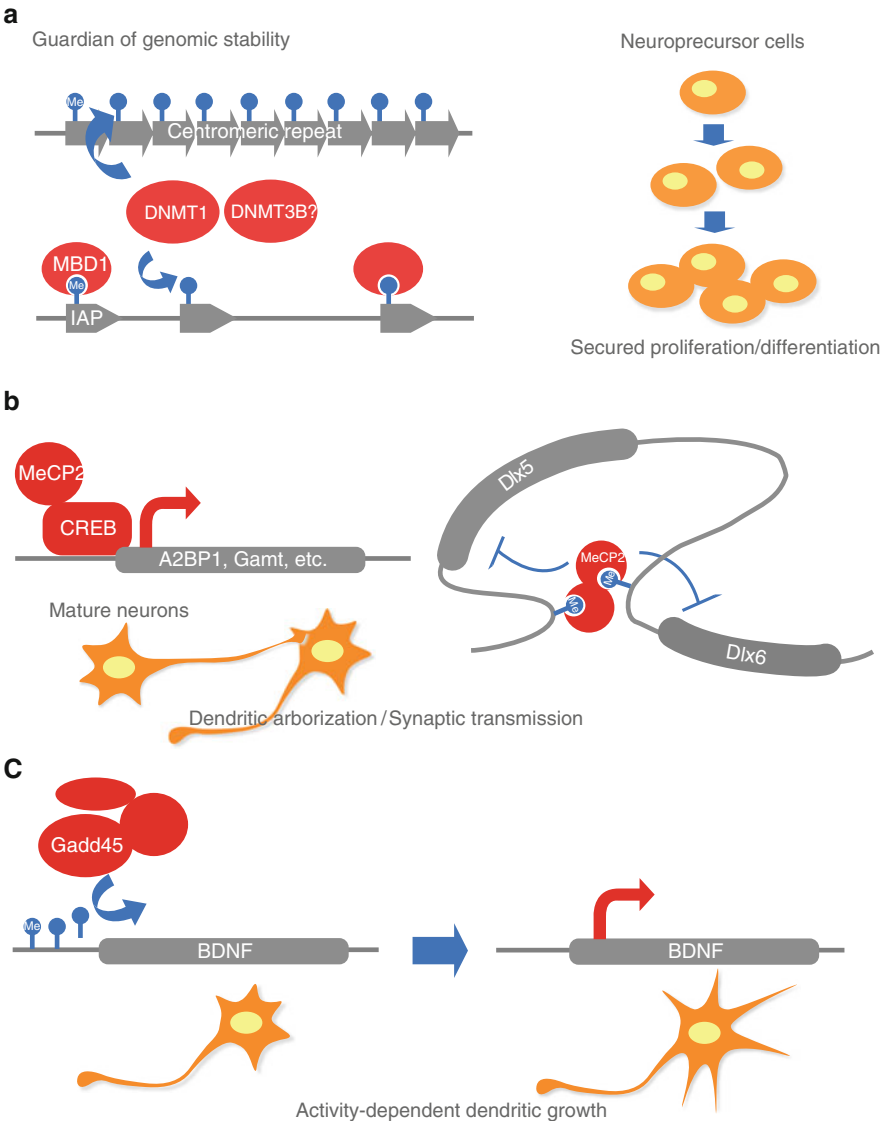


Fig. 2 CpG methylation in MR-related neuronal functions. (**a**) Proposed function of DNMT1/3B and MBD in neurogeneration and differentiation. Centromeric DNA (*upper left*) repeats and retroviral IAP (*lower left*) genes are methylated by DNMT1. The de novo methylase DNMT3B may also be involved in this process. MBD1 suppresses IAP gene expression (*lower left*). CpG methylation-mediated silencing of centromeric repeats and viral genes protect genomic integrity, which in turn facilitates faithful proliferation and differentiation of neuroprecursor cells (*right*). (**b**) Multiple facets of MeCP2 action. MeCP2 activates MR and autism-related genes, such as A2BP1 and gamt, in concert with CREB (*left*). MeCP2 regulates genomic imprinting at the Dlx5/6 locus likely through the recognition of CpG methylation in the intergenic regions and forming repressive chromatin loop (*right*). MeCP2 plays key roles in the refinement of connectivity

in vivo. However, it remains unclear whether bona fide DNA demethylases exist to directly remove the methyl groups without involving base or nucleotide excision and repair.

Nonetheless, active DNA demethylation in postmitotic neuron has been observed in several studies [22, 23], and dynamic regulation of DNA methylation has been shown to be important for learning and memory [24]. For instance, forced membrane depolarization of cultured cortical neurons by potassium chloride (KCl) induces transcription of neuronal synaptic plasticity-related genes, such as BDNF [25, 26]. Importantly, a membrane depolarization-responsive promoter of the BDNF gene has been found to be demethylated upon KCl treatment [22, 23]. In addition to BDNF, the promoter of Reelin, a positive regulator of memory formation, is demethylated in the rat hippocampus during the learning process in the fear-conditioning test [24]. Finally, it was reported that Gadd45, an integral component of the deamination coupled G/T mismatch repair-mediated demethylation [27], promotes demethylation of plasticity-related gene promoters including BDNF, thereby activating their transcription and inducing activity-dependent dendritic growth [28]. These observations suggest that reversal of DNA methylation is engaged in synaptic plasticity and in turn long-term memory formation, which requires de novo synthesis of mRNA/protein [22–24] (Fig. 2c). Taken together, balancing the level of DNA methylation *in vivo* is important for a variety of important biological processes.

2.3 *Methyl CpG "Readers"*

Rett Syndrome is caused by mutations of the MeCP2 gene, which encodes a methyl CpG binding protein [14]. Rett Syndrome patients appear to develop normally from birth until about 6–18 months of age, and then start showing regression in speech and purposeful hand movements. These patients also exhibit severe MR, autistic features, ataxia, apraxia, and hyperventilation [29]. Neuropathological studies suggested that loss of MeCP2 causes disruption of critical postnatal synaptic refinements but not gross change of brain architecture [30, 31]. Mecp2 null male mice display no phenotype early in development, but by 3–8 weeks of age, they develop a stiff, uncoordinated gait, reduced spontaneous activity, irregular breathing with most animals dying at 10 weeks of age [32, 33]. Mecp2 heterozygous females start exhibiting irregular breathing around 9 months old [32]. These phenotypes are reminiscent of Rett syndrome symptoms. MeCP2 expression in

Fig. 2 (Continued) maturating/mature neuronal circuitry (*lower left*). (**c**) Active DNA demethylation of BDNF gene promoter. Stimulation of neurons evokes transcription of a subset of genes involved in dendritic growth and synaptic plasticity such as BDNF (*top*). In this process, a deamination coupled G/T mismatch repair-mediated demethylation complex including Gadd45 protein, remove methyl group from DNA (*top*). Gadd45 is required for the activity-dependent dendritic growth (*bottom*)

neurons increases progressively with maturation throughout the CNS [34–36], suggesting a role in maturating neurons. In agreement with this observation, loss of MeCP2 causes an accumulation of neurons stalled at a transitional stage, judged by the expression of proteins associated with both immature and mature neurons, suggesting delayed maturation [37–39]. In addition to a role in terminal differentiation, MeCP2 also has well documented synaptic functions. Mecp2-deficient mice show impaired dendritic outgrowth [33, 34], basal synaptic transmission [40–42], balance between excitatory and inhibitory circuit [43], and hippocampal- and amygdalar-mediated learning [41, 44]. Taken together, MeCP2 plays important roles in the refinement of connectivity maturating/mature neuronal circuitry (Fig. 2b). Importantly, postnatal expression of wild-type Mecp2 in MeCP2-deficient mice restores many symptoms, including longevity, irregular breathing, and synaptic transmission, suggesting that the damage caused by the loss of MeCP2 is not irreversible. Given the important roles of MeCP2 in neuronal wiring, amelioration of Rett-like symptoms revealed surprising plasticity of postnatal neuronal connectivity. Moreover, these data suggest that human Rett-syndrome individuals might not have irrevocable alterations in their neurons, raising the hope for therapeutic intervention through restoration of MeCP2 functioning [45, 46].

How does MeCP2 carry out the biological functions discussed above? Based on its ability to bind methylated CpG and to recruit HDAC-containing corepressors, it was speculated that MeCP2 acts as a global transcriptional silencer [47]. However, recent global gene expression analyses uncovered an unexpected feature of Mecp2, which is an ability to act as a transcriptional activator, as well as a repressor [48]. Specifically, MeCP2 functions as a transcriptional activator through its physical interaction with the sequence-specific DNA-binding transcription factor CREB [48]. Genes that are activated by MeCP2 include A2BP1 and Gamt, which have been implicated in MR and autism, and may therefore account for some facets of Rett phenotypes [48]. Interestingly, the genes directly activated by MeCP2 lack methylated CpG in their promoters, suggesting that the activation function is independent of CpG methylation [48] (Fig. 2b, left). Its interaction with CREB suggests that MeCP2 is recruited by CREB to gene promoters for activation. However, at present it is unclear how MeCP2 contributes to transcriptional activation.

Mecp2 appears to regulate genomic imprinting as well. MeCP2 has been shown to occupy an imprinted gene cluster harboring Dlx5 and Dlx6 at mouse chromosome 6, and loss of Mecp2 resulted in biallelic expression of both genes, possibly due to disruption of specific three-dimensional chromatin loops [49]. It is noteworthy that the Dlx genes are implicated in GABA synthesis whose misregulation can cause the Angelman Syndrome, which is a Rett-like neurodevelopmental disorder [50]. Genome-wide MeCP2 localization analysis utilizing ChIP-Chip revealed that 59.4% of MeCP2 binding sites were intergenic, and of these, 58.4% were >10 kb away from the transcriptional start or transcriptional termination sites [51]. The imprinted gene regulation and this periodic occupation of intergenic regions suggest that organization of chromatin loops is perhaps one of the major mechanisms of MeCP2 function.

A more recent study supports the notion that MeCP2 does not function as a gene-specific regulator but rather globally suppresses transcriptional noise, such as transcription of repetitive sequences, in a methylation-dependent manner [52]. In sum, recent evidences have argued that the classical role of MeCP2, CpG methylation-dependent, gene-specific transcriptional repressor may not be how it regulates biological processes important for brain function. Since there appears to be some correlation between CpG methylation and MeCP2 occupancy in the intergenic regions [51], effects on long-range chromosomal organization, such as the formation of chromatin loops, may be the most relevant hypothesis for how CpG methylation contributes to MR (Fig. 2b, right).

In contrast to MeCP2 function in mature neurons, MBD1-deficient mice have lower neuron density and exhibited reduced survival of the adult neural stem cells (ANCs) during neuronal differentiation in the hippocampus [53]. MBD1-deficient ANCs have higher probabilities of aneuploidy, mainly gaining of chromosome numbers, and de-repression of the IAP retroviral genes. Furthermore, LTP in hippocampus of MBD1 KO mice was significantly attenuated, and mice showed leaning deficits in the Morris water maze trials [53]. Thus, DNMT1 and MBD1 suggest a possible role of CpG methylation as a guardian of genomic stability in postmitotic neuronal survival and cognitive function (Fig. 2a), but the exact mechanism still remains to be understood.

Taken together, it appears that DNA methylation regulators exert two major roles in CNS. One is protection of genome stability of neuroprecursors in turn assuring neuronal survival in development. Another role is gene regulation in mature neurons represented by MeCP2 and Gadd45. With respect to the gene regulatory role, it is noteworthy that DNA methylation is not irreversible in neurons. This dynamic nature CpG methylation may be the key to achieve synaptic plasticity. We anticipate that continued investigation of dynamic regulation of CpG methylation will provide significant new insights into molecular mechanisms underlying learning and memory, and their connections with MR.

3 Covalent Histone Modifications and MR

3.1 Histone Acetyl Transferases/Histone Deacetylase

Acetylation is one of the first histone posttranslational modifications to be discovered to impact gene transcription [54, 55]. Acetylation of the histone tails of all four core histones is generally correlated with open chromatin and active transcription [56], which is consistent with the molecular function ascribed to the majority of histone acetyltransferases (HATs), i.e., transcriptional coactivators. Various mechanisms can account for the effect of acetylation on chromatin structure and gene expression. Acetylation of lysine residues leads to a reduction of positive charge, which can weaken histone-DNA [54] or nucleosome–nucleosome

interaction, but might also alter specific contacts between nucleosomes and regulatory proteins [57–59]. Any such changes can lead to changes in higher-order chromatin folding. Importantly, the histone acetyltransferase CBP (CREB-binding protein) [60] has been genetically defined as an MR gene, whose mutations are correlated with the Rubinstein–Taybi Syndrome (RTS).

CBP was originally identified as a coactivator of CREB (cAMP responsive element binding protein) [61]. Importantly, CREB had been characterized as a key player for long-term memory [62], which is one of the cognitive deficits observed in MR patients. Several decades ago, neuroscientists became aware of the fact that unlike short-term memory, long-term memory requires synthesis of new mRNAs and proteins [63]. The behavioral and anatomical studies of the sea slug *Aplysia* suggested that the formation of long-term memory and learning results from changes in the strength or effectiveness of preexisting interconnected cells [64]. This flexible nature of synaptic strength is termed synaptic plasticity and is now proven to be conserved among species. In this context, CREB induces transcription of genes necessary for long-term memory formation, in response to cAMP signaling caused by neuronal stimulation [62, 64], and this in turn increases the efficacy of synapses [65].

Consistent with a cofactor role of CBP for CREB, a genetic study identified heterozygous deletions of the CBP gene in the human cognitive disorder RTS. RTS patients are characterized by growth and psychomotor development delay, skeletal abnormality, and severe MR [66]. Genetic manipulations have generated several RTS model mice. Heterozygous deletion of the CBP gene results in long-term memory deficits, assessed by behavioral tests, as well as defective long-term potentiation in the late phase (L-LTP) in the hippocampus through electrophysiological studies [67]. Analysis of this mouse model supports the idea that haploinsufficiency of the CBP gene causes cognitive deficits in the RTS patients [67]. Another mouse model harboring an inducible expression system of a catalytically inactive CBP mutant revealed the importance of HAT activity for synaptic plasticity *in vivo* [68]. These authors suggested the idea that local changes of chromatin structure by histone acetylation allow a prolonged elevation of CREB target gene transcription, although CREB activation by neuronal stimulation is transient in nature [68]. These studies support a role of epigenetic regulation in cognitive functions including learning and memory.

Histone acetylation is dynamically regulated by both HATs and histone deacetylases (HDACs). HDACs are commonly found in corepressor complexes [69] that antagonize the transcriptional activation function of HATs [70–74]. Using the RTS mouse model discussed above, investigators found that administration of HDAC inhibitors ameliorated long-term memory formation deficits in the heterozygous CBP mutant [67]. Similarly, an earlier study of *Aplysia* showed that inhibition of HDACs by the general HDAC inhibitor Trichostatin A (TSA) facilitated long-term memory formation by blocking long-term depression, a manifestation of negative synaptic plasticity [75]. Recently, a mouse genetic study unveiled negative regulation of synaptic plasticity by HDAC2 [76]. Specifically, over-expression of HDAC2 in neurons led to impaired memory formation, while targeted disruption of the HDAC2 gene, which was correlated with increased H3 and H4 acetylation,

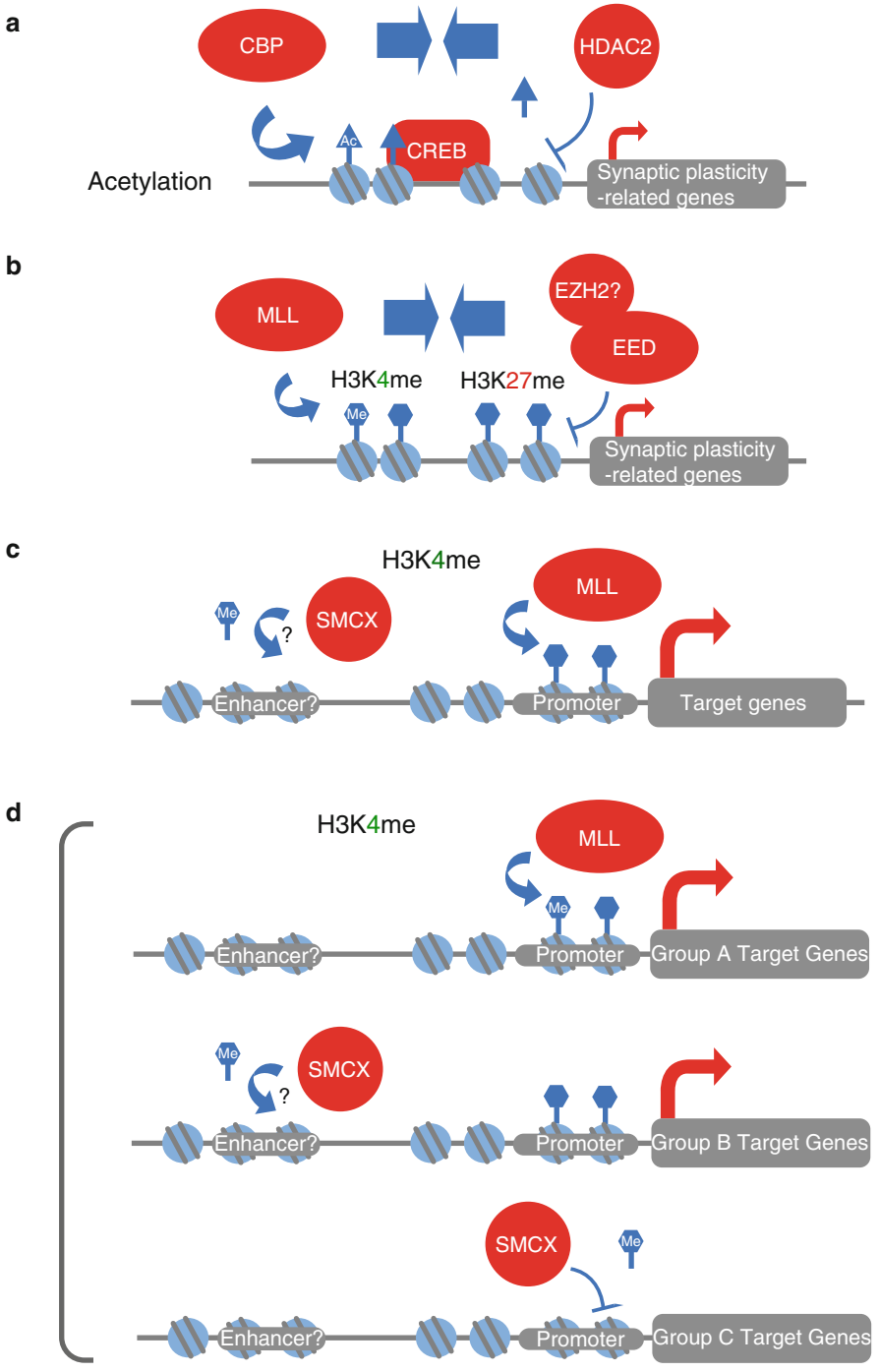
facilitated memory formation [76]. Consistently, HDAC2 has been shown to bind regulatory regions of memory-related genes in the brain, including some CREB targets. These results provide a conceptual framework that memory formation is under the regulation of opposing epigenetic forces, such as acetylation and deacetylation, and suggest that a proper balance between the actions of the opposing epigenetic regulators is critical for memory formation (Fig. 3). This raises the possibility that restoring the altered epigenetic balance by modulating the opposing regulators (by small molecules, for instance) may provide therapeutic benefits.

3.2 Histone Lysine Methyltransferases

In addition to acetylation, histones are methylated on both lysine (K) and arginine (R) residues. Six lysine residues, K4, K9, K27, K36 and K79 on histone H3, and K20 on histone H4, are known to be important for gene regulation, heterochromatin formation, and DNA damage response [77, 78]. Importantly, the same lysine residue can be differentially methylated to mono-, di-, or tri-methylated states and increasing evidence suggests that the different degrees of methylation on the same lysine residue may also play differential roles [79–81]. Given that histone methylation has a slower turnover rate than acetylation, methylation is more likely to be involved in longer-term regulatory events even with respect to neuronal function.

While histone methylation is catalyzed by histone methyltransferases (HMTs) [82], removal of the methyl group is carried out by two histone demethylase (HDMs) families of proteins including LSD1/LSD2 as well as multiple members of the JmjC family [83–86]. Importantly, a number of HMTs and HDMs have been implicated in MR and neuronal functions. For instance, haploinsufficiency of the HMT NSD1 has been implicated in the Sotos Syndrome, a neurological disorder characterized by overgrowth from the prenatal stage through childhood, advanced bone age, large skull, acromegalic features, occasional brain anomalies and seizures, and MR [87, 88]. There are no defined mouse models as yet for Sotos Syndrome, since homozygous deletion of NSD1 in mice caused gastrulation failure and a high incidence of apoptosis [89], and no remarkable morphological abnormalities were manifest in heterozygous mice [89]. However, more careful analysis may yet reveal neurological disorders in heterozygous mice, given that Sotos Syndrome is semidominant in man.

There are conflicting data with respect to NSD1 substrates, but a recent study demonstrated that NSD1 predominantly methylates H3K36 to the dimethyl state (H3K36me2) *in vitro* when the more physiologically relevant, nucleosomal substrates were presented for the methylation reaction [90]. A role for H3K36 methylation is the suppression of inappropriate initiation from cryptic start sites within the coding region [91–93]. Although the exact function of NSD1 in the Sotos Syndrome remains unclear, the demonstration that NSD1 is a HMT that methylates H3K36 suggests that it may play a positive role in gene expression by suppressing spurious gene transcription. It may be particularly interesting to investigate NSD1


a
CBP
HDAC2
Ac
CREB
Acetylation
Synaptic plasticity
-related genes
b
MLL
EZH2?
EED
H3K4me
H3K27me
Me
Synaptic plasticity
-related genes
c
H3K4me
MLL
Me
SMCX
?
Enhancer?
Promoter
Target genes
d
H3K4me
MLL
Me
Enhancer?
Promoter
Group A Target Genes
Me
SMCX
?
Enhancer?
Promoter
Group B Target Genes
SMCX
Me
Enhancer?
Promoter
Group C Target Genes

in the context of retinoic acid signaling (where NSD1 was initially identified [94]), since retinoic acid appears to facilitate neurogeneration in both embryonic CNS development and adult brain plasticity [95, 96].

In addition to NSD1, the euchromatic histone-lysine *N*-methyltransferase 1 (EHMT1; also known as GLP[97]), which mediates H3K9 dimethylation [98, 99], is implicated in the chromosome 9q subtelomere deletion syndrome (9qSTDS). Haploinsufficiency of GLP causes patients to suffer from severe hypotonia, speech and motor delay, a compromised facial feature, epilepsy, and behavioral disturbances including antisocial/autistic behavior and aggressive outbursts [97, 100]. Targeted disruption of Ehmt1 in mice causes embryonic lethality, indicating a crucial role in early development [99]. Recent behavioral analysis using the heterozygous Ehmt1/Glp mice recapitulated hypoactivity and the autistic-like features of 9qSTDS [101]. However, similar to NSD1, little is known about the cellular and molecular functions of EHMT1/GLP relevant to the neurological abnormality observed in these patients. Importantly, HMT1/GLP is primarily responsible for generating H3K9me2 in euchromatin [98, 99]. Thus, EHMT1/GLP is likely to be involved in gene regulatory mechanisms rather than maintenance of chromosome integrity via regulation of the centromeric/telomeric heterochromatin. Consistently, EHMT1/GLP has been implicated in the transcriptional repression mediated by E2F6 [98], although the neuronal functions of EHMT1/GLP and the underlying molecular mechanisms remain to be elucidated.

The third HMT implicated in cognitive function is MLL, which catalyzes H3K4me0 to H3K4me3, a mark that is enriched at the transcription start sites (TSS) of active (or poised) genes[102]. Recent studies showed that hippocampal LTP is significantly attenuated in the Mll heterozygous mice, suggesting a positive role for MLL and H3K4 methylation in memory formation [103]. In contrast, heterozygous deficiency of Eed, which encodes a component of the PRC2 complex that mediates H3K27 trimethylation [104], resulted in a robust enhancement of LTP [103]. Interestingly, mice bearing heterozygous deletions for both Mll and Eed

Fig. 3 Balancing epigenetic forces in cognitive function. (**a**) Balancing histone acetylation. CBP acetylates histones and activates synaptic genes, which facilitate long-term potentiation (LTP). HDAC2 negatively regulates learning and memory. HDAC2 occupies CREB-regulated gene promoter in mouse brain and represses their transcription via histone deacetylation. (**b**) Balancing in histone methylation. MLL is a histone H3K4 methyltransferase that positively regulates LTP. Eed, an integral component of the Polycomb complex, negatively regulates LTP and transcription. The polycomb complex contains the repressive H3K27 methyltransferase EZH2, thus EZH2 and Eed may work in concert. Given the opposing phenotype of MLL and Eed-deficiency in LTP, they may counteract on the regulatory regions of the same target genes. (**c**) Loss of function of SMCX results in MR, suggesting that SMCX is a positive regulator of cognitive function. SMCX and MLL may cooperatively activate common target genes at different regulatory regions such as enhancers and promoters. (**d**) Alternatively, MLL and SMCX may regulate different sets of genes. While MLL likely methylates H3K4 at promoters to activate a set of genes (group A target genes, on *top*), SMCX may demethylate H3K4 at enhancers and thereby activate another set of genes (group B target genes, in *middle*). When SMCX works on promoter, the enzymatic activity is predicted to negatively regulate target gene transcription (group C target genes, *bottom*)

showed normal LTP [103], suggesting an antagonistic relationship between H3K4me3 and H3K27me3 in the regulation of memory formation, reminiscent of the relationship between CBP and HDAC discussed above. Detailed behavioral studies of the Mll and Eed mutant mice will provide novel insights into how a proper balance between K4 and K27 methylation on histone H3 contributes to cognitive function (Fig. 3).

3.3 Histone Lysine Demethylases

Similar to the other posttranslational modifications, histone methylation is also dynamically regulated by both HMTs and HDMs [84–86]. Recent studies showed that two XLMR genes, SMCX/JARID1C [105, 106] and PHF8 [107], encode HDMs. Importantly, patient mutations significantly compromise the demethylase activities of SMCX and PHF8, thus linking histone methylation regulation to MR. XLMR-associated point mutations in SMCX/JARID1C compromise nuclear localization, chromatin binding, and enzymatic activity, suggesting multiple effects of these mutations on the functions of SMCX [105, 106]. Loss of SMCX results in increased apoptosis in the zebrafish brain during development and significantly compromised dendritic growth in the rat cerebellar granule neurons, respectively. These findings suggest that SMCX plays both a developmental role in regulating neuronal survival as well as in postmitotic neurons where it is essential for dendritic development [106]. The role of SMCX in dendritic development is particularly intriguing and warrants further investigation given that dendritic spine development is believed to be important for cognitive function and therefore highly relevant to MR [108, 109].

What might be the molecular mechanism that underlies SMCX's involvement in cognitive function regulation? A recent study shows that SMCX represses neuron-specific gene repression in nonneuronal cells, in concert with REST/NRSF, a zinc-finger containing DNA binding factor [105]. REST is mainly expressed in nonneuronal cells [110, 111], but is also found at a low level in neurons. It remains to be investigated whether SMCX also works with REST in neurons to repress gene expression and how this may impact cognitive function. More recently, a second XLMR gene, PHF8, has been shown to also encode a demethylase [107] (Qi and Shi unpublished). Missense mutations associated with XLMR abolish catalytic activity, suggesting the importance of catalysis in neuronal development/functioning. However, the cellular/molecular functions of PHF8 are largely unknown.

In summary, both histone acetyltransferases/deacetylases and HMTs/HDMs have been shown to play a role in regulating brain function, but the underlying mechanisms are far from clear. In some cases, the biochemically opposing enzymes (such as CBP and HDAC2) also have opposite biological functions, suggesting that balancing histone acetylation is important for brain function. In other cases, defects in the antagonistic methyltransferases and demethylases (such as the H3K4 methylase MLL vs. the H3K4 demethylase SMCX) both lead to malfunction of the brain. In

the latter case, it is possible that MLL and SMCX collaborate to regulate cognitive functions via regulation of either the same sets of target genes when they are positioned at different regulatory regions (Fig. 3c), such as enhancers vs. promoters, or distinct sets of genes (Fig. 3d). Thus, determination of direct target genes by genome-wide localization study and gene expression analysis using the relevant cell types, including hippocampal and cortical neurons, will be necessary to understand the molecular etiology of XLMR caused by HMTase/HDMase mutations.

3.4 Histone Binding Factors ("Readers")

Modified histone tails are proposed to serve as recruitment signals for various molecular and biochemical activities (such as ATP-dependent remodeling proteins discussed below). Consistent with this idea, a large number of proteins have been identified to carry distinct protein modules dedicated to the recognitions of various histone modifications ("reader modules") [112, 113]. Interestingly, multiple MR genes appear to encode proteins carrying those reader modules, indicating the importance of the ability to recognize histone modifications in human diseases. For instance, the XLMR gene products SMCX and PHF8 both contain PHD domains (plant homeo domain), which play a role in recognizing H3K9me and H3K4me, respectively [106, 114]. The ATRX protein has the ADD domain, which also serves to bind histone H3 tail (Iwase and Shi unpublished). It is believed that histone binding via these modules is important for the proper genomic localization of the corresponding enzymes [115, 116]. Other potential histone modification readers implicated in MR are discussed below.

WSTF is one of several candidate genes for the Williams syndrome, which is a complex developmental disorder including vascular and heart malfunction, dysmorphic facial features, growth, and MR [117, 118]. The WSTF protein is found in at least two functionally distinct multiprotein complexes, WINAC [119] and WICH [120], both of which contain ATP-dependent chromatin remodeling factors, Brg1/Brm and SNF2h, respectively. However, WSTF itself does not have remodeling activity, but instead functions as both a histone tyrosine kinase and a reader of specific histone modifications [119–121].

WSTF carries two potential reader modules, a PHD finger and a bromodomain, which has an intrinsic affinity for acetylated lysines [122]. The WSTF bromodomain seems to have an important role in VDR (Vitamin D receptor)-mediated transactivation through its association with acetylated histones [123]. Although many PHD fingers have been shown to recognize methylated or unmethylated histone tails [124–127], the substrate for the WSTF PHD finger is unknown. Another bromodomain-containing XLMR gene is BRWD3, which was found to be mutated in two XLMR families where patients show macrocephaly, noted in infancy, and prominent ears [128]. Notably, both mutations cause frame-shift and result in the loss of the bromodomain [128]. However, once again the molecular and cellular roles of BRWD3 are unknown and await analysis.

A final case is the Börjeson–Forssman–Lehmann syndrome (BFLS), which is characterized by moderate to severe MR, epilepsy, hypogonadism, hypometabolism, obesity with marked gynecomastia, swelling of subcutaneous tissue of the face, narrow palpebral fissure, and large but not deformed ears [129]. Mutations in *PHF6*, which encodes a protein carrying two tandemly arranged PHD fingers, are considered causal for BFLS [129]. Consistent with the role for PHF6 in MR, Mouse *Phf6* was most highly expressed in the embryonic central nervous system (CNS) and at lower levels in other tissues. However, only low levels of *Phf6* transcripts were detected in the adult brain, suggesting that it is involved in neurogenesis [129]. In one patient, a missense mutation is located in the first PHD finger, suggesting a role for this putative histone-binding module, yet the substrates for either PHD finger are not known.

Finally, dominant mutations of BCOR (BLC6 corepressor) in females have been found in the Oculofaciocardiodental (OFCD) syndrome, a congenital disorder involving cataracts, microphthalmia, cardiac, dental and digital anomalies, and MR [130]. BCOR is a transcriptional corepressor originally identified by its ability to interact with the transcriptional repressor BCL6, which plays critical roles in specific immunological processes [131]. Inhibition of BCOR expression in zebrafish results in colobomatous eye defects and perturbations in somite and skeletal anomalies, reminiscent of the human OFCD symptoms [130]. BCOR morpholinos also severely affect CNS development, causing disruption and irregularities of the cerebellum, optic tectum, and the boundaries of the fourth ventricle [130], suggesting that defects in neural development may cause the MR.

Recent studies provided insights into the mechanism of BCOR-mediated transcriptional repression. BCOR forms a multisubunit complex with the PcG proteins including RING1, RYBP, NSPC1, RNF2, and the HDM FBXL10/JHDM1B [132]. RNF2 is a mono-ubiquitin ligase of histone H2A [133, 134], which serves as a repressive histone mark [134, 135]. FBXL10/JHDM1B is capable of removing mono- and di-methylation on H3K36, modifications that are correlated with transcriptional elongation and suppression of intragenic transcription [136]. Thus, BCOR seems to function as a platform to recruit a repertoire of histone modifying enzymes onto specific genomic loci. Consistently, decreased occupation by FBXL10/JHDM1B and increased K36me2 have been observed in a particular gene promoter in the mesenchymal stem cells upon the loss of BCOR [137]. The C-terminus of BCOR contains the conserved ankyrin repeats (Fig. 1) and is required for the association of BCOR with NSPC1 and FBXL10/JHDM1B [132]. Strikingly, most patient mutations are nonsense, resulting in the generation of premature stop codons before the ankyrin repeats [130]. Interestingly, the ankyrin repeats of G9a and EHMT/GLP have recently been shown to recognize mono- and dimethylated histone H3K9 (H3K9me1/2) [138, 139]. It is tempting to speculate that the BCOR ankyrin repeats bind methylated histones and play a role in the coordination of the various histone-modifying enzymes in regulating gene expression. Once again, identification of relevant BCOR target genes in neuronal lineage development will allow further investigation of this corepressor in CNS development associated with human MR.

Understanding the role of "reader" proteins in MR is still at an early stage, as their involvement in MR has only been suggested by human genetic studies. Biochemical screenings aimed at identifying their binding substrates will provide insights into their mechanism of action, while their functions in the nervous system will have to be investigated by cellular and animal models. These efforts to clarify the roles of "readers" will provide the platform of understanding how epigenetic signals are translated into functional outcomes in neurons.

4 Noncovalent, ATP-Dependent Chromatin Remodeling in MR

In addition to covalent DNA and histone modifications, chromatin can also be regulated by noncovalent mechanisms such as ATP-dependent remodeling events mediated by ATP-dependent remodeling proteins (ATPases) that alter higher-order chromatin structure and transcriptional regulation (reviewed in [140–142]). *In vitro* activities of these enzymes include nucleosome sliding, eviction/deposition, and DNA-translocation, and their activities *in vivo* are often coordinated with enzymes that mediate covalent modifications, such as HATs and HDACs [141, 142]. To date, two ATP-dependent chromatin remodelers have been implicated in MR-associated human diseases. α-Thalassemia/mental retardation X-linked (ATRX syndrome) is a rare congenital disorder that inflicts severe MR, anemia, facial hypomorphism, and genital abnormality primarily on males [143, 144]. Unlike in conventional α-thalassemia, the α-globin genes are not mutated in these patients; instead, the α-globin mRNA level is significantly down-regulated, implicating the involvement of transcriptional or posttranscriptional mechanisms. The causative mutation resides in a gene encoding an SNF2 (Sucrose Non Fermentation)-like ATP-dependent remodeler, ATRX [145]. The majority of the missense mutations (55 out of 66) were found either in the N-terminal ADD (ATRX-DNMT3-DNMT3L) domain or the C-terminal enzymatic domain, indicating functional importance of both domains. However, essentially nothing is known about the function of either the ADD or the catalytic domain.

ATRX is known to be localized in the pericentromeric heterochromatin (PCH) [146], acrocentric chromosome where rDNA repeats reside [146] and the inactive X chromosome in female [147], suggesting a general function for ATRX in heterochromatin. Loss of ATRX results in sister chromatid cohesion and condensation defects, which in turn lead to chromosomal mis-segregation in neuroprecursors, as well as increased apoptosis and hypocellularity in the mouse forebrain [148]. This is consistent with the fact that the integrity of the PCH is crucial for sister chromatid cohesion and faithful chromosome segregation [149, 150]. These findings form the basis for the current model where ATRX syndrome is associated with reduced number of neurons due to apoptosis caused by the cohesion/condensation defect.

Interestingly, there may be a functional connection between ATRX and DNA CpG methylation via MeCP2, whose mutation causes Rett Syndrome, as discussed

above. It has been reported that ATRX physically interacts with MeCP2 [151] and that there is an alteration of CpG methylation level in certain area of the ATRX patient genome [152]. Specifically, patient lymphocytes showed either hypomethylation or hypermethylation at the rDNA loci and Y chromosome-specific repeat (DYZ2), respectively, suggesting a potential regulatory role of ATRX in CpG methylation [152]. Importantly, MeCP2 appears to be necessary for ATRX PCH localization [151], suggesting CpG methylation-dependent recruitment of ATRX to heterochromatin. These findings suggest reciprocal regulation between ATRX and DNA CpG methylation and MeCP2, representing one of the few examples where an interplay between two MR gene products is documented. Finally, a recent study identified ATRX as a chaperone for histone H3.3, and demonstrated that ATRX is important for the deposition of H3.3 in the telomeric regions [153]. Taken together, ATRX may play a role in heterochromatin-specific nucleosomal organization in concert with DNA methylation and deposition of specific histone variants.

Mutations in the second ATP-dependent remodeling protein, CHD7 (Chromodomain Helicase DNA binding), cause an MR-related disorder, CHARGE syndrome (Coloboma of the eye, heart defects, atresia of choanae, severe retardation of growth and development, genital and ear abnormalities) [154, 155]. CHARGE syndrome is a rare, usually sporadic dominant autosomal disorder and MR is found in more than 70% of the CHARGE cases where patients show low adaptive behavior skills and motor impairments [156]. Several mouse models recapitulate some of the patient phenotypes such as defects in inner ear and heart ultrastructure [157]. There is a broad and abundant expression of CHD7 in the developing CNS, including the neopallial (future frontal) cortex, tectum, and the ventricular zone of the medulla, yet the MR-relevant phenotypes of CHD7 mutation have not yet been examined.

Recent genome-wide mapping efforts of CHD7-bound loci revealed that CHD7 predominantly localizes distal to the TSS (Transcription Start Site), and in a smaller number of cases, to TSS. CHD7 localization most often coincides with H3K4me1/2 or H3K4me3 at the distal sites and TSS, respectively, where genes nearby are expressed at relatively high levels. Based on this observation together with other studies demonstrating the H3K4me1 is a hallmark modification of enhancers [80, 81], it has been suggested that CHD7 activates genes by facilitating long-range communications between enhancers and promoters [158]. However, important questions still remain; what are the MR-relevant target genes for CHD7? What is the impact of CHD7 loss on the regulation of these genes? What is the role of the ATP-dependent remodeling activity of CHD7 in gene regulation and in the CHARGE syndrome? To address these issues, further studies using the *in vivo* mouse models and *in vitro* enzymatic assay will be required.

There is very limited knowledge regarding how higher-order structure of chromatin is regulated that affects neuronal functions. Unlike histone modifying enzymes, ATP-dependent motors directly impact chromatin structure, which influences gene transcription and maintenance of genome integrity. Thus, ATRX and CHD7 are attractive models that provide opportunities to investigate how chromatin structure contributes to proper brain development and functioning.

5 Potential Therapeutic Implications

The demonstration that MeCP2 expression in the mouse Rett syndrome model can resurrect many of the cognitive function raised the hope of therapeutic intervention by targeting epigenetic regulators [45, 46]. As discussed above, a number of MR gene products, such as SMCX and PHF8, are epigenetic enzymes whose activities are compromised by the disease-causing genetic mutations. Thus, it is conceivable to develop small molecules that enhance and restore the enzymatic activities of the mutant enzymes, which in turn may have a beneficial effect on MR patients carrying these mutations. Alternatively, given the importance of proper epigenetic mechanism in learning and memory discussed above (Fig. 3), another therapeutic strategy would be to restore the balance in patients by small molecules that are designed to inhibit the negative regulators in cognitive functions. Furthermore, MR-associated mutations compromise not only enzymatic activity but also histone-binding ability of epigenetic regulators. For instance, missense mutations in the ADD domain of ATRX severely decrease its affinity to histone (Iwase and Shi unpublished). Thus, it is possible that small molecules can be found that will restore the ability of the mutant ADD domain to bind histone tails.

