

# DNA Methylation, Nuclear Structure, Gene Expression and Cancer

Heinrich Leonhardt\* and M. Cristina Cardoso

Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany

---

**Abstract** DNA methylation, chromatin structure, transcription, and cancer have traditionally been studied as separate phenomena. Recent data provide now direct physical and functional links between these processes revealing a complex network of interactions and mutual dependences. Methylated DNA is bound by methyl-CpG binding protein (MeCP) complexes that include histone deacetylases (HDACs). This recruitment of HDACs is suggested to promote local chromatin condensation and thereby repress gene expression. Most recently, also complexes of DNA methyltransferase (Dnmt1) with transcriptional repressors, DMAP1 and pRB, have been described providing a direct link to transcriptional regulation and tumor suppression. Inactivation of the DNA methyltransferase genes (Dnmt1, 3a, and 3b) was found to be lethal in mice and several human diseases (ICF and Rett syndrome) turned out to be linked to DNA methylation. In particular, global hypomethylation has been found in tumor samples together with cancer-type-specific, local hypermethylation. Taken together, these lines of evidence clearly underscore the central role of DNA methylation in the regulation of gene expression and chromatin structure during normal development and diseases like cancer. *J. Cell. Biochem. Suppl.* 35: 78–83, 2000. © 2001 Wiley-Liss, Inc.

**Key words:** chromatin structure; DNA replication; proliferating cell nuclear antigen (PCNA); MeCP2; Transcription; histone deacetylation; histone deacetylase (HDAC); transcriptional silencing; CpG island; DNA methyltransferase (Dnmt)

---

The mammalian genome contains in addition to the well known four bases also methylated cytosine residues. This fifth base is generated by a post-replicative modification by DNA methyltransferases that occurs predominantly at CpG sites. Both strands are symmetrically methylated at these CpG sites but only about 50–70% of all CpG sites are methylated. The pattern of methylated and unmethylated sites is precisely maintained over many cell division cycles, however, at specific stages during development and disease the methylation pattern changes. Until now three DNA methyltransferases, Dnmt1, Dnmt3a, and Dnmt3b, have been identified in mammalian cells [Bestor et al., 1988; Okano et al., 1999]. A fourth gene, Dnmt2, shows the conserved sequence motifs of DNA methyltransferases but neither enzyme activity nor function could be determined until now [Okano et al., 1998].

The by far most abundant and ubiquitous enzyme, Dnmt1, shows a preference for hemimethylated DNA suggesting a role in the maintenance of a given DNA methylation pattern after each round of DNA replication by specifically methylating cytosine residues at the newly synthesized strand [Bestor et al., 1988]. Indeed, Dnmt1 was found to be directly associated with the replication machinery indicating that the epigenetic information, meaning the DNA methylation pattern, is replicated along with the genetic information [Leonhardt et al., 1992].

The other two enzymes, Dnmt3a and Dnmt3b, are tissue specific and were shown to be involved in de novo methylation [Okano et al., 1999]. In addition to these methyltransferases there is now also growing evidence for demethylation in mammalian cells [Wolffe et al., 1999].

Beyond the identification of these enzymatic activities very little is known about the regulation of DNA methylation during development and about the changes occurring in diseases. However, it was clearly demonstrated by targeted disruption of Dnmt genes in mice that DNA methylation is crucial for mammalian

---

\*Correspondence to: Heinrich Leonhardt, Franz Volhard Clinic, Wiltbergstr. 50, 13125 Berlin, Germany. E-mail: leonhardt@fvk-berlin.de

Received 12 January 2001; Accepted 16 January 2001

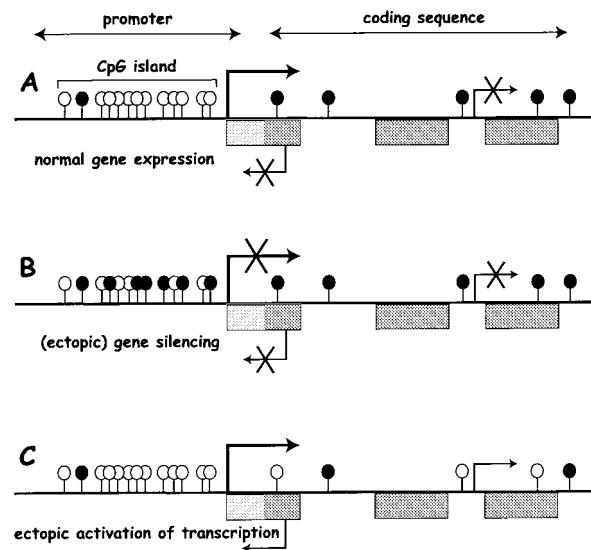
© 2001 Wiley-Liss, Inc.

