Programming of DNA Methylation Patterns

Howard Cedar and Yehudit Bergman

Department of Developmental Biology and Cancer Research, Hebrew University Medical School, Ein Kerem, Jerusalem, Israel 91120; email: cedar@cc.huji.ac.il, yehuditb@ekmd.huji.ac.il

Annu. Rev. Biochem. 2012. 81:97–117

First published online as a Review in Advance on February 23, 2012

The *Annual Review of Biochemistry* is online at biochem.annualreviews.org

This article's doi: 10.1146/annurev-biochem-052610–091920

0066-4154/12/0707-0097$20.00

Keywords

maintenance, repression, chromatin, development, reprogramming, imprinting

Abstract

DNA methylation represents a form of genome annotation that mediates gene repression by serving as a maintainable mark that can be used to reconstruct silent chromatin following each round of replication. During development, germline DNA methylation is erased in the blastocyst, and a bimodal pattern is established anew at the time of implantation when the entire genome gets methylated while CpG islands are protected. This brings about global repression and allows housekeeping genes to be expressed in all cells of the body. Postimplantation development is characterized by stage- and tissue-specific changes in methylation that ultimately mold the epigenetic patterns that define each individual cell type. This is directed by sequence information in DNA and represents a secondary event that provides long-term expression stability. Abnormal methylation changes play a role in diseases, such as cancer or fragile X syndrome, and may also occur as a function of aging or as a result of environmental influences.

Contents

INTRODUCTION

Although the existence of methylated nucleotides has been known for a long time, studies of its function in vivo mostly concentrated on its role as a restriction modification mechanism in bacteria, where 6-methyladenine or 5-methylcytidine (5mC) appear at fixed nucleotide sequences throughout the genome. Nucleotide analysis of animal and plant DNA indicated that these organisms also have a fair amount of 5mC. Strikingly, when this DNA was subjected to nearest-neighbor analysis, a method that detects and quantifies the bases located 5′ adjacent to any labeled nucleotide, it was determined that almost all of this methylation is concentrated in the dinucleotide sequence CpG (1). Furthermore, in any particular cell type, only a portion of the CpGs is actually methylated (2). Taken together, these observations on the placement and distribution of methylation strongly suggested that this modification must play a different biological role in animals and plants as compared to bacteria.

5mC: 5-methylcytidine

PRINCIPLES OF METHYLATION

Methylation Metabolism

In order to gain some insight into the significance of this modification system, it was important to ask whether DNA methyl groups are freely metabolized throughout the genome or, alternatively, are located at fixed positions in each cell type. This was accomplished by using bacterial restriction enzymes to assay methylation at specific sites in the DNA (3). Enzymes such as HpaII (CCGG) or HhaI (GCGC), which have a CpG dinucleotide in their recognition sequences, are inhibited by methylation on the internal C residue. This unmethylated site can be cut, whereas the same site in a methylated form remains undigested. Using this assay, it was shown that different CpG residues in the genome are either highly methylated or present in an unmethylated form, strongly suggesting that methyl groups have fixed positions (4).

One of the major steps in understanding how methyl groups are organized and managed in the genome came about through the use of DNA-mediated gene transfer to stably insert foreign sequences into the endogenous genome

of fibroblast cells in culture. Plasmid DNA derived from *Escherichia coli* is unmethylated at all its HpaII (CCGG) recognition sequences, but these same sites could be artificially methylated in vitro using the specific methylases found in the same *Haemophilus* parainfluenza bacteria (5). Strikingly, unmethylated DNA remained unmodified in these cells, but methylated DNA retained its original methyl groups, even after many generations of growth in culture. These studies clearly showed that methylation is not the result of ongoing transient metabolism, but rather methylation marks must be located at specific sites in the genome. Furthermore, these patterns appeared to be stably maintained through cell division (6–8).

Maintenance of DNA Methylation Patterns

In addition to demonstrating that methylation patterns can be faithfully maintained, these transfection experiments also revealed important principles about the mechanism for this process. Because methylation was added artificially in vitro prior to the introduction of DNA into cells, maintenance of the methylation state clearly had nothing to do with the DNA sequence. This clearly suggested that modification patterns in somatic cells do not come about by the recruitment of factors through specific local sequence motifs. Rather, it appeared that maintenance must be accomplished by some sort of autonomous mechanism that can actually read and copy the modification pattern per se at the time of replication.

The key to understanding this process evolved from studies of the enzymatic DNA methylation activity found in crude extracts from somatic cells. These experiments showed that, although completely unmodified DNA is a rather poor substrate for methylation, hemimethylated DNA (i.e., methylated on one strand) works extremely efficiently with a 100-fold better K_m for this reaction (9). As previously noted, methyl moieties are located within CpG residues, and because these sites have strand symmetry, these results implied that DNA must be normally methylated on both strands. During replication, synthesis of the new strand generates a hemimethylated site that is then recognized by the maintenance methylase and therefore is methylated on the opposite strand, thereby regenerating the original bimethylated state originally present in the mother cell. In contrast, unmethylated CpG residues still appear completely unmethylated during replication and therefore do not constitute a substrate for the enzyme. In this way, the methylation pattern on the native strands serves as a template for regenerating methylation patterns during replication (10).

It is clear that the basis for this semiconservative mechanism derives from the fact that methyl groups in CpG residues are symmetrically disposed on both strands of the DNA. Strong support for this idea came from studies of plant DNA. Nearest-neighbor analysis of this DNA originally indicated that all C-containing dinucleotides are partially methylated. Upon closer examination, however, it could be seen that all of these methyl groups are actually located in CpNpG trinucleotide symmetrical sequences where N can be any of the four bases. Thus, every instance of modified C on one strand of the DNA is opposed by an equivalent methyl group two nucleotides over on the other strand (11). At the time of replication, methylation on the native DNA strand can thus serve as a template for complementing the methylation pattern on the newly synthesized strand, thus reproducing the methyl profile present in the mother cell.

It is now known that in animal cells the maintenance reaction is carried out by the enzyme Dnmt1 (12), which is perpetually localized to replication foci (13) and therefore is constantly available to provide CpG maintenance function. Exactly how hemimethylated DNA is recognized is not well understood (14, 15), but recent studies on Dnmt1 indicate that it does not operate alone. Rather, it probably works as part of a larger complex that includes other essential factors, such as Np95 (UHRF1) (16–18). Ultimately, these complexes constitute the biochemical basis for epigenetic

memory by providing an enzymatic platform for copying methylation patterns in a semiconservative manner in the same way that DNA sequence itself is reproduced from generation to generation.

CpG island: a region that contains a high density of CpGs and is located at the promoters of many genes

Methylation and Gene Repression

Once it became possible to use restriction enzymes for analyzing the DNA methylation patterns of specific endogenous genes, it became immediately obvious that this modification is correlated with gene repression. Tissue-specific genes were found to be highly methylated in most tissue samples but undermethylated in their tissue of expression (19). At the same time, housekeeping genes were shown to have a unique CpG island promoter structure, which is constitutively unmethylated in every cell (20, 21). Furthermore, early genomic studies based on analysis of total mRNA even demonstrated that active genes, in general, are undermethylated as compared to inactive DNA (22). This correlation has also been confirmed over and over again using more comprehensive assays, such as bisulfite analysis to measure endogenous gene DNA methylation (23).

Although these studies demonstrate a strong association with repression, it was still necessary to actually test whether DNA methylation has a causal effect on gene expression, and this was ultimately accomplished by employing DNA-mediated gene transfer. These experiments showed that unmethylated genes are actively transcribed when inserted into the genome, whereas the exact same sequences are repressed if the inserted DNA had been premethylated in vitro (6, 7, 24). Because the only difference between these templates is the presence of methyl groups, these studies provided convincing proof that DNA methylation itself is responsible for gene inhibition. Even though these experiments were based exclusively on the analysis of exogenous gene sequences, the implication of these results is that DNA methylation can explain why the homologous endogenous genes are completely silenced in these same exact cells, and this concept has been reinforced by subsequent transfection studies in vitro (25) and transgenic studies in vivo (26, 27), demonstrating that methylated DNA molecules are repressed.

When DNA methylation was first being explored as a mechanism of gene repression, it was commonly thought that methyl groups work by preventing the binding of key transcription factors much in the same way that bacterial restriction enzymes cannot cut when their recognition site is methylated. Although this may be true for a small number of specific regulatory factors, this is certainly not a general phenomenon. Many transcription factors do not have CpGs in their binding sites, and even when present, as is the case for Sp1 (28), DNA methylation does not necessarily inhibit their binding.