6 Concluding Remarks

In this review, we discussed evidence that supports an emerging role for epigenetic regulation in CNS functions that are relevant to MR. The developmental defects in these patients are often attributable to compromised neuronal differentiation and survival, as in the cases of DNMT1, MBD1, and ATRX. Other types of MR are accounted for by attenuated functions of postmitotic neurons, shown most clearly in Rett syndrome caused by MeCP2 deletion. Developmental defects likely involve the inheritance of disease-associated epigenetic modifications through successive cell divisions. In addition to the developmental roles, in mature, postmitotic neurons, epigenetic mechanisms may play a role in neuronal plasticity regulation relevant to learning and memory. In this regard, we noted the importance of balancing histone modifications via the opposing activities of the epigenetic regulators (CBP vs. HDAC2, and MLL vs. EED) in regulating LTP. These modifications are reversible due to the antagonistic actions of both the forward and the reverse enzymes (DNMT vs. the Gadd45 complex, HATs vs. HDACs, and HMTases vs. HDMases). Thus, the regulation of dynamic and reversible nature of DNA and histone modifications may be a key mechanism that underlies the unique nature of neurons; i.e., plasticity, which give rise to higher-order cognitive functioning. We have also described examples of histone modification "reader" proteins and their roles as integral components of the histone modifying enzymes in regulating neuronal functions, possibly functioning to guide the histone modifying enzymes and ATP-dependent remodelers to their respective genomic sites of

action. From the point of view of a molecular biologist, future challenges will include understanding the relationship, if any, of these various epigenetic regulators in MR, the biological basis for their involvement, and the molecular mechanism by which they regulate cognitive function. One attractive model is that there may be an intrinsic connection among various MR genes. For instance, the epigenetic regulators described here may be direct regulators of transcription of other MR genes whose products are directly involved in the regulation of synaptic plasticity, while MR genes encoding signaling molecules may play regulatory roles of either the epigenetic regulators and/or the gene products at the synapses. To decipher the principles of epigenetic regulation in neuronal and cognitive function, it will be necessary to employ an interdisciplinary effort involving animal models, electrophysiology, and behavior studies, as well as biochemical, molecular, structural, and chemical biology approaches.

References

1. Inlow JK, Restifo LL (2004) Molecular and comparative genetics of mental retardation. Genetics 166:835–881
2. Tarpey PS, Smith R, Pleasance E, Whibley A, Edkins S, Hardy C, O'Meara S, Latimer C, Dicks E, Menzies A et al (2009) A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. Nat Genet 41:535–543
3. Ropers HH, Hamel BC (2005) X-linked mental retardation. Nat Rev Genet 6:46–57
4. Chiurazzi P, Schwartz CE, Gecz J, Neri G (2008) XLMR genes: update 2007. Eur J Hum Genet 16:422–434
5. Akbarian S, Huang HS (2009) Epigenetic regulation in human brain-focus on histone lysine methylation. Biol Psychiatry 65:198–203
6. Bird A (2002) DNA methylation patterns and epigenetic memory. Genes Dev 16:6–21
7. Ooi SK, O'Donnell AH, Bestor TH (2009) Mammalian cytosine methylation at a glance. J Cell Sci 122:2787–2791
8. Illingworth RS, Bird AP (2009) CpG islands – 'a rough guide'. FEBS Lett 583:1713–1720
9. Ono T, Uehara Y, Kurishita A, Tawa R, Sakurai H (1993) Biological significance of DNA methylation in the ageing process. Age Ageing 22:S34–S43
10. Wilson VL, Smith RA, Ma S, Cutler RG (1987) Genomic 5-methyldeoxycytidine decreases with age. J Biol Chem 262:9948–9951
11. Tawa R, Ono T, Kurishita A, Okada S, Hirose S (1990) Changes of DNA methylation level during pre- and postnatal periods in mice. Differentiation 45:44–48
12. Ballestar E, Wolffe AP (2001) Methyl-CpG-binding proteins. Targeting specific gene repression. Eur J Biochem 268:1–6
13. Wade PA (2001) Methyl CpG-binding proteins and transcriptional repression. Bioessays 23:1131–1137
14. 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
15. Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CM, Gartler SM (1999) The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. Proc Natl Acad Sci USA 96:14412–14417
16. Ehrlich M (2003) The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. Clin Immunol 109:17–28

17. Jin B, Tao Q, Peng J, Soo HM, Wu W, Ying J, Fields CR, Delmas AL, Liu X, Qiu 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
18. Fan G, Beard C, Chen RZ, Csankovszki G, Sun Y, Siniaia M, Biniszkiewicz D, Bates B, Lee PP, Kuhn R et al (2001) DNA hypomethylation perturbs the function and survival of CNS neurons in postnatal animals. J Neurosci 21:788–797
19. Zhu JK (2009) Active DNA demethylation mediated by DNA glycosylases. Annu Rev Genet 43:143–166
20. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324:930–935
21. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324:929–930
22. Chen WG, Chang Q, Lin Y, Meissner A, West AE, Griffith EC, Jaenisch R, Greenberg ME (2003) Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. Science 302:885–889
23. Martinowich K, Hattori D, Wu H, Fouse S, He F, Hu Y, Fan G, Sun YE (2003) DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. Science 302:890–893
24. Miller CA, Sweatt JD (2007) Covalent modification of DNA regulates memory formation. Neuron 53:857–869
25. Greenberg ME, Xu B, Lu B, Hempstead BL (2009) New insights in the biology of BDNF synthesis and release: implications in CNS function. J Neurosci 29:12764–12767
26. Greer PL, Greenberg ME (2008) From synapse to nucleus: calcium-dependent gene transcription in the control of synapse development and function. Neuron 59:846–860
27. 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
28. Ma DK, Jang MH, Guo JU, Kitabatake Y, Chang ML, Pow-Anpongkul N, Flavell RA, Lu B, Ming GL, Song H (2009) Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. Science 323:1074–1077
29. Rett A (1966) On a unusual brain atrophy syndrome in hyperammonemia in childhood. Wien Med Wochenschr 116:723–726
30. Armstrong D, Dunn JK, Antalffy B, Trivedi R (1995) Selective dendritic alterations in the cortex of Rett syndrome. J Neuropathol Exp Neurol 54:195–201
31. Belichenko PV, Hagberg B, Dahlstrom A (1997) Morphological study of neocortical areas in Rett syndrome. Acta Neuropathol 93:50–61
32. Guy J, Hendrich B, Holmes M, Martin JE, Bird A (2001) A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat Genet 27:322–326
33. Chen RZ, Akbarian S, Tudor M, Jaenisch R (2001) Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. Nat Genet 27:327–331
34. Kishi N, Macklis JD (2004) MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions. Mol Cell Neurosci 27:306–321
35. Mullaney BC, Johnston MV, Blue ME (2004) Developmental expression of methyl-CpG binding protein 2 is dynamically regulated in the rodent brain. Neuroscience 123:939–949
36. 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
37. Fukuda T, Itoh M, Ichikawa T, Washiyama K, Goto Y (2005) Delayed maturation of neuronal architecture and synaptogenesis in cerebral cortex of Mecp2-deficient mice. J Neuropathol Exp Neurol 64:537–544

38. Matarazzo V, Cohen D, Palmer AM, Simpson PJ, Khokhar B, Pan SJ, Ronnett GV (2004) The transcriptional repressor Mecp2 regulates terminal neuronal differentiation. Mol Cell Neurosci 27:44–58
39. Smrt RD, Eaves-Egenes J, Barkho BZ, Santistevan NJ, Zhao C, Aimone JB, Gage FH, Zhao X (2007) Mecp2 deficiency leads to delayed maturation and altered gene expression in hippocampal neurons. Neurobiol Dis 27:77–89
40. Asaka Y, Jugloff DG, Zhang L, Eubanks JH, Fitzsimonds RM (2006) Hippocampal synaptic plasticity is impaired in the Mecp2-null mouse model of Rett syndrome. Neurobiol Dis 21:217–227
41. Moretti P, Levenson JM, Battaglia F, Atkinson R, Teague R, Antalffy B, Armstrong D, Arancio O, Sweatt JD, Zoghbi HY (2006) Learning and memory and synaptic plasticity are impaired in a mouse model of Rett syndrome. J Neurosci 26:319–327
42. Nelson ED, Kavalali ET, Monteggia LM (2006) MeCP2-dependent transcriptional repression regulates excitatory neurotransmission. Curr Biol 16:710–716
43. Dani VS, Chang Q, Maffei A, Turrigiano GG, Jaenisch R, Nelson SB (2005) Reduced cortical activity due to a shift in the balance between excitation and inhibition in a mouse model of Rett syndrome. Proc Natl Acad Sci USA 102:12560–12565
44. Pelka GJ, Watson CM, Radziewic T, Hayward M, Lahooti H, Christodoulou J, Tam PP (2006) Mecp2 deficiency is associated with learning and cognitive deficits and altered gene activity in the hippocampal region of mice. Brain 129:887–898
45. Giacometti E, Luikenhuis S, Beard C, Jaenisch R (2007) Partial rescue of MeCP2 deficiency by postnatal activation of MeCP2. Proc Natl Acad Sci USA 104:1931–1936
46. Guy J, Gan J, Selfridge J, Cobb S, Bird A (2007) Reversal of neurological defects in a mouse model of Rett syndrome. Science 315:1143–1147
47. Willard HF, Hendrich BD (1999) Breaking the silence in Rett syndrome. Nat Genet 23:127–128
48. Chahrour M, Jung SY, Shaw C, Zhou X, Wong ST, Qin J, Zoghbi HY (2008) MeCP2, a key contributor to neurological disease, activates and represses transcription. Science 320:1224–1229
49. 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
50. Lalande M, Minassian BA, DeLorey TM, Olsen RW (1999) Parental imprinting and Angelman syndrome. Adv Neurol 79:421–429
51. Yasui DH, Peddada S, Bieda MC, Vallero RO, Hogart A, Nagarajan RP, Thatcher KN, Farnham PJ, Lasalle JM (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
52. Skene PJ, Illingworth RS, Webb S, Kerr AR, James KD, Turner DJ, Andrews R, Bird AP (2010) Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. Mol Cell 37:457–468
53. Zhao X, Ueba T, Christie BR, Barkho B, McConnell MJ, Nakashima K, Lein ES, Eadie BD, Willhoite AR, Muotri AR et al (2003) Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function. Proc Natl Acad Sci USA 100:6777–6782
54. Allfrey VG (1966) Structural modifications of histones and their possible role in the regulation of ribonucleic acid synthesis. Proc Can Cancer Conf 6:313–335
55. Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD (1996) Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell 84:843–851
56. Zhang Y, Reinberg D (2001) Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev 15:2343–2360
57. Luger K, Richmond TJ (1998) The histone tails of the nucleosome. Curr Opin Genet Dev 8:140–146

58. Edmondson DG, Smith MM, Roth SY (1996) Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. Genes Dev 10:1247–1259
59. Hecht A, Laroche T, Strahl-Bolsinger S, Gasser SM, Grunstein M (1995) Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. Cell 80:583–592
60. Bannister AJ, Kouzarides T (1996) The CBP co-activator is a histone acetyltransferase. Nature 384:641–643
61. Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy MR, Goodman RH (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP. Nature 365:855–859
62. Carlezon WA Jr, Duman RS, Nestler EJ (2005) The many faces of CREB. Trends Neurosci 28:436–445
63. Pinsker HM, Hening WA, Carew TJ, Kandel ER (1973) Long-term sensitization of a defensive withdrawal reflex in Aplysia. Science 182:1039–1042
64. Kandel ER (2001) The molecular biology of memory storage: a dialogue between genes and synapses. Science 294:1030–1038
65. Cohen S, Greenberg ME (2008) Communication between the synapse and the nucleus in neuronal development, plasticity, and disease. Annu Rev Cell Dev Biol 24:183–209
66. Petrij F, Giles RH, Dauwerse HG, Saris JJ, Hennekam RC, Masuno M, Tommerup N, van Ommen GJ, Goodman RH, Peters DJ et al (1995) Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. Nature 376:348–351
67. Alarcon JM, Malleret G, Touzani K, Vronskaya S, Ishii S, Kandel ER, Barco A (2004) Chromatin acetylation, memory, and LTP are impaired in CBP+/− mice: a model for the cognitive deficit in Rubinstein-Taybi syndrome and its amelioration. Neuron 42:947–959
68. Korzus E, Rosenfeld MG, Mayford M (2004) CBP histone acetyltransferase activity is a critical component of memory consolidation. Neuron 42:961–972
69. Taunton J, Hassig CA, Schreiber SL (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. Science 272:408–411
70. Grozinger CM, Schreiber SL (2002) Deacetylase enzymes: biological functions and the use of small-molecule inhibitors. Chem Biol 9:3–16
71. Ayer DE (1999) Histone deacetylases: transcriptional repression with SINers and NuRDs. Trends Cell Biol 9:193–198
72. Chinnadurai G (2002) CtBP, an unconventional transcriptional corepressor in development and oncogenesis. Mol Cell 9:213–224
73. Jones PL, Shi YB (2003) N-CoR-HDAC corepressor complexes: roles in transcriptional regulation by nuclear hormone receptors. Curr Top Microbiol Immunol 274:237–268
74. Hildmann C, Riester D, Schwienhorst A (2007) Histone deacetylases – an important class of cellular regulators with a variety of functions. Appl Microbiol Biotechnol 75:487–497
75. Guan Z, Giustetto M, Lomvardas S, Kim JH, Miniaci MC, Schwartz JH, Thanos D, Kandel ER (2002) Integration of long-term-memory-related synaptic plasticity involves bidirectional regulation of gene expression and chromatin structure. Cell 111:483–493
76. Guan JS, Haggarty SJ, Giacometti E, Dannenberg JH, Joseph N, Gao J, Nieland TJ, Zhou Y, Wang X, Mazitschek R et al (2009) HDAC2 negatively regulates memory formation and synaptic plasticity. Nature 459:55–60
77. Morrison AJ, Shen X (2005) DNA repair in the context of chromatin. Cell Cycle 4:568–571
78. Altaf M, Saksouk N, Cote J (2007) Histone modifications in response to DNA damage. Mutat Res 618:81–90
79. Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. Cell 128:707–719
80. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA et al (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39:311–318

81. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP et al (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448:553–560
82. Bannister AJ, Schneider R, Kouzarides T (2002) Histone methylation: dynamic or static? Cell 109:801–806
83. Kubicek S, Jenuwein T (2004) A crack in histone lysine methylation. Cell 119:903–906
84. Bannister AJ, Kouzarides T (2005) Reversing histone methylation. Nature 436:1103–1106
85. Shi Y, Whetstine JR (2007) Dynamic regulation of histone lysine methylation by demethylases. Mol Cell 25:1–14
86. Klose RJ, Zhang Y (2007) Regulation of histone methylation by demethylimination and demethylation. Nat Rev Mol Cell Biol 8:307–318
87. Kurotaki N, Imaizumi K, Harada N, Masuno M, Kondoh T, Nagai T, Ohashi H, Naritomi K, Tsukahara M, Makita Y et al (2002) Haploinsufficiency of NSD1 causes Sotos syndrome. Nat Genet 30:365–366
88. Ellison J (2008) Gene symbol: NSD1. Disease: Sotos syndrome. Hum Genet 124:311
89. Rayasam GV, Wendling O, Angrand PO, Mark M, Niederreither K, Song L, Lerouge T, Hager GL, Chambon P, Losson R (2003) NSD1 is essential for early post-implantation development and has a catalytically active SET domain. Embo J 22:3153–3163
90. Li Y, Trojer P, Xu CF, Cheung P, Kuo A, Drury WJ 3rd, Qiao Q, Neubert TA, Xu RM, Gozani O et al (2009) The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. J Biol Chem 284:34283–34295
91. Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP et al (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell 123:581–592
92. Joshi AA, Struhl K (2005) Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Mol Cell 20:971–978
93. Keogh MC, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ et al (2005) Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell 123:593–605
94. Huang N, vom Baur E, Garnier JM, Lerouge T, Vonesch JL, Lutz Y, Chambon P, Losson R (1998) Two distinct nuclear receptor interaction domains in NSD1, a novel SET protein that exhibits characteristics of both corepressors and coactivators. Embo J 17:3398–3412
95. Bremner JD, McCaffery P (2008) The neurobiology of retinoic acid in affective disorders. Prog Neuropsychopharmacol Biol Psychiatry 32:315–331
96. McCaffery P, Zhang J, Crandall JE (2006) Retinoic acid signaling and function in the adult hippocampus. J Neurobiol 66:780–791
97. Kleefstra T, van Zelst-Stams WA, Nillesen WM, Cormier-Daire V, Houge G, Foulds N, van Dooren M, Willemsen MH, Pfundt R, Turner A et al (2009) Further clinical and molecular delineation of the 9q subtelomeric deletion syndrome supports a major contribution of EHMT1 haploinsufficiency to the core phenotype. J Med Genet 46:598–606
98. Ogawa H, Ishiguro K, Gaubatz S, Livingston DM, Nakatani Y (2002) A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. Science 296:1132–1136
99. Tachibana M, Ueda J, Fukuda M, Takeda N, Ohta T, Iwanari H, Sakihama T, Kodama T, Hamakubo T, Shinkai Y (2005) Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. Genes Dev 19:815–826
100. Stewart DR, Kleefstra T (2007) The chromosome 9q subtelomere deletion syndrome. Am J Med Genet C Semin Med Genet 145C:383–392
101. Balemans MC, Huibers MM, Eikelenboom NW, Kuipers AJ, Summeren RC, Pijpers MM, Tachibana M, Shinkai Y, van Bokhoven H, Zee CE (2010) Reduced exploration, increased anxiety, and altered social behavior: autistic-like features of Euchromatin histone methyltransferase 1 heterozygous knockout mice. Behav Brain Res 208(1):47–55

102. Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH, Allis CD (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell 121:859–872
103. Kim SY, Levenson JM, Korsmeyer S, Sweatt JD, Schumacher A (2007) Developmental regulation of Eed complex composition governs a switch in global histone modification in brain. J Biol Chem 282:9962–9972
104. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev 16:2893–2905
105. Tahiliani M, Mei P, Fang R, Leonor T, Rutenberg M, Shimizu F, Li J, Rao A, Shi Y (2007) The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. Nature 447:601–605
106. Iwase S, Lan F, Bayliss P, de la Torre-Ubieta L, Huarte M, Qi HH, Whetstine JR, Bonni A, Roberts TM, Shi Y (2007) The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. Cell 128:1077–1088
107. Loenarz C, Ge W, Coleman ML, Rose NR, Cooper CD, Klose RJ, Ratcliffe PJ, Schofield CJ (2009) PHF8, a gene associated with cleft lip/palate and mental retardation, encodes for an N{varepsilon}-dimethyl lysine demethylase. Hum Mol Genet 19(2):217–222
108. Carlisle HJ, Kennedy MB (2005) Spine architecture and synaptic plasticity. Trends Neurosci 28:182–187
109. Halpain S, Spencer K, Graber S (2005) Dynamics and pathology of dendritic spines. Prog Brain Res 147:29–37
110. Jones FS, Meech R (1999) Knockout of REST/NRSF shows that the protein is a potent repressor of neuronally expressed genes in non-neural tissues. Bioessays 21:372–376
111. Coulson JM (2005) Transcriptional regulation: cancer, neurons and the REST. Curr Biol 15: R665–R668
112. Ruthenburg AJ, Li H, Patel DJ, Allis CD (2007) Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol 8:983–994
113. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol 14:1025–1040
114. Chan DW, Wang Y, Wu M, Wong J, Qin J, Zhao Y (2009) Unbiased proteomic screen for binding proteins to modified lysines on histone H3. Proteomics 9:2343–2354
115. Nottke A, Colaiacovo MP, Shi Y (2009) Developmental roles of the histone lysine demethylases. Development 136:879–889
116. Lan F, Nottke AC, Shi Y (2008) Mechanisms involved in the regulation of histone lysine demethylases. Curr Opin Cell Biol 20:316–325
117. Lu X, Meng X, Morris CA, Keating MT (1998) A novel human gene, WSTF, is deleted in Williams syndrome. Genomics 54:241–249
118. Schubert C (2009) The genomic basis of the Williams-Beuren syndrome. Cell Mol Life Sci 66:1178–1197
119. Kitagawa H, Fujiki R, Yoshimura K, Mezaki Y, Uematsu Y, Matsui D, Ogawa S, Unno K, Okubo M, Tokita A et al (2003) The chromatin-remodeling complex WINAC targets a nuclear receptor to promoters and is impaired in Williams syndrome. Cell 113:905–917
120. Xiao A, Li H, Shechter D, Ahn SH, Fabrizio LA, Erdjument-Bromage H, Ishibe-Murakami S, Wang B, Tempst P, Hofmann K et al (2009) WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity. Nature 457:57–62
121. Poot RA, Bozhenok L, van den Berg DL, Steffensen S, Ferreira F, Grimaldi M, Gilbert N, Ferreira J, Varga-Weisz PD (2004) The Williams syndrome transcription factor interacts with PCNA to target chromatin remodelling by ISWI to replication foci. Nat Cell Biol 6:1236–1244
122. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM (1999) Structure and ligand of a histone acetyltransferase bromodomain. Nature 399:491–496

123. Kato S, Fujiki R, Kim MS, Kitagawa H (2007) Ligand-induced transrepressive function of VDR requires a chromatin remodeling complex, WINAC. J Steroid Biochem Mol Biol 103:372–380
124. Baker LA, Allis CD, Wang GG (2008) PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. Mutat Res 647:3–12
125. Adams-Cioaba MA, Min J (2009) Structure and function of histone methylation binding proteins. Biochem Cell Biol 87:93–105
126. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD et al (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 448:714–717
127. Lan F, Collins RE, De Cegli R, Alpatov R, Horton JR, Shi X, Gozani O, Cheng X, Shi Y (2007) Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. Nature 448:718–722
128. Field M, Tarpey PS, Smith R, Edkins S, O'Meara S, Stevens C, Tofts C, Teague J, Butler A, Dicks E et al (2007) Mutations in the BRWD3 gene cause X-linked mental retardation associated with macrocephaly. Am J Hum Genet 81:367–374
129. Lower KM, Turner G, Kerr BA, Mathews KD, Shaw MA, Gedeon AK, Schelley S, Hoyme HE, White SM, Delatycki MB et al (2002) Mutations in PHF6 are associated with Borjeson-Forssman-Lehmann syndrome. Nat Genet 32:661–665
130. Ng D, Thakker N, Corcoran CM, Donnai D, Perveen R, Schneider A, Hadley DW, Tifft C, Zhang L, Wilkie AO et al (2004) Oculofaciocardiodental and Lenz microphthalmia syndromes result from distinct classes of mutations in BCOR. Nat Genet 36:411–416
131. Dent AL, Shaffer AL, Yu X, Allman D, Staudt LM (1997) Control of inflammation, cytokine expression, and germinal center formation by BCL-6. Science 276:589–592
132. Gearhart MD, Corcoran CM, Wamstad JA, Bardwell VJ (2006) Polycomb group and SCF ubiquitin ligases are found in a novel BCOR complex that is recruited to BCL6 targets. Mol Cell Biol 26:6880–6889
133. Cao R, Tsukada Y, Zhang Y (2005) Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol Cell 20:845–854
134. Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y (2004) Role of histone H2A ubiquitination in Polycomb silencing. Nature 431:873–878
135. de Napoles M, Mermoud JE, Wakao R, Tang YA, Endoh M, Appanah R, Nesterova TB, Silva J, Otte AP, Vidal M et al (2004) Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev Cell 7:663–676
136. Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, Zhang Y (2006) Histone demethylation by a family of JmjC domain-containing proteins. Nature 439:811–816
137. Fan Z, Yamaza T, Lee JS, Yu J, Wang S, Fan G, Shi S, Wang CY (2009) BCOR regulates mesenchymal stem cell function by epigenetic mechanisms. Nat Cell Biol 11:1002–1009
138. Brent MM, Marmorstein R (2008) Ankyrin for methylated lysines. Nat Struct Mol Biol 15:221–222
139. Collins RE, Northrop JP, Horton JR, Lee DY, Zhang X, Stallcup MR, Cheng X (2008) The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. Nat Struct Mol Biol 15:245–250
140. Winston F, Carlson M (1992) Yeast SNF/SWI transcriptional activators and the SPT/SIN chromatin connection. Trends Genet 8:387–391
141. Narlikar GJ, Fan HY, Kingston RE (2002) Cooperation between complexes that regulate chromatin structure and transcription. Cell 108:475–487
142. Kingston RE, Narlikar GJ (1999) ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. Genes Dev 13:2339–2352
143. Gibbons R (2006) Alpha thalassaemia-mental retardation, X linked. Orphanet J Rare Dis 1:15

144. Gibbons RJ, Higgs DR (2000) Molecular-clinical spectrum of the ATR-X syndrome. Am J Med Genet 97:204–212
145. 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
146. McDowell TL, Gibbons RJ, Sutherland H, O'Rourke DM, Bickmore WA, Pombo A, Turley H, Gatter K, Picketts DJ, Buckle VJ 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
147. 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
148. Ritchie K, Seah C, Moulin J, Isaac C, Dick F, Berube NG (2008) Loss of ATRX leads to chromosome cohesion and congression defects. J Cell Biol 180:315–324
149. Pidoux AL, Allshire RC (2005) The role of heterochromatin in centromere function. Philos Trans R Soc Lond B Biol Sci 360:569–579
150. Allshire RC, Nimmo ER, Ekwall K, Javerzat JP, Cranston G (1995) Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. Genes Dev 9:218–233
151. Nan X, Hou J, Maclean A, Nasir J, Lafuente MJ, Shu X, Kriaucionis S, Bird A (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
152. Gibbons RJ, McDowell TL, Raman S, O'Rourke DM, Garrick D, Ayyub H, Higgs DR (2000) Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. Nat Genet 24:368–371
153. Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, Dewell S, Law M, Guo X, Li X et al (2010) Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell 140:678–691
154. Pagon RA, Graham JM Jr, Zonana J, Yong SL (1981) Coloboma, congenital heart disease, and choanal atresia with multiple anomalies: CHARGE association. J Pediatr 99:223–227
155. Vissers LE, van Ravenswaaij CM, Admiraal R, Hurst JA, de Vries BB, Janssen IM, van der Vliet WA, Huys EH, de Jong PJ, Hamel BC et al (2004) Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. Nat Genet 36:955–957
156. Sanlaville D, Verloes A (2007) CHARGE syndrome: an update. Eur J Hum Genet 15:389–399
157. Bosman EA, Penn AC, Ambrose JC, Kettleborough R, Stemple DL, Steel KP (2005) Multiple mutations in mouse Chd7 provide models for CHARGE syndrome. Hum Mol Genet 14:3463–3476
158. Schnetz MP, Bartels CF, Shastri K, Balasubramanian D, Zentner GE, Balaji R, Zhang X, Song L, Wang Z, Laframboise T et al (2009) Genomic distribution of CHD7 on chromatin tracks H3K4 methylation patterns. Genome Res 19:590–601
159. Jensen LR, Amende M, Gurok U, Moser B, Gimmel V, Tzschach A, Janecke AR, Tariverdian G, Chelly J, Fryns JP et al (2005) Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. Am J Hum Genet 76:227–236
160. Laumonnier F, Holbert S, Ronce N, Faravelli F, Lenzner S, Schwartz CE, Lespinasse J, Van Esch H, Lacombe D, Goizet C et al (2005) Mutations in PHF8 are associated with X linked mental retardation and cleft lip/cleft palate. J Med Genet 42:780–786
161. Borjeson M, Forssman H, Lehmann O (1962) An X-linked, recessively inherited syndrome characterized by grave mental deficiency, epilepsy, and endocrine disorder. Acta Med Scand 171:13–21

HDAC Inhibitors and Cancer Therapy

Peter W. Atadja

Abstract Maintenance of normal cell growth and differentiation is highly dependent on coordinated and tight transcriptional regulation of genes. In cancer, genes encoding growth regulators are abnormally expressed. Particularly, silencing of tumor suppressor genes under the control of chromatin modifications is a major underlying cause of unregulated cellular proliferation and transformation. Thus mechanisms, which regulate chromatin structure and gene expression, have become attractive targets for anticancer therapy. Histone deacetylases are enzymes that modify chromatin structure and contribute to aberrant gene expression in cancer. Research over the past decade has led to the development of histone deacetylase inhibitors as anticancer agents. In addition to their effect on chromatin and epigenetic mechanisms, HDAC inhibitors also modify the acetylation state of a large number of cellular proteins involved in oncogenic processes, resulting in antitumor effects. The current monograph will review the role of histone deacetylases in protumorigenic mechanisms and the current developmental status and prospects for their inhibitors in cancer therapy.

1 Introduction

Aberrant expression of genes, such as increased expression of cellular growth and survival genes and/or silencing of tumor suppressor genes, constitutes a major molecular hallmark of cancer. Therefore, targeting the gene expression machinery to restore normal cellular gene expression promises to be a useful anticancer therapeutic strategy (see recent reviews in [1] and [2]).

P.W. Atadja
Novartis Institute for Biomedical Research, 898 Halei Rd. Buliding 8, Shanghai, China
e-mail: peter.atadja@novartis.com

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_9,

Due to the protein–protein interactions and multiprotein complexes required to regulate transcription, attempts to target the transcriptional machinery with small molecules with reasonable potency and specificity had been challenging and mostly unsuccessful. Recent studies have unveiled another level of transcriptional regulation through enzymatic modification of chromatin structure, a mechanism that is providing more successful opportunities to modulate gene expression with small molecule inhibitors for therapeutic purposes [2]. This chapter reviews current information on inhibition of one such chromatin modifying enzyme class, histone deacetylases (HDACs) as an anticancer therapeutic strategy.

HDACs are a class of enzymes that regulate protein function through the removal of acetyl groups from lysine residues. As the name indicates, histones are the major substrate class for HDACs, and by deacetylating the lysine tails of core histones, they modify chromatin structure. Such chromatin modifications result in changes to gene transcription and expression [3]. Although acetylated histones were the earliest substrates identified for HDACs, a large number of additional nuclear and cytoplasmic proteins mediating diverse biological processes have been identified, whose structure and function could be regulated through acetylation [4, 5]. Thus, as in other posttranslational protein modifications, acetylation is emerging as a major mechanism of regulating protein and cell function.

2 Opposing Activities of Histone Acetylation and Deacetylation

Histones are a fundamental component of nucleosomes, the basic packaging unit of nuclear DNA. Each nucleosome contains an octamer of histones H2a, H2b, H3, and H4 variants forming a globular core around which about 146 base pairs of DNA are wrapped. Internucleosomal interactions further package nuclear DNA into the chromatin fiber. Accessibility of the DNA for gene transcription is regulated by changes in chromatin structure. Condensed chromatin, in which the nucleosomes are tightly compacted, may not be as easily accessible to the transcriptional machinery as chromatin in which the histones and the DNA are not as tightly packed. Thus, the histone component of chromatin functions to organize the DNA as well as provide mechanisms for regulating transcription factor accessibility and gene expression [1].

Chromatin condensation and packing is promoted by the ionic interaction between positively charged histones and a negatively charged DNA backbone. Posttranslational modifications, such as acetylation, neutralize the positive charge of histone tails and decrease the strength of interaction with the DNA backbone, leading to a more open chromatin structure. Conversely, when acetyl groups are removed from histones by HDACs, chromatin is more tightly packed leading to gene silencing [6]. In this way, gene expression is regulated in part by the opposing activities of histone acetylation and histone deacetylation through dynamic chromatin changes.

3 Classification of HDACs and Networks of Protein Regulation by Acetylation and Deacetylation

There are four identified classes of HDACs (Classes I–IV). Class I, II, and IV HDACs are all dependent upon Zn^{2+} cations for catalytic activity (Table 1). Class I HDACs include HDACs 1, 2, 3, and 8 and are found primarily in the nuclei of cells. The principal target substrates for Class I HDACs are histones. Class II HDACs act on histone and nonhistone proteins and are of two types. The Class IIa HDACs (HDACs 4, 5, 7, and 9) can shuttle between the nucleus and cytoplasm, and Class IIb HDACs (HDACs 6 and 7) are located primarily in the cytoplasm [7, 8]. HDAC11 is the sole member of Class IV, as it shares attributes of both Class I and II HDACs [9, 10]. A third class of HDACs (Class III) is comprised of a family of structurally distinct, NAD+-dependent enzymes known as sirtuins, which are not Zn^{2+} dependent and will not be discussed in this chapter.

The diverse effects of HATs and HDACs on cellular function and behavior have long suggested that acetylation plays fundamental roles in contexts other than histone and DNA-dependent processes. HDACs have been found to possess general deacetylase activity against a broad spectrum of histone and nonhistone protein substrates. An increasing number of nonhistone proteins have been identified and functionally characterized. These include transcription factors, nuclear hormone receptors, and other factors that have significant downstream effects in pathways relating to tumor development and survival [8, 11]. The developing concept of the "acetylome" represents a growing list of proteins dynamically regulated by acetylation and deacetylation [5, 8, 12]. A diverse array of cellular functions is thus affected by the balance between HAT and HDAC activities acting upon proteins that make up the "acetylome." With continued research into the scope of the "acetylome," numerous proteins have been identified as being acetylated in human cells and for many, the effects and regulation of their acetylation status remain unknown. The specific effects of acetylation and deacetylation on proteins may vary on an individual basis with effects on cellular function and behavior [4, 11].

Table 1 Representative HDAC inhibitors and some areas of clinical development

Compound	Type	Cellular potency	Isoform selectivity	Clinical phase	Some major clinical studies ongoing
Phenyl butyrate	Carboxylic acid	μM	Pan	I/II	Heme malignancies, prostate, endometrial
SAHA	Hydroxamate	μM	Pan	Approved for CTCL	CTCL, lung, prostate, mesothelioma, breast, AML
Panobinostat	Hydroxamate	nM	Pan	I/II	Multiple myeloma, Hodgkins lymphoma, CTCL, prostate, breast, AML
PXD101	Hydroxamate	nM	Pan	I/II	Breast, prostate
SNDX275	Benzamide	nM	Class 1	I/II	Hematological malignancies
MGCD0103	Benzamide	nM	Class I	I/II	Hodgkins lymphoma, AML
Rhomidepsin	Depsipeptide	nM	Class I	I/II	CTCL, PCTL, Melanoma

4 Rationale for Targeting HDACs for Anticancer Therapy

Deregulation of HDAC levels and their functions have been discovered in cancer and in the process of tumorigenesis [2, 7, 13]. Abnormal HDAC activities have been shown to play major roles in mechanisms regulating a number of the hallmarks of cancer, including tumor suppressor silencing, aberrant cell-cycle control and growth signaling, differentiation, angiogenesis, cell adhesion, tissue evasion, and metastasis [14]. These discoveries have formed the basis of a vigorous investigation and use of HDAC inhibitors in anticancer therapy.

Although no major cancer-linked mutations of HDACs have so far been identified, over-expression of HDAC protein has been identified in multiple cancers. In one study by Zhu et al. [15], increased HDAC2 expression was found in the majority of human colon cancer explants, as well as in intestinal mucosa and polyps of *APC*-deficient mice. They compared 57 samples of mostly moderately differentiated colon tumors vs. patient-matched normal tissues with respect to alterations in HDAC2, and elevated HDAC2 expression was observed in 47 samples (82%), as compared with normal adjacent tissues. As an example of a cross-talk with other tumorigenic pathway, HDAC2 and β-catenin appeared to be regulated in the same way in the vast majority of the tumor samples, suggesting that both are, in most cases, related events in the development of colon cancer [15]. Similar increases of HDAC expression were described in other tumor types, including prostate, gastric, and breast cancers.

4.1 Role of HDACs in Tumor Suppressor Silencing

Aberrant recruitment of HDACs has been correlated with decreased levels of acetylated histones at specific genes associated with the control of cell growth [4]. Particular genes involved in tumor suppression and cell growth regulation that are known to be silenced through HDAC-mediated epigenetic mechanisms include p21, p27, p16, p19, among many others (Fig. 1). Using chromatin precipitation assays, Gui and colleagues [16] showed that the class 1 HDACs 1 and 2 are directly bound to the p21 promoter and undergo changes in promoter occupancy with HDAC inhibitors treatment [16].

The link between cancer and HDACs is best characterized in leukemias, in which chromosomal translocations that promote oncogenesis are associated with increased recruitment of HDAC to gene promoters. In acute promyelocytic leukemias (APL), the chromosomal translocations t(15;17) and t(11;17) result in the fusion proteins RARα-PML and RARα-PLZF, respectively. These oncogenic fusion proteins form complexes with HDACs to repress transcription of RAR-targeted genes required for normal myeloid differentiation (reviewed in [17]). The very first clinical antitumor activity associated with an HDAC inhibitor was when butyric acid was combined with retinoic acid to achieve remission in a patient with refractory APL [18]. Other examples of HDAC-associated translocations in leukemiogenesis include AML1-ETO t(8;21) and CBFβ-MYH11 containing chromosome inversion 16 [17,

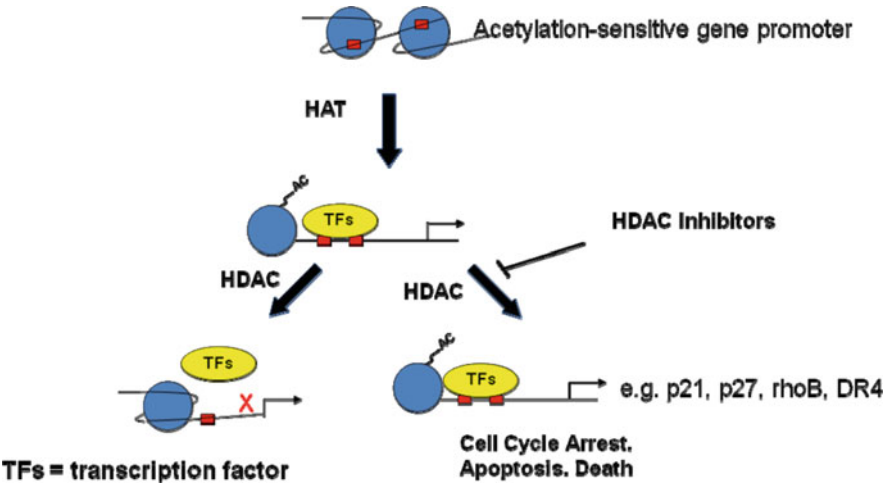


Fig. 1 A model of transcriptional control of growth regulatory genes by acetylation and deacetylation. The possible effect of HDAC inhibitors on gene expression is depicted

19]. In both cases, the fusion proteins combine the DNA binding activities of critical regulators of hematopoiesis with the HDAC-recruiting capabilities of corepressors resulting in the silencing of genes involved in myeloid differentiation [17, 19]. In yet another example, the proliferation and survival of malignant B-cells in non-Hodgkin's lymphoma (NHL) was shown to be driven by HDAC-containing complexes of the Bcl-6 oncoprotein [17, 19, 20].

HDACs may also regulate transcription of growth regulatory genes through their effect on the acetylation state of transcription factors. This has been demonstrated for the transcriptional activity of the tumor suppressor p53. In studies by Gu and colleagues [21–23], they demonstrated that direct p53 acetylation, specifically in the c-terminal domain enhanced its sequence-specific DNA binding *in vitro* and that endogenous p53 can be fully acetylated in response to DNA damage when HDAC is inhibited. These investigators further demonstrated that the activity of the p53 tumor suppressor is modulated by protein stabilization and posttranslational modifications including acetylation in response to DNA damage. Interestingly, they showed that both acetylation and ubiquitination modified the same lysine residues at the C terminus of p53, and that this lysine residue is the target for ubiquitination by the anti-p53 oncogene mdm2 [21, 24]. In another example, the non-Hodgkin lymphoma-associated transcription factor BCL6 is also reported to be inactivated through direct acetylation of the protein [19].

4.2 Role for HDACs in Deregulated Cell Growth and Survival Pathways

HDAC activity has been shown to be required for cell-cycle progression, as well as many other cellular functions required for cell proliferation and survival. The

earliest noticeable effect of HDAC inhibition on cell proliferation is the arrest of the cell cycle at both the G1/S and G2/M phases. An important axis of control for the G1 phase of the mammalian cell cycle involves the interactions of the E2F family of DNA binding proteins with the tumor suppressor retinoblastoma (Rb) and Rb-related "pocket protein" (p107 and p130) family of tumor proteins and not surprisingly, alterations in this pathway are present in a large number of human malignancies [25]. Although the exact mechanism of Rb-mediated cell-cycle deregulation in cancer is not fully understood, in vitro studies have identified HDACs in E2F and Rb complexes. Since hyperphosphorylation of Rb by cyclin-dependent kinases (cdks) is associated with and might be required for exit through the G1/S restriction point, HDAC-mediated silencing of cdk inhibitors, such as p21, p27, and p16, may allow uncontrolled Rb phosphorylation and cell proliferation. Interestingly, some studies have also shown that HDACs may be required for Rb-mediated gene repression, and whether this activity of HDACs is involved in Rb's tumor suppressor function is not clear [25].

HDAC1 has been shown to be required for cell proliferation. In a study by Lagger et al. [26], targeted disruption of both HDAC1 alleles in mice resulted in embryonic lethality before E10.5 due to severe proliferation and developmental defects. Proliferation rates of HDAC1-deficient embryonic stem cells were slower than HDAC1 wild-type cells and correlated with decreased cyclin-associated kinase activities and elevated levels of the cdk inhibitors p21(WAF1/CIP1) and p27(KIP1) [26]. This study provided the evidence that a HDAC may be essential for cell proliferation, disregulation of which may play a role in the process of tumorigenesis.

