# DNA METHYLATION IN HEALTH AND DISEASE

*Keith D. Robertson* and Alan P. Wolffe‡*

DNA methylation has recently moved to centre stage in the aetiology of human neurodevelopmental syndromes such as the fragile X, ICF and Rett syndromes. These diseases result from the misregulation of genes that occurs with the loss of appropriate epigenetic controls during neuronal development. Recent advances have connected DNA methylation to chromatin-remodelling enzymes, and understanding this link will be central to the design of new therapeutic tools.

HETEROCHROMATIN
The densely staining regions of the nucleus that generally contain condensed, transcriptionally inactive regions of the genome.

EUCHROMATIN
The lightly staining regions of the nucleus that generally contain decondensed, transcriptionally active regions of the genome.

BLASTOCYST
An early stage of embryonic development at which cells begin to commit to certain developmental lineages.

**Epigenetic Gene Regulation and Cancer Section, NCI, NIH, Building 41, 41 Library Drive, Bethesda, Maryland 20892-5431, USA.*
*‡Sangamo Biosciences, Point Richmond Tech Center, 501 Canal Boulevard, Suite A100, Richmond, California 94804, USA.*
*e-mail: awolffe@sangamo.com*
*Correspondence to: A.P. W.*

Today almost all of the genes in the human genome have been sequenced. The challenge now is to understand the molecular mechanisms that allow these genes to be selectively expressed. Although all genes are transcribed in an organism at some stage of its life cycle, a more restricted number are required for the differentiation of a specialized cell type. Remarkably, it is essential to select not only the correct genes to turn on, but also those that need to be inactivated. Failure to repress genes appropriately has been connected to many human diseases, including neurodevelopmental disorders and cancer. Notable among the mechanisms that stably inactivate genes in a heritable manner is DNA methylation and the associated assembly of repressive HETEROCHROMATIN.

The genomes of many animals are compartmentalized by being packaged into either transcriptionally competent EUCHROMATIN or repressive, transcriptionally silent heterochromatin. This compartmentalization arises progressively during early embryonic development. Mouse primordial germ cells, embryonic stem (ES) cells and the cells of the BLASTOCYST can progress through the cell cycle and divide without detectable DNA methylation[1]. Once differentiation begins, however, DNA methylation becomes essential for individual cell viability[2,3]. DNA methylation seems to function as a method of partitioning the genome, and the chromosomal infrastructure within which it is packaged, into active and inactive compartments[4]. Much of development relies upon the imposition of progressively more stable states of transcriptional repression[5,6]. This enables the transcription machinery to search out the limited number of genes that are essential for cell viability and differentiation.

In this article, we first summarize the basics of DNA methylation in the genome and discuss recent advances in our understanding of the cellular methylation machinery and its role in development. We also discuss the functions of DNA methylation in host defence and in transcriptional repression. Finally, we put all this together with known molecular events in human disease and discuss possible approaches to therapeutic intervention. The aim of this article is to illustrate how these processes go awry in the neurodevelopmental disorders of ICF (for immunodeficiency, centromeric instability and facial anomalies), Rett and fragile X syndromes. These disorders of defective DNA methylation highlight the importance of chromosomal modification for gene control and development. Accumulating observations emphasize the need to map not only the sequence of the human genome, but also the functional compartmentalization of this information in healthy and diseased cells. There are many excellent reviews that discuss methylation in disease, including the silencing of tumour suppressor genes in cancer[7,8], genomic IMPRINTING deficiencies[9–11], and aspects of ageing and heart disease[12], and so we will not focus on these areas here.

## DNA methylation and the genome

DNA methylation in mammalian cells occurs at the 5-position of cytosine within the CpG dinucleotide.

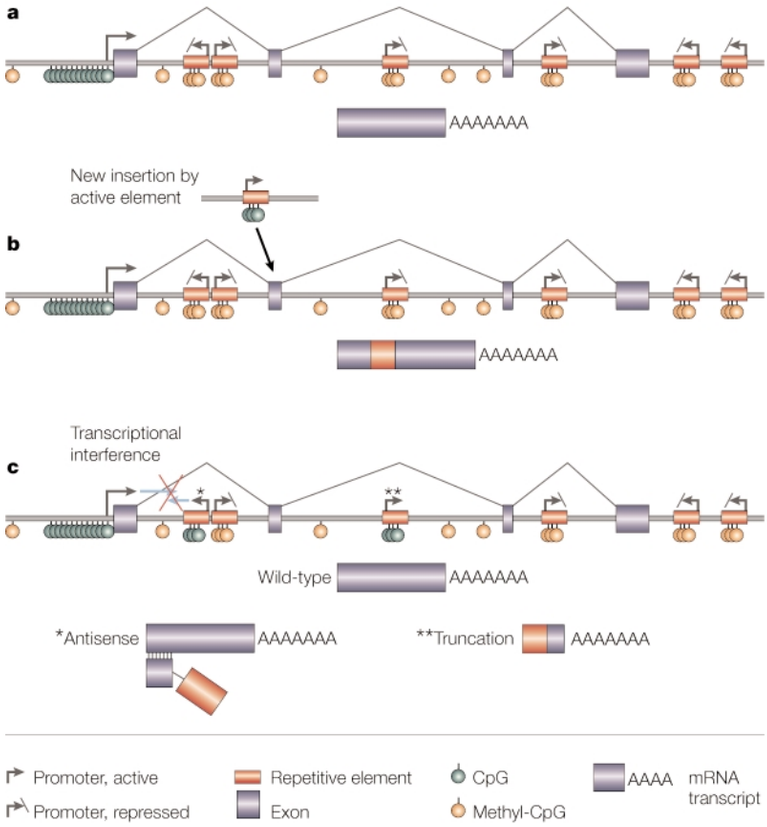


Figure 1 | **DNA methylation, CpG islands and genome defence. a** | Spatial distribution of the CpG dinucleotide in the genome and the concentration of 5-methylcytosine in repetitive or parasitic elements. Methylation of such elements prevents the expression of the genes required for their transposition. CpG islands are often associated with promoters and are unmethylated except on the inactive X chromosome in females. **b** | Dangers to genomic integrity of unregulated transposition. An active transposable element may integrate into the coding region of a gene and disrupt its function. **c** | Alternatively, the presence of active promoters within retrotransposons (because of reduced DNA methylation) may interfere with the normal transcription of the gene. This might occur through internal initiation that results in: a chimaeric messenger-RNA molecule, reduced mRNA levels or translation capacity through an antisense mechanism, or disrupted transcription initiation. The last two effects will occur if the retroelement is integrated in the antisense orientation relative to the direction of gene transcription.

IMPRINTING
A genetic mechanism by which genes are selectively expressed from the maternal or paternal chromosomes.

CPG ISLAND
A genomic region of about one kilobase that contains close to the theoretical, expected frequency of the CpG dinucleotide.

NUCLEOSOME
The basic structural subunit of chromatin, which consists of roughly 200 base pairs of DNA and an octamer of histone proteins.

