

Annotating the genome by DNA methylation

HOWARD CEDAR* and AHARON RAZIN

Department of Developmental Biology & Cancer Research, Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel

ABSTRACT DNA methylation plays a prominent role in setting up and stabilizing the molecular design of gene regulation and by understanding this process one gains profound insight into the underlying biology of mammals. In this article, we trace the discoveries that provided the foundations of this field, starting with the mapping of methyl groups in the genome and the experiments that helped clarify how methylation patterns are maintained through cell division. We then address the basic relationship between methyl groups and gene repression, as well as the molecular rules involved in controlling this process during development *in vivo*. Finally, we describe ongoing work aimed at defining the role of this modification in disease and deciphering how it may serve as a mechanism for sensing the environment.

KEY WORDS: *DNA methylation, chromatin, epigenetics, gene regulation, repression*

Introduction

It is now well accepted that there is a sophisticated chemical system to annotate the genetic text and in this way regulate the way genes are read and interpreted in each cell. Nonetheless, in the early seventies very little was known about how the animal genome is controlled. The elegant work of Jacob and Monod had paved the way for understanding the fundamental idea of how cis-acting elements and trans-acting factors control transcription in bacteria, but animal cells appeared to have complex patterns of gene expression which could not easily be explained on the basis of this principle alone.

Early studies indicated that the DNA in animal cells is not free in solution, but rather packaged within an ordered protein structure made up of well-defined histones and this served to compress the genetic material (Felsenfeld, 1978). Using RNA polymerase to probe this structure, it was demonstrated that individual genes are available for transcription in chromatin from the cell type of expression, while being closed in other tissues (Axel *et al.*, 1973; Cedar and Felsenfeld, 1973). This was confirmed by similar experiments using nucleases to probe for regions that are open to digestion (Weintraub and Groudine, 1976; Panet and Cedar, 1977). These studies helped establish a brand new concept that pictured the 3D-genome map as being mostly inaccessible to the transcription machinery while specific loci were available for factor interactions. It thus appeared that chromatin structure may play a big role in gene regulation.

Methylation and expression

It was always assumed that the genetic text is identical in every cell of the body, having been derived directly from the parental gametes. It had already been shown in the 1940's, however, that at least one pyrimidine base, Cytosine, was often found to be methylated in animal DNA, although literally nothing was known about its distribution. Understanding the overall structure of chromatin allowed us, for the first time, to probe the possible relevance of DNA methylation to gene expression. In a very simple experiment using micrococcal nuclease digestion to isolate core nucleosomal DNA, we demonstrated that this protected fraction is relatively methylated compared to open regions (Razin and Cedar, 1977; Solage and Cedar, 1978). This indicated that DNA methylation is not distributed evenly across the genome, representing the first hint that this modification may be correlated with gene repression.

By employing restriction enzymes to measure DNA methylation levels at specific sites in the genome, it was demonstrated that a number of individual genes are preferentially unmethylated in their tissue of expression while being highly modified in other cell types (van der Ploeg and Flavell, 1980). In our laboratory, we used DNase I to carry out nick translation on chromatin, thereby selectively labeling active regions of the genome, which were then shown to be preferentially hypomethylated (Fig. 1). This was then confirmed

Abbreviations used in this paper: HBB, β globin; IE, island element; Po, polyp.

*Address correspondence to: Howard Cedar, Dept. Developmental Biology & Cancer Research, Faculty of Medicine, Hebrew University, P.O.B. 12272 Ein Kerem, Jerusalem 91120 Israel. Fax: 972-2-641-5848. Tel: 972-2-675-8167. E-mail: cedar@mail.huji.ac.il - web: http://haimcedar.ekmd.huji.ac.il

Accepted: 10 August, 2016.

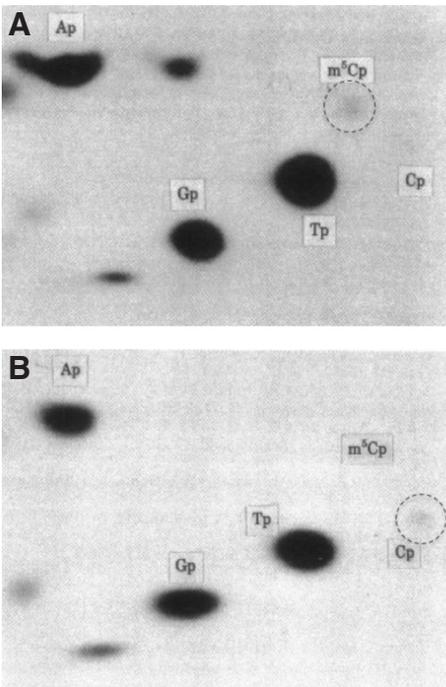


Fig. 1. Active DNA is undermethylated. Nearest neighbor analysis shows the distribution of all nucleotides (A, G, T and C) located near G in total DNA (A). Over 75% of the CpG dinucleotide is methylated. In contrast, specifically labeled (by nick translation) active DNA (B) has mostly unmethylated CpGs. This indicates that active regions of the genome are undermethylated. Adapted from Naveh-Many and Cedar (1981).

by isolating DNA complementary to the total mRNA population, which again was found to be undermethylated compared to the rest of the genome (Naveh-Many and Cedar, 1981). Taken together these studies demonstrated that DNA methylation is correlated with repression, strongly suggesting that it may actually play a role in the control of transcription.

DNA methylation patterns

The observation that some sites in a given cell population are consistently methylated while others are unmethylated strongly suggested that there are fixed patterns of modification that are maintained in dividing cells. In order to prove this concept, we adopted the newly invented technology of DNA mediated gene transfer (transfection) (Wigler *et al.*, 1977). Bacterially-cloned DNA inserted into animal cells becomes integrated into the genome and remains completely unmethylated. In contrast, when this same DNA was first methylated *in vitro* using the HpaII methylase (CCGG), these specific sites retained their modified state regardless of the underlying sequence, even many cell generations later (Pollack *et al.*, 1980; Stein *et al.*, 1982a). This seminal experiment demonstrated that DNA methylation patterns are not generated *de novo* in somatic cells, and are not necessarily directed by transacting factors, but once formed, can be faithfully maintained in dividing cells.

These transfection studies strongly implied that there must be an autonomous mechanism for preserving methylation patterns, one that is not dependent on sequence. It had been suggested that DNA methylation may actually be “copied” during replication in a manner similar to the sequence text itself, perhaps by

utilizing DNA methyltransferases found in the nucleus (Razin and Riggs, 1980). This idea was consistent with the observation that all methyl groups in animal-cell DNA are located at CpG sites and are therefore symmetrically placed on both strands (Cedar *et al.*, 1979). To test this concept, we extracted methylase activity from animal cells and used this to methylate various forms of DNA *in vitro*. Strikingly, this enzyme was found to preferentially modify a hemi-methylated substrate with specificity over 100-fold greater than completely unmethylated DNA (Gruenbaum *et al.*, 1982) (Fig. 2).

This experiment suggested that during replication, synthesis of the new strand would automatically generate a hemi-methylated site which would then be recognized by the DNA methyltransferases that would fill in a methyl group opposite that of the native DNA. In contrast, unmethylated CpG residues would not be recognized by this enzyme. In this way, the methylation pattern present in the mother cell would be faithfully copied to the daughter cells in an autonomous manner. The foundation for this mechanism is the dinucleotide symmetry of CpG residues that enable the methylation on one strand to serve as a template for methylating the second strand using Dnmt1 together with additional protein factors (Bostick *et al.*, 2007).

Support for this concept also came from our studies on methylation in plant DNA. Nearest-neighbor experiments indicated that as opposed to animal DNA, methylation in plants was observed at additional non-symmetric C-containing dinucleotides, including CpC, CpA and CpT. While this initial observation did not appear to be compatible with a maintenance mechanism, extended analysis revealed that all of these methylations were actually associated with the more general sequence family, CXG, containing a trinucleotide symmetry (Gruenbaum *et al.*, 1981) and subsequent experiments later showed that plants indeed contain enzymes capable of recognizing hemi-methylated DNA at these sites, thus providing the machinery for autonomous maintenance in plant cells. Taken together, these experiments provide the basis for understanding

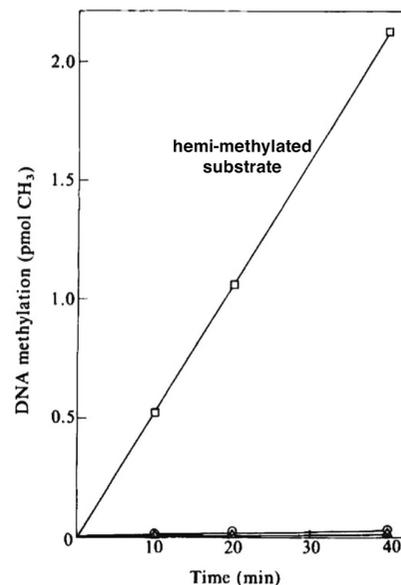


Fig. 2. Specificity of methylase for hemi-methylated DNA. DNA methylase activity was extracted from cells and used on unmethylated DNA (Δ), methylated DNA (○) or a hemi-methylated DNA substrate (□). This experiment explains how maintenance methylation works. Adapted from (Gruenbaum *et al.*, 1982).

