The diverse roles of DNA methylation in mammalian development and disease

*Maxim V. C. Greenberg and Deborah Bourc'his**

Abstract | DNA methylation is of paramount importance for mammalian embryonic development. DNA methylation has numerous functions: it is implicated in the repression of transposons and genes, but is also associated with actively transcribed gene bodies and, in some cases, with gene activation per se. In recent years, sensitive technologies have been developed that allow the interrogation of DNA methylation patterns from a small number of cells. The use of these technologies has greatly improved our knowledge of DNA methylation dynamics and heterogeneity in embryos and in specific tissues. Combined with genetic analyses, it is increasingly apparent that regulation of DNA methylation erasure and (re-)establishment varies considerably between different developmental stages. In this Review, we discuss the mechanisms and functions of DNA methylation and demethylation in both mice and humans at CpG-rich promoters, gene bodies and transposable elements. We highlight the dynamic erasure and re-establishment of DNA methylation in embryonic, germline and somatic cell development. Finally, we provide insights into DNA methylation gained from studying genetic diseases.

Pericentromeric satellite repeats
Tandem repeats enriched in heterochromatin modifications such as DNA methylation and histone H3 Lys9 trimethylation.

The methylation of the fifth carbon of cytosines (5-methylcytosine (5mC)) originated in bacteria and was present in the first eukaryote[1]. As eukaryotic 5mC is mostly found in the context of symmetrical CpG dinucleotides[1–3], it was anticipated decades ago that a mechanism was in place to recognize the hemimethylated CpG site after DNA replication and faithfully methylate the daughter strand[4]; elegant experiments have subsequently validated this hypothesis[5]. Early studies in vitro and in vivo indicated that 5mC was associated with transcriptional repression[6–8]. Accordingly, DNA methylation has since been implicated in the classical epigenetic phenomena of genomic imprinting[9–12] and X-chromosome inactivation (XCI)[13,14] (see below).

Despite its ancient origins, DNA cytosine methylation has been lost in several eukaryotic lineages, including in many animals[15,16]; common model organisms such as *Drosophila melanogaster*, *Caenorhabditis elegans*, fission yeasts and bakers' yeasts exhibit virtually no 5mC. In fact, cytosine methylation comes at a cost: 5mC is inherently mutagenic because it can spontaneously undergo deamination, leading to C → T transitions[17]. Thus, organisms with CpG methylation also have reduced CpG content[18,19]. For example, mammals have roughly 5-fold fewer CpG dinucleotides than expected from the nucleotide composition of their genome. Furthermore, DNA methyltransferases (DNMTs) were revealed to introduce toxic 3-methylcytosine lesions into DNA[20]. The systematic co-evolution of DNMTs with a specific alkylation repair enzyme (ALKBH2) may have allowed eukaryotes to tolerate DNA methylation.

Nevertheless, mammalian genomes exhibit particularly high CpG methylation levels; although there are some tissue-specific differences, 70–80% of CpGs are methylated[21]. Moreover, DNMT-deficient mice exhibit severe developmental abnormalities, culminating in early embryonic lethality[22,23]. Deregulation of DNA methylation is also a defining feature of virtually all cancer types[24]. In addition to XCI and genomic imprinting, DNA methylation has a major role in repressing transposons[25,26] and germline-specific genes[27]. DNA methylation is also highly enriched in pericentromeric satellite repeats[28] and in the bodies of transcribed genes[29], although the precise function of 5mC in both of these contexts is unclear.

Remarkably, the mammalian genome undergoes two extensive waves of reprogramming of CpG methylation patterns during embryogenesis — following fertilization and after germline cell specification[30,31]. Much progress has been made recently in understanding the genetic requirements for these epigenome reprogramming processes. Genome-wide methods with base-pair resolution have been developed to elucidate the nuances of DNA methylation dynamics during embryonic development,

Genetics and Developmental Biology Department, Institut Curie, Paris Sciences Lettres University, INSERM, CNRS, Paris, France.

**e-mail: deborah.bourchis@curie.fr*

https://doi.org/10.1038/s41580-019-0159-6

and precision epigenome editing tools are increasingly being used to ascertain the function of DNA methylation on a locus-specific basis.

In this Review, we discuss the latest major advances in our understanding of the functions of DNA methylation and its establishment, maintenance and erasure during mammalian development. We also discuss new insights gained into the patterning of DNA methylation during development. Finally, we discuss the most current views of the diverse functions of DNA methylation in genetic diseases.

Cellular functions of DNA methylation

Cytosine methylation is pervasive throughout mammalian genomes, but likely carries out distinct functions in different genomic regions. Moreover, even when DNA methylation is associated with transcription silencing, the underlying mechanisms are not necessarily identical at gene promoters, gene bodies or repeated sequences. In this section, we discuss the latest evidence for the genomic effects and functional mechanisms of DNA methylation.

The writers and erasers of methylation

There are three phases of DNA methylation: establishment (de novo DNA methylation), maintenance and demethylation. In mammals, there are two major de novo DNA methylation enzymes, DNMT3A and DNMT3B[32,33], which contain a highly conserved DNMT domain (the MTase domain) in the carboxy terminus and two chromatin reading domains, ATRX-DNMT3-DNMT3L (ADD) and PWWP (FIG. 1; TABLE 1). There is also a catalytically inactive DNMT, DNMT3L, which interacts with and stimulates the activity of DNMT3A and DNMT3B specifically in the germline[34,35] (TABLE 1). DNA methylation is usually excluded from CpG-rich promoters of actively transcribed genes, which are typically enriched in trimethylated histone H3 Lys4 (H3K4me3)[36]. The ADD domain, which binds to the K4 residue of H3 tails, is repelled by increasing numbers of methyl moieties at K4, with H3K4me3 being the most repelling of ADD domains[34,37,38] (FIG. 1a). When not binding H3K4, the ADD binds the MTase domain and auto-inhibits the activity of the DNMT3 enzymes; binding of the ADD domain to unmethylated H3K4 releases the MTase domain and enables DNA methylation[39] (FIG. 1a,b). In contrast to active promoters, the bodies of actively transcribed genes are enriched with DNA methylation[1,2,29]. As RNA polymerase II (Pol II) transcribes, the histone methyltransferase SETD2 concomitantly trimethylates H3K36 (REFS[40,41]); the PWWP domain binds to H3K36me3 in vitro[42], which strongly suggests there is a mechanistic link between transcription and DNA methylation at gene bodies (FIG. 1b). Indeed, a DNMT3B protein with a PWWP mutation loses affinity for H3K36me3-marked gene bodies in mouse embryonic stem cells (ESCs), and DNMT3B is lost from gene bodies in *Setd2* mutants[43]. Additionally, mouse *Setd2*-mutant oocytes exhibit grossly deregulated targeting of DNA methylation[44].

Whereas de novo DNA methylation can occur in any sequence context, only symmetrical CpG methylation is maintained upon DNA replication. This depends on the activity of the methylation maintenance enzyme DNMT1 in concert with another multidomain protein, E3 ubiquitin-protein ligase UHRF1 (TABLE 1). UHRF1 specifically binds hemimethylated CpG dinucleotides at replication forks through its SET- and RING-associated (SRA) domain[45,46], and H3K9me2 and H3K9me3 through its tandem TUDOR-PHD (TTD-PHD) domain[47–50] (FIG. 1c). By itself, DNMT1 exists in an auto-inhibitory configuration, as its replication foci targeting sequence (RFTS) is buried in the catalytic MTase domain, akin to the ADD domain of the DNMT3 enzymes[51–53] (FIG. 1c). UHRF1 recruits DNMT1 through its ubiquitin-like (UBL) domain, thereby releasing its auto-inhibition and allowing RFTS binding to histone H3 tails that were previously ubiquitylated by the RING finger domain of UHRF1 (REFS[54,55]). DNMT1 then methylates the daughter DNA strand[53]. Accordingly, mouse ESCs expressing a mutated *Uhrf1* recapitulate the DNA methylation phenotype of the *Dnmt1* mutant[45,46].

Active DNA demethylation is carried out by the TET methylcytosine dioxygenases, which progressively oxidize 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC)[56–60] (TABLE 1). All the oxidized forms can promote DNA demethylation during replication[61,62]; in the case of 5fC and 5caC, demethylation can also occur through base removal by thymine DNA glycosylase (TDG) followed by the activity of the base excision repair pathway[60,63,64]. A more detailed description of DNA methylation and demethylation mechanisms can be found in Supplementary Box 1 and Supplementary Figure 1.

DNA methylation represses transcription

The repressive role of DNA methylation in transcription has long been recognized with a correlation between DNA methylation and gene silencing that increases with the density of CpG dinucleotides at promoters[65]. However, how this leads to transcription inhibition is still not entirely resolved, as the methyl mark per se does not seem to confer silencing. Regions of accessible chromatin are frequently lowly methylated or unmethylated, indicating that binding of transcription factors and DNA methylation are mutually exclusive[66]. Certain transcription factors are sensitive to CpG methylation: a recent survey of 542 human transcription factors found that 117 (22%) exhibited decreased binding to their motifs when methylated compared with unmethylated[67]. By preventing the binding of such transcription factors, DNA methylation can therefore impede transcription activation of CGI promoters that contain their sequence-recognition motifs. Using the same principle, methylated cytosines can also serve as binding modules for transcription activators (BOX 1).