Roughly 70% of all CpGs are methylated, but neither 5-methylcytosine distribution nor the spatial distribution of the CpG dinucleotide itself is random[13]. In general, the CpG dinucleotide is greatly under-represented throughout the mammalian genome (also termed CpG suppression) but it can be found at close to its expected frequency in small genomic regions of about one kilobase, called CPG ISLANDS (FIG. 1a)[14]. Although CpG islands account for only about 1% of the genome and for 15% of the total genomic CpG sites, these regions contain over 50% of the unmethylated CpG dinucleotides. There are about 45,000 CpG islands, most of which reside within or near the promoters or first exons of genes and are unmethylated in normal cells, with the exception of CpG islands on the inactive X chromosome in females[15]. CpG islands also have a markedly open chromatin structure that is deficient in the linker histone H1 and contain NUCLEOSOMES enriched in acetylated forms of histones H3 and H4 (REF. 16). CpG depletion is restricted to organisms that have heavily methylated genomes and is believed to arise from both the increased spontaneous DEAMINATION rate of 5-methylcytosine versus cytosine, and the differential repair efficiency of the deamination product of 5-methylcytosine (which is thymine, a naturally occurring DNA base) versus that of cytosine (which is uracil and is efficiently removed by the abundant uracil DNA glycosylase)[17]. These mutations result in transitions of CpG to TpG. CpG sites have been shown to act as hot-spots for germline mutations, contributing to 30% of all point mutations in the germ line[18], and for acquired somatic mutations that lead to cancer[19,20]. For example, methylated CpG sites in the *TP53* coding region contribute to as many as 50% of all inactivating mutations in colon cancer and to 25% of cancers in general[20].

### The cellular methylation machinery

Cellular DNA methylation patterns seem to be established by a complex interplay of at least three independent DNA methyltransferases: DNMT1, DNMT3A and DNMT3B (FIG. 2). DNMT1 was the first methyltransferase to be discovered[21]. Pioneering work has established that DNMT1 has a 10–40-fold preference for hemimethylated DNA[22,23]. DNMT1 is the most abundant methyltransferase in somatic cells[24], localizes to replication foci[25] and interacts with the proliferating cell nuclear antigen (PCNA)[26]. It is often referred to as the 'maintenance' methyltransferase because it is believed to be the enzyme responsible for copying methylation patterns after DNA replication. DNMT1 is required for proper embryonic development, imprinting and X-inactivation[2,27,28]. *Drosophila melanogaster*, which lacks detectable DNA methylation, also has a DNMT1-like protein, as well as a *DNMT2* homologue, indicating that methyltransferases may have an important and conserved function in a process other than DNA methylation[29,30].

More recently the DNMT3 family of methyltransferases (FIG. 2) from mouse and human have been characterized. They seem to be highly conserved: homologous genes have been identified in zebrafish, *Arabidopsis thaliana* and maize[31,32]. These enzymes are required for the wave of *de novo* methylation that occurs in the genome following embryo implantation and for the *de novo* methylation of newly integrated retroviral sequences in mouse ES cells[3]. *Dnmt3a* knockout mice are born live but become runted and die at about four weeks of age[3]. *Dnmt3b* knockout mice, in contrast, are not viable, and mutant embryos show numerous developmental defects and growth impairment after embryonic day (E) 9.5, close to the time at which the *Dnmt1* knockout mice begin to show growth defects[2,3]. These observations, coupled with *in vitro* data indicating that the DNMT3 enzymes have an equal preference for hemi- and unmethylated DNA substrates, have led to them being termed the '*de novo* methyltransferases'[31].

The idea that different methylating activities can be completely accounted for by different DNA methyltransferases is clearly far too simplistic. The search for the DNMT3 family was originally motivated by the

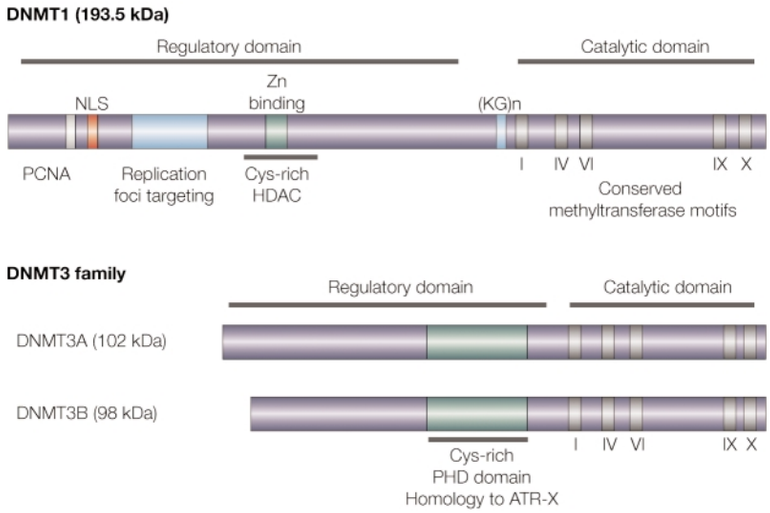


Figure 2 | **Schematic structure of the three catalytically active DNA methyltransferases in mammals showing the N-terminal regulatory and C-terminal catalytic domains, and other regions with known or proposed functions.** The catalytic domains of the three enzymes are conserved, but there is little similarity between their N-terminal regulatory domains. PCNA, domain that interacts with proliferating cell nuclear antigen; NLS, nuclear localization signal; (KG)n, lysine-glycine repeat hinge region; HDAC, histone deacetylase interaction domain; PHD, plant homeodomain motif that shows homology to the *ATR-X* (α-thalassaemia, mental retardation, X-linked) gene.

observation that *Dnmt1* knockout ES cells retained *de novo* methylating activity, supporting the classification of *Dnmt1* as the 'maintenance' enzyme[1]. This has been supported by recent experiments in which *Dnmt1* was expressed in *Drosophila* and showed no *de novo* methylating activity[33]. Furthermore, homozygous deletion of *Dnmt3a* and *Dnmt3b* does not alter pre-existing methylation patterns in mouse ES cells[3], whereas homozygous deletion of *Dnmt1* causes a roughly 70% reduction in 5-methylcytosine content[2].

Other observations, however, indicate that DNMT1 can act as a *de novo* methyltransferase. Enforced overexpression of *DNMT1* in cancer cell lines leads to *de novo* methylation of endogenous CpG islands[34]. An alternative interpretation of these results[34] is that the increased CpG methylation might occur because of the enhanced ability of these cells to fix the methylation pattern set down by the less efficient *de novo* methyltransferases. A recent report that somatic cells lacking *DNMT1* retain about 80% of their normal methylation levels has further complicated the '*de novo* versus maintenance' classification. This indicates that the DNMT3s may act as maintenance enzymes or that there are additional, undiscovered methyltransferases that can compensate for the loss of DNMT1 (REF. 35). It is far more likely that all three enzymes possess both *de novo* and maintenance functions and that, at least in somatic cells, specific methyltransferases will be responsible for the methylation of certain genomic regions by their interaction with other nuclear proteins or DNA-binding factors. This idea is supported by two recent studies: purification of a DNMT1 complex that contains the retinoblastoma (*Rb*) gene product, E2F1 and histone deacetylase 1 (HDAC1)[36]; and yeast two-hybrid experiments that show that DNMT1 can form a complex with HDAC2 and the co-repressor proteins DMAP1 and tumour susceptibility gene 101 (TSG101)[37].

DEAMINATION
(of cytosine) The reaction of a water molecule with the amino-group on position 4 of the pyrimidine ring of cytosine, which results in the conversion of cytosine to uracil.

RETROTRANSPOSON
A mobile genetic element; its DNA is transcribed into RNA, which is reverse-transcribed into DNA and then is inserted into a new location in the genome.

L1 OR LINE ELEMENT
Long interspersed sequences generated by RNA polymerase II transcripts.

ALU OR SINE ELEMENT
Short interspersed sequences generated by RNA polymerase III transcripts. An Alu is one of several different SINEs and it requires factors encoded by other retrotransposons (reverse transcriptase) to proliferate.

INTRACISTERNAL A PARTICLE (IAP)
A mouse L1-like element.