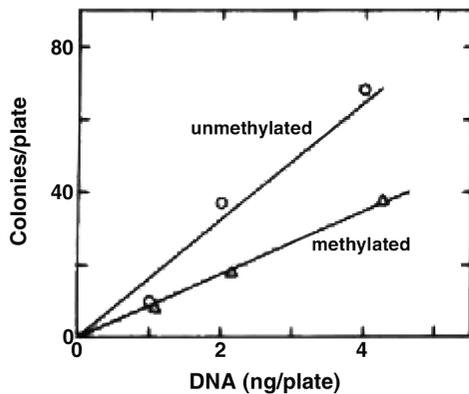


Fig. 3. Methylation inhibits gene expression. Transfection efficiency of an unmethylated (○) or methylated (△) *tk* gene at various concentrations into *Ltk⁻* cells in culture, with selection for *tk⁺*. Note that methylation inhibits *tk* gene activity. Adapted from (Pollack *et al.*, 1980).

how DNA methylation patterns are maintained through replication, thereby serving as an efficient memory mechanism.

DNA methylation represses gene expression

All of the initial studies on DNA methylation indicated a good correlation between gene expression and undermethylation, suggesting, but not proving, that the presence of methylation may operate to inhibit transcription. We designed experiments to actually test this hypothesis using DNA-mediated gene transfer. We stably introduced the HSV *tk* gene into mouse L-cells in culture and then compared this to the same *tk* gene that had been methylated *in vitro* using the *HpaII* methylase, which modifies all CCGG sites (Pollack *et al.*, 1980). Strikingly, the methylated gene was much less active (Fig. 3) and similar results were obtained for a variety of different co-transfected genes. Since in this experiment the only difference between the two templates was their *in-vitro* added methylation, it was clear that it must be the methyl groups themselves that repress gene activity. Housekeeping genes that are normally unmethylated in almost all cells were inhibited (Stein *et al.*, 1982b) as well as tissue-specific genes, such as β -globin (Yisraeli *et al.*, 1988). Since the endogenous β -globin gene in these mouse cells is in a methylated state, this experiment suggested indirectly that normal tissue-specific genes of this nature are probably repressed in most cells by virtue of their methylation.

While these studies had shown that the presence of DNA methylation inhibits transcription, it was not clear how this might be carried out. Considering that expression, in general, appeared to be controlled by accessibility, we did a simple experiment to test whether DNA methylation itself may affect chromatin structure. To this end, we stably inserted bacterial DNA into animal cells and measured its sensitivity to DNaseI once it had integrated in the endogenous chromosomal material. As expected, this DNA was shown to be sensitive to nucleases. In contrast, when this same DNA was first methylated *in vitro*, it still integrated into the DNA, but in a form that was resistant to DNaseI treatment (Keshet *et al.*, 1986) (Fig. 4), and became wrapped in nucleosomes carrying deacetylated histones, another indication of inactive DNA (Eden *et al.*, 1998). This effect was exclusively dependent on DNA methylation itself and occurred despite the fact that these were foreign

templates carrying no eukaryotic regulatory sequences, and taken together, these results clearly showed that DNA methylation plays a role in controlling chromatin structure and gene accessibility.

DNA methylation in development

After having derived the basic rules of DNA methylation, we next turned our attention to understanding how DNA methylation patterns are established during development. Studies in the 80s indicated that in the very early embryo the general levels of DNA methylation were low, but these experiments did not give information about specific gene sequences (Monk *et al.*, 1987). The main problem was that a more sensitive assay was needed to measure DNA methylation in samples with extremely low cell numbers. With this in mind, we developed a method for carrying out methylation-sensitive restriction enzyme analysis in combination with PCR to amplify the DNA. Our results showed that many different gene sequences originally methylated in oocyte and sperm DNA appear to undergo a slow process of erasure in the preimplantation embryo (Kafri *et al.*, 1992). These sequences ultimately become methylated again at about the time of implantation, while CpG islands remain unmethylated throughout this entire process. These studies demonstrated that DNA methylation apparently undergoes a process of reprogramming, whereby methylation marks present in the gametes are removed and a new pattern then generated in each individual (Fig. 5). This made it clear that methylation patterns are not inherited from parents, and thus represent an epigenetic, as opposed to a genetic mark. More recent studies using high throughput bisulfite sequencing analysis have confirmed these results both in mouse and in man (Smith *et al.*, 2012; Smith *et al.*, 2014).

The studies on DNA methylation patterns during early embryogenesis indicated that implantation is accompanied by a wave of *de novo* methylation that appears to modify many sites, but leaves CpG islands unmethylated. Since this was clearly a major defining step in the generation of methylation profiles, we sought to understand how this worked. To this end, we developed an *in vitro* system employing ES cells. Although these cells initially appeared to be derived from the blastula, in terms of DNA methylation, they

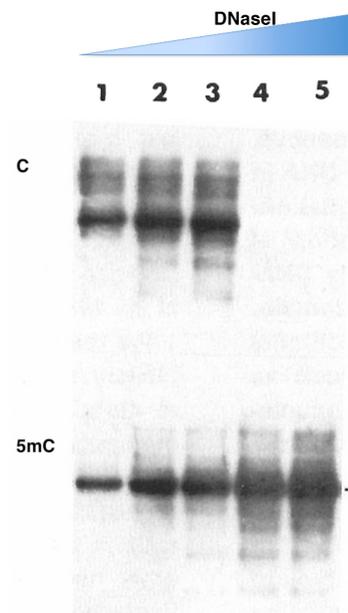
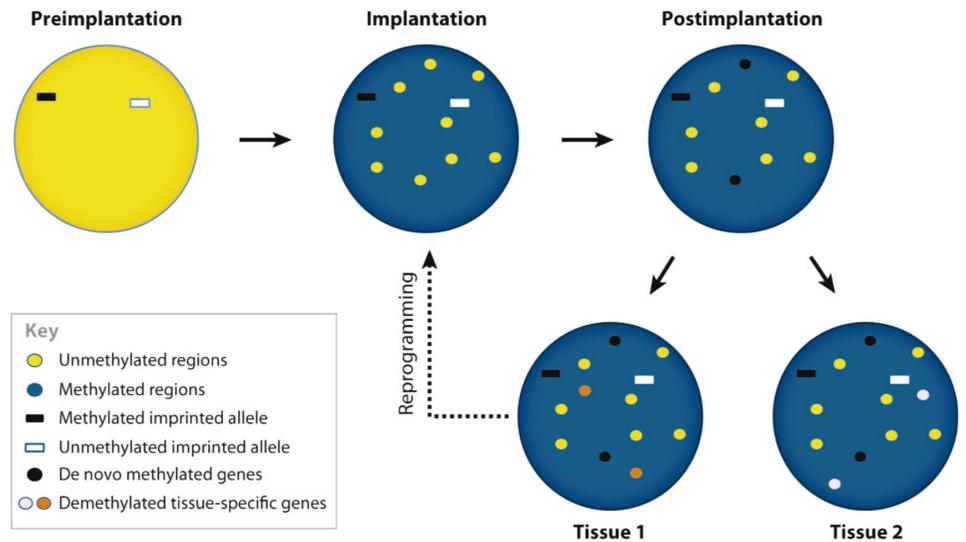


Fig. 4. Methylation induces closed chromatin structure.

Unmethylated (top) or methylated (bottom) DNA were transfected into cells and then tested by Southern blotting for DNaseI sensitivity in the nucleus. While the unmethylated DNA is digested by DNaseI at relatively low concentrations, the methylated DNA is relatively resistant, indicating it is in a more closed conformation. This shows that DNA methylation probably works by affecting chromatin structure. Adapted from (Keshet *et al.*, 1986).

Fig. 5. Programming 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. See (Cedar and Bergman, 2012).



turned out to be an excellent model for the implantation stage of embryogenesis. Using transfection, we demonstrated that all DNA inserted into these cells becomes de novo methylated, a result that is strikingly different than what happens in somatic cells where there is very little de novo activity. Strikingly, when CpG island DNA was inserted into these cells, no methylation took place, exactly mimicking what happens at the time of implantation (Frank *et al.*, 1991; Brandeis *et al.*, 1994).

In vivo, CpG islands are literally protected from the de novo methylation that takes place at the time of implantation, and this occurs without any need for demethylation. In contrast, ES cells have a unique ability to recognize and demethylate CpG islands in a reaction that apparently utilizes a defined molecular complex that contains all the components needed to remove methyl groups (Sabag *et al.*, 2014). It is very likely that this mechanism, which is not involved in the molding of CpG islands during normal development, actually plays an important role in the ability of ES cells to reprogram somatic nuclei by demethylating all those CpG islands that underwent de novo methylation during differentiation and

aging, thus resetting the global methylation pattern to its original bimodal state.

Cis acting sequences control DNA methylation

This ES-cell system also enabled us to evaluate what specific sequence elements are involved in molding the bimodal methylation profile at the time of implantation. To our surprise, it turned out that the high concentration of CpG is not the direct signal for protection of islands. Instead, it appears that it is the recognition of simple cis-acting sequences at these loci, and the transcription factor Sp1 binding-motif, for example, is sufficient to drive this reaction (Brandeis *et al.*, 1994). Indeed, even using this sequence alone it was possible to make any piece of DNA resistant to de novo methylation and this was then proven *in vivo* by its ability to control the methylation state in transgenic mice (Siegfried and Cedar, 1997). Using bioinformatics to analyze the sequence composition of all constitutively unmethylated CpG islands in the genome, it was possible to write an algorithm that predicts with high accuracy which regions of the DNA will remain unmodified at the time of implantation (Straussman *et al.*, 2009). This mathematical formulation clearly implies that the pattern of methylation is actually directed by sequence elements in the DNA.