A high proportion of growth signals in tumor cells are transduced through amplified or mutant oncoproteins. Molecular chaperones, such as hsp90, are needed to help the correct folding and stability of such oncoproteins. Recent evidence shows that HDAC inhibitors increase the acetylation status of hsp90 where the chaperone activity of the acetylated form is inactivated, leading to increased degradation of its client proteins including Her-2/neu, AKT, c-Raf-1, mutant p53, and Bcr-Abl [27–29]. Studies by Nimmanapalli et al. reveal increased acetylation of Hsp90 by HDAC inhibitor treatment of CML cells and proteasomal degradation of wild-type Bcr-Abl as well as imatinib refractory mutant Bcr-Abl. Degradation of Bcr-Abl and c-Raf-1, Src, as well as activated AKT promoted CMC-BL cell apoptosis [28]. Furthermore, in studies involving breast cancer, exposure of cells to HDAC inhibitors increased acetylated Hsp90 leading to inefficient binding of ATP. This event also caused increased degradation of the Her2/neu oncoprotein leading to reduced growth and apoptosis [29].

Resistance to death is another hallmark of tumor cells, and HDACs have been implicated in cell survival pathways. In studies to elucidate the mechanism of action of HDAC inhibitors, both the trail-mediated external and mitochondrial apoptosis pathways have been implicated in HDAC inhibitor-mediated apoptosis. Guo and colleagues exposed leukemia cells to HDAC inhibitors and showed that mRNA and protein expression of the proapoptotic death receptors DR5 and/or DR4 increased, but the mRNA and protein of the antiapoptotic protein

c-FLIP was greatly reduced [30]. When they compared treatment HDAC inhibitor and TRAIL or HDAC inhibitor alone, they observed increased assembly of Fas-associated death domain and caspase-8, but not of c-FLIP, into the Apo-2L/TRAIL-induced death-inducing signaling complex. This increased the processing of caspase-8 and BID augmented cytosolic accumulation of the prodeath molecules such as cytochrome-c, Smac, and Omi and led to increased activity of caspase-3 and apoptosis [30]. Additional studies have shown that HDAC inhibitors induce the expression of TRAIL (Apo2L, TNFSF10) by directly activating the *TNFSF10* promoter leading to tumor-selective death signaling in acute myeloid leukemia (AML) cells and the blasts of individuals with AML [31, 32]. Furthermore, RNA interference revealed that the induction of p21, TRAIL, and differentiation are separable activities of HDAC inhibitors. In other studies, Ellis and colleagues using the Eμ-myc models demonstrated that death receptor pathway, but not the mitochondrial pathway, was dispensable in apoptosis induced by HDAC inhibitors in lymphoma cells. Ectopic expression of antiapoptotic proteins Bcl2 and BclXL prevented apoptosis induced by HDAC inhibitors in Eμ-myc lymphoma models [2]. Thus, it appears that HDAC inhibitors might operate through both external and mitochondrial pathways together or independently in a cell-type or context-dependent manner.

Protein acetylation/deacetylation processes have also been implicated in cell motility and invasion. A major target of HDACs is acetylated alpha-tubulin. Reversible acetylation of alpha-tubulin is involved in regulating microtubule stability and function. In a bid to identify the HDAC isoform (s) that regulate microtubule function, yeast two hybrid and biochemical techniques were applied to identify HDAC6 as the major tubulin deacetylase [33, 34]. Further studies showed that overexpression of HDAC6 promoted chemotactic cell movement, suggesting a potential involvement in cellular motility, invasion, and metastasis [35].

4.3 A Role for HDACs in Angiogenesis

Studies done by Pili and his colleagues have demonstrated antiangiogenic properties of HDAC inhibitors through effects on HIF-1α and VEGF [36–38]. The proangiogenic vascular endothelial growth factor (VEGF) is a transcriptional target of the hypoxic inducible factor-1 alpha (HIF-1alpha) whose expression is in turn increased by the hypoxic conditions associated with tumor angiogenesis. Direct posttranslational modifications including acetylation/deacetylation of the HIF-α protein have been shown to regulate the stability of this angiogenesis factor. To elucidate the mechanism of antiangiogenic effects of HDAC inhibitors, coimmunoprecipitation techniques were employed to demonstrate that class II HDAC4 and HDAC6 were associated with the HIF-1 alpha protein [38]. Furthermore, knockdown of HDAC4 and HDAC6 levels by RNA interference resulted in reduced HIF-1 alpha protein expression and transcriptional activity [38]. Additionally, the expression of proangiogenic factors, angiopoietins 1 and 2, was also shown to be

reduced by HDAC inhibitors. In light of the potential role played by HDACs in angiogenesis, synergistic antitumor activities were achieved by combining HDAC inhibitors and antiangiogenic agents in experimental therapy [39].

5 HDAC Inhibitors

Crystallographic studies have shown that the catalytic site of HDACs contains a zinc ion that is responsible for the transfer of the acetyl group. The lysine groups of histones reach this catalytic site through a long hydrophobic tunnel. The amino acid sequence of different isoforms of HDAC shares high homology at the active site and the lysine tunnel. Immediately adjacent to the tunnel is a shallow hydrophobic pocket; it is generally believed that this pocket is quite distinct among different classes of HDAC enzymes. Based on their chemical structure, HDAC inhibitors can be categorized into four subtypes: (1) short chain fatty acid; (2) hydroxamic acid; (3) benzamides; and (4) cyclic peptides. Despite their structural distinctiveness, HDAC inhibitors are generally considered to share common pharmacophore that include three key motifs, a "war-head" such as hydroxamic acid that interacts with the Zinc ion in the active site, which is necessary for the histone deacetylation activity [40], a hydrophobic cap that covers the entrance to the active site, and a hydrophobic spacer about 10 Å long that links the cap and the war-head (See examples in Fig. 2).

Sodium butyrate was one of the first HDAC inhibitors discovered when it increased histone acetylation in HELA and Friend erythroleukemia cells following treatment. Subsequent studies revealed suppression of histone deacetylation *in vivo* and *in vitro* by this compound [41]. Other short chain fatty acids, such as phenylbutyrate and valproic acid, have been reported as HDAC inhibitors with antitumor effects. Phenyl butyrate, although clinically approved for treating certain hematological disorders, has poor potency as an HDAC inhibitor and a short plasma half-life. Its development as an anticancer agent was hampered due to numerous side effects and poor pharmaceutical properties [42, 43].

The hydroxamic acid-based inhibitors are the most potent reported to date. Their high potency can be attributed to the strong chelating affinity of the hydroxamic acid to the zinc cation in the active site. In addition, it is proposed that the hydroxamic acid can form hydrogen bonds with active site hydrophilic residues such as tyrosine and histidine (Fig. 3). Through chemical manipulations of HDAC inhibitors, some degree of selectivity can be achieved against other zinc containing enzyme such as matrix metalloproteases. The discovery of the natural product trichostatin A (TSA), a nanomolar HDAC inhibitor, has stimulated great efforts in identifying hydroxamic acid-based inhibitors with better pharmacological properties. As a result, a broad set of inhibitors with potent HDAC activities has been reported, including suberoylanilide hydroxamic acid (SAHA), NVP-LAQ824, NVP-LBH589, pyroxamide, oxamflatin, and PXD-101. The hydroxamic acids such as TSA, NVP-LAQ824, NVP-LBH589, and SAHA inhibit HDAC isoforms

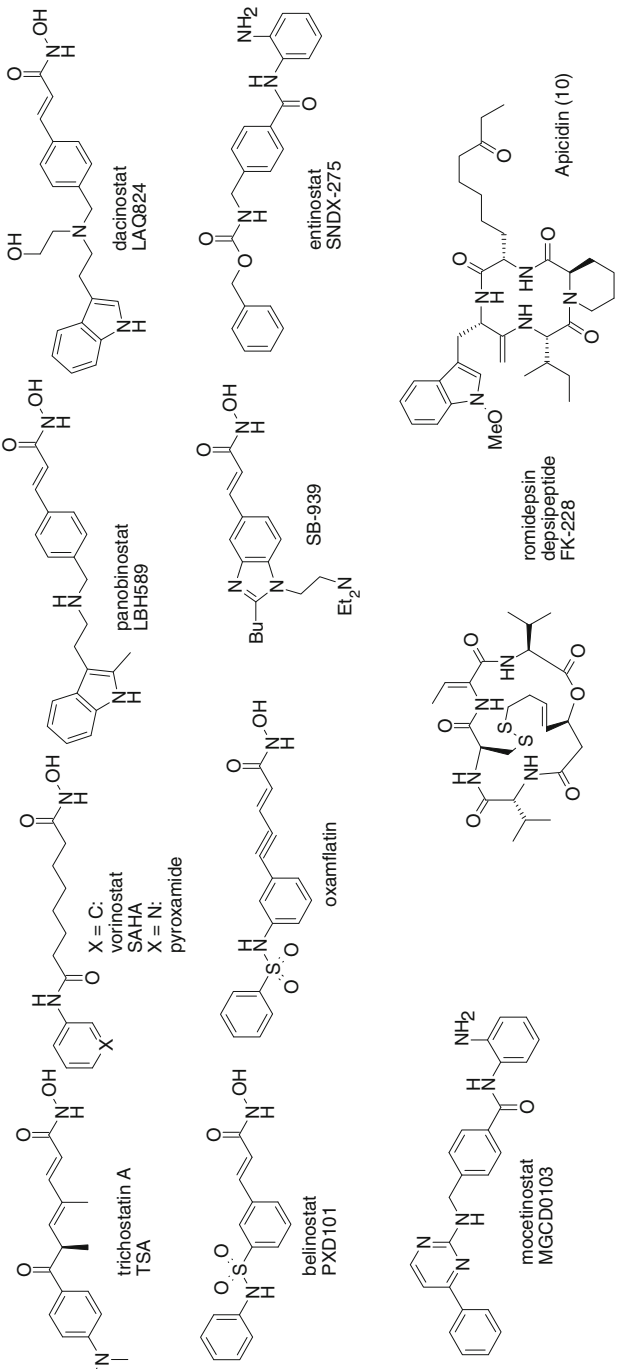


Fig. 2 Structure of representative HDAC inhibitors

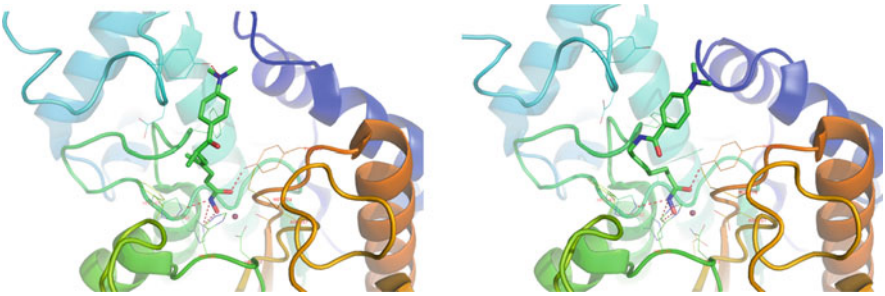


Fig. 3 X-ray structure of trichostatin and HDAC8 complex

in a nonselective manner at nanomolar concentrations and have been shown to produce potent antitumor activity *in vitro* and *in vivo*.

A number of natural product cyclic tetrapeptides have also been identified as HDAC inhibitors. Examples include trapoxin, HC-toxin, chlamydocin, FK-228, and apidicin. The cyclic tetrapeptides are also potent nanomolar inhibitors of HDACs and possess *in vitro* antiproliferative activity. However, there is insufficient *in vivo* antitumor efficacy data, most likely resulting from metabolic instability of these molecules [44]. Interestingly, the cyclic tetrapeptides appear to have some isoform selectivity. Trapoxin and apicidin selectively inhibit class I HDACs, but are inactive against some class II HDACs such as HDAC6. The cyclic depsipeptide, FK-228, is a nanomolar HDAC inhibitor with selectivity against class I enzymes [45]. The mechanism of action of HDAC inhibition for FK228 is thought to involve intracellular reduction of the disulfide bond resulting in a thiol that interacts with the zinc cation at the active site. Contrary to the other cyclic peptides, FK-228 has shown *in vivo* antitumor efficacy both in animal models and in clinical trials.

A third class of HDAC inhibitors have a benzamide group, and of this group SNDX-275 (formally known as MS-275) and MGCD003 are class 1 selective HDAC inhibitors. The mechanism of HDAC inhibition is thought to involve interaction of the two substituted amides with the zinc cation at the active site of the enzyme. Despite the relative *in vitro* low potency of HDAC inhibition by some of the benzamide derivatives such as SNDX-275 compared with the hydroxamates and the cyclic peptides, they have produced marked *in vivo* tumor efficacy in animal models [46].

6 Molecular Antitumor Effects of HDAC Inhibitors

As compounds which directly affect transcription of genes, it was originally feared that HDAC inhibitors would produce pleiotropic and global effects on gene expression. However, microarray experiments showed that treatment of cells with HDAC inhibitors modulated the expression of only about 1% of the total genes [47, 48]. These results suggested that only specific gene promoters are regulated by HDACs

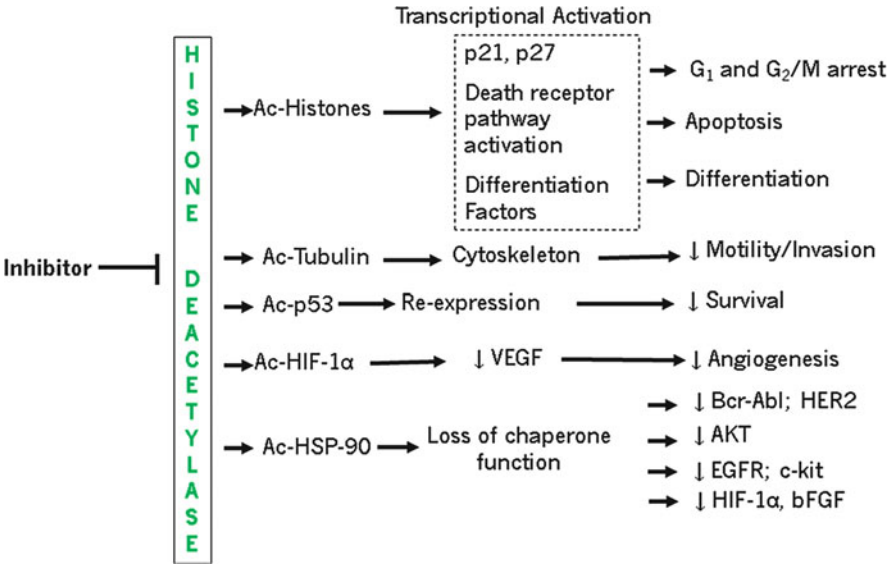


Fig. 4 Multipronged effects of HDAC inhibitors on the molecular mechanisms of cancer

and these may be selectively modulated by HDAC inhibitors. In additional gene expression profiling experiments where different cell lines were treated with structurally different HDAC inhibitors, a core set of genes that can be classified by function emerged as HDAC inhibitor-modulated genes. Products of the HDAC-modulated genes play various cell growth and regulatory roles, including cell cycle, apoptosis, signal transduction, metabolism, transcription, cytoskeletal structure, and cell adhesion (Fig. 4).

The recent characterization of the cellular acetylome resulting in the discovery of nonhistone acetylated proteins has generated great interest and has added an additional level of complexity to protein regulation. Transcription factors constitute a major class of nonhistone acetylated proteins. These include p53, E2F1, TCF, GATA1, NF-κB, HIF-1α whose functions are modulated through one or more of the following ways: binding to DNA, protein–protein interaction, cellular localization, and proteasomal degradation [22, 27, 49–52].

Recent evidence shows that HDAC inhibitors increase the acetylation status of nonnuclear proteins as well which leads to effects in growth signaling pathways. For example, treatment of tumor cells leads to increased acetylation and inactivation of the molecular chaperone hsp90, resulting in the destabilization of its client proteins, including Her-2/neu, AKT, c-Raf-1, mutant p53, and Bcr-Abl [27–29]. Studies by Bhalla and colleagues indicate that treatment of CML cells with the HDAC inhibitor NVP-LAQ824 increases acetylation of Hsp90 and directs proteasomal degradation of wild-type Bcr-Abl as well as imatinib refractory mutant Bcr-Abl. This Hsp90-mediated proteasomal degradation of Bcr-Abl in conjunction with degradation of c-Raf-1, Src, and AKT appears to promote CMC-BL cell

apoptosis. Similar studies in Her2+ breast cancer cells have led to the degradation of the Her2/neu oncoprotein [29]. These results have led to clinical trials of HDAC inhibitors as single agents or in combination in CML and Her2+ breast cancer, respectively.

7 Clinical Experience with HDAC Inhibitors

HDAC inhibitors belonging to different structural classes have entered clinical development and a number have demonstrated efficacy in at least one tumor type at tolerable doses. Since no mutant versions of HDACs have been associated with the molecular epidemiology of any particular tumor types, most of the early clinical trials with HDAC inhibitors have been "all comer" types. Early clinical results indicate that HDAC inhibitors are well tolerated in the clinic and some anticancer activity has been observed.

8 Activities of HDAC Inhibitors in Hematological Malignancies

The first clinical efficacy with a HDAC inhibitor was observed when an acute promyelocytic leukemia patient who had experienced multiple relapses was treated with retinoic acid combination with phenylbutyrate. In this study, the APL patient had proved clinically resistant to all-trans retinoic acid (ATRA) as a single agent but a combination of ATRA with phenylbutyrate produced a complete clinical and cytogenetic remission [18]. HDAC inhibitors are proving effective against a number of hematological malignancies in the clinic as described below.

8.1 Cutaneous T-Cell Lymphoma

The first such HDAC inhibitor-sensitive tumor type to emerge from clinical trials was cutaneous T-cell lymphoma (CTCL). In one of the earliest phase 1 trials, three patients with CTCL who had failed all other therapy were found to exhibit dramatic responses to the depsipeptide HDAC inhibitor FK-228 [53]. Subsequently, the phase 1 trials of hydroxamate HDAC inhibitors SAHA (Vorinostat) and LBH589 (panobinostat) also demonstrated high response rates (>50%) in CTCL [54, 55]. Based on the promising early results, additional phase II results were conducted with all three compounds in larger numbers of patients and the phase further confirmed the phase 1 results. Based on these phase II data, SAHA (Vorinostat) was approved by the US Food and Drug Administration for the

treatment of the cutaneous manifestations of CTCL in 2008 [56]. In 2009, the FDA also approved the depsipetide FK-227 (Rhomidepsin) for the tratment of CTCL. The mechanism by which HDAC inhibitors induce antitumor activity in CTCL is currently unknown. However, a number of studies have shown that apoptosis in CTCL cells is associated with the mitochondrial death pathway [57]. In one study, LBH589 was shown to inhibit the mRNA and protein levels of HDAC7 and induces expression and translocation of Nur77 to the mitochondria where Nur77 converts death resistance protein Bcl-2 into a killer protein, promoting cell death of cultured and patient-derived human CTCL cells [58]. Additionally, the JAK/STAT pathway has been associated with resistance of CTCL cells to SAHA [59]. Interestingly, previous, independent studies have implicated the Jak/STAT pathway in CTCL. Mention that also romidepsin has been recently FDA-approved for CTCL treatment.

8.2 Hodgkins Lymphoma

Two structurally different HDAC inhibitors have shown very promising antitumor clinical efficacy in Hodgkins lymphoma. In a phase IA/II multicenter clinical study, 13 Hodgkins lymphoma patients who had failed up to six previous therapies, including bone marrow transplantation, were treated with escalating doses of panobinostat and followed by metabolic (PET) or computer tomographic (CT) scanning. In that study, complete and partial responses were obtained in close to 60% of this refractory population. In another phase II study, where Hodgkins lymphoma patients were treated with the class I selective benzamide HDAC inhibitor MGCD103, 8 out of 22 patients enrolled showed objective responses [60]. The mechanism underlying such single agent antitumor activity in Hodgkins lymphoma is not known. However, our unpublished data indicate that the inhibitors might be interfering with the JAK/STAT pathway in this disease also.

8.3 AML/MDS and Other Acute Hemeatological Malignancies

Following encouraging responses seen with some HDAC inhibitors and the rationale based on a possible role in AML1/ETO and other acute leukemias, a number of phase I/II clinical studies were undertaken in a number of acute hematological malignancies [61–63]. In a study with panobinostat reported by Ottmann et al. (2008), where patients with various hematological malignancies were enrolled, dose-dependent efficacy was observed in patients with advanced refractory AML. No antileukemic activity was observed in patients treated below 30 mg 3x/wk oral dose; however, a number of partial and complete responses were observed in patients treated with 40 mg 3x/wk and above. A yet inexplicable phenomenon

where reduction in bone marrow blasts continued months after cessation of drug treatment was observed in some patients [62]. Other HDAC inhibitors have also demonstrated activity in acute leukemias. For example, in a phase 1 study with vorinostat to evaluate the safety and activity in patients with relapsed or refractory leukemias or myelodysplastic syndromes (MDS), 7 of 41 patients had hematologic improvement responses, including two complete responses and two complete responses with incomplete blood count recovery [63].

The mechanism of action of HDAC inhibitors in acute leukemias is not known. However, preclinical studies are yielding some clues. Bali et al. [64] reported that mutant FLT3 kinase, known to be a leukemic oncogene, is a client protein of hsp90 that is degraded in leukemia cells by HDAC inhibitor treatment. Other studies have also implicated mechanisms involving the Ezh2 polycomp complex and HOXA9 and MEIS1 oncogenic factors [65].

8.4 Multiple Myeloma

Based on high sensitivity of myeloma cells to HDAC inhibitors, a number of multiple myeloma patients have been treated. However, with the exception of panobinostat where one durable response was seen with single agent treatment in a patient highly refractory to previous multiple treatments, none of the other HDAC inhibitors have demonstrated single agent activity in multiple myeloma. Since myeloma cells are very sensitive to HDAC inhibitors *in vitro*, the apparent lack of single agent activity in this indication might be due to decreased bone marrow exposure or resistance of myeloma cells in the bone marrow microenvironment. In combination with standard therapy however, greater success was achieved with HDAC inhibitors in multiple myeloma. In a phase I/II clinical study to investigate the safety and efficacy of various combination doses of panobinostat and the current proteasome inhibitor standard of care bortezomib, 26 responses were observed in 36 evaluable patients (4 complete responses, 2 very good partial responses, 16 partial responses, and 4 minor responses) [66]. Patients enrolled in this combination therapy had received up to six prior lines of therapy, including patients not having responded to or who are insensitive to bortezomib at study entry (8 of 13 patients responded in this bortezomib population). Similar results obtained with vorinostat and rhomidepsin further strengthen the potential of combining HDA inhibitors with proteasome inhibitors in myeloma therapy. In a phase 1 vorinostat + bortezomib combination trial in myeloma patients who have received a median of seven prior regimens, an overall response rate of 42% was obtained. These included three partial responses among nine bortezomib refractory patients [67]. Based on these promising early results with HDAC inhibitor and proteasome inhibitor combinations in multiple myeloma, a number of phase II and phase III studies have been initiated with this combination (clinicaltrials.gov). Although the exact mechanism accounting for the enhanced efficacy of HDAC and proteasome inhibitor combinations is not clear, recent preclinical findings indicate that alternative protein

processing mechanisms, such as the aggresomes, are upregulated in myeloma cells that have become refractory to HDAC inhibitors [68]. Since myeloma cells make extremely high levels of proteins, efficient functioning of the protein processing machinery that prevents accumulation of cytotoxic misfolded proteins may be required for cell survival. This phenomenon may account for the efficacy of proteasome inhibitors in myeloma and the upregulation of possible compensatory protein processing mechanisms such as the aggresomes in relapsed/refractory cells. Recent reports indicate that HDAC6 is required for the proper formation and functioning of aggresomes [69]. Thus, combining HDAC inhibitors with proteasome inhibitors may be a two-pronged attack on two protein processing mechanisms necessary for myeloma cell survival.

9 Activity of HDAC Inhibitors in Solid Tumors

Although preclinical efficacy has been achieved with HDAC inhibitors in a number of solid tumor xenograft models, clinical efficacy as single agents has been less promising. In early phase 1 trials, one or two partial responses in different tumor types were achieved. The most promising single agent activity achieved with an HDAC inhibitor was in mesothelioma. In a phase 1 study with vorinostat where a number of previously treated mesothelioma patients were enrolled, among other tumor types, some partial responses were achieved [70]. These encouraging results led to a currently ongoing phase III randomized, double-blind, placebo-controlled trial of oral SAHA in patients with advanced malignant pleural mesothelioma previously treated with systemic chemotherapy. This study is still enrolling patients.

The more common antitumor response seen in solid tumor patients with single agent HDAC inhibitor treatment is tumor stabilization. Solid tumor types where stabilization of disease has been reported include mesothelioma, head and neck cancer, NSCLC (Pivanex, vorinostat) and melanoma (MS-27-275), breast cancer, renal cell carcinoma, and prostate cancer (panobinostat). Since HDAC inhibitors clearly exhibit dose-dependent apoptotis in tumor cell lines *in vitro* and in tumor xenografts, whether or not it is the inability to achieve enough intratumoral exposure to the drugs that accounts for the general lack of clinical regression of solid tumors is currently a matter of investigation. There is, however, a great potential for combination therapy with HDAC inhibitors in solid tumors. For example, HDAC inhibitors have been shown to synergize with numerous standard chemotherapeutic or targeted therapies, including antimetabolites, antibiotics, alkylating agents, plant alkyloids, topoisomerase inhibitors, immunomodulators, corticosteroids, and radiation therapy, tyrosine kinase inhibitors, or nuclear receptor antagonists in preclinical experiements ([7, 71] and references therein). Current understanding of the basis for such synergy revolves around the effect of HDAC inhibition on inducing DNA damage responses in cells that lack mechanisms to manage them, inhibition of antiapoptotic and survival mechanisms induced in

tumor cells to resist chemotherapy, or combinatorial effects of targeting multiple aspects of the hallmarks of cancer in combination therapy [72–75].

Combination of HDAC inhibitors in the clinical setting in solid tumors has begun to show some initial promising results. Notably, a phase 1 study combining vorinostat with paclitaxel and carboplatin in patients with advanced solid tumors resulted in partial responses in 11 of 25 patients, 10 of whom are with nonsmall cell lung cancer [76]. With the reported increased acetylation and destabilization of HIF1-α leading to decreased angiogenesis, vorinostat also demonstrated promising activities in combination with paclitaxel and bevacizumab in patients with renal cell carcinoma (53% overall response rate), a tumor type that has shown good sensitivity to antiangiogenic agents in the clinic. Panobinostat has also demonstrated promising activity in combination with standard agents. In a phase I dose-finding trial of intravenous (i.v.) panobinostat with docetaxel in patients (pts) with castration-resistant prostate cancer (CRPC), 4 of 38 patients achieved partial responses and 4 of 7 patients showed a >90% PSA decrease from baseline by cycle 4 [77]. Encouraging results are also being obtained with panobinostat in combination with herceptin or with vorinostat and tamoxifen in breast cancer [78]. HDAC inhibitors have been shown to cause degradation of the androgen receptor, a major target in prostate cancer, estrogen, and Her2/neu breast cancer targets for tamoxifen and herceptin respectively, as well as cause increased acetylation and stabilization of microtubules, the target of docetaxel. How all these effects might combine to induce therapeutic synergy in the relevant tumor types is currently being investigated. Similarly, a sequence-specific combination of the HDAC inhibitor Valproic acid and the topoisomerase II inhibitor epirubicin, patients with advanced solid tumors resulted in partial responses in 9 of 41 patients and stable disease in a further 16 of the 41 patients. In light of the encouraging results being obtained with HDAC inhibitors in combination therapy in solid tumors, a number of phase II studies are currently in progress in a variety of tumor types to further confirm and expand the early results. Table 1 is a list of the different HDAC inhibitors currently under development for cancer therapy.

10 Challenges in Developing HDAC Inhibitors for Anticancer Therapy and Future Outlook

Although encouraging data are emerging preclinically and clinically with HDAC inhibitors as antitumor therapy, several challenges remain in their development. One major challenge regards that little is known about the mechanism of action of these drugs. It is not known which of the 11 HDAC isoforms or which subsets are the most relevant for the activity of the inhibitors in antitumor therapy. Neither is it known which HDAC substrates play the most critical roles in the tumors being treated. Thus, a lot of the clinical trials with HDAC inhibitors are relying on trial and error and serendipity in selecting indications and patients. Similarly, as no

mutations or aberrations of specific HDAC isoforms have been associated with particular tumor types, there are very little clues for stratifying patients to deepen the therapeutic effect of these drugs. Because of the aforementioned issues, it is also not clear whether isoform selective inhibitors or pan-deacetylase inhibitors will be most beneficial for antitumor therapy. Although HDAC inhibitors are fairly well tolerated in cancer therapy, thrombocytopenia and fatigue remain the most common adverse events associated with their use. Other toxicities that have been reported include nausea, vomiting, anorexia, diarrhea, and neutropenia. All these adverse events have been seen with both pan-deacetylase and isoform-selective inhibitors. Thus, it is not known which class of inhibitors will produce the most beneficial risk-benefit profile. Another challenge facing the development of HDAC inhibitors is the lack of a reliable efficacy biomarker that could be used to calibrate PK/PD relationship to identify the most beneficial dosing levels and schedules. Also, since the most proximal and principal substrate of HDAC inhibitors happen to be histones, increased histone acetylation has been used as a marker of cell exposure to these drugs. However, the relationship between histone acetylation and antitumor efficacy has eluded any calibration. The foregoing challenges notwithstanding, HDAC inhibitors offer a major hope for anticancer therapy due to their multi-pronged attack on multiple mechanisms underlying the hallmarks of cancer as described in this chapter. Tackling these developmental challenges will further enable them fulfill their tremendous therapeutic promise.

Acknowledgment I would like to thank Lei Jiang of the China Novartis Institute for Biomedical Research for assistance in drawing chemical structures and reviewing the chemistry sections.

References

1. Hager GL, McNally JG, Misteli T (2009) Transcription dynamics. Mol Cell 5:741–753
2. Ellis L, Atadja PW, Johnstone RW (2009) Epigenetics in cancer: targeting chromatin modifications. Mol Cancer Ther 8:1409–1420
3. Smith CL (2008) A shifting paradigm: histone deacetylases and transcriptional activation. Bioessays 30:15–24
4. Glozak MA, Sengupta N, Zhang X, Seto E (2005) Acetylation and deacetylation of non-histone proteins. Gene 363:15–23
5. Kim SC, Sprung R, Chen Y et al (2006) Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. Mol Cell 23:607–618
6. Kouzarides T (2007) Chromatin modifications and their function. Cell 128:693–705
7. Bolden JE, Peart MJ, Johnstone RW (2006) Anticancer activities of histone deacetylase inhibitors. Nat Rev Drug Discov 5:769–784
8. Minucci S, Pelicci PG (2006) Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. Nat Rev Cancer 6:38–51
9. Gao L, Cueto MA, Asselbergs F, Atadja P (2002) Cloning and functional characterization of HDAC11, a novel member of the human histone deacetylase family. J Biol Chem 277 (28):25748–25755
10. Gregoretti IV, Lee YM, Goodson HV (2004) Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. J Mol Biol 33:17–31

11. Spange S, Wagner T, Heinzel T, Krämer OH (2009) Acetylation of non-histone proteins modulates cellular signalling at multiple levels. Int J Biochem Cell Biol 41:185–198
12. Smith KT, Workman JL (2009) Introducing the acetylome. Nat Biotechnol 27:917–919
13. Atadja P (2009) Development of the pan-DAC inhibitor panobinostat (LBH589): successes and challenges. Cancer Lett 280:233–241
14. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100:57–70
15. Zhu P, Martin E, Mengwasser J, Schlag P, Janssen KP, Gottlicher M (2004) Induction of HDAC2 expression upon loss of APC in colorectal tumorigenesis. Cancer Cell 5:455–463
16. Gui CY, Ngo L, Xu WS, Richon VM, Marks PA (2004) Histone deacetylase (HDAC) inhibitor activation of p21WAF1 involves changes in promoter-associated proteins, including HDAC1. Proc Natl Acad Sci USA 101:1241–1246
17. Bhalla KN (2005) Epigenetic and chromatin modifiers as targeted therapy of hematologic malignancies. J Clin Oncol 23:3971–3993
18. Warrell RP Jr, He LZ, Richon V, Calleja E, Pandolfi PP (1998) Therapeutic targeting of transcription in acute promyelocytic leukemia by use of an inhibitor of histone deacetylase. J Natl Cancer Inst 90:1621–1625
19. Wang Y, Wang SY, Zhang XH et al (2007) FK228 inhibits Hsp90 chaperone function in K562 cells via hyperacetylation of Hsp70. Biochem Biophys Res Commun 356:998–1003
20. Bereshchenko OR, Gu W, Dalla-Favera R (2002) Acetylation inactivates the transcriptional repressor BCL6. Nat Genet 32:606–613
21. Li M, Luo J, Brooks CL, Gu W (2002) Acetylation of p53 inhibits its ubiquitination by Mdm2. J Biol Chem 277:50607–50611
22. Gu W, Roeder RG (1997) Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell 90:595–606
23. Luo J, Li M, Tang Y, Laszkowska M, Roeder RG, Gu W (2002) Acetylation of p53 augments its site-specific DNA binding both *in vitro* and *in vivo*. Proc Natl Acad Sci 101:2259–2264. Li M, Luo J, Brooks CL, Gu W. (2002) Acetylation of p53 inhibits its ubiquitination by Mdm2. J Biol Chem 277:50607–50611
24. Tang Y, Zhao W, Chen Y, Zhao Y, Gu W (2008) Acetylation is indispensable for p53 activation. Cell 133:612–626
25. Zhang HS, Gavin M, Dahiya A, Postigo AA, Ma D, Luo RX, Harbour JW, Dean DC (2000) Exit from G1 and S phase of the cell cycle is regulated by repressor complexes containing HDAC-Rb-hSWI/SNF and Rb-hSWI/SNF. Cell 101(1):79–89
26. Lagger G, O'Carroll D, Rembold M, Khier H, Tischler J, Weitzer G, Schuettengruber B, Hauser C, Brunmeir R, Jenuwein T, Seiser C (2002) Essential function of histone deacetylase 1 in proliferation control and CDK inhibitor repression. EMBO J 21(11):2672–2681
27. Yu C, Subler M, Rahmani M et al (2003) Induction of apoptosis in BCR/ABL+ cells by histone deacetylase inhibitors involves reciprocal effects on the RAF/MEK/ERK and JNK pathways. Cancer Biol Ther 2:544–551
28. Nimmanapalli R, Fuino L, Bali P et al (2003) Histone deacetylase inhibitor LAQ824 both lowers expression and promotes proteasomal degradation of Bcr-Abl and induces apoptosis of imatinib mesylate-sensitive or -refractory chronic myelogenous leukemia-blast crisis cells. Cancer Res 63:5126–5135
29. Fuino L, Bali P, Wittmann S, Donapaty S, Guo F, Yamaguchi H, Wang HG, Atadja P, Bhalla K (2003) Histone deacetylase inhibitor LAQ824 down-regulates Her-2 and sensitizes human breast cancer cells to trastuzumab, taxotere, gemcitabine, and epothilone B. Mol Cancer Ther 2:971–984
30. Guo F, Sigua C, Tao J, Bali P, George P, Li Y, Wittmann S, Moscinski L, Atadja P, Bhalla K (2004) Cotreatment with histone deacetylase inhibitor LAQ824 enhances Apo-2L/tumor necrosis factor-related apoptosis inducing ligand-induced death inducing signaling complex activity and apoptosis of human acute leukemia cells. Cancer Res 64:2580–2589
31. Nebbioso A, le Clarke N, Voltz E, Germain E, Ambrosino C, Bontempo P, Alvarez R, Schiavone EM, Ferrara F, Bresciani F, Weisz A, de Lera AR, Gronemeyer H, Altucci L

(2005) Tumor-selective action of HDAC inhibitors involves TRAIL induction in acute myeloid leukemia cells. Nat Med 11:77–84
32. Insinga A, Monestiroli S, Ronzoni S, Gelmetti V, Marchesi F, Viale A, Altucci L, Nervi C, Minucci S, Pelicci PG (2004) Inhibitors of histone deacetylases induce tumor-selective apoptosis through activation of the death receptor pathway. Nat Med 11:71–76
33. Hubbert C, Guardiola A, Shao R et al (2002) HDAC6 is a microtubule-associated deacetylase. Nature 417:455–458
34. Zhang X, Yuan Z, Zhang Y et al (2007) HDAC6 modulates cell motility by altering the acetylation level of cortactin. Mol Cell 27:197–213
35. Tran AD, Marmo TP, Salam AA et al (2007) HDAC6 deacetylation of tubulin modulates dynamics of cellular adhesions. J Cell Sci 120:1469–1479
36. Ellis L, Hammers H, Pili R (2009) Targeting tumor angiogenesis with histone deacetylase inhibitors. Cancer Lett 280:145–153
37. Qian DZ, Wang X, Kachhap SK, Kato Y, Wei Y, Zhang L, Atadja P, Pili R (2004) The histone deacetylase inhibitor NVP-LAQ824 inhibits angiogenesis and has a greater antitumor effect in combination with the vascular endothelial growth factor receptor tyrosine kinase inhibitor PTK787/ZK222584. Cancer Res 64:6626–6634
38. Qian DZ, Kachhap SK, Collis SJ, Verheul HM, Carducci MA, Atadja P, Pili R (2006) Class II histone deacetylases are associated with VHL-independent regulation of hypoxia-inducible factor 1 alpha. Cancer Res 66:8814–8821
39. Verheul HM, Salumbides B, Van Erp K, Hammers H, Qian DZ, Sanni T, Atadja P, Pili R (2008) Combination strategy targeting the hypoxia inducible factor-1 alpha with mammalian target of rapamycin and histone deacetylase inhibitors. Clin Cancer Res 14(11):3589–3597
40. Finnin MS, Donigian JR, Cohen A, Richon VM, Rifkind RA, Marks PA, Breslow R, Pavletich NP (1999) Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. Nature 401:188–193
41. Boffa LC, Vidali G, Mann RS, Allfrey VG (1978) J Biol Chem. Suppression of histone deacetylation *in vivo* and *in vitro* by sodium butyrate 253:3364–3366
42. Gore SD, Carducci MA (2000) Modifying histones to tame cancer: clinical development of sodium phenylbutyrate and other histone deacetylase inhibitors. Expert Opin Investig Drugs 9:2923–2934
43. Santini V, Gozzini A, Scappini B, Grossi A, Rossi Ferrini P (2001) Searching for the magic bullet against cancer: the butyrate saga. Leuk Lymphoma 42:275–289
44. Remiszewski SW (2003) The discovery of NVP-LAQ824: from concept to clinic. Curr Med Chem 10:2393–2402
45. Nakajima H, Kim YB, Terano H, Yoshida M, Horinouchi S (1998) FR901228, a potent antitumor antibiotic, is a novel histone deacetylase inhibitor. Exp Cell Res 241:126–133
46. Saito A, Yamashita T, Mariko Y, Nosaka Y, Tsuchiya K, Ando T, Suzuki T, Tsuruo T, Nakanishi O (1999) A synthetic inhibitor of histone deacetylase, MS-27-275, with marked *in vivo* antitumor activity against human tumors. Proc Natl Acad Sci USA 96:4592–4597
47. Van Lint C, Emiliani S, Verdin E (1996) The expression of a small fraction of cellular genes is changed in response to histone hyperacetylation. Gene Expr 5:245–253
48. Glaser KB, Staver MJ, Waring JF, Stender J, Ulrich RG, Davidsen SK (2003) Gene expression profiling of multiple histone deacetylase (HDAC) inhibitors: defining a common gene set produced by HDAC inhibition in T24 and MDA carcinoma cell lines. Mol Cancer Ther 2:151–163
49. Boyes J, Byfield P, Nakatani Y, Ogryzko V (1998) Regulation of activity of the transcription factor GATA-1 by acetylation. Nature 396:594–598
50. Waltzer L, Bienz M (1998) Drosophila CBP represses the transcription factor TCF to antagonize Wingless signalling. Nature 395:521–525
51. Martínez-Balbás MA, Bauer UM, Nielsen SJ, Brehm A, Kouzarides T (2000) Regulation of E2F1 activity by acetylation. EMBO J 19:662–671
52. Marzio G, Wagener C, Gutierrez MI, Cartwright P, Helin K, Giacca M (2000) E2F family members are differentially regulated by reversible acetylation. J Biol Chem 275:10887–10892