## Functions of DNA methylation

*Genome defence and structural integrity.* Much attention in the methylation field has focused on CpG islands, primarily because of the propensity of such sequences to become aberrantly hypermethylated in tumours, resulting in the transcriptional silencing of the associated gene[7,8]. Most CpG dinucleotides, however, reside in the relatively CpG-poor 'bulk' genomic DNA and are hypermethylated[13]. A closer examination of the distribution of this fraction of CpG dinucleotides within the genome has revealed that most reside within parasitic DNA elements or RETROTRANSPOSONS, such as endogenous retroviruses, L1 ELEMENTS and ALU ELEMENTS, which are quite CpG-rich[38,39] (FIG. 1a). Such parasitic DNA elements account for almost 40% of the human genome (exons account for about 5%) and it has been proposed that DNA methylation may have arisen as a genome-defence system to silence expression of these elements and limit their spread through the genome[38]. Parasitic DNA elements represent a significant threat to the structural integrity of the genome because they can mediate recombination between non-allelic repeats, which causes chromosome rearrangements or translocations, and active retrotransposons can integrate into and disrupt genes (FIG. 1b)[40,41]. Many retrotransposons contain strong promoters that, if integrated within a transcriptional unit, could result in internal initiation. If integrated in the 'sense' orientation this could produce a truncated transcript. Conversely, if integrated in the 'antisense' orientation relative to the normal direction of transcription of the targeted gene, this could could inhibit gene expression by transcriptional interference or an antisense mechanism (FIG. 1c)[38]. Expression of genes encoded by retrotransposons, such as reverse transcriptase, is essential for their mobility, and methylation of retrotransposon promoters has been shown to silence their transcription[42].

Evidence supporting the genome-defence hypothesis, although indirect, comes from several sources. *Dnmt1* homozygous knockout ES cells, which retain only 30% of normal methylation levels, have a tenfold increase in the rate of mutations involving gene rearrangements[43]. These cells also show a large increase in transcription from an endogenous transposable element (INTRACISTERNAL A PARTICLES) compared with wild-type ES cells, and it has been proposed that this massive increase in transcription may cause the observed increase in genomic instability[44]. Furthermore, interspecific kangaroo hybrids show a massive expansion of an endogenous retrovirus-like element concomitant with marked hypomethylation of the hybrid genome[45]. Although the mechanism for the failure of methylation in these hybrids is unclear, the finding supports the idea that DNA methylation suppresses expansion of parasitic DNA elements. Further evidence comes from human tumour samples in which global hypomethylation of the genome is a common event, generally

**Box 1 | Chromatin assembly**

- Chromatin is assembled from arrays of nucleosomes, each of which contains 180–200 base pairs (bp) of DNA wrapped around the histone proteins.
- Acetylated histones are normally enriched for transcriptionally competent chromatin.
- Transcriptionally silent chromatin is mostly deacetylated and may be methylated as well.
- Methyl-CpG binding proteins (MBDs) bind to methylated regions, recruit histone deacetylase, and mediate transcriptional silencing by deacetylating histones in the vicinity of the promoter (FIG. 4).
- DNA methyltransferase itself interacts with histone deacetylase, so methylation patterns may dictate histone acetylation patterns.

occurring in conjunction with the more widely studied CpG island hypermethylation[7]. Tumour cells have less methylation than normal cells, and this loss appears to occur primarily from parasitic and repetitive DNAs, which are usually heavily methylated. In conjunction with the loss of methylation is a region-specific gain of methylation at CpG islands that are normally unmethylated. The reason for the loss of methylation in tumours is not known, but its effects on endogenous transposable elements are similar to those seen in *Dnmt1* knockout ES cells: these elements become demethylated and begin to re-express[46,47]. The reactivation of many strong promoters might globally alter transcription patterns by altering transcription factor levels or by negatively affecting specific growth-regulatory genes in which the reactivated elements reside.

In terms of genome integrity, DNA methylation might stabilize the genomes of organisms that contain large amounts of repetitive DNA by 'masking' or inhibiting homologous recombination between such repeats[39]. Such recombination can occur in the human population, with deleterious consequences[48–50]. Direct evidence that DNA methylation suppresses homologous recombination has come from work with the fungus *Ascobolus immersus*, in which it was shown that methylation of a known meiotic recombination hot-spot reduced the frequency of crossing-over within this region several-hundred-fold[51]. Furthermore, V(D)J RECOMBINATION in mammalian cells is reduced more than 100-fold when the recombination substrate is methylated[52]. Other evidence that DNA methylation protects the genome against homologous recombination events in mammalian cells is more indirect and includes observations that imprinted regions show significant differences in recombination frequency between male and female meioses[53]; *Dnmt1* knockout ES cells have a tenfold increase in the rate of mutations involving gene rearrangements[43]; and both people with ICF syndrome (in which *DNMT3B* is mutated, as discussed below) and cultured cells treated with 5-AZA-2′-DEOXYCYTIDINE (5-AZA-CDR) show increased numbers of chromosomal translocations[54,55]. How DNA methylation suppresses homologous recombination remains unknown, but potential mechanisms involve masking of the recombination initiation site, destabilization of the recombination intermediate and interference with the assembly of the recombination machinery.

V(D)J RECOMBINATION
A specialized form of recombination that assembles the genes that encode lymphocyte antigen receptors from variable (V), diversity (D) and joining (J) gene segments. DNA double-strand breaks are introduced between the V, D and J segments and DNA repair proteins then join the segments together.

5-AZA-2′-DEOXYCYTIDINE (5-AZA-CDR)
A potent and specific inhibitor of DNA methylation.

HETEROLOGOUS GAL4 DNA-BINDING DOMAIN
A protein that is fused to the DNA-binding domain of the yeast GAL4 protein to determine its effect on transcription.

TFIIB
Transcription factor IIB, a critical component of the basal transcription machinery.

*Transcriptional repression*. The connection between CpG methylation and transcriptional silencing in vertebrates has been recognized for the past twenty years. Nevertheless, definitive biochemical evidence connecting the two has only recently been obtained[56,57]. Early experiments established that local cytosine methylation of a particular sequence could directly interfere with transcription-factor binding[58]. Such interference could not easily account for the wide range of biological phenomena that rely on methylation for the global silencing of a chromosome, as seen in X-chromosome inactivation. In 1992, a transcriptional repressor that selectively recognizes methylated DNA, methyl-CpG binding protein 2 (MECP2), was characterized[59]. MECP2 can be divided into two structural domains: the methyl-CpG-binding domain (MBD), which recognizes a symmetrically methylated-CpG dinucleotide through contacts in the major groove of the double helix[60], and a transcriptional repression domain (TRD), which interacts with several other regulatory proteins (FIG. 3)[61]. The properties of MECP2 have been central to understanding the mechanisms of DNA methylation-dependent silencing and further focus has been placed on MECP2 with the recent discovery that its gene is mutated in individuals with Rett syndrome[62].