DNA Methylation and Chromatin Structure

A great deal of evidence points to the idea that the major effect of methyl groups is to model chromatin structure, and this may be carried out at many different levels. Microinjection experiments, for example, have demonstrated that a gene template inserted directly into the nucleus of cells in culture is initially unaffected by DNA methylation. Only after these substrates have had a chance to get packaged into a chromosomal structure does one begin to see the effects of this modification on transcription (29). The most convincing evidence for this idea comes from DNA-mediated cell transfection studies showing that unmethylated substrates are packaged into an open chromatin structure following their integration into the genome, whereas the exact same DNA remains completely resistant to DNase I if it is methylated (30). Because these experiments were carried out using bacterial DNA sequences that do not harbor any eukaryotic regulatory information, one can conclude that methyl moieties themselves must be responsible for generating a closed chromatin structure regardless of sequence context.

Although the precise mechanisms by which DNA methylation affects chromatin packaging

have not been completely elucidated, many studies have concentrated on nucleosome structure. Early experiments indicated that most methylated CpGs are concentrated within the central core of nucleosomes, as opposed to internucleosomal regions, suggesting that the positioning of this modification may have an intrinsic effect on where nucleosomes reside on the DNA (31, 32), and these results have recently been confirmed by sophisticated genome-wide analysis (33). In addition to this intrinsic effect, methylation may also play a role in regulating other factors that ultimately impinge on nucleosome displacement.

Another way that methylation mediates gene repression is through methyl-binding proteins, such as MeCP2, MBD2, and MBD3 (34). These factors can specifically recognize methyl moieties and, once bound, may model local chromatin structure, perhaps by recruiting modifying enzymes that bring about overall histone deacetylation (35, 36) or the methylation of specific lysine residues (37). Alternatively, methylation may actually prevent the binding of chromatin proteins, such as boundary-forming CTCF (38) or Cfp1 (39), which is known to be exclusively located on unmethylated CpG islands.

These observations on chromatin structure provide important insight on the significance of DNA methylation as a mechanism for stable gene silencing. Although gene expression profiles are mediated by the initial interactions with transcriptional regulatory factors, it is most likely the state of gene accessibility at the level of chromatin structure that actually determines gene activity or repression. In principle then, the chromatin conformation could provide a mold for preserving long-term gene expression profiles. The problem with this idea is that these basic structural features get disrupted in the wake of DNA polymerase advancement during every round of replication, and the newly made DNA must then be repackaged into chromatin (40). There is, as yet, no well-established mechanism for reproducing chromatin profiles at the replication fork (see Reference 41). Because the underlying DNA methylation pattern is autonomously copied following replication, it is possible that this modification pattern may play a role both in the generation of active chromatin over undermethylated regions as well as in the formation of closed chromatin over methylated DNA (42). It is probably in this manner that DNA methylation mediates stable gene repression.

Imprinting centers: regulatory regions that are marked in the gametes and that control domain-wide allele-specific expression

DNA METHYLATION PATTERNS DURING DEVELOPMENT

Erasure

Studies in tissue culture have been very helpful in characterizing basic concepts of DNA methylation metabolism and function, but these experiments do not reveal very much about the precise role of this modification during normal development in vivo. To understand this process, it was necessary to first map the dynamic pattern of DNA methylation as a function of embryogenesis and organogenesis (**Figure 1**). Early studies using restriction enzymes indicated that methyl groups inherited from parental gametes are largely erased in the preimplantation morula and blastula (43–45). This process of erasure appears to take place in two distinct stages, with much of the paternal genome undergoing active demethylation, which begins in the zygote (46), and further demethylation takes place during the first few early embryonic replication cycles, perhaps passively as a result of Dnmt1 relocation from the nucleus to the cytoplasm (47). Although the level of DNA methylation in the blastula is very low, the exact pattern of this modification in these preimplantation cells has not yet been accurately determined. Differentially methylated regions located at imprinting centers must have a mechanism to preserve them through this erasure stage (48, 49), and other studies indicate that additional specific regions and a number of different repeated sequences may also be partially protected from this process (50, 51).

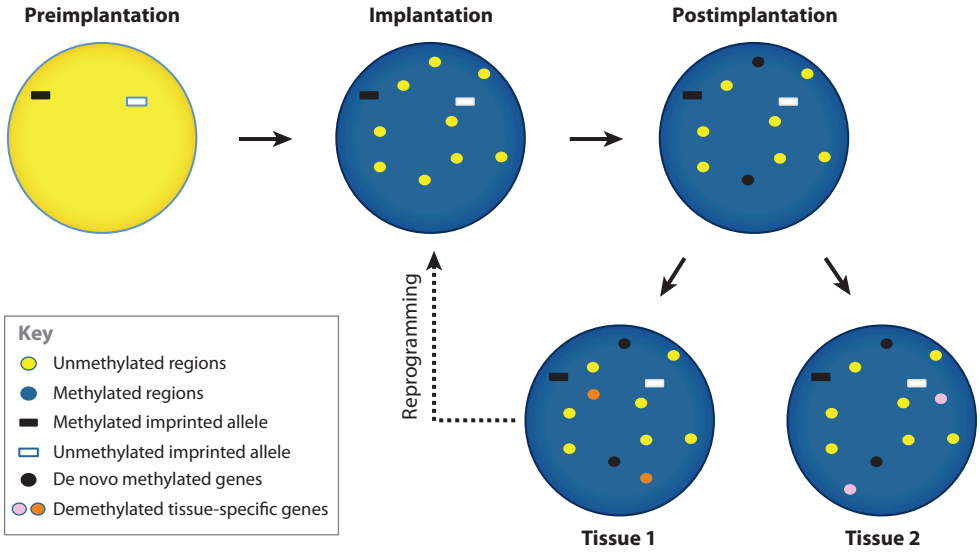

Figure 1

The generation of DNA methylation patterns during development. Almost all methylation in the gametes is erased (*yellow*) in the preimplantation embryo, but imprinting centers retain methylation on one allele (*black*). At the time of implantation, the entire genome gets methylated (*blue*), but the CpG islands are protected (*yellow circles*). Postimplantation, pluripotency genes are de novo methylated (*black*). Tissue-specific genes undergo demethylation (*orange* in Tissue 1, *pink* in Tissue 2) in their cell type of expression. Imprinting centers remain differentially methylated throughout development. Somatic cell reprogramming by induced pluripotent stem cells or fusion resets the methylation pattern of somatic cells to the stage of implantation.

Generation of a Bimodal Methylation Pattern

In the next stage of embryogenesis at about the time of implantation, the entire genome gets remodified through a dramatic wave of de novo methylation (**Figure 1**), and genetic analyses indicate that this is mediated by Dnmt3a and -3b, which are present at high concentrations at this stage of development (52). Strikingly, this process appears to generate a bimodal pattern of methylation, with most sequences becoming methylated to high levels (>80%), while CpG island-like windows are protected and therefore remain unmethylated (23). This concept of what happens in the early embryo derives from the observation that CpG island-like windows are constitutively unmodified in a large number of adult cell types, whereas other regions are constitutively methylated (53). This includes repeated sequences in satellite DNA that make up over 50% of all methyl moieties in the nucleus (32). Because all of these tissues are ultimately derived from the early inner cell mass at the time of implantation, it is reasonable to assume that the bimodal pattern is generated at this early stage and then maintained largely intact over succeeding cell divisions. This overall picture is also supported by the observation that mouse and human embryonic stem (ES) cells derived directly from this stage of development already show this same bimodal profile of DNA methylation (53, 54).

Mechanism of CpG Island Protection

Although the precise mechanism for CpG island protection is not known, it is quite clear that this is mediated by common sequence motifs present within the islands. This was initially deduced from transfection studies in ES cells. Unlike somatic cells, these cells can actually de novo methylate exogenously

introduced DNA sequences and, at the same time, also have the ability to recognize and protect CpG islands (55). Experiments using these cells in vitro indicated that Sp1 motifs play a role in this protection process, and this was further confirmed in vivo using transgenic mice (56, 57). Furthermore, these elements were even able to protect a non-CpG island sequence from de novo methylation at the time of implantation after being introduced as a transgene in one-cell embryos (26).