DNA methylation can also contribute to heterochromatin formation through the recruitment to chromatin of chromatin remodellers and modifiers by DNMT proteins: de novo DNMTs function in complex with the chromatin remodeller lymphocyte-specific helicase (LSH) and with H3K9 methyltransferases and histone deacetylases[68–75]. Protein recruitment can also occur

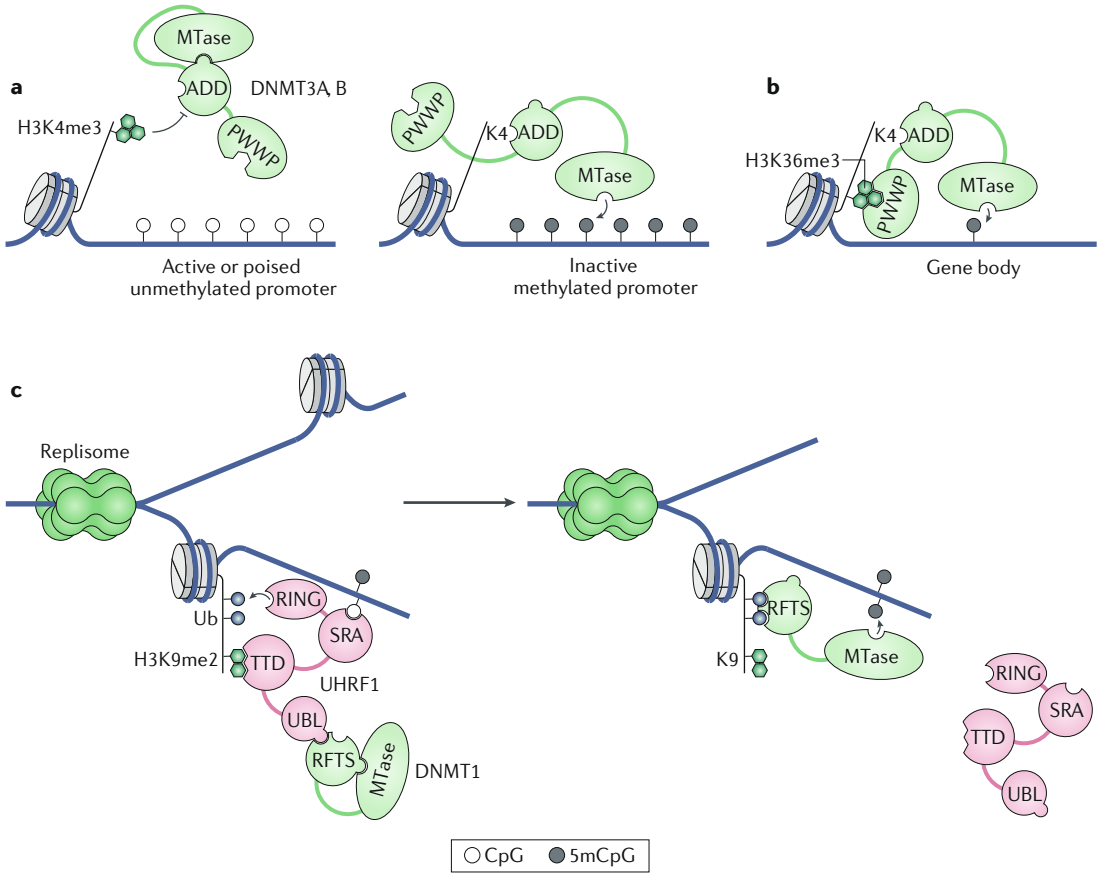


Fig. 1 | **DNA methylation machinery and mechanisms. a** | Mechanism of DNA methylation at promoters. Left: trimethylated histone H3 Lys4 (H3K4me3), which marks active and poised promoters, prevents binding to chromatin of the ATRX-DNMT3-DNMT3L (ADD) domain of DNA (cytosine-5)-methyltransferase 3A (DNMT3A) and DNMT3B (and also of DNMT3L), thereby causing it to bind to the methyltransferase (MTase) domain and auto-inhibit the DNMT3 enzymes. Right: in the absence of H3K4 methylation, the ADD domain binds to H3K4 and the auto-inhibition is relieved, thereby allowing the MTase domain to methylate the DNA. **b** | DNA methylation at gene bodies. In gene bodies, the ADD domain binds unmethylated H3K4, thereby releasing the auto-inhibition of the DNMT3 enzymes. H3K36me3 is deposited in gene bodies of actively transcribed genes and serves as a recruitment module for the DNMT3 PWWP domain. **c** | Maintenance of DNA methylation. Left: E3 ubiquitin-protein ligase UHRF1 is recruited to replicating DNA through its SET- and RING-associated (SRA) domain, which binds hemimethylated CpG sites, and through its TUDOR (TTD) domain, which binds H3K9me2. The RING domain of UHRF1 ubiquitylates the histone H3 tails (Ub). The replication foci targeting sequence (RFTS) of DNMT1 folds into the MTase domain, thereby preventing its catalytic activity. UHRF1 recruits DNMT1 through an interaction between its ubiquitin-like (UBL) domain and the DNMT1 RFTS. Right: the auto-inhibition of DNMT1 is released when the RFTS binds to ubiquitylated H3 tails, which enables the maintenance of symmetrical DNA methylation at CpG sites. Some domains of UHRF1 and DNMT1 were omitted from the figure for simplicity. 5m, methylation of the fifth carbon.

through 5mC and the readers, methyl-CpG-binding domain (MBD) proteins[76,77]. Mammals have five MBD proteins: MBD1–MBD4 and methyl-CpG-binding protein 2 (MeCP2). Four of the MBD proteins exhibit a linear relationship between an increase in CpG binding and increased CpG methylation[78], but MBD3 does not have preference for methylated cytosines[79]. All MBDs interact with nucleosome remodelling and histone deacetylase complexes, which leads to gene silencing[80,81]. It should be noted that genetic evidence in vivo for MBD function is complicated by redundancy between the proteins. Finally, there are several zinc finger proteins that recognize and bind DNA methylated sequences (reviewed elsewhere[82]). Such factors could contribute to DNA methylation-based silencing independently of, or redundantly with, MBDs.

CpG-island promoters

The mammalian genome is generally CpG poor, with the exception of CpG islands (CGIs), which are relatively small genomic regions of roughly 1 kb on average. Over two-thirds of mammalian promoters are CGIs[83,84]: virtually all housekeeping genes have CGI promoters, and so do several developmentally regulated genes[85]. CGIs are very rarely methylated[86]; they are particularly unmethylated in the dividing male germline cells, explaining why they are not subject to CpG erosion by deamination during evolution and their remarkably high CpG content. Most inactive CGI promoters are silenced by Polycomb repressive complex 2-mediated H3K27 methylation, which is a more plastic mode of silencing than DNA methylation and therefore more amenable to gene (re)activation in response to

Table 1 | **DNA methylation and demethylation factors and their functions**

Factor	Function	Mouse loss-of-function phenotype	Human diseases associated with genetic mutations
DNMT1	Maintenance DNA methyltransferase	• Low global DNA methylation • Derepression of IAP transposons • Early embryonic lethality	• Hereditary sensory autonomic neuropathy 1E (HSAN1E; OMIM 614116) • Autosomal-dominant cerebellar ataxia, deafness and narcolepsy (ADAC-DN; OMIM 604121)
UHRF1	DNMT1 cofactor	• Low global DNA methylation • Early embryonic lethality	
DNMT3A	De novo DNA methyltransferase	• Constitutive knockouts die ~4 weeks after birth[a] • Sterility in both males and females in germline-specific knockouts	• Microcephalic dwarfism • Tatton-Brown–Rahman syndrome (TBRS; OMIM 602729) • Acute myeloid leukaemia (AML; OMIM 601626)
DNMT3B	De novo DNA methyltransferase	Constitutive knockouts die mid-gestation[a]. More important for embryonic DNA methylation than for germline DNA methylation	Immunodeficiency, centromeric instability and facial anomalies syndrome (ICF; OMIM 602900)
DNMT3C	De novo DNA methyltransferase (Muroidea specific)	Males are infertile likely owing to defect in methylating transposon promoters during spermatogenesis	
DNMT3L	De novo DNA methyltransferase cofactor	• Male germline cells unable to undergo meiosis • Females unable to establish maternal imprinting, leading to mid-gestation lethality of progeny	
TET1	DNA demethylation via oxidation of methylcytosine	Loss has subtle effects and the embryos are viable[b,c]	
TET2	DNA demethylation via oxidation of methylcytosine	Increased self-renewal of haematopoetic stem cells[b,c]	• AML (OMIM 601626) • Chronic myelomonocytic leukaemia • Lymphomas • Myeloproliferative neoplasms
TET3	DNA demethylation via oxidation of methylcytosine	Germline conditional knockout leads to impaired paternal demethylation and reduced fecundity[c]	

DNMT, DNA (cytosine-5)-methyltransferase; IAP, intracisternal A particle. [a]*Dnmt3a*$^{-/-}$;*Dnmt3b*$^{-/-}$ double mutants exhibit early embryonic lethality; [b]*Tet1*$^{-/-}$;*Tet2*$^{-/-}$ double mutants display a range of developmental defects and partial lethality; [c]*Tet1*$^{-/-}$;*Tet2*$^{-/-}$;*Tet3*$^{-/-}$ triple mutant embryos exhibit gastrulation failure.

developmental or environmental cues (reviewed elsewhere[87]). Nevertheless, there are three major classes of genes, in which stable, lifelong DNA-methylation-based silencing in somatic tissues is very important: genes on the inactive X chromosome, imprinted genes and germline-specific genes. Here, we discuss the processes by which DNA methylation can be targeted to these specific classes of CGI promoters.