53. Piekarz RL, Robey R, Sandor V, Bakke S, Wilson WH, Dahmoush L, Kingma DM, Turner ML, Altemus R, Bates SE (2001) Inhibitor of histone deacetylation, depsipeptide (FR901228), in the treatment of peripheral and cutaneous T-cell lymphoma: a case report. Blood 98:2865–2868
54. Duvic M, Talpur R, Ni X, Zhang C, Hazarika P, Kelly C, Chiao JH, Reilly JF, Ricker JL, Richon VM, Frankel SR (2007) Phase 2 trial of oral vorinostat (suberoylanilide hydroxamic acid, SAHA) for refractory cutaneous T-cell lymphoma (CTCL). Blood 109:31–39
55. Ellis L, Pan Y, Smyth GK, George DJ, McCormack C, Williams-Truax R, Mita M, Beck J, Burris H, Ryan G, Atadja P, Butterfoss D, Dugan M, Culver K, Johnstone RW, Prince HM (2008) Histone deacetylase inhibitor panobinostat induces clinical responses with associated alterations in gene expression profiles in cutaneous T-cell lymphoma. Clin Cancer Res 14:4500–4510
56. Mann BS, Johnson JR, Cohen MH, Justice R, Pazdur R (2007) FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. Oncologist 12 (10):1247–1252
57. Zhang C, Richon V, Ni X, Talpur R, Duvic M (2005) Selective induction of apoptosis by histone deacetylase inhibitor SAHA in cutaneous T-cell lymphoma cells: relevance to mechanism of therapeutic action. J Invest Dermatol 125:1045–1052
58. Chen J, Fiskus W, Eaton K, Fernandez P, Wang Y, Rao R, Lee P, Joshi R, Yang Y, Kolhe R, Balusu R, Chappa P, Natarajan K, Jillella A, Atadja P, Bhalla KN (2009) Cotreatment with BCL-2 antagonist sensitizes cutaneous T-cell lymphoma to lethal action of HDAC7-Nur77-based mechanism. Blood 113:4038–4048
59. Fantin VR, Loboda A, Paweletz CP, Hendrickson RC, Pierce JW, Roth JA, Li L, Gooden F, Korenchuk S, Hou XS, Harrington EA, Randolph S, Reilly JF, Ware CM, Kadin ME, Frankel SR, Richon VM (2008) Constitutive activation of signal transducers and activators of transcription predicts vorinostat resistance in cutaneous T-cell lymphoma. Cancer Res 68: 3785–3794
60. Dickinson M, Ritchie D, DeAngelo DJ, Spencer A, Ottmann OG, Fischer T, Bhalla KN, Liu A, Parker K, Scott JW, Bishton M, Prince HM (2009) Preliminary evidence of disease response to the pan deacetylase inhibitor panobinostat (LBH589) in refractory Hodgkin Lymphoma. Br J Haematol 147:97–101
61. Giles F, Fischer T, Cortes J, Garcia-Manero G, Beck J, Ravandi F, Masson E, Rae P, Laird G, Sharma S, Kantarjian H, Dugan M, Albitar M, Bhalla KA (2006) Phase I study of intravenous LBH589, a novel cinnamic hydroxamic acid analogue histone deacetylase inhibitor, in patients with refractory hematologic malignancies. Clin Cancer Res 12:4628–4635
62. Otmann OG, Spencer A, Prince MH, Bhalla KN, Fischer T, Liu A, Parker K, Jalaluddin M, Laird G, Woo M, Scott JW, DeAngelo DJ (2008) Phase IA/II study of oral panobinostat (LBH589), a novel pan-deacetylase inhibitor (DACi) demonstrating efficacy in patients with advanced hematologic malignancies. American Society of Hematology annual conference 112, Abstract 958
63. Garcia-Manero G, Yang H, Bueso-Ramos C, Ferrajoli A, Cortes J, Wierda WG, Faderl S, Koller C, Morris G, Rosner G, Loboda A, Fantin VR, Randolph SS, Hardwick JS, Reilly JF, Chen C, Ricker JL, Secrist JP, Richon VM, Frankel SR, Kantarjian HM (2008) A phase 1 study of the histone deacetylase inhibitor vorinostat (suberoylanilide hydroxamic acid [SAHA]) in patients with advanced leukemias and myelodysplastic syndromes. Blood 111 (3):1060–1066
64. Bali P, George P, Cohen P, Tao J, Guo F, Sigua C, Vishvanath A, Scuto A, Annavarapu S, Fiskus W, Moscinski L, Atadja P, Bhalla K (2004) Superior activity of the combination of histone deacetylase inhibitor LAQ824 and the FLT-3 kinase inhibitor PKC412 against human acute myelogenous leukemia cells with mutant FLT-3. Clin Cancer Res 10:4991–4997
65. Fiskus W, Wang Y, Sreekumar A, Buckley KM, Shi H, Jillella A, Ustun C, Rao R, Fernandez P, Chen J, Balusu R, Koul S, Atadja P, Marquez VE, Bhalla KN (2009) Combined epigenetic

therapy with the histone methyltransferase EZH2 inhibitor 3-deazaneplanocin A and the histone deacetylase inhibitor panobinostat against human AML cells. Blood 114:2733–2743
66. San-Miguel JF, Sezer O, Siegel D, Guenther A, Mateos M-V, Blade J, Prosser I, Cavo M, Boccadoro M, Goebeler M, Bengoudifa BR, Hazell K, Klebsattel M, Bourquelot PM, Anderson KC (2009) A phase Ib, multicenter, open-label, dose-escalation study of oral panobinostat (LBH589) and I.V. bortezomib in patients with relapsed multiple myeloma. American Hematological Society meeting X, Abstract 3852
67. Badros A, Burger AM, Philip S, Niesvizky R, Kolla SS, Goloubeva O, Harris C, Zwiebel J, Wright JJ, Espinoza-Delgado I, Baer MR, Holleran JL, Egorin MJ, Grant S (2009) Phase I study of vorinostat in combination with bortezomib for relapsed and refractory multiple myeloma. Clin Cancer Res 15:5250–5257
68. Hideshima T, Bradner JE, Wong J, Chauhan D, Richardson P, Schreiber SL, Anderson KC (2005) Small-molecule inhibition of proteasome and aggresome function induces synergistic antitumor activity in multiple myeloma. Proc Natl Acad Sci USA 24:8567–8572
69. Kawaguchi Y, Kovacs JJ, McLaurin A, Vance JM, Ito A, Yao TP (2003) The deacetylase HDAC6 regulates aggresome formation and cell viability in response to misfolded protein stress. Cell 115:727–738
70. Krug LM, Curley T, Schwartz L, Richardson S, Marks P, Chiao J, Kelly WK (2006) Potential role of histone deacetylase inhibitors in mesothelioma: clinical experience with suberoylanilide hydroxamic acid. Clin Lung Cancer 7:257–261
71. Carraway HE, Gore SD (2007) Addition of histone deacetylase inhibitors in combination therapy. J Clin Oncol 25(15):1955–1956
72. Lee JH, Park JH, Jung Y, Kim JH, Jong HS, Kim TY, Bang YJ (2006) Histone deacetylase inhibitor enhances 5-fluorouracil cytotoxicity by down-regulating thymidylate synthase in human cancer cells. Mol Cancer Ther 5(12):3085–3095
73. Ocker M, Alajati A, Ganslmayer M et al (2005) The histone-deacetylase inhibitor SAHA potentiates proapoptotic effects of 5-fluorouracil and irinotecan in hepatoma cells. J Cancer Res Clin Oncol 131:385–394
74. Maiso P, Carvajal-Vergara X, Ocio EM, López-Pérez R, Mateo G, Gutiérrez N, Atadja P, Pandiella A, San Miguel JF (2006) The histone deacetylase inhibitor LBH589 is a potent antimyeloma agent that overcomes drug resistance. Cancer Res 66(11):5781–5789
75. Mitsiades N, Mitsiades CS, Richardson PG, McMullan C, Poulaki V, Fanourakis G, Schlossman R, Chauhan D, Munshi NC, Hideshima T, Richon VM, Marks PA, Anderson KC (2003) Molecular sequelae of histone deacetylase inhibition in human malignant B cells. Blood 101 (10):4055–4062
76. Ramalingam SS, Parise RA, Ramanathan RK, Lagattuta TF, Musguire LA, Stoller RG, Potter DM, Argiris AE, Zwiebel JA, Egorin MJ, Belani CP (2007) Phase I and pharmacokinetic study of vorinostat, a histone deacetylase inhibitor, in combination with carboplatin and paclitaxel for advanced solid malignancies. Clin Cancer Res 13:3605–3610
77. Rathkpf D, Wong BY, Ross RW, Anand A, Tanaka E, Woo MM, Hu J, Dzik-Jurasz A, Yang W, Scher HI (2010) A phase I study of oral panobinostat alone and in combination with docetaxel in patients with castration-resistant prostate cancer. Cancer Chemother Pharmacol 66:181–189
78. Capone M, Conte PF, Amadori D, Pronzato P, Wardley A, Dey H, Dzik-Jurasz A, Maclean N, McBride K, Fandi A (2009) Phase I trial of panobinostat (LBH589) in combination with trastuzumab in pretreated HER2-positive metastatic breast cancer: preliminary safety and tolerability results San Antonio Breast Cancer Symposium, Abstract #6101
79. Mehnert JM, Kelly WK (2007) Histone deacetylase inhibitors: biology and mechanism of action. Cancer J 13:23–29

Epigenetic Mechanisms in Acute Myeloid Leukemia

Antoine H.F.M. Peters and Juerg Schwaller

Abstract Acute leukemia is characterized by clonal expansion of hematopoietic stem and progenitor cells with blocked differentiation. Clinical and experimental evidences suggest that acute myeloid leukemia (AML) is the product of several functionally cooperating genetic alterations including chromosomal translocations leading to expression of leukemogenic fusion proteins. Several AML-associated lesions target chromatin regulators like histone methyltransferases or histone acetyltransferases, including mixed-lineage leukemia 1 (MLL1) or CREB bindung protein/p300. Molecular and biochemical studies start to provide useful insights into the mechanisms of targeting and mode-of-action of such leukemogenic fusion proteins resulting in aberrant gene expression programs and AML. Chromatin modulating mechanisms are also mediating the transforming activity of key drivers of leukemogenesis by aberrant recruitment of corepressors. Recent large-scale screening efforts demonstrated that both aberrant DNA promoter methylation and aberrantly expressed microRNAs play an important role in the pathogenesis of AML as well. Current efforts to therapeutically exploit the potential reversibility of epigenetic mechanisms are focused on small molecules that inhibit DNA methyltransferases or histone deacetylases. Several phase I/II clinical trials using such compounds have reported promising, but mostly transient, clinical responses. This underscores the need to further dissect the molecular players of epigenetic mechanisms driving induction, maintenance, and potential reversibility of leukemic state to develop efficient and long-lasting targeted therapeutic strategies.

A.H.F.M. Peters
Friedrich Miescher Institute for Biomedical Research, CH-4058 Basel, Switzerland
e-mail: antoine.peters@fmi.ch

J.Schwaller
University Hospital Basel, Department Biomedicine, CH-4031 Basel, Switzerland
e-mail: j.schwaller@unibas.ch

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_10,

1 Introduction

Acute leukemia is characterized by clonal expansion of hematopoietic stem and progenitor cells with blocked differentiation. Clinical and experimental evidences suggest that acute myeloid leukemia (AML) is the product of several functionally cooperating genetic alterations [1]. Best-studied examples are fusion proteins (mostly resulting from balanced chromosomal translocations) and point mutations in protein tyrosine kinases or classical cellular signaling mediators such as FLT3 or RAS that support proliferation and survival of hematopoietic cells. For example, expression of mutations like FLT3-ITD in murine hematopoiesis induces a lethal disorder characterized by extensive myelo-proliferation with normal maturation. The genetic translocations in AML often affect transcriptional regulators that are key players in normal hematopoiesis. Well-known examples involve the core-binding factor (CBF), mixed-lineage leukemia (MLL), or the retinoid acid receptor (RARα), and the expression of translocation-generated fusion proteins like PML-RARα, MLL-fusions, or CBF-fusions in mouse bone marrow cells generally results in aberrant self-renewal capacity and blocked differentiation *in vitro*. This is associated with a clonal AML-like disease *in vivo* after prolonged latency. Studies of these molecular epigenetic regulators have provided convincing evidence that a significant number of driver leukemogenic mutations mediate their oncogenic potential directly or indirectly through chromatin or epigenetic modifications. Finally, about a third of AML cases do not harbor cytogenetically detectable chromosomal alterations. Interestingly, large-scale molecular screening efforts suggest that mechanisms such as DNA promoter methylation or microRNA expression might be key to the pathobiology of these AML cases. Here we review the role of epigenetic mechanisms focusing on histone modifying activities for the pathogenesis of AML that might result in novel targeted therapeutic approaches in the near future.

2 Epigenetic Regulation by Histone Methyltransferases

2.1 Mixed-Lineage Leukemia 1

Several AML-associated chromosomal translocations result in disruption of genes encoding histone modifiers. The best-studied example is the mixed-lineage leukemia 1 (MLL1) gene on 11q23 for which over 60 different fusion partners have been identified. Partial tandem duplications (PTDs) and gene amplifications have also been found in AML blasts [2, 3]. The most common MLL fusions in human acute leukemia are MLL/AF4, MLL/AF9, and MLL/ENL resulting from t(4;11), t(9;11), and t(11;19) translocations that are associated with acute lymphoblastic, myeloid, or mixed-lineage acute leukemia, respectively. Several studies demonstrated that expression of human or mouse MLL-fusion genes during murine hematopoiesis

transforms hematopoietic stem and progenitor cells by enabling aberrant self-renewal and blocking differentiation, resulting in the development of a lethal hematological disease that closely mimicks the leukemic phenotype observed in human patients [4, 5]. Comparative gene expression profiling studies revealed extensive overlap in target genes of MLL-fusion proteins in murine and human hematopoiesis, making the mouse a perfect system to study the molecular mechanisms of MLL-mediated leukemogenesis [6].

MLL, also known as ALL1 or KMT2A, is a SET domain containing histone lysine methyltransferase (HKMT). Together with WDR5, RbBP5, and ASH2L, it catalyzes mono-, di, and tri-methylation on lysine 4 of histone H3 (H3K4) around transcriptional start sites of genes [7, 8]. The protein is a member of a multiprotein family and is structurally and functionally conserved among flies and mammals [9]. Like the *D. melanogaster* TRITHORAX protein, MLL1 protein is a key developmental regulator [10]. Loss of *Mll1* function in the mouse results in altered expression of homeobox (HOX) genes and causes embryonic lethality [11]. In the hematopoietic system, Mll1 is an essential regulator of self-renewal of hematopoietic stem cells [12, 13].

2.2 Targeting of MLL-Fusion Complexes

Importantly, the C-terminal SET domain is absent in all leukemic MLL-fusion proteins. Instead, the fusions retain the N-terminal fragment of MLL1, containing three AT-hook DNA-binding motifs and a CxxC DNA-binding domain. Furthermore, the N-terminal end interacts with MENIN, the gene product of multiple endocrine neoplasia type 1 gene (MEN1) [14], and with Lens-epithelial growth factor (LEDGF) and Myb. These interactions seem not only to be critical for MLL's normal function but also for the leukemogenic activity of MLL-fusion proteins [14–16]. Interestingly, LEDGF contains a PWWP domain, a motif belonging to Tudor domain "Royal Family" [17], that is thought to mediate binding to chromatin [18]. The integrity of the PWWP domain of LEDGF is required for leukemogenesis [16]. Furthermore, direct fusion of LEDGF's PWWP domain to MLL1, a configuration naturally existing in MLL1 orthologs in plants [19], rescues the leukemogenic potential of MLL-fusion proteins lacking the MENIN and LEDGF interaction domains [16]. This observation suggested that MENIN's predominant role in MLL-fusion leukemia is to recruit MLL-fusions to chromatin by promoting the interaction between MLL and LEDGF [16]. The exact nature of chromatin bound by LEDGF's PWWP motif is unknown. In contrast to the AT-hooks, oncogenic MLL1-fusions require the CxxC domain for transformation. Similar to those of MLL2 and CFP1, the CxxC domain of MLL1 binds to CpG-rich unmethylated DNA [20–22]**.** Although closely related, MLL2-based translocations have never been identified in human AMLs. Domain swapping experiments showed that the presence of the CxxC domain of MLL1 in the backbone of MLL2 protein is sufficient for leukemogenesis [20]. Interestingly, by impairing the DNA-binding

ability of MLL1's CxxC domain through the introduction of one specific point mutation, the leukemogenic potential of the MLL/AF9 fusion on bone marrow progenitor cells was fully abrogated. This observation coincided with reduced expression of *Hoxa9*. In non hematopoietic cells, this mutation resulted in increased levels of DNA methylation at specific CpG nucleotides of the *Hoxa9* promoter and of repressive histone H3 lysine 9 trimethylation [23]. Promoter occupancy of MLL/AF9 fusion protein was, however, largely unaltered, arguing that the CxxC domain may play a critical role in the regulation of chromatin states and gene expression levels of MLL-fusion target genes while the PWWP domain of LEDGF would direct binding of the oncogenic fusion to chromatin.

2.3 Enhancing Transcriptional Elongation

In most human and mouse MLL-fusion-based leukemias, a number of direct target genes are commonly upregulated. Prime examples are HOXA7, HOXA9, HOXA10, and MEIS1 [6, 24–28]. Elevated expression of the HOXA cluster genes is necessary for maintenance of the transformed state mediated by a MLL fusion [29]. HOX and MEIS1 proteins function as heterodimers to activate downstream genes [30]. Accordingly, overexpression of HOXA9 and MEIS1 is sufficient to induce an aggressive AML phenotype *in vivo* [31]. Surprisingly, transformation potential is not limited to HOXA9, with the exception of HOXA2 and HOXA5, overexpression of any HOXA gene immortalized murine granulocytic–monocytic progenitor cells *in vitro* [32].

Nonetheless, despite sharing common targets, it is currently unknown why different MLL-fusions cause leukemias in different hematopoietic lineages [33, 34]. An answer may lie in the cellular and biochemical properties of the particular fusion partners. Biochemical studies have shown that MLL-fusion partners like AF4, ENL, and AF9 form multiprotein complexes that serve a role in transcriptional elongation [35–40]. The AF4 and ENL/AF9 protein families account for two-thirds of MLL-associated leukemias [41]. The AF4 family comprises four paralogous proteins including AF4, AF5q31, LAF4, and FMR2. The paralogous ENL and AF9 proteins show homology to the yeast Anc1 protein, which is implicated in transcriptional regulation. All but FMR2 have been identified as MLL-fusion partners in leukemia. Recent work by Yokoyama and colleagues [40] showed that AF4 and AF5q31 preferentially form heterodimers (over homodimers) that interact with ENL as well as with the CDK9 and cyclinT1 subunits of pTEFb, the Positive Transcription Elongation Factor b, which phosphorylates the carboxy-terminal domain (CTD) of RNA polymerase II to facilitate transcriptional elongation [42, 43]. Furthermore, they showed that ENL interacts with AF4 (or AF5q31) and with DOT1L, the H3 lysine 79-specific methyltransferase, although in a mutually exclusive manner.

These data establish an order of action of the different protein complexes involved in transcriptional elongation, with the AF4–ENL–pTEFb (AEP) complex

functioning upstream of the ENL–DOT1L complex. Furthermore, it provides a rationale of how MLL-fusions enhance expression of target genes by bypassing the quality control of recruitment of elongation complexes during the course of transcription. Consistent with this model, target genes of MLL-fusions are extensively marked by the elongation chromatin marker H3K79me2 along their gene bodies, and DOT1L is required for leukemogenesis [26, 27, 44]. Nevertheless, initiation of transcription of MLL-fusion target genes requires the function of the second nonmutated M*ll1* allele, arguing that MLL-fusion enhanced transcriptional elongation is still controlled at the level of transcriptional initiation [45]. Since all major MLL-fusion partners function in the transcriptional elongation process, the development of a lymphoid, mixed, or myeloid-specific leukemia may result from the ability of the different MLL-fusion partners to interact with distinct proteins in a lineage-specific manner during hematopoiesis, thereby potentially regulating different target genes.

2.4 A Role for PHD Fingers in Leukemia

Characterization of a complex cytogenetic abnormality in cells from a patient with AML harboring a t(11;12)(p15;p13) revealed a novel fusion between nucleoporin-98 (NUP98) and the lysine-specific demethylase 5A (KDM5A), also known as JARID1A or retinoblastoma binding protein 2 (RBP2) [46]. JARID1A belongs to a family of histone demethylases including PLU1 (KDM5B) and SMCX (KDM5C) that harbor di- and tri-demethylation activity toward H3K4 [47, 48]. In *D. melanogaster*, the ortholog LID controls homeotic gene expression during development [49]. In mouse ES cells, Jarid1A interacts with members of the Polycomb Repressive Complex 2 (PRC2) and colocalizes to a large number of PRC2 target genes in a largely PRC2-dependent manner. Jarid1A is thought to contribute to PRC2-mediated silencing by counteracting the transcriptional promoting role of H3K4 methyltransferases like those of the MLL family [50, 51]. Surprisingly, *Jarid1a* deficient mice are viable and showed no major phenotypes [48].

Expression of the NUP98/JARID1A fusion blocked differentiation of murine bone marrow cells *in vitro* and induced an AML-like phenotype *in vivo* [52]. Besides the enzymatic JmjC domain, Jarid1A contains three PhD fingers [53]. Interestingly, whereas the first N-terminal PHD finger has affinity toward unmethylated H3K4, the third C-terminally located PHD finger binds to di- and tri-methylated H3K4 [52]. In the leukemic NUP98/JARID1A fusion, only a nuclear localization signal and the third PHD finger of JARID1A are fused to the transactivating phenylalanine-gylcine (FG)-repeats of NUP98. Expression of this fusion protein in hematopoietic progenitor cells caused up-regulation of expression of a number of developmental regulators like *Hoxa5*/*a7*/*a9*/*a10*, *Gata3*, *Meis1*, and *Pbx1* that are repressed by Polycomb group proteins in ES cells. Likewise, NUP98-JARID1A expression resulted in decreased H3K27me3 and increased

H3K4me3 levels throughout the *Hoxa* cluster. Importantly, the H3K4me2/3-binding affinity of the PHD finger was required for binding of the NUP98-JARID1A fusion to classical MLL-fusion targets like *Hoxa7* and *Hoxa9* and for *in vitro* transformation. Oncogenic transformation potential was retained when JARID1's third PHD finger was replaced by H3K4me2/3-specific PHD fingers from other proteins but not when replaced by PHD fingers recognizing unmethylated H3K4. Mechanistically, NUP98/PHD fusions are proposed to act as "chromatin boundary factors" counteracting PcG-mediated repression to "lock-in" critical loci into an active chromatin state [52]. In this model, the PHD finger would act as a molecular glue to direct NUP98's transactivation function to target genes (see below), presumably in a positive feedback loop involving H3K4 methylation. Interestingly, MLL1 harbors a number of PHD fingers adjacent to its CxxC domain. Inclusion of the third PHD finger into MLL-fusions abrogates their leukemogenic potential [54, 55], arguing that the targeting and activity of transactivating fusion proteins is primarily determined by having the right combination of DNA and chromatin-binding modules.

3 NSD-Fusions Target Histone Methyltransferase Activity

Another family of histone methyltransferases that are targets of genetic alterations in hematologic malignancies are the Nuclear receptor binding SET Domain (NSD) proteins composed by three members, NSD1 to NSD3. NSD1 was discovered in a screen for interacting partners of nuclear hormone receptors, acting as corepressor or coactivator depending on the cellular context [56]. NSD1 harbors intrinsic histone methyltransferase activity with specificity for H3K36 and H4K20 and is an important regulator of early postimplantation development [57]. NSD1 and NSD3 are both involved in genetic alterations like t(5;11)(q35;p15), cryptic insertions add(11)(p15) or t(8;11)(p11;p15) leading to expression of NUP98/NSD1 or NUP98/NSD3 fusion genes, respectively [58–60]. The resulting fusions contain the FG-repeats of NUP98 fused to the PHD fingers, PWWP, SET, and C5HCH domains of NSD1 and NSD3, respectively. NUP98/NSD1 was demonstrated to be a potent oncoprotein that efficiently transforms murine bone marrow cells *in vitro* and induces an AML-like phenotype in a transplant model *in vivo* [61]. Interestingly, transformation of the NUP98/NSD1 fusion was associated with its H3K36 methylation activity and HOXA gene cluster binding and activation. Structure functional analysis revealed that a region containing the fifth PHD finger and the adjacent C5HCH domain was essential for binding to the *Hoxa9* promoter and for the transforming activity of NUP98/NSD1. The region containing the first four PHD fingers and the PWWP domain was not required for *Hoxa9* promoter binding nor transactivation of *Hoxa7 in vitro*. In contrast to JARID1's third PHD finger, none of NSD1's PHD fingers seem to have the critical H3K4-interacting residues, suggesting that NUP98/NSD1 may interact with chromatin through another histone modification [62, 63]. Deletion analysis showed that the NUP98 domain is not required for *Hoxa9* promoter binding; instead it controls transcriptional activation

by targeting p300/CBP and regulating H3 acetylation and is thereby essential for myeloid progenitor immortalization. It will be important to determine whether the NUP98 domain is sufficient for transactivation by NUP98/NSD1 or whether the catalytic activity of the SET domain is required as well.

Interestingly, germline missense mutations of the SET, PWWP, PHD, and cysteine/histidine-rich C5HCH domains are also found in SOTOS and Weaver syndromes, two childhood overgrowth syndromes characterized by facial dysmorphism, advanced bone age, seizures, and mental retardation [64, 65]. Mutations associated with overgrowth syndromes seem to cluster toward the fifth PHD and the adjacent C5HCH motif, suggesting at least in part overlaps in targeting mechanisms. Interestingly, epidemiological data recently suggested a link between male gender and elevated risk for the development of hematological malignancies of SOTOS patients [66].

The role of NSD proteins in human cancer is not limited to AML, as NSD2 (also called MMSET) is found in a chromosomal translocation t(4;14)(p16;q23) associated with a significant fraction of patients with multiple myeloma [67]. NSD2 is also known as Wolf–Hirschhorn candidate 1 (WHSC1), as the gene maps to a 165 kb critical region on chromosomal 4 that is targeted by a hemizygous deletion in this malformation syndrome [68]. Accordingly, mice deficient for *Nsd2* display phenotypes reminiscent of the human disease [69]. There is increasing evidence that alterations of NSD1 and NSD3 could be involved in the pathogenesis of diverse forms of human cancers like neuroblastoma, glioma, colon carcinoma, or breast cancer. Convincing functional evidence is, however, lacking [70, 71].

4 Epigenetic Regulation by Histone Acetyltransferases

The second class of chromatin modifying enzymes involved in leukemia-associated genetic alterations are histone acetyltransferases (HATs). Based on their catalytic domains, HATs are classified into three classes composed of (1) GCN5 N-acetyltransferases, (2) p300 and CBP (cAMP response element binding (CREB) protein), and (3) MYSTs including MOZ, MORF, Ybf2, Sas2, and Tip60 members. Several chromosomal translocations affecting class 2 and 3 HATs have been identified in AML blasts [72].

Both the transcriptional coactivator CBP and its paralogue p300 are involved in cellular fate decisions in hematopoiesis. Whereas self-renewal of hematopoietic stem cells is dependent on CBP, p300 is crucial for proper hematopoietic differentiation [73,74]. p300 but not CBP regulates hematopoiesis through its KIX domain, a known protein/protein interaction interface for c-Myb and CREB [75]. By interacting with p300, c-Myb controls proliferation and differentiation of hematopoietic stem and progenitor cells [76]. Two translocations (t(11,16)(q23, p13) and t(11;22)(q23;q13)) mostly associated with secondary therapy-related AML result in fusions of the CBP or p300 to MLL [77, 78]. Transforming activity of MLL/CBP has been demonstrated *in vitro* and in bone marrow transplantation assays as well as by

generation of a conditional knock-in allele in mice [11, 79]. Furthermore, knockdown of MLL/CBP expression in the SN-1 cell line carrying this translocation resulted in cellular differentiation. Interestingly, the fusion-mediated block in differentiation can be overcome with RXR agonists or a clinical more applicable combination of all-trans retinoic acid (ATRA) and histone deacetylase (HDAC) inhibitors (see below) [80].

The monocytic leukemia zinc finger (MOZ; also called MYST1) and the monocytic leukemia zinc finger-related factor (MORF; also called MYST4) are both involved in rare translocations t(8;16)(p11;p13) and t(10;16)(q22;p13) leading to expression of MOZ-CBP and MORF-CBP, respectively [81, 82]. MOZ, but not the MOZ-CBP fusion, is a coactivator of AML1 (a subunit of the core binding factor) and is essential for the establishment of definitive hematopoiesis [83]. MOZ deficient mice die around embryonic day 15 with a severely reduced number of hematopoietic stem and progenitor cells that are incapable of reconstituting hematopoiesis upon transplantation [84, 85]. There is growing evidence that these fusion might induce the leukemic phenotype not only through interference with AML1 but also by blocking p53-mediated transcription upon DNA damage and by blocking apoptosis through NF-κB regulation [86, 87]. Interestingly, gene expression profiling of AML harboring a MOZ/CBP fusion revealed a distinct signature characterized by up-regulation of HOXA9, HOXA10, MEIS1, and FLT3 resembling in part expression programs found in AML with MLL rearrangements [88].

Inv(8)(p11;q13) associated with AML leads to fusion of MOZ to the transcriptional intermediary factor 2 (TIF2). This fusion couples the HAT domain of MOZ to the trans-activation and CBP-interaction domain of TIF2 [89, 90]. Expression of MOZ/TIF2 in murine hematopoietic stem and progenitor cells *in vitro* provided aberrant self-renewal capacity with a block in normal differentiation and caused AML in a murine bone marrow transplant assay. Interestingly, whereas the HAT domain of MOZ was dispensable for leukemia induction, interaction of MOZ/TIF2 with CBP was essential for transformation [91, 92]. Functionally, MOZ/TIF2 seems to act as a dominant inhibitor of CBP-dependent activators such as nuclear receptors or p53 and to alter cofactor recruitment and histone modification at the retinoid acid receptor beta 2 (RARβ2) [93, 94]. In summary, both direct and indirect mistargeting of CBP or p300 via various fusion proteins are associated with AML in humans and mouse.

5 Indirect Epigenetic Deregulation

Chromosomal translocations involving the core binding factor (CBF) such as t(8;21)(q22;q22) or inv16(p13q22) resulting in the AML1/ETO and CBFβ/MYH11 fusion protein, respectively, are the most prevalent cytogenetic alterations in human AML. Gene knockout studies in mice have shown that the CBF subunits

AML1 and CBFβ are both essential regulators of definitive hematopoiesis. Whereas AML1 acts primarily as a transcriptional activator, leukemia-associated fusion proteins are dominant-negative silencers of genes essential for normal myeloid differentiation, acting through direct cooperation with nuclear corepressors including N-COR1/2, SIN3A, as well as HDACs [95]. In addition, they also recruit DNA methyltransferase 1 (DNMT1), suggesting that aberrant repression might also be mediated by promoter hypermethylation [96, 97].

Similar molecular mechanisms seem to underlie malignant transformation by translocations involving the retinoid acid receptor alpha (RARα), comprising the second most frequent genetic alteration that is almost exclusively associated with AML M3, also called acute promyelocytic leukemia (APL). Whereas wild-type RARα acts as transcriptional activator, X-RARα fusions (with PML-RARα being by far the most prevalent fusion) function as transcriptional repressors, also through recruitment of chromatin modifiers like HDACs, N-COR1/2, DNMT1/3a, and the H3K9 HMT SUV39H1[98].

PML/RARα recruits PRC2 to the promoter of the RARβ2 target gene [87]. siRNA-mediated knockdown of the PRC2 component SUZ12 released in part the PML/RARα-mediated differentiation block in APL cells. Importantly, the PML/RARα-mediated differentiation block can be overcome by ATRA. In contrast, blocked differentiation and leukemogenesis by the PLZF-RARα variant are insensitive to ATRA treatment. This may be due to direct interaction of the PLZF-RARα variant with the Polycomb group protein Bmi1 resulting in the recruitment of the repressive PRC1 complex to RA-responsive elements [99].

The importance of DNA methylation, H3K9 and H3K27 tri-methylation in PML/RARα-mediated gene silencing has recently, however, been challenged by an extensive genome-wide localization study of PML/RARα, RXR, RNA polymerase II and various histone modifications and DNA methylation in PML/RARα expressing cell lines as well as in primary APL cells [100]. The authors identified over 2,700 sites co-occupied by the fusion and RXR. Importantly, levels of H3K9me3, H3K27me3, and DNA methylation were generally low at PML/RARα-RXR binding sites. Moreover, the level of these modifications did not change upon treatment with pharmacological doses of ATRA, whereas acetylation at H3K9 and H3K14 dramatically increased. Upon ATRA treatment, H3K9K14ac levels increased at many PML/RARα-RXR targets, yet the increase was stronger for genes that were induced rather than repressed [100]. Another ChIP-chip study reported an inverse correlation between PML-RARα occupancy and H3 acetylation, yet observed a positive correlation for H3K9me3 and PML-RARα [101]. Finally, Wang and colleagues [102] identified through genome-wide chromatin and computational analyses that the PU.1 transcription factor binds to promoters of genes occupied and repressed by PML-RARα. While the crucial changes in transcription remain unclear, it is through such genome-wide studies that we will be able to evaluate the importance of different transcription and chromatin factors during leukemic oncogenesis. Furthermore, these recent studies underscore the importance of HDAC inhibitors as a potential therapeutic approach for PML/RARα leukemias.

6 Aberrant Promoter DNA Methylation

Methylation at position carbon 5 of the cytidine ring in the context of a CpG dinucleotide is mediated by DNA methyltransferases (DNMTs) including DNMT1, 3a and 3b. Whereas DNMT3a/b act as *de novo* methylases binding to both unmethylated and hemimethylated CpGs, DNMT1 is responsible for maintaining DNA methylation patterns and preferentially binds to hemimethylated DNA. In mammalian genomes, hypomethylated CpG dinucleotides are generally located in higher frequencies around promoters of more than half of genes (in so-called CpG-islands). DNA methylation is essential for development, as demonstrated by embryonic lethality of mice deficient for Dnmt1 and Dnmt3b [103,104]. Conditional disruption in the hematopoietic system revealed an essential role of Dnmt1 in regulating HSC self-renewal, niche retention, and early multilineage differentation [105]. Interestingly, the dosage of DNA methylation seems also to be important to protect HSC multipotency from myeloerythroid restriction [106]. Conditional ablation of Dnmt3a, Dnmt3b, or both in Kit+/Sca1+/Cd34low/lineage-negative hematopoietic cells revealed an essential role in HSC self-renewal [107].

Aberrant DNA methylation is a hallmark of many human cancers including AML, which are typically characterized by hypermethylation of CpG islands in promoter regions of many tumor suppressor genes, such as CDKN2B (p15), HIC1, and CTNNA1, while being associated with overall DNA hypomethylation [108–110]. Interestingly, AML blast cells with methylated p15(INK4B) tended to express higher levels of DNMT1 and DNMT3B [111]. Aberrant DNA methylation seems to be a dominant mechanism for silencing of tumor suppressors during clonal progression of myelodysplastic syndromes (MDS) to AML [112]. Quantitative screening of DNA methylation by mass spectrometry of almost 100 genomic regions in cells from over 200 patients revealed a novel outcome predictor for AML [113]. Similarly, large-scale DNA methylation profiling of tumor samples of 344 AML patients using the HELP assay (*HpaII* tiny fragment Enrichment by Ligation-mediated PCR) in combination with high-density microarray hybridization revealed DNA methylation signatures for biologically distinct AML subtypes with genetic alterations in e.g., CEPBA, NPM1, CBF, or RARα alterations [114].

In infant acute leukemia, specific promoter DNA methylation patterns correlated with the underlying MLL rearrangement. MLL translocations t(4;11) and t(11;19) showed extensive hypermethylation, whereas leukemias with the t(9;11) translocation or carrying wild-type MLL lacked aberrant DNA methylation. Again, the degree of hypermethylation appeared to influence relapse-free survival, furthermore demonstrating the prognostic impact [115]. Another study using a different screening platform observed global hypermethylation in MLL-rearranged infant leukemia compared with both normals and common childhood ALL [116]. In both studies, treatment with demethylating agents reversed aberrant DNA methylation and induced apoptosis in leukemic cell lines carrying MLL rearrangments, suggesting that inhibition of aberrant DNA methylation might provide a functional therapeutic strategy [115, 116].

Cloning of the translocation t(10;11)(q22;q23) associated with several cases of AML revealed a novel fusion of MLL1 to TET1 (Ten-Eleven Translocation; also known as Leukemia-associated protein with a CXXC domain, LCX) [117,118]. TET1 is the prototype of a larger family of so-called kinetoplastid base J binding proteins (JBP) that are able to catalyze *in situ* hydroxylation of bases in nucleic acids [119]. TET1 has been identified as a 2-oxoglutarate (2OG)-and Fe(II)-dependent enzyme that is able to convert 5-methylcytosine into 5-hydroxymethylcytosine (5-hmC) *in vitro*. In mouse embryonic stem cells, 5-hmC levels decreased upon differentiation or by TET1 depletion suggested that this novel enzyme might function as an epigenetic regulator [120]. Cloning of candidate tumor suppressor genes common in a large cohort of patients with myeloid cancers revealed point mutations and deletions in TET2 in 15% of patients including 5 out of 21 secondary AML cases developed from chronic myeloproliferative disorders (MPD) [121]. TET2 mutations were also found in 12/119 patients with AML without any correlation to specific cytogenetic alterations or history of antecedent MPD or MDS, but with a decreased overall survival compared with TET2 wild-type patients. The same study did not find any alterations of TET1 or TET3 in 96 patients with myeloproliferative neoplasia [122]. Functional contribution and the underlying molecular mechanisms of alterations of the TET gene family for induction and maintenance of leukemic disorders remain to be elucidated.

7 MicroRNAs in AML

Regulatory noncoding RNAs including microRNAs (miRNAs), small interfering RNAs, Piwi-interacting RNAs, and various types of long noncoding RNAs regulate gene expression at the transcriptional and posttranscriptional level (reviewed in [123]). An avalanche of recent reports has provided convincing evidence that aberrant miRNA expression controls the development and maintenance of various human disorders including cancer (reviewed in [124]). Since the expression of miRNAs is highly regulated during normal hematopoiesis, it is not surprising that deregulated miRNA expression contributes to leukemogenesis. The role of aberrantly expressed miRNAs as potential oncogenes or tumor suppressor genes in human leukemia was the subject of recent reviews [125, 126]; therefore, we highlight here only the most recent findings that underscore the role of miRNAS in AML.

Specific subsets of aberrantly expressed miRNAs have been identified in AML cases harboring distinct cytogenetic aberrations such as t(15;17) leading to PML/RARα, CBF alterations resulting from inv(16) or t(8;21), or balanced translocations of 11q23 mostly composed of MLL-fusions [127]. Transformation by MLL fusion genes was associated with overexpression of miR-196 and the miR-17-92 cluster. Interestingly, treatment with specific antagomirs abrogated the self-renewal capacity of MLL-fusion expressing bone marrow cells, suggesting an active role for aberrant miRNA expression in leukemic transformation by MLL fusions [128]. Expression of the miR-17-92 cluster functionally cooperated with

MLL fusions increasing clonogenic growth of murine bone marrow cells. Expression profiling of miR-17-92 expressing cells resulted in over 300 potential targets that were significantly enriched in regulators of cellular differentiation, hematopoiesis, cell cycle, and apoptosis [129]. MiR-128b and 221 were found to be downregulated in MLL-rearranged ALL. Reexpression of these miRNAs cooperatively sensitized two MLL/AF4 positive cell lines to glucocorticoid response. Interestingly, mutations targeting miR-128b have been identified in a MLL/AF4 cell line and in primary cells. One mutation seems to significantly reduce the processing of miR-128b contributing to glucocorticoid resistance [130, 131].

Leukemia-associated chromosomal translocation can also directly lead to aberrant miRNA expression and transformation of hematopoietic cells. Cloning of a recurrent translocation t(2;11)(p21;q23) associated with MDS and AML revealed that this translocation resulted in significant upregulation of miR-125b-1 located near the breakpoint on chromosome 11. Expression of miR-125b-1 in primary human CD34+ cell and leukemic cell lines resulted in impaired cellular differentiation, suggesting that aberrant expression of miR-125b-1 might contribute to malignant transformation [132]. Likewise, miR-29a that is highly expressed in HSCs and downregulated during progenitor differentiation was also overexpressed in blasts from AML patients. Ectopic expression of miR-29a in mouse bone marrow stem and progenitor cells resulted in aberrant self-renewal *in vitro* and the development of a myeloproliferative disorder that progressed into AML *in vivo* [133].

Another miRNA with potential active contribution to leukemogenesis is miR-125b-2 located on chromosome 21, which is overexpressed in acute megakaryoblastic leukemia associated with trisomy 21/Down syndrome (DS-AMKL). Overexpression of miR-125b-2 increased proliferation and self-renewal of human and mouse megakaryocytic and megakaryocytic/erythroid progenitor cells, with accentuation in presence of the GATA1s mutation associated with trisomy 21. Transcriptome analysis revealed several potential targets that were downregulated in DS-AMKL highly expressing miR-125b, suggesting that miR-125b-2 might act as potential oncomir in DS-AMKL [134].

In summary, these studies assign an important role to miRNAs in AML. To what extent functional interference with aberrantly expressed miRNAs could be of therapeutic benefit in MLL leukemia remains to be determined.