Although these studies indicate, in principle, that *cis* acting sequences are involved in protecting specific CpG islands from de novo methylation at the time of implantation, the general sequence rules for this mechanism have not been deciphered. Genome-wide studies indicate that almost all unmethylated CpG islands contain transcription start sites, are marked with H3K4me3 in ES cells (58, 59), and harbor many transcription factor–binding motifs (53). This suggests a model whereby sites of RNA polymerase binding in the blastocyst may serve as a mark for preventing de novo methylation during implantation. According to this idea, generation of a bimodal DNA methylation pattern in early development essentially serves to perpetuate the factor-mediated basal transcription profile of the preimplantation embryo.

Sequence Information Generates Methylation Patterns

By comparing constitutively unmethylated and methylated DNA sequences, it has been possible to derive a well-defined accurate algorithm that can distinguish between these two types of sequences, and by means of this mathematical formula, it is possible to accurately predict the basal methylation state of any sequence window in the genome (53). These studies serve to emphasize the idea that, although DNA methylation patterns represent an epigenetic mark that is not in itself inherited from the parents, the ultimate modification profile is fully determined by underlying sequence information within the DNA itself. Once this epigenetic code is deciphered, it should be possible to predict the full dynamics of DNA methylation during development.

Global Repression

De novo establishment of a bimodal methylation pattern in the early embryo has far-reaching implications for the management of gene regulation in the organism. Unlike simple biological systems where almost the entire genome is transcribed, even though only a relatively small number of genes are specifically recognized and repressed, animals have a large genome in which over 50% of the genes may be silenced in any given cell type. This type of repression pattern clearly requires a different mechanism that can carry out gene repression in a global manner without the need to recognize specific motifs on every target gene. According to this scheme, at the time of implantation, the entire genome undergoes global modification while CpG islands are protected. This insures that housekeeping genes that have CpG island promoters will be kept unmethylated and may also provide a mechanism for setting up regulatory modules, such as enhancers, in a permanent active configuration so they can bind key transcription factors at later stages of development (60).

At the same time, this wave of methylation guarantees that genes with nonisland promoters will automatically get modified and therefore be repressed in most tissues of the body (26, 27). While this relatively sparse background methylation may only have a small influence on the transcription levels of tissue-specific genes in their nonexpressing cell types, it also brings about a dramatic repression of endogenous viral sequences and foreign elements throughout the genome (61). It may also cause blanket inactivation of many cryptic promoters. Because, at this early stage, the entire genome undergoes DNA modification in a nonspecific manner, this event probably represents the only time in development when methylation itself actually serves as the primary cause of silencing. Even in this case, however, DNA methylation may

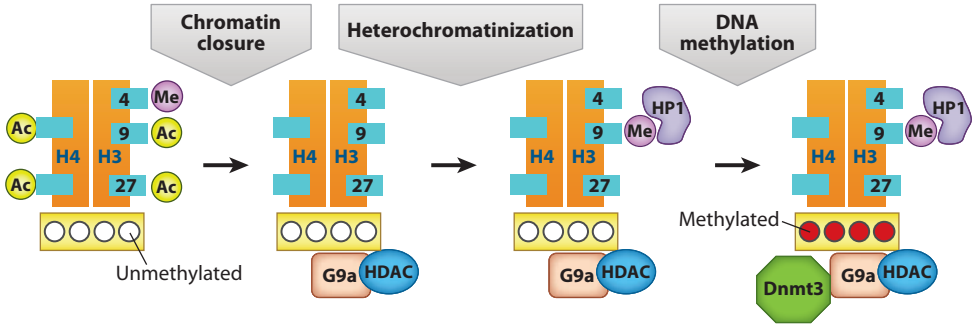

Figure 2

Inactivation of pluripotency genes. The *Oct-3/4* gene is unmethylated (*white circles*) at the implantation stage (ES cells) and active. With the onset of differentiation, transcription is first turned off by repressor proteins. The histone methylase G9a, together with a histone deacetylase (HDAC) and other chromatin modifying enzymes, is then recruited and brings about chromatin inactivation. In the next step, G9a methylates K9, and this serves as a binding site for heterochromatin protein 1 (HP1), thereby generating local heterochromatin. In the last step, G9a recruits de novo methylases (Dnmt3a and -3b) to cause promoter methylation (*red circles*).

represent only one of multiple factors that can mediate long-term repression through its effects on chromatin structure (62).

Postimplantation Methylation Changes

Following implantation, the animal genome can undergo additional changes in methylation, but these events are all of a tissue-specific or gene-specific nature. Perhaps one of the most significant of these modifications is that involved in the silencing of pluripotency genes (**Figure 2**). Genes, such as *Oct-3/4* and *Nanog*, for example, are active in the early embryo and still maintain an unmethylated transcribed promoter at the time of implantation. Following this stage, however, these genes undergo inactivation, thereby setting the stage for embryonic differentiation (63). Using ES cells as a model system, it has been possible to learn about the mechanism of this repression process, which occurs in three steps.

Pluripotency genes: genes, such as *Oct-3/4* and *Nanog*, that are required for increased potency in the early embryo

With the onset of differentiation, *Oct-3/4* transcription is initially turned off through a simple repression-factor mechanism. In the second step, the histone methyltransferase G9a is specifically recruited to the promoter of these genes, thus facilitating histone deacetylation and subsequent methylation of H3K9, which then binds HP1, forming heterochromatin (64). Finally, this same G9a complex can also recruit the de novo methylases Dnmt3a and -3b, thus bringing about methylation of the pluripotency-gene promoters themselves (**Figure 2**) (65–67). Although this de novo methylation represents a secondary event, it still plays an important role in stabilizing the silent state. Indeed, experiments in vitro, using mutant ES cells, clearly show that both the transcriptional repression and heterochromatinization steps are easily reversed in culture, but once DNA methylation has occurred, it is no longer possible to return to the pluripotent state (64, 65). In a similar manner, it has been demonstrated that viruses, initially inactivated by transient mechanisms in ES cells, may then become permanently silenced following differentiation-mediated de novo methylation (68).

Another major event that occurs soon after implantation throughout all cells of the embryo is the inactivation of one X chromosome in female animals. Here too, repression takes place as a multistep process, beginning with rapid chromosome-wide changes in replication timing, gene expression, and chromatin structure (69), followed by de novo methylation of CpG island promoters (70). In this case, as well, DNA

modification is probably mediated by histone methylases capable of generating heterochromatin and then recruiting Dnmts that carry out targeted local methylation many days after the initial inactivation event (71). These findings once again serve to emphasize the important concept that DNA methylation often plays a role as a secondary mechanism programmed to insure long-term silencing.

Tissue-Specific Methylation Patterns

Other alterations of the basic bimodal methylation pattern occur in a cell-type-specific manner. Many genes that are silenced throughout the organism and expressed specifically in a single tissue have non-CpG island promoters that automatically undergo de novo methylation at the time of implantation. During tissue development, these genes have to be specifically recognized by cell-type-specific factors that apparently recruit the molecules needed for demethylating their promoters (72), decondensing the overlying chromatin structures and making them accessible to the transcription machinery (**Figure 1**). A number of experiments in different cell types have demonstrated that this demethylation occurs in an active manner that does not require DNA replication (73, 74) and is mediated by specific *cis* acting sequences (75) and *trans* acting factors (76). Because this type of demethylation is specific and requires prior recognition, it cannot be considered the primary underlying cause of gene activation. Once demethylation has occurred, however, it is possible that this serves to stably maintain chromatin accessibility.

In a manner similar to demethylation, genes may also undergo tissue-specific de novo methylation at CpG island sequences that were originally protected at the time of implantation (77). Interestingly, these targets are not necessarily associated with promoters, and many are actually located within coding regions, where their methylation is associated with gene activation (53). Although the mechanism for this reverse effect is not known, one possibility is that these internal CpG islands carry promoters for antisense transcripts whose methylation would inhibit the production of these repressive RNA molecules. Alternatively, these regions may simply contain binding sites for transcriptional repressors. A large number of these targets have been shown to undergo de novo methylation in brain tissue, and genetic studies demonstrate that this may have profound effects on gene expression and function (78).

Although the precise mechanism for targeting de novo methylation in vivo is not yet known, it appears that this may be largely mediated by the polycomb complex (**Figure 3**). Almost all of the sites that undergo this type of modification are known polycomb targets (53), and it has been shown that the polycomb complex has an inherent ability to recruit Dnmt3a and -3b (79). It is still not known, however, which *cis* acting elements and *trans* acting factors may be involved in determining cell-type and gene-specific modification.