MECP2 selectively represses transcription of methylated templates in the absence of an organized chromatin structure[61,63] and its TRD, when tethered to a specific HETEROLOGOUS GAL4 DNA-BINDING DOMAIN, confers transcriptional repression by interacting with TFIIB, a component of the basal transcription machinery[63]. Transcriptional repression *in vitro* might occur as a result of large aggregates of MECP2 forming on the

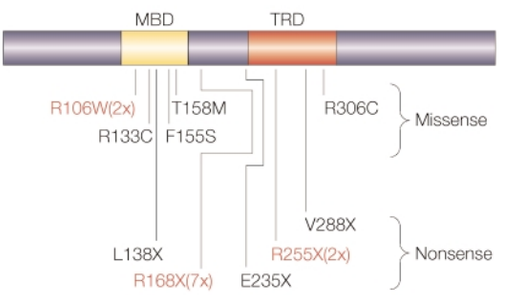


Figure 3 | **Summary of mutations identified in the MECP2 protein in people with Rett syndrome[62,88].** The methyl-CpG binding domain (MBD) and the transcriptional repression domains (TRD) are indicated[100]. All missense mutations involve highly conserved amino acids within defined functional domains. Nonsense mutations would result in truncated proteins (if stable) that would delete some or all of the conserved domains. The functional consequences of all mutations have yet to be investigated *in vitro*. Most *MECP2* mutations arise from C-to-T transitions within CpG dinucleotides owing to the increased mutability of 5-methylcytosine, implying that these sites are methylated *in vivo*. Several mutational hot-spots have also been identified (R106W, R168X and R255X)[88], as shown in red. The bracketed numbers indicate how many times these mutations have been found.

template, preventing the assembly of functional transcription complexes[63]. Similar clustering of MECP2 occurs within nuclei at foci that contain methylated DNA[64]. Although MECP2 can repress transcription on naked DNA templates within the nucleus, it must function in the context of the chromatin infrastructure that assembles the chromosomes (BOX 1). The association of DNA with histones prevents many components of the transcriptional machinery from binding nucleosomal DNA. MECP2 is an exception. The MBD can recognize methylated CpG dinucleotides in the nucleosome[65] and, under certain circumstances, MECP2 can displace histone H1 from chromatin[61]. So MECP2 is ideally placed to modify chromatin structure directly or through the recruitment of other repressive enzymes or structural components (FIG. 4).

Chromatin assembly facilitates the repression of methylated DNA[66,67]. Methyl-CpG binding proteins, including MECP2, associate with co-repressor complexes that include histone deacetylases[56,57,68,69] (BOX 1). Recruitment of a histone deacetylase by MECP2 occurs indirectly through its interaction with the Sin3A adaptor protein, which causes transcriptional silencing, in part by deacetylating the histones[56,57]. This directs the formation of stable repressive chromatin structures (FIG. 4)[70]. Recent findings link the four different methyl-CpG binding domain proteins, MECP2, MBD1, MBD2 and MBD3 with the chromatin-remodelling machinery. In *Xenopus* eggs, MBD3 is a component of the Mi-2 chromatin-remodelling complex, along with the histone deacetylase binding proteins Rpd3 and RbAp46/48, and methyl-CpG binding activity and histone deacetylase activity fractionate together in egg extracts[68]. MBD2, HDAC1, HDAC2 and RbAp46/48 purify together in HeLa cell nuclear extracts and are components of the MECP1 repressor complex[69].

The transcriptional repression mediated by both such complexes is partially sensitive to the histone deacetylase inhibitor TRICHOSTATIN A (TSA)[56,69,71]. MBD1, although not yet identified as a part of a repressor complex, can also act as a transcriptional silencer in a TSA-sensitive manner[71,72]. The importance of methyl-CpG binding proteins to gene regulation has been emphasized by the finding of MBD2- and MBD3-like proteins in *Drosophila*. As *Drosophila* DNA is not methylated, the binding specificity of these MBD homologues for methyl-CpG has not been preserved, but the interaction with histone deacetylases, presumably in the *Drosophila* version of the Mi-2 complex, has remained[30].

Two important additional links between DNA methylation and chromatin structure have recently come to light. First, DNMT1 can interact with histone deacetylase and repress transcription[36,37,73]. The minimal interaction domain (CXXC) within the amino-terminal regulatory domain of DNMT1 mediates transcriptional repression when fused to a heterologous DNA-binding domain, and this repression is partly alleviated by TSA treatment[36,37,73]. This interaction indicates that histone deacetylase or acetylation patterns might target methylation. Conversely, DNMT1 might directly target deacetylation to regions that are to be methylated. In either scenario, methylation and deacetylation would act together to potentiate the repressed state.

The second link between chromatin structure and methylation comes from patients with mutations in a putative ATP-dependent chromatin-remodelling factor of the SNF2 family, termed *ATR-X*. Patients with ATR-X syndrome (for α-thalassaemia, mental retardation, X-linked) have subtle defects in cellular methylation patterns, which include both hypo- and hypermethylation at certain repetitive elements[74], and show some

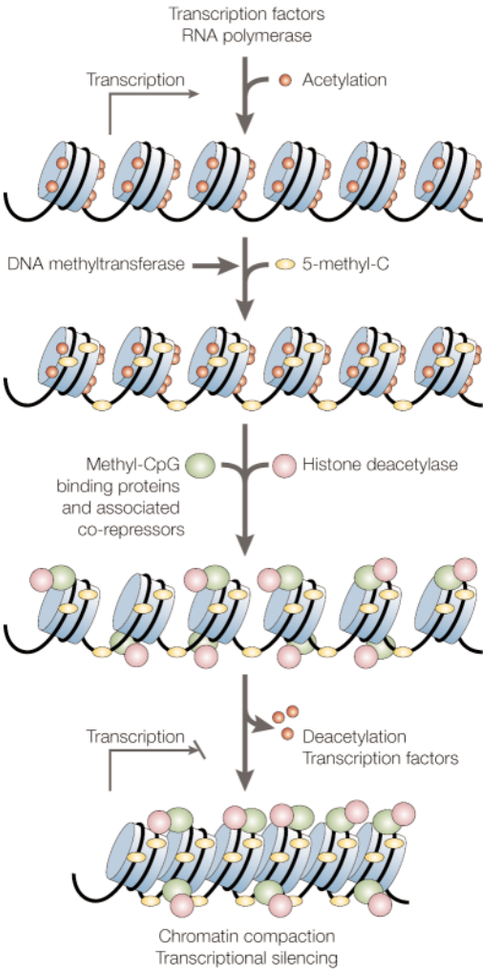


Figure 4 | **The mechanism whereby DNA methylation and histone deacetylation cooperate to repress transcription.** A transcriptionally active region targeted for silencing is proposed to acquire DNA methylation first, which then recruits the methyl-CpG binding proteins and their associated co-repressors and histone deacetylases (HDACs). As DNA methyltransferase 1 (DNMT1) can interact directly with histone deacetylase, it is also possible that transcription is first silenced by deacetylation by other tethering factors, after which the methylation machinery and the methyl-CpG binding proteins are recruited to 'cement' the promoter in the silent state. In either case, the deacetylated nucleosomes adopt a more tightly packed structure that inhibits the access of transcription factors to their binding sites.

*Animated online*

TRICHOSTATIN A (TSA)
A specific inhibitor of histone deacetylase.

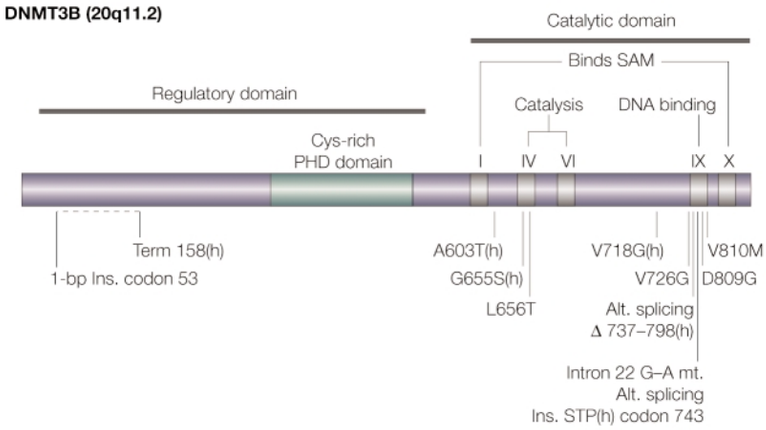


Figure 5 | **Summary of mutations in DNA methyltransferase 3B (DNMT3B) identified in ICF syndrome patients**[3,83,84]**.** Many individuals were compound heterozygotes for two independent mutations (denoted with an h). Nearly all mutations affect the catalytic domain. Loss of catalytic activity was confirmed for the D809G mutation in cell line overexpression studies[83]. The conserved methyltransferase motifs within the catalytic domain are indicated with roman numerals. The cysteine-rich PHD (plant homeodomain) region in the N terminus is highly homologous to a PHD motif within the α-thalassaemia, mental retardation, X-linked (ATR-X) protein[85]. The *S*-adenosyl-L-methionine (SAM) binding domain is indicated, as well as motifs involved in catalysis (IV, VI) and DNA binding (IX). Several mutations (mt.) result in altered splicing patterns (Alt.) and the insertion (Ins.) or deletion (Δ) of coding sequence.