Interestingly, an overall look at the sites protected from de novo methylation suggests that the selection of these regions is based on their association with RNA polymerase. Not only are these loci enriched for many transcription factor binding sites, but mapping experiments even indicate that these regions generally serve as gene promoters and literally all of them contain histones that are methylated at H3K4, another clear-cut sign of transcriptional potential. These results suggest that CpG islands are already marked at transcription start sites in the preimplantation embryo and that this serves as the mechanism for protection from de novo methylation at the time of implantation (Fig. 6). Although appearing counter-intuitive, these results actually demonstrate that

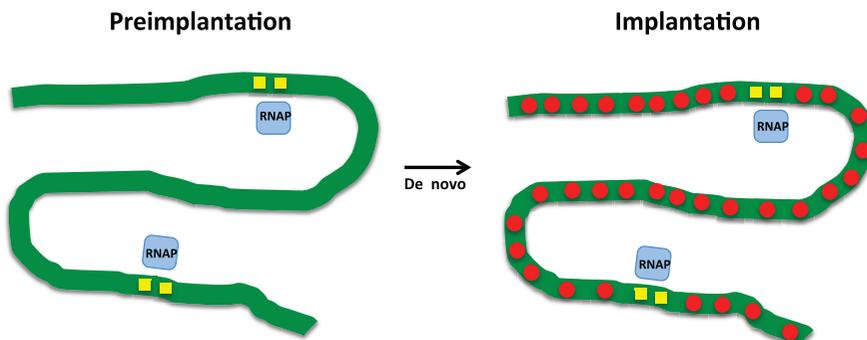


Fig. 6. Establishment of a bimodal methylation pattern. Preimplantation, the genome is largely unmethylated. Cis acting sequences (yellow squares) in CpG islands serve as binding sites for RNA polymerase. At the time of implantation there is a wave of de novo methylation (red circles) that modifies all CpG sites in the genome while regions marked with RNA polymerase binding are protected. This pattern is then maintained through every cell division even if RNA polymerase is no longer bound to the islands. According to this model, de novo DNA methylation does not bring about repression of active genes, but can serve to perpetuate the repression of genes originally inactive in the preimplantation embryo.

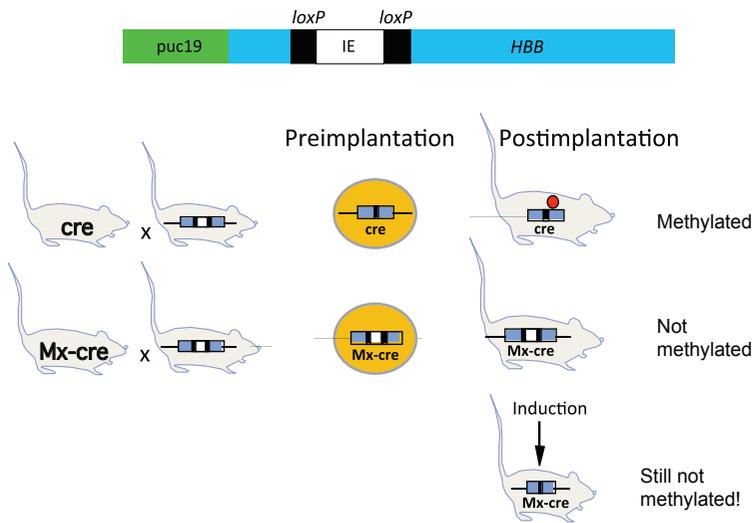


Fig. 7. Cis-acting sequences protect CpG islands. A human β globin (*HBB*) transgenic animal containing an island element (IE), which protects against *de novo* methylation, flanked by loxP sites was crossed with a mouse which expresses cre constitutively. Since the IE is removed preimplantation, the *HBB* sequence gets *de novo* methylated in vivo. When crossed with an inducible Cre (*Mx-Cre*), however, the IE is still present at implantation and this protects against *de novo* methylation. If *Mx-cre* is induced in the adult, the IE is removed, but this does not affect methylation. This experiment demonstrates that cis acting sequences control CpG island protection from *de novo* methylation and also shows that once the methylation pattern is formed at implantation, it is maintained, even in the absence of the original cis acting sequences (IE).

de novo methylation does not itself bring about repression, but rather serves to mark and then perpetuate the inactive state by making it difficult to turn on these genes in the soma.

By understanding the sequence code that controls DNA methylation patterns, it then became possible to finally test in a conclusive manner whether DNA methylation has a real function in vivo. To this end, we constructed a 120 bp island-inducing sequence element (IE) flanked by LoxP sites, embedded it within the human β -globin gene and generated a mouse transgene (Siegfried and Cedar, 1997). Using Cre to remove the “island sequence” prior to implantation, we showed that the transgene gets *de novo* methylated and remains that way in all tissues. In contrast, if the IE is not removed, the genomic region behaves like a CpG island and is protected from *de novo* methylation. In these mice, Cre-induced removal of the IE in the adult has no effect on DNA methylation. These studies clearly showed that the overall methylation pattern is indeed established as a one-time event in the early embryo and is then maintained autonomously even though the cis-acting sequences that drive this specificity at the time of implantation are no longer present (Fig. 7). When tissues from these mice were analyzed for gene expression, we found that the presence of DNA methylation caused anywhere from a 20-50 fold reduction of RNA synthesis, clearly substantiating the claim that DNA methylation

represses gene transcription in vivo. In addition, methylation also had a big effect on chromatin structure in vivo, mainly by affecting histone acetylation (Hashimshony *et al.*, 2003).

Genomic imprinting

Genomic imprinting is a phenomenon whereby some genes are expressed from only a single parental genome, either the maternal or paternal. Obviously, this implies that there must be a mechanism by which these sites remember where they come from, and it was hypothesized that these genes were differentially methylated in the gametes and then retained these patterns throughout development. Indeed, methylation analysis confirmed that this is indeed the case and through the use of methylase knockouts, it was shown that this modification is required for imprinting (Li *et al.*, 1993). We realized that in order for this model to work, these imprinting marks had to be preserved during early embryogenesis at a time when the entire genome methylation pattern was undergoing erasure (Fig. 5).

Using the sensitive PCR assays that we had developed, we demonstrated that these marks are established at the end of gametogenesis and are then maintained in a monoallelic manner in the pre-implantation embryo (Brandeis *et al.*, 1993; Kafri *et al.*, 1993; Stoger *et al.*, 1993). We also showed that these marks are

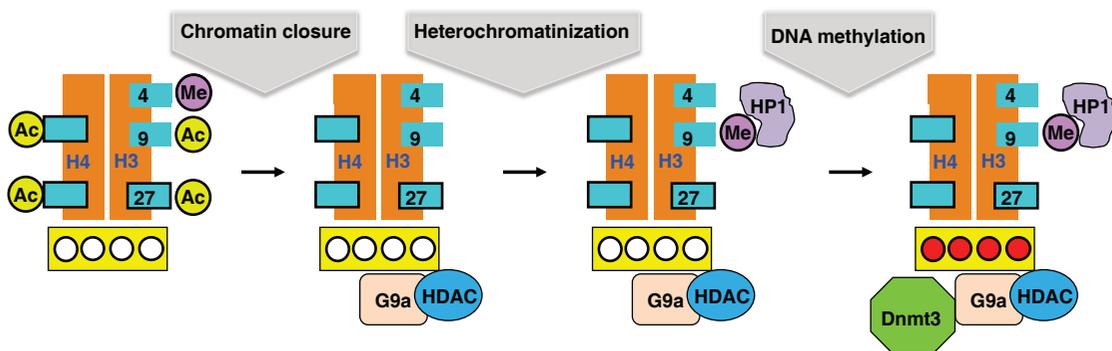


Fig. 8. Epigenetic inactivation of pluripotency genes. Pluripotency gene promoters are unmethylated (white ovals) and carry nucleosomes with acetylated histones and H3K4me3. Differentiation brings about gene repression in 3 steps. 1) transcriptional repression, 2) heterochromatinization through the action of G9a which recruits HDACs together with histone H3K4 demethylases and then methylates H3K9 which serves as a binding site for heterochromatin protein -1 (HP1) and 3) G9a recruits Dnmt3 which methylates (red ovals) the DNA. Adapted from Cedar and Bergman (2012).

erased during early gametogenesis in the next generation only to be reset in the mature gamete, thus completing the imprinting cycle. In order to better understand this process, we developed a zygote injection method for testing the maintenance mechanism, and, in this way, discovered that this protection required specific cis-acting control sequences (Birger *et al.*, 1999), and this ultimately led to the idea of an imprinting control region (ICR) which turns out to be the master regulator of the entire process (Shemer *et al.*, 2000; Perk *et al.*, 2002).

The basic bimodal methylation pattern

Putting all of these studies together, DNA methylation emerges as a global repression mechanism. Erasure in the early embryo serves to reprogram the epigenetic information present in the gametes and then, using sequence cues from the genome, a new methylation pattern is established which can subsequently be maintained in every cell of the body. This serves to set up an expression template which is programmed to allow housekeeping genes with CpG island promoters to be unmethylated and “open” while other regions of the genome and many tissue-specific regulatory elements are modified and “closed,” thus preventing their inadvertent expression in somatic cells.