Polycomb complex: a complex of proteins bound at specific genes where they cause repression, probably through histone modification

ABNORMAL DNA METHYLATION

De Novo Methylation in Cancer

In the same way that normal development is dependent on the proper programming of methylation, abnormal cell behavior is also associated with alterations in DNA modification patterns. A prime example of this phenomenon is cancer. All tumors that have ever been examined show changes in DNA methylation, suggesting that this may represent a basic element of cancer biology that has a significant effect on tumor pathology. Early studies indicated, for example, that cancer is characterized by widespread demethylation (80), and using more advanced technologies, it has been possible to map this effect to specific blocks covering a large portion of the genome (81). In addition, almost all tumors undergo de novo methylation at specific sites (82). It has been commonly assumed that this modification process occurs in a completely random manner, with the final pattern seen in the tumor being determined mainly by

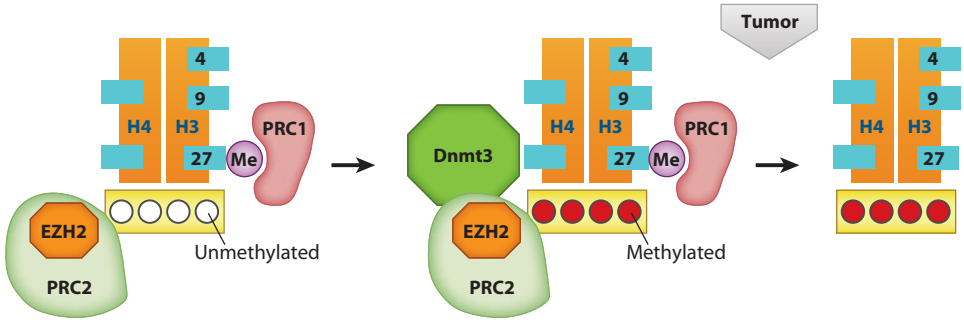

Figure 3

Targeted de novo methylation. Genes targeted by the polycomb complex, PRC2, normally have unmethylated CpG island promoters (*white circles*) but are repressed by virtue of the histone methylase EZH2-mediated methylation of H3K27, which then recruits the chromodomain-containing complex, PRC1, generating a form of heterochromatin. Targeted de novo methylation (*red circles*) can occur at specific sites during normal development or abnormally in cancer, and this probably occurs because EZH2 itself recruits Dnmt3a and Dnmt3b. In some tumors, it has been found that PRC2 may no longer remain bound to the methylated islands, leading to a situation where flexible polycomb repression is replaced with a more stable form of methylation-mediated repression (139).

selection, thus explaining why many tumor suppressor genes are found methylated in various cancer types. Although this may represent a small part of the methylation picture, more recent results have shown that DNA methylation in cancer is actually quite extensive, with hundreds of CpG islands in the genome getting methylated in a largely biallelic manner (83).

According to this selection idea, DNA methylation in the tumor serves to inactivate genes that would normally be expressed in the normal tissue, but transcription analysis showed that most methylation targets are actually inactive at the beginning (84). It appears that many of these sites harbor polycomb, and it may be this complex that actually recruits the methylases that then bring about this abnormal modification (**Figure 3**) (85–87). From this point of view, de novo methylation in cancer behaves in a programmed manner, in the sense that target sites are already predetermined independent of whether these genes are active or inactive or whether they actually play some role in tumorigenesis (88). Studies on human colon cancer suggest that de novo methylation occurs very early in the process of tumor evolution and that some of the polycomb target sites are associated with genes essential for epithelial cell differentiation (85). It is thus possible that DNA methylation plays a role in tumorigenesis by replacing the normally flexible polycomb repression mechanism with a more permanent silencing mode, thus inhibiting crypt cells from undergoing final differentiation and promoting the survival of proliferating precursors.

Role of Methylation in Cancer

The overall part played by DNA methylation in cancer may best be studied by using a genetics approach in mouse model systems. Min$^-$ mice carry a deletion in one allele of *Apc*, a key tumor suppresser gene that is deleted in almost every case of human colon cancer. These animals develop thousands of intestinal microadenomas owing to spontaneous deletion of the second *Apc* allele, but only about 5% of these actually develop into full-blown adenomas (89). When these mice are genetically manipulated to express lower levels of Dnmts, or when they are treated from birth with low doses of 5 azacytidine, the number of adenomas that is generated is reduced by almost 100-fold (90), whereas the number of microadenomas is unaffected (91). These observations suggest that by preventing de novo methylation after birth, it is

possible to reduce the incidence of tumor formation. Similar results have been obtained with other mouse tumor models (92).

When taken together, these observations support a model for intestinal cancer that involves two separate and independent molecular events; both of these are necessary for tumor formation. Spontaneous deletion of *Apc* is what brings about the generation of microadenomas, but abnormal de novo methylation is required to convert this into an authentic adenoma. In keeping with this concept, min^{-} mice that overexpress the de novo methylase, Dnmt3b, have increased de novo methylation at tumor target genes in several of their tissues and develop twofold more adenomas than are normally seen in this model system (89).

It has been known for a long time that cells growing in culture have abnormal methylation patterns, largely characterized by excess de novo methylation at CpG islands (93, 94). This is much different than the profile observed in ES cells growing in culture, probably because these stem cells have the molecular machinery required for setting up and preserving the correct bimodal pattern characteristic of the early embryo. Indeed, once these stem cells undergo differentiation, thereby losing this ability to protect CpG islands, they also become subject to abnormal sequence-directed de novo methylation (59, 95). In these cases, as well, polycomb-binding sites represent the major targets for new methylation (59, 96).

Fragile X Syndrome

Abnormal de novo methylation also plays a role in the pathogenesis of fragile X syndrome, causing the *FMR1* gene, carrying an expanded triplet repeat, to become repressed at a very early stage in development. Although the precise mechanism for this process has not been elucidated, inactivation is known to be accompanied by H3K9me3 heterochromatinization (97, 98). Taken together, these findings suggest that programming of abnormal modification may involve a common mechanism (71) whereby various histone methylases mediate unscheduled DNA methylation and repression.

DEMETHYLATION

Demethylation is a common event that occurs at a number of different stages during normal development. Early in preimplantation development, for example, the entire genome is subject to a wave of demethylation that apparently erases almost all of the methyl marks inherited from the parents, and a very similar process also takes place during early stages of gametogenesis (43, 44). In addition to this global form, there are also many instances of site-specific demethylation. This usually involves tissue-specific genes that become automatically methylated at the time of implantation and then undergo demethylation in their cell type of expression (19).

Active Demethylation

Extensive efforts over several decades have been invested in deciphering the biochemical mechanisms that may be involved in this demethylation process. Considering the way DNA methylation patterns are maintained throughout cell division, it was originally suggested that demethylation may occur through a passive mechanism whereby selected modified sites may not be efficiently recapitulated during replication. Although this appears to occur in an artificial manner when cells are treated with 5 azacytidine, early studies already demonstrated that demethylation in vivo largely occurs in an active manner, even in the absence of replication and cell division (73, 74).

The most obvious biochemical mechanism for demethylation would be the reverse of DNA methylation, involving direct removal of methyl groups from the 5′ position of cytosine. This type of demethylation has indeed been observed in extracts from cancer cells, where it was proposed that MBD2 may have an enzymatic activity that can combine methyl groups with H_2O to generate methanol (99). Even though this type of reaction may be chemically and

Fragile X syndrome: a developmental disease caused by a triplet repeat (CGG) expansion in the *FMR1* gene that brings bout mental retardation

energetically feasible (100), other laboratories have not yet been able to reproduce these studies (101).

Demethylation by Repair

An alternative mechanism for demethylation was suggested by experiments showing that 5mC may be a weak substrate for glycosylases. On this basis, it was proposed that the apyrimidinic base product of this reaction may then undergo DNA patch repair, which would result in the replacement of this damaged nucleotide with unmethylated cytosine, thereby completing the demethylation process (102). This idea was originally confirmed in studies showing nucleotide substitution both in vivo (103) and in vitro (104) and has now received wide support from a number of different cellular and biological systems showing that glycosylases (105–107) and excision-repair enzymes may indeed be essential for demethylation (108, 109).