phenotypic similarity to mutations in a similar SNF2-like protein in *Arabidopsis*, called ddm1 (REF. 75). Mutations in the *ddm1* gene result in a roughly 70% reduction in 5-methylcytosine levels in *Arabidopsis*, and this loss occurs primarily at repetitive elements. Both of these observations further emphasize the functionally important and highly conserved relationship between chromatin remodelling and DNA methylation.

### DNA methylation and disease

Several genetic diseases have been described that cause methylation defects, including the ICF, Rett and fragile X syndromes. One common aspect to all these diseases is a variable degree of mental impairment, implicating DNA methylation-dependent gene control pathways as being particularly important for brain development. Interestingly, patients with ATR-X syndrome, who frequently show severe mental retardation, also have methylation defects, although the exact function of the ATR-X protein remains unknown[74]. It has been reported that DNA methyltransferase activity is high in neurons despite their terminally differentiated state[76] and that DNA methyltransferase activity might contribute to induced ischaemic brain damage in mice[77]. Here, we summarize what is known about the three best characterized genetic diseases of DNA methylation (ICF, Rett and fragile X syndromes).

*ICF syndrome and DNMT3B.* ICF syndrome is a rare autosomal recessive disease. Patients show variable immunodeficiency that consists of an absence or severe reduction in at least two immunoglobulin isotypes, and some, but not all, have defective cell-mediated immunity. Developmental defects include delayed developmental milestones, mental retardation and peculiar facial features such as roundness, widely spaced eyes (hypertelorism), flat nasal bridge, small jaw (micrognathia), and an enlarged, protruding tongue (macroglossia). Furthermore, ICF patients often suffer from severe respiratory tract infections[78,79]. Perhaps the most remarkable characteristic of ICF syndrome is the marked elongation of centromeric or juxtacentromeric heterochromatin in phytohaemagglutinin (PHA)-stimulated lymphocytes from such patients. These same cells also have multibranched configurations (multiradials), deletions or duplication of entire chromosome arms, isochromosomes and centromeric breakage[79]. These abnormalities occur almost exclusively on chromosomes 1, 9 and 16, which contain particularly large blocks of alphoid and classical satellite long-tandem-repeat arrays[80]. These regions are normally heavily methylated in somatic cells but ICF patients show marked hypomethylation of such regions[80,81], indicating that DNA methylation is essential for proper centromeric structure and stability. Repetitive elements elsewhere in the genome[82] and single-copy sequences on the inactive X chromosome[54] are also hypomethylated in ICF cells.

Three studies in 1999 linked the observed methylation defect in ICF syndrome to mutations in the *DNMT3B* gene[3,83,84], which had been mapped to 20q11.2 (REFS 24,85), the ICF susceptibility locus[86]. Several mutations that have been identified in *DNMT3B* are summarized in FIG. 5. Although no 'hotspot' has been identified and many of the mutations are heterozygous, all seem to affect the carboxy-terminal catalytic domain of DNMT3B. Interestingly, mice with a homozygous knockout of the *Dnmt3b* gene have similar patterns of pericentromeric demethylation and may serve as a model for ICF syndrome[3]. Given that complete loss of *Dnmt3b* function causes embryonic lethality in mice and that most of the ICF mutations leave the N-terminal regulatory domain of DNMT3B intact, this region of the protein may be critical for survival, and mutations in, or the complete loss of, the N-terminal domain may be lethal to human embryos. Alternatively, some of the mutations occurring in the catalytic domain may retain some residual enzymatic activity, although this has been ruled out for a small subset of such mutations using overexpression studies in cell lines[83].

It is not yet known how a defect in centromeric methylation contributes to improper brain development or how DNMT3B is targeted to centromeres. It is probable that targeting is accomplished by the interaction of DNMT3B with another DNA-binding protein whose binding sites lie in pericentromeric heterochromatin. Improper centromeric chromosome structure may globally alter gene transcription in subtle ways and this may affect the development of the brain more than that of other organs simply because of its extreme complexity. Alternatively, DNMT3B may be required for the proper methylation, and therefore proper gene expression patterns and chromosomal structure, of a select group of unidentified genes required for normal brain development.

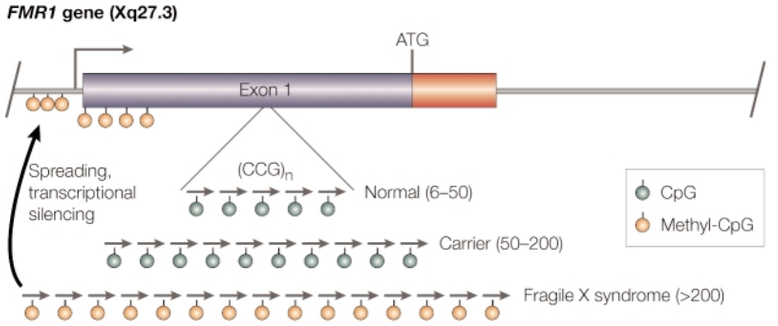


Figure 6 | **Schematic of the fragile X mental retardation 1 (*FMR1*) gene showing the location of the CGG repeat within the 5′-untranslated region of exon 1.** Normal people have between 6 and 50 CGG repeats and the region is unmethylated. Fragile X carriers show an expanded repeat of between 50 and 200 copies but the region remains unmethylated and the gene is expressed. In fragile X syndrome patients, the repeat has expanded further and become methylated *de novo*. This methylation extends into the promoter region of *FMR1* as well (denoted by the bent arrow) and silences transcription[95,101].

*Rett syndrome and MECP2*. Rett syndrome is an X-linked dominant disorder and is one of the most common causes of sporadic mental retardation in females (incidence of 1 in 10,000–15,000). It is characterized by a period of normal development, followed by progressive degeneration in speech and acquired motor skills as well as seizures, autism, loss of motor coordination (ataxia) and stereotypical, repetitive hand movements. The condition often stabilizes after the initial regression period and patients usually survive into adulthood[87]. In 1999, the molecular defect in a large fraction of Rett syndrome patients was found to result from mutations in the *MECP2* gene, located on human chromosome Xq28 (REF. 62) and further mutations have since been described[88]. Most mutations are missense or truncating mutations that affect the integrity of the two main domains of MECP2, the MBD and the TRD (FIG. 3). Several mutational hot-spots at CpG dinucleotides, involving C-to-T transitions, have been identified[88].

Four of the Rett-associated mutations within the MBD, described by Amir *et al.*[62], have been introduced into the *Xenopus* MECP2 protein. Three of the mutations severely reduced or abolished binding to a methylated probe in *in vitro* binding assays[89]. Chimaeric mice, in which the fraction of *Mecp2*-deficient ES cells is low, do develop to term although they show developmental defects[90]. Such mice may be useful as models of Rett syndrome, although their defects need to be further characterized. Results from mouse models imply that the Rett-associated *Mecp2* mutations do not cause complete loss of function, an idea also supported by the marked skewing in sex of affected offspring: Rett females are heterozygous for the *Mecp2* mutation and mosaic owing to X-chromosome inactivation, and so some show skewed patterns of X-chromosome inactivation; conversely, the hemizygous condition in males is lethal[91].

