**Functions of DNA methylation: islands, start sites, gene bodies and beyond**

**Peter A. Jones**

Abstract | DNA methylation is frequently described as a ‘silencing’ epigenetic mark, and indeed this function of 5-methylcytosine was originally proposed in the 1970s. Now, thanks to improved genome-scale mapping of methylation, we can evaluate DNA methylation in different genomic contexts: transcriptional start sites with or without CpG islands, in gene bodies, at regulatory elements and at repeat sequences. The emerging picture is that the function of DNA methylation seems to vary with context, and the relationship between DNA methylation and transcription is more nuanced than we realized at first. Improving our understanding of the functions of DNA methylation is necessary for interpreting changes in this mark that are observed in diseases such as cancer.

Two key papers in 1975 independently suggested that methylation of cytosine residues in the context of CpG dinucleotides could serve as an epigenetic mark in vertebrates. These papers proposed that sequences could be methylated *de novo*, that methylation can be inherited through somatic cell divisions by a mechanism involving an enzyme that recognizes hemimethylated CpG palindromes, that the presence of methyl groups could be interpreted by DNA-binding proteins and that DNA methylation directly silences genes. Although several of these key tenets turned out to be correct, the relationship between DNA methylation and gene silencing has proved to be challenging to unravel.

Most work in animals has focused on 5-methylcytosine (5mC) in the CpG sequence context. Methylation of other sequences is widespread in plants and some fungi and has recently been reported in mammals. In mammals, the function of non-CpG methylation is currently unknown. Here I primarily focus on CpG methylation in mammalian genomes, including some discussion of the differences observed in other animals and in plants.

Understanding the functions of DNA methylation requires consideration of the distribution of methylation across the genome. More than half of the genes in vertebrate genomes contain short (approximately 1 kb) CpG-rich regions known as CpG islands (CGIs), and the rest of the genome is depleted for CpGs. As 5mC can be converted to thymine by spontaneous or enzymatic deamination, it is thought that the loss of genomic CpGs is due to deamination of methylated sequences in the germline; CGIs are thought to exist because they are probably never or only transiently methylated in the germline. However, there is a lot of discussion as to exactly what the definition of the CGI should be, and although the CpG density of promoters in mammalian genomes has a bimodal distribution, regions with intermediate CpG densities also exist. Until recently, much of the work on DNA methylation focused on CGIs at transcriptional start sites (TSSs), and it is this work that has tended to shape general perceptions about the function of DNA methylation.

Recent approaches that enable genome-wide studies of the methylome (BOX 1) — for example, using bisulphite-treated DNA (which detects 5mC and hydroxymethylcytosine; see BOX 1) — have emphasized that the position of the methylation in the transcriptional unit influences its relationship to gene control. For example, methylation in the immediate vicinity of the TSS blocks initiation, but methylation in the gene body does not block and might even stimulate transcription elongation, and exciting new evidence suggests that gene body methylation may have an impact on splicing. Methylation in repeat regions such as centromeres is important for chromosomal stability (for example, chromosome segregation at mitosis) and is also likely to suppress the expression of transposable elements and thus to have a role in genome stability. The role of methylation in altering the activities of enhancers, insulators and other regulatory elements is only just...
Ten-eleven translocation (TET). Proteins of this type were recently shown to catalyse the conversion of 5-methylcytosine to 5-hydroxymethylcytosine.

Activation-induced cytidine deaminase (AID). An enzyme that removes the amino group from cytosine (5-methylcytosine). It is involved in class-switch recombination and DNA demethylation.

Thymine DNA glycosylase A protein that is involved in the repair of GC mismatches that are often caused by 5-methylcytosine deamination and that participates in DNA demethylation.

Nucleosome-depleted regions (NDRs). Regions of DNA that are not extensively wrapped up in nucleosomes. They can be seen at transcription start sites and other regulatory regions such as enhancers.

Polycomb proteins Polycomb proteins participate in the silencing of genes by mechanisms that do not involve DNA methylation. They often silence genes that are key regulators of differentiation.

Patterns at CpG island transcription start sites. Most CGIs remain unmethylated in somatic cells. When genes with CGIs at their TSS are active, their promoters are usually characterized by nucleosome-depleted regions (NDRs) at the TSS, and these NDRs are often flanked by nucleosomes containing the histone variant H2A.Z and are marked with trimethylation of histone H3 at lysine 4 (H3K4me3). The levels of gene expression are controlled by transcription factors. CGI promoters can be repressed by various mechanisms, such as repression mediated by Polycomb proteins. For example, genes encoding master regulators of embryonic development, such as myogenic differentiation 1 (MYOD1) or paired box 6 (PAX6), are repressed by the Polycomb complex both in ESCs and in differentiated cells that are not expressing these genes; they have nucleosomes at the TSS and are marked by H3K27me3, which is generally associated with inactive genes.