development. Dnmt1 and Dnmt3b deficient embryos die around mid-gestation and Dnmt3a deficient ones die soon after birth [Li et al., 1992; Okano et al., 1999]. With this genetic approach it was furthermore shown that DNA methylation is required for X-inactivation and parental imprinting [Li et al., 1993]. Later, mutations in the human Dnmt3b were linked to the ICF syndrome [Okano et al., 1999] and mutations in the MeCP2 gene to the Rett syndrome [Amir et al., 1999] indicating a direct connection between DNA methylation and human disease.

### DNA METHYLATION, CHROMATIN STRUCTURE AND GENE EXPRESSION

Soon after methylated cytosine residues were mapped in mammalian genomes changes in the methylation pattern were observed during cellular differentiation and were correlated with changes in gene expression [reviewed in Doerfler, 1983]. In general, actively transcribed genes were found to be hypomethylated, while transcriptionally inactive sequences are usually highly methylated. This mere correlation, though reproduced with countless genes, stirred up a chicken- and -egg discussion about what comes first, whether DNA methylation is only a consequence of transcriptional inactivation or may actually play an active role and cause transcriptional silencing and whether DNA methylation is required to begin with. Thus the dead Dnmt1 deficient embryos came as a relief, as they unequivocally demonstrated the crucial role of DNA methylation during mammalian development [Li et al., 1992].

Methylated and unmethylated CpG sites are not randomly distributed in mammalian genomes. CpG sites are clustered in the promoter region of all housekeeping genes and of about 40% of the tissue-specific genes while they are underrepresented in the coding part of genes [Bird, 1986]. The few CpG sites in the coding part of genes are usually methylated, while the clusters of CpG sites, the CpG islands, in the promoter region are unmethylated in actively transcribed genes (see Fig. 1A). DNA methylation outside of promoter regions probably contributes to the silencing of cryptic transcriptional start sites. As mentioned above, transcriptional inactivation usually correlates with methylation of the CpG islands in the promoter region. This is a normal event during cellular differentiation but ectopic gene silencing may



**Fig. 1.** DNA methylation and transcriptional regulation. This scheme summarizes the typical role of DNA methylation in normal gene expression and compares it with detrimental consequences of methylation errors. The transcriptional start site and the promoter of all house keeping genes and about 40 % of all tissue-specific genes fall in or near a CpG island. The typical elements of vertebrate genes are outlined: exons (grey rectangles and open reading frame in dark grey), methylated cytosine residues (●), unmethylated sites (○), transcription start site and direction (bent arrows). **A:** CpG islands in promoters of actively expressed genes are usually unmethylated, while the rest of the gene containing the coding sequence contains less but mostly methylated CpG sites. This distribution of methylated and unmethylated sites promotes transcription initiation at proper sites and suppresses initiation at random, cryptic sites throughout the genome. **B:** Local hypermethylation at CpG islands may cause gene silencing. **C:** Alternatively, general hypomethylation may activate cryptic transcriptional start sites generating anti-sense and truncated mRNA with potentially detrimental consequences.

also be caused by errors in the regulation of DNA methylation causing an unscheduled methylation of CpG islands (Fig. 1B). Finally, errors in the regulation of DNA methylation may also lead to a loss of methylation and may cause ectopic gene activation, including regular and cryptic transcriptional start sites (Fig. 1C). The ectopic activation of cryptic transcriptional start sites may lead to the expression of truncated proteins or may prevent expression of essential genes via an anti-sense RNA repression mechanism, both of which may be detrimental for the cell or even the entire organism.