8 Epigenetic Therapy of AML

Aberrant genetic programs mediated by epigenetic mechanisms, such as promoter hypermethylation and histone deacetylation, can be theoretically reverted by pharmacological inhibitors. The prototype of successful epigenetic therapy is reversion of the PML/RARα-induced myeloid differentiation block in APL by all-trans retinoid acid (ATRA) and/or arsenic trioxide (AsO_3) [135]. It is currently unclear why non-APL AML only poorly responds to ATRA given the importance of the RAR/RXR signaling pathway in myelomonocytic differentiation [136], although

the nature of cooperating genetic lesions in non-APL AML could be responsible for failure of ATRA in these cases. In addition, non-APL AML cells often express low levels of RARα [137]. The product of t(8;21) the AML1/ETO fusion seem to epigenetically modify RAR/RXR signaling by inducing aberrant DNA methylation of the promoter of the ATRA target gene RARβ2 [138]. AML-associated fusions like AML1/ETO or MN1/TEL directly block RAR/RXR-mediated transcription. Meningioma1 (MN1) overexpression often observed in AML with inv(16) seems to abrogate ATRA-induced transcription and is able to induce a leukemic phenotype in mice alone and in cooperation with leukemogenic oncogenes like CBFβ/MYH11 or MLL/ENL [139,140]. More research is needed to evaluate the molecular clues defining the ATRA responsiveness of non-APL AML.

Current epigenetic therapeutic approaches are based on targeting aberrant DNA methylation and inhibition of HDACs. As outlined above, leukemic blasts (and cancer cells in general) often show aberrant promoter hypermethylation resulting in silencing of a significant number of potential tumor suppressor genes. Hypomethylation of these promoters may restore expression of genes, thereby impairing growth and survival and/or promoting cellular differentiation. Two nucleoside analogs, 5-azacitidine (VIDAZA) and 5-Aza-2′-deoxycytidine (Decitabine, DACOGEN), are currently being evaluated for their antileukemic potential in several clinical trials. Both of these compounds are incorporated into DNA and form a covalent complex with the DNMT enzyme that results in trapping and progressive loss of DNMT activity. Hematological response rates with clinical benefits were reported in up to 40% of patients with MDS (including some patients with signs of progression toward AML) in trials using aza-nucleoside as single agents (reviewed in [141]). Phase III trials comparing aza-nucleosides to supportive care in MDS patients reported encouraging overall response rates of 30–47% with complete response rates of less than 10% [142, 143]. Several studies using decitabine have also provided promising results for AML therapy [144–146]. A recent phase II clinical trial with single agent decitabine therapy in older patients (>60y) with previously untreated AML showed high activity with a complete remission rate of 47% after a median of three cycles [147]. However, the effects of these drugs on DNA methylation are not permanent and chronic application is required. Whether the observed clinical activity of aza-nucleoside is purely based on reversion of epigenetic silencing is currently not clear given the ability of these compounds to induce a DNA damage response [148, 149].

As outlined above, HDACs are epigenetic key elements in mediating the function of different fusion proteins (like PML/RARα, AML1/ETO, or CBFβ/SMMHC) derived from chromosomal translocations associated with acute leukemia. However, little is known about the role of the 11 individual HDACs for the development of a leukemic phenotype. Conditional isoform-specific knockout mice will be useful tools to further dissect the role of HDACs in normal and leukemic hematopoiesis. The use of HDAC inhibitors for the therapy of leukemia was driven by the idea that such compounds would relieve HDAC-mediated transcriptional repression resulting in differentiation and/or death of leukemic blasts. There is accumulating evidence that most currently used HDAC inhibitors

have pleiotropic effects in leukemic cells, beyond the block of repression associated with the fusion. Notably these involve induction of key apoptosis regulators and proteasomal degradation of oncogenic fusion proteins leading to differentiation and subsequent death of the leukemic blasts [150–152].

Despite the lack of a complete understanding of their activities, several small molecule HDAC inhibitors have been evaluated in clinical trials for AML therapy with encouraging results. HDAC inhibitors in clinical trials are often classified based on their structure and on the proposed activity profile, including hydroxamic acid-based compounds like SAHA (Vorinostat), PXD101 (Belinostat), and LBH-589 (Panbinostat); synthetic benzamide derivates like MS275 (Entinostat) and MGCD0103 (Mocetinostat); 2-propylpentanoic acid (Valproate) or cyclic peptides like FK228 (Istodax) [153]. A phase I study with MGCD0103 for monotherapy of AML or MDS demonstrated some clinical activity inducing complete remissions in 3 out of 21 patients [154]. Transient antileukemic activities of monotherapies with HDAC inhibitors initiated several trials exploring the combinations of HDAC inhibitors with chemotherapy and/or demethylating agents. Indeed, preclinical studies demonstrated synergistic anticancer activity of HDAC inhibitors with DNMT inhibitors resulting in reexpression of genes silenced in cancer [155]. However, various combinations of these compounds have been associated with increased toxicity, and no definitive clinical benefit over monotherapies was observed [144, 156].

To overcome the current limitations of epigenetic cancer therapy, it will be essential to dissect the critical epigenetic mechanisms and molecular players that maintain a potentially reversible leukemic state. Further molecular and structural analyses of key enzymatic players, their interaction partners, and their modes of chromatin binding will lead to novel and more specific small molecules that will hopefully revolutionize leukemia therapy following the success of small molecule kinase inhibitors.

References

1. Kelly LM, Gilliland DG (2002) Genetics of myeloid leukemias. Annu Rev Genomics Hum Genet 3:179–198
2. Meyer C, Kowarz E, Hofmann J, Renneville A, Zuna J, Trka J, Ben Abdelali R, Macintyre E, De Braekeleer E, De Braekeleer M et al (2009) New insights to the MLL recombinome of acute leukemias. Leukemia 23:1490–1499
3. Meyer C, Schneider B, Jakob S, Strehl S, Attarbaschi A, Schnittger S, Schoch C, Jansen MW, van Dongen JJ, den Boer ML et al (2006) The MLL recombinome of acute leukemias. Leukemia 20:777–784
4. Krivtsov AV, Armstrong SA (2007) MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer 7:823–833
5. Slany RK (2009) The molecular biology of mixed lineage leukemia. Haematologica 94:984–993
6. Li Z, Luo RT, Mi S, Sun M, Chen P, Bao J, Neilly MB, Jayathilaka N, Johnson DS, Wang L et al (2009) Consistent deregulation of gene expression between human and murine MLL rearrangement leukemias. Cancer Res 69(3):1109–1116

7. Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, Hess JL (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell 10:1107–1117
8. Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, Wassell R, Dubois G, Mazo A, Croce CM, Canaani E (2002) ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol Cell 10:1119–1128
9. Hublitz P, Albert M, Peters AH (2009) Mechanisms of transcriptional repression by histone lysine methylation. Int J Dev Biol 53:335–354
10. Ringrose L, Paro R (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu Rev Genet 38:413–443
11. Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ (1995) Altered Hox expression and segmental identity in Mll-mutant mice. Nature 378:505–508
12. Jude CD, Climer L, Xu D, Artinger E, Fisher JK, Ernst P (2007) Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. Cell Stem Cell 1:324–337
13. McMahon KA, Hiew SY, Hadjur S, Veiga-Fernandes H, Menzel U, Price AJ, Kioussis D, Williams O, Brady HJ (2007) Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. Cell Stem Cell 1:338–345
14. Yokoyama A, Somervaille TC, Smith KS, Rozenblatt-Rosen O, Meyerson M, Cleary ML (2005) The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. Cell 123:207–218
15. Jin S, Zhao H, Yi Y, Nakata Y, Kalota A, Gewirtz AM (2010) c-Myb binds MLL through menin in human leukemia cells and is an important driver of MLL-associated leukemogenesis. J Clin Invest 120:593–606
16. Yokoyama A, Cleary ML (2008) Menin critically links MLL proteins with LEDGF on cancer-associated target genes. Cancer Cell 14:36–46
17. Maurer-Stroh S, Dickens NJ, Hughes-Davies L, Kouzarides T, Eisenhaber F, Ponting CP (2003) The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. Trends Biochem Sci 28:69–74
18. Shun MC, Botbol Y, Li X, Di Nunzio F, Daigle JE, Yan N, Lieberman J, Lavigne M, Engelman A (2008) Identification and characterization of PWWP domain residues critical for LEDGF/p75 chromatin binding and human immunodeficiency virus type 1 infectivity. J Virol 82:11555–11567
19. Alvarez-Venegas R, Avramova Z (2001) Two Arabidopsis homologs of the animal trithorax genes: a new structural domain is a signature feature of the trithorax gene family. Gene 271:215–221
20. Bach C, Mueller D, Buhl S, Garcia-Cuellar MP, Slany RK (2009) Alterations of the CxxC domain preclude oncogenic activation of mixed-lineage leukemia 2. Oncogene 28:815–823
21. Birke M, Schreiner S, Garcia-Cuellar MP, Mahr K, Titgemeyer F, Slany RK (2002) The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. Nucleic Acids Res 30:958–965
22. Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, Webb S, Kerr AR, Deaton A, Andrews R, James KD et al (2010) CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature 464:1082–1086
23. Cierpicki T, Risner LE, Grembecka J, Lukasik SM, Popovic R, Omonkowska M, Shultis DD, Zeleznik-Le NJ, Bushweller JH (2010) Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. Nat Struct Mol Biol 17:62–68
24. Krivtsov AV, Twomey D, Feng Z, Stubbs MC, Wang Y, Faber J, Levine JE, Wang J, Hahn WC, Gilliland DG et al (2006) Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature 442:818–822
25. Chen W, Kumar AR, Hudson WA, Li Q, Wu B, Staggs RA, Lund EA, Sam TN, Kersey JH (2008) Malignant transformation initiated by Mll-AF9: gene dosage and critical target cells. Cancer Cell 13:432–440

26. Guenther MG, Lawton LN, Rozovskaia T, Frampton GM, Levine SS, Volkert TL, Croce CM, Nakamura T, Canaani E, Young RA (2008) Aberrant chromatin at genes encoding stem cell regulators in human mixed-lineage leukemia. Genes Dev 22:3403–3408
27. Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S, Sinha AU, Xia X, Jesneck J, Bracken AP, Silverman LB et al (2008) H3K79 methylation profiles define murine and human MLL-AF4 leukemias. Cancer Cell 14:355–368
28. Somervaille TC, Matheny CJ, Spencer GJ, Iwasaki M, Rinn JL, Witten DM, Chang HY, Shurtleff SA, Downing JR, Cleary ML (2009) Hierarchical maintenance of MLL myeloid leukemia stem cells employs a transcriptional program shared with embryonic rather than adult stem cells. Cell Stem Cell 4:129–140
29. Faber J, Krivtsov AV, Stubbs MC, Wright R, Davis TN, van den Heuvel-Eibrink M, Zwaan CM, Kung AL, Armstrong SA (2009) HOXA9 is required for survival in human MLL-rearranged acute leukemias. Blood 113:2375–2385
30. Shen WF, Rozenfeld S, Kwong A, Kom ves LG, Lawrence HJ, Largman C (1999) HOXA9 forms triple complexes with PBX2 and MEIS1 in myeloid cells. Mol Cell Biol 19:3051–3061
31. Kroon E, Krosl J, Thorsteinsdottir U, Baban S, Buchberg AM, Sauvageau G (1998) Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. EMBO J 17:3714–3725
32. Bach C, Buhl S, Mueller D, Garcia-Cuellar MP, Maethner E, Slany RK (2010) Leukemogenic transformation by HOXA cluster genes. Blood 115:2910–2918
33. Drynan LF, Pannell R, Forster A, Chan NM, Cano F, Daser A, Rabbitts TH (2005) Mll fusions generated by Cre-loxP-mediated de novo translocations can induce lineage reassignment in tumorigenesis. EMBO J 24:3136–3146
34. Metzler M, Forster A, Pannell R, Arends MJ, Daser A, Lobato MN, Rabbitts TH (2006) A conditional model of MLL-AF4 B-cell tumourigenesis using invertor technology. Oncogene 25:3093–3103
35. Bitoun E, Oliver PL, Davies KE (2007) The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. Hum Mol Genet 16:92–106
36. Lin C, Smith ER, Takahashi H, Lai KC, Martin-Brown S, Florens L, Washburn MP, Conaway JW, Conaway RC, Shilatifard A (2010) AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. Mol Cell 37:429–437
37. Mohan M, Herz HM, Takahashi YH, Lin C, Lai KC, Zhang Y, Washburn MP, Florens L, Shilatifard A (2010) Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom). Genes Dev 24:574–589
38. Mueller D, Bach C, Zeisig D, Garcia-Cuellar MP, Monroe S, Sreekumar A, Zhou R, Nesvizhskii A, Chinnaiyan A, Hess JL et al (2007) A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. Blood 110:4445–4454
39. Mueller D, Garcia-Cuellar MP, Bach C, Buhl S, Maethner E, Slany RK (2009) Misguided transcriptional elongation causes mixed lineage leukemia. PLoS Biol 7:e1000249
40. Yokoyama A, Lin M, Naresh A, Kitabayashi I, Cleary ML (2010) A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. Cancer Cell 17:198–212
41. Huret JL, Dessen P, Bernheim A (2001) An atlas of chromosomes in hematological malignancies. Example: 11q23 and MLL partners. Leukemia 15:987–989
42. Peterlin BM, Price DH (2006) Controlling the elongation phase of transcription with P-TEFb. Mol Cell 23:297–305
43. Saunders A, Core LJ, Lis JT (2006) Breaking barriers to transcription elongation. Nat Rev Mol Cell Biol 7:557–567
44. Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, Su L, Xu G, Zhang Y (2005) hDOT1L links histone methylation to leukemogenesis. Cell 121:167–178

45. Thiel AT, Blessington P, Zou T, Feather D, Wu X, Yan J, Zhang H, Liu Z, Ernst P, Koretzky GA et al (2010) MLL-AF9-induced leukemogenesis requires coexpression of the wild-type Mll allele. Cancer Cell 17:148–159
46. van Zutven LJ, Onen E, Velthuizen SC, van Drunen E, von Bergh AR, van den Heuvel-Eibrink MM, Veronese A, Mecucci C, Negrini M, de Greef GE et al (2006) Identification of NUP98 abnormalities in acute leukemia: JARID1A (12p13) as a new partner gene. Genes Chromosomes Cancer 45:437–446
47. Christensen J, Agger K, Cloos PA, Pasini D, Rose S, Sennels L, Rappsilber J, Hansen KH, Salcini AE, Helin K (2007) RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. Cell 128:1063–1076
48. Klose RJ, Yan Q, Tothova Z, Yamane K, Erdjument-Bromage H, Tempst P, Gilliland DG, Zhang Y, Kaelin WG Jr (2007) The retinoblastoma binding protein RBP2 is an H3K4 demethylase. Cell 128:889–900
49. Lee N, Zhang J, Klose RJ, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y (2007) The trithorax-group protein Lid is a histone H3 trimethyl-Lys4 demethylase. Nat Struct Mol Biol 14:341–343
50. Lopez-Bigas N, Kisiel TA, Dewaal DC, Holmes KB, Volkert TL, Gupta S, Love J, Murray HL, Young RA, Benevolenskaya EV (2008) Genome-wide analysis of the H3K4 histone demethylase RBP2 reveals a transcriptional program controlling differentiation. Mol Cell 31:520–530
51. Pasini D, Hansen KH, Christensen J, Agger K, Cloos PA, Helin K (2008) Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. Genes Dev 22:1345–1355
52. Wang GG, Song J, Wang Z, Dormann HL, Casadio F, Li H, Luo JL, Patel DJ, Allis CD (2009) Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. Nature 459:847–851
53. Cloos PA, Christensen J, Agger K, Helin K (2008) Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. Genes Dev 22:1115–1140
54. Chen J, Santillan DA, Koonce M, Wei W, Luo R, Thirman MJ, Zeleznik-Le NJ, Diaz MO (2008) Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. Cancer Res 68:6199–6207
55. Muntean AG, Giannola D, Udager AM, Hess JL (2008) The PHD fingers of MLL block MLL fusion protein-mediated transformation. Blood 112:4690–4693
56. Huang N, vom Baur E, Garnier JM, Lerouge T, Vonesch JL, Lutz Y, Chambon P, Losson R (1998) Two distinct nuclear receptor interaction domains in NSD1, a novel SET protein that exhibits characteristics of both corepressors and coactivators. EMBO J 17:3398–3412
57. Rayasam GV, Wendling O, Angrand PO, Mark M, Niederreither K, Song L, Lerouge T, Hager GL, Chambon P, Losson R (2003) NSD1 is essential for early post-implantation development and has a catalytically active SET domain. EMBO J 22:3153–3163
58. Brown J, Jawad M, Twigg SR, Saracoglu K, Sauerbrey A, Thomas AE, Eils R, Harbott J, Kearney L (2002) A cryptic t(5;11)(q35;p15.5) in 2 children with acute myeloid leukemia with apparently normal karyotypes, identified by a multiplex fluorescence in situ hybridization telomere assay. Blood 99:2526–2531
59. Jaju RJ, Fidler C, Haas OA, Strickson AJ, Watkins F, Clark K, Cross NC, Cheng JF, Aplan PD, Kearney L et al (2001) A novel gene, NSD1, is fused to NUP98 in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia. Blood 98:1264–1267
60. Rosati R, La Starza R, Veronese A, Aventin A, Schwienbacher C, Vallespi T, Negrini M, Martelli MF, Mecucci C (2002) NUP98 is fused to the NSD3 gene in acute myeloid leukemia associated with t(8;11)(p11.2;p15). Blood 99:3857–3860
61. Wang GG, Cai L, Pasillas MP, Kamps MP (2007) NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. Nat Cell Biol 9:804–812
62. Baker LA, Allis CD, Wang GG (2008) PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. Mutat Res 647:3–12

63. Ruthenburg AJ, Li H, Patel DJ, Allis CD (2007) Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol 8:983–994
64. Douglas J, Hanks S, Temple IK, Davies S, Murray A, Upadhyaya M, Tomkins S, Hughes HE, Cole TR, Rahman N (2003) NSD1 mutations are the major cause of Sotos syndrome and occur in some cases of Weaver syndrome but are rare in other overgrowth phenotypes. Am J Hum Genet 72:132–143
65. Kurotaki N, Imaizumi K, Harada N, Masuno M, Kondoh T, Nagai T, Ohashi H, Naritomi K, Tsukahara M, Makita Y et al (2002) Haploinsufficiency of NSD1 causes Sotos syndrome. Nat Genet 30:365–366
66. Martinez-Glez V, Lapunzina P (2007) Sotos syndrome is associated with leukemia/lymphoma. Am J Med Genet A 143A:1244–1245
67. Chesi M, Nardini E, Lim RS, Smith KD, Kuehl WM, Bergsagel PL (1998) The t(4;14) translocation in myeloma dysregulates both FGFR3 and a novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts. Blood 92:3025–3034
68. Stec I, Wright TJ, van Ommen GJ, de Boer PA, van Haeringen A, Moorman AF, Altherr MR, den Dunnen JT (1998) WHSC1, a 90 kb SET domain-containing gene, expressed in early development and homologous to a Drosophila dysmorphy gene maps in the Wolf–Hirschhorn syndrome critical region and is fused to IgH in t(4;14) multiple myeloma. Hum Mol Genet 7:1071–1082
69. Nimura K, Ura K, Shiratori H, Ikawa M, Okabe M, Schwartz RJ, Kaneda Y (2009) A histone H3 lysine 36 trimethyltransferase links Nkx2-5 to Wolf–Hirschhorn syndrome. Nature 460:287–291
70. Angrand PO, Apiou F, Stewart AF, Dutrillaux B, Losson R, Chambon P (2001) NSD3, a new SET domain-containing gene, maps to 8p12 and is amplified in human breast cancer cell lines. Genomics 74:79–88
71. Berdasco M, Ropero S, Setien F, Fraga MF, Lapunzina P, Losson R, Alaminos M, Cheung NK, Rahman N, Esteller M (2009) Epigenetic inactivation of the Sotos overgrowth syndrome gene histone methyltransferase NSD1 in human neuroblastoma and glioma. Proc Natl Acad Sci USA 106:21830–21835
72. Yang XJ (2004) The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases. Nucleic Acids Res 32:959–976
73. Kung AL, Rebel VI, Bronson RT, Ch'ng LE, Sieff CA, Livingston DM, Yao TP (2000) Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. Genes Dev 14:272–277
74. Rebel VI, Kung AL, Tanner EA, Yang H, Bronson RT, Livingston DM (2002) Distinct roles for CREB-binding protein and p300 in hematopoietic stem cell self-renewal. Proc Natl Acad Sci USA 99:14789–14794
75. Kasper LH, Boussouar F, Ney PA, Jackson CW, Rehg J, van Deursen JM, Brindle PK (2002) A transcription-factor-binding surface of coactivator p300 is required for haematopoiesis. Nature 419:738–743
76. Sandberg ML, Sutton SE, Pletcher MT, Wiltshire T, Tarantino LM, Hogenesch JB, Cooke MP (2005) c-Myb and p300 regulate hematopoietic stem cell proliferation and differentiation. Dev Cell 8:153–166
77. Ida K, Kitabayashi I, Taki T, Taniwaki M, Noro K, Yamamoto M, Ohki M, Hayashi Y (1997) Adenoviral E1A-associated protein p300 is involved in acute myeloid leukemia with t(11;22)(q23;q13). Blood 90:4699–4704
78. 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
79. Wang J, Iwasaki H, Krivtsov A, Febbo PG, Thorner AR, Ernst P, Anastasiadou E, Kutok JL, Kogan SC, Zinkel SS et al (2005) Conditional MLL-CBP targets GMP and models therapy-related myeloproliferative disease. EMBO J 24:368–381

80. Niitsu N, Hayashi Y, Honma Y (2001) Downregulation of MLL-CBP fusion gene expression is associated with differentiation of SN-1 cells with t(11;16)(q23;p13). Oncogene 20:375–384
81. Borrow J, Stanton VP Jr, Andresen JM, Becher R, Behm FG, Chaganti RS, Civin CI, Disteche C, Dube I, Frischauf AM et al (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
82. Panagopoulos I, Fioretos T, Isaksson M, Samuelsson U, Billstrom R, Strombeck B, Mitelman F, Johansson B (2001) Fusion of the MORF and CBP genes in acute myeloid leukemia with the t(10;16)(q22;p13). Hum Mol Genet 10:395–404
83. Kitabayashi I, Aikawa Y, Nguyen LA, Yokoyama A, Ohki M (2001) Activation of AML1-mediated transcription by MOZ and inhibition by the MOZ-CBP fusion protein. EMBO J 20:7184–7196
84. Katsumoto T, Aikawa Y, Iwama A, Ueda S, Ichikawa H, Ochiya T, Kitabayashi I (2006) MOZ is essential for maintenance of hematopoietic stem cells. Genes Dev 20:1321–1330
85. Thomas T, Corcoran LM, Gugasyan R, Dixon MP, Brodnicki T, Nutt SL, Metcalf D, Voss AK (2006) Monocytic leukemia zinc finger protein is essential for the development of long-term reconstituting hematopoietic stem cells. Genes Dev 20:1175–1186
86. Rokudai S, Aikawa Y, Tagata Y, Tsuchida N, Taya Y, Kitabayashi I (2009) Monocytic leukemia zinc finger (MOZ) interacts with p53 to induce p21 expression and cell-cycle arrest. J Biol Chem 284:237–244
87. Villa R, Pasini D, Gutierrez A, Morey L, Occhionorelli M, Vire E, Nomdedeu JF, Jenuwein T, Pelicci PG, Minucci S et al (2007) Role of the polycomb repressive complex 2 in acute promyelocytic leukemia. Cancer Cell 11:513–525
88. Camos M, Esteve J, Jares P, Colomer D, Rozman M, Villamor N, Costa D, Carrio A, Nomdedeu J, Montserrat E et al (2006) Gene expression profiling of acute myeloid leukemia with translocation t(8;16)(p11;p13) and MYST3-CREBBP rearrangement reveals a distinctive signature with a specific pattern of HOX gene expression. Cancer Res 66:6947–6954
89. Carapeti M, Aguiar RC, Goldman JM, Cross NC (1998) A novel fusion between MOZ and the nuclear receptor coactivator TIF2 in acute myeloid leukemia. Blood 91:3127–3133
90. Liang J, Prouty L, Williams BJ, Dayton MA, Blanchard KL (1998) Acute mixed lineage leukemia with an inv(8)(p11q13) resulting in fusion of the genes for MOZ and TIF2. Blood 92:2118–2122
91. Deguchi K, Ayton PM, Carapeti M, Kutok JL, Snyder CS, Williams IR, Cross NC, Glass CK, Cleary ML, Gilliland DG (2003) MOZ-TIF2-induced acute myeloid leukemia requires the MOZ nucleosome binding motif and TIF2-mediated recruitment of CBP. Cancer Cell 3:259–271
92. Huntly BJ, Shigematsu H, Deguchi K, Lee BH, Mizuno S, Duclos N, Rowan R, Amaral S, Curley D, Williams IR et al (2004) MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. Cancer Cell 6:587–596
93. Collins HM, Kindle KB, Matsuda S, Ryan C, Troke PJ, Kalkhoven E, Heery DM (2006) MOZ-TIF2 alters cofactor recruitment and histone modification at the RARbeta2 promoter: differential effects of MOZ fusion proteins on CBP- and MOZ-dependent activators. J Biol Chem 281:17124–17133
94. Kindle KB, Troke PJ, Collins HM, Matsuda S, Bossi D, Bellodi C, Kalkhoven E, Salomoni P, Pelicci PG, Minucci S et al (2005) MOZ-TIF2 inhibits transcription by nuclear receptors and p53 by impairment of CBP function. Mol Cell Biol 25:988–1002
95. Wang J, Hoshino T, Redner RL, Kajigaya S, Liu JM (1998) ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. Proc Natl Acad Sci USA 95:10860–10865
96. Barbetti V, Gozzini A, Rovida E, Morandi A, Spinelli E, Fossati G, Mascagni P, Lubbert M, Dello Sbarba P, Santini V (2008) Selective anti-leukaemic activity of low-dose histone deacetylase inhibitor ITF2357 on AML1/ETO-positive cells. Oncogene 27:1767–1778

97. Liu S, Shen T, Huynh L, Klisovic MI, Rush LJ, Ford JL, Yu J, Becknell B, Li Y, Liu C et al (2005) Interplay of RUNX1/MTG8 and DNA methyltransferase 1 in acute myeloid leukemia. Cancer Res 65:1277–1284
98. Lin RJ, Nagy L, Inoue S, Shao W, Miller WH Jr, Evans RM (1998) Role of the histone deacetylase complex in acute promyelocytic leukaemia. Nature 391:811–814
99. Boukarabila H, Saurin AJ, Batsche E, Mossadegh N, van Lohuizen M, Otte AP, Pradel J, Muchardt C, Sieweke M, Duprez E (2009) The PRC1 Polycomb group complex interacts with PLZF/RARA to mediate leukemic transformation. Genes Dev 23:1195–1206
100. Martens JH, Brinkman AB, Simmer F, Francoijs KJ, Nebbioso A, Ferrara F, Altucci L, Stunnenberg HG (2010) PML-RARalpha/RXR alters the epigenetic landscape in acute promyelocytic leukemia. Cancer Cell 17:173–185
101. Hoemme C, Peerzada A, Behre G, Wang Y, McClelland M, Nieselt K, Zschunke M, Disselhoff C, Agrawal S, Isken F et al (2008) Chromatin modifications induced by PML-RARalpha repress critical targets in leukemogenesis as analyzed by ChIP-Chip. Blood 111:2887–2895
102. Wang K, Wang P, Shi J, Zhu X, He M, Jia X, Yang X, Qiu F, Jin W, Qian M et al (2010) PML/RARalpha targets promoter regions containing PU.1 consensus and RARE half sites in acute promyelocytic leukemia. Cancer Cell 17:186–197
103. Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99:247–257
104. Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69:915–926
105. Trowbridge JJ, Snow JW, Kim J, Orkin SH (2009) DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. Cell Stem Cell 5:442–449
106. Broske AM, Vockentanz L, Kharazi S, Huska MR, Mancini E, Scheller M, Kuhl C, Enns A, Prinz M, Jaenisch R et al (2009) DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. Nat Genet 41:1207–1215
107. Tadokoro Y, Ema H, Okano M, Li E, Nakauchi H (2007) De novo DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells. J Exp Med 204:715–722
108. Melki JR, Vincent PC, Clark SJ (1999) Concurrent DNA hypermethylation of multiple genes in acute myeloid leukemia. Cancer Res 59:3730–3740
109. Toyota M, Kopecky KJ, Toyota MO, Jair KW, Willman CL, Issa JP (2001) Methylation profiling in acute myeloid leukemia. Blood 97:2823–2829
110. Wong IH, Ng MH, Huang DP, Lee JC (2000) Aberrant p15 promoter methylation in adult and childhood acute leukemias of nearly all morphologic subtypes: potential prognostic implications. Blood 95:1942–1949
111. Mizuno S, Chijiwa T, Okamura T, Akashi K, Fukumaki Y, Niho Y, Sasaki H (2001) Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia. Blood 97:1172–1179
112. Jiang Y, Dunbar A, Gondek LP, Mohan S, Rataul M, O'Keefe C, Sekeres M, Saunthararajah Y, Maciejewski JP (2009) Aberrant DNA methylation is a dominant mechanism in MDS progression to AML. Blood 113:1315–1325
113. Bullinger L, Ehrich M, Dohner K, Schlenk RF, Dohner H, Nelson MR, van den Boom D (2010) Quantitative DNA methylation predicts survival in adult acute myeloid leukemia. Blood 115:636–642
114. Figueroa ME, Lugthart S, Li Y, Erpelinck-Verschueren C, Deng X, Christos PJ, Schifano E, Booth J, van Putten W, Skrabanek L et al (2010) DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. Cancer Cell 17:13–27
115. Stumpel DJ, Schneider P, van Roon EH, Boer JM, de Lorenzo P, Valsecchi MG, de Menezes RX, Pieters R, Stam RW (2009) Specific promoter methylation identifies different subgroups of MLL-rearranged infant acute lymphoblastic leukemia, influences clinical outcome, and provides therapeutic options. Blood 114:5490–5498

116. Schafer E, Irizarry R, Negi S, McIntyre E, Small D, Figueroa ME, Melnick A, Brown P (2010) Promoter hypermethylation in MLL-r infant acute lymphoblastic leukemia: biology and therapeutic targeting. Blood 115(23):4798–4809
117. Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR (2003) TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). Leukemia 17:637–641
118. Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H, Hayashi Y (2002) LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). Cancer Res 62:4075–4080
119. Iyer LM, Tahiliani M, Rao A, Aravind L (2009) Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. Cell Cycle 8:1698–1710
120. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324:930–935
121. Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A, Kosmider O, Le Couedic JP, Robert F, Alberdi A et al (2009) Mutation in TET2 in myeloid cancers. N Engl J Med 360:2289–2301
122. Abdel-Wahab O, Mullally A, Hedvat C, Garcia-Manero G, Patel J, Wadleigh M, Malinge S, Yao J, Kilpivaara O, Bhat R et al (2009) Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. Blood 114:144–147
123. Kurokawa R, Rosenfeld MG, Glass CK (2009) Transcriptional regulation through noncoding RNAs and epigenetic modifications. RNA Biol 6:233–236
124. Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS (2010) Non-coding RNAs: regulators of disease. J Pathol 220:126–139
125. Chen J, Odenike O, Rowley JD (2010) Leukaemogenesis: more than mutant genes. Nat Rev Cancer 10:23–36
126. Wieser R, Scheideler M, Hackl H, Engelmann M, Schneckenleithner C, Hiden K, Papak C, Trajanoski Z, Sill H, Fonatsch C (2010) microRNAs in acute myeloid leukemia: expression patterns, correlations with genetic and clinical parameters, and prognostic significance. Genes Chromosomes Cancer 49:193–203
127. Li Z, Lu J, Sun M, Mi S, Zhang H, Luo RT, Chen P, Wang Y, Yan M, Qian Z et al (2008) Distinct microRNA expression profiles in acute myeloid leukemia with common translocations. Proc Natl Acad Sci USA 105:15535–15540
128. Popovic R, Riesbeck LE, Velu CS, Chaubey A, Zhang J, Achille NJ, Erfurth FE, Eaton K, Lu J, Grimes HL et al (2009) Regulation of mir-196b by MLL and its overexpression by MLL fusions contributes to immortalization. Blood 113:3314–3322
129. Mi S, Li Z, Chen P, He C, Cao D, Elkahloun A, Lu J, Pelloso LA, Wunderlich M, Huang H et al (2010) Aberrant overexpression and function of the miR-17-92 cluster in MLL-rearranged acute leukemia. Proc Natl Acad Sci USA 107:3710–3715
130. Kotani A, Ha D, Hsieh J, Rao PK, Schotte D, den Boer ML, Armstrong SA, Lodish HF (2009) miR-128b is a potent glucocorticoid sensitizer in MLL-AF4 acute lymphocytic leukemia cells and exerts cooperative effects with miR-221. Blood 114:4169–4178
131. Kotani A, Ha D, Schotte D, den Boer ML, Armstrong SA, Lodish HF (2010) A novel mutation in the miR-128b gene reduces miRNA processing and leads to glucocorticoid resistance of MLL-AF4 acute lymphocytic leukemia cells. Cell Cycle 9:1037–1042
132. Bousquet M, Quelen C, Rosati R, Mansat-De Mas V, La Starza R, Bastard C, Lippert E, Talmant P, Lafage-Pochitaloff M, Leroux D et al (2008) Myeloid cell differentiation arrest by miR-125b-1 in myelodysplastic syndrome and acute myeloid leukemia with the t(2;11)(p21;q23) translocation. J Exp Med 205:2499–2506
133. Han YC, Park CY, Bhagat G, Zhang J, Wang Y, Fan JB, Liu M, Zou Y, Weissman IL, Gu H (2010) microRNA-29a induces aberrant self-renewal capacity in hematopoietic progenitors, biased myeloid development, and acute myeloid leukemia. J Exp Med 207:475–489

134. Klusmann JH, Li Z, Bohmer K, Maroz A, Koch ML, Emmrich S, Godinho FJ, Orkin SH, Reinhardt D (2010) miR-125b-2 is a potential oncomiR on human chromosome 21 in megakaryoblastic leukemia. Genes Dev 24:478–490
135. Wang ZY, Chen Z (2008) Acute promyelocytic leukemia: from highly fatal to highly curable. Blood 111:2505–2515
136. Petrie K, Zelent A, Waxman S (2009) Differentiation therapy of acute myeloid leukemia: past, present and future. Curr Opin Hematol 16:84–91
137. Glasow A, Barrett A, Petrie K, Gupta R, Boix-Chornet M, Zhou DC, Grimwade D, Gallagher R, von Lindern M, Waxman S et al (2008) DNA methylation-independent loss of RARA gene expression in acute myeloid leukemia. Blood 111:2374–2377
138. Fazi F, Zardo G, Gelmetti V, Travaglini L, Ciolfi A, Di Croce L, Rosa A, Bozzoni I, Grignani F, Lo-Coco F et al (2007) Heterochromatic gene repression of the retinoic acid pathway in acute myeloid leukemia. Blood 109:4432–4440
139. Carella C, Bonten J, Sirma S, Kranenburg TA, Terranova S, Klein-Geltink R, Shurtleff S, Downing JR, Zwarthoff EC, Liu PP et al (2007) MN1 overexpression is an important step in the development of inv(16) AML. Leukemia 21:1679–1690
140. Liu T, Jankovic D, Brault L, Ehret S, Baty F, Stavropoulou V, Rossi V, Biondi A, Schwaller J (2010) Functional characterization of high levels of meningioma 1 as collaborating oncogene in acute leukemia. Leukemia 24:601–612
141. Jain N, Rossi A, Garcia-Manero G (2009) Epigenetic therapy of leukemia: an update. Int J Biochem Cell Biol 41:72–80
142. Kantarjian H, Issa JP, Rosenfeld CS, Bennett JM, Albitar M, DiPersio J, Klimek V, Slack J, de Castro C, Ravandi F et al (2006) Decitabine improves patient outcomes in myelodysplastic syndromes: results of a phase III randomized study. Cancer 106:1794–1803
143. Silverman LR, McKenzie DR, Peterson BL, Holland JF, Backstrom JT, Beach CL, Larson RA (2006) Further analysis of trials with azacitidine in patients with myelodysplastic syndrome: studies 8421, 8921, and 9221 by the Cancer and Leukemia Group B. J Clin Oncol 24:3895–3903
144. Blum W, Klisovic RB, Hackanson B, Liu Z, Liu S, Devine H, Vukosavljevic T, Huynh L, Lozanski G, Kefauver C et al (2007) Phase I study of decitabine alone or in combination with valproic acid in acute myeloid leukemia. J Clin Oncol 25:3884–3891
145. Cashen AF, Schiller GJ, O'Donnell MR, DiPersio JF (2010) Multicenter, phase II study of decitabine for the first-line treatment of older patients with acute myeloid leukemia. J Clin Oncol 28:556–561
146. Issa JP, Garcia-Manero G, Giles FJ, Mannari R, Thomas D, Faderl S, Bayar E, Lyons J, Rosenfeld CS, Cortes J et al (2004) Phase 1 study of low-dose prolonged exposure schedules of the hypomethylating agent 5-aza-2′-deoxycytidine (decitabine) in hematopoietic malignancies. Blood 103:1635–1640
147. Blum W, Garzon R, Klisovic RB, Schwind S, Walker A, Geyer S, Liu S, Havelange V, Becker H, Schaaf L et al (2010) Clinical response and miR-29b predictive significance in older AML patients treated with a 10-day schedule of decitabine. Proc Natl Acad Sci USA 107(16):7473–7478
148. Jiemjit A, Fandy TE, Carraway H, Bailey KA, Baylin S, Herman JG, Gore SD (2008) p21 (WAF1/CIP1) induction by 5-azacytosine nucleosides requires DNA damage. Oncogene 27:3615–3623
149. Palii SS, Van Emburgh BO, Sankpal UT, Brown KD, Robertson KD (2008) DNA methylation inhibitor 5-Aza-2′-deoxycytidine induces reversible genome-wide DNA damage that is distinctly influenced by DNA methyltransferases 1 and 3B. Mol Cell Biol 28:752–771
150. Insinga A, Monestiroli S, Ronzoni S, Gelmetti V, Marchesi F, Viale A, Altucci L, Nervi C, Minucci S, Pelicci PG (2005) Inhibitors of histone deacetylases induce tumor-selective apoptosis through activation of the death receptor pathway. Nat Med 11:71–76

151. Kramer OH, Muller S, Buchwald M, Reichardt S, Heinzel T (2008) Mechanism for ubiquitylation of the leukemia fusion proteins AML1-ETO and PML-RARalpha. FASEB J 22:1369–1379
152. Nebbioso A, Clarke N, Voltz E, Germain E, Ambrosino C, Bontempo P, Alvarez R, Schiavone EM, Ferrara F, Bresciani F et al (2005) Tumor-selective action of HDAC inhibitors involves TRAIL induction in acute myeloid leukemia cells. Nat Med 11:77–84
153. Mercurio C, Minucci S, Pelicci PG (2010) Histone deacetylases and epigenetic therapies of hematological malignancies. Pharmacol Res 62(1):18–34
154. Garcia-Manero G, Assouline S, Cortes J, Estrov Z, Kantarjian H, Yang H, Newsome WM, Miller WH Jr, Rousseau C, Kalita A et al (2008) Phase 1 study of the oral isotype specific histone deacetylase inhibitor MGCD0103 in leukemia. Blood 112:981–989
155. Cameron EE, Bachman KE, Myohanen S, Herman JG, Baylin SB (1999) Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. Nat Genet 21:103–107
156. Gore SD, Baylin S, Sugar E, Carraway H, Miller CB, Carducci M, Grever M, Galm O, Dauses T, Karp JE et al (2006) Combined DNA methyltransferase and histone deacetylase inhibition in the treatment of myeloid neoplasms. Cancer Res 66:6361–6369

The Liver-Specific MicroRNA miR-122: Biology and Therapeutic Potential

Witold Filipowicz and Helge Großhans

Abstract MicroRNAs (miRNAs) are small noncoding RNAs that regulate the expression of a large fraction of genes in animals, plants, and protozoa. miRNA-mediated gene repression occurs posttranscriptionally, generally by base-pairing to the 3′-untranslated regions of target mRNAs, which inhibits protein synthesis and destabilizes the mRNA. In this chapter, we discuss the biological functions of miR-122, a highly abundant, liver-specific miRNA. We will review how studies of miR-122 helped to establish important new paradigms of miRNA-mediated regulation, as well as identifying miR-122 as a factor implicated in important human diseases, including cancer and hepatitis C. We discuss antisense strategies targeting miR-122 as a potential therapeutic approach to treat hepatitis C and possibly other diseases.