In order to prove this point, we took a close look at the γ globin gene that is normally methylated in every cell of the body except for the fetal liver where it is expressed. Using targeting technology, we converted the promoter of this gene to a CpG island-like sequence, causing it to be unmethylated in every cell type (Goren *et al.*, 2006). Strikingly, this gene now undergoes abnormal transcription in almost every tissue. In adult β -globin expressing red

blood cells where this gene is usually methylated and inactive, the unmethylated γ globin gene is expressed at high levels, presumably aided by the rich abundance of globin-type transcription factors present in these cells. This example provides a much clearer picture for how global methylation operates *in vivo* to suppress genes by preventing inadvertent activation by a variety of different transcription factors present in somatic cells. It should be noted that the overall strategy used in animal cells differs markedly from the system present in simpler organisms where gene-specific repressors modulate gene activity. In the animal-cell system, a very large number of genes are silenced in each cell type and this is controlled by a global mechanism that does not require that each gene be recognized by a specific repressor.

Post implantation

In keeping with our goal of understanding the role DNA methylation plays during development, we next asked what happens following implantation, when the processes of differentiation and organogenesis actually begin. One of the first events that occur at this stage is that pluripotency genes, such as Oct4, Nanog and Sox2, get turned off and then remain silenced throughout life in every cell of the body. Using differentiating ES cells as a model system, we found that this silencing takes place in a three-step process (Feldman *et al.*, 2006). Initially, transcription is turned off by trans acting factors. This is followed by binding of the G9a protein that recruits histone deacetylase and then tri-methylates histone H3K9 that consequently binds HP1, thereby heterochromatinizing the gene. Finally, this same G9a recruits the DNA methylases, Dnmt3a and 3b, causing *de novo* methylation (Fig. 8).

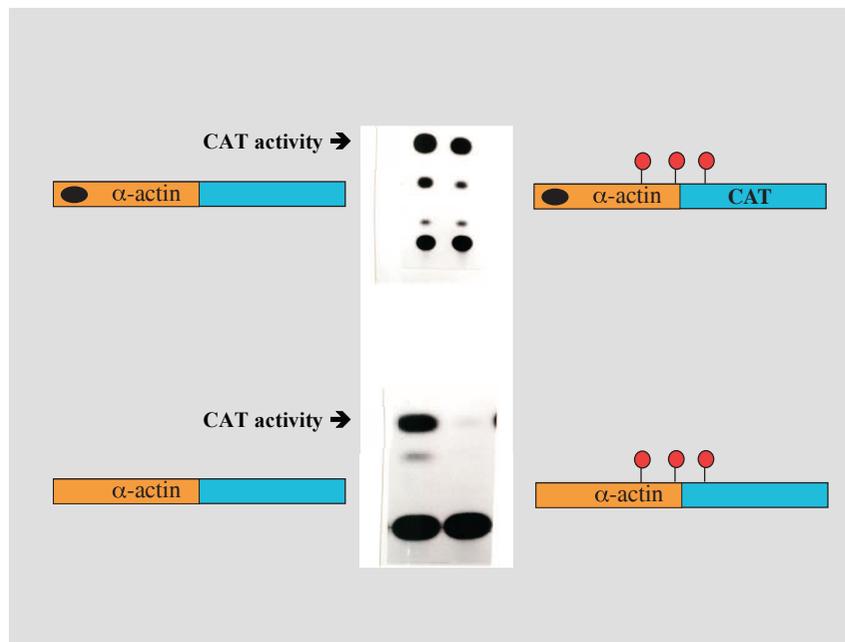


Fig. 9. Tissue-specific demethylation. An α -actin reporter construct transfected into myoblasts showed CAT activity (using chromatography to detect the product), whether it was methylated or not (top panel). This is because the gene has a regulatory region (black dot) which causes demethylation in myoblasts. If the regulatory region is deleted, the methylation cannot be removed and the gene remains inactive. This demonstrates that tissue-specific demethylation is controlled by cis acting regulatory sequences and that methylation must be removed for the gene to be active.

These experiments clearly show that methylation represents a late event in the inactivation of pluripotency genes, raising the question of what DNA methylation actually does. To address this issue, we used ES cells to test whether differentiated cells were capable of going back into a pluripotent state. Although WT cells were unable to do this, a large percentage of cells deleted for G9a could recover a pluripotent phenotype when returned to growth in LIF medium, and fine genetic analysis indicated that this was because these cells were not able to recruit the DNA methylases (Epsztejn-Litman *et al.*, 2008). It thus appears that *de novo* methylation does not initiate gene repression, but is necessary for stably keeping a gene in its silenced state.

The fact that G9a is able to recruit DNA methylases shows that there must be a physical link between these two epigenetic mechanisms, histone modification and DNA methylation. It turns out that this relationship is actually part of a general concept in epigenetics, the linking of repression mechanisms (see (Cedar and Bergman, 2009) for review). Thus, it has been shown, for example, that another H3K9 methylase, Suv39h is also able to recruit DNA methylase and this is responsible for DNA methylation of satellite sequences in mammalian cells. The enzyme Ezh2, which methylates H3K27 as part of the polycomb complex, serves

as a repressor of many developmental genes. It too has the ability to recruit DNA methylases, and it is this process that brings about aberrant *de novo* methylation in cancer (Schlesinger *et al.*, 2007). Finally, it should be noted that there is a negative relationship between the presence of the activating modification, H3K4me3 and DNA methylation which is based on the idea that this histone modification prevents the recruitment of DNA methylases as occurs during early development (Ooi *et al.*, 2007; Otani *et al.*, 2009). Taken together, these observations suggest that there is an evolutionary connection between histone and DNA-methylation that is built into the structure of the underlying proteins involved in these processes, thus generating a well-designed serial mechanism of repression.

Demethylation

With this perspective of how the basic methylation pattern is established during early embryogenesis, it became clear that the undermethylation of tissue-specific genes must have come about through a process of demethylation (Fig. 5), but the picture of how this occurs took a long time to decipher. The first hints came from transfection experiments into myoblasts (Yisraeli *et al.*, 1986). In previous studies, *in vitro*-methylated test genes inserted into fibroblasts were found to show inhibited transcription. When the α -actin gene was transfected into muscle cells in culture, however, it was equally active whether or not it had been previously methylated *in vitro*. Although puzzling at first, it soon became evident that this could be explained by the observation that this gene had actually undergone specific demethylation in these cells. In other words, the inserted exogenous copy of this gene must be recognized by the demethylation machinery in muscle cells and this is what allows it to be expressed, probably in the same manner that the endogenous gene is turned on during normal development. Using reverse genetics, we were able to identify the sequence elements required to direct demethylation and these were located within the 5' regulatory region of the α -actin gene (Paroush *et al.*, 1990) (Fig. 9). We later succeeded in deciphering similar elements in the Igk gene region (Lichtenstein *et al.*, 1994), and, in this case, we were also able to show that NF κ B may represent the trans-acting protein factor that activates this process (Kirillov *et al.*, 1996). It thus appears that tissue-specific demethylation takes place in a well-ordered preprogrammed manner, with unique elements being able to bind developmental factors designated to recruit the demethylation machinery.

The actual biochemical mechanism of demethylation has long been an enigma. Early workers had suggested that this may occur in a passive manner as a result of DNA replication, simply by blocking the maintenance reaction. While this type of mechanism may contribute to demethylation at early stages of development, many studies have shown that tissue-specific demethylation usually occurs in an active manner in the absence of DNA synthesis. This was first demonstrated for the δ -crystallin gene in the eye lens (Sullivan and Grainger, 1986), but was then confirmed in other systems. Using the α -actin gene transfected into rat myoblasts, we showed that demethylation even takes place on the original template molecules themselves, clearly indicating that the reaction must be active (Paroush *et al.*, 1990).

It has been very difficult to understand the demethylation reaction, but in retrospect, many facets of this process were already known, although not fully appreciated. Earlier experiments showed

quite convincingly that Friend Leukemia cells undergo extensive demethylation when these cells are induced to differentiate and using labeled substrates, it was found that the methylated C itself is replaced with a new (labeled) Cytosine in this process, suggesting that the reaction does not involve removal of the methyl groups from the base (Razin *et al.*, 1986). This concept was supported by experiments showing that glycosylases could carry out part of this reaction (Jost, 1993; Jost *et al.*, 1995). Finally, studies from our laboratory documented demethylation activity *in vitro* whereby 5mC is replaced by cytosine and, by using different labeled substrates, it was possible to show that this replacement involved more than one nucleotide (Weiss *et al.*, 1996; Swisher *et al.*, 1998). Despite these experiments, the full molecular picture of demethylation was still not clear.