However, some repressed genes do have methylated promoter CGIs. Methylated promoter CGIs are usually restricted to genes at which there is long-term stabilization of repressed states. Examples include imprinted genes, genes located on the inactive X chromosome and genes that are exclusively expressed in germ cells and that would presumably be inappropriate for expression in somatic cells. The stabilization of suppression by DNA methylation of CGIs can last over a 100-year lifespan and has no effect on the existence
Box 2 | Known and unknown features of DNA methylation in mammals

This box summarizes the key points regarding our knowledge and lack of knowledge of DNA methylation in mammals.

**Known**
- Most CpG islands (CGIs) are not methylated when located at transcription start sites (TSSs).
- CGI methylation of the TSS is associated with long-term silencing (for example, X-chromosome inactivation, imprinting, genes expressed predominantly in germ cells and some tissue-specific genes).
- CGIs in gene bodies are sometimes methylated in a tissue-specific manner.
- Non-CGI methylation is more dynamic and more tissue-specific than CGI methylation.
- Methylation blocks the start of transcription not elongation (note that this is different in the fungus Neurospora crassa).
- Methylation of transposable elements silences these elements but allows the host gene to undergo transcriptional elongation.
- Gene body methylation contributes to cancer-causing somatic and germline mutations.

**Unknown**
- Does non-CGI methylation silence genes (that is, is it a cause or a consequence)?
- The function of methylation in the context CHG (where H is A, C or T).
- The roles of active and passive demethylation in activating genes.
- The function of ‘shore’ methylation.
- Does gene body methylation control splicing?
- The role of 5-hydroxymethylation in the brain and other tissues.
- The role of methylation in enhancer or insulator function.
- Does silencing always precede methylation?

of CGIs, because any deamination events within these regions in somatic cells would not be passed on through the germline to subsequent generations. We still do not completely understand why a minority of CpG islands become methylated, whereas most do not.

**Patterns at non-CGI island TSSs.** In contrast to genes with CGIs at their TSSs, substantial fluctuations occur in the promoter methylation levels of genes that are CpG-poor at the TSS. Genes with non-CGI TSSs that are expressed in primordial germ cells are unmethylated at the TSS, whereas genes that are exclusively expressed in ESCs or tissue-specific genes often show methylation in sperm but not in oocytes or in expressing somatic cells. Well-known examples are the genes encoding the OCT4 and NANOG transcription factors, which are essential for the maintenance of the stem cell state. Recent studies have suggested that the OCT4 and NANOG promoters may undergo active demethylation by AID1,2 and/or by TET3 (REF. 27). Some tissue-specific genes, however, show methylation in sperm and in ESCs and only show demethylation in the specific tissues in which these genes are expressed.

One genome-wide study postulated that no inverse relationship existed between methylation of non-CGIs and expression, but re-analysis of the data suggested that such a relationship between expression and methylation is in fact apparent genome-wide. Because of the long-standing focus on CGIs, we still do not know the details of the role of methylation in controlling non-CGI TSSs.

---

**Imprinted genes**
Imprinted genes show parent-of-origin expression and are controlled by epigenetic processes, including DNA methylation.

**X-chromosome inactivation**
One of the two X chromosomes in female mammalian somatic cells is stably silenced by epigenetic processes, including DNA methylation, to achieve dosage compensation.

**Does methylation silence transcription initiation?** Given the observations of methylation at some repressed TSSs described above, what is the functional relationship between DNA methylation and transcription initiation? There is incontrovertible evidence that methylated CGIs at TSSs cannot initiate transcription after the DNA has been assembled into nucleosomes. However, the issue of whether silencing or methylation comes first has long been a discussion in the field. Early experiments by Lock et al. clearly showed that methylation of the Hprt gene on the inactive X chromosome occurred after the chromosome had been inactivated. In other words, methylation appeared to serve as a ‘lock’ to reinforce a previously silenced state of X-linked genes. Although most CGIs on autosomal genes remain unmethylated in somatic cells, a small number of them (<10%) become methylated in normal tissues and cells, but the timing of the de novo methylation with respect to silencing has not been studied in depth. As mentioned above, recent findings regarding the role of DNMT3A in haematopoietic stem cell differentiation raise doubts about the universality of the long-term ‘locking’ model. Because the authors of this study showed that the methylase was essential for differentiation of a fairly short-lived cell type, it seems possible that DNA methylation has a more instructive role in initiating rather than reinforcing the silencing.