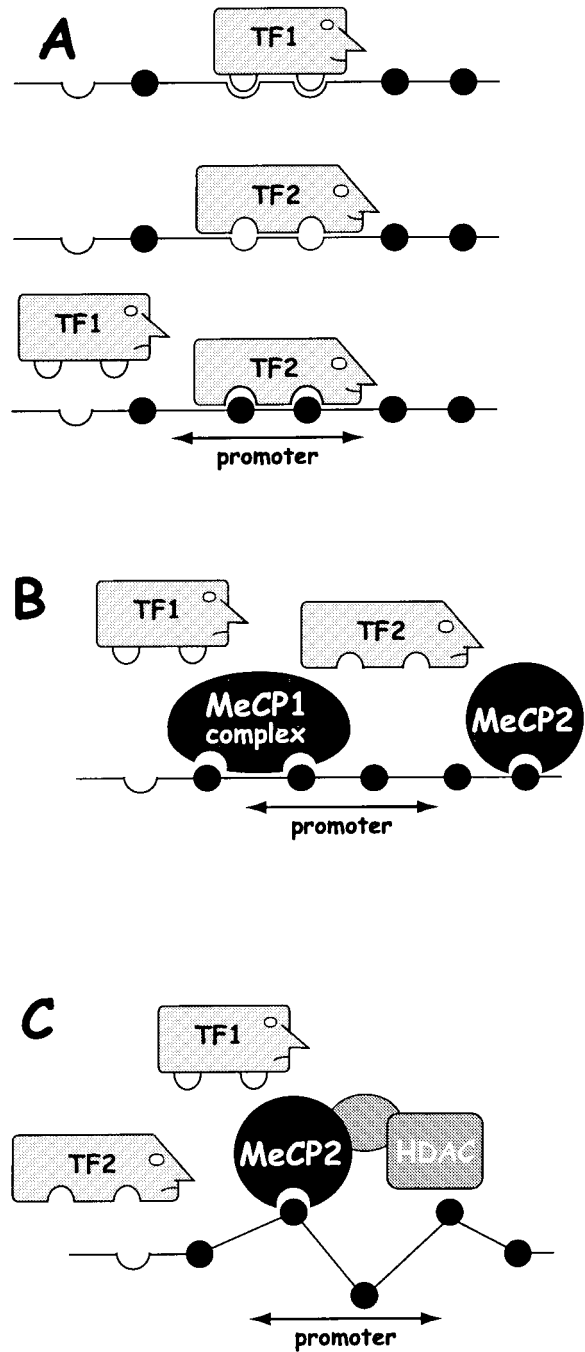
These correlations raise the question how DNA methylation can prevent transcription and vice versa. On one hand, early studies on the promoter region of the tyrosine aminotrans-

ferase gene showed that DNA methylation can directly interfere with protein-DNA interactions [Becker et al., 1987]. On the other hand, many DNA binding proteins like, e.g., the Sp1 transcription factor, are not directly sensitive to DNA methylation [Höller et al., 1988]. This is summarized in Figure 2A showing one type of transcription factor (TF1) that is prevented from binding to the promoter sequence by DNA methylation and another type (TF2) that does not discriminate. For the regulation of gene expression in mammalian cells this direct interference with DNA binding, however, seems to be of less importance as it is superseded by higher order structures as outlined in Figure 2B and C. Over the past years a number of different methyl-CpG binding proteins (MBD1-4 and MeCP2) were identified [Bird and Wolffe, 1999] that compete with transcription factors and prevent them from binding to promoter sequences (Fig. 2B). More important than this direct competition is the recently discovered ability of these methyl-CpG binding factors to recruit histone deacetylases (HDACs), which then cause a condensation of local chromatin structure and thereby render the methylated DNA less accessible to transcription factors (Fig. 2C). MeCP1 is a multi-protein complex that contains MBD2 together with HDACs and HDAC-binding factors [Ng et al., 1999] while MeCP2 is a methyl-CpG binding protein itself that binds to Sin3a and HDACs [Nan et al., 1998]. Both MeCP complexes specifically bind to methylated DNA, cause chromatin remodelling and repress transcription [Bird and Wolffe, 1999].

These recent results demonstrate a functional link between DNA methylation, chromatin structure, and gene expression in mammalian cells. This methylation directed chromatin condensation divides the complex mammalian genomes into transcriptionally active parts that are accessible to transcription factors and silenced parts that are less accessible to diffusing factors. This functional division of the genome facilitates the regulation of gene expression as transcription factors have to scan only a subset of the entire genome and it further enhances their precision by quenching initiation at cryptic sites.