The key to understanding demethylation came with the discovery of the Tet family of enzymes that can hydroxylate 5mC to produce 5hmC (Tahiliani *et al.*, 2009; Ito *et al.*, 2010). This conversion probably represents the first step in the demethylation pathway. According to this idea, the conversion of 5mC to 5hmC and further

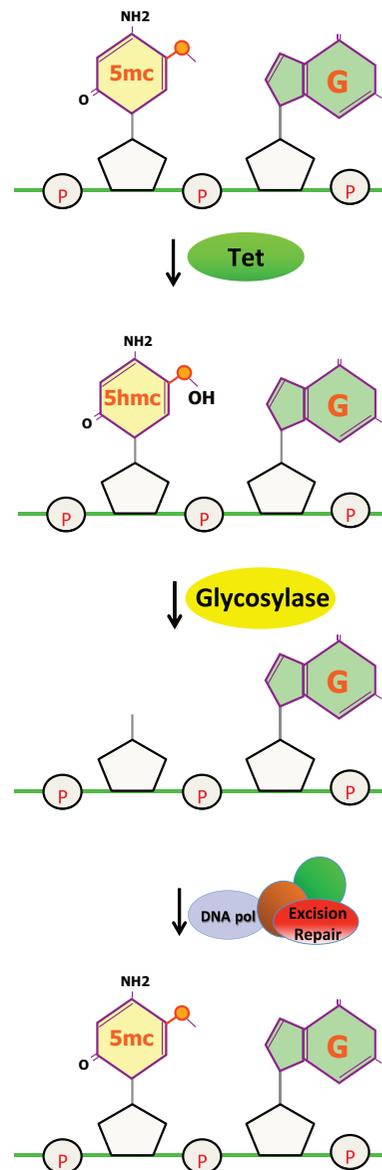


Fig. 10. DNA demethylation pathway. The methylated C residue in the CpG dinucleotide is hydroxylated by Tet proteins, generating 5hmC. This unusual base can then be removed by a glycosylase generating an apyrimidinic acid residue which may undergo nucleotide excision repair, thereby introducing an unmethylated C. 5hmC can also be diluted out by replication of the DNA.

oxidation products (5fmC and 5mC) makes this site into a preferred substrate for glycosylation enzymes that can recognize and remove these bases. This is then followed by a repair process that introduces an unmodified cytosine, thus completing the reaction (Wu and Zhang, 2014) (Fig. 10). The general principle appears to be that 5mC must be reconfigured to an “unusual” or mismatched base in order for it to be recognized as being “different” so that it can then be removed by the repair machinery or simply diluted out by DNA replication (Guo *et al.*, 2014). In a similar manner there is evidence that the deaminase, AID, may also be involved in demethylation by converting 5mC to T, which would then be subject to mismatch repair (Dominguez *et al.*, 2015).

Very recently, we were finally able to identify the key molecular components that carry out tissue-specific demethylation and thus provide genetic proof for a two-step pathway. We made a conditional Tet2/Tet3 mouse and generated animals specifically knocked out at the pro-B cell stage of lymphogenesis *in vivo*. During normal development, B cells go through a process whereby hundreds of regulatory regions undergo demethylation, and this was almost completely prevented in the knockout mouse, clearly indicating that Tet2 and Tet3 are necessary for this process (Orlanski *et al.*, 2016) (Fig. 11). As a result, transcription of many key genes regulated by these enhancers was significantly reduced, greatly inhibiting proper B-cell differentiation and strongly affecting B-cell function *in vivo*.

Ever since the initial observations indicating that DNA methylation is inversely correlated with gene expression, there has been much skepticism about the role of this modification, with many scientists claiming that all gene regulation is mediated exclusively by trans-acting factors and that DNA methylation does not really play any

role, other than serving to mark inactive genes (Ptashne, 2007). Although there have been a number of good artificial experiments showing that DNA methylation inhibits gene expression both *in vitro* and *in vivo*, definitive proof for this idea was lacking, mainly because there was no way to specifically inhibit demethylation without affecting other factors involved in the regulatory process. Since Tet enzymes appear to be only capable of carrying out the biochemical aspects of this process, the experiment described above using a Tet2/Tet3 knockout in B cells represents the first real proof for the role of methylation in gene regulation. Thus, even though all the proper trans-acting factors are still readily available during B-cell development in the knockout, programmed gene activation cannot take place, simply because the methyl groups are still present, somehow preventing proper transcription of adjacent genes.

DNA methylation in cancer

Cancer cells have an altered methylation pattern, with many regions showing hypomethylation while others, in particular CpG

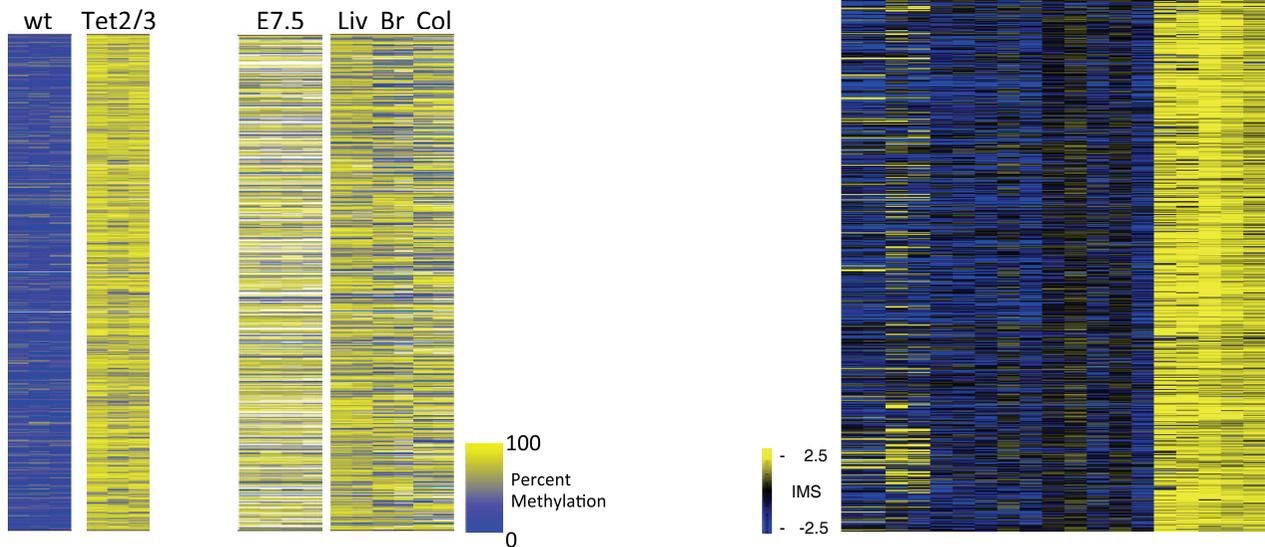


Fig. 11 (Left). Tet mediates tissue-specific methylation. The heat map shows methylation analysis (yellow, methylated; blue, unmethylated) of 1,400 tiles (100 bp) that undergo demethylation specifically in follicular B cells. These sequences are methylated in the embryo (E 7.5 d) right after implantation and remain methylated in all tissues (e.g. liver, brain, colon). They are hypo-methylated in wt B-cells. Knockout of Tet2 and Tet3 at an early stage of B-cell development prevents this demethylation. These tiles are mainly nearby enhancers of B-cell specific genes and preventing demethylation inhibits the activation of these genes. Adapted from (Orlanski *et al.*, 2016).

Fig. 12 (Right). Aberrant methylation in cancer. Heatmap showing methylation analysis (yellow, methylated; blue, unmethylated) of 2,000 CpG islands that are aberrantly methylated in colon cancer (CRC) as compared to normal tissues (ES cells, brain, liver, muscle, colon). Almost all of these islands are bound by the polycomb complex. Note that these islands are already highly methylated in colon polyps (Po), a benign tumor that precedes cancer. Adapted from (Nejman *et al.*, 2013).

islands, become de novo methylated compared to normal cells of the same tissue type (Baylin and Jones, 2011). Initially, it was thought that these changes represented an alternative pathway towards affecting “cancer genes,” with de novo methylation turning off tumor suppressors, while demethylation serving to turn on normally-silenced oncogenes. It turned out, however, that this represented an oversimplified view of the role of DNA methylation in cancer. Further studies have now clarified, to some extent, how methylation is altered in cancer and how this affects tumor biology.

As a first step in this direction, microarray technology revealed that a large number of CpG islands actually undergo de novo methylation, including many that are not associated with tumor suppressor genes (Keshet *et al.*, 2006) (Fig. 12). This observation implied that this de novo process was not driven by selection, but rather was taking place as a result of an instructive mechanism. Further studies demonstrated that almost all of these methylation events occur at CpG islands that are bound by polycomb (Schlesinger *et al.*, 2007). This complex is a repressor that works by virtue of the enzyme EZH2 to methylate histone H3 Lysine 27 on local nucleosomes, thereby bringing about a state of local heterochromatin. This represents a general system for silencing almost all the key developmental and differentiation genes in the cell from the time of early embryogenesis. This polycomb-mediated repression does not normally involve DNA methylation and occurs exclusively on promoter and regulatory sequences that have CpG islands and are therefore constitutively unmethylated.

In cancer, these sites appear to undergo de novo methylation (Nejman *et al.*, 2013). Although the trigger for this is not known, it is probably carried out by the enzyme EZH2 itself, which, in addition to its activity as a histone methylase, can also bind DNMT3, thus exposing these sites to DNA methylation (Vire *et al.*, 2006). Earlier studies had noted that many of the sites aberrantly methylated in tumors are found to be slightly methylated in normal tissues, as well, in a manner that appears to increase with age (Teschendorff *et al.*, 2010; Issa, 2014; Kulis *et al.*, 2015) and it turns out that these patterns are actually very similar in all cell types with the degree of methylation at each locus apparently controlled by the level of bound polycomb (Nejman *et al.*, 2013). In addition to de novo methylation at CpG islands, it was originally noted that tumors are also characterized by widespread hypomethylation and this evidently comes about as a result of demethylation at regions associated with the nuclear lamina covering over 40% of the genome (Berman *et al.*, 2012). These observations have now been elegantly confirmed by genome-wide studies showing the changes in methylation that take place in hematopoietic or skin stem cells during aging *in vivo* (Raddatz *et al.*, 2013; Sun *et al.*, 2014).