1 Introduction

1.1 History of miRNAs: What Are miRNAs and How Were They Found

MicroRNAs are a large class of genomically encoded, regulatory RNAs of ~22 nucleotides (nt). The first miRNA, *lin-4*, was identified in the nematode *Caenorhabditis elegans* through a mutation that caused specific alterations in developmental cell fates [1]. At the time, no homologues of *lin-4* could be identified in other organisms, suggesting the possibility of a worm-specific function. However, the subsequent discovery of another *C. elegans* miRNA, *let-7* [2], renewed interest in these small RNAs, particularly when orthologues of *let-7* were found in other animals including humans [3]. Extensive searches for additional miRNAs, using

W. Filipowicz and H. Großhans
Friedrich Miescher Institute for Biomedical Research, Basel CH-4002, Switzerland
e-mail: witold.filipowicz@fmi.ch, helge.grosshans@fmi.ch

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_11,

both bioinformatic strategies and cloning and sequencing approaches, quickly identified dozens of additional miRNAs in *C. elegans*, but also humans and flies [4–6]. Today, we know thousands of miRNAs from various animals and plants, including unicellular algae and anemones. In humans, 721 miRNAs have been validated (miRBase 14.0; [7]), with others presumably still awaiting discovery. Among hundreds of miRNAs identified in metazoan animals, there are many that are expressed in a tissue- or a developmental stage-specific manner [8]. Individual miRNAs also vary substantially in their expression levels. Some are present at the level of 100 or fewer molecules but expression of others can exceed tens of thousands molecules per cell [9]. Vertebrate miR-122, the subject of this chapter, belongs to the most abundant known tissue-specific miRNAs. It is expressed almost exclusively in liver cells (hepatocytes) at over 50,000 copies per cell [10, 11]. As each miRNA silences dozens or even hundreds of different target mRNAs [12], this abundance of miRNAs permits regulation of various developmental and cellular processes [13]. Accordingly, miRNA dysregulation has also been implicated in various diseases, particularly cancers [14].

1.2 The Multistep Process of miRNA Biogenesis

Before a mature miRNA can silence target genes in the cytoplasm, it has to proceed along a complex biogenesis pathway [15]. Initially, a capped and polyadenylated primary miRNA (pri-miRNA) of hundreds or even thousands of nucleotides is transcribed by RNA polymerase II. In vertebrates, many miRNAs are located in the introns of protein-coding "host genes," whereas others are independent "intergenic" transcripts, but production of either type may be transcriptionally regulated by diverse transcription factors. The pri-miRNA is cleaved by the RNase Drosha and its cofactor Pasha/DGCR8 ("DiGeorge syndrome critical region gene 8") to release the precursor-miRNA (pre-miRNA). This 60–70 nucleotide long RNA, characterized by a stem-loop structure, is subsequently exported into the cytoplasm by exportin-5, where it is processed by the RNase Dicer. Dicer, along with its cofactors TRBP2 or PACT, excises a ~22 nt duplex RNA from the pre-miRNA, from which one strand will subsequently be selected as the mature miRNA and incorporated into an Argonaute (AGO) protein. The resulting miRNA-induced silencing complex (miRISC) additionally contains a protein of the GW182/TNRC6 family as an essential effector of mRNA silencing.

1.3 miRNAs Silence Target mRNAs Through an Antisense Mechanism

MicroRNAs in miRISC repress gene expression posttranscriptionally by binding to complementary sequences in the 3′ untranslated regions (UTRs) of their target mRNAs. Although miRNAs that are incorporated into mammalian AGO2 have the

capacity to direct endonucleolytic target cleavage ("slicing"), this requires perfect complementarity between the miRNA and its target, which is highly unusual in animals. Almost exclusively, animal miRNAs exhibit only partial complementarity to their targets and, regardless of the AGO with which they are associated, cause target gene repression through mechanisms distinct from slicing [16, 17]. In animals, the most stringent requirement for association with mRNA targets is frequently a contiguous and perfect Watson–Crick base-pairing of the miRNA 5′ nucleotides 2–8 [12]. These nucleotides represent the so-called "seed" region, which nucleates the miRNA–mRNA interaction. Complementarity of the miRNA 3′ half is quite relaxed although it stabilizes the interaction, particularly when the seed matching is suboptimal.

The details of miRNAs function in repressing protein synthesis are not well understood. Although it is now well established that miRNAs can cause repression of translation and/or deadenylation and subsequent degradation of target mRNA in vivo, in cells grown in culture, and in cell-free systems, mechanistic details remain largely unknown [16, 17]. In addition, the results from studies conducted in different systems and different laboratories have often been contradictory, particularly regarding the mechanism of translational inhibition [16–18]. Likewise, much remains to be learned about cellular localization of the repressive events. Components of miRISC and repressed mRNAs are enriched in different cytoplasmic structures such as processing bodies (P-bodies, also known as GW-bodies), stress granules (SGs), or multivesicular bodies (MVBs), but the precise role these cellular compartments play in mediating miRNA function is not well established [19].

Animal miRNA target sites appear to be most abundant in the 3′UTRs of target genes, but can also be in exons or even 5′UTRs [12]. However, the limited degree of complementarity between miRNAs and their targets (see above) means that target prediction algorithms are far from perfect, yielding both large numbers of false positive and false negative hits, with usually little overlap among results from different algorithms [20]. Experimental target validation is thus essential, and reporter assays are commonly used as a first step in the validation process. A diagnostic feature of a true miRNA target in these assays is that fusing its 3′UTR to a constitutively expressed reporter confers repression to the reporter. Thus, reporter repression is enhanced when miRNA levels are elevated, but impaired when miRNA levels are decreased and when the predicted target sites are mutated. If the endogenous protein levels of the target similarly change in response to modulated miRNA levels, the target can be considered validated, although functional assays will still be needed to demonstrate physiological (or pathophysiological) relevance of the interaction.

1.4 miRNA-Mediated Gene Silencing Is Reversible

miR-122, the miRNA that is the subject of this chapter, is not only important for liver physiology and pathology, but has also been instrumental in investigating the miRNA

mechanism, and in particular, the reversibility of miRNA-mediated repression. Until very recently, miRNAs have been primarily identified as negative regulators of expression of cellular mRNAs, and it remained unknown whether the inhibition of a specific mRNA can be effectively reversed. Clearly, the ability to disengage miRNPs from the repressed mRNA, or render them inactive, would make miRNA regulation much more dynamic and also more responsive to specific cellular needs.

One of the targets of miR-122 in liver cells is the mRNA encoding the high-affinity cationic amino acid transporter 1, CAT-1, which facilitates uptake of arginine and lysine in mammalian cells. Both human and mouse CAT-1 mRNAs contain in their 3′UTR several target sites for miR-122, and assays involving either chimeric mRNA reporters or endogenous human CAT-1 mRNA indicated that these sites mediate the repressive effect of miR-122 on protein synthesis [11, 21]. Repression of CAT-1 in liver cells is important to avoid hydrolysis of the plasma arginine by arginase, which is highly expressed in hepatocytes. However, under certain conditions, e.g., when urea cycle enzymes are downregulated or during liver regeneration after partial hepatectomy, CAT-1 expression is induced, most likely to sustain the import of cationic amino acids required for protein synthesis in hepatocytes [22]. Consistent with the above considerations, CAT-1 is expressed fairly ubiquitously but its levels vary significantly in different cells and tissues and are subject to extensive regulation at both transcriptional and posttranscriptional levels (reviewed in [23]). For example, in rat C6 glioma cells, transcription of the CAT-1 gene and stability and translation of the mRNA are strongly upregulated in response to different types of cellular stress, including amino acid deprivation [23].

Experiments carried out in human hepatoma Huh7 cells revealed that endogenous CAT-1 mRNA and reporters bearing its 3′UTR can be relieved from miR-122 repression by subjecting the cells to different stress conditions, i.e., amino acid starvation, oxidative, and endoplasmic reticulum stress [21]. The derepression was accompanied by the release of CAT-1 mRNA from P-bodies, cytoplasmic aggregates known to function in storage of translationally repressed mRNAs [24]. The CAT-1 mRNA was preferentially recruited to large polysomes, consistent with miR-122 inhibition occurring at the level of translational initiation [21]. Investigation of the mechanism of the stress-induced relief from miR-122 repression pointed to a role of the ELAV family RNA binding protein HuR in this process. In response to different types of cellular stress, HuR is known to be mobilized from the nucleus to the cytosol, where it modulates translation and/or stability of different mRNAs by binding to AU-rich elements in the 3′UTR [25]. It appears that HuR, by binding to the AU-rich sequences positioned in the 3′UTR even far away from the miR-122 recognition sites in the CAT-1 mRNA, either prevents the repressive function of miRISC or displaces it from the CAT-1 mRNA. More recently, other examples of proteins binding to mRNA 3′UTRs and modulating activity of miRNAs have been described [26–28]. Hence, it appears that regulation of miRNA repression by RNA binding proteins is probably a widespread phenomenon. It is also likely that, conversely, miRISC will influence the activity of tRNA binding proteins that bind to the 3′UTR and regulate mRNA stability or localization.

2 miR-122 and Cancer

2.1 *Prelude: The* hcr *RNA and Hepatocellular Carcinoma in Woodchucks*

Twenty years ago, an unusual RNA was discovered through its association with liver cancer or hepatocellular carcinoma (HCC) in woodchucks that suffered from chronic infection with the woodchuck hepatitis virus. In some of the tumors, the *MYC* proto-oncogene was overexpressed, and in one case, this was due to a DNA rearrangement that had translocated *c-MYC* into the *hepatocellular carcinoma-related* (*hcr*) locus [29, 30]. The normal function of *hcr* remained mysterious, as it produced two long RNAs of 4.5 and 4.7 kilobasepairs (kb), respectively, which seemed to have little protein-coding potential [31]. Whereas the 4.7 kb transcript was of low abundance and polyadenylated, the smaller transcript was truncated at the 3′ end, lacking the poly(A) tail, yet much more abundant. Both transcripts were specifically and abundantly expressed in liver, but not in other tissues such as spleen, kidney, lung, heart, or intestine [31].

It was only 13 years later that a function of the *hcr* transcript emerged, when Tuschl and colleagues cloned miRNAs from different mouse tissues. One of the newly identified miRNAs was miR-122, which accounted for some 70% of cloned liver miRNAs, and which mapped to the conserved *hcr* locus [10]. Intriguingly, miR-122 mapped to the 3′ end of the transcript, the sequence that is present in the 4.7 kb but not the 4.5 kb transcript [11]. It thus appears that one of these transcripts is the pri-miR-122, the other a processing product, from which the pre-miR-122 has been released, and which is stable for unknown reasons. Consistent with the liver-specific expression of *hcr* in woodchucks, mature miR-122 was found to accumulate almost exclusively, and abundantly, in the livers of humans, mouse [10, 11], and zebrafish [32]. miR-122 targets hundreds of mRNAs expressed in liver cells [33, 34], including a large fraction of genes that accumulate in a circadian fashion [35]. Accordingly, transcription of the miR-122 gene is also under circadian regulation, although the functional relevance of this observation is not completely clear, as accumulation of the mature miRNAs appears to be stable over the 24-h period that makes up a day [35]. Nonetheless, it appears that miR-122 is of primary importance for maintaining liver homeostasis.

2.2 *Misexpression of miR-122 in HCC*

Although the connection to woodchuck HCC had implicated miR-122 in liver carcinogenesis, there was little evidence for a misexpression of (pri-) *miR-122* in these tumors, and it appears that woodchuck HCC is a consequence of overexpression of *MYC* from a strong, liver-specific promoter rather than reduced miR-122 levels [11, 31].

More recently, however, several lines of research have provided support for a function of miR-122 in HCC in humans and mice, with most reports suggesting a role for miR-122 as a tumor suppressor [36–43]. Specifically, reduced expression of miR-122 has been observed at high frequency in HCC primary tumors relative to cirrhotic liver [40], and in HCC primary tumors relative to normal adjacent liver tissue [36, 37, 41]. Tsai et al. further distinguished advanced (Tumor node metastasis stage 3, T3; intrahepatic metastasis) from less advanced (T1 stage) liver tumors and found that miR-122 was only reduced in the more advanced tumors [36].

A general heterogeneity of miR-122 expression in primary tumors was also reported by others [38, 41], and Coulouarn et al. demonstrated the expression of miR-122 to be prognostic for tumor invasiveness; time to recurrence; and overall patient survival, with high miR-122 levels signaling a better prognosis [38]. Low miR-122 levels in primary tumors also predicted reduced time to recurrence in an independent study, although in that case overall patient survival was not affected in a statistically significant manner [43]. Finally, miR-122 is also part of a larger miRNA prognostic signature, comprising 20 miRNAs, that was found to predict patient survival and metastasis [44]. Indeed, in most of the studies profiling miRNA expression in HCC, larger panels of miRNAs were found to be deregulated; however, only miR-122 appeared to be deregulated consistently across studies, whereas levels of other miRNAs appeared much more variable.

Jointly, the expression studies and the clinical data strongly suggest that miR-122 could have a tumor-suppressive function in HCC. However, one report observed that in Hepatitis C Virus (HCV)-derived HCCs, miR-122 levels tended to be upregulated when fixed tissue samples were queried [42]. As this report specifically investigated HCV-induced HCC, an intriguing possibility is that alterations in miR-122 levels in HCC might depend on tumor etiology. Supporting this view, Coulouarn et al. observed in their study [38] that reduced miR-122 levels were a specific property of HCC arising in HBV, but not in HCV-infected livers, where they saw neither a decrease nor an increase. Such differences are particularly intriguing when viewed in the context of miR-122's function in HCV replication, which we discuss below.

When considering the data of Varnholt et al. [42] together with the other studies, it would appear that some HCCs can evade the tumor-suppressing function of miR-122, or, more intriguingly, that miR-122 might be Janus-faced, acting as tumor suppressor in some settings, but as "oncomiR" in others. A distinction between these possibilities will require more extended analysis of HCCs of different etiologies and, importantly, functional and clinical data linking elevated miR-122 levels in HCV-derived HCCs to clinical outcomes.

Although miR-122 levels seem to be reduced in many HCCs, the cause(s) remain generally unknown. Integrity of the *miR-122* genomic locus in HCC has not been investigated so that loss of an miR-122 allele remains possible. However, given the complexity of miRNA biogenesis, other, posttransciptional defects are just as likely. Indeed, recent work in mouse has shown that miR-122 undergoes posttranscriptional modification, with the cytoplasmic poly(A) polymerase mGLD2 adding one or few adenosines specifically to the 3′ end of mature miR-122 [45].

Deletion of *mGLD2* leads to a marked reduction in mature miR-122 levels and activity. It would thus be of interest to determine whether *mGLD2* knockout mice, with reduced levels of functional miR-122, are particularly susceptible to HCC and, conversely, whether human GLD2 is mutated in HCC, specifically in those where miR-122 levels are reduced.

At this point, it is also not clear when during HCC development miR-122 expression starts to be deregulated. A study of rats put at risk for HCC through an appropriate diet revealed that miR-122 expression was still normal in preneoplastic nodules, but decreased in the tumors, indicating miR-122 misregulation as a later event in tumorigenesis [41]. Nonetheless, in NASH (nonalcoholic steatohepatitis), a risk factor and potential precursor to HCC, miR-122 levels are reduced relative to controls, both in mouse [46] and humans [47]. However, in the case of humans, both NASH patients and controls suffered from metabolic syndrome, so that it is less clear how the observation relates to healthy people, with healthy livers.

2.3 *miR-122 as a Tumor Suppressor: Evidence from In Vitro and In Vivo Mechanistic Studies*

To identify the mechanisms by which miR-122 affects HCC formation, its function has been studied extensively in vitro and in vivo. Different hepatic cell lines were found to differ remarkably in their levels of miR-122 accumulation, with most cell lines such as Hep G2, SK-Hep1, Mahlavu, and others exhibiting low or undetectable expression but Huh-1, -6, -7, and Hep40 exhibiting high expression levels [11, 38]. Beyond comparing the cell lines, it is thus possible to modulate miR-122 expression by repressing its function using antisense oligonucleotide in Huh-7 or other highly expressing cells, or by elevating its levels through transfection into poorly expressing cell lines. Such studies have been performed in several laboratories, and the consistent result was that elevated levels of miR-122 reduce migration and invasiveness in vitro, whereas reduced levels promote these properties [36–38, 43]. Moreover, this association also holds in vivo, at least with ectopically and orthotopically transplanted tumors, which are more locally invasive and grow faster when miR-122 levels are low [36, 37]. It is less clear whether miR-122 directly affects cell proliferation, as in vitro results have been split on this issue [36, 37], perhaps owing to different miR-122 concentrations achieved in these experiments. Beyond any potential direct function on cell replication, reduction in tumor size in vivo appears to be mediated at least in part through modulation of angiogenesis, which is decreased with elevated miR-122 levels [36, 37].

What are the targets through which miR-122 achieves its functions? As for any miRNA, target identification has been difficult so that many relevant targets may yet await discovery. However, a number of interesting candidates are known, in particular cyclin G1 [40], a cyclin of little known function that appears to regulate

transcription of the p53 tumor suppressor gene [43]; Serum Response Factor SRF, IGF1 receptor tyrosin kinase Igf1R, and A Disintegrin and A Metalloproteinase protein (ADAM) 10 [37]; and ADAM-17 (also known as TACE, TNFα-converting enzyme) [36]. Expression of these proteins is inversely correlated with miR-122 levels, reporter genes carrying their 3′UTRs are repressed by miR-122, suggesting direct regulation, and, most importantly, overexpression of these proteins can copy aspects of reduced miR-122 expression, whereas their decrease can suppress some of the miR-122 phenotypes in vitro and in vivo. Given the potential metastasis-inhibiting function of miR-122, ADAM proteins are interesting for their role in cell adhesion, although functional data for an involvement of these proteins in cancer are still sparse [48]. Apoptosis inhibitor Bcl-w is another apparent target, but has so far only been validated in vitro, where its modulation through miR-122 does indeed appear to alter cell survival and apoptosis [49].

3 miR-122 and HCV

3.1 miR-122 Has Stimulatory Effects on Accumulation and Translation of HCV RNA

As discussed above, some evidence suggests that miR-122 levels might be differently affected in tumors arising from HBV and HCV infection. Intriguingly then, Jopling et al. discovered that miR-122 is essential for accumulation of HCV RNA in cultured human hepatoma Huh7 cells expressing an HCV replicon [50]. HCV is a positive strand single-stranded RNA virus belonging to Flaviviridae family. Its 9.6-kb genomic RNA is translated to a single long polyprotein, which is subsequently processed to about ten viral polypeptides by cellular and viral proteases. HCV contains a 5′-noncoding region (5′-NCR) of 320 nt, part of which functions as an Internal Ribosome Entry Site (IRES), responsible for the cap-independent initiation of translation of HCV RNA in infected cells. Although in cellular mRNAs most of the functional miRNA-binding sites are present in 3′-noncoding regions, HCV RNA contains two conserved hexamer sequences complementary to the seed region of miR-122 in the 5′-NCR, close to the 5′ end of HCV RNA. Jopling et al. demonstrated that simultaneous recognition of these sites by miR-122 is essential for viral RNA accumulation as demonstrated by mutational analysis (including rescue of the HCV RNA mutations by compensatory mutations in miR-122) and experiments involving sequestration of miR-122 with antisense oligonucleotides [50, 51]. The two miR-122 HCV sites are separated from each other by a highly conserved 14-nt-long spacer, and both the length of this spacer and its sequence were also found to be important for accumulation of HCV RNA.

Additional evidence of the stimulatory effect of miR-122 on HCV RNA yields was provided by experiments performed with human embryonic kidney epithelial

(HEK293) cells, which normally do not express this miRNA. Exogenously supplied miR-122 strongly increased accumulation of HCV RNA in HEK293 cells expressing the HCV replicon [52]. Importantly, miR-122 was also found to have a stimulatory effect on HCV replication and infectious virus production in Huh7 cells infected with a J6/JFH-1 chimeric HCV, which produces the full replication cycle in cultured cells [53]. In this system, Randall et al. also established that RNAi-mediated knockdown of genes encoding factors involved in miRNA biogenesis and function such as Dicer, Drosha, and Ago proteins markedly decreased HCV replication [53], providing additional, although indirect, support for the role of miR-122 in HCV replication. The mechanism through which the interaction of HCV RNA with miR-122 leads to the increased viral RNA level remains unknown. It is unlikely to be due to the effect of miR-122 on HVC RNA translation or stability ([50, 52]; but see below). In addition, miR-122 sequestration was recently found to have only moderate effect on the rate of HCV RNA synthesis in Huh7 cells [54]. Possibly, miR-122 modulates the structure of HCV RNA, making it a better template for replication, or it helps to localize the RNA to a specific cellular compartment that promotes replication or some posttranscriptional regulatory steps.

Since miRNAs generally repress protein synthesis by binding to the 3′UTRs of their targets (see above), the observation that miR-122 stimulates HCV replication when interacting with the 5′-end-proximal sequences in HCV RNA was surprising and raised questions regarding the position-dependence of miRNA effects. Two groups have investigated the effect of miR-122 target site placement on the translation of reporter mRNAs. Their findings indicated that placement of miR-122 sites in the 3′UTR of a reporter resulted in marked, miR-122-dependent, inhibition of translation, both for sites corresponding to a short artificial sequence complementary to the miR-122 seed region and an HCV 5′-NCR fragment encompassing one or both authentic miR-122 sites [51, 55]. In contrast, translation of reporters containing the HCV 5′-NCR sequence (which includes the miR-122 sites and the IRES region) in their 5′UTRs was either not affected [51] or stimulated [55] by miR-122. More detailed analysis of the stimulatory effect of the 5′-proximal placement of miR-122 sites indicated that miR-122 stimulation was not confined to Huh7 endogenously expressing miR-122 but also occurred in HeLa cells and rabbit reticulocyte lysates supplemented with exogenous miR-122; in addition, translational stimulation was also observed in the context of a full-length HCV genome expressed in Huh7 cells. In reticulocyte lysates, miR-122 appeared to stimulate 48S complex formation, an early step of the initiation reaction [55]. Interestingly, the stimulation of HCV RNA accumulation reported by Jopling et al. [51] and the enhancing effect on translation reported by Henke et al. [55] required the presence of both miR-122 sites conserved in the HCV 5′-NCR. This contrasts with the situation at the 3′UTR where either miR-122 site was found to be active in translational repression, independently of the other. This distinction underscores a rather specialized role of the HCV 5′-NCR miR-122 sites. It will be interesting to elucidate a precise molecular mechanism of their function. Recently, stimulatory effects on translation of other miRNAs

acting at either at the 5′UTR or 3′UTR of cellular mRNAs have been reported, but as in the case of miR-122 the mechanism of their stimulatory function is not known [56–58].

3.2 *Silencing of miR-122 in the Liver of Rodents and Nonhuman Primates*

The demonstration that miR-122 is required for replication of HCV RNA in cultured cells either expressing an HCV replicon or infected with HCV virus raised much interest in a potential use of antisense oligonucleotides (often referred to as anti-miRs or antagomirs) to target miR-122 in liver cells, as potential therapeutic agents against HCV infection. The fact that anti-miRs can effectively inhibit specific miRNAs when transfected into cells grown in culture or injected into mice in vivo has been documented in many studies [33, 34, 59–61]. Generally, oligonucleotides used as anti-miRs contain modified nucleotides such as LNA (locked nucleic acid) derivatives or 2′-*O*-methyl or 2′-*O*-methoxyethyl ribonucleotides to both strengthen the base paring of oligonucleotides with target miRNAs and to decrease their sensitivity to cellular and serum nucleases. For the administration in vivo, antagomirs can be conjugated to lipophilic compounds such as cholesterol to facilitate cellular uptake.

The first successful studies demonstrating a potent activity of antagomirs in vivo were performed with oligonucleotides targeting miR-122, which were administered to mice by intravenous or intraperitoneal injection. Injection of either cholesterol-conjugated 2′-*O*-methyl phosphorothioate-containing [33] or nonconjugated 2′-*O*-methoxyethyl phosphorothioate-modified [34] anti-miR-122 oligonucleotides resulted in efficient and specific elimination or blocking of miR-122 in liver. The effect of the injected anti-miRs lasted several weeks and was not associated with changes in body weight and liver morphology or in elevation of serum markers (the transaminases ALT and AST) indicative of liver toxicity. Profiling of mRNAs isolated from livers of anti-miR-122-treated animals demonstrated a 1.4–4-fold elevation in the level of several hundreds of mRNAs, many of them containing hexamer or heptamer sequences complementary to the miR-122 seed region [33, 34]. Hence, these mRNA likely represent primary targets of miR-122 in liver cells, and this has been confirmed for selected upregulated mRNAs by performing reporter assays in cultured cells [33].

Interestingly, a similar number of mRNAs (~300) was downregulated upon anti-miR-122-treatment. These mRNAs thus likely represent secondary targets, expression of which may depend on activity of primary miR-122 targets, such as specific inhibitors or transcription factors which are under direct control of miR-122. Several enzymes of the mevalonate pathway, involved in cholesterol biosynthesis, were identified among the secondary targets of miR-122. Consistently, antagomir-122-treated animals were found to have plasma cholesterol levels reduced by 30–40% when compared with control animals.

Inhibition of miR-122 with antagomir had also a beneficial effect on hepatic fatty acid metabolism, improving liver steatosis (lipid droplet accumulation in liver cells) and decreasing triglyceride accumulation in high-fat-diet animals [34]. More recently, similar results were obtained with unconjugated LNA-modified anti-miRs systematically administered to mice [59]. In summary, these findings not only provided a proof of concept for the feasibility of antagomir-based miRNA inhibition in vivo, but also extended potential therapeutic applications of these compounds beyond the treatment of hepatitis C, to conditions such as cholesterolemia or obesity.

The HCV virus does not infect mouse cells and there is no rodent model available for studies of hepatitis C. However, effects of anti-miR122 antagomirs were more recently investigated in nonhuman primates, i.e., African green monkeys [62] and chimpanzees [63], with the latter being susceptible to chronic infection with HCV. Systemic administration of unconjugated LNA-modified anti-miR-122 by repeated intravenous infusions (three infusions over 5 days) to African green monkeys resulted in dose-dependent, up to 40% decrease in plasma cholesterol level [62]. This effect lasted for 7 weeks and afterwards the level of cholesterol gradually returned to normal over a period of 3 months, with parallel normalization of the miR-122 level. As in mice, administration of LNA-modified anti-miR-122 was not associated with any liver or other organ toxicity or histopathological changes.

Yet more relevant findings were obtained with chimpanzees [63]. Intravenous administration into animals chronically infected with HCV of an LNA anti-miR-122 (5 mg kg^{-1} body weight per dose) for 12 weeks at weekly intervals led to marked suppression of viral RNA levels in both serum and liver biopsy material. The maximum decrease, 2.6 and 2.3 orders of magnitude for serum and liver respectively, was observed 2 weeks after the end of treatment. Importantly, no resistance to the therapy emerged during the study period as evidenced by absence of a rebound of viremia during the treatment and of viral RNA with mutations in miR-122 sites through the end of the follow-up period. These observations are consistent with the conservation, and thus presumably importance for viral fitness, of both miR-122 complementary sites in all HCV genotypes and subtypes [63]. The high barrier to developing resistance seen with the LNA anti-miR-122 therapy contrasts with treatment of infected chimpanzees with other drugs, directly targeting viral enzymes, such as inhibitors of HCV RNA polymerase [64].

Like in mice and African green monkeys, administration of anti-miR-122 to chimpanzees led to 30–40% decrease in serum cholesterol [63]. Also similar to other animal systems, the treatment was not associated with appreciable toxic effects. Clearly, the studies with nonhuman primates point to the feasibility and safety of using LNA, and perhaps other formulations of anti-miR oligonucleotides, to antagonize specific miRNAs in vivo with the aim to treat hepatitis C and other diseases, to which miRNAs may contribute.

3.3 miR-122 Status in Patients with Chronic Hepatitis C

HCV infection and Chronic Hepatitis C (CHC) are major causes of liver cirrhosis and liver cancer. Approximately 170 million people world-wide are chronically

infected with HCV but current therapy, involving combination of pegylated interferon-α (IFN-α) and ribavirin (a nucleoside analog), is successful in only ~50% patients and strongly depends on the genotype of the infecting virus [65].

Since the therapy requires administration of IFN-α and ribavirin over many months and is associated with strong side effects, efforts were undertaken in recent years to identify factors, in addition to HCV genotype, that could help to predict the therapy outcome and distinguish responders from nonresponders prior to the initiation of therapy. Profiling of gene expression in liver biopsies collected before the therapy revealed that patients found later to be nonresponders had increased expression of IFN stimulated genes (ISGs) already prior to the administration of exogenous IFN-α ([66], and references therein). In a most recent study of this type, expression of ISGs was analyzed in paired liver biopsies collected from CHC patients before treatment and 4 h after the first injection of pegylated IFN-α. Analysis of 16 patients, among them 10 representing future rapid responders (>2 $\log^{10}$ decrease in serum HCV level at week 4) and 6 showing no rapid response to therapy, demonstrated that nonresponders had a preactivated IFN system in the liver prior to the treatment and showed no significant changes in expression of ISGs in liver upon pegIFN-α administration [66]. In contrast, future responders generally lacked significant preactivation of the IFN system and responded to pegIFN-α administration with increased activity of ISGs. Taken together, these results indicate that patients who do not benefit from the therapy have their IFN system turned on even in the absence of exogenous IFN administration but, due to some downstream defects in the IFN defense pathway, they cannot effectively combat the viral infection.

Profiling of mRNAs in biopsies collected prior to therapy made it also possible to identify sets of genes, the expression of which is predictive, with a relatively high confidence, of the therapy outcome. For example, one of the studies identified 29 genes, 76% of which represent ISGs, which predict responsiveness to the therapy with an error rate of 4.3% [66]. Analysis of larger cohorts of CHC patients may eventually lead to the development of predictive tests which in the future could be used for directing the pegIFN-α therapy only to patients who will benefit from it.

What are the levels of miR-122 in CHC patients that do or do not respond to the pegIFN-α therapy? Is there a correlation between miR-122 levels and HCV RNA titer in the liver and serum of CHC patients? Availability of liver biopsies from CHC patients, including paired biopsies collected before and after the first injection of pegylated IFN-α, made it possible to measure miR-122 levels in patients' liver and correlate them with responsiveness to therapy. The measurements provided also an opportunity to assess the effect of IFN on expression of miR-122 in human liver. This was important since Pedersen et al. have recently reported that treatment of human Huh7 cells and mouse primary hepatocytes with IFN markedly down-regulates miR-122 levels, and that miR-122 down-regulation may represent one of the mechanisms underlying the anti-HCV effect of IFN [67].

Since miR-122 is important for effective HCV replication, intuitively one would expect lower levels of miR-122 in CHC patients responding to the therapy and higher levels in nonresponders. Surprisingly, the opposite was found to be true [68].

Patients who did not respond to therapy (no decrease in viral load of more than 2 $\log_{10}$ at week 12) had several fold lower miR-122 levels than patients with a strong response to IFN-α (no detectable HCV-RNA at week 12). The difference in miR-122 levels between responders and nonresponders was also apparent when only patients infected with difficult-to-treat HCV genotypes 1 and 4 were examined, indicating that the difference is not due to the biased distribution of HCV genotypes between the two response groups. Moreover, measurements of HCV-RNA in liver and serum of patients showed no positive correlation between miR-122 and viral load [68]. Since miR-122 is important for HCV replication in Huh7 cells and in the chimpanzee liver, the finding of low miR-122 levels in patients not responding to therapy and a lack of correlation between miR-122 expression and viral load is rather unexpected. It is possible that even the low miR-122 levels found in nonresponders are sufficient to support HCV replication. Alternatively, measurements of total miR-122 in the biopsy extracts may not reflect the miR-122 levels in the fraction of hepatocytes (generally considered to be small) that are infected with HCV.

Measurements of miR-122 in paired liver biopsies collected prior to and 4 h after administration of pegIFN-α revealed no IFN-induced decrease of the miR-122 level in both responding or nonresponding groups of patients. In addition, no pronounced decrease of the miR-122 level was found in livers of mice even after 4 days following multiple injections with mouse IFN-α [68]. Hence, miR-122 is not an early IFN-responsive gene in vivo and it is unlikely that the antiviral effects of IFN during CHC therapy can be explained by changes in the level of miR-122 or other miRNAs [68]. However, it cannot be excluded that prolonged activation of the IFN system in patients not responding to the therapy contributes to the miR-122 down-regulation observed in these patients. The finding that miR-122 is significantly lower in nonresponders than in responders makes it a convenient marker, together with preactivated ISGs, for predicting the outcome of IFN therapy. If miR-122 is indeed required for HCV persistence in CHC patients, its low levels in non-responders to IFN therapy might yet turn into a therapeutic benefit in the future, providing these patient with a “head-start” in potential interventions involving the use of miR-122 antagomirs.

4 Outlook

The recent demonstration that inhibition of miR-122 in chimpanzees chronically infected with HCV leads to significant suppression of HCV RNA levels without apparent toxicity or emergence of resistant viral RNA mutants indicates that this approach has the potential to replace or enhance other therapies, particularly for treatment of patients infected with a difficult-to-treat genotype 1 HCV. The anti-miR might be combined with a classical pegIFN-α/ribovirin therapy or inhibitors of viral proteases or RNA polymerases currently undergoing clinical trials. It will be important to establish how well the anti-miR-122 oligonucleotides are tolerated in

humans. Currently ongoing Phase 1 clinical trials with healthy volunteers should soon provide an answer to this question (http://www.santaris.com). Likewise, it will be important to further investigate the potential of anti-miR-122 oligonucleotides to treat disorders associated with lipid metabolism. Measurements of levels of miR-122, and possibly also other miRNAs, could also be used, in combination with mRNA profiling, for predicting patients' response to IFN therapy. A recent demonstration that measurements of miR-122 in plasma represent a very sensitive and specific readout for liver injury are promising in this regard [69].

Clinical and experimental data from both in vitro and in vivo systems further provide support for a role of miR-122 as a tumor suppressor in the liver. In a next step, it would now be interesting to study its function in *miR-122* knockout mice, or perhaps *miR-122* heterozygotes, asking in particular whether these mice would be more susceptible to development of autochthonous liver cancers than wild-type animals. Interestingly, mice with a liver-specific knockout of the pre-miRNA processing enzyme dicer are indeed more susceptible to HCC development [70]. If further studies support a tumor-suppressive function of miR-122, it might become a valid target for miRNA replacement studies. However, given that *miR-122* overexpression carries its own issues with regard to lipid metabolism and, possibly, HCV susceptibility, this would have to be approached very carefully. Determining the causes of pathologically reduced miR-122 levels in HCC, and reverting them specifically by curing the underlying mechanism, might then be a more palatable alternative. Indeed, since increases in miR-122 levels sensitize cancer cells to the known chemotherapeutics sorafenib [37] and doxorubicin [43], less than complete restoration of miR-122 levels to wild-type amounts might suffice for therapeutic gain.

References

1. 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
2. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G (2000) The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 403:901–906
3. Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, Hayward DC, Ball EE, Degnan B, Muller P, Spring J, Srinivasan A, Fishman M, Finnerty J, Corbo J, Levine M, Leahy P, Davidson E, Ruvkun G (2000) Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 408:86–89
4. Lee RC, Ambros V (2001) An extensive class of small RNAs in Caenorhabditis elegans. Science 294:862–864
5. Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 294:858–862
6. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. Science 294:853–858
7. Griffiths-Jones S, Saini HK, Dongen SV, Enright AJ (2007) miRBase: tools for microRNA genomics. Nucleic Acids Res 36:D154–D158

8. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst AO, Landthaler M, Lin C, Socci ND, Hermida L, Fulci V, Chiaretti S, Foà R, Schliwka J, Fuchs U, Novosel A, Müller RU, Schermer B, Bissels U, Inman J, Phan Q, Chien M, Weir DB, Choksi R, De Vita G, Frezzetti D, Trompeter HI, Hornung V, Teng G, Hartmann G, Palkovits M, Di Lauro R, Wernet P, Macino G, Rogler CE, Nagle JW, Ju J, Papavasiliou FN, Benzing T, Lichter P, Tam W, Brownstein MJ, Bosio A, Borkhardt A, Russo JJ, Sander C, Zavolan M, Tuschl T (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. Cell 129:1401–1414
9. Lim LP, Lau NC, Weinstein EG, Abdelhakim A, Yekta S, Rhoades MW, Burge CB, Bartel DP (2003) The microRNAs of Caenorhabditis elegans. Genes Dev 17:991–1008
10. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T (2002) Identification of tissue-specific microRNAs from mouse. Curr Biol 12:735–739
11. Chang J, Nicolas E, Marks D, Sander C, Lerro A, Buendia MA, Xu C, Mason WS, Moloshok T, Bort R, Zaret KS, Taylor JM (2004) miR-122, a mammalian liver-specific microRNA, is processed from hcr mRNA and may downregulate the high affinity cationic amino acid transporter CAT-1. RNA Biol 1:106–113
12. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136:215–233
13. Bushati N, Cohen SM (2007) microRNA functions. Annu Rev Cell Dev Biol 23:175–205
14. Chang TC, Mendell JT (2007) microRNAs in vertebrate physiology and human disease. Annu Rev Genomics Hum Genet 8:215–239
15. Kim VN, Han J, Siomi MC (2009) Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol 10:126–139
16. Carthew RW, Sontheimer EJ (2009) Origins and mechanisms of miRNAs and siRNAs. Cell 136:642–655
17. Filipowicz W, Bhattacharyya SN, Sonenberg N (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet 9:102–114
18. Hurschler BA, Ding XC, Großhans H (2010) Translational control of endogenous MicroRNA target genes in C. elegans. Prog Mol Subcell Biol 50:21–40
19. Fabian MR, Sonenberg N, Filipowicz W (2010) Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem 79:351–379
20. Alexiou P, Maragkakis M, Papadopoulos GL, Reczko M, Hatzigeorgiou AG (2009) Lost in translation: an assessment and perspective for computational microRNA target identification. Bioinformatics 25:3049–3055
21. Bhattacharyya SN, Habermacher R, Martine U, Closs EI, Filipowicz W (2006) Relief of microRNA-mediated translational repression in human cells subjected to stress. Cell 125:1111–1124
22. Wu JY, Robinson D, Kung HJ, Hatzoglou M (1994) Hormonal regulation of the gene for the type C ecotropic retrovirus receptor in rat liver cells. J Virol 68:1615–1623
23. Hatzoglou M, Fernandez J, Yaman I, Closs E (2004) Regulation of cationic amino acid transport: the story of the CAT-1 transporter. Annu Rev Nutr 24:377–399
24. Eulalio A, Behm-Ansmant I, Izaurralde E (2007) P bodies: at the crossroads of post-transcriptional pathways. Nat Rev Mol Cell Biol 8:9–22
25. López de Silanes I, Quesada MP, Esteller M (2007) Aberrant regulation of messenger RNA 3′-untranslated region in human cancer. Cell Oncol 29:1–17
26. Kedde M, Strasser MJ, Boldajipour B, Vrielink JA, Slanchev K, le Sage C, Nagel R, Voorhoeve PM, van Duijse J, Orom UA, Lund AH, Perrakis A, Raz E, Agami R (2007) RNA-binding protein dnd1 inhibits MicroRNA access to target mRNA. Cell 131:1273–1286
27. Kim HH, Kuwano Y, Srikantan S, Lee EK, Martindale JL, Gorospe M (2009) HuR recruits let-7/RISC to repress c-Myc expression. Genes Dev 23:1743–1748
28. Huang J, Liang Z, Yang B, Tian H, Ma J, Zhang H (2007) Derepression of micro-RNA-mediated protein translation inhibition by apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G (APOBEC3G) and its family members. J Biol Chem 282:33632–33640