Genome-wide studies in cancer cells have, however, shown that genes with CGI promoters that are already silenced by Polycomb complexes are much more likely than other genes to become methylated in cancer: that is, the silent state precedes methylation. Therefore, it seems likely that silencing preceding methylation is a general mechanism, but the data are not yet mature enough to be sure. In addition to alterations at the CpG islands themselves, tissue-specific changes occur in ‘shores’ surrounding them. However, the implications of these alterations is not yet understood. The evidence regarding the timing of DNA methylation is consistent with the idea that methylation adds an additional level of stability to epigenetic states. Intriguingly, it is not required for this purpose in some species, including Drosophila melanogaster and yeast.

**The relationship between transcription and de novo methylation.** The reasons why DNA methylation is probably not used as an initial silencing mechanism are now starting to be understood. The pioneering work of Ooi et al. showed that the process of de novo methylation in cells expressing DNMT3L (which is a catalytically inactive homologue of DNMT3A and DNMT3B) is achieved by a tetrameric complex of two molecules each of DNMT3A2 and DNMT3L and requires a nucleosome. Active TSSs are depleted of nucleosomes and therefore lack this substrate for de novo methylation. Recently, we have directly tested the role of the nucleosome in triggering de novo methylation by examining the kinetics of OCT4 silencing in embryonal carcinoma cells induced to differentiate with retinoic acid. These experiments showed that at the OCT4 distal enhancer and the NANOG promoter, following differentiation, first a
nucleosome becomes present, and then this is followed by the recruitment of DNMT3A to this nucleosome and, subsequently, de novo methylation occurs. Whether a similar sequence of events occurs in cells that are not expressing DNMT3L is not yet known.

Furthermore, Ooi et al.\textsuperscript{12} showed that de novo methylation could not occur on a nucleosome bearing the H3K4me3 or H3K4me3 marks, which are associated with active genes. The nucleosomes flanking the nucleosome-depleted start site often contain both the histone mark H3K4me3 and the histone variant H2A.Z, both of which are strongly anti-correlated with DNA methylation\textsuperscript{46-49}. The occurrence of the H3K4me3 mark in mice is possibly maintained by the presence of CXXC finger protein 1 (CXXC1; also known as CFP1), which recruits the H3K4 methyltransferase to the region, thus ensuring that the +1 and −1 nucleosomes contain marks that are incompatible with de novo DNA methylation\textsuperscript{49}. The unmethylated state of the CpG island is also presumably ensured by the presence of the TET1 protein, which is found at a high proportion of the TSSs of high-CpG-content promoters. Presumably, TET1 converts any 5mC that might be in this region into 5-hydroxymethylcytosine\textsuperscript{49}. The molecular anatomy of active CGIs can therefore explain why they are resistant to methylation (FIG. 1).

Of course, not all CGI-promoter genes are expressed in ESCs, and many are suppressed by the Polycomb complex, so why are these not de novo methylated? The answer probably lies in the fact that they contain the antagonistic H3K4me3 (REF. 12) and H2A.Z marks\textsuperscript{46-47} and are also bound by TET1, which would ensure that they remain 5mC-free. Interestingly, this protection seems to break down during immortalization\textsuperscript{50}, and these CGIs become highly susceptible to de novo methylation, which increases after oncogenic transformation\textsuperscript{41-43}.

This model predicts that the higher the level of expression is, the less likely it is that a CGI is to become de novo methylated. Direct evidence in support of this prediction has recently come from several exciting papers that have shown that monoallelic methylation of CGIs preferentially occurs on the allele that is less highly expressed. For example, Hitchins et al.\textsuperscript{51} showed that an allele of the MLH1 gene containing a single-nucleotide variant in the promoter, which was less active than the more common allele in transfection experiments, was more likely to become methylated in the somatic cells of cancer-affected families. In other words, the less active allele was the one that was more likely to acquire de novo methylation. An alternative scenario was shown by Bounbec et al.\textsuperscript{52}, who found that an allele of RIL (also known as PDLIM4) bearing a polymorphism in the promoter that created an additional binding site for the transcription factor SP1 or SP3 was much less likely to become de novo methylated than the allele without this polymorphism. The extra SP1 site therefore confers resistance of this allele to de novo methylation, although the authors could not demonstrate that the extra transcription factor binding site increased gene expression.