These wide-ranging experiments have put together an interesting picture of how DNA methylation may be involved in tumorigenesis. Firstly, it should be noted that the methylation pattern observed in tumor cells might already exist in the normal cell population. Analysis of individual DNA molecules in normal tissues carried out in our laboratory indicates that for any particular island, almost all molecules are completely unmethylated, with only a few being highly modified. This suggests that the aberrant methylation pattern may actually be present at relatively high levels, but only in select cells (Nejman *et al.*, 2013). We would propose that it is these cells that are prone to cancer.

According to the original concept that aberrant DNA methylation in cancer cells operates to inactivate tumor suppresser genes, it

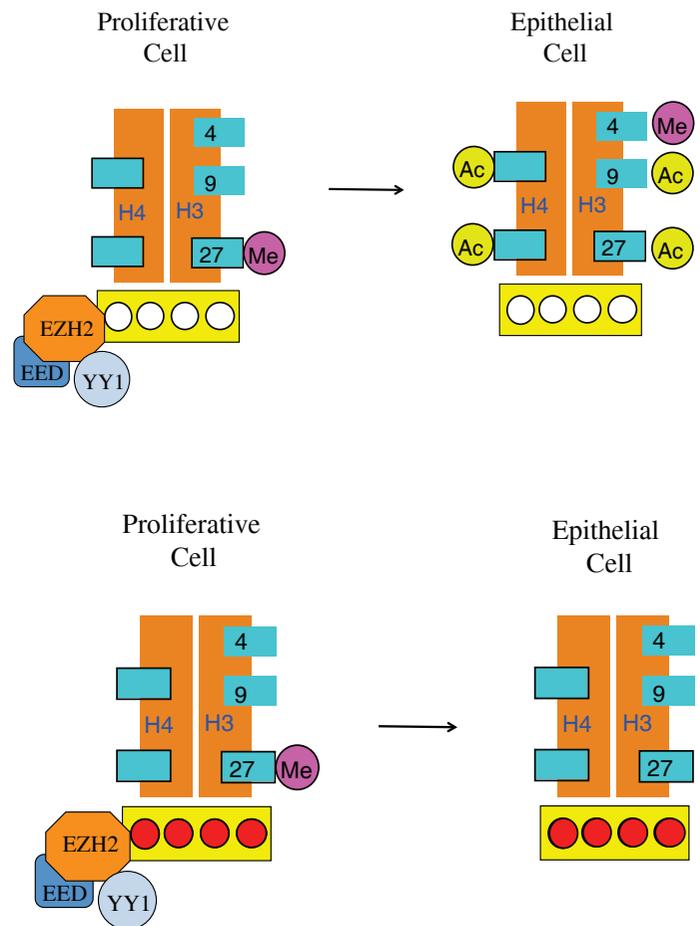


Fig. 13. Effect of DNA methylation in colon tumorigenesis. Normal colon epithelium is derived from numerous crypts which contain stem cells that constantly generate proliferative progenitor cells. These climb up the crypt and at some point undergo differentiation to mature epithelial cells. In normal proliferative cells of the colon (top), polycomb-target genes are repressed by the presence of a repressive complex, which contains the histone methyltransferase EZH2 and other proteins. Although the expression of these genes is inhibited, most of them have unmethylated (white circles) CpG island promoters and therefore maintain the potential for being activated (note acetylation of histones) during differentiation to epithelium. Cancer may initiate in stem cells in which polycomb target genes have already become de novo methylated (red circles) during aging. Although these genes are able to shed their polycomb complex as the cells advance up the crypt, they are prevented from being activated during differentiation because they are stably DNA methylated.

was assumed that the pattern of modification in the tumor brings about repression of genes that were initially active in the normal tissue. RNA analyses, however, demonstrated that this is not really the case. Rather, over 90% of the target genes were found to be inactive even prior to tumorigenesis (Keshet *et al.*, 2006), and this is in keeping with our findings that these genes are marked by polycomb. In light of this, we proposed that while mutations are essential to initiate the transformation process, the high level of aberrant methylation in target cells could be needed for the tumor to expand. This may happen because the presence of methyl groups on key developmental genes could prevent them from undergoing proper differentiation, thereby leaving them in a relatively prolifera-

tive state (Fig. 13). Alternatively, this pre-existing methylation may prevent the activation of many repair and apoptosis genes that the cell would normally use to prevent tumorigenesis.

This model is consistent with everything we know about the basic biology of DNA methylation. In every instance of site-specific de novo methylation, the initial repression itself always takes place through the involvement of repression factors, with DNA methylation coming as a secondary mechanism to provide long-term silencing stability. Indeed, even the initial global methylation at the time of implantation only takes place at sites that are already inactive. Thus, while DNA methylation, itself, cannot directly cause gene repression once this modification occurs it becomes very difficult to activate the underlying gene without removing the methyl groups.

Postnatal changes in DNA methylation

While it is well accepted that DNA methylation patterns are established during development, one of the major questions in epigenetics is whether further changes can occur postnatally and whether these may be brought about through the influence of environment. As a first step in approaching this problem, we asked whether there may be differences in methylation between males and females that appear post adolescence when sex hormones beginning to be secreted (Reizel *et al.*, 2015). We used a genome-wide assay (RRBS) to analyze DNA from mouse liver at various times after birth into adulthood. In a striking manner, we observed hundreds of differences in DNA methylation with almost all of them being hypomethylated in the male. These differentially methylated regions (DMRs) are formed starting at about three weeks of age by a process of demethylation and are programmed to occur at enhancer-like sequences associated with liver genes

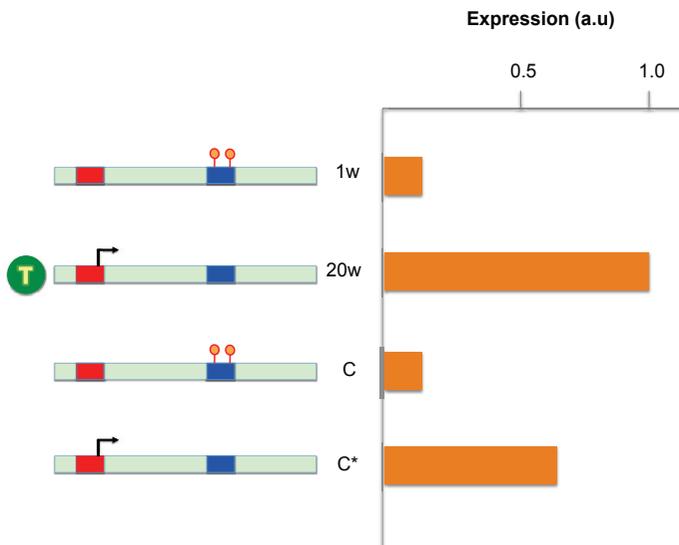


Fig. 14. Testosterone activates genes through demethylation. A large number of gene promoters (red) are controlled by enhancers (blue) in the liver. These genes have low expression profiles at one week of age, but get promoter (red) activated postnatally and are highly expressed at 20 wk of age after testosterone brings about demethylation at the enhancer. In castrated animals (C), there is no testosterone, so the gene does not get activated. Castration late in life removes the testosterone after the enhancer has already undergone demethylation, but the liver retains a memory (undermethylation) and therefore expresses this gene at a slightly lower level.

specifically activated in the male, but not the female. This entire process appears to occur as a result of testosterone secretion and in its absence male specific gene expression is not seen (Fig. 14). Furthermore, treatment of females with this hormone can also induce these changes. These results indicate that postnatal methylation changes can occur in a programmed manner directed by trans acting protein factors.

Recent studies in our laboratory have demonstrated that postnatal demethylation is very prevalent in the mouse and in humans, with thousands of enhancer-like sequences becoming hypomethylated in a variety of different tissue types during the first few weeks of life and these have a wide-ranging effect on specific gene expression in these tissues. Preliminary experiments using a Tet2/Tet3 knockout approach show that these changes are essential for normal expression. It is clear that these demethylation events must be driven by the appearance of hormone-like substances, growth factors or metabolites. Many of these alterations occur in the brain. It thus appears that tissues of the body undergo adaptive changes presumably in response to internal influences. In the same manner, it seems likely that outside influences from the environment could trigger the release of internal hormonal-like substances that could then set in motion a programmed process of methylation changes that may first have an immediate effect on tissue metabolism but then remain this way in a stable manner. Any changes in methylation that occur would undoubtedly be “remembered” as a result of the ability of DNA methylation to be maintained for long periods of time even after the original inducer is no longer present.

Conclusions

When we first started working on DNA methylation, the main focus of molecular biology was to decipher the genetic text and understand how this book of instructions is used to generate an organism and navigate its biology. As the methylation story emerged, it has become clear that this epigenetic marker plays a key role in controlling how the DNA text is actually used and interpreted, especially in long-living eukaryotic organisms like man. In a sense, we have learned that the DNA, much like other languages, has a system of annotation, with methylation being the main player. Our studies have also revealed a brand new regulatory mechanism. Unlike classical repression brought about by trans acting factors that are plastic in nature, DNA methylation produces a permanent state of gene silencing which can be maintained for many years, even in dividing cells. This same modification can also serve as a memory mechanism for propagating decisions made at earlier stages of development.