#### DNA METHYLATION AND CANCER

These potential benefits of DNA methylation in mammalian cells, however, come at a high



**Fig. 2.** Transcriptional repression through DNA methylation. Different mechanisms have been described how DNA methylation can prevent transcription factor binding and lead to gene silencing. **A:** DNA binding of some transcription factors (represented by TF1) is directly affected by DNA methylation while others (TF2) are not inhibited. **B:** More frequent, however, are indirect effects through competition with methyl-CpG binding protein complexes (MeCP1 and 2) for binding sites. **C:** Most importantly, these MeCP complexes recruit histone deacetylases (HDACs) that cause deacetylation of adjacent histones leading to chromatin condensation and gene silencing.

price. Its central role in the regulation of gene expression naturally entails the risk of detrimental consequences if errors in the DNA methylation pattern occur. Virtually all tumors are characterized by more or less extensive alterations in their methylation pattern. Matters, however, get complicated by the fact that transformed cells show both, global hypomethylation and local hypermethylation, at the same time. In addition, the major methyltransferase, Dnmt1, is often upregulated in transformed cells which is hard to reconcile with the simultaneous, genome-wide hypomethylation. These complex issues have been reviewed elsewhere [Baylin et al., 1998; Jones and Laird, 1999]. It is still unknown whether hypomethylation is just a consequence of the transformation process or whether hypomethylation may actually cause transformation. It has been speculated that hypomethylation may lead to an elevated mutation and recombination frequency that would eventually cause inactivation of tumor suppressors and/or upregulation of oncogenes and thereby trigger cellular transformation. Much better understood is the role of local hypermethylation in the cellular transformation process.

First evidence came from studies on the founding member of the tumor suppressor family, the retinoblastoma gene. In some tumors the retinoblastoma was found to be silenced by ectopic methylation in the promoter region [Ohtani-Fujita et al., 1993; Greger et al., 1994]. Subsequent studies on different tumors identified a variety of genes that were silenced by methylation and play a role in the control of cell proliferation and migration. The list of affected genes ranges from cyclin-dependent kinase inhibitors (p15 and p16) and cell adhesion factors (E-cadherin) to differentiation factors like the MyoD transcription factor. In several of these cases it was possible to reverse cellular transformation by treating cells with an inhibitor of DNA methylation (5-azacytosine) causing re-expression of the ectopically silenced genes [reviewed in Baylin et al., 1998].

These results have encouraged speculations whether it might be feasible to treat this type of tumors with inhibitors of DNA methylation. Any such approach, however, would face the dilemma that demethylation might lead to an overexpression of growth promoting factors and that tumors are usually characterized by irre-

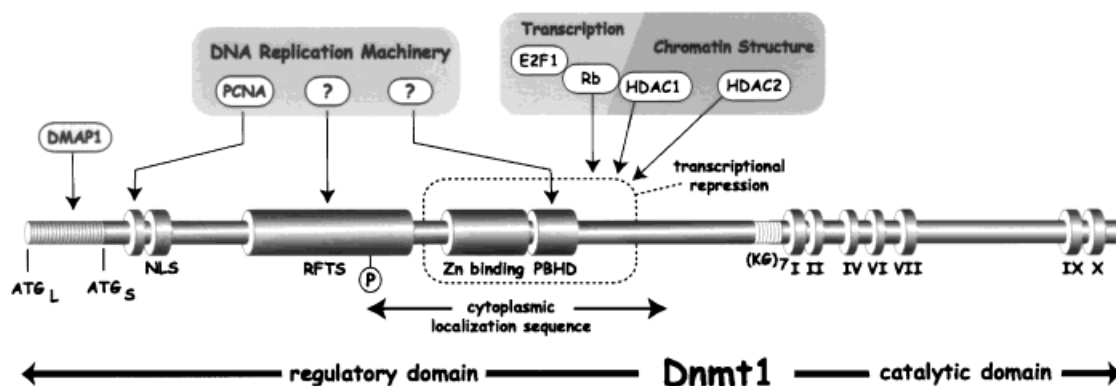
versible, secondary changes like genomic rearrangements.

Although the occurrence of alteration in DNA methylation in all types of tumors is now well documented it is still unclear how they arise. The upregulation of Dnmt1 in tumors can hardly explain both, hypo- and hypermethylation, and might simply be a reflection of the higher proliferative state of tumors. Although forced overexpression of Dnmt1 in cultured cells was found to cause transformation [Wu et al., 1993] this can hardly explain tumor formation in general since Dnmt1 is expressed at very different levels in different cell types. Most notably, one of the cell types expressing high levels of Dnmt1 are postmitotic neurons and mature oocytes. Given the complex alterations in DNA methylation during development and disease it is likely that the regulation and deregulation of DNA methylation involves interactions with other cellular factors, in particular the transcription and chromatin remodeling machinery.