X-chromosome inactivation. In female mammals, one X chromosome in each cell is randomly silenced by the activity in *cis* of the non-coding RNA X-inactive specific transcript (*XIST*). In this process of XCI, DNA methylation of X-linked CGIs appears to occur relatively late and to function as a final lock added after the genes have already been silenced[14,88–90]. In mice, X-linked CGI-promoter silencing mostly depends on DNMT3B — the other DNMT3s are dispensable[91] — and in a subset of X-linked CGIs it also requires structural maintenance of chromosomes flexible hinge domain-containing 1 (SMCHD1)[91–93] (FIG. 2a). In humans, SMCHD1 is also involved in XCI, but its link with DNA methylation has not been demonstrated[94]. The precise mechanism of DNMT3B recruitment during XCI remains unclear, and so does the reason why SMCHD1 evolved such an important XCI-specific function. Perhaps the hierarchical and highly regulated heterochromatization of the X chromosome provides a unique chromatin environment that facilitates DNA methylation of X-linked CGIs.

Genomic imprinting. In genomic imprinting, DNA methylation is differentially established in the two parental germlines; these epigenetically imprinted patterns withstand the genomic reprogramming of DNA methylation that takes place during early embryogenesis. In both mouse and human, around 20 genomic regions known as imprinting control regions (ICRs) withstand this reprogramming and force mono-allelic expression of neighbouring genes[95–101]. The majority of ICRs are methylated in the oocyte and these all coincide with extremely CpG-rich CGIs[101], whereas the three paternal ICRs map to CpG-poor intergenic sequences. During oocyte growth, DNMT3A assisted by DNMT3L methylates oocyte-expressed gene bodies, including their intragenic CGIs in a transcription-dependent manner, whereas the rest of the genome

Retrotransposons
Transposable elements that propagate in the genome through RNA intermediates and reverse transcription.

remains hypomethylated[35,102–106] (FIG. 2b). Mammalian oocyte transcription typically emanates from alternative promoters that are located upstream from canonical CGI promoters (which are used after fertilization) and often coincide with sequences of retrotransposons[106–109].

What distinguishes maternal (and also paternal) ICRs from other sequences that are methylated in the gametes is their enrichment in specific genetic motifs (TGCCGC), which, when methylated, are recognized by the Krüppel-associated box (KRAB)-containing zinc-finger protein 57 (ZFP57)[110–112] (FIG. 2b). ZFP57 recruits KRAB-associated protein 1 (KAP1; also known as TIF1β) and other silencing factors, including DNMTs[111,113,114]. This selective DNA binding of ZFP57–KAP1 in oocytes allows ICRs to maintain allele-specific methylation in post-fertilization embryos, which undergo global DNA methylation erasure and re-establishment. It should be noted that mutations in *Zfp57* are not fully penetrant in mice: a recent study demonstrated that a similar protein, ZFP445, cooperates with ZFP57 at nearly all ICRs to maintain methylation imprints in mice, and perhaps has an even more important role than ZFP57 in humans[115]. In summary, the ability of maternally imprinted CGIs to undergo DNA methylation relies on a combination of the unusual transcriptional landscape of the oocyte and on specific genetic sequences.

Germline-specific genes. CGI promoters of germline-specific genes are silenced by DNMT3B-mediated DNA methylation with the onset of somatic differentiation during embryo implantation[27,116]. Germline genes are acutely sensitive to loss of DNA methylation and are derepressed in *Dnmt* triple-knockout mouse ESCs[117], *DNMT1*-depleted human fibroblasts[118], *Dnmt3B*-mutant mouse embryos[27,119] and human diseases linked with *DNMT3B* mutations[120]. What makes CGI promoters of germline genes, which account for only 5% of the total number of CGIs, dependent on DNA methylation, whereas the great majority of autosomal CpG-rich promoters remain unmethylated? The key may be a non-canonical Polycomb repressive complex 1 (PRC1) known as PRC1.6, which is a regulator of germline-gene repression in mouse ESCs[121,122]. The complex contains DNA binding proteins[123] that provide sequence-specific targeting, such as MAX[121,124], MGA[121,122] and E2F6 (REF.[125]). Furthermore, PRC1.6 associates with L3MBTL2 (REFS[126,127]), which interacts with the H3K9 methyltransferase G9A (also known as EHMT2)[126]. Mouse *G9a*-mutant embryos fail to acquire DNA methylation specifically at a large subset of germline-specific genes[119] (FIG. 2c). Many aspects of the role of PRC1.6 in repressing germline-specific genes remain unclear, such as the means by which CGIs of germline genes are targeted by DNMT3B but not by DNMT3A. Nevertheless, contrary to most CGI promoters, which resist de novo DNA methylation at implantation, germline CGIs have evolved sequence-specific means of recruiting the DNA methylation machinery to ensure lifelong somatic silencing.

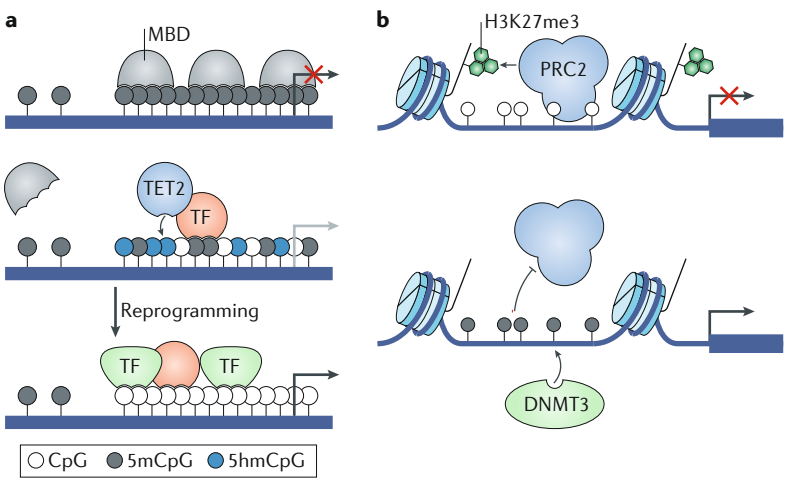
Harnessing the repetitive genome

Genome defence against transposable elements has been proposed as a major driver of the evolution of DNA methylation[128,129]. Indeed, the main targets of DNA methylation in mammalian genomes are not genes but transposable elements, and in particular retrotransposons. There are millions of copies of retrotransposons in the mouse and human genomes, which occupy roughly half of the genomic space[130]. The expression of the most active retrotransposons is controlled by CpG-rich promoters, and DNA methylation is important for their

Box 1 | **DNA methylation-mediated transcription activation**

DNA methylation can be read by methyl CpG binding domain (MBD)-containing transcriptional repressors (see the figure, part **a**), but methylated DNA motifs can also be specifically recognized by transcription activators. Various methods, including stable isotope labelling with amino acids[238], microarrays[246] and systematic evolution of ligands by exponential enrichment[67] — all of which are designed to discriminate between methylated and unmethylated DNA motifs recognized by DNA binding factors — collectively revealed that humans and mice express dozens of transcription factors with binding preferences of specific methylated sequences. These include the cell pluripotency factors KLF4 (REFS[238,246–248]) and OCT4 (also known as POU5F1)[67], the homeobox proteins HOXB13 (REF.[67]) and the NKX neural patterning factors[67]. C/EBPα also uses a methyl-specific binding motif, which is important for keratinocyte differentiation[249]. It is intriguing that several transcription factors (TFs) involved in cell-type transitions exhibit methylcytosine binding specificity and can function at chromatin that is otherwise refractory to transcription activation in order to facilitate these transitions. A recent study showed that C/EBPα and KLF4 (and another transcription factor) recruit the methylcytosine dioxygenase TET2 to enhancers for demethylation during cell-type reprogramming[201] (see the figure, part **a**).

DNA methylation can have a counterintuitive role in activating genes. The genomic distribution of DNA methylation is largely mutually exclusive from that of histone H3 Lys27 trimethylation (H3K27me3), which is a gene-repressive histone modification that is catalysed by Polycomb repressive complex 2 (PRC2)[250–253]. Whereas gene silencing is usually preserved when replacing one repressive modification by another, displacement of PRC2 by DNA methylation and, consequently, loss of H3K27me3 occasionally correlates with gene activation, in normal physiological conditions and in cancer[223,254]. During neurogenesis, DNA (cytosine-5)-methyltransferase 3A (DNMT3A)-mediated de novo DNA methylation evicts PRC2 and, consequently, H3K27me3 from the regulatory regions of neural genes[208,228] (see the figure, part **b**). The human *FOXA2* gene appears to be regulated in a similar fashion during endoderm development[255]. Moreover, during mouse embryonic de novo DNA methylation, the imprinted *Zdbf2* gene is activated by DNA methylation upstream of its promoter, which disrupts H3K27 trimethylation[183]. In the absence of the Polycomb-to-DNA methylation switch, *Zdbf2* remains silenced throughout life, resulting in reduced body size.



5hm, hydroxymethylation of the fifth carbon; 5m, methylation of the fifth carbon.