This research has opened up exciting new vistas. Understanding how methylation works in development has many implications for health and disease. Fragile X syndrome, for example, is a developmental disease in which a single gene is abnormally turned off by DNA methylation. Because the gene itself is intact, however, there is a distinct possibility that this disease can actually be prevented or reversed by demethylation agents such as 5azaC. Similarly, 5azaC may be able to derepress the γ -globin gene in red blood cells of patients with β -Thalassemia, thereby bypassing the defect. Our knowledge of the role of DNA methylation in cancer has spurred clinical research aimed at using demethylating agents to treat a variety of different tumors, and this approach may even provide

a new route to prevention (Laird *et al.*, 1995). Finally, it should be noted that DNA methylation can serve as a potent diagnostic tool as well. Recent studies done in collaboration with Yuval Dor and Ruth Shemer (Lehmann-Werman *et al.*, 2016) show that DNA methylation profiles can be used to identify the tissue origin of free circulating DNA in plasma, thus providing an exciting new way to detect cell death derived from specific cell types in the body. Using this approach, it may now be possible to employ periodic blood tests to pick up signs of disease, leading the way to the early diagnosis of Type I diabetes, multiple sclerosis, myocardial infarction, sports-induced chronic traumatic encephalopathy (CTE), and other pathologic processes characterized by cell death.

The main challenge of DNA methylation research in the future is to understand the possible role of DNA methylation as a memory mechanism for the effect of environment on gene expression. There is strong circumstantial evidence that what we eat, exposure to different surroundings and life-changing events may all bring about long-term changes in body function, and there is strong suspicion that this may be mediated by DNA methylation. Even more exciting is the possibility that environmentally-induced epigenetic changes that occur in the parent may be “inherited” by the offspring as a shortcut to evolution, and here too, this might involve some form of DNA-methylation-derived memory.

References

- AXEL R, CEDAR H, FELSENFELD G (1973). The synthesis of globin RNA from duck reticulocyte chromatin *in vitro*. *Proc Natl Acad Sci USA* 70: 2029-2032.
- BAYLIN S B, JONES P A (2011). A decade of exploring the cancer epigenome biological and translational implications. *Nature Rev Cancer* 11: 726-734.
- BERMAN B P, WEISENBERGER D J, AMAN J F, HINOUE T, RAMJAN Z, LIU Y, NOUSHMEHR H, LANGE C P, VAN DIJK C M, TOLLENAAR R A *et al.*, (2012). Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nature Genet* 44: 40-46.
- BIRGER Y, SHEMER R, PERK J, RAZIN A (1999). The imprinting box of the mouse *Igf2r* gene. *Nature* 397: 84-88.
- BOSTICK M, KIM J K, ESTEVE P O, CLARK A, PRADHAN S, JACOBSEN S E (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317: 1760-1764.
- BRANDEIS M, FRANK D, KESHET I, SIEGFRIED Z, MENDELSON M, NEMES A, TEMPER V, RAZIN A, CEDAR H (1994). Sp1 elements protect a CpG island from de novo methylation. *Nature* 371: 435-438.
- BRANDEIS M, KAFRI T, ARIEL M, CHAILLET J R, MCCARREY J, RAZIN A, CEDAR H (1993). The ontogeny of allele-specific methylation associated with imprinted genes in the mouse. *EMBO J* 12: 3669-3677.
- CEDAR H, BERGMAN Y (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nature Rev Genet* 10: 295-304.
- CEDAR H, BERGMAN Y (2012). Programming of DNA Methylation Patterns. *Annu Rev Biochem* 81: 97-117.
- CEDAR H, FELSENFELD G (1973). Transcription of chromatin *in vitro*. *J Mol Biol* 77: 237-254.
- 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: 2125-2132.
- DOMINGUEZ P M, TEATER M, CHAMBWE N, KORMAKSSON M, REDMOND D, ISHII J, VUONG B, CHAUDHURI J, MELNICK A, VASANTHAKUMAR A *et al.*, (2015). DNA Methylation Dynamics of Germinal Center B Cells Are Mediated by AID. *Cell reports* 12: 2086-2098.
- EDEN S, HASHIMSHONY T, KESHET I, THORNE A W, CEDAR H (1998). DNA methylation models histone acetylation. *Nature* 394: 842-843.
- EPSZTEJN-LITMAN S, FELDMAN N, ABU-REMAILEH M, SHUFARO Y, GERSON A, UEDA J, DEPLUS R, FUKS F, SHINKAI Y, CEDAR H *et al.*, (2008). De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nature Struct Mol Biol* 15: 1176-1183.
- 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. *Nature Cell Biol* 8: 188-194.
- FELSENFELD G (1978). Chromatin. *Nature* 271: 115-122.
- FRANK D, KESHET I, SHANI M, LEVINE A, RAZIN A, CEDAR H (1991). Demethylation of CpG islands in embryonic cells. *Nature* 351: 239-241.
- GOREN A, SIMCHEN G, FIBACH E, SZABO P E, TANIMOTO K, CHAKALOVA L, PFEIFER G P, FRASER P J, ENGEL J D, CEDAR H (2006). Fine Tuning of Globin Gene Expression by DNA Methylation. *PLoS ONE* 1: e46.
- GRUENBAUM Y, CEDAR H, RAZIN A (1982). Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature* 295: 620-622.
- GRUENBAUM Y, NAVEH-MANY T, CEDAR H, RAZIN A (1981). Sequence specificity of methylation in higher plant DNA. *Nature* 292: 860-862.
- GUOF, LI X, LIANG D, LI T, ZHUP, GUOH, WU X, WENL, GU TP, HU B *et al.*, (2014). Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell* 15: 447-458.
- HASHIMSHONY T, ZHANG J, KESHET I, BUSTIN M, CEDAR H (2003). The role of DNA methylation in setting up chromatin structure during development. *Nature Genet* 34: 187-192.
- ISSA J P (2014). Aging and epigenetic drift: a vicious cycle. *J Clin Invest* 124: 24-29.
- ITO S, D'ALESSIO A C, TARANOVA O V, HONG K, SOWERS L C, ZHANG Y (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466: 1129-1133.
- JOST J P (1993). Nuclear extracts of chicken embryos promote an active demethylation of DNA by excision repair of 5-methyldeoxycytidine. *Proc Natl Acad Sci USA* 90: 4684-4688.
- JOST J P, SIEGMANN M, SUN L, LEUNG R (1995). Mechanism of DNA demethylation in chicken embryos. *J Biol Chem* 270: 9734-9739.
- KAFRI T, ARIEL M, BRANDEIS M, SHEMER R, URVEN L, MCCARREY J, CEDAR H, RAZIN A (1992). Developmental pattern of gene-specific DNA methylation in the mouse embryo and germline. *Genes Dev* 6: 705-714.
- KAFRI T, GAO X, RAZIN A (1993). Mechanistic aspects of genome-wide demethylation in the preimplantation mouse embryo. *Proc Natl Acad Sci USA* 90: 10558-10562.
- KESHET I, LIEMAN-HURWITZ J, CEDAR H (1986). DNA methylation affects the formation of active chromatin. *Cell* 44: 535-543.
- KESHET I, SCHLESINGER Y, FARKASH S, RANDE, HECHT M, SEGAL E, PIKARSKI E, YOUNG R A, NIVELEAU A, CEDAR H *et al.*, (2006). Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nature Genet* 38: 149-153.
- KIRILLOV A, KISTLER B, MOSTOSLAVSKY R, CEDAR H, WIRTH T, BERGMAN Y (1996). A role for nuclear NF- κ B in B-cell-specific demethylation of the Igk locus. *Nature Genet* 13: 435-441.
- KULIS M, MERKEL A, HEATH S, QUEIROSO A C, SCHUYLER R P, CASTELLANO G, BEEKMAN R, RAINERI E, ESTEVE A, CLOT G *et al.*, (2015). Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nature Genet* 47: 746-756.
- LAIRD P W, JACKSON-GRUSBY L, FAZELI A, DICKINSON S L, JUNG W E, LI E, WEINBERG R A, JAENISCH R (1995). Suppression of intestinal neoplasia by DNA hypomethylation. *Cell* 81: 197-205.
- LEHMANN-WERMAN R, NEIMAN D, ZEMMOUR H, MOSS J, MAGENHEIM J, VAKNINDEMBINSKY A, RUBERTSSON S, NELLGARD B, BLENNOW K, ZETTERBERG H *et al.*, (2016). Identification of tissue-specific cell death using methylation patterns of circulating DNA. *Proc Natl Acad Sci USA* 113: E1826-1834.
- LI E, BEARD C, JAENISCH R (1993). Role for DNA methylation in genomic imprinting. *Nature* 366: 362-365.
- LICHTENSTEIN M, KEINI G, CEDAR H, BERGMAN Y (1994). B-cell specific demethylation: a novel role for the intronic k-chain enhancer sequence. *Cell* 76: 913-923.
- 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: 371-382.
- NAVEH-MANY T, CEDAR H (1981). Active gene sequences are undermethylated. *Proc Natl Acad Sci USA* 78: 4246-4250.
- NEJMAN D, STRAUSSMAN R, STEINFELD I, RUVOLO M, ROBERTS D, YAKHINI