**Gene body methylation**

Most gene bodies are CpG-poor, are extensively methylated and contain multiple repetitive and transposable elements. Methylation of the CpG sites in gene exons is a major cause of C→T transition mutations, leading to disease-causing mutations in the germline and cancer-causing mutations in somatic cells\textsuperscript{53}. It is important to realize that although many CGIs are located at gene promoters, CGIs also exist within the bodies of genes\textsuperscript{54} and within gene deserts. Although their functions here remain unknown, Adrian Bird has proposed that these regions may represent ‘orphan promoters’
that might be used at early stages of development and have escaped methylation in the germline so that their high CpG density is maintained.\cite{54}

**Gene body methylation is not associated with repression.** It has been known from the early days of DNA methylation research that gene body methylation is a feature of transcribed genes. Extensive positive correlations between active transcription and gene body methylation have recently been confirmed on the active X chromosome\cite{55} and by shotgun bisulphite sequencing of plant and animal genomes.\cite{53,56} Most gene bodies are not CGIs, and when CGIs are situated in intragenic regions, they were, with a few exceptions\cite{57}, thought to remain unmethylated. However, recent experiments\cite{58,59} have changed this perception: for example, as many as 34\% of all intragenic CGIs are methylated in the human brain.\cite{60} The role of this methylation, which is tissue-specific, is not yet clear. It is intriguing, especially because TSSs largely remain unmethylated. Intragenic CGIs can also be preferential sites for de novo methylation in cancer.\cite{61}

Even though gene body CGIs can become extensively methylated, this does not block transcription elongation. This is despite the fact that the methylated CGIs are marked by H3K9me3 and are bound by methyl-CpG-binding protein 2 (MECP2), which are chromatin features that are associated with repressed transcription when they are present at the TSS.\cite{62} This leads to an apparent paradox in which methylation in the promoter is inversely correlated with the expression, whereas methylation in the gene body is positively correlated with expression.\cite{63} Thus, in mammals, it is the initiation of transcription but not transcription elongation that seems to be sensitive to DNA methylation silencing. By contrast, cytosine methylation in CpG and other sequence contexts in *Neurospora crassa* blocks elongation but not initiation.\cite{64} Therefore, it is not simply the presence of a 5mC mark itself that governs its relationship to transcription but rather the interpretation of the mark in a particular genomic and cellular context.

**Possible functions of gene body methylation.** What is the function of the gene body methylation outside CGIs? Initially, it was thought that this methylation was primarily a mechanism for silencing repetitive DNA elements, such as retroviruses, LINE1 elements, *Alu* elements and others, and evidence has been obtained to substantiate this idea.\cite{65} Methylation blocks initiation of transcription at these elements while at the same time allowing transcription of the host gene to run through them.

It has also been proposed that the process of transcript elongation could itself stimulate DNA methylation and that H3K36me3, which is also associated with elongation but not initiation, might be involved in the recruitment of DNMTs.\cite{66} However, whole-genome studies have shown that there might be alternative functions for DNA methylation in gene bodies. This work has shown that exons are more highly methylated than introns, and transitions in the degree of methylation occur at exon–intron boundaries, possibly suggesting a role for methylation in regulating splicing. Indeed, genome-wide nucleosome-positioning data suggest that exons also show increased nucleosome occupancy levels compared to introns and nucleosomes are preferential sites for DNA methylation.\cite{67} A recent study has suggested that binding of CTCF (which can be regulated

![Figure 2](nature-revs-genetics/figures/Fig2.png)
by DNA methylation; see below) results in the pausing of RNA polymerase II (RNAPII) and, as the kinetics of RNAPII movement influences splicing, this might link DNA methylation to splicing. These observations suggest a previously unrecognized role for DNA methylation at the transcriptional level, possibly resulting in alternative splicing. Therefore, it seems likely that DNA methylation in gene bodies will have outcomes beyond the recognized function in the silencing of intragenic repetitive DNA sequences.

**When is a body a start site?** It is often assumed that TSSs and gene bodies are two separate genomic features. However, most genes have at least two TSSs, so the downstream start sites are within the ‘bodies’ of the transcriptional units of the upstream promoters. These alternative promoters can be CGIs or non-CGIs, or there can be combinations of an upstream non-CGI and a downstream CGI, or vice versa. These alternative start sites complicate the interpretation of experiments linking expression to methylation, because probes that are used to measure expression often detect the output of all of the promoters, yet only one might be active in a given cell type. Methylation of a downstream promoter would only block transcription from that promoter — it would allow the elongation of a transcript that emanates from an upstream promoter — leading to an apparent discordance between methylation and expression. Indeed, DNA methylation may well be a mechanism for controlling alternative promoter usage.