How exactly mutations in *MECP2* lead to developmental defects in the brain remains unclear, but it is probably linked to the ability of MECP2 to silence transcription. MECP2 is in fact more abundant in the brain than in any other tissue examined[61]. One of the proposed functions of DNA methylation in vertebrates is to reduce 'transcriptional noise', and the methyl-CpG binding protein family members are excellent candidates for such a role[92]. Perhaps brain development is more sensitive to such 'noise' than other tissues and brain tissue is therefore more dependent on MECP2 to carry out this global silencing. Alternatively, *MECP2* mutations may alter transcription of specific genes necessary for brain development. Such issues may be resolved by examining global patterns of gene expression in Rett cells, in addition to determining whether mutations in other methyl-CpG binding proteins are responsible for the Rett cases in which no mutations in the *MECP2* coding-region were identified.

*Fragile X syndrome and* FMR1. Fragile X syndrome is the most common form of inherited mental retardation (incidence of around 1 in 5,000 males) after Down syndrome, affects primarily males, and has been mapped to human chromosome Xq27.3. Additional features of the disease can include a long face, large everted ears, autism, hand biting, hyperactivity and enlarged testicles (macro-orchidism)[93]. The mode of inheritance of fragile X syndrome is unusual in that the gene becomes penetrant only when it is maternally transmitted, and the chance of penetrance and the severity of the disease increase with successive generations. Several reports in 1991 identified the molecular defect responsible for fragile X syndrome by the cloning and characterization of the *FMR1* gene (fragile X mental retardation-1)[94,95].

*FMR1* contains a highly polymorphic CGG repeat within the 5′-untranslated region of exon 1 with an average length of 29 repeats (within a range of 6–52 repeats) in normal people (FIG. 6). In fragile X patients, the repeat length increases markedly to 200–600 or more copies. Concomitant with this expansion is aberrant *de novo* methylation and histone deacetylation of the CpG island upstream of the gene and silencing of *FMR1* transcription (FIG. 6)[94–96]. Carriers show an intermediate repeat length (50–200 copies) but lack the methylation of the upstream CpG island[97]. The mechanism of the expansion is believed to result from polymerase slippage during DNA replication. The mechanism driving *de novo* methylation of the CpG island is not clear but may result from the potential of the repeats to form hairpin structures. Such hairpin loops and other 'odd' DNA structures are good substrates for DNMT1 (REF. 98). Alternatively, the *de novo* methylation may in some way be related to the genome defence/ repetitive element methylation system in that the expanded repeat may begin to resemble or adopt one or more of the features of parasitic elements that target them for methylation. Interestingly, treatment of fragile X cells with 5-Aza-CdR results in re-expression of *FMR1*, whereas treatment with TSA does not, indicating that methylation is the dominant silencing mechanism[96]. The FMR1 protein is abundant in neurons, contains several RNA-binding motifs, associates with translating ribosomes

in an RNA-dependent manner and it has been proposed to be important in protein synthesis in neurons[93,99]. Lack of proper protein expression in neurons during development could therefore give rise to the observed neurological defects.

### Future directions

The past few years have seen an explosion in knowledge of how DNA methylation affects such diverse processes as transcriptional regulation, chromatin structure, genome stability and tumorigenesis. Of critical importance will be to determine the exact functions, target-site specificity, nuclear localization and interaction partners of each of the DNA methyltransferases. Recent evidence also indicates that more DNA methyltransferases may exist and their identification will no doubt be facilitated by the imminent completion of the human genome sequencing project. The identification of defined molecular defects in DNA methylation in several neurodevelopmental disorders combined with a huge increase in our knowledge of the methylation machinery makes this an opportune time to begin exploring methods to correct these defects. Furthermore, re-examination of the roles of DNA methylation defects in other neurodevelopmental disorders with less defined origins could be highly informative. Both DNA methylation and histone deacetylation are reversible modifications, and inhibitors of each process exist. What will be required, however, is an improvement in the potency and specificity of the current agents, and a focus on the design of other small-molecule inhibitors. Improvements in antisense and gene therapy procedures may also allow correction of the molecular defects in these diseases. Given the current pace of research in the field of epigenetics, it is likely that great strides will continue to be made towards these goals in the next decade, which may revolutionize how we think about and treat the disease process.

### Links

DATABASE LINKS ICF syndrome | Rett syndrome | Fragile X syndrome | TP53 | DNMT1 | DNMT3A | DNMT3B | DNMT1-like protein in *Drosophila* | DNMT2 homologue in *Drosophila* | Dnmt3a | Dnmt3b | E2F1 | HDAC1 | HDAC2 | DMAP1 | TSG101 | MECP2 | MBD1 | MBD2 | MBD3 | RbAp46/48 | MECP1 | MBD2/3 in *Drosophila* | ATRX gene | ATRX syndrome | FMR1 gene