One of the major problems with this proposed demethylation pathway is that 5mC is actually only a very weak substrate for the known glycosylases. It now appears, however, that demethylation may take place through a more complex process that involves first modifying 5mC to prepare it for glycosylation. 5-hydroxymethylcytidine (5hmC), for example, has been detected at a number of sites known to be undergoing demethylation in vivo and in vitro (110–112), suggesting that this may represent a necessary intermediate in the demethylation pathway. It has also been shown that the cytosine deaminase, Aid, may be required for demethylation during spermatogenesis (113).

Taken together, these studies suggest that demethylation may occur through a series of biochemical steps, ultimately leading to cytosine substitution by repair (**Figure 4**). In one possible scenario, 5mC is first hydroxylated, thereby generating 5hmC (114). This modified nucleoside may then be recognized by a deaminase, which coverts it to 5-hydroxymethyluridine (5hmU), a nucleotide variant that is normally removed by specific glycosylases to generate apyrimidinic acid sites, which are subsequently repaired. In keeping with this idea, many of the enzymatic components involved in this pathway have been found together in distinct complexes, apparently making up a type of demethylation machine (107, 115). The recent discovery of other ten-eleven-translocation (Tet)-catalyzed chemical changes at the 5′ position raises the possibility that demethylation may even involve additional intermediate reaction steps (116).

Evidence for this type of multistep repair process has been derived from a number of different demethylation examples in vivo, including the global removal of methyl groups seen during gametogenesis (108), the general paternal-genome demethylation that takes

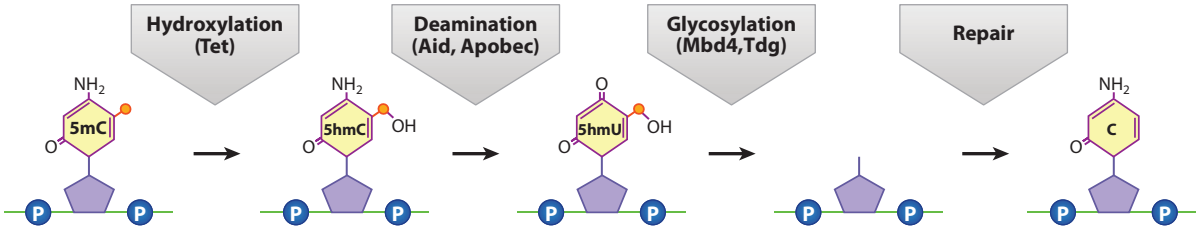

Figure 4

The demethylation pathway. Active demethylation may take place through a series of biochemical steps that modify 5-methylcytidine (5mC) to make it a recognized substrate for removal and replacement by repair with unmethylated cytosine. In the first step, a Tet enzyme brings about hydroxylation of the methyl group to form 5-hydroxymethylcytidine (5hmC). Deamination then occurs through the involvement of activation-induced deaminase (Aid) or Apobec family proteins to generate 5-hydroxymethyluridine (5hmU), which then becomes a substrate for a glycosylase (Mbd4 or Tdg). The resulting apyrimidinic acid residue is then replaced with C by means of path repair base or nucleotide excision repair (BER or NER).

place in the zygote (111), as well as tissue-specific demethylation (105, 114). This process evidently works in a modular manner. Thus, even though all demethylation events may occur through the same sequence of biochemical reactions, the enzymes that mediate these steps may vary in different cell types and at different stages of development. Indeed, hydroxymethylation (Tet1, Tet2, Tet3), deamination (Aid, Apobec), and glycosylation (Tdg, Mdb4) can all be carried out by different family members.

Reprogramming

As noted above, genes required for pluripotency, such as *Oct-3/4* or *Nanog*, undergo targeted methylation postimplantation and, in this way, prevent differentiated cells from undergoing dedifferentiation back to their pluripotent state. There is no question that the methylated state of these genes is what provides stability to the differentiated phenotype during normal development (65). Despite this layer of protection, it is still possible to reprogram somatic cells artificially either in vitro (117) or by nuclear transfer in vivo (118). How does this occur? A wide variety of studies in a number of different organisms have shown that somatic cell nuclei transplanted into primed denucleated oocytes undergo reprogramming to totipotency and then serve as the genetic source for generating an entire organism, including extraembryonic tissues (118). It appears that this is made possible by the global demethylation that takes place in the preimplantation embryo, which first enables trophectoderm differentiation by turning on the initially silenced *Elf5* gene (119), and then activates key early embryonic genes, such as *Oct-3/4* and *Nanog* (120, 121).

Reprogramming of somatic nuclei in vitro can be carried out in two different ways. The most efficient method is to fuse somatic cells with ES cells. When this is done, *Oct-3/4*, *Nanog*, and other pluripotency CpG island-like genes undergo rapid demethylation, thus setting in motion the regulatory network that defines the early ES cell phenotype (122). Previous experiments have shown that mouse ES cells have an activity capable of demethylating a wide variety of CpG islands in a reaction that is targeted by specific *cis* acting sequences, such as Sp1 (56, 57), and this probably includes the regulatory regions of *Oct-3/4* and other pluripotency genes (123). A similar process may occur in induced pluripotent stem (iPS) cells, where the addition of key stem-cell transcription factors apparently serve to turn on a set of endogenous ES master genes, including the factors needed to carry out CpG island-specific demethylation (**Figure 4**). It is interesting that, although demethylation of CpG islands generates a pluripotency phenotype, it is not sufficient for totipotency. This may be because genes like *Elf5* that are required for trophoblast formation (119) in the preimplantation embryo have a non-CpG island promoter, which remains methylated in ES cells.

Induced pluripotent stem (iPS) cells: these cells are generated by adding pluripotency factors to somatic cells, causing them to be reprogrammed

METHYLATION DURING GAMETOGENESIS

Erasure of Methylation Patterns

In addition to the DNA methylation changes that occur during early embryonic development, the process of gametogenesis is also characterized by an elaborate program of modification dynamics. Germ line cells probably emerge from the postimplantation embryo carrying the full bimodal methylation pattern and migrate along the genital ridge to ultimately take part in forming the gamete (124). These primitive germ cells undergo global demethylation at about 10–12 days post coitum (44). One of the important consequences of this event is that it facilitates the removal of all methyl sites associated with imprinted genes, and it is this erasure that allows reestablishment of new parent-specific methyl imprints at late stages of gametogenesis, either in the oocytes or sperm (125).

Following this global demethylation, the entire genome gets remethylated, while CpG islands are protected in a manner reminiscent of the process that occurs during implantation

in the early embryo (44). It thus appears that the cycle of erasure followed by reestablishment of a new methylation pattern is a basic aspect of epigenetic regulation. Both in man and in mouse, there are a large number of genes that emerge from spermatogenesis in an unmethylated form, even though they are fully methylated in all somatic cells, and many of these indeed have a testes-specific pattern of expression (53, 58). This profile of undermethylation may be generated by one of two possible mechanisms. Some genes clearly get methylated as part of the remethylation process and then undergo tissue-specific demethylation later in spermatogenesis (126). It is also possible, however, that some genes are actually protected from de novo methylation during early gametogenesis and then remain this way during the formation of a haploid genome in sperm. Recent evidence indicates that CpG islands also undergo specific methylation changes during oogenesis (51).

Imprinting

One of the important events that occurs during gametogenesis is the establishment of imprinting. Over 100 imprinted genes have been identified in mammals, and these are all clustered at well-defined loci in the genome, where *cis*-acting imprinting centers regulate their allele-specific expression. DNA methylation plays a major role in orchestrating this process. At most imprinted loci, the center undergoes de novo methylation during late oogenesis but remains unmodified during spermatogenesis. However, there are also a few loci where the imprinting center actually becomes methylated postmitotically in the testis, while remaining unmethylated in the oocytes. These parental-specific methylation profiles are generated by a combination of paternal or maternal gametic factors that either promote or prevent de novo methylation (127).

Once formed, differential methyl imprints are preserved in an allele-specific form throughout early embryogenesis and then serve as an epigenetic mark that either turns on or turns off the imprinting center's regulatory activity controlling all of the genes in the entire locus. It is not yet clear why these methylation patterns do not get erased in the preimplantation embryo (49) and do not lose their allele specificity during the wave of de novo methylation at the time of implantation, but it is possible that other maintenance mechanisms, such as replication timing (128) or histone modification (129), may play a supporting role. Following implantation, however, there is no question that allele-specific methylation is carried on through every replication cycle by the maintenance methylase, Dnmt1 (130).