29. Möröy T, Marchio A, Etiemble J, Trépo C, Tiollais P, Buendia MA (1986) Rearrangement and enhanced expression of c-myc in hepatocellular carcinoma of hepatitis virus infected woodchucks. Nature 324:276–279
30. Etiemble J, Möröy T, Jacquemin E, Tiollais P, Buendia MA (1989) Fused transcripts of c-myc and a new cellular locus, hcr in a primary liver tumor. Oncogene 4:51–57
31. Möröy T, Etiemble J, Bougueleret L, Hadchouel M, Tiollais P, Buendia MA (1989) Structure and expression of hcr, a locus rearranged with c-myc in a woodchuck hepatocellular carcinoma. Oncogene 4:59–65
32. Wienholds E, Kloosterman WP, Miska E, Alvarez-Saavedra E, Berezikov E, de Bruijn E, Horvitz HR, Kauppinen S, Plasterk RH (2005) MicroRNA expression in zebrafish embryonic development. Science 309:310–311
33. Krutzfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M, Stoffel M (2005) Silencing of microRNAs in vivo with 'antagomirs'. Nature 438:685–689
34. Esau C, Davis S, Murray SF, Yu XX, Pandey SK, Pear M, Watts L, Booten SL, Graham M, McKay R, Subramaniam A, Propp S, Lollo BA, Freier S, Bennett CF, Bhanot S, Monia BP (2006) miR-122 regulation of lipid metabolism revealed by in vivo antisense targeting. Cell Metab 3:87–98
35. Gatfield D, Le Martelot G, Vejnar CE, Gerlach D, Schaad O, Fleury-Olela F, Ruskeepää AL, Oresic M, Esau CC, Zdobnov EM, Schibler U (2009) Integration of microRNA miR-122 in hepatic circadian gene expression. Genes Dev 23:1313–1326
36. Tsai WC, Hsu PW, Lai TC, Chau GY, Lin CW, Chen CM, Lin CD, Liao YL, Wang JL, Chau YP, Hsu MT, Hsiao M, Huang HD, Tsou AP (2009) MicroRNA-122, a tumor suppressor microRNA that regulates intrahepatic metastasis of hepatocellular carcinoma. Hepatology 49:1571–1582
37. Bai S, Nasser MW, Wang B, Hsu SH, Datta J, Kutay H, Yadav A, Nuovo G, Kumar P, Ghoshal K (2009) MicroRNA-122 inhibits tumorigenic properties of hepatocellular carcinoma cells and sensitizes these cells to Sorafenib. J Biol Chem 284:32015–32027
38. Coulouarn C, Factor VM, Andersen JB, Durkin ME, Thorgeirsson SS (2009) Loss of miR-122 expression in liver cancer correlates with suppression of the hepatic phenotype and gain of metastatic properties. Oncogene 28:3526–3536
39. Murakami Y, Yasuda T, Saigo K, Urashima T, Toyoda H, Okanoue T, Shimotohno K (2006) Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. Oncogene 25:2537–2545
40. Gramantieri L, Ferracin M, Fornari F, Veronese A, Sabbioni S, Liu CG, Calin GA, Giovannini C, Ferrazzi E, Grazi GL, Croce CM, Bolondi L, Negrini M (2007) Cyclin G1 is a target of miR-122a, a MicroRNA frequently down-regulated in human hepatocellular carcinoma. Cancer Res 67:6092–6099
41. Kutay H, Bai S, Datta J, Motiwala T, Pogribny I, Frankel W, Jacob ST, Ghoshal K (2006) Downregulation of miR-122 in the rodent and human hepatocellular carcinomas. J Cell Biochem 99:671–678
42. Varnholt H, Drebber U, Schulze F, Wedemeyer I, Schirmacher P, Dienes HP, Odenthal M (2008) MicroRNA gene expression profile of hepatitis C virus-associated hepatocellular carcinoma. Hepatology 47:1223–1232
43. Fornari F, Gramantieri L, Giovannini C, Veronese A, Ferracin M, Sabbioni S, Calin GA, Grazi GL, Croce CM, Tavolari S, Chieco P, Negrini M, Bolondi L (2009) MiR-122/cyclin G1 interaction modulates p53 activity and affects doxorubicin sensitivity of human hepatocarcinoma cells. Cancer Res 69:5761–5767
44. Budhu A, Jia HL, Forgues M, Liu CG, Goldstein D, Lam A, Zanetti KA, Ye QH, Qin LX, Croce CM, Tang ZY, Wang XW (2008) Identification of metastasis-related microRNAs in hepatocellular carcinoma. Hepatology 47:897–907
45. Katoh T, Sakaguchi Y, Miyauchi K, Suzuki T, Kashiwabara S, Baba T, Suzuki T (2009) Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Genes Dev 23:433–438

46. Wang B, Majumder S, Nuovo G, Kutay H, Volinia S, Patel T, Schmittgen TD, Croce C, Ghoshal K, Jacob ST (2009) Role of microRNA-155 at early stages of hepatocarcinogenesis induced by choline-deficient and amino acid-defined diet in C57BL/6 mice. Hepatology 50:1152–1161
47. Cheung O, Puri P, Eicken C, Contos MJ, Mirshahi F, Maher JW, Kellum JM, Min H, Luketic VA, Sanyal AJ (2008) Nonalcoholic steatohepatitis is associated with altered hepatic MicroRNA expression. Hepatology 48:1810–1820
48. Duffy MJ, McKiernan E, O'Donovan N, McGowan PM (2009) Role of ADAMs in cancer formation and progression. Clin Cancer Res 15:1140–1144
49. Lin CJ, Gong HY, Tseng HC, Wang WL, Wu JL (2008) miR-122 targets an anti-apoptotic gene, Bcl-w, in human hepatocellular carcinoma cell lines. Biochem Biophys Res Commun 375:315–320
50. Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P (2005) Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. Science 309:1577–1581
51. Jopling CL, Schütz S, Sarnow P (2008) Position-dependent function for a tandem MicroRNA miR-122-binding site located in the hepatitis C virus RNA genome. Cell Host Microbe 4:77–85
52. Chang J, Guo JT, Jiang D, Guo H, Taylor JM, Block TM (2008) Liver-specific microRNA miR-122 enhances the replication of hepatitis C virus in nonhepatic cells. J Virol 82:8215–8223
53. Randall G, Panis M, Cooper JD, Tellinghuisen TL, Sukhodolets KE, Pfeffer S, Landthaler M, Landgraf P, Kan S, Lindenbach BD, Chien M, Weir DB, Russo JJ, Ju J, Brownstein MJ, Sheridan R, Sander C, Zavolan M, Tuschl T, Rice CM (2007) Cellular cofactors affecting hepatitis C virus infection and replication. Proc Natl Acad Sci USA 104:12884–12889
54. Norman KL, Sarnow P (2009) Modulation of Hepatitis C virus RNA abundance and the isoprenoid biosynthesis pathway by microRNA miR-122 involves distinct mechanisms. J Virol 84:666–670
55. Henke JI, Goergen D, Zheng J, Song Y, Schüttler CG, Fehr C, Jünemann C, Niepmann M (2008) microRNA-122 stimulates translation of hepatitis C virus RNA. EMBO J 27:3300–3310
56. Vasudevan S, Steitz JA (2007) AU-rich-element-mediated upregulation of translation by FXR1 and argonaute 2. Cell 128:1105–1118
57. Vasudevan S, Tong Y, Steitz JA (2007) Switching from repression to activation: microRNAs can up-regulate translation. Science 318:1931–1934
58. Orom UA, Nielsen FC, Lund AH (2008) MicroRNA-10a binds the 5′UTR of ribosomal protein mRNAs and enhances their translation. Mol Cell 30:460–471
59. Elmén J, Lindow M, Silahtaroglu A, Bak M, Christensen M, Lind-Thomsen A, Hedtjärn M, Hansen JB, Hansen HF, Straarup EM, McCullagh K, Kearney P, Kauppinen S (2008) Antagonism of microRNA-122 in mice by systemically administered LNA-antimiR leads to up-regulation of a large set of predicted target mRNAs in the liver. Nucleic Acids Res 36:1153–1162
60. Hutvagner G, Simard MJ, Mello CC, Zamore PD (2004) Sequence-specific inhibition of small RNA function. PLoS Biol 2:E98
61. Meister G, Landthaler M, Dorsett Y, Tuschl T (2004) Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. RNA 10:544–550
62. Elmén J, Lindow M, Schütz S, Lawrence M, Petri A, Obad S, Lindholm M, Hedtjärn M, Hansen HF, Berger U, Gullans S, Kearney P, Sarnow P, Straarup EM, Kauppinen S (2008) LNA-mediated microRNA silencing in non-human primates. Nature 452:896–899
63. Lanford RE, Hildebrandt-Eriksen ES, Petri A, Persson R, Lindow M, Munk ME, Kauppinen S, Orum H (2009) Therapeutic silencing of microRNA-122 in primates with chronic Hepatitis C virus infection. Science 327:198–201
64. Chen CM, He Y, Lu L, Lim HB, Tripathi RL, Middleton T, Hernandez LE, Beno DW, Long MA, Kati WM, Bosse TD, Larson DP, Wagner R, Lanford RE, Kohlbrenner WE, Kempf

DJ, Pilot-Matias TJ, Molla A (2007) Activity of a potent hepatitis C virus polymerase inhibitor in the chimpanzee model. Antimicrob Agents Chemother 51:4290–4296
65. Sarasin-Filipowicz M (2010) Interferon therapy of hepatitis C: molecular insights into success and failure. Swiss Med Wkly 140:3–11
66. Sarasin-Filipowicz M, Oakeley EJ, Duong FH, Christen V, Terracciano L, Filipowicz W, Heim MH (2008) Interferon signaling and treatment outcome in chronic hepatitis C. Proc Natl Acad Sci USA 105:7034–7039
67. Pedersen IM, Cheng G, Wieland S, Volinia S, Croce CM, Chisari FV, David M (2007) Interferon modulation of cellular microRNAs as an antiviral mechanism. Nature 449:919–922
68. Sarasin-Filipowicz M, Krol J, Markiewicz I, Heim MH, Filipowicz W (2009) Decreased levels of microRNA miR-122 in individuals with hepatitis C responding poorly to interferon therapy. Nat Med 15:31–33
69. Laterza OF, Lim L, Garrett-Engele PW, Vlasakova K, Muniappa N, Tanaka WK, Johnson JM, Sina JF, Fare TL, Sistare FD, Glaab WE (2009) Plasma MicroRNAs as sensitive and specific biomarkers of tissue injury. Clin Chem 55:1977–1983
70. Sekine S, Ogawa R, Ito R, Hiraoka N, McManus MT, Kanai Y, Hebrok M (2009) Disruption of Dicer1 induces dysregulated fetal gene expression and promotes hepatocarcinogenesis. Gastroenterology 136:2304–2315.e1-4

Transcriptional Regulatory Networks in Embryonic Stem Cells

Yun Shen Chan, Lin Yang, and Huck-Hui Ng

Abstract Transcriptional regulation is one of the most fundamental processes in biology, governing the morphology, function, and behavior of cells and thus the survival of organisms. The embryonic stem cell (ESC) provides a good model for the understanding of transcriptional regulation in vertebrate systems. Recent efforts have led to the identification of molecular events, which confer upon these cells the unique properties of pluripotency and self renewal. The core regulatory network maintaining the ESC identity involves three master regulators: Oct4, Sox2, and Nanog. Large-scale mapping studies interrogating the binding sites of these and other transcription factors showed co-occupancy of distinct sets of transcription factors. The assembly of multitranscription factor complexes could serve as a mechanism for providing specificity in regulating ESC-specific gene expression. These studies are also beginning to unravel the transcriptional regulatory networks that govern the ESC identity. Loss-of-function RNAi screens also identified novel regulatory molecules involved in the stable propagation of the ESC state. This argues for an ESC transcriptional regulation program in which interconnected transcriptional regulatory networks involving large numbers of transcription factors and epigenetic modifiers work in concert on ESC- and differentiation-specific genes to achieve cell state stability. This chapter traces the major efforts made over the past decade in dissecting the transcriptional regulatory network governing ESC identity and offers perspectives on the future directions of the field.

Y.S. Chan and H.-H. Ng
Gene Regulation Laboratory, Genome Institute of Singapore, 60 Biopolis Street, #02-01, Genome Building, Singapore, Singapore 138672
NUS Graduate School for Integrative Sciences and Engineering, Singapore, Singapore 117597
e-mail: nghh@gis.a-star.edu.sg

L. Yang
Gene Regulation Laboratory, Genome Institute of Singapore, 60 Biopolis Street, #02-01, Genome Building, Singapore, Singapore 138672

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_12,

1 Transcriptional Regulation

1.1 The Transcriptome and Cell State

The multicellular organism is made up of a diverse array of cell types that play unique roles in ensuring the survival and propagation of the organism. Despite vast differences in morphology, intracellular biochemical processes and functions across cells of various types, the set of genetic information packaged within each nucleus is almost identical. The ability of cells to become highly specialized despite similarities at the genomic level is facilitated by the unique complement of mRNA molecules, known as the transcriptome, synthesized by individual cells. Each cell's transcriptome reflects the set of actively transcribed genes and is governed by a unique class of DNA binding proteins known as transcription factors.

Transcriptional regulation is a complex and dynamic process. While transcription factors can autoregulate their expression levels by binding to their own promoters, the interactions between transcription factors and genes are seldom restricted to one-to-one correspondences. Studies of transcriptional circuitry in model organisms, such as *Saccharomyces cerevisiae*, revealed that clusters of transcription factors work in unison to regulate gene expression [1]. It is common for transcription factors to closely associate (e.g., through the formation of dimers or larger complexes) with several different proteins (transcription factors or otherwise) in various combinations and thus coregulate the expression of target genes common to them. In addition, these regulatory molecules also form modules in which one or a set of factors control the transcription of a group of genes with similar binding patterns. Such interactions have been extensively documented and mapped into what is known as transcriptional regulatory networks [1]. These networks have been dissected to reveal recurring network motifs, such as feed-forward and multicomponent loops that are the basic building blocks of large transcriptional networks. The intricate web of protein–protein and protein–DNA interactions creates regulatory responses that extend beyond that of the simple linear relationships between transcription factors and the genes they regulate into nonlinear and temporal behaviors. This interplay between the components of the transcriptional regulatory circuitry is crucial for achieving stable cell states and maintaining cell identities.

1.2 Dissecting Transcriptional Regulation

The chromatin immunoprecipitation (ChIP) assay has been widely used in transcriptional regulatory network studies to characterize the *in vivo* binding activities of transcription factors (Fig. 1). This method uses formaldehyde-mediated cross-links to provide a snapshot of direct interactions between proteins and DNA in specific cell states. Subsequent nuclease treatment or sonication shears chromatin

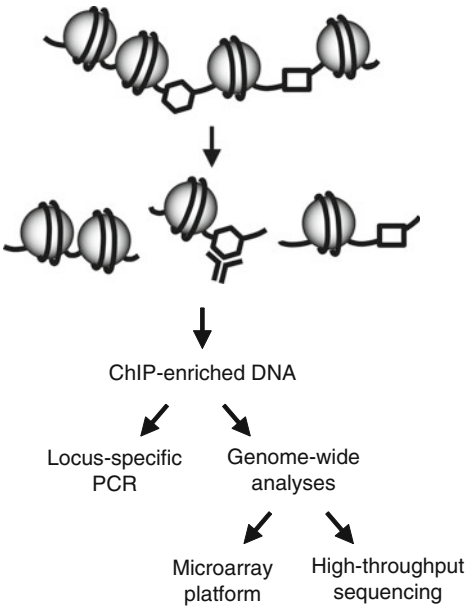


Fig. 1 The use of ChIP to map transcriptional regulatory network. Transcription factors (Hexagon drawn as *hexagons* and *rectangles*) are cross-linked to DNA with formaldehyde and subsequently sheared via mechanical or chemical means. Antibodies recognizing the specific transcription factors are used to pull down the protein–DNA complexes. Enriched DNA is quantified via conventional qPCR for locus-specific enrichments or with high-throughput detection platforms to generate global mapping profiles

into small fragments. The chromatin is probed with antibodies to enrich for DNA bound to the transcription factor of interest. Quantitative methods are employed to define the abundance of DNA fragments obtained; these vary between applications and range from the simple quantitative polymerase chain reaction (qPCR), which focuses on specific genomic loci, to high throughput hybridization arrays and sequencing platforms which map binding sites across the entire genome. Data from genome-wide ChIP assays unravel the transcription factor-DNA interactomes and provide new insights into the architecture of gene control networks.

1.3 From Model Organisms to Complex Vertebrates

Early studies of cellular processes in eukaryotes were undertaken with *S. cerevisiae* as a model organism. With a small genome and unicellular organization, yeast presents an ideal starting point for establishing the fundamental principles governing metabolism, proliferation, and differentiation in eukaryotes. Subsequent discoveries that certain metabolic pathways and transcription network motifs in yeast were conserved in higher vertebrates prompted further research into such regulatory processes in yeast [2]. However, in comparison to unicellular yeast, the developmental processes of multicellular eukaryotes are much more complex. Initial efforts to dissect gene regulatory networks controlling development were pursued in model organisms such as the sea urchin [3]. The complexity of the gene regulation

program in this organism served as a reminder of the even greater breadth and depth of information that remain to be gathered and analyzed for analogous studies in more complex vertebrates such as mammals. A full understanding of cellular processes in higher vertebrates required the complete decoding, dissecting, and mapping of biochemical interactions and transcriptional regulatory networks in diverse cell types – a formidable challenge, given currently available technologies. Importantly, projects of such scales call for large quantities of homogenous cells, which may not be readily available.

Embryonic stem cells (ESCs) provide a feasible alternative model system in this case. ESCs are isolated from the inner cell mass (ICM) of the early embryo and have the ability to self renew indefinitely *in vitro*. These cells retain the memory of an *in vivo* pluripotent state as they retain the ability to differentiate into all lineages of the adult organism. These unique characteristics make ESCs ideal systems for the study of transcriptional networks in vertebrates.

2 Unraveling the Transcriptional Regulatory Networks in ESCs

2.1 Master Regulators in ESC

The core regulatory network governing ESC pluripotency is based on the key transcription factors Oct4, Sox2, and Nanog [4–8]. These factors are preferentially expressed in the pluripotent cells of the ICM and epiblast, as well as in primordial germ cells. They play crucial roles in developmental processes by promoting and maintaining pluripotency in the early embryo. Oct4, encoded by the *Pou5f1* gene, is a homeodomain transcription factor belonging to the POU family. Oct4-null mouse embryos do not survive beyond implantation due to the inability to form pluripotent ICM [9]. Instead, the cells spontaneously differentiate into the trophoblast lineage. In ESCs, the level of Oct4 is critical for maintenance of the undifferentiated state. Depletion of Oct4 in ESCs leads to differentiation into trophoblast cells, while the overexpression of Oct4 induces preferential differentiation of ESCs into the primitive endoderm and mesoderm [10]. The consequences of perturbations in Oct4 expression levels on ESC fates highlight the critical role played by Oct4 in the ESC transcription program.

Similar to Oct4, the HMG-box transcription factor Sox2 is also essential for early embryonic development in mice [5]. Sox2 was originally implicated in pluripotency when it was discovered to interact synergistically with Oct4 to regulate the pluripotency-related gene *Fgf4* [11]. Subsequent genome-wide mapping studies revealed the presence of Oct-Sox binding motifs in enhancer regions of many ESC-specific genes, suggesting that Sox2 works cooperatively with Oct4 to preserve the ESC identity.

Nanog was identified in a screen for novel pluripotency regulators functioning independently of the LIF/STAT3 pathway [8]. The Nanog protein is critical for the formation of the epiblast *in vivo* [7, 12] but is dispensable in the maintenance of pluripotency in cultured ESCs [13]. It has thus been proposed that Nanog is the gateway to the pluripotent ground state – crucial for the establishment of pluripotency but not for its maintenance. Importantly, the overexpression of Nanog can sustain mouse ESCs in their pluripotent states without the need for leukemia inhibitory factor (LIF) [8] and allows human ESCs to bypass the need for both FGF and TGFβ signaling [14]. This suggests that Nanog establishes the pluripotent cell state via transcriptional regulatory pathways independently of external signals.

2.2 *Mapping the Core Transcriptional Regulatory Network in ESCs*

With the identification of master regulators of ESC pluripotency, mapping studies have been undertaken in both human and mouse ESCs to locate genome-wide binding sites for these factors. Coupling ChIP with whole genome promoter array (ChIP-on-chip) analysis, Boyer et al. identified high confidence binding sites for Oct4, Sox2, and Nanog in human ESCs [15]. Of the 17,917 human genes studied, 623, 1,271, and 1,687 genes were bound by Oct4, Sox2, and Nanog, respectively. More importantly, data from the ChIP-chip assay revealed that Oct4, Sox2, and Nanog co-occupied the promoters of at least 353 genes in human ESCs. In particular, Nanog bound to the regulatory sequences of more than 90% of the genes bound by the Oct4–Sox2 heterodimer. In comparing the ESC expression profiles with 158 published expression data representing 79 different tissues and cell types, the pool of genes regulated by some or all of the three factors could be grouped into two major classes: those preferentially expressed in ESCs and those expressed in specific lineages. These observations support a model for ESC transcriptional regulation in which the core regulators worked in a concerted and context-specific manner (i.e., either activating or repressing) to bring about stable propagation of the ESC state.

Besides regulating the transcription of downstream genes, Oct4, Sox2, and Nanog also bind to their own and each other's promoters to establish autoregulatory, feed-forward, and feedback loops. Such transcription network motifs set up a tripartite system in which the levels of Oct4, Sox2, and Nanog are monitored and adjusted by the three factors themselves to achieve homeostasis (Fig. 2). This adaptive control at the top of the transcription factor pyramid creates a sensitive, responsive, and self-stabilizing molecular circuitry, which ensures accurate regulation of subsequent downstream pluripotency genes.

Using ChIP-PET technology, Loh et al. characterized the genome-wide binding profiles of Oct4 and Nanog in mouse ESCs [16]. ChIP-PET involves the extraction of 5′ and 3′ ends of ChIP enriched fragments to form pair-end ditags (PETs), which

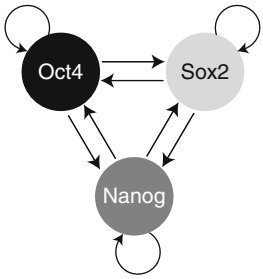


Fig. 2 Regulatory interactions between the transcription factor trio (Oct4, Sox2, and Nanog). The master regulators form intrinsic self regulatory, feed-forward, and feed-back loops that ensure high expression levels and provide self-stabilizing system of the factors for the maintenance of pluripotency in mouse and human ESCs

were then cloned into vectors for subsequent sequencing. This method identified 1,083 and 3,006 high confidence binding sites of Oct4 and Nanog, respectively. As in the case of human ESCs, self-regulatory and interregulatory loops were identified as mechanisms governing murine Oct4 and Nanog expression. In addition, the mouse study revealed divergences in the downstream target genes of Oct4 and Nanog regulatory networks in mouse and human systems. This was to be expected, since external signals which regulate human and mouse ESC pluripotency are fundamentally different. For instance, mouse ESCs depend on growth factors such as LIF and BMP while human ESCs require FGF and TGFβ for propagation. The difference between human and mouse ESCs may be due to different developmental stage from which cells were derived. In addition, technical differences between the two mapping platforms, such as the wider genome coverage offered by ChIP-PET, may also account for the discrepancy.

A number of genes commonly targeted by both Oct4 and Nanog are implicated in the regulation of pluripotency. One candidate gene was the orphan nuclear receptor Esrrb. Depletion of Esrrb in ESCs using RNA interference (RNAi) induced spontaneous ESC differentiation, indicating that Esrrb is crucial for the maintenance of the ESC state. More significantly, Esrrb could, together with Oct4 and Sox2, confer pluripotency to mouse embryonic fibroblasts (MEFs) via direct reprogramming [17]. These two lines of evidence indicate that Esrrb plays an important role to promote and maintain ESC identity. As such, dissecting the core transcriptional regulatory network not only unravels the machinery that drives ESC pluripotency but also identifies other factors which may wield key roles in driving the establishment of the pluripotent cell state.

2.3 Expanding the ESC Transcriptional Regulatory Networks

Data from genome-wide mappings of the master regulators described above have revealed that a significant percentage of downstream targets of Oct4, Sox2, and

Nanog encode for transcription factors. It is highly likely that there exist other factors within the transcriptional regulation program of ESCs which could either take on supporting roles to maintain master regulator expression levels at a stable equilibrium or directly target and control the expression of certain downstream gene clusters.

Kim et al. further expanded the pluripotency network of ESCs with genome-wide mapping studies of factors previously found to associate with Nanog [18]. Employing the ChIP-on-chip platform with promoter arrays, nine pluripotency-associated factors were studied [19]: Oct4, Sox2, c-Myc, Klf4, Nanog, Dax1, Nac1, Rex1, and Zpf281. In agreement with previous mapping studies of the master regulators, this study showed that 50% of the 6,632 target genes were bound by more than one transcription factor. The activities of target genes were shown to correlate with the number of transcription factors bound to the gene promoters. Genes whose promoters were bound by multiple factors tended to be preferentially expressed in ESCs, while those with single transcription factor binding were transcriptionally silent. While the binding profiles of Oct4, Sox2, Nanog, Klf4, Dax1, Nac1, and Zpf281 showed degrees of overlap, c-Myc and Rex1 binding sites formed a separate gene cluster. c-Myc and Rex1 target genes were implicated more frequently in protein metabolism than in developmental processes. Downstream genes bound by c-Myc were predominantly marked with the activating H3K4me3 histone modification, which accounted for their high expression levels. It was hypothesized that c-Myc and Rex1 maintain the expressions of housekeeping genes, which play key roles in sustaining the high proliferative capacities of ESCs. The extended pluripotency map emphasized a specialized and segregated network in which regulatory factors take on different responsibilities to support the ESC state.

2.4 Wiring Components of Signaling Pathways to the Core Transcriptional Regulatory Network

Chen et al. further investigated the transcriptional regulatory network for 13 transcription factors (Oct4, Sox2, Nanog, STAT3, Smad1, Zfx, c-Myc, n-Myc, Klf4, Esrrb, Tcfcp2l1, E2f1, and CTCF) through comprehensive unbiased mapping of transcription factor binding sites [20]. Using the Solexa/Illumina sequencing platform, the group was able to acquire mapping data with high specificity and sensitivity and quantify ChIP enrichment with greater sequencing depth. This represented an improvement over techniques such as ChIP-SAGE [21] or ChIP-PET [16], which require ChIP fragments to be modified before quantification. This study demonstrated that known components of the external signaling pathways such as Smad1 and STAT3, the downstream effectors of the BMP4 and LIF

pathways, respectively, are connected with and integrated into the core transcriptional regulatory network.

Another signaling pathway implicated in the maintenance of pluripotency in both human and mouse ESCs is the Wnt pathway [22]. Work by Cole et al. examined the binding profile of the Wnt downstream effector Tcf3 in mouse ESCs and found that Tcf3 co-occupied with Oct4 and Nanog at a significant number of ESC-specific and differentiation-associated genes [23]. This study proposed that the role of Tcf3 in mouse ESCs is context-specific: active Wnt signaling induces Tcf3 to partner with beta-catenin for activation of pluripotency genes, while the absence of beta-catenin promotes its association with groucho to repress the expression of differentiation genes in ESCs.

Such binding studies present the evidence for the integration of downstream targets of signaling pathways into the core transcription network of ESCs (Fig. 3). Further studies are needed to fully appreciate the implications of such interactions on the establishment and propagation of the ESC state.

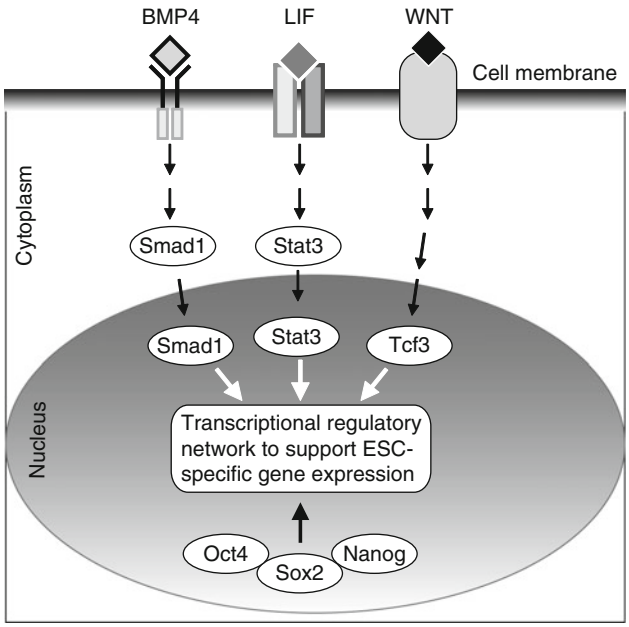


Fig. 3 Integration of the signaling pathways with the core ES cell transcriptional regulatory network. Signaling pathways shown to play a role in the maintenance of ESC include the BMP/Smad, LIF/ STAT, and Wnt/TCF pathways. Downstream effectors of the cell signaling cascades translocate to the nucleus and bind to the genome to mediate transcriptional regulation. Smad1, STAT3, and Tcf3 genome-wide binding site profiles reveal extensive overlap of these factors with the master regulators, suggesting that the communication between external signals and the core regulators is crucial in maintaining the ESC state. The maintenance of pluripotent state involves specific up-regulation of "stemness" genes and repression of lineage-specific genes

2.5 Enhanceosomes in ESCs

The study by Chen et al. also revealed that the transcription factors studied could be divided into two clusters [20]. Oct4, Sox2, Nanog, Smad1, and STAT3 were found to co-occupy genomic loci frequently, while n-Myc, c-Myc, Zfx, and E2f1 clustered together, away from the former group. An examination of the binding site locations of factors in the Oct4-centric cluster relative to regulated genes revealed that the majority of these sites fell within enhancers and not proximal promoter regions. Furthermore, a high degree of overlap was detected between the binding sites of the Oct4-centric cluster and the histone acetyltransferase p300, a chromatin regulator known to be widely recruited to gene enhancer sites to promote transactivation. Depletion of Oct4, Sox2, or Nanog using RNAi led to a significant reduction in p300 binding at these sites, indicating that p300 is recruited by the three transcription factors to gene enhancers. More importantly, the genomic loci from the Oct4-centric cluster show ESC-specific enhancer activity, indicating that these loci may be involved in the positive regulation of gene expression [20]. The discovery of short genomic segments densely packed with Oct4, Sox2, Nanog, and other transcription factor binding sites suggests that ESC transcriptional regulation involves ESC-specific enhanceosomes. These are short enhancer sequences in the genome that are bound either directly or indirectly at higher-than-normal frequencies by transcriptional regulators. Enhanceosomes could potentially contain gene regulatory codes essential for the maintenance of ESCs, and they provide the platforms for imparting specificity for ESC-specific gene expression.

2.6 Identification of Novel Nodes in the ESC Transcriptional Regulatory Networks

Ronin was identified as a novel key regulator of ESC identity in a screen for proteins downregulated by caspases during ESC differentiation. Ronin deficiency results in preimplantation lethality due to defects in the ICM. Conditional knockout of Ronin in ESCs also leads to growth arrest [24]. Ronin is neither LIF-dependent nor regulated by the three master regulators. A THAP domain-containing repressor, Ronin interacts with the transcription regulator HCF-1 and is thought to be involved in the epigenetic silencing of lineage-specific genes in ES cells. The genome-wide binding profile of Ronin remains to be elucidated. It would be of interest to determine whether the Ronin network functions in parallel or interactively with that of the other three master regulators. The discovery of Ronin indicates that more regulators of the ESC remain to be uncovered.

To obtain a more comprehensive model of the ESC transcriptional program, extensive efforts have been undertaken to expand the list of regulators governing ESC pluripotency and self renewal through RNAi screens. An initial screen conducted by Ivanova et al. on 65 transcription factors/DNA-binding proteins identified

three molecules that regulate ESC self-renewal: Tcl1, Tbx3, and Esrrb [25]. A comparison of the changes in global expression profiles of ESCs after shRNA-induced depletion of each transcription factor demonstrated that the responses of downstream genes to Oct4, Sox2, Nanog, Esrrb, Tbx3, and Tcl1 knockdowns were dichotomous. One gene cluster was sensitive to changes in Oct4, Sox2, and Nanog levels and indifferent to that of Esrrb, Tbx3, and Tcl1, while the other cluster responded only to fluctuations in Esrrb, Tbx3, and Tcl1 levels. This further supports the hypothesis that multiple distinct transcription networks exist in ESCs to maintain cell identity. It is thought that Oct4, Sox2, and Nanog simultaneously promote pluripotency and suppress differentiation, while the Tbx3, Tcl1, and Esrrb network contributes to the ESC state mainly via the suppression of differentiation into epiblast-derived lineages [26]. The loss-of-function approach not only identified crucial factors which could be manipulated to induce differentiation along specific lineages but also mapped out two interconnected global pathways that could potentially be responsible for the maintenance of the ESC state. In a similar RNAi screen, Gaspar-Maia et al. discovered a new chromatin regulator Chd1 responsible for maintaining the open chromatin state of ESCs [27]. Chd1 knockdown resulted in a loss of pluripotency and induced preferential differentiation of ESCs into the neural lineage. Chd1 has also been implicated in somatic reprogramming whereby significant decreases in reprogramming efficiency were observed following Chd1 depletion via RNAi. Apart from Chd1, other chromatin regulators have also been implicated in ESC pluripotency. Using endoribonuclease-prepared siRNAs (esiRNAs) targeting 1,008 transcripts encoding for chromatin proteins, Fazzio et al. identified several subunits of the Tip60-p400 chromatin remodeling complex to play a role in maintaining ESC identity [28].

The search for novel ESC regulators has, in recent years, expanded to genome-wide RNAi screens. Through the construction of Oct4 reporter lines, candidate genes of interest were evaluated for their ability to directly or indirectly modulate Oct4 activity and therefore promote the ESC identity. Using an extensive esiRNA library, Ding et al. identified an initial pool of 296 ESC regulators [29]. The top hits were further narrowed down to 16 potential regulators, 14 of which were known transcription factors or chromatin remodelers. Ctr9 and RTF1, components of the Paf1 complex, were shown to be important in the activation of promoters of key ESC regulators. In an independent study using a siRNA library, Hu et al. identified 148 genes, which may have roles in maintaining the ESC state [30]. Focusing on the top hits, the group illustrated that Cnot3 and Trim28 were essential transcriptional regulators of ESC self renewal. Cnot3 and Trim28 genome-wide binding sites overlapped significantly with that of the c-Myc and Zfx self-renewal modules but not with binding sites of Oct4 or Nanog. These genome-wide studies highlighted that the complexity of the ES transcription program could only be achieved through the concerted efforts of large numbers of regulatory molecules. In the context of ESC transcriptional regulatory network, key nodes specified by transcription factors and chromatin regulators are being identified.

3 From Pluripotent Stem Cell to Induced Pluripotent Stem Cell Networks

3.1 Transcription Factor-Mediated Reversion of Terminal Cell Fate

Before the demonstration of factor-mediated direct reprogramming, somatic cell fate had always been deemed irreversible. The reversal of cell fate was difficult to achieve and required drastic measures such as cell fusion or somatic nuclear transfer into enucleated oocytes [31]. In 2006, a groundbreaking work by Takahashi and Yamanaka demonstrated that somatic cell fate in mouse embryonic and adult fibroblasts could easily be reversed via the viral delivery of four transcription factors, commonly known as the Yamanaka factors: Oct4, Sox2, Klf4, and c-Myc [32]. This technique was rapidly adapted to the human system [33]. Rapid progress has been made on using different factors or chemicals to induce reprogramming of mouse and human somatic cells [34, 35]. Induced pluripotent stem cells (iPSCs) not only resembled ESCs in morphology, but they also adopted ESC-like epigenetic landscapes and were able to differentiate into all three germ lineages both *in vitro* and *in vivo*. Importantly, iPSCs were proven to be truly pluripotent: they were germline transmissible and could give rise to an entire organism via tetraploid complementation [36]. iPSC technology thus proves to be a promising source of patient-specific stem cells. At the same time, direct reprogramming also provides a new platform for dissecting the mechanism of cell fate decision.

3.2 Dissecting the Mechanism of Factor-Mediated Reprogramming

In a bid to understand how the Yamanaka factors induced pluripotency in somatic cells, Sridharan et al. examined the promoter occupancy of these factors at different stages of the reprogramming process [37]. Through the characterization and comparison of Oct4, Sox2, Klf4, and c-Myc binding profiles in ESCs, fibroblast-derived iPSCs and partially reprogrammed somatic cells (pre-iPSCs), molecular barriers (and therefore key requisites) in the establishment of the pluripotent cell state were identified. The gene expression profiles in pre-iPSCs were found to deviate significantly from that of ESCs. In particular, the occupancies of Oct4, Sox2, and cMyc at ESC specific genes in pre-iPSCs were much lower than that in iPSCs. Many ESC-specific genes which lacked Oct4, Sox2, and Klf4 binding in pre-iPSCs were identified to be target genes of Nanog. Since pre-iPSCs cannot activate endogenous *Nanog* expression, it is highly probable that the binding of the three factors to these target genes was Nanog-dependent. This observation further

supports the hypothesis that Nanog is required for the establishment of pluripotency [12].

This study also found that the virus-induced expression of reprogramming factors in MEFs was three- to fivefold higher than their levels in ESCs. This increases the probability of factors binding to non-ESC-specific genes with conserved binding motifs at their promoters. As such, genes detrimental to the maintenance of the ESC state may also be activated during the course of reprogramming. Both Klf4 and Sox2 are known to take on cell specification roles in certain cell types, and the inappropriate binding of these factors could create additional reprogramming barriers.

3.3 iPSC Transcriptome Highlights Potential Difference in Pluripotency Status

In a separate study, Chin et al. reported that the iPSCs and ESCs differed in their defining gene signatures [38]. The global expression profiles of early passage iPSCs showed greater deviations from that of ESCs as compared with late passage iPSCs, suggesting a gradual adaptation of the cells to selection pressure during *in vitro* culture. However, even with this adaptive process, the transcriptome of late passage iPSCs shows some distinctions from that in ESCs. Furthermore, marked deviations exist in the histone methylation patterns at the promoters of ESC-specific genes in iPSCs. Hence, while there is a clear preference for the use of iPSCs rather than nondonor-specific human ESCs in bids to avoid graft rejection and ethical complications, there is still much work to be done before iPSC technology can attain the level of safety and reliability needed to make its clinical application a reality.

4 Conclusions

ESCs provide a limitless source of pluripotent cells for regenerative therapy and offer opportunities for dissecting transcriptional regulation in vertebrates. Recent studies have made much headway in understanding the combinatorial nature of ESC transcriptional regulation through the identification and characterization of binding profiles of regulatory molecules. These insights into the modus operandi of ESC transcriptional regulation, coupled with the possibilities of iPSC technology, will benefit the development of safe and reliable clinical applications from ESCs for use in regenerative medicine.

References

1. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I, Zeitlinger J, Jennings EG, Murray HL, Gordon DB, Ren B, Wyrick JJ, Tagne JB, Volkert TL, Fraenkel E, Gifford DK, Young RA (2002) Transcriptional regulatory networks in Saccharomyces cerevisiae. Science 298:799–804

2. Kusari AB, Molina DM, Sabbagh W Jr, Lau CS, Bardwell L (2004) A conserved protein interaction network involving the yeast MAP kinases Fus3 and Kss1. J Cell Biol 164:267–277
3. Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yuh CH, Minokawa T, Amore G, Hinman V, Arenas-Mena C, Otim O, Brown CT, Livi CB, Lee PY, Revilla R, Rust AG, Pan Z, Schilstra MJ, Clarke PJ, Arnone MI, Rowen L, Cameron RA, McClay DR, Hood L, Bolouri H (2002) A genomic regulatory network for development. Science 295:1669–1678
4. Schöler HR, Ruppert S, Suzuki N, Chowdhury K, Gruss P (1990) New type of POU domain in germ line-specific protein Oct-4. Nature 344:435–439
5. Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R (2003) Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev 17:126–140
6. Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, Okochi H, Okuda A, Matoba R, Sharov A, Ko M, Niwa H (2007) Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol 9:625–635
7. Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 113:631–642
8. Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell 113:643–655
9. Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius KD, Chambers I, Schöler HR, Smith A (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell 95:379–391
10. Niwa H, Miyazaki J, Smith A (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet 24:372–376
11. Yuan H, Corbi N, Basilico C, Dailey L (1995) Developmental-specificity activity of the FGF-4 enhancer requires the synergistic action of Sox-2 and Oct-3. Genes Dev 9:2635–2645
12. Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, Barrandon O, Wray J, Yamanaka S, Chambers I, Smith A (2009) Nanog is the gateway to the pluripotent ground state. Cell 138:722–737
13. Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, Smith A (2007) Nanog safeguards pluripotency and mediates germline development. Nature 450:1230–1234
14. Xu RH, Sampsell-Barron TL, Gu F, Root S, Peck RM, Pan G, Yu J, Antosiewicz-Bourget J, Tian S, Stewart R, Thomson JA (2008) NANOG is a direct target of TGFb/activin-mediated SMAD signaling in human ESCs. Cell Stem Cell 3:196–206
15. Boyer LA, Lee TL, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122:947–956
16. Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J et al (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet 38:431–440
17. Feng B, Jiang JM, Kraus P, Ng JH, Dominic HJC, Chan YS, Yaw LP, Zhang WW, Loh YH, Han JY, Vega VB, Cacheux-Rataboul V, Lim B, Lufkin T, Ng HH (2009) Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. Nat Cell Biol 11:197–203
18. Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH (2006) A protein interaction network for pluripotency of embryonic stem cells. Nature 444:364–368
19. Kim J, Chu JL, Shen XH, Wang JL, Orkin SH (2008) An extended transcriptional network for pluripotency of embryonic stem cells. Cell 132:1049–1061
20. Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133:1106–1117

21. Impey S, McCorkle SR, Cha-Molstad H, Dwyer JM, Yochum GS, Boss JM, McWeeney S, Dunn JJ, Mandel G, Goodman RH (2004) Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. Cell 119:1041–1054
22. Sato N, Meijer L, Skaltsounis L, Greengard P, Brivanlou AH (2004) Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat Med 10:55–63
23. Cole MF, Johnstone SE, Newman JJ, Kagey MH, Young RA (2008) Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev 22:746–755
24. Dejosez M, Krumenacker JS, Zitur LJ, Passeri M, Chu LF, Songyang Z, Thomson JA, Zwaka TP (2008) Ronin is essential for embryogenesis and the pluripotency of mouse embryonic stem cells. Cell 133:1162–1174
25. Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, DeCoste C, Schafer X, Lun Y, Lemischka IR (2006) Dissecting self-renewal in stem cells with RNA interference. Nature 442:533–538
26. Rao S, Orkin SH (2006) Unraveling the transcriptional network controlling ES cell pluripotency. Genome Biol 7:230
27. Gaspar-Maia A, Alajem A, Polesso F, Sridharan R, Mason MJ, Heidersbach A, Ramalho-Santos J, McManus MT, Plath K, Meshorer E, Ramalho-Santos M (2009) Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature 460:863–868
28. Fazzio TG, Huff JT, Panning B (2008) An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. Cell 134:162–174
29. Ding L, Paszkowski-Rogacz M, Nitzsche A, Slabicki MM, Heninger AK, de Vries I, Kittler R, Junqueira M, Shevchenko A, Schulz H, Hubner N, Doss MX, Sachinidis A, Hescheler J, Iacone R, Anastassiadis K, Stewart AF, Pisabarro MT, Caldarelli A, Poser I, Theis M, Buchholz F (2009) A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. Cell Stem Cell 4:403–415
30. Hu G, Kim J, Xu QK, Leng Y, Orkin SH, Elledge SJ (2009) A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev 23:837–848
31. Hochedlinger K, Jaenisch R (2006) Nuclear reprogramming and pluripotency. Nature 441:1061–1067
32. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2006) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131:861–872
33. Yamanaka S (2009) A fresh look at iPS cells. Cell 137:13–17
34. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA (2007) Induced pluripotent stem cell lines derived from human somatic cells. Science 318:1917–1920
35. Feng B, Ng JH, Heng JC, Ng HH (2009) Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells. Cell Stem Cell 4:301–312
36. Zhao XY, Li W, Zhuo L, Liu L, Tong M, Hai T, Hao J, Guo C, Ma QW, Wang L, Zeng F, Zhou Q (2009) iPS cells produce viable mice through tetraploid complementation. Nature 461:86–90
37. Sridharan R, Tchieu J, Mason MJ, Yachechko R, Kuoy E, Horvath S, Zhou Q, Plath K (2009) Role of the murine reprogramming factors in the induction of pluripotency. Cell 136:364–377
38. Chin MH, Mason MJ, Xie W, Volinia S, Singer M, Peterson C, Ambartsumyan G, Aimiuwu O, Richter L, Zhang J, Khvorostov I, Ott V, Grunstein M, Lavon N, Benvenisty N, Croce CM, Clark AT, Baxter T, Pyle AD, Teitell MA, Pelegrini M, Plath K, Lowry WE (2009) Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell 5:111–123

Small Molecules in Cellular Reprogramming and Differentiation

Xu Yuan, Wenlin Li, and Sheng Ding

Abstract Recent advances in somatic cell reprogramming and directed differentiation make it possible to generate patient-specific pluripotent cells and further derive functional tissue-specific cells for biomedical research and future therapies. Chemical compounds targeting enzymes or signaling proteins are powerful tools to regulate and reveal complex cellular processes and have been identified and applied to controlling cell fate and function, including stem cell maintenance, differentiation, and reprogramming. Not only are small molecules useful in generating desired cell types *in vitro* for various applications, but also such small molecules could be further developed as conventional therapeutics to target patient's own cells residing in different tissues/organs for treating degenerative diseases, injuries, and cancer. Here, we will review recent studies of small molecules in controlling cell fate.