FURTHER INFORMATION Histone sequence database | DNA methylation society


1. Lei, H. *et al. De novo* DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122,** 3195–3205 (1996).
2. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
3. Okano, M., Bell, D. W., Haber, D. A. & Li, W. DNA methyltransferases *Dnmt3a* and *Dnmt3b* are essential for *de novo* methylation and mammalian development. *Cell* **99**, 247–257 (1999).
**References 2 and 3 describe the effects of loss of *Dnmt1*, *Dnmt3a* and *Dnmt3b* on mouse development.**
4. Bird, A. P. Gene number, noise reduction and biological complexity. *Trends Genet.* **11**, 94–99 (1995).
5. Steinbach, O. C., Wolffe, A. P. & Rupp, R. A. Somatic linker histones cause loss of mesodermal competence in *Xenopus*. *Nature* **389**, 395–399 (1997).
6. Mannervik, M., Nibu, Y., Zhang, H. & Levine, M. Transcriptional coregulators in development. *Science* **284**, 606–609 (1999).
7. Baylin, S. B., Herman, J. G., Herman, J. R., Vertino, P. M. & Issa, J.-P. Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv. Cancer Res.* **72**, 141–196 (1998).
8. Jones, P. A. & Laird, P. W. Cancer epigenetics comes of age. *Nature Genet.* **21**, 163–166 (1999).
**References 7 and 8 are good reviews on the roles of DNA methylation in cancer.**
9. Jaenisch, R. DNA methylation and imprinting: why bother? *Trends Genet.* **13**, 323–329 (1997).
10. Jirtle, J. L., Sander, M. & Barrett, J. C. Genomic imprinting and environmental disease susceptibility. *Environ. Health Perspect.* **108**, 271–278 (2000).
11. Tyko, B. & Ashkenas, J. Epigenetics and its role in disease. *J. Clin. Invest.* **105**, 245–246 (2000).
12. Post, W. S. *et al.* Methylation of the estrogen receptor gene is associated with aging and atherosclerosis in the cardiovascular system. *Cardiovasc. Res.* **43**, 985–991 (1999).
13. Cooper, D. N. & Krawczak, M. Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. *Hum. Genet.* **83**, 181–188 (1989).
14. Bird, A., Taggart, M., Frommer, M., Miller, O. J. & Macleod, D. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* **10**, 91–99 (1985).
15. Antequera, F. & Bird, A. Number of CpG islands and genes in human and mouse. *Proc. Natl Acad. Sci. USA* **90**, 11995–11999 (1993).
16. Tazi, J. & Bird, A. Alternative chromatin structure at CpG islands. *Cell* **60**, 909–920 (1990).
17. Jiricny, J. in *Cancer Surveys: Genetic Instability in Cancer* Vol. 28 47–68 (Imperial Cancer Research Fund, 1996).
18. Cooper, D. N. & Youssoufian, H. The CpG dinucleotide and human genetic disease. *Hum. Genet.* **78**, 151–155 (1988).
19. Rideout, W. M. I., Coetzee, G. A., Olumi, A. F. & Jones, P. A. 5-methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science* **249**, 1288–1290 (1990).
20. Greenblatt, M. S., Bennett, W. P., Hollstein, M. & Harris, C. C. Mutations in the p53 tumor suppressor gene: Clues to cancer etiology and molecular pathogenesis. *Cancer Res.* **54**, 4855–4878 (1994).
21. Bestor, T., Laudano, A., Mattaliano, R. & Ingram, V. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. *J. Mol. Biol.* **203**, 971–983 (1988).
22. Pradhan, S., Bacolla, A., Wells, R. D. & Roberts, R. J. Recombinant human DNA (cytosine-5) methyltransferase I. Expression, purification, and comparison of *de novo* and maintenance methylation. *J. Biol. Chem.* **274**, 33002–33010 (1999).
23. Pradhan, S. *et al.* Baculovirus-mediated expression and characterization of the full-length murine DNA methyltransferase. *Nucleic Acids Res.* **25**, 4666–4673 (1997).
24. Robertson, K. D. *et al.* The human DNA methyltransferases (DNMTs) 1, 3a, and 3b: Coordinate mRNA expression in normal tissues and overexpression in tumors. *Nucleic Acids Res.* **27**, 2291–2298 (1999).
25. Leonhardt, H., Page, A. W., Weier, H. & Bestor, T. H. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71**, 865–873 (1992).
26. Chuang, L. S.-H. *et al.* Human DNA-(cytosine-5) methyltransferase-PCNA complex is a target for p21$^{Waf1}$. *Science* **277**, 1996–2000 (1997).
27. Li, E., Beard, C. & Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature* **366**, 362–365 (1993).
28. Beard, C., Li, E. & Jaenisch, R. Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes Dev.* **9**, 2325–2334 (1995).
29. Hung, M.-S. *et al. Drosophila* proteins related to vertebrate DNA (5-cytosine) methyltransferases. *Proc. Natl Acad. Sci. USA* **96**, 11940–11945 (1999).
30. Tweedie, S. *et al.* Vestiges of DNA methylation system in *Drosophila melanogaster*. *Nature Genet.* **23**, 389–390 (1999).
31. Okano, M., Xie, S. & Li, E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genet.* **19**, 219–220 (1998).
32. Cao, X. *et al.* Conserved plant genes with similarity to mammalian *de novo* DNA methyltransferases. *Proc. Natl Acad. Sci. USA* **97**, 4979–4984 (2000).
33. Lyko, L. *et al.* Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in Drosophila. *Nature Genet.* **23**, 363–366 (1999).
34. Vertino, P. M., Yen, R.-W. C., Gao, J. & Baylin, S. B. De novo methylation of CpG island sequences in human fibroblasts overexpressing DNA (cytosine-5)-methyltransferase. *Mol. Cell. Biol.* **16**, 4555–4565 (1996).
**This paper uniquely describes the use of somatic cell knockout technology to delete the *DNMT1* gene in a cancer cell line.**
35. Rhee, I. *et al.* CpG methylation is maintained in human cancer cells lacking DNMT1. *Nature* **404**, 1003–1007 (2000).
36. Robertson, K. D. *et al.* DNMT1 forms a complex with Rb, E2F1, and HDAC1 and represses transcription from E2F-responsive promoters. *Nature Genet.* **25**, 338–342 (2000).
37. Roundtree, M. R., Bachman, K. E. & Baylin, S. B. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature Genet.* **25**, 269–277 (2000).
38. Yoder, J. A., Walsh, C. P. & Bestor, T. H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**, 335–340 (1997).
39. Colot, V. & Rossignol, J.-L. Eukaryotic DNA methylation as an evolutionary device. *BioEssays* **21**, 402–411 (1999).
40. Montagna, M. *et al.* Identification of a 3 kb Alu-mediated BRCA1 gene rearrangement in two breast/ovarian cancer families. *Oncogene* **18**, 4160–4165 (1999).
41. Kazazian, J. H. H. & Moran, J. V. The impact of L1 retrotransposons on the human genome. *Nature Genet.* **19**, 19–24 (1998).
42. Kochanek, S., Renz, D. & Doerfler, W. Transcriptional silencing of human Alu sequences and inhibition of protein binding in the B box regulatory elements by 5'-CG-3' methylation. *FEBS Lett.* **360**, 115–120 (1995).
43. Chen, R. Z., Pettersson, U., Beard, C., Jackson-Grusby, L. & Jaenisch, R. DNA hypomethylation leads to elevated mutation rates. *Nature* **395**, 89–93 (1998).
44. Walsh, C. P., Chaillet, J. R. & Bestor, T. H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nature Genet.* **20**, 116–117 (1998).
45. Waugh-O'Neill, R. J., O'Neill, M. J. & Marshall-Graves, J. A. Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* **393**, 68–72 (1988).
46. Flori, A. R., Lower, R., Schmitz-Drager, B. J. & Schulz, W. A. DNA methylation and expression of LINE-1 and HERV-K provirus sequences in urothelial and renal cell carcinomas. *Br. J. Cancer* **80**, 1312–1321 (1999).
47. Grassi, M., Girault, J. M., Wang, W. P., Thiery, J. P. & Jouanneau, J. Metastatic rat carcinoma cells express a new retrotransposon. *Gene* **233**, 59–66 (1999).
48. Puget, N. *et al.* A 1-kb Alu-mediated germ-line deletion removing *BRCA1* exon 17. *Cancer Res.* **57**, 828–831 (1997).
49. Rouyer, F., Simmler, M. C., Page, D. & Weissenbach, J. A sex chromosome rearrangement in a human XX male caused by Alu-Alu recombination. *Cell* **51**, 417–425 (1987).
50. Small, K., Iber, J. & Warren, S. T. Emerin deletion reveals a common X chromosome inversion mediated by inverted repeats. *Nature Genet.* **16**, 96–99 (1997).
51. Maloisel, L. & Rossignol, J.-L. Suppression of crossing-over