NONPROGRAMMED INFLUENCES ON DNA METHYLATION

It has already been demonstrated that many CpG islands are subject to creeping de novo methylation as a function of aging in a variety of different cell types (131), suggesting that, in addition to built-in developmental events, the basic pattern of modification may also be influenced by nonprogrammed changes during the lifetime of the organism. The major question, however, is whether methylation patterns can also be influenced by environmental cues, either during embryogenesis or even afterward in the newborn or adult organism. A number of pioneering studies on this topic appear to indicate that this may indeed be possible.

Using a mouse model that contains a retrotransposon integrated upstream of the agouti gene, it was shown that offspring from each litter have variable coat color, ranging from dark brown to yellow, even though all the animals are isogenic, and molecular analysis indicated that this probably occurs because the agouti gene promoter is methylated to different degrees in each individual (132). This methylation variability appears to occur stochastically and probably takes place at about the time of implantation. The degree of methylation appears to be under the control of a number of different modifier genes located throughout the genome (133, 134), but the coat color in the offspring

can also be modulated by exposing pregnant mothers to ethanol (135) or to a methyl-rich diet (136). In general, there appear to be many natural sequences in the genome that are also subject to variable stochastic methylation during development, and it is clear that this effect can generate trait variability in isogenic mice (137). Similar stochastic effects may also be responsible for causing tumor variability (81).

These results on a mouse model confirm the idea that DNA methylation patterns can be influenced by environment, either during gestation or perhaps even after birth. Nonetheless, these changes do not seem to be fully inherited through the germ line into the next generation (136). In the case of the agouti trait, for example, all parent mice, regardless of coat color, still generate approximately the same range of variable offspring. This is consistent with the idea that methylation patterns are largely erased between generations first during gametogenesis, and then again in early embryogenesis (44). This picture is very different from the *trans*-generational epigenetic effects observed in plants, and this is undoubtedly caused by the lack of embryonic erasure in this organism (138).

SUMMARY POINTS

1. DNA methylation patterns are erased during preimplantation and then re-established throughout development via sequence information in the DNA.
2. Once established, DNA methylation patterns can be maintained autonomously through many cell divisions.
3. DNA methylation inhibits gene expression by affecting chromatin structure.
4. Changes in methylation during postimplantation development are usually secondary to factor-mediated gene activation or repression, but this subsequent methylation pattern provides long-term stability.
5. Active demethylation takes place through a multistep biochemical pathway that involves hydroxylation, deamination, glycosylation, and subsequent repair.
6. Somatic cell reprogramming involves resetting the methylation pattern to that of an early implantation embryo.
7. Abnormal methylation in cell lines and cancer takes place through a programmed process that involves the recruitment of de novo methylases by the polycomb complex.

FUTURE ISSUES

1. What is the mechanism involved in setting up methylation patterns? How is local sequence information translated in epigenetic information?
2. How does methylation play a role in lineage determination during development?
3. What are the roles of environment and aging on DNA methylation? What are methylation's effects on long-term physiology and disease susceptibility?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work was supported by research grants from the Israel Academy of Sciences (H.C., Y.B.), the National Institutes of Health (Y.B.), the Israel Cancer Research Foundation (H.C., Y.B.), the European Community 5th Framework Quality of Life Program (Y.B.), and the European Research Council (H.C.).

LITERATURE CITED

1. Sinsheimer RL. 1955. The action of pancreatic deoxyribonuclease. II. Isomeric dinucleotides. *J. Biol. Chem.* 215:579–83
2. Cedar H, Solage A, Glaser G, Razin A. 1979. Direct detection of methylated cytosine in DNA by use of the restriction enzyme MspI. *Nucleic Acids Res.* 6:2125–32
3. Bird AP. 1978. Use of restriction enzyme to study eukaryotic DNA methylation. II. The symmetry of methylation sites supports semiconservative copying of the methylation pattern. *J. Mol. Biol.* 118:49–60
4. van der Ploeg LHT, Flavell RA. 1980. DNA methylation in the human γ-globin locus in erythroid and nonerythroid tissues. *Cell* 19:947–58
5. Quint A, Cedar H. 1981. In vitro methylation of DNA with Hpa II methylase. *Nucleic Acids Res.* 9:633–46
6. Pollack Y, Stein R, Razin A, Cedar H. 1980. Methylation of foreign DNA sequences in eukaryotic cells. *Proc. Natl. Acad. Sci. USA* 77:6463–67
7. Wigler M, Levy D, Perucho M. 1981. The somatic replication of DNA methylation. *Cell* 24:33–40
8. Stein R, Gruenbaum Y, Pollack Y, Razin A, Cedar H. 1982. Clonal inheritance of the pattern DNA methylation in mouse cells. *Proc. Natl. Acad. Sci. USA* 79:61–65
9. Gruenbaum Y, Cedar H, Razin A. 1982. Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature* 295:620–22
10. Razin A, Riggs AD. 1980. DNA methylation and gene function. *Science* 210:604–10
11. Gruenbaum Y, Naveh-Many T, Cedar H, Razin A. 1981. Sequence specificity of methylation in higher plant DNA. *Nature* 292:860–62
12. Li E, Bestor TH, Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69:915–26
13. Leonhardt H, Page AW, Weier HU, Bestor TH. 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71:865–73
14. Song J, Rechkoblit O, Bestor TH, Patel DJ. 2011. Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. *Science* 331:1036–40
15. Rajakumara E, Wang Z, Ma H, Hu L, Chen H, et al. 2011. PHD finger recognition of unmodified histone H3R2 links UHRF1 to regulation of euchromatic gene expression. *Mol. Cell* 43:275–84
16. 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* 317:1760–64
17. Sharif J, Koseki H. 2011. Recruitment of Dnmt1: roles of the SRA protein Np95 (Uhrf1) and other factors. In *Progress in Molelcular Bioliolgy and Translational Science*, Vol. 101: *Modifications of Nuclear DNA and its Regulatory Proteins*, ed. X Cheng, RM Blumenthal, pp. 289–310. London: Acad. Press
18. Achour M, Jacq X, Ronde P, Alhosin M, Charlot C, et al. 2008. The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of *VEGF* gene expression. *Oncogene* 27:2187–97
19. Yisraeli J, Szyf M. 1984. Gene methylation patterns and expression. In *DNA Methylation: Biochemistry and Biological Significance*, ed. A Razin, H Cedar, AD Riggs, pp. 352–70. New York: Springer-Verlag
20. Stein R, Sciaky-Gallili N, Razin A, Cedar H. 1983. Pattern of methylation of two genes coding for housekeeping functions. *Proc. Natl. Acad. Sci. USA* 80:2422–26
21. Bird AP, Taggart M, Frommer M, Miller OJ, Macleod D. 1985. A fraction of the mouse genome that is derived from islands of nonmethylated CpG-rich DNA. *Cell* 40:91–99
22. Naveh-Many T, Cedar H. 1981. Active gene sequences are undermethylated. *Proc. Natl. Acad. Sci. USA* 78:4246–50