Finally, methylated cytosine by itself is highly mutagenic and thus potentially carcinogenic. It was estimated that about 35% of all point mutations in human genetic diseases are C to T transitions at CpG sites [Cooper and Youssoufian, 1988]. The reason for this high mutation rate is the similarity of methylated cytosine and thymine (notice both have a methyl group at the 5 position) so that hydrolytic deamination of 5-methylcytosine directly generates thymine. This spontaneous deamination could be enhanced by DNA methyltransferases in the presence of limiting amounts of S-adenosylmethionine, which is an essential cofactor [Shen et al., 1992]. These results provide an interesting link between DNA methylation, cancer and nutrition.

#### DNA METHYLATION - A COMPLEX PROBLEM WITH CONSEQUENCES

The complex regulation and biological role of DNA methylation were hard to explain with just a few methyltransferases being up or down regulated during development and disease. The first challenge was to explain how DNA replication and methylation are coordinated during the cell cycle to guarantee the precise maintenance of a given methylation pattern. Studies on the subcellular localization of Dnmt1 revealed that Dnmt1 is redistributed to replication



**Fig. 3.** Overview of DNA methyltransferase binding factors. The structure of the major mammalian DNA methyltransferase (Dnmt1) is schematically outlined. Dnmt1 contains a C-terminal domain of about 500 amino acids including the ten conserved sequence motifs of pro- and eukaryotic C5 cytosine methyltransferases. The N-terminal, regulatory domain exists in two variants generated by alternative transcriptional start sites and translation initiation at either ATG-L (large isoform) or ATG-S (short isoform). Several functional domains have been mapped including a Zn binding region [Bestor, 1992], a nuclear

localization sequence (NLS) [Cardoso and Leonhardt, 1999], a replication foci targeting sequence (RFTS) [Leonhardt et al., 1992], a phosphorylation site [Glickman et al., 1997], a polybromo homology domain (PBHD) that also targets to replication foci [Liu et al., 1998], a transcriptional repression domain [Fuks et al., 2000] and a cytoplasmic localization sequence [Cardoso and Leonhardt, 1999]. Several proteins have been reported to interact with Dnmt1 including DMAP1 and HDAC1 [Rountree et al., 2000], PCNA [Chuang et al., 1997], and a complex of E2F1, pRB, and HDAC1 [Robertson et al., 2000].

sites during S-phase. This association with the cellular replication machinery was found to be mediated by a distinct replication foci targeting sequence (RFTS) located in the N-terminal, regulatory domain of Dnmt1 (Fig. 3). These results suggested that Dnmt1 is part of a complex DNA replication and methylation machinery that replicates the epigenetic along with the genetic information in a processive, assembly-line-like mode [Leonhardt et al., 1992]. Further protein binding assays showed that Dnmt1 binds to PCNA. The minimal region required for this interaction is located at the N-terminus of Dnmt1 [Chuang et al., 1997]. Later, it was shown that also the region similar to the Polybromo-1 protein (PBHD) could direct Dnmt1 to replication sites [Liu et al., 1998]. This means that the association with sites of DNA replication is independently mediated by three parts of the N-terminal domain of Dnmt1 (Fig. 3). It remains to be investigated which components of the replication machinery RFTS and PBHD bind to, whether they direct Dnmt1 to different subsets of nuclear replication sites and how their activity is coordinated. Recently, a complex with MBD2 and MBD3 was reported that specifically binds hemi-methylated DNA and contains Dnmt1 [Tatematsu et al., 2000]. This complex could specifically direct Dnmt1 to hemimethylated DNA at replication sites and

could thus achieve the maintenance of DNA methylation patterns after each round of DNA replication.

Biochemical purification experiments and transcription assays identified a repressor domain in the N-terminal domain of Dnmt1 that forms a complex with pRB, E2F, and HDAC1 [Fuks et al., 2000; Robertson et al., 2000]. An independent genetic screen for interacting proteins revealed the interaction with another histone deacetylase (HDAC2) and a novel transcriptional repressor (DMAP1) [Rountree et al., 2000]. These results suggest that Dnmt1 may repress transcription in a direct and an indirect way. The indirect repression is mediated by methyl-CpG binding factors that then recruit HDACs (see Fig. 2C). The direct repression is caused by Dnmt1 binding to hemi-methylated sites and direct recruitment of transcriptional repressors and HDACs (see Fig. 3). Both mechanisms would contribute to the restoration of repressive chromatin structures at methylated sequences on both daughter strands after DNA replication.