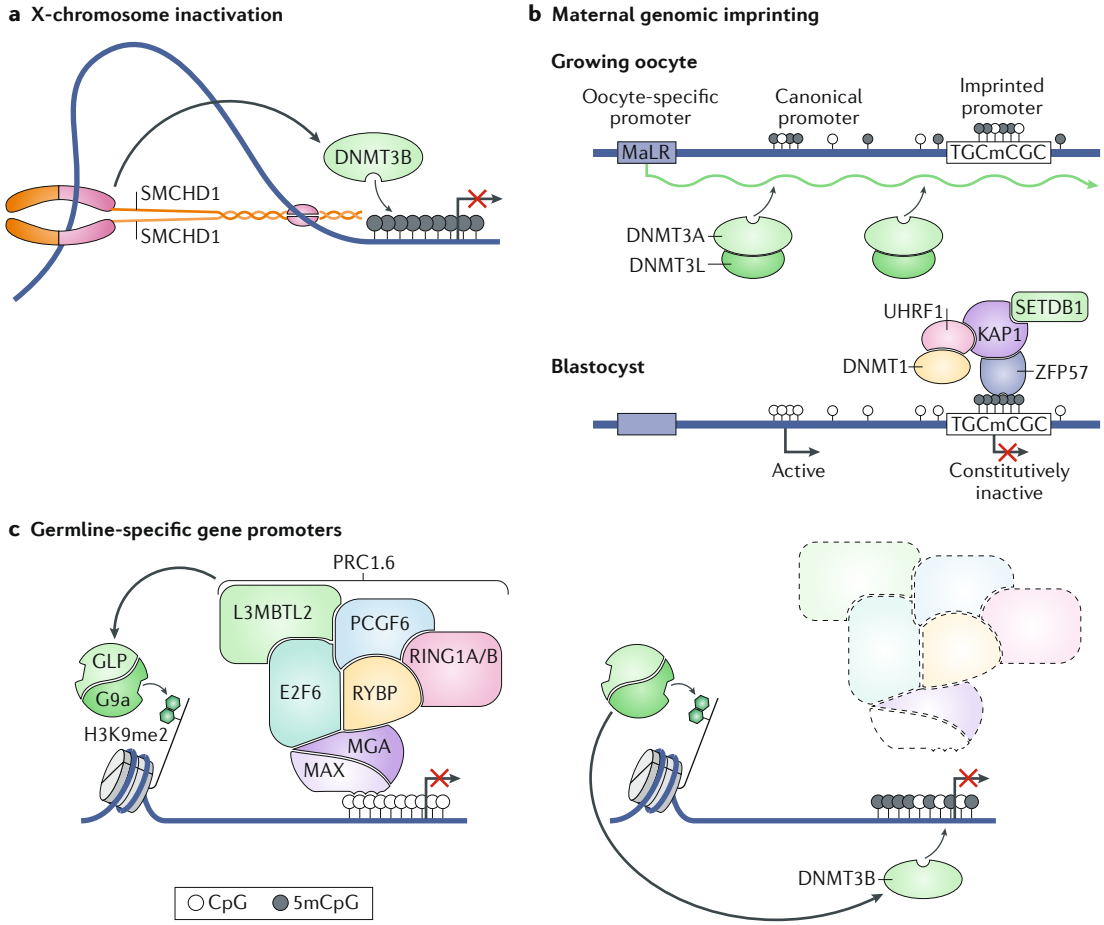


Fig. 2 | **Targeting DNA methylation to CGI promoters. a** | Structural maintenance of chromosome flexible hinge domain-containing 1 (SMCHD1)-dependent de novo DNA methylation at CpG islands (CGIs) during X-chromosome inactivation. SMCHD1 forms homodimers and may be involved in condensing the chromatin of the X chromosome that is undergoing inactivation. At a subset of CGIs, SMCHD1 is required for DNA (cytosine-5)-methyltransferase 3B (DNMT3B)-mediated DNA methylation through an unclear mechanism. **b** | Establishment of maternal imprinting. In the growing oocyte, DNA methylation is strongly correlated with transcription elongation in a DNMT3A–DNMT3L-dependent manner. In mice, many of these transcripts arise from mammalian apparent long terminal repeat retrotransposons (MaLR). After fertilization, imprinted promoters withstand the early embryonic global DNA methylation erasure through the binding to methylated TGCCGC motifs of the zinc-finger protein 57 (ZFP57)–KRAB-associated protein 1 (KAP1) complex, which recruits the histone methyltransferase SETDB1 and the maintenance DNA methylation factors DNMT1 and UHRF1, thereby maintaining gene silencing. **c** | Model for the establishment of DNA methylation at CGI promoters of germline-specific genes. A subset of germline genes is bound by the non-canonical Polycomb repressive complex 1 (PRC1), PRC1.6. The complex also consists of L3MBTL2, which can recruit the heterodimeric H3K9 methyltransferase complex, G9a–GLP. G9a is required for DNMT3B-mediated deposition of DNA methylation specifically at germline genes. 5m, methylation of the fifth carbon.

transcriptional silencing. This is exemplified in mouse *Dnmt1* knockout embryos, where intracisternal A particle (IAP) retrotransposons, which constitute a class of evolutionarily young retrotransposons that can still actively mobilize in rodent genomes, are massively derepressed[131] (FIG. 3a).

Two independent mouse genetic screens have recently discovered DNMT3C — a previously unknown de novo DNMT with a specific role in controlling retrotransposons[129,132] (TABLE 1). Originally annotated as a pseudogene, *Dnmt3C* originates from a tandem duplication of the *Dnmt3B* gene in the Muroidea lineage — the rodent superfamily that contains mice and rats. During evolution, DNMT3C lost its PWWP domain, but maintained its ADD and its de novo methylation activity. Unlike DNMT3A and DNMT3B, which are broadly expressed in various developmental contexts in both sexes, DNMT3C expression is confined to male fetal germ cells. Although homozygous *Dnmt3C*-mutant mice are viable, males have small testes and are infertile (FIG. 3a).

Whole-genome methylation analysis revealed that DNMT3C selectively methylates and represses the promoters of evolutionarily young transposable elements, which account for only 1% of the mouse genome[129]. The molecular phenotype of the *Dnmt3C* mutation is identical to that of mutants of the germline methylation cofactor DNMT3L[35] and also to that of mutants of the piwi-interacting RNA pathway; piwi-interacting RNAs

constitute a highly conserved class of small RNAs specifically dedicated to the silencing of transposable elements in the germline[133–139]. Thus, although mechanistic confirmation awaits, it is highly likely that DNMT3C is the tip of the small-RNA-directed DNA methylation spear (FIG. 3b). Given the sterility of the mouse *Dnmt3C* mutant, one wonders how the human male germline copes with transposable elements given that the human