23. Laurent L, Wong E, Li G, Huynh T, Tsirigos A, et al. 2010. Dynamic changes in the human methylome during differentiation. *Genome Res.* 20:320–31
24. Stein R, Razin A, Cedar H. 1982. In vitro methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells. *Proc. Natl. Acad. Sci. USA* 79:3418–22
25. Schubeler D, Lorincz MC, Cimbora DM, Telling A, Feng YQ, et al. 2000. Genomic targeting of methylated DNA: influence of methylation on transcription, replication, chromatin structure, and histone acetylation. *Mol. Cell. Biol.* 20:9103–12
26. Siegfried Z, Eden S, Mendelsohn M, Feng X, Tzubari B, Cedar H. 1999. DNA methylation represses transcription in vivo. *Nat. Genet.* 22:203–6
27. Goren A, Simchen G, Fibach E, Szabo PE, Tanimoto K, et al. 2006. Fine tuning of globin gene expression by DNA methylation. *PLoS ONE* 1:e46
28. Höller M, Westin G, Jiricny J, Schaffner W. 1988. Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. *Genes Dev.* 2:1127–35
29. Buschhausen G, Wittig B, Graessmann M, Graessmann A. 1987. Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene. *Proc. Natl. Acad. Sci. USA* 84:1177–81
30. Keshet I, Lieman-Hurwitz J, Cedar H. 1986. DNA methylation affects the formation of active chromatin. *Cell* 44:535–43
31. Razin A, Cedar H. 1977. Distribution of 5-methylcytosine in chromatin. *Proc. Natl. Acad. Sci. USA* 74:2725–28
32. Solage A, Cedar H. 1978. Organization of 5-methylcytosine in chromosomal DNA. *Biochemistry* 17:2934–38
33. Chodavarapu RK, Feng S, Bernatavichute YV, Chen PY, Stroud H, et al. 2010. Relationship between nucleosome positioning and DNA methylation. *Nature* 466:388–92
34. Klose RJ, Bird AP. 2006. Genomic DNA methylation: the mark and its mediators. *Trends Biochem. Sci.* 31:89–97
35. Nan X, Ng H-H, Johnson CA, Laherty CD, Turner BM, et al. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393:386–89
36. Jones PL, Veenstra GJC, Wade PA, Vermaak D, Kass SU, et al. 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* 19:187–91
37. 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.* 278:4035–40
38. Bell AC, Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 405:482–85
39. Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, et al. 2010. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* 464:1082–86
40. Lucchini R, Sogo JM. 1995. Replication of transcriptionally active chromatin. *Nature* 374:276–80
41. Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, et al. 2008. A model for transmission of the H3K27me3 epigenetic mark. *Nat. Cell. Biol.* 10:1291–300
42. Hashimshony T, Zhang J, Keshet I, Bustin M, Cedar H. 2003. The role of DNA methylation in setting up chromatin structure during development. *Nat. Genet.* 34:187–92
43. Monk M, Boubelik M, Lehnert S. 1987. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* 99:371–82
44. Kafri T, Ariel M, Brandeis M, Shemer R, Urven L, et al. 1992. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germline. *Genes Dev.* 6:705–14
45. Chaillet JR, Vogt TF, Beier DR, Leder P. 1991. Parental-specific methylation of an imprinted transgene is established during gametogenesis and progressively changes during embryogenesis. *Cell* 66:77–83
46. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. 2000. Demethylation of the zygotic paternal genome. *Nature* 403:501–2
47. Carlson LL, Page AW, Bestor TH. 1992. Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes Dev.* 6:2536–41

48. Birger Y, Shemer R, Perk J, Razin A. 1999. The imprinting box of the mouse *Igf2r* gene. *Nature* 397:84–88
49. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. 2001. Dnmt3L and the establishment of maternal genomic imprints. *Science* 294:2536–39
50. Sanford JP, Clark HJ, Chapman VM, Rossant J. 1987. Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes Dev.* 1:1039–46
51. Smallwood SA, Tomizawa SI, Krueger F, Ruf N, Carli N, et al. 2011. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat. Genet.* 43:811–14
52. Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99:247–57
53. Straussman R, Nejman D, Roberts D, Steinfeld I, Blum B, et al. 2009. Developmental programming of CpG island methylation profiles in the human genome. *Nat. Struct. Mol. Biol.* 16:564–71
54. Mohn F, Weber M, Schübeler D, Roloff TC. 2009. Methylated DNA immunoprecipitation (MeDIP). *Methods Mol. Biol.* 507:55–64
55. Frank D, Keshet I, Shani M, Levine A, Razin A, Cedar H. 1991. Demethylation of CpG islands in embryonic cells. *Nature* 351:239–41
56. Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, et al. 1994. Sp1 elements protect a CpG island from de novo methylation. *Nature* 371:435–38
57. Macleod D, Charlton J, Mullins J, Bird AP. 1994. Sp1 sites in the mouse *aprt* gene promoter are required to prevent methylation of the CpG island. *Genes Dev.* 8:2282–92
58. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, et al. 2007. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* 39:457–66
59. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, et al. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454:766–70
60. Xu J, Watts JA, Pope SD, Gadue P, Kamps M, et al. 2009. Transcriptional competence and the active marking of tissue-specific enhancers by defined transcription factors in embryonic and induced pluripotent stem cells. *Genes Dev.* 23:2824–38
61. Walsh CP, Chaillet JR, Bestor TH. 1998. Transcription of IAP endogenous retrovirus is constrained by cytosine methylation. *Nat. Genet.* 20:116–17
62. Lande-Diner L, Zhang J, Ben-Porath I, Amariglio N, Keshet I, et al. 2007. Role of DNA methylation in stable gene repression. *J. Biol. Chem.* 282:12194–200
63. Scholer HR. 1991. Octamania: the POU factors in murine development. *Trends Genet.* 7:323–29
64. Feldman N, Gerson A, Fang J, Li E, Zhang Y, et al. 2006. G9a-mediated irreversible epigenetic inactivation of *Oct-3/4* during early embryogenesis. *Nat. Cell Biol.* 8:188–94
65. Epsztejn-Litman S, Feldman N, Abu-Remaileh M, Shufaro Y, Gerson A, et al. 2008. De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat. Struct. Mol. Biol.* 15:1176–83
66. Dong KB, Maksakova IA, Mohn F, Leung D, Appanah R, et al. 2008. DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J.* 27:2691–701
67. Tachibana M, Matsumura Y, Fukuda M, Kimura H, Shinkai Y. 2008. G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO J.* 27:2681–90
68. Gautsch JW, Wilson MC. 1983. Delayed de novo methylation in teratocarcinoma suggests additional tissue-specific mechanisms for controlling gene expression. *Nature* 301:32–37
69. Keohane AM, Lavender JS, O'Neill LP, Turner BM. 1998. Histone acetylation and X inactivation. *Dev. Genet.* 22:65–73
70. Lock LF, Takagi N, Martin GR. 1987. Methylation of the *Hprt* gene on the inactive X occurs after chromosome inactivation. *Cell* 48:39–46
71. Cedar H, Bergman Y. 2009. Linking DNA methylation and histone modification: patterns and paradigms. *Nat. Rev. Genet.* 10:295–304
72. Yisraeli J, Adelstein RS, Melloul D, Nudel U, Yaffe D, Cedar H. 1986. Muscle-specific activation of a methylated chimeric actin gene. *Cell* 46:409–16
73. Sullivan CH, Grainger RM. 1986. δ-Crystallin genes become hypomethylated in postmitotic lens cells during chicken development. *Proc. Natl. Acad. Sci. USA* 83:329–33

74. Paroush Z, Keshet I, Yisraeli J, Cedar H. 1990. Dynamics of demethylation and activation of the α-actin gene in myoblasts. *Cell* 63:1229–37
75. Lichtenstein M, Keini G, Cedar H, Bergman Y. 1994. B cell-specific demethylation: a novel role for the intronic κ-chain enhancer sequence. *Cell* 76:913–23
76. Kirillov A, Kistler B, Mostoslavsky R, Cedar H, Wirth T, Bergman Y. 1996. A role for nuclear NF-κB in B-cell-specific demethylation of the Igκ locus. *Nat. Genet.* 13:435–41
77. Illingworth R, Kerr A, Desousa D, Jørgensen H, Ellis P, et al. 2008. A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol.* 6:e22
78. Wu H, Coskun V, Tao J, Xie W, Ge W, et al. 2010. Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* 329:444–48
79. Viré E, Brenner C, Deplus R, Blanchon L, Fraga M, et al. 2006. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439:871–74
80. Feinberg AP, Vogelstein B. 1983. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* 301:89–92
81. Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, et al. 2011. Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet.* 43:768–75
82. Baylin SB, Höppener JW, de Bustros A, Steenbergh PH, Lips CJ, Nelkin BD. 1986. DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. *Cancer Res.* 46:2917–22
83. Zardo G, Tiirikainen MI, Hong C, Misra A, Feuerstein BG, et al. 2002. Integrated genomic and epigenomic analyses pinpoint biallelic gene inactivation in tumors. *Nat. Genet.* 32:453–58
84. Keshet I, Schlesinger Y, Farkash S, Rand E, Hecht M, et al. 2006. Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nat. Genet.* 38:149–53
85. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, et al. 2007. Polycomb mediated histone H3(K27) methylation pre-marks genes for de novo methylation in cancer. *Nat. Genet.* 39:232–36
86. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, et al. 2007. Epigenetic stem cell signature in cancer. *Nat. Genet.* 39:157–58
87. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, et al. 2007. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nat. Genet.* 39:237–42
88. Luckow B, Schütz G. 1987. CAT constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoter and regulatory elements. *Nucleic Acids Res.* 15:5490
89. Linhart HG, Lin H, Yamada Y, Moran E, Steine EJ, et al. 2007. Dnmt3b promotes tumorigenesis in vivo by gene-specific de novo methylation and transcriptional silencing. *Genes Dev.* 21:3110–22
90. Laird PW, Jackson-Grusby L, Fazeli A, Dickinson SL, Jung WE, et al. 1995. Suppression of intestinal neoplasia by DNA hypomethylation. *Cell* 81:197–205
91. Lin H, Yamada Y, Nguyen S, Linhart H, Jackson-Grusby L, et al. 2006. Suppression of intestinal neoplasia by deletion of Dnmt3b. *Mol. Cell. Biol.* 26:2976–83
92. McCabe MT, Low JA, Daignault S, Imperiale MJ, Wojno KJ, Day ML. 2006. Inhibition of DNA methyltransferase activity prevents tumorigenesis in a mouse model of prostate cancer. *Cancer Res.* 66:385–92
93. Jones PA, Wolkowicz MJ, Rideout WM 3rd, Gonzales FA, Marziasz CM, et al. 1990. De novo methylation of the MyoD1 CpG island during the establishment of immortal cell lines. *Proc. Natl. Acad. Sci. USA* 87:6117–21
94. Antequera F, Boyes J, Bird A. 1990. High levels of de novo methylations and altered chromatin structure at CpG islands in cell lines. *Cell* 62:503–14
95. Shen Y, Chow J, Wang Z, Fan G. 2006. Abnormal CpG island methylation occurs during in vitro differentiation of human embryonic stem cells. *Hum. Mol. Genet.* 15:2623–35
96. Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, et al. 2008. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* 30:755–66
97. Pietrobono R, Tabolacci E, Zalfa F, Zito I, Terracciano A, et al. 2005. Molecular dissection of the events leading to inactivation of the *FMR1* gene. *Hum. Mol. Genet.* 14:267–77
98. Coffee B, Zhang F, Ceman S, Warren ST, Reines D. 2002. Histone modifications depict an aberrantly heterochromatinized *FMR1* gene in fragile X syndrome. *Am. J. Hum. Genet.* 71:923–32