1 Introduction

Pluripotent stem cells can replicate indefinitely through symmetric cell divisions and are able to give rise to all the cell types of the three germ layers *in vitro* and *in vivo*. They can be typically derived from several different cell sources, including inner cell mass (ICM) of preimplantation blastocyst, late epiblast of postimplantation embryos, germline stem cells, and somatic cells through reprogramming. As stem cells hold significant potentials in biomedical research and regenerative medicine, they have been attracting increased interests in recent years. Chemical compounds targeting enzymes or signaling proteins are powerful tools to regulate and reveal complex cellular processes and have been identified and applied to controlling cell fate and function, including maintaining pluripotency of murine

X. Yuan, W. Li, and S. Ding
Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA
e-mail: sding@scripps.edu

S.M. Gasser and E. Li (eds.), *Epigenetics and Disease*, Progress in Drug Research 67,
DOI 10.1007/978-3-7643-8989-5_13,

embryonic stem (mES) cells in the absence of growth factors and cytokines [1, 2], promoting survival of dissociated human embryonic stem (hES) cells [3], facilitating directed differentiation of ES cells into different cell types, e.g., cardiomyocytes and neural cells, and enhancing the somatic cell reprogramming. In this review, we will discuss the recent progresses of using small molecules in cellular reprogramming and directed differentiation.

2 Small Molecules in Inducing Pluripotency

Cellular differentiation from less specialized cells to more functional cells with more restricted developmental potential is characterized molecularly by specific changes in epigenome. Once regarded irreversible, differentiation status (e.g., corresponding epigenetic changes of cells) can be reversed with cells reprogrammed to an earlier developmental state, such as totipotent or pluripotent state as demonstrated by somatic cell nuclear transfer (SCNT), or fusion of somatic cells with pluripotent cells; or changed to different lineages via genetic manipulation of master gene expression in various cell types. Recent breakthrough has established a more convenient and defined reprogramming condition to generate induced pluripotent stem (iPS) cells from mammalian somatic cells by ectopic expression of Oct4, Sox2, Klf4, and c-Myc (or Nanog and Lin28 instead of Klf4 and c-Myc) via viral infection [4–10].

Although the iPS cell technique may reduce certain ethical concerns and technical challenges of generating patient-specific pluripotent stem cells from adults, it has several critical limitations, e.g., genome modifications of target cells by exogenous oncogenes or other sequences, as well as more subtle genetic and epigenetic changes in the target cells generated and possibly preferentially selected during the nonspecific reprogramming process. To date, a number of approaches have been developed to address those challenges associated with the conventional iPS cell method, including using less transcription factors and taking the advantage of endogenous gene expression in specific cell types [7, 11–19], delivering transcription factors with various nonintegrating and removable systems [20–26], direct delivery of four recombinant proteins (Oct4, Sox2, Klf4, and c-Myc) [27, 28], and use of small molecules to facilitate the reprogramming process (Table 1).

2.1 Epigenetic Modifiers in Inducing Pluripotent Cells

In a simple view, as the reprogramming process in generating iPS cells fundamentally entails epigenetic changes, small molecules that directly modulate the epigenetic enzymes or mechanisms (e.g., DNA or histone modifications) may have an effect on facilitating reprogramming.

Table 1 Small molecules facilitating transcription factor-induced somatic cell reprogramming

Category	Small molecules	Working conditions and effects
HMTase G9a inhibitor	BIX-01294	Enhances reprogramming efficiency of K/O infected NPCs and MEFs [11, 12]; enables K/S/M induced reprogramming of NPCs [12]
LSD1 inhibitor	Parnate	Induces reprogramming of K/O infected human keratinocytes together with CHIR99021 [29]
DNMT inhibitor	5-aza	Increases the reprogramming efficiency of 4-factor-infected MEFs [30]
	RG108	In combination with BIX to improve K/O induced MEF reprogramming [11]
HDAC inhibitor	TSA, SAHA, VPA	Improve iPSC generation from 4 factor-infected MEFs [30]
	VPA	Enhances K/S/O induced reprogramming and induces iPSC from O/S infected human fibroblasts [15, 30]
GSK3 inhibitor	CHIR99021	Induces iPSC from K/O infected MEFs [29]; combined with PD0325901 to facilitate maturation of mouse iPSC [16]
	Kenpaullone	Facilitates MEF reprogramming in the presence of O/S/M [31]
TGFβ receptor inhibitor	A-83-01	Combined with MEK and Rock inhibitors to improve human fibroblast reprogramming [32]
	E-616451	Induces iPSC from K/M/O infected MEFs in the presence of VPA [33]
	E-616452 (RepSox)	Functionally substitutes ecotopic expression of Sox2 and c-Myc in MEF reprogramming [33]
MEK inhibitor	PD0325901	Promotes transitioning to mature iPSC [12, 16]; cooperates with TGFβ inhibition to generate human iPSC [32]
L-calcium channel agonist	BayK8644	Synergizes with BIX in K/O induced MEF reprogramming [11]
Src kinase inhibitor	EI-275	Cooperates with VPA in K/M/O induced MEF reprogramming [33]

iPSC Induced pluripotent stem cell; *K* Klf4; *O* Oct4; *M* c-Myc; *S* Sox2

As a proof of concept, a small molecule inhibitor of histone methyltransferase (HMTase) G9a, BIX-01294 (BIX) [34], was first identified to significantly enhance the reprogramming efficiency of mouse neural progenitor cells (NPCs) and fibroblasts that were transduced with only two reprogramming transcription factors (Oct4 and Klf4) [11, 12]. The studies also demonstrated that BIX can functionally replace Sox2 and c-Myc in generating iPS cells from somatic cells. More interestingly and significantly, the BIX treatment combined with the viral transduction of

Klf4/Sox2/c-Myc could enable generation of iPS cells from NPCs in the absence of Oct4, the only gene (of the four reprogramming genes) not expressed in any somatic cells [12]. Methylation of H3K9 by G9a is implicated in silencing of target genes, including Oct4 [35]. It is conceivable that G9a inhibition by BIX may facilitate relieving the repression of endogenous Oct4 during reprogramming and therefore bypassed the need of exogenous Oct4. More recently, a small molecule inhibitor of LSD1 (a H3K4 demethylase), parnate, was shown to enhance reprogramming of human somatic cells that were transduced with only Oct4 and Klf4 [29]. Those studies highlight the important roles of dynamic histone methylation and demethylation processes in epigenetic reprogramming and the utility of small molecule modifiers of histone methylation.

DNA methylation is another common epigenetic mechanism involved in gene silencing. As pluripotent stem cells have a distinct DNA methylation signature in comparison to somatic cells, remodeling of DNA methylation during reprogramming is likely to be an important epigenetic barrier. In addition, pluripotent cells are globally less methylated at the DNA level. Therefore, inhibition of DNA methyltransferases (DNMTs), which establish and maintain DNA methylation, would have a positive impact on reprogramming. Consistently, 5-azacytidine (5-aza), a DNMT inhibitor, could promote the overall efficiency in the 4-factor-induced reprogramming of MEFs [30] and induce a rapid and stable transition of certain intermediate/partially reprogrammed cells to a fully reprogrammed state [36]. However, 5-aza may have undesirable side-effects, as it incorporates into DNA to covalently trap the enzyme, as well as it can demethylate centromeric satellite elements, which could result in chromosomal instability. In our studies, RG108, a more potent and noncovalent DNMT inhibitor [37], was shown to enhance the reprogramming efficiency of MEFs transduced with Oct4/Klf4 and treated with BIX [11]. Compared with 5-aza, RG108 may have a more specific demethylation effect as it does not appear to demethylate centromeric satellite elements.

Similarly, small molecules that modulate histone acetylation, another epigenetic mechanism in regulating gene expression, would be predicted to have an impact on reprogramming process. In previous SCNT studies, inhibition of histone deacetylases (HDACs) by Trichostatin A (TSA) was shown to have a modest effect on the reprogramming efficiency [38, 39]. More recently, various HDAC inhibitors, including TSA, suberoylanilide hydroxamic acid (SAHA), and valproic acid (VPA), were shown to improve the transcription factor-induced reprogramming [15, 30]. It is suggested that the global histone hyperacetylation resulted from HDAC inhibition may create a more relaxed chromatin state for binding of exogenous/endogenous transcription factors and therefore benefit the reprogramming. Particularly, VPA significantly improved induction of mouse and human iPS cells in the absence of c-Myc and further enabled human fibroblasts to be reprogrammed with only two factors (Oct4 and Sox2) [15, 30]. Zhou et al. reported that VPA also facilitated the generation of recombinant protein-induced pluripotent stem (piPS) cells from MEFs without the use of any genetic material and genetic manipulation [28]. As VPA is a nonspecific HDAC inhibitor and may have other effects, it would be useful to genetically dissect out which HDAC(s) inhibition has

major contribution to reprogramming, therefore allowing design and use of more specific HDAC inhibitors.

Although small molecule epigenetic modifiers have been shown to improve the reprogramming efficiency, the precise molecular mechanisms underlying their effects remain elusive. They are presumed to facilitate shifting epigenetic balances among different states during the nonspecific reprogramming process that involves random events. Because those small molecules relatively nonspecifically affect the epigenome of target cells, they have limitations in orchestrating a more directed reprogramming process. Small molecules modulating specific signaling pathways to facilitate/direct reprogramming may complement those epigenetic modifiers. Furthermore, identifying such molecules may allow better elucidating the molecular mechanisms of reprogramming.

2.2 Signaling Modulators in Reprogramming

Several signaling pathways, including Wnt-β-catenin, MEK-ERK, Calcium-cAMP, TGFβ, Rho-Rock pathways, have been identified to impact on the reprogramming process, as demonstrated by the effects of pathway modulating small molecules in generation of iPS cells [12, 16, 29, 32, 33, 40, 41]. Activation of Wnt signaling pathway by Wnt3a protein can promote iPS cell generation from 3-factor transduced MEFs in the absence of c-Myc [40]. In addition, a glycogen synthase kinase-3β (GSK-3β) inhibitor CHIR99021, which strongly activates the β-catenin pathway, was shown to be particularly useful in enhancing generation of mouse and human iPS cells in the absence of exogenous Sox2 and c-Myc [29]. Via a high throughput screen, another GSK-3 inhibitor, kenpaullone, which also inhibits cyclin-dependent kinase (CDK) and other kinases, was found able to reprogram MEFs transduced with Oct4/Sox2/c-Myc but in the absence of Klf4 [31]. Interestingly, neither a more specific GSK3 inhibitor (such as CHIR99021) nor a CDK inhibitor was able to replace Klf4 in the same context, indicating the involvement of other mechanisms independent of GSK-3 or CDK inhibition by kenpaullone [31].

MEK-ERK pathway inhibition by small molecules, such as pluripotin (inhibiting ERK and RasGap) and PD0325901 (inhibiting MEK), has been shown to promote self-renewal of mES cells [1, 2]. Interestingly, MEK inhibition by reversine was shown to be critical in reprogramming muscle lineage restricted cells back to more primitive, mesenchymal precursor-like cells [42]. It appears that the mechanism of MEK inhibition in muscle lineage reprogramming is also shared in iPS cell generation. Recently, we have shown that PD0325901 can enhance reprogramming of mouse NPCs to iPS cells by transitioning early iPS cells to fully reprogrammed mature iPS cells and promoting/stabilizing their growth [12], as well as has a synergistic effect with TGFβ pathway inhibition in promoting generation of human iPS cells [32]. Furthermore, it was shown that combination of PD0325901 and CHIR99021 can facilitate transitioning certain late-stage incompletely reprogrammed cells to the pluripotent state with stable endogenous Oct4 expression [16].

A role of TGFβ signaling in reprogramming was first revealed by our study of generating chimerism-competent rat iPS cells. The inhibition of the TGFβ pathway using the small molecule A-83-01 was critical for the ultimate reprogramming and maintenance of stable rat iPS cell clones in conjunction with the MEK and GSK-3 inhibitors [43]. More recently, studies have further demonstrated the importance of TGFβ signaling inhibition in generating mouse and human iPS cells [32, 33, 41]. Inhibiting TGFβ pathway cooperated with the four transcription factors in reprogramming of mouse fibroblasts [41] and functionally substituted ecotopic expression of Sox2 or c-Myc [33, 41]. Further studies revealed that TGFβ pathway inhibition ultimately facilitated induction of endogenous pluripotency gene expression, including Nanog [33], and drove complete conversion of partially reprogrammed cells to iPS cells [33, 41]. More importantly, we have shown that TGFβ pathway inhibition by small molecules, alone or in conjunction with MEK and Rock inhibitors, can dramatically increase reprogramming efficiency and accelerate reprogramming kinetics of human fibroblasts [32]. The process of reprogramming from fibroblasts to iPS cells represents an ultimate mesenchymal to epithelial transition (MET), during which the mesenchymal type fibroblasts undergo dramatic morphological changes that result in iPS cells with distinct cell polarity, boundaries, cell–cell interactions, and high E-cadherin expression. TGFβ is a prototypical cytokine for induction of epithelial to mesenchymal transition (EMT) and maintenance of the mesenchymal state. Conversely inhibition of TGFβ signaling can result in derepression of epithelial fate and would benefit the reprogramming process. The demonstration that TGFβ and MAPK pathway inhibition improves efficiency and kinetics of fibroblast reprogramming highlights critical roles of MET mechanisms and its players in the process.

Small molecules modulating signaling pathways can also synergize with direct epigenetic modifiers in reprogramming. For example, BayK8644, an L-calcium channel agonist [44], significantly increased the number and size of iPS cell colonies derived from Oct4/Klf4 transduced MEFs in the presence of G9a inhibitor [11]. Src kinase inhibitor EI-275 and TGFβ receptor inhibitor E-616451 were reported to cooperate with VPA in Oct4/Klf4/cMyc induced reprogramming of mouse fibroblasts [33]. It is conceivable that a precise combinatorial action of signaling and epigenetic modifiers may direct a more specific and efficient reprogramming process *in vitro* or *in vivo*.

3 Small Molecules in Directed Differentiation

To model/study development and diseases or to generate functional cell types for replacing lost and damaged cells to treat degenerative diseases and tissue injuries, a corresponding specific and efficient differentiation process of stem cells is essential. Conventional stem cell differentiation (e.g., spontaneous differentiation of ES cells via embryoid body/EB formation) is typically nonselective and inefficient and

would require selection of cell types of interest from a heterogeneous cell population. Recent advances in understanding of development and high throughput small molecule screens in stem cell differentiation have allowed construction and further improvement of more specific and robust differentiation process.

Various small molecules have been identified and characterized to direct or enhance stem cell differentiation. For example, TWS119 (a GSK3 inhibitor), Hedgehog pathway agonist Hh-Ag1.3, neuropathiazol, phosphoserine (P-Ser), HDAC inhibitor VPA, and a Ca^{2+}-triggering small molecule isoxazole were identified to promote neuronal differentiation of ES cells or neural precursor cells [45–50]. In screening for cardiogenic molecules, ascorbic acid, cardiogenol, and an isoxazolyl-serine-based PPAR agonist were found to promote cardiac marker gene expression and enhance differentiation of ES cells to cardiac myocytes [51–53]. These small molecules were discussed in previous reviews [54–56]. It is worth to mention that some small molecules, such as VPA, GSK3 inhibitor, and TGFβ receptor inhibitor, are functioning in both reprogramming and differentiation. Depending on their working contexts (the presence of growth factors and transcription factors), these small molecules facilitate cell fate determinations. In the following, we will focus on the recent discovery of small molecules in cardiac, pancreatic, and neural lineage specifications of ES cells.

3.1 Small Molecules for Cardiac Induction

Generation of functional cardiomyocytes from ES cells typically involves sequential inductions of mesendoderm, cardiogenic mesoderm, cardiovascular precursor cells, cardiomyocyte differentiation, and maturation. While the TGFβ family member Nodal was shown to efficiently induce mesoderm differentiation from ES cells, Wnt/β-catenin signaling (e.g., Wnt3a or GSK3 inhibitors) played biphasic roles in cardiogenesis depending on the specific developmental stage [57, 58]: canonical Wnt signaling induced mesoderm specification of ES cells at the early stage, but inhibited cardiomyocyte differentiation from cardiovascular precursor cells at later stage. Consistently, activation of Wnt using GSK3 inhibitor allowed expansion of multipotent cardiovascular precursor cells derived from ES cells or isolated from adult hearts, representing an alternative strategy for cardiac regenerative medicine [59]. Another key regulator of cardiogenic differentiation is Notch signaling: inactivation of Notch favors cardiomyocyte differentiation of ES cells [60, 61].

Nkx2.5 is an early cardiac master gene expressed in cardiovascular precursor cells. In a screen for chemical activators of Nkx2.5, Sadek et al. identified a class of small molecules called sulfonyl-hydrazones (Shz), which induced Nkx2.5 expression and additional cardiac markers including myocardin and sarcomeric α-tropomyosin (SαTM) in P19CL6 cells and mES cells [62]. Interestingly, human adult progenitor cells (M-PBMCs) treated with Shz could express cardiac genes and showed

improved cardiac function after engraftment in a rat myocardial cryoinjury model. Shz functioned independent of BMP, FGF, and Wnt pathways, suggesting a potentially novel mechanism to regulate cardiac differentiation in stem cells [62].

In addition, some other signaling modulating small molecules, including p38 MAPK inhibitor (SB203580), an L-type Ca^{2+} channel blocker (Verapamil), and Cyclosporin, were identified to promote cardiac differentiation of murine or human ES cells [63, 64]. Further mechanistic characterizations of those compounds may provide better understanding and devising of *in vitro* and *in vivo* cardiogenesis.

3.2 Small Molecules for Endoderm Induction

Type I diabetes is caused by the autoimmune destruction of insulin-producing pancreatic β cells. Deriving insulin-producing cells or developing renewable pancreatic progenitors from ES cells represents an alternative strategy for cell-replacement therapy for type I diabetes. Application of developmental principles to stem cell biology has allowed construction of a directed stepwise differentiation process from ES cells to insulin-producing cells through sequential inductions of mesendoderm and definitive endoderm mainly by Activin A and Wnt treatment of ES cells, followed by a series of cytokine and small molecule treatments, including FGF10, Hedgehog antagonists (i.e., cyclopamine), retinoic acids, Notch inhibitors (i.e., DAPT), nicotinamide, directing cells through stages resembling posterior foregut, pancreatic endoderm, and endocrine precursors, and ultimately to endocrine cells that secrete hormones [65, 66]. While such a protocol is labor intensive and still inefficient, it provides a basis for identifying additional modulators (especially small molecule enhancers) for each step to improve pancreatic induction.

In a screen for small molecules that can induce Sox17 expression (an endoderm-specific marker) from mES cells using a Sox17-GFP reporter mES cell line in the presence of low serum but absence of Activin A, Borowiak et al. identified IDE1 and IDE2 from 4,000 compounds and further confirmed their effects on inducing definitive endoderm from both hES and mES cells in monolayer cultures [67]. While the molecular targets and precise mechanisms of action of IDE1 and IDE2 remain unknown, they were shown to at least partially function through activating TGFβ signaling pathway and inducing Smad2 phosphorylation in mES cells, mimicking the function of Activin A. The IDE-induced definitive endoderm cells appeared to have developmental potentials similar to their *in vivo* counterparts and could be further induced to pancreatic precursor cells by another small molecule (-)-indolactam V (ILV) [68]. ILV was identified in a separate screen for small molecules that could induce Pdx1-expressing cells from hES cells-derived definitive endoderm [68]. Pdx1 is a master regulator of pancreatic development and begins to express since the pancreatic precursor stage. It was shown that the ILV-induced Pdx1 positive cells were able to further differentiate into multiple pancreatic lineages including insulin-producing β cells. Further mechanistic studies suggested that ILV functions at least partially through activation of PKC signaling

to induce gut tube endoderm into pancreatic precursors [68]. Consistently, another two PKC agonists could mimic the effect of ILV on pancreatic induction, and inhibiting PKC signaling with antagonists blocked ILV's effects and decreased the percentage of Pdx1 expressing cells.

In another small molecule screen for definitive endoderm differentiation of ES cells in the presence of low concentration of Activin A, Zhu et al. identified a staurosporine analog, named stauprimide, which can synergize with Activin A and potentiate definitive endoderm differentiation from murine and human ES cells [69]. With stauprimide treatment, more efficient endodermal differentiation could be achieved using less Activin A and in the absence of serum. Interestingly, it was shown that Activin A was necessary for the differentiation and stauprimide alone was not able to promote the specific endodermal differentiation. Further analysis revealed that stauprimide functioned to prime ES cells in conjunction with specific differentiation cues (such as Activin A or BMP) for the corresponding lineage specification. A putative target of stauprimide was identified through affinity pull-down as NME2, which mediates c-Myc activity. The inhibition of NME2 by stauprimide resulted in a rapid down-regulation of c-Myc and destabilization of the pluripotency of ES cells and consequently promoted their differentiation.

3.3 Small Molecules for Neural Induction

Chemically defined media (e.g., N2 and B27 supplements) and specific small molecule modulators of developmental pathways (e.g., Hedgehog pathway agonists or antagonists, RA) have been incredibly useful in neural specification and patterning of ES cells [50, 70–72]. A particular challenge for neural induction of hES cells is the heterogeneity and slow kinetics of the differentiation even under monolayer and chemically defined medium conditions. To address this, a method of applying the combination of BMP and TGFβ pathway inhibition using Noggin (a secreted protein that binds to and inhibits BMP4) and SB431542 (TGFβ receptor inhibitor) was developed [73]. This is based on the previous findings that Noggin treatment could enhance the differentiation from ES cells to neural precursor cells, and SB431542 treatment increased neural induction in EB-mediated differentiation of hES cells [74–76]. Chamber et al. found that the dual inhibition of TGFβ and BMP receptors led to a rapid and more complete neural conversion [73]. In their study, TGFβ pathway inhibition resulted in rapid loss of pluripotency gene Nanog expression in hES cells associated with an up-regulation of CDX2, a marker for trophoblast differentiation. Meanwhile, Noggin played key roles in suppressing trophoblast and primitive endoderm differentiation mainly driven by the BMP signaling [73]. The derived transient neural progeny could be further patterned into motor neurons and dopamine neurons after exposing to corresponding differentiation cues [73].

4 Perspectives

Recent advances in somatic cell reprogramming and directed differentiation make it possible to generate patient-specific pluripotent cells and further derive functional tissue-specific cells for biomedical research and future therapies. As reviewed above, small molecules have played increasingly important roles in those processes, including stem cell maintenance, differentiation, and reprogramming. Not only are small molecules useful in generating desired cell types *in vitro* for various applications, but also such small molecules could be further developed as conventional therapeutics to target patients' own cells residing in different tissues/organs for treating degenerative diseases, injuries, and cancer. Clearly, identification, characterization, and further development of additional small molecules for various targets/mechanisms controlling cell fate and function with improved activity, specificity, and desirable *in vivo* tissue distribution/availability represent a fertile area for stem cell research and regenerative medicine.

References

1. Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A (2008) The ground state of embryonic stem cell self-renewal. Nature 453:519–523
2. Chen S, Do JT, Zhang Q, Yao S, Yan F, Peters EC, Scholer HR, Schultz PG, Ding S (2006) Self-renewal of embryonic stem cells by a small molecule. Proc Natl Acad Sci USA 103:17266–17271
3. Watanabe K, Ueno M, Kamiya D, Nishiyama A, Matsumura M, Wataya T, Takahashi JB, Nishikawa S, Nishikawa S, Muguruma K et al (2007) A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nat Biotechnol 25:681–686
4. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R et al (2007) Induced pluripotent stem cell lines derived from human somatic cells. Science 318:1917–1920
5. Lowry WE, Richter L, Yachechko R, Pyle AD, Tchieu J, Sridharan R, Clark AT, Plath K (2008) Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc Natl Acad Sci USA 105:2883–2888
6. Aoi T, Yae K, Nakagawa M, Ichisaka T, Okita K, Takahashi K, Chiba T, Yamanaka S (2008) Generation of pluripotent stem cells from adult mouse liver and stomach cells. Science 321:699–702
7. Park IH, Zhao R, West JA, Yabuuchi A, Huo H, Ince TA, Lerou PH, Lensch MW, Daley GQ (2008) Reprogramming of human somatic cells to pluripotency with defined factors. Nature 451:141–146
8. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126:663–676
9. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131:861–872
10. Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K, Bernstein BE, Jaenisch R (2007) *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature 448:318–324

11. Shi Y, Desponts C, Do JT, Hahm HS, Scholer HR, Ding S (2008) Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell Stem Cell 3:568–574
12. Shi Y, Do JT, Desponts C, Hahm HS, Scholer HR, Ding S (2008) A combined chemical and genetic approach for the generation of induced pluripotent stem cells. Cell Stem Cell 2:525–528
13. Kim JB, Sebastiano V, Wu G, Arauzo-Bravo MJ, Sasse P, Gentile L, Ko K, Ruau D, Ehrich M, van den Boom D et al (2009) Oct4-induced pluripotency in adult neural stem cells. Cell 136:411–419
14. Kim JB, Zaehres H, Wu G, Gentile L, Ko K, Sebastiano V, Arauzo-Bravo MJ, Ruau D, Han DW, Zenke M et al (2008) Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. Nature 454:646–650
15. Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, Chen S, Muhlestein W, Melton DA (2008) Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol 26:1269–1275
16. Silva J, Barrandon O, Nichols J, Kawaguchi J, Theunissen TW, Smith A (2008) Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol 6:e253
17. Wernig M, Meissner A, Cassady JP, Jaenisch R (2008) c-Myc is dispensable for direct reprogramming of mouse fibroblasts. Cell Stem Cell 2:10–12
18. Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochiduki Y, Takizawa N, Yamanaka S (2008) Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol 26:101–106
19. Kim JB, Greber B, Arauzo-Bravo MJ, Meyer J, Park KI, Zaehres H, Scholer HR (2009) Direct reprogramming of human neural stem cells by OCT4. Nature 461:649–643
20. Abujarour R, Ding S (2009) Induced pluripotent stem cells free of exogenous reprogramming factors. Genome Biol 10:220
21. Stadtfeld M, Nagaya M, Utikal J, Weir G, Hochedlinger K (2008) Induced pluripotent stem cells generated without viral integration. Science 322:945–949
22. Okita K, Nakagawa M, Hyenjong H, Ichisaka T, Yamanaka S (2008) Generation of mouse induced pluripotent stem cells without viral vectors. Science 322:949–953
23. Yu J, Hu K, Smuga-Otto K, Tian S, Stewart R, Slukvin II, Thomson JA (2009) Human induced pluripotent stem cells free of vector and transgene sequences. Science 324:797–801
24. Kaji K, Norrby K, Paca A, Mileikovsky M, Mohseni P, Woltjen K (2009) Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature 458:771–775
25. Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, Hamalainen R, Cowling R, Wang W, Liu P, Gertsenstein M et al (2009) piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature 458:766–770
26. Gonzalez F, Barragan Monasterio M, Tiscornia G, Montserrat Pulido N, Vassena R, Batlle Morera L, Rodriguez Piza I, Izpisua Belmonte JC (2009) Generation of mouse-induced pluripotent stem cells by transient expression of a single nonviral polycistronic vector. Proc Natl Acad Sci USA 106:8918–8922
27. Kim D, Kim CH, Moon JI, Chung YG, Chang MY, Han BS, Ko S, Yang E, Cha KY, Lanza R et al (2009) Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. Cell Stem Cell 4:472–476
28. Zhou H, Wu S, Joo JY, Zhu S, Han DW, Lin T, Trauger S, Bien G, Yao S, Zhu Y et al (2009) Generation of induced pluripotent stem cells using recombinant proteins. Cell Stem Cell 4:381–384
29. Li W, Zhou H, Abujarour R, Zhu S, Joo JY, Lin T, Hao E, Scholer HR, Hayek A, Ding S (2009) Generation of human induced pluripotent stem cells in the absence of exogenous Sox2. Stem Cells 27:2992–3000
30. Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, Chen AE, Melton DA (2008) Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol 26:795–797

31. Lyssiotis CA, Foreman RK, Staerk J, Garcia M, Mathur D, Markoulaki S, Hanna J, Lairson LL, Charette BD, Bouchez LC et al (2009) Reprogramming of murine fibroblasts to induced pluripotent stem cells with chemical complementation of Klf4. Proc Natl Acad Sci USA 106:8912–8917
32. Lin T, Ambasudhan R, Yuan X, Li W, Hilcove S, Abujarour R, Lin X, Hahm HS, Hao E, Hayek A et al (2009) A chemical platform for improved induction of human iPSCs. Nat Methods 6:805–808
33. Ichida JK, Blanchard J, Lam K, Son EY, Chung JE, Egli D, Loh KM, Carter AC, Di Giorgio FP, Koszka K et al (2009) A small-molecule inhibitor of Tgf-beta signaling replaces Sox2 in reprogramming by inducing nanog. Cell Stem Cell 5:491–503
34. Kubicek S, O'Sullivan RJ, August EM, Hickey ER, Zhang Q, Teodoro ML, Rea S, Mechtler K, Kowalski JA, Homon CA et al (2007) Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. Mol Cell 25:473–481
35. Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y, Cedar H, Bergman Y (2006) G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat Cell Biol 8:188–194
36. Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, Bernstein BE, Jaenisch R, Lander ES, Meissner A (2008) Dissecting direct reprogramming through integrative genomic analysis. Nature 454:49–55
37. Brueckner B, Garcia Boy R, Siedlecki P, Musch T, Kliem HC, Zielenkiewicz P, Suhai S, Wiessler M, Lyko F (2005) Epigenetic reactivation of tumor suppressor genes by a novel small-molecule inhibitor of human DNA methyltransferases. Cancer Res 65:6305–6311
38. Kishigami S, Mizutani E, Ohta H, Hikichi T, Thuan NV, Wakayama S, Bui HT, Wakayama T (2006) Significant improvement of mouse cloning technique by treatment with trichostatin A after somatic nuclear transfer. Biochem Biophys Res Commun 340:183–189
39. Rybouchkin A, Kato Y, Tsunoda Y (2006) Role of histone acetylation in reprogramming of somatic nuclei following nuclear transfer. Biol Reprod 74:1083–1089
40. Marson A, Foreman R, Chevalier B, Bilodeau S, Kahn M, Young RA, Jaenisch R (2008) Wnt signaling promotes reprogramming of somatic cells to pluripotency. Cell Stem Cell 3:132–135
41. Maherali N, Hochedlinger K (2009) Tgfbeta signal inhibition cooperates in the induction of iPSCs and replaces Sox2 and cMyc. Curr Biol 19:1718–1723
42. Chen S, Zhang Q, Wu X, Schultz PG, Ding S (2004) Dedifferentiation of lineage-committed cells by a small molecule. J Am Chem Soc 126:410–411
43. Li W, Wei W, Zhu S, Zhu J, Shi Y, Lin T, Hao E, Hayek A, Deng H, Ding S (2009) Generation of rat and human induced pluripotent stem cells by combining genetic reprogramming and chemical inhibitors. Cell Stem Cell 4:16–19
44. Schramm M, Thomas G, Towart R, Franckowiak G (1983) Novel dihydropyridines with positive inotropic action through activation of Ca2+ channels. Nature 303:535–537
45. Warashina M, Min KH, Kuwabara T, Huynh A, Gage FH, Schultz PG, Ding S (2006) A synthetic small molecule that induces neuronal differentiation of adult hippocampal neural progenitor cells. Angew Chem Int Ed Engl 45:591–593
46. Saxe JP, Wu H, Kelly TK, Phelps ME, Sun YE, Kornblum HI, Huang J (2007) A phenotypic small-molecule screen identifies an orphan ligand-receptor pair that regulates neural stem cell differentiation. Chem Biol 14:1019–1030
47. Ding S, Wu TY, Brinker A, Peters EC, Hur W, Gray NS, Schultz PG (2003) Synthetic small molecules that control stem cell fate. Proc Natl Acad Sci USA 100:7632–7637
48. Schneider JW, Gao Z, Li S, Farooqi M, Tang TS, Bezprozvanny I, Frantz DE, Hsieh J (2008) Small-molecule activation of neuronal cell fate. Nat Chem Biol 4:408–410
49. Yu IT, Park JY, Kim SH, Lee JS, Kim YS, Son H (2009) Valproic acid promotes neuronal differentiation by induction of proneural factors in association with H4 acetylation. Neuropharmacology 56:473–480

50. Wichterle H, Lieberam I, Porter JA, Jessell TM (2002) Directed differentiation of embryonic stem cells into motor neurons. Cell 110:385–397
51. Takahashi T, Lord B, Schulze PC, Fryer RM, Sarang SS, Gullans SR, Lee RT (2003) Ascorbic acid enhances differentiation of embryonic stem cells into cardiac myocytes. Circulation 107:1912–1916
52. Wu X, Ding S, Ding Q, Gray NS, Schultz PG (2004) Small molecules that induce cardiomyogenesis in embryonic stem cells. J Am Chem Soc 126:1590–1591
53. Wei ZL, Petukhov PA, Bizik F, Teixeira JC, Mercola M, Volpe EA, Glazer RI, Willson TM, Kozikowski AP (2004) Isoxazolyl-serine-based agonists of peroxisome proliferator-activated receptor: design, synthesis, and effects on cardiomyocyte differentiation. J Am Chem Soc 126:16714–16715
54. Willems E, Bushway PJ, Mercola M (2009) Natural and synthetic regulators of embryonic stem cell cardiogenesis. Pediatr Cardiol 30:635–642
55. Xu Y, Shi Y, Ding S (2008) A chemical approach to stem-cell biology and regenerative medicine. Nature 453:338–344
56. Chen S, Hilcove S, Ding S (2006) Exploring stem cell biology with small molecules. Mol Biosyst 2:18–24
57. Ueno S, Weidinger G, Osugi T, Kohn AD, Golob JL, Pabon L, Reinecke H, Moon RT, Murry CE (2007) Biphasic role for Wnt/beta-catenin signaling in cardiac specification in zebrafish and embryonic stem cells. Proc Natl Acad Sci USA 104:9685–9690
58. Naito AT, Shiojima I, Akazawa H, Hidaka K, Morisaki T, Kikuchi A, Komuro I (2006) Developmental stage-specific biphasic roles of Wnt/beta-catenin signaling in cardiomyogenesis and hematopoiesis. Proc Natl Acad Sci USA 103:19812–19817
59. Qyang Y, Martin-Puig S, Chiravuri M, Chen S, Xu H, Bu L, Jiang X, Lin L, Granger A, Moretti A et al (2007) The renewal and differentiation of Isl1+ cardiovascular progenitors are controlled by a Wnt/beta-catenin pathway. Cell Stem Cell 1:165–179
60. Nemir M, Croquelois A, Pedrazzini T, Radtke F (2006) Induction of cardiogenesis in embryonic stem cells via downregulation of Notch1 signaling. Circ Res 98:1471–1478
61. Schroeder T, Meier-Stiegen F, Schwanbeck R, Eilken H, Nishikawa S, Hasler R, Schreiber S, Bornkamm GW, Nishikawa S, Just U (2006) Activated Notch1 alters differentiation of embryonic stem cells into mesodermal cell lineages at multiple stages of development. Mech Dev 123:570–579
62. Sadek H, Hannack B, Choe E, Wang J, Latif S, Garry MG, Garry DJ, Longgood J, Frantz DE, Olson EN et al (2008) Cardiogenic small molecules that enhance myocardial repair by stem cells. Proc Natl Acad Sci USA 105:6063–6068
63. Graichen R, Xu X, Braam SR, Balakrishnan T, Norfiza S, Sieh S, Soo SY, Tham SC, Mummery C, Colman A et al (2008) Enhanced cardiomyogenesis of human embryonic stem cells by a small molecular inhibitor of p38 MAPK. Differentiation 76:357–370
64. Sachinidis A, Schwengberg S, Hippler-Altenburg R, Mariappan D, Kamisetti N, Seelig B, Berkessel A, Hescheler J (2006) Identification of small signalling molecules promoting cardiac-specific differentiation of mouse embryonic stem cells. Cell Physiol Biochem 18:303–314
65. D'Amour KA, Bang AG, Eliazer S, Kelly OG, Agulnick AD, Smart NG, Moorman MA, Kroon E, Carpenter MK, Baetge EE (2006) Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. Nat Biotechnol 24:1392–1401
66. Kroon E, Martinson LA, Kadoya K, Bang AG, Kelly OG, Eliazer S, Young H, Richardson M, Smart NG, Cunningham J et al (2008) Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells *in vivo*. Nat Biotechnol 26:443–452
67. Borowiak M, Maehr R, Chen S, Chen AE, Tang W, Fox JL, Schreiber SL, Melton DA (2009) Small molecules efficiently direct endodermal differentiation of mouse and human embryonic stem cells. Cell Stem Cell 4:348–358

68. Chen S, Borowiak M, Fox JL, Maehr R, Osafune K, Davidow L, Lam K, Peng LF, Schreiber SL, Rubin LL et al (2009) A small molecule that directs differentiation of human ESCs into the pancreatic lineage. Nat Chem Biol 5:258–265
69. Zhu S, Wurdak H, Wang J, Lyssiotis CA, Peters EC, Cho CY, Wu X, Schultz PG (2009) A small molecule primes embryonic stem cells for differentiation. Cell Stem Cell 4:416–426
70. Tropepe V, Hitoshi S, Sirard C, Mak TW, Rossant J, van der Kooy D (2001) Direct neural fate specification from embryonic stem cells: a primitive mammalian neural stem cell stage acquired through a default mechanism. Neuron 30:65–78
71. Ying QL, Stavridis M, Griffiths D, Li M, Smith A (2003) Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. Nat Biotechnol 21:183–186
72. Kim M, Habiba A, Doherty JM, Mills JC, Mercer RW, Huettner JE (2009) Regulation of mouse embryonic stem cell neural differentiation by retinoic acid. Dev Biol 328:456–471
73. Chambers SM, Fasano CA, Papapetrou EP, Tomishima M, Sadelain M, Studer L (2009) Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat Biotechnol 27:275–280
74. Smith JR, Vallier L, Lupo G, Alexander M, Harris WA, Pedersen RA (2008) Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm. Dev Biol 313:107–117
75. Itsykson P, Ilouz N, Turetsky T, Goldstein RS, Pera MF, Fishbein I, Segal M, Reubinoff BE (2005) Derivation of neural precursors from human embryonic stem cells in the presence of noggin. Mol Cell Neurosci 30:24–36
76. Elkabetz Y, Panagiotakos G, Al Shamy G, Socci ND, Tabar V, Studer L (2008) Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes Dev 22:152–165

Index