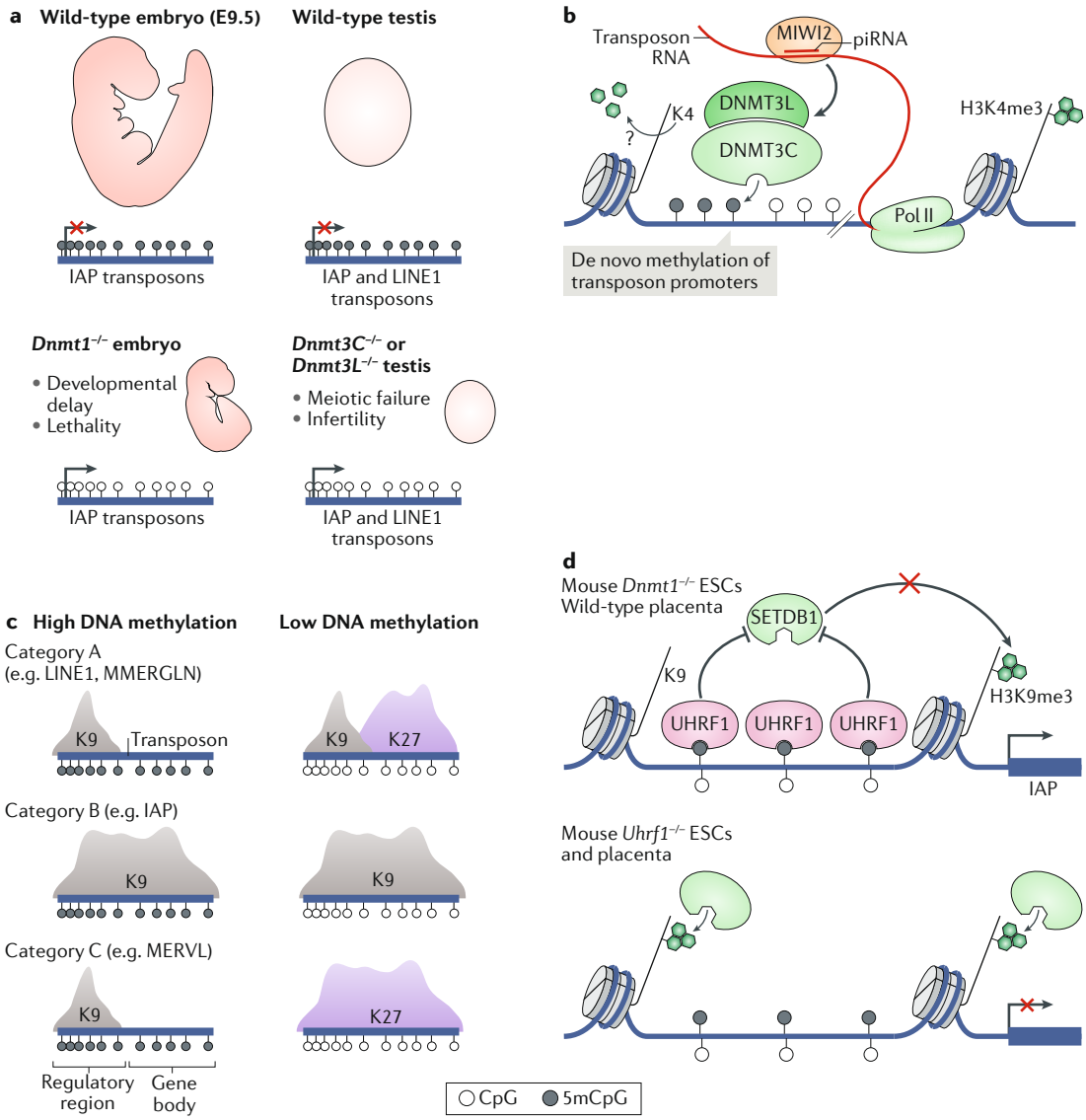


Fig. 3 | **DNA methylation-based regulation of transposons. a** | In vivo consequences of DNA methylation loss at transposons. In the mouse embryo, loss of DNA (cytosine-5)-methyltransferase 1 (DNMT1) leads to drastic upregulation of intracisternal A particle (IAP) retrotransposons. This is associated with severe developmental defects and prenatal lethality in embryonic day 9.5 (E9.5). *Dnmt3C* or *Dnmt3L* mutants have small testes and exhibit specific loss of DNA methylation at IAP and at long interspersed nuclear element 1 (LINE1) retrotransposons in the male germline, which leads to meiotic catastrophe and infertility. **b** | Model of piwi-interacting RNA (piRNA)-directed DNA methylation in mice. piRNAs generated from transposon transcripts that are loaded onto MIWI2 (also known as PIWIL4) localize to the nucleus and recruit DNMT3C–DNMT3L through an unknown mechanism. As DNMT3C contains a chromatin-reading ATRX-DNMT3-DNMT3L (ADD) domain, there likely is a mechanism to circumvent the gene-activating effects of trimethylated histone H3 Lys4 (H3K4me3), such as its removal by a lysine demethylase. **c** | Compensation for loss of DNA methylation in mouse embryonic stem cells (ESCs). During conversion from culturing mouse ESCs in media supporting high global DNA methylation to media supporting low global DNA methylation, transposons can be generally divided into three classes. In class A, H3K9 methylation remains unperturbed, but H3K27me3 invades the transposon body; in class B, H3K9 methylation remains unchanged; in class C, there is an epigenetic switch from H3K9-methylation-based silencing to H3K27me3-based silencing. **d** | UHRF1 prevents SETDB1-mediated silencing of IAPs at hemimethylated DNA. In *Dnmt1* conditional knockouts, UHRF1 remains bound to post-replication hemimethylated DNA and prevents the H3K9 methyltransferase SETDB1 from silencing IAP transposons. This is physiologically relevant in wild-type mouse placenta, where IAP is derepressed owing to UHRF1 binding to hemimethylated DNA. In *Uhrf1* mutant ESCs and placenta, SETDB1 catalyses the trimethylation of H3K9 and consequently IAPs are repressed. 5m, methylation of the fifth carbon; MERVL, mouse endogenous retrovirus with leucine tRNA primer; MMERGLN, mouse endogenous retrovirus with glutamine tRNA primer; Pol II, RNA polymerase II.

genome lacks the *DNMT3C* gene. Perhaps human PWWP-less DNMT3B isoforms perform a similar function to rodent DNMT3C; alternatively, small-RNA-directed DNA methylation may simply not exist in the same manner in humans, and instead we have evolved a different mechanism for controlling transposable elements.

In addition to the short-term effects of repressing transposable elements, DNA methylation also promotes their irreversible genetic inactivation through mutagenic deamination. Finally, DNA methylation may support genome stability by limiting recombination between non-allelic copies of transposable elements with high sequence similarity. This possibility is supported by positive correlations between elevated levels of chromosomal rearrangements and genome-wide DNA hypomethylation — mostly involving transposable elements— in various human cancers[140]. Also, during meiosis, DNA methylation may limit transposable elements from engaging in homology-dependent search and recombination[141]. It should be mentioned that DNA methylation may exert genomic stability protective functions not only through targeting interspersed transposon repeats, but also through targeting tandem satellite repeats found in telomeric, centromeric and pericentromeric regions (reviewed elsewhere[142]).

In striking contrast to in vivo situations, transposable elements are not reactivated in mouse ESC models of constitutive DNA methylation deficiency, except IAPs, and their upregulation is only modest in *Dnmt1* knockout or *Dnmt* triple-knockout cells[117,143,144]. However, by conditionally triggering the ablation of DNA methylation, independent studies have recently demonstrated that several classes of transposable elements — including IAPs — are indeed reactivated upon withdrawal of DNA methylation[145–147]. Following this initial response, histone methylation restores long-term transcriptional silencing in hypomethylated genomes. Depending on the class of transposable element, silencing involves deposition of H3K9me3, H3K27me3 or a combination of the two[145] (FIG. 3c).

Paradoxically, the initial burst of IAP reactivation that follows conditional *Dnmt1* deletion is not observed upon removal of its DNA-targeting cofactor UHRF1 (REF.[146]). The basis of this discrepancy has been genetically dissected: suppression of post-replicative maintenance of DNA methylation in *Dnmt1* mutants initially generates an excess of hemimethylated DNA, which transiently prolongs UHRF1 residency on DNA, thereby inhibiting the recruitment of the histone methyltransferase histone-lysine *N*-methyltransferase SETDB1, its catalysis of H3K9 trimethylation and silencing of IAPs (FIG. 3d). Following longer culture and progressive dilution of hemimethylated DNA, UHRF1-based inhibition of histone methylation is relieved and SETDB1-mediated silencing is enabled. Interestingly, accumulation of hemimethylated DNA naturally happens in vivo, in specific developmental contexts[146]. Whether this leads to IAP reactivation depends on the nuclear availability of UHRF1, for example in the mouse placenta. This does not occur in the early embryo or in migrating primordial germ cells (PGCs), where UHRF1 is sequestered in the cytoplasm and cannot counteract SETDB1-mediated repression through H3K9 methylation[146].

The puzzling case of gene bodies

The enrichment of DNA methylation in gene bodies presents a paradox: on the one hand, gene-body methylation is highly conserved across eukaryotes — more than it is conserved at transposable elements, for example — indicating it has an important function[1,2]. On the other hand, DNA methylation is mutagenic, so why is it so prominent in coding sequences? Importantly, gene-body DNA methylation is positively correlated with transcription[29,148,149], hence, its role is not linked to gene silencing. Two hypotheses have been proposed for the function of DNA methylation in gene bodies: that it facilitates transcription elongation and/or co-transcriptional splicing, and that it represses intragenic cryptic promoters.

DNA methylation is enriched at exons relative to introns[29,150] and may affect the processivity of Pol II and, through this, splicing. Nucleosomes are also enriched at exons, and de novo DNA methyltransferase activity requires histone binding, which would explain the enrichment of DNA methylation at exons. Several studies indicate that DNA methylation affects splicing and gene expression, not vice versa[151–153]. At the human *CD45* gene, CCCTC-binding factor (CTCF) slows Pol II elongation rates at exon 5, thereby facilitating its inclusion; DNA methylation prevents CTCF binding, thereby leading to exon exclusion[154]. By contrast, at other loci, DNA methylation can facilitate exon inclusion by recruiting MeCP2 (REF.[155]), in line with the observation that alternative exons exhibit lower levels of DNA methylation on average than constitutive ones[151]. Additionally, heterochromatin protein 1 can regulate exon inclusion by recruiting splicing factors to H3K9me3-modified nucleosomes wrapped in methylated DNA[156]. However, these mechanisms only account for a fraction of alternative splicing events.

The second hypothesis — that DNA methylation inhibits intragenic promoters — is attractive for numerous reasons. First, it is consistent with DNA methylation being a transcriptional repressor. Moreover, as discussed above (FIG. 1b), the de novo DNA methylation machinery is recruited to DNA through binding to H3K36me3, which is known to prevent the use of cryptic promoters in baker's yeast[157]. Indeed, methylation of intragenic CGIs prevents promoter activity, and differential methylation can regulate transcription initiation in a tissue-specific and cell-type-specific manner in mammals[158]. However, intragenic CGIs are often conserved from mouse to human, and are more likely alternative promoters rather than illegitimate, cryptic promoters. A recent study in mouse ESCs showed that DNMT3B-mediated gene-body methylation restricts the activity of cryptic promoters downstream of H3K36me3 (REF.[159]). Although the finding is appealing, the observed effect, although statistically significant, occurs in a very small proportion of cells within a cell population. Moreover, *Dnmt* triple-knockout ESCs do not exhibit inhibition of intragenic cryptic promoters as the *Dnmt3B* single-mutant cells do[117,160]. More functional analyses are needed to confirm the role of DNA methylation in

repressing cryptic promoters, particularly in cell types other than mouse ESCs, in which DNA methylation is not required for cell viability and other safeguards may be in place in gene bodies.