99. Bhattacharya SK, Ramchandani S, Cervoni N, Szyf M. 1999. A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* 397:579–83
100. Cedar H, Verdine GL. 1999. Gene expression. The amazing demethylase. *Nature* 397:568–69
101. Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, et al. 1999. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat. Genet.* 23:58–61
102. Jost JP. 1993. Nuclear extracts of chicken embryos promote an active demethylation of DNA by excision repair of 5-methyldeoxycytidine. *Proc. Natl. Acad. Sci. USA* 90:4684–88
103. Razin A, Szyf M, Kafri T, Roll M, Giloh H, et al. 1986. Replacement of 5-methylcytosine by cytosine: a possible mechanism for transient DNA demethylation during differentiation. *Proc. Natl. Acad. Sci. USA* 83:2827–31
104. Weiss A, Keshet I, Razin A, Cedar H. 1996. DNA demethylation in vitro: involvement of RNA. *Cell* 86:709–18
105. Kim MS, Kondo T, Takada I, Youn MY, Yamamoto Y, et al. 2009. DNA demethylation in hormone-induced transcriptional derepression. *Nature* 461:1007–12
106. Cortazar D, Kunz C, Selfridge J, Lettieri T, Saito Y, et al. 2011. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* 470:419–23
107. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, et al. 2011. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 146:67–79
108. Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, Surani MA. 2010. Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* 329:78–82
109. Barreto G, Schafer A, Marhold J, Stach D, Swaminathan SK, et al. 2007. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 445:671–75
110. Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, et al. 2011. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* 473:389–93
111. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, et al. 2011. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun.* 2:241
112. Iqbal K, Jin SG, Pfeifer GP, Szabo PE. 2011. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl. Acad. Sci. USA* 108:3642–47
113. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, et al. 2010. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 463:1101–5
114. Guo JU, Su Y, Zhong C, Ming GL, Song H. 2011. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* 145:423–34
115. 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* 135:1201–12
116. Ito S, Shen L, Dai Q, Wu SC, Collins LB, et al. 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333:1300–3
117. Yamanaka S. 2009. A fresh look at iPS cells. *Cell* 137:13–17
118. Jaenisch R, Hochedlinger K, Blelloch R, Yamada Y, Baldwin K, Eggan K. 2004. Nuclear cloning, epigenetic reprogramming, and cellular differentiation. *Cold Spring Harb. Symp. Quant. Biol.* 69:19–27
119. Ng RK, Dean W, Dawson C, Lucifero D, Madeja Z, et al. 2008. Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. *Nat. Cell. Biol.* 10:1280–90
120. Boiani M, Eckardt S, Scholer HR, McLaughlin KJ. 2002. Oct4 distribution and level in mouse clones: consequences for pluripotency. *Genes Dev.* 16:1209–19
121. Bortvin A, Eggan K, Skaletsky H, Akutsu H, Berry DL, et al. 2003. Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. *Development* 130:1673–80
122. Cowan CA, Atienza J, Melton DA, Eggan K. 2005. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* 309:1369–73
123. Gidekel S, Bergman Y. 2002. A unique developmental pattern of *Oct-3/4* DNA methylation is controlled by a *cis*-demodification element. *J. Biol. Chem.* 277:34521–30
124. McCarrey JR. 1993. Development of the germ cell. In *Cell and Molecular Biology of the Testis*, ed. C Desjardins, LL Ewing, pp. 58–89. New York: Oxford Univ. Press

125. Shemer R, Razin A. 1996. Establishment of imprinted methylation patterns during development. In *Epigenetic Mechanisms of Gene Regulation*, ed. VEA Russo, RA Martienssen, AD Riggs, pp. 215–29. Cold Spring Harbor, NY: Cold Spring Harb. Lab. Press
126. Ariel M, Cedar H, McCarrey JR. 1994. Developmental changes in methylation of spermatogenesis-specific genes include reprogramming in the epididymis. *Nat. Genet.* 7:59–63
127. Hudson QJ, Kulinski TM, Huetter SP, Barlow DP. 2010. Genomic imprinting mechanisms in embryonic and extraembryonic mouse tissues. *Heredity* 105:45–56
128. Simon I, Tenzen T, Reubinoff BE, Hillman D, McCarrey JR, Cedar H. 1999. Asynchronous replication of imprinted genes is established in the gametes and maintained during development. *Nature* 401:929–32
129. Ng RK, Gurdon JB. 2008. Epigenetic inheritance of cell differentiation status. *Cell Cycle* 7:1173–77
130. Li E, Beard C, Jaenisch R. 1993. Role for DNA methylation in genomic imprinting. *Nature* 366:362–65
131. Maegawa S, Hinkal G, Kim HS, Shen L, Zhang L, et al. 2010. Widespread and tissue specific age-related DNA methylation changes in mice. *Genome Res.* 20:332–40
132. Morgan HD, Sutherland HG, Martin DI, Whitelaw E. 1999. Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* 23:314–18
133. Chong S, Vickaryous N, Ashe A, Zamudio N, Youngson N, et al. 2007. Modifiers of epigenetic reprogramming show paternal effects in the mouse. *Nat. Genet.* 39:614–22
134. Blewitt ME, Vickaryous NK, Hemley SJ, Ashe A, Bruxner TJ, et al. 2005. An *N*-ethyl-*N*-nitrosourea screen for genes involved in variegation in the mouse. *Proc. Natl. Acad. Sci. USA* 102:7629–34
135. Kaminen-Ahola N, Ahola A, Maga M, Mallitt KA, Fahey P, et al. 2010. Maternal ethanol consumption alters the epigenotype and the phenotype of offspring in a mouse model. *PLoS Genet.* 6:e1000811
136. Daxinger L, Whitelaw E. 2010. Transgenerational epigenetic inheritance: more questions than answers. *Genome Res.* 20:1623–28
137. Whitelaw NC, Chong S, Whitelaw E. 2010. Tuning in to noise: epigenetics and intangible variation. *Dev. Cell* 19:649–50
138. Feng S, Jacobsen SE, Reik W. 2010. Epigenetic reprogramming in plant and animal development. *Science* 330:622–27
139. Gal-Yam EN, Egger G, Iniguez L, Holster H, Einarsson S, et al. 2008. Frequent switching of Polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. *Proc. Natl. Acad. Sci. USA* 105:12979–84

Annual Review of Biochemistry
Volume 81, 2012

Contents

Recent Advances in Biochemistry

Indexes

Errata

An online log of corrections to *Annual Review of Biochemistry* articles may be found at http://biochem.annualreviews.org/errata.shtml