#### ACKNOWLEDGMENTS

We thank the Deutsche Forschungsgemeinschaft and the Max Delbrück Center for their support.

## REFERENCES

- Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. 1999. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 23:185–188.
- Baylin SB, Herman JG, Graff JR, Vertino PM, Issa JP. 1998. Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv Cancer Res* 72:141–196.
- Becker PB, Ruppert S, Schutz G. 1987. Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors. *Cell* 51:435–443.
- Bestor TH. 1992. Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *EMBO J* 11:2611–2617.
- Bestor T, Laudano A, Mattaliano R, Ingram V. 1988. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J Mol Biol* 203:971–983.
- Bird AP. 1986. CpG-rich islands and the function of DNA methylation. *Nature* 321:209–213.
- Bird AP, Wolffe AP. 1999. Methylation-induced repression-belts, braces, and chromatin. *Cell* 99:451–454.
- Cardoso MC, Leonhardt H. 1999. DNA methyltransferase is actively retained in the cytoplasm during early development. *J Cell Biol* 147:25–32.
- Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF. 1997. Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* 277:1996–2000.
- Cooper DN, Youssoufian H. 1988. The CpG dinucleotide and human genetic disease. *Hum Genet* 78:151–155.
- Doerfler W. 1983. DNA methylation and gene activity. *Annu Rev Biochem* 52:93–124.
- Fuks F, Burgers WA, Brehm A, Hughes-Davies L, Kouzarides T. 2000. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat Genet* 24:88–91.
- Glickman JF, Pavlovich JG, Reich NO. 1997. Peptide mapping of the murine DNA methyltransferase reveals a major phosphorylation site and the start of translation. *J Biol Chem* 272:17851–17857.
- Greger V, Debus N, Lohmann D, Hopping W, Passarge E, Horsthemke B. 1994. Frequency and parental origin of hypermethylated RB1 alleles in retinoblastoma. *Hum Genet* 94:491–496.
- 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* 2:1127–1135.
- Jones PA, Laird PW. 1999. Cancer epigenetics comes of age. *Nat Genet* 21:163–167.
- Leonhardt H, Page AW, Weier HU, Bestor TH. 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71:865–873.
- Li E, Bestor TH, Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69:915–926.
- Li E, Beard C, Jaenisch R. 1993. Role for DNA methylation in genomic imprinting. *Nature* 366:362–365.
- Liu Y, Oakeley EJ, Sun L, Jost JP. 1998. Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. *Nucleic Acids Res* 26:1038–1045.
- Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393:386–389.
- Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, Tempst P, Reinberg D, Bird A. 1999. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat Genet* 23:58–61.
- Ohtani-Fujita N, Fujita T, Aoike A, Osifchin NE, Robbins PD, Sakai T. 1993. CpG methylation inactivates the promoter activity of the human retinoblastoma tumor-suppressor gene. *Oncogene* 8:1063–1067.
- Okano M, Xie S, Li E. 1998. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res* 26:2536–2540.
- Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99:247–257.
- Robertson KD, Ait-Si-Ali S, Yokochi T, Wade PA, Jones PL, Wolffe AP. 2000. DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat Genet* 25:338–342.
- Rountree MR, Bachman KE, Baylin SB. 2000. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat Genet* 25:269–277.
- Shen JC, Rideout WMD, Jones PA. 1992. High frequency mutagenesis by a DNA methyltransferase. *Cell* 71:1073–1080.
- Tatematsu KI, Yamazaki T, Ishikawa F. 2000. MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase. *Genes Cells* 5:677–688.
- Wolffe AP, Jones PL, Wade PA. 1999. DNA demethylation. *Proc Natl Acad Sci USA* 96:5894–5896.
- Wu J, Issa JP, Herman J, Bassett DE Jr., Nelkin BD, Baylin SB. 1993. Expression of an exogenous eukaryotic DNA methyltransferase gene induces transformation of NIH 3T3 cells. *Proc Natl Acad Sci USA* 90:8891–8895.