In summary, both of the above hypotheses are logical and sound, and not necessarily mutually exclusive. In both cases, DNA methylation appears to be at most a fine-tuner, as loss of gene-body methylation does not lead to drastic molecular phenotypes. Future work may reveal hidden mechanisms and/or functions that will help explain the highly conserved prevalence of DNA methylation in gene bodies.

Methylation patterning in development

The discovery of DNA methylation reprogramming in mammalian development was made over three decades ago. However, it was not until the advent of whole-genome bisulfite sequencing that DNA methylation patterns could be assessed with single-nucleotide resolution, and, more recently, using a small number of cells (Supplementary Box 2). Subsequently, techniques have been developed to assess the oxidized forms of 5mC with the same precision. In this section, we describe the recent nuanced findings and revised insights these techniques have provided into the role of DNA methylation in development.

Early embryos and the germlines

The traditional view of mammalian epigenetic reprogramming predicates an extensive expunction of DNA methylation during both early embryonic and germline development. This may be required for acquiring epigenetic plasticity at important developmental stages, but also for limiting the inheritance of acquired epimutations. However, the mechanisms of DNA methylation reprogramming in early embryogenesis and in germline development have key differences, although in both cases it is now clear that the process is neither complete nor totally linear: a limited number of sequences resist DNA methylation loss, and de novo DNA methylation may be occurring during this process.

Global passive and active demethylation and local de novo methylation. Genome-wide DNA demethylation has been a contentious area of research. In germline reprogramming, a two-step demethylation accompanies the acquisition of germ cell identity in PGCs[161,162]. The first phase consists of passive demethylation of the bulk of the genome as a default route[163–165]. This is followed by TET1-dependent and TET2-dependent demethylation, which mostly affects imprinted loci and germline-specific genes (FIG. 4a) and coincides with the appearance of 5hmC in PGC nuclei in the mouse[164–166].

During post-fertilization reprogramming, the embryo loses gamete-specific DNA methylation patterns inherited from the oocyte and the sperm as it progresses towards pluripotency; this also occurs in two phases. In mouse one-cell-stage embryos (zygotes), joint appearance of TET3-mediated 5hmC and 5mC loss indicates that the paternal genome is initially actively demethylated by TET3 (REFS[167–169]); the two parental genomes then undergo rounds of passive, DNA replication-dependent dilution of DNA methylation, as the maintenance enzyme DNMT1 provided by the oocyte is excluded from the nucleus during subsequent cell divisions[170] (FIG. 4a). This model was challenged by three recent whole-genome bisulfite sequencing analyses of embryos from hybrid crosses between two mouse strains (to infer parental specificity with available single-nucleotide polymorphism) in conjunction with genome-wide 5hmC mapping, which revealed that the maternal methylome also undergoes limited TET3-dependent oxidation in the zygote[171–173]. Other studies also indicated that most zygotic DNA methylation is lost independently of replication[174] and that, at the paternal genome, this may occur in a TET3-independent manner through an unclear mechanism[174,175]. Although human data are more limited, they indicate that generally similar dynamics take place, with some species-specific differences. An initial rapid demethylation also primarily affects the paternal genome of human embryos, but it occurs from fertilization to the two-cell stage, whereas it is completed at the one-cell stage in the mouse[176–178]. There is then a gradual genome-wide loss of DNA methylation until the blastocyst stage, although the maternal genome retains higher DNA methylation levels in humans compared with mice[179].

Single-cell whole-genome bisulfite sequencing revealed prevalent occurrences of de novo DNA methylation in the human embryonic genome while it undergoes global demethylation: first in the paternal genome at the one-cell stage, and then genome-wide at the eight-cell stage, coincidently with embryonic genome activation[179], mainly targeting transposable elements. During mouse development, some de novo DNA methylation may also occur concurrently with global demethylation of the embryonic or the germline genomes. Indeed, 5mC is lost before 5hmC is gained in the paternal zygotic genome[174]; therefore, de novo DNA methylation must occur in this time window to provide 5mC for TET enzymes to oxidize. Similarly, during the active germline demethylation, TET1 may help to maintain the hypomethylated state against spurious events of de novo DNA methylation[180]. In summary, during global DNA methylation erasure both in embryos and in the germline, the major role of TET may be protection from ectopic DNA methylation rather than driving active demethylation per se.

Methylation reprogramming and the biological implications of resisting it. Although embryonic and germline demethylation is global, a substantial amount of DNA methylation persists at the end of both processes. This observation has raised the interesting possibility of the existence of intergenerational epigenetic inheritance and even transgenerational epigenetic inheritance (reviewed elsewhere[181]).

In the inner cell mass of preimplantation embryos, approximately 20% of CpGs retain gamete-inherited methylation in both mice[171] and humans[179] (FIG. 4a). These notably map to ICRs, as expected from the intergenerational nature of genomic imprinting, which is linked to the sequence-specific DNA demethylation resistance through KRAB–ZFP recruitment of KAP1 (REFS[111,113,114]). However, collectively, the ICRs comprise a negligible proportion of the genome. In addition to these

Whole-genome bisulfite sequencing
Sodium bisulfite treatment converts unmodified cytosines — but not (hydroxy)methylated cytosines — into uracils (thymines following PCR). Paired with next-generation sequencing, this technique generates genome-wide, single-nucleotide resolution maps of DNA methylation.

Intergenerational epigenetic inheritance
Epigenetic information that is inherited from the parents (for example, genomic imprinting).

Transgenerational epigenetic inheritance
Epigenetic information that is inherited from generations that were not exposed to the initial cue that caused the epigenetic change.

Inner cell mass
Refers to the pluripotent cells in the blastocyst of preimplantation embryos, which can be derived and cultured as embryonic stem cells.