by DNA methylation in *Ascobolus*. *Genes Dev.* **12**, 1381–1389 (1998).
52. Hsieh, C.-L. & Lieber, M. R. CpG methylated minichromosomes become inaccessible for V(D)J recombination after undergoing replication. *EMBO J.* **11**, 3115–325 (1992).
53. Paldi, A., Gyapay, G. & Jami, J. Imprinted chromosomal regions of the human genome display sex-specific meiotic recombination frequencies. *Curr. Biol.* **5**, 1030–1035 (1995).
54. Miniou, P. *et al.* Abnormal methylation pattern in constitutive and facultative (X inactive chromosome) heterochromatin of ICF patients. *Hum. Mol. Genet.* **3**, 2093–2102 (1994).
55. Ji, W. *et al.* DNA demethylation and pericentromeric rearrangements of chromosome 1. *Mutat. Res.* **379**, 33–41 (1997).
56. Jones, P. L. *et al.* Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genet.* **19**, 187–191 (1998).
57. Nan, X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 invloves a histone deacetylase complex. *Nature* **393**, 386–389 (1998).
**References 56 and 57 were the first reports to link DNA methylation to methyl-CpG binding proteins and chromatin-remodelling factors, such as histone deacetylase.**
58. Tate, P. H. & Bird, A. P. Effects of DNA methylation on DNA-binding proteins and gene expression. *Curr. Opin. Genet. Dev.* **3**, 226–231 (1993).
59. Lewis, J. D. *et al.* Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69**, 905–914 (1992).
60. Wakefield, R. I. D. *et al.* The solution structure of the domain from MeCP2 that binds to methylated DNA. *J. Mol. Biol.* **291**, 1055–1065 (1999).
61. Nan, X., Campoy, F. J. & Bird, A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**, 471–481 (1997).
**This paper was the first to report that Rett syndrome was associated with mutations in the *MeCP2* gene.**
62. Amir, R. E. *et al.* Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein. *Nature Genet.* **23**, 185–188 (1999).
63. Kaludov, N. & Wolffe, A. P. MeCP2 driven transcriptional repression *in vitro*: selectivity for methylated DNA, action at a distance and contacts with the basal transcription machinery. *Nucleic Acids Res.* **28**, 1921–1928 (2000).
64. Nan, X., Tate, P., Li, E. & Bird, A. DNA methylation specifies chromosomal localization of MeCP2. *Mol. Cell. Biol.* **10**, 414–421 (1996).
65. Chandler, S. P., Guschin, D., Landsberger, N. & Wolffe, A. P. The methyl CpG binding transcriptional repressor MeCP2 stably associates with nucleosomal DNA. *Biochemistry* **38**, 7008–7018 (1999).
66. Buschhausen, G., Wittig, B., Graessmann, M. & Graessmann, A. Chromatin structure is required to block transcription from the methylated herpes simplex virus thymidine kinase gene. *Proc. Natl Acad. Sci. USA* **84**, 1177–1181 (1986).
67. Kass, S. U., Landsberger, N. & Wolffe, A. P. DNA methylation directs a time-dependent repression of transcription initiation. *Curr. Biol.* **7**, 157–165 (1997).
68. Wade, P. A. *et al.* The Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nature Genet.* **23**, 62–66 (1999).
69. Ng, H.-H. *et al.* MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nature Genet.* **23**, 58–61 (1999).
70. Tse, C., Sera, T., Wolffe, A. P. & Hansen, J. C. Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Mol. Cell. Biol.* **18**, 4629–4638 (1998).
71. Ng, H.-H., Jeppesen, P. & Bird, A. Active repression of methylated genes by the chromosomal protein MBD1. *Mol. Cell. Biol.* **20**, 1394–1406 (2000).
72. Fujita, N. *et al.* Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms. *Mol. Cell. Biol.* **19**, 6415–6426 (1999).
73. Fuks, F., Bergers, W. A., Brehm, A., Hughes-Davies, L. & Kouzarides, T. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nature Genet.* **24**, 88–91 (2000).
74. Gibbons, R. J. *et al.* Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the patterns of DNA methylation. *Nature Genet.* **24**, 368–371 (2000).
75. Jeddeloh, J. A., Stokes, T. L. & Richards, E. J. Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nature Genet.* **22**, 94–97 (1999).
76. Goto, K. *et al.* Expression of DNA methyltransferase gene in mature and immature neurons as well as proliferating cells in mice. *Differentiation* **56**, 39–44 (1993).
77. Endres, M. *et al.* DNA methyltransferase contributes to delayed ischemic brain injury. *J. Neurosci.* **20**, 3175–3181 (2000).
78. Smeets, D. F. C. M. *et al.* ICF syndrome: a new case and review of the literature. *Hum. Genet.* **94**, 240–246 (1994).
79. Franceschini, P. *et al.* Variability of clinical and immunological phenotype in immunodeficiency-centromeric instability-facial anamolies syndrome. *Eur. J. Pediatr.* **154**, 840–846 (1995).
80. Tagarro, I., Fernandez-Peralta, A. M. & Gonzales-Aguilera, J. J. Chromosomal localization of human satellites 2 and 3 by FISH method using oligonucleotides as probes. *Hum. Genet.* **93**, 383–388 (1994).
81. Jeanpierre, M. *et al.* An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Hum. Mol. Genet.* **2**, 731–735 (1993).
82. Kondo, T. *et al.* Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats *D4Z4* and *NBL2*. *Hum. Mol. Genet.* **9**, 597–604 (2000).
83. Xu, G.-L. *et al.* Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402**, 187–191 (1999).
84. Hansen, R. S. *et al.* The *DNMT3B* DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc. Natl Acad. Sci. USA* **96**, 14412–14417 (1999).
85. Xie, S. *et al.* Cloning, expression and chromosome locations of the human DNMT3 gene family. *Gene* **236**, 87–95 (1999).
**References 83–85 were the first to report that ICF syndrome is associated with mutations in the *DNMT3B* gene.**
86. Wijmenga, C. *et al.* Localization of the ICF syndrome to chromosome 20 by heterozygosity mapping. *Am. J. Hum. Genet.* **63**, 803–809 (1998).
87. Hagberg, B., Aicardi, J., Dias, K. & Ramos, O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann. Neurol.* **14**, 471–479 (1983).
88. Wan, M. *et al.* Rett syndrome and beyond: recurrent spontaneous familial *MECP2* mutations at CpG hotspots. *Am. J. Hum. Genet.* **65**, 1520–1529 (2000).
89. Ballestar, E., Yusufzai, T. M. & Wolffe, A. P. The effects of Rett syndrome mutations of the methyl-CpG binding domain of the transcriptional repressor MeCP2 on selectivity for association with methylated DNA. *Biochemistry* **39**,7100–7106 (2000).
90. Tate, P., Skarnes, W. & Bird, A. The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. *Nature Genet.* **12**, 205–208 (1996).
91. Willard, H. F. & Hendrich, B. D. Breaking the silence in Rett syndrome. *Nature Genet.* **23**, 127–128 (1999).
92. Bird, A. & Tweedie, S. Transcriptional noise and the evolution of gene number. *Phil. Trans. R. Soc. Lond.* **349**, 249–253 (1995).
93. Warren, S. T. & Nelson, D. L. Trinucleotide repeat expansions in neurological disease. *Curr. Opin. Neurobiol.* **3**, 752–759 (1993).
94. Kremer, E. J. *et al.* Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science* **252**, 1711–1714 (1991).
95. Oberle, I. *et al.* Instability of a 550-base pair segment and abnormal methylation in Fragile X Syndrome. *Science* **252**, 1097–1102 (1991).
**References 94 and 95 were among the first to show that lack of expression of the *FMR1* gene was associated with abnormal expansion and methylation of a trinucleotide repeat.**
96. Coffee, B., Zhang, F., Warren, S. T. & Reines, D. Acetylated histones are associated with *FMR1* in normal but not fragile X-syndrome cells. *Nature Genet.* **22**, 98–101 (1999).
97. Fu, Y. H. *et al.* Variation of the CGG repeat at the Fragile X site results in genetic instability: Resolution of the Sherman paradox. *Cell* **67**, 1047–1058 (1991).
98. Smith, S. S., Laayoun, A., Lingeman, R. G., Baker, D. J. & Riley, J. Hypermethylation of telomere-like foldbacks at codon 12 of the human c-Ha ras gene and the trinucleotide repeat of the *FMR-1* gene of fragile X. *J. Mol. Biol.* **243**, 143–151 (1994).
99. Feng, Y. *et al.* FMRP associates with polyribosomes as an mRNP, and the 1304N mutation of severe fragile X syndrome abolishes this association. *Mol. Cell* **1**, 109–118 (1997).
100. Hendrich, B. & Bird, A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.* **10**, 6538–6547 (1998).
101. Hornstra, I. K., Nelson, D. L., Warren, S. T. & Yang, T. P. High resolution methylation analysis of the FMR1 gene trinucleotide repeat region in fragile X syndrome. *Hum. Mol. Genet.* **2**, 1659–1665 (1993).