- Z, CEDAR H (2013). Molecular rules governing de novo methylation in cancer. *Cancer Res* 74: 1475-1483.
- OOI S K, QIU C, BERNSTEIN E, LI K, JIA D, YANG Z, ERDJUMENT-BROMAGE H, TEMPST P, LIN S P, ALLIS C D *et al.*, (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* 448: 714-717.
- ORLANSKI S, LABI V, REIZEL Y, SPIRO A, LICHTENSTEIN M, LEVIN-KLEIN R, KORALOV S, SKVERSKY Y, RJEWSKY K, CEDAR H *et al.*, (2016). Tissue-specific DNA demethylation is required for proper B-cell differentiation and function. *Proc Natl Acad Sci USA* 113: 5018-5023.
- OTANI J, NANKUMO T, ARITA K, INAMOTO S, ARIYOSHI M, SHIRAKAWA M (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A/ATRX-DNMT3-DNMT3L domain. *EMBO Rep* 10: 1235-1241.
- PANET A, CEDAR H (1977). Selective degradation of integrated murine leukemia proviral DNA by deoxyribonucleases. *Cell* 11: 933-940.
- PAROUSH Z, KESHET I, YISRAELI J, CEDAR H (1990). Dynamics of demethylation and activation of the α actin gene in myoblasts. *Cell* 63: 1229-1237.
- PERK J, MAKEDONSKI K, LANDE L, CEDAR H, RAZIN A, SHEMER R (2002). The imprinting mechanism of the Prader-Willi/Angelman regional control center. *EMBO J* 21: 5807-5814.
- POLLACK Y, STEIN R, RAZIN A, CEDAR H (1980). Methylation of foreign DNA sequences in eukaryotic cells. *Proc Natl Acad Sci USA* 77: 6463-6467.
- PTASHNE M (2007). On the use of the word 'epigenetic'. *Curr Biol* 17: R233R236.
- RADDATZ G, HAGEMANN S, ARAN D, SOHLE J, KULKARNI P P, KADERALI L, HELLMAN A, WINNEFELD M, LYKO F (2013). Aging is associated with highly defined epigenetic changes in the human epidermis. *Epigenetics Chromatin* 6: 36.
- RAZIN A, CEDAR H (1977). Distribution of 5-methylcytosine in chromatin. *Proc Natl Acad Sci USA* 74: 2725-2728.
- RAZINA, RIGGSAD (1980). DNA methylation and gene function. *Science* 210: 604-610.
- RAZINA, SZYF M, KAFRI T, ROLL M, GILOH H, SCARPA S, CAROTTI D, CANTONI G L (1986). Replacement or 5-methylcytosine by cytosine: a possible mechanism for transient DNA demethylation during differentiation. *Proc Natl Acad Sci USA* 83: 2827-2831.
- REIZEL Y, SPIRO A, SABAG O, SKVERSKY Y, HECHT M, KESHET I, BERMAN B-P, CEDAR H (2015). Gender-specific postnatal demethylation and establishment of epigenetic memory. *Genes Dev* 29: 923-933.
- SABAG O, ZAMIR A, KESHET I, HECHT M, LUDWIG G, TABIB A, MOSS J, CEDAR H (2014). Establishment of methylation patterns in ES cells. *Nature Struct Mol Biol* 21: 110-112.
- SCHLESINGER Y, STRAUSSMAN R, KESHET I, FARKASH S, HECHT M, ZIMMERMAN J, EDEN E, YAKHINI Z, BEN-SHUSHAN E, REUBINOFF B E *et al.*, (2007). Polycomb mediated histone H3(K27) methylation pre-marks genes for de novo methylation in cancer. *Nature Genet* 39: 232-236.
- SHEMER R, HERSHKO A Y, PERK J, MOSTOSLAVSKY R, TSUBERI B-Z, CEDAR H, BUITING K, RAZIN A (2000). The imprinting box of the Prader-Willi/Angelman Syndrome domain. *Nature Genet* 26: 440-443.
- SIEGFRIED Z, CEDAR H (1997). DNA methylation: A molecular lock. *Curr Biol* 7: 305-307.
- SMITH Z D, CHAN M M, HUMM K C, KARNIK R, MEKHOUBAD S, REGEVA, EGGAN K, MEISSNER A (2014). DNA methylation dynamics of the human preimplantation embryo. *Nature* 511: 611-615.
- SMITH Z D, CHAN M M, MIKKELSEN T S, GU H, GNIIRKE A, REGEVA, MEISSNER A (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* 484: 339-344.
- SOLAGE A, CEDAR H (1978). Organization of 5-methylcytosine in chromosomal DNA. *Biochemistry* 17: 2934-2938.
- STEIN R, GRUENBAUM Y, POLLACK Y, RAZIN A, CEDAR H (1982a). Clonal inheritance of the pattern DNA methylation in mouse cells. *Proc Natl Acad Sci USA* 79: 61-65.
- STEIN R, RAZINA, CEDAR H (1982b). *In vitro* methylation of the hamster APRT gene inhibits its expression in mouse L-cells. *Proc Natl Acad Sci USA* 79: 3418-3422.
- STOGER R, KUBICKA P, LIU C-G, KAFRI T, RAZIN A, CEDAR H, BARLOW D P (1993). Maternal-specific methylation of the imprinted mouse Igf2r locus identifies the expressed locus as carrying the imprinting signal. *Cell* 73: 61-71.
- STRAUSSMAN R, NEJMAND, ROBERTS D, STEINFELD I, BLUM B, BENVENISTY N, SIMON I, YAKHINI Z, CEDAR H (2009). Developmental programming of CpG island methylation profiles in the human genome. *Nature Struct Mol Biol* 16: 564-571.
- SULLIVAN C H, GRAINGER R M (1986). d-Crystallin genes become hypomethylated in postmitotic lens cells during chicken development. *Proc Natl Acad Sci USA* 83: 329-333.
- SUN D, LUO M, JEONG M, RODRIGUEZ B, XIA Z, HANNAH R, WANG H, LE T, FAULL K F, CHEN R *et al.*, (2014). Epigenomic Profiling of Young and Aged HSCs Reveals Concerted Changes during Aging that Reinforce Self-Renewal. *Cell Stem Cell* 14: 673-688.
- SWISHER J F, RAND E, CEDAR H, PYLE A M (1998). Analysis of putative RNase sensitivity and protease insensitivity of demethylation activity in extracts from rat myoblasts. *Nucleic Acids Res* 26: 5573-5580.
- TAHLILANI M, KOH K P, SHEN Y, PASTOR W A, BANDUKWALA H, BRUDNO Y, AGARWAL S, IYER L M, LIU D R, ARAVIND L *et al.*, (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324: 930-935.
- TESCHENDORFF A E, MENON U, GENTRY-MAHARAJ A, RAMUS S J, WEISENBERGER D J, SHEN H, CAMPAN M, NOUSHMEHR H, BELL C G, MAXWELL A P *et al.*, (2010). Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res* 20: 440-446.
- VAN DER PLOEG L H T, FLAVELL R A (1980). DNA methylation in the human gamma delta beta-globin locus in erythroid and nonerythroid tissues. *Cell* 19: 947-958.
- VIRE E, BRENNER C, DEPLUS R, BLANCHON L, FRAGA M, DIDELOT C, MOREY L, VAN EYNDE A, BERNARD D, VANDERWINDEN J M *et al.*, (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439: 871-874.
- WEINTRAUB H, GROUDINE M (1976). Chromosomal subunits in active genes have an altered conformation. *Science* 193: 848-856.
- WEISS A, KESHET I, RAZIN A, CEDAR H (1996). DNA demethylation in vitro: Involvement of RNA. *Cell* 86: 709-718.
- WIGLER M, SILVERSTEIN S, LEE L S, PELLICER A, CHENG Y, AXEL R (1977). Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. *Cell* 11: 223-232.
- WU H, ZHANG Y (2014). Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell* 156: 45-68.
- YISRAELI J, ADELSTEIN R S, MELLOUL D, NUDEL U, YAFFE D, CEDAR H (1986). Muscle-specific activation of a methylated chimeric actin gene. *Cell* 46: 409-416.
- YISRAELI J, FRANK D, RAZIN A, CEDAR H (1988). Effect of *in vitro* DNA methylation on beta globin gene expression. *Proc Natl Acad Sci USA* 85: 4638-4642.

Further Related Reading, published previously in the *Int. J. Dev. Biol.*

Epigenetics and imprinting in human disease

Jennifer M. Kalish, Connie Jiang and Marisa S. Bartolomei
Int. J. Dev. Biol. (2014) 58: 291-298

DNA methylation and its role in the trophoblast cell lineage

Satoshi Tanaka, Momo O. Nakanishi and Kunio Shiota
Int. J. Dev. Biol. (2014) 58: 231-238

DNA methylation establishment during oocyte growth: mechanisms and significance

Shin-Ichi Tomizawa, Joanna Nowacka-Woszek and Gavin Kelsey
Int. J. Dev. Biol. (2012) 56: 867-875

DNA methylation reprogramming and DNA repair in the mouse zygote

Konstantin Lepikhov, Mark Wossidlo, Julia Arand and Jörn Walter
Int. J. Dev. Biol. (2010) 54: 1565-1574

Epigenetic asymmetry in the zygote and mammalian development

Robert Feil
Int. J. Dev. Biol. (2009) 53: 191-201