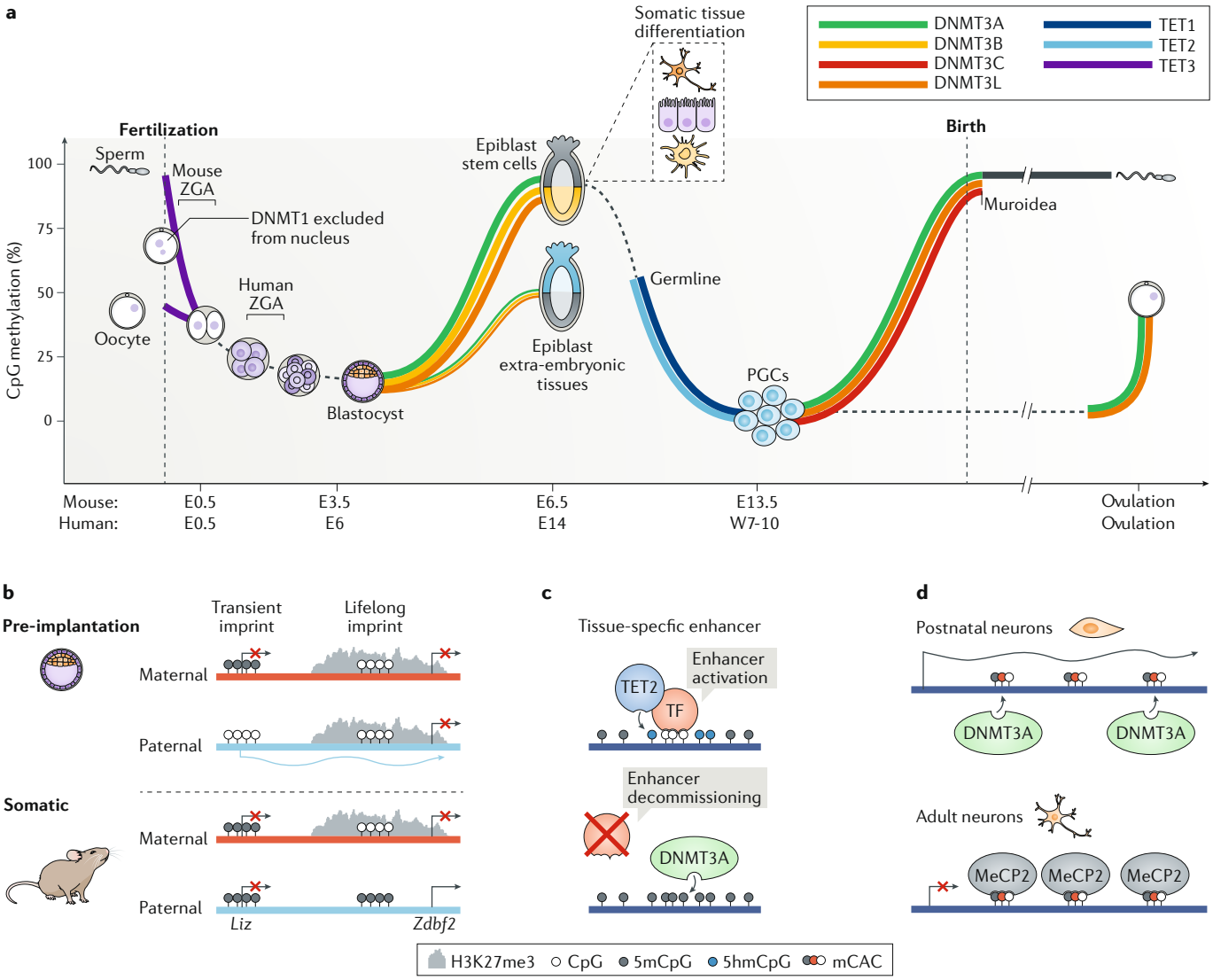


Fig. 4 | **DNA methylation reprogramming in development. a** | Embryonic and germline DNA methylation erasure and establishment. The methylcytosine dioxygenase TET3 is active in the fertilized zygote, leading to hydroxymethylation and active DNA demethylation. Following passive demethylation (dashed line), DNA methylation reaches a low point at the blastocyst stage, which is followed by DNA (cytosine-5)-methyltransferase 3A (DNMT3A)-mediated and DNMT3B-mediated de novo DNA methylation after blastocyst implantation. DNMT3L is also expressed in this time window, but is not absolutely required to methylate the embryonic genome. In extra-embryonic tissues, DNMT3A, DNMT3B and DNMT3L are expressed, but to a lesser extent than in the embryo proper, which correlates with relative DNA hypomethylation. Post implantation, in the epiblast, a subset of stem cells is specified for the germline, where they undergo two waves of DNA demethylation: one passive and one mediated by TET1 and TET2. Male gametes become highly methylated before birth through the activity of DNMT3A and DNMT3L, and, in the case of Muroidea (the superfamily that includes mice and rats), DNMT3C. The oocyte gains methylation after meiosis and prior to ovulation through the activity of DNMT3A and DNMT3L in mice, and likely through DNMT3A in humans. **b** | Transient imprinting at the *Zdbf2* locus. During pre-implantation development, the expression of long isoform of Zdbf2 (*Liz*) is imprinted: the maternal allele is methylated and silenced, and the paternal allele is unmethylated and expressed. The *Zdbf2* promoters on both alleles are silenced through Polycomb-mediated trimethylation of histone H3 Lys27 (H3K27me3). During implantation, the expression of paternal *Liz* leads to DNA methylation upstream of the canonical *Zdbf2* promoter and eviction of H3K27me3, whereas the maternal allele remains repressed by H3K27me3. Thus, although the imprinting of *Liz* is only transient, it causes permanent imprinting of *Zdbf2*, because the post-implantation methylation of the paternal locus enables lifelong expression of paternal *Zdbf2* by antagonizing Polycomb-mediated repression. Failure to activate *Zdbf2* in the embryo leads to a postnatal growth defect. **c** | DNA methylation turnover at tissue-specific enhancers. Dynamic DNA methylation in somatic tissues is frequently found at regulatory elements. TET enzymes in collaboration with transcription factors (TF) lead to DNA demethylation and enhancer activation. Inactive enhancers are bound less by transcription factors and are more accessible to de novo DNA methyltransferases. **d** | The role of CpA methylation in neurons. After birth, DNMT3A deposits DNA methylation in gene bodies. Methylated CAC sequences (mCAC) are recognized by methyl-CpG-binding protein 2 (MeCP2), the binding of which establishes a lifelong epigenetic memory of gene silencing. 5hm, hydroxymethylation of the fifth carbon; 5m, methylation of the fifth carbon; E, embryonic day; PGC, primordial germ cell; W, weeks.

classical ICRs, there are hundreds of 'transient' methylation imprints: these are mainly inherited from the oocyte, maintain maternal-specific DNA methylation until the blastocyst stage, probably through KRAB–ZFP-dependent protection, similarly to classical ICRs[101,115], and lose it through DNA demethylation or remethylation after implantation[101,176,179]. This phenomenon may be particularly pervasive in humans, where the maternal genome exhibits substantially higher preservation of DNA methylation than the paternal genome prior to implantation, but also after implantation, exclusively in placental tissues[130,179]. Investigation of the *Zdbf2* locus, whose transient imprinting is conserved in mice and humans, demonstrated that transient hypomethylation of the paternal allele can cause long-lasting imprinting through a cascade of downstream epigenetic changes that affect postnatal size[182,183] (FIG. 4b). Future work will be needed to assess whether other transient imprints have developmental roles.

The bulk of the residual gametic methylation in the blastocyst is not related to single-copy sequences but instead to transposable elements. Young retrotransposons of the IAP class are particularly resistant to pre-implantation demethylation in mice[171,184,185]. In human blastocysts, the highest retention of DNA methylation is also observed at evolutionarily young and potentially active transposable elements, in particular the hominid-specific SINE-VNRT-Alu elements[178]. Indications are that DNA methylation is retained through KRAB–ZFP-mediated sequence-specific recruitment of KAP1 (REF.[186]) in a manner analogous to the selective DNA binding of KAP1 in ICRs. KRAB–ZFPs require time to develop a match for recently emerged transposable elements[187], and thus the most recent IAP elements, which are often polymorphic between mouse strains, are not resistant to DNA methylation reprogramming during preimplantation development[188].

Compared with preimplantation development, the germline wave of demethylation is more extensive. By day 13.5 of mouse embryogenesis, PGCs exhibit only 6–8% of methylated CpGs[171] (FIG. 4a). Similar basal levels are retained in human fetal germ cells after 9 weeks of pregnancy[189,190]. Contrary to embryonic development, ICRs undergo demethylation in developing PGCs, as a prerequisite for subsequent acquisition of sex-specific DNA methylation patterns during male and female germline differentiation. Similarly to methylation reprogramming in preimplantation development, germline retention of DNA methylation is mostly observed at young and potentially harmful retrotransposons. In the mouse, this includes IAPs and long terminal repeat elements of the endogenous retrovirus 1 class of retrotransposons (ERV1)[161,191]. In humans, SINE-VNRT-Alu (SVA) elements and human-specific elements of the long interspersed nuclear element (LINE) family remain partially methylated compared with the rest of the genome and comparatively to older LINE types with a wider and therefore older distribution in the primate lineage[190]. Owing to difficulties in mapping repetitive elements in the genome, it is not clear whether the same retrotransposon copies resist both embryonic and germline demethylation or whether they represent distinct sets of elements. Therefore, whether young transposable elements contribute to epigenetic inheritance is unclear. However, although very rare, some single-copy sequences also escape germline DNA methylation reprogramming in mice and humans; intriguingly, these sequences exhibit very low levels of sequence and syntenic conservation between the two species[161,166,190]. Even more intriguingly, the DNA methylation of human germline escapee sequences was generally not erased also during embryonic methylation reprogramming. Whether these sparse sequences could serve as vectors of epigenetic inheritance is an alluring possibility.

The de novo DNA methylation programme

Following reprogramming in PGCs, sex-specific patterns of DNA methylation are established in the male and female germlines. At the end of gametogenesis, the sperm genome exhibits ~80% CpG methylation, with a genomic distribution roughly similar to that of somatic cells, although with a slightly higher DNA methylation content. By contrast, the oocyte genome is only ~50% methylated, and nearly exclusively in gene bodies[105,171,192] (FIG. 4a). The relative hypomethylation of the oocyte is associated with cytoplasmic retention of DNMT1, itself secondary to the sequestration of UHRF1 to the cytoplasm by the protein STELLA[193]. In *Stella*-mutant oocytes, DNMT1 enters the nucleus, resulting in ectopic DNA methylation deposition, a twofold increase in DNA methylation and female infertility. Puzzlingly, this suggests that DNMT1 is capable of de novo DNA methylation in the oocyte, although this is not the case in embryonic cells, for example[33]. The gene-body specificity of oocyte DNA methylation is linked to a tight coupling of de novo DNA methylation with gene transcription, in both mice and humans[106,177]. However, notable discrepancies exist between the two species. First, whereas gene-body methylation strictly requires DNMT3L-mediated stimulation of DNMT3A in mouse oocytes[192], DNMT3L is not expressed in human oocytes[177]. Second, thousands of syntenic regions show different DNA methylation patterns in mouse and in human oocytes. These differences were recently linked to species-specific insertions of long terminal repeat retrotransposons, whose transcription is particularly high during oogenesis[109]. By defining new transcription units during oocyte maturation, these elements appear to be essential contributors of oocyte methylation during development and mammalian evolution, and potential drivers of the emergence of new maternally imprinted loci (FIG. 2b).

At the end of embryonic reprogramming — at embryonic day 4.5 in the mouse — the blastocyst consists of the inner cell mass, which will go on to form the epiblast and then embryonic tissues, and of the trophoectoderm and the primitive endoderm, which contribute to the formation of extra-embryonic tissues: the extra-embryonic endoderm (ExE) and the visceral endoderm, respectively. Two recent genome-wide profiling studies have confirmed that the ExE and the visceral endoderm are both hypomethylated relative to the epiblast[194,195] (FIG. 4a). DNA hypomethylation persists in the differentiated placenta, which is largely derived from ExE cells[196,197]. Hypomethylation is correlated with

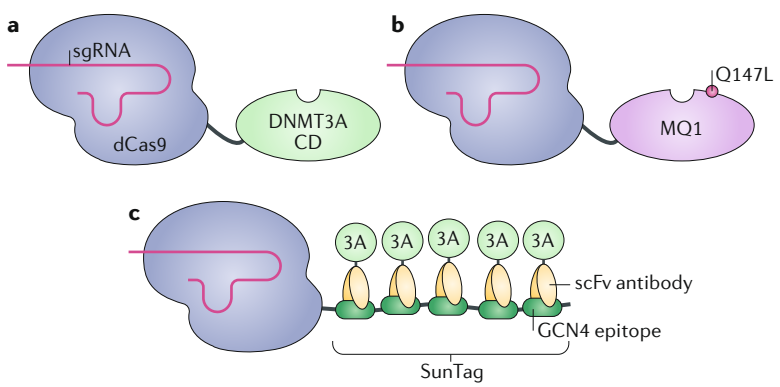
Box 2 | **Experimentally editing DNA methylation**

There is no universal mode of gene regulation by DNA methylation, as the methylation modification is associated with genomic regions harbouring inactive and active genes with protein-coding genes and repetitive elements alike; there are also broad genomic regions that are methylated but where the modification serves no obvious function. When analysing the effects of changes in DNA methylation in DNA methylation mutants (or in conditions of induced hypomethylation), there is always the risk of confusing primary effects with confounding and/or secondary effects. Methods exist for analysing the role of DNA methylation at specific loci. For example, CpG sequences can be inserted into endogenous promoters to trigger de novo DNA methylation[256]. However, the most promising avenue for altering the methylation state of a given genomic locus is the emerging approach of epigenome editing.