**Other regulatory sites**

**Methylation at enhancers.** Enhancers are situated at variable distances from promoters and are key to controlling gene expression in development and cell function. They are mostly CpG-poor, and their methylation status has been examined by whole-methylome analysis (in plants and mammals). In general, these regions tend to have fairly variable methylation. Indeed, Stadler et al. identified enhancers in the mouse genome on the basis that they are regions that are not 100% methylated or unmethylated and termed these ‘low-methylated regions’ (LMRs). Because a given cytosine can either be completely methylated or unmethylated, ‘variable methylation’ is the outcome of averaging these binary states. This might suggest that the CpG sites are in a dynamic state and that at a given time some are methylated and others are not, owing to competing methylation and demethylation events. Alternatively, the DNA methylation status of each CpG might not be accurately maintained during cell division, and so the LMR state might be due to inefficient inheritance. In different subsets of T cells, Schmidl et al. also found a large number of differentially methylated regions (DMRs) within the enhancers of differentiation specific genes. In terms of function, this study showed that methylation of these CpG sites could result in reduced activity of the enhancer in reporter assays.

The idea that the methylation status of an enhancer and enhancer function are closely connected is supported by several observations of proteins that modulate methylation at these regions. For example, analysis of the binding of the glucocorticoid receptor to distal regulatory elements showed that CpGs could become demethylated and that the enhancer could be activated by the presence of this receptor. Similar findings had originally been reported more than 25 years ago by Saluz et al., who demonstrated demethylation of the overlapping oestradiol and glucocorticoid receptor binding sites in roosters that had been treated with oestradiol. In addition, 5-hydroxymethylcytosine and the TET proteins can be detected at these elements. However, the relationship between CpG methylation and transcription factor binding is complex (see below), and so we have a long way to go before we understand the mechanisms by which the methylation of CpG-poor enhancers is involved in the regulation of these regions.

**Methylation at insulators.** Insulators can be defined as elements that block the interaction between an enhancer and a promoter. The most well-studied examples are DNA sequences bound by the CTCF protein, which binds to a somewhat heterogeneous sequence motif. A well-studied case is CFCF binding to a site within the imprinted IGF2–H19 locus, at which the presence or absence of CTCF binding controls enhancer–promoter interactions. It has been shown that methylation of a CTCF-binding site at this locus blocks the binding of CTCF, so DNA methylation has an important role in controlling this locus. More recent studies have likewise shown that CTCF binding to exons of the gene encoding CD45 is inhibited by DNA methylation, leading to effects on splicing. However, global studies in mouse ESCs and differentiated cells have suggested that CTCF binding within CpG-poor regions is generally not affected by the methylation status of the binding sites, but rather that the binding itself initiates local demethylation. Therefore, is it possible that there are no universal rules for the effects of methylation on CTCF sites (which tend to be degenerate) and binding? In this regard, it is important to note that there are seven potential CTCF-binding sites in the human H19 promoter, and only one of them shows differential parent-of-origin methylation.

**Possible mechanisms**

The mechanisms by which an inactive CGI promoter is held in a stably repressed state by DNA methylation are fairly well understood and have been extensively reviewed. The methylated promoter has nucleosomes at the TSS that bear the repressive H3K9me3 mark and that are stabilized by methylated DNA-binding proteins, which in turn recruit histone deacetylases to the region.

The issue of causality of methylation changes in stabilizing inactive gene expression states of non-CGI promoters has been the subject of much controversy, and the issue has not yet been adequately resolved. Because transcription factors can bind strongly to methylated DNA sequences, subsequently resulting in the passive...
demethylation of these regions, it is not always clear whether the methylation changes are a result of transcription or whether they stabilize transcriptionally incompetent states. Methylation of non-CGI regions can have a direct impact on the binding of transcription factors to target sites. Indeed, it has been known for some time that the binding of MYC to its cognate sequence is directly inhibited by the presence of 5mC₂, whereas the binding of SP1 does not show such a relationship. However, methylation of transcription factor binding sites — for example, in the laminin beta 3 (LAMB3), runt-related transcription factor 2 (RUNX2) or Oct4 (REF. 86) promoters — can decrease gene expression in transfection experiments.

More recent genome-wide studies have confirmed that transcription factor binding can be strongly influenced by methylation of CpG sites within their recognition sequences. These experiments point to a cause-and-effect relationship between CpG methylation at the TSS and gene repression, but there are still issues relating to the probable mechanisms. For example, a puzzling observation is our finding that in human ESCs, there is almost always no OCT4 bound at OCT4 target sites when there is DNA methylation within 100 bp on each side of the target sequence. As the OCT4 recognition sequence does not contain a CpG sequence, it is difficult to postulate a mechanism that could explain this strong correlation between CpG methylation in the vicinity of the site and a lack of binding. Nevertheless, the fact that methylation at the CpG sites flanking the OCT4-binding sites is inherently more variable than at CGIs, the neighbourhood could be a fruitful area for investigation in the future and has largely been overlooked until now.

**Conclusions**

Whole-genome approaches have given us a detailed view of the methylome and have shown that methylation patterns beyond TSSs are far more dynamic than was previously appreciated. Now that we have these global patterns, we need to resolve their potential roles and mechanisms — methylated sites beyond TSSs clearly are not simply ‘passengers’ as had been previously assumed (FIG. 1). Compared to CGIs at TSSs, which are generally unmethylated and seem mostly to be methylated to ensure long-term silencing, the patterns of modification in the CpG-depleted regions of the genome are more interesting. Although we clearly do not understand the detailed mechanisms by which methylation of enhancers, insulators and gene bodies influence the binding and function of regulatory proteins, there seems to be little doubt that this is crucial to development, differentiation and indeed cellular viability. However, some species are able to survive without methylation, yet mammals require three DNMTs; finding out why this is the case will help to explain its function. The identification of the TET genes and the localization of the TET proteins to regulatory regions is highly suggestive of a dynamic turnover of 5mC, which is consistent with it having a control function in gene expression.

The potential involvement of methylation beyond CGI promoters in human disease has been largely overlooked because of the focus on abnormal CGI methylation in cancer. Future work that aims to provide detailed maps of epigenomes in normal and diseased states is crucial to our understanding of many human diseases (BOX 3). This will be essential if we are to develop strategies and drugs to target the epigenome and to treat these diseases.

---

**Box 3 | DNA methylation and disease**

Methylation of cytosine strongly increases the rate of C→T transition mutations and is thought to be responsible for about one-third of all disease-causing mutations in the germline. In somatic cells, gene body methylation is a major cause of cancer gene mutations in tumour suppressor genes, such as TP53, which encodes p53 (REF. 55). The transcriptional start sites of many genes encoding tumour suppressors, such as retinoblastoma-associated protein 1 (RB1), MLH1, p16 and BRCA1, among others, lie within CGIs. Factors that reduce expression of these genes increase the likelihood of de novo methylation and irreversible silencing. These gene promoters have been found to be extensively methylated in a large number of tumours, such as retinoblastoma, colon, lung and ovarian cancers. The methylation changes are present in tumours before they are placed into culture but become enhanced during passing in vitro. In general, the tumour genome is hypomethylated, yet the potential role of methylation of CpG-poor promoters, enhancers, insulators, repetitive elements and gene bodies in cancer is almost completely unknown. The relevance of DNA methylation to human cancer has become even more evident with the identification of mutations in DNMT3A and isocitrate dehydrogenase 1 (IDH1) and IDH2 in leukemias. Several diseases, such as fragile X syndrome and immunodeficiency, centromere instability and facial anomalies syndrome (ICF syndrome) can be caused by mutations in DNA methyltransferase 3B (DNMT3B) and leads to the disease phenotype. This was the first paper to provide single-base DNA methylation and disease epigenomic differences.
This paper provided a structural basis to the mechanism of DNA demethylation and showed how active histone marks could exclude methylation of DNA.

This was a key paper in defining the need for DNA cytosine methylation in mammals.

This is a key paper that unequivocally established that the coherent application of methyl groups to DNA could result in silencing and is involved in X-chromosomal inactivation.

This paper unexpected showed that methylation of cytosine was not the primary silencing mechanism for X inactivation.

This paper pointed out the crucial role of 5mC in parasitic DNA methylation-induced gene silencing in cancer cells.

This paper showed the importance of histone deacetylation in the reprogramming of pluripotency genes.

This paper revealed nucleosome shifting on mitotically silenced promoters.

This study demonstrated that single-nucleotide variants that decrease promoter activity can lead to preferential allele-specific methylation.

This was a key paper showing how methylation of CTCF binding sites could regulate promoter activity by directly blocking binding of CTCF.