Epigenome editing tools preceded the CRISPR–Cas era; indeed, modifying DNA methylation has been successfully achieved by targeting chromatin modifiers to DNA through zinc-finger proteins[257] or transcription activator-like effectors (TALEs)[258,259]. However, zinc-finger constructs are cumbersome to generate, and although TALEs are relatively straightforward to assemble, they are sensitive to DNA methylation[260] and thus not ideal for manipulating DNA methylation. Given its ease of use, the tool of choice for targeting DNA methylation to specific loci is now CRISPR–dCas9, in which catalytically inactive Cas9 (dCas9) is fused with the catalytic domain of a DNA methyltransferase (DNMT) or a TET methylcytosine dioxygenase and targeted to genomic sequences by single guide RNAs (sgRNAs). Successful targeting of methyltransferases at mammalian genomes was achieved by fusions of dCas9 with the catalytic domain of mammalian DNMT3A[261,262] (DNMT3A CD; see the figure, part **a**) and with the MQ1 CpG DNA methylase of the bacterium *Mollicutes spiroplasma*[263] (see the figure, part **b**). MQ1 is highly active and can generate considerable off-target methylation; this was mitigated by generating a Q147L mutation, which reduces the intrinsic DNA binding capacity of MQ1.

A caveat of dCas9–methyltransferase fusion proteins is the limited range of DNA methylation from the targeted site, which can only be mitigated using multiple sgRNAs for targeting. In an interesting attempt to utilize DNMT biology, a dCas9 fused to chimeric DNMT3A catalytic domain–DNMT3L was demonstrated to effectively methylate DNA at a greater distance from the target sequence, likely through the formation of multimeric dCas9–DNMT3A–DNMT3L complexes[264]. Perhaps the most successful DNA methylome editing strategy thus far was derived from the dCas9–SunTag system, in which an array of GCN4 epitopes is fused to dCas9 (REFS[265,266]). By expressing the DNMT3A catalytic domain linked to a single-chain antibody that recognizes the epitope (scFv–DNMT3A), dCas9–SunTag recruits multiple enzymes to one locus (see the figure, part **c**). Furthermore, modular recruitment of DNMT3A can decrease off-target DNA methylation levels[266]. Similarly, using dCas9–SunTag to recruit scFv–TET1 has proved efficient for targeting demethylation activity[267].

Several studies are already putting epigenome editing to the test. Both DNMT3A and TET1 have been successfully targeted in vivo to alter gene expression and alter genome architecture[262], and targeted DNA methylation was used to assess the role of DNA methylation at neuronal genes during differentiation[223]. An exceptional study that set out to determine the extent of epigenetic memory of transient epigenome editing demonstrated that using a combination of dCas9–KRAB (a transcription repression domain), dCas9–DNMT3A and dCas9–DNMT3L proffered long-term gene silencing[268]. This approach was notably employed to study the role of deregulated genes in breast cancer[269]. We expect a torrent of forthcoming research using these precision tools to assess locus-specific functions of DNA methylation.



decreased expression of the DNMT3 enzymes compared with the epiblast, and even retrotranposons are relatively hypomethylated[194]. By contrast, a specific class of CGI promoters are methylated in the ExE and the visceral endoderm but not in the epiblast[194,195]. Genes associated with this set of promoters are developmentally important: their increased methylation either reflects a need to keep them repressed in extra-embryonic tissues or indicates that embryonic tissues have more stringent mechanisms to prevent aberrant DNA methylation and long-term silencing.

The remethylation of the epiblast genome after implantation is very rapid: somatic-tissue levels of DNA methylation are attained already by mouse embryonic day 6.5 (REFS[171,198]) (FIG. 4a). This is notable, as the stem cells in the epiblast are still pluripotent, so DNA methylation patterns established in epiblast stem cells have the potential to be propagated through life in all tissues and maintain epigenetic memory of early embryogenesis. Nevertheless, whole-genome surveys in both mice[101,197] and humans[149,199,200] revealed widespread yet focal discrepancies in DNA methylation patterns between various adult tissues, indicative of microwaves of de novo DNA methylation and demethylation occurring during tissue differentiation. In mice, the majority (75%) of tissue-specific differential DNA methylation regions are at enhancers or other regulatory elements, in line with studies showing that enhancers undergo active turnover of DNA methylation, likely resulting from the combined activities of TET2 and tissue-specific transcription factors[201,202] (FIG. 4c). In humans, a minority of DNA methylation regions (<50%) map to regulatory elements; a large subset is found in undefined regions, where the functional relevance of differences in DNA methylation is unclear. Some of the differences between the mouse and human datasets may be due to the greater sequencing depth of the human study[199], which may have allowed more nuanced extrapolations from the data.

The brain exhibits a peculiar methylome among somatic tissues. In both mice and humans, there is a postnatal spike in CpA methylation, which is generated by DNMT3A[203–205]. In fact, total CpA methylation levels in neural tissues are roughly equivalent to those of CpG methylation, although because CpA is a much more prevalent dinucleotide, mCpA is found at a lower frequency than mCpG. Recent work indicates that the MBD protein MeCP2 binds mCAC sequences in gene bodies of young mice neurons, and that MeCP2 binding maintains the silencing of these genes in later stages of life[206,207] (FIG. 4d). This mechanism could help us understand the role of DNA methylation, and in particular non-CpG methylation, in regulating neurological and behavioural traits. Future studies will hopefully further clarify how defective DNA methylation early in life can lead to neuronal abnormalities. Epigenome editing may prove to be a useful avenue to explore these questions[208] (BOX 2).

Genetic diseases, epigenetic outcomes

Disease-associated mutations have been characterized in DNMT and TET genes. Their effects on DNA methylation patterns are very diverse, as are their tissue and pathological manifestations, ranging from congenital

syndromes of immunodeficiency, growth phenotypes or neurodegeneration to haematological cancers. Other, phenotypically overlapping, syndromes have been linked to genes with known but also to those with unknown functions in DNA methylation and demethylation. The characterization of genetic mutations in the DNA methylation pathways has greatly enriched our understanding of the complexity and specificity of DNA methylation-based gene regulation.

DNMT1 and neurological disorders

Heterozygous mutations in *DNMT1* have been identified in a spectrum of autosomal-dominant forms of progressive cognitive and behavioural deterioration, including hereditary sensory autonomic neuropathy 1E with dementia and hearing loss (HSNA1E; OMIM 614116)[209] and autosomal-dominant cerebellar ataxia, deafness and narcolepsy (ADAC-DN; OMIM 604121)[210] (TABLE 1). The mean age of onset of HSNA1E is around 37 years, with hearing and sensory loss being the initial symptoms followed by decline in cognitive function, ataxia and brain atrophy[211]. The average life expectancy is 50 years and dementia is the main driver of morbidity.

Strikingly, all *DNMT1* mutations occur at the RFTS domain, with HSNA1E mutations clustering in the amino-terminal and middle part of the domain and ADAC-DN mutations distinctively segregating to the carboxy terminus. As discussed above, proper folding of the RFTS is crucial for DNMT1 function[52,53]. Genome-wide analysis of the peripheral blood of individuals with HSNA1E or ADAC-DN indicated that disturbances to DNA methylation were very moderate compared with non-affected siblings, with hypomethylation of intergenic regions and hypermethylation of some CGIs[212,213]. These DNA methylation signatures may be used as pathological biomarkers in accessible cell types, but they may not necessarily relate to the brain-specific manifestation of the diseases. Interestingly, DNMT1 proteins with mutated RFTS domains form cytoplasmic aggregates in cell-based overexpression assays, whereas wild-type DNMT1 is normally nuclear[211]. The induction of cellular stress by protein aggregation could result in toxicity and age-dependent degeneration of the central and peripheral nervous systems.

DNMT3B and immunodeficiency

The immunodeficiency, centromeric instability and facial anomalies (ICF) syndrome (OMIM 602900) was the first genetic disease to be linked to congenital DNA methylation defects[214] (TABLE 1). Individuals with ICF suffer from atypical immunoglobulin deficiency that causes recurrent and often life-threatening infections. The diagnosis of ICF typically includes a karyotype with multiradiated chromosomes 1, 9 and 16, in association with specific hypomethylation of pericentromeric satellite repeats II and III. Recessive mutations in the *DNMT3B* gene define ICF1, which accounts for the majority of ICF cases[215]. As expected given the lethality of mouse models of *Dnmt3B* loss-of-function mutations[33], most ICF1 mutations cause single amino acid substitutions or deletions and are thought to result in reduced rather than abolished DNMT3B activity. Despite the scarcity of ICF — only 70 cases have been reported — it has been linked to germline mutations in three other genes, whose products are postulated to facilitate DNA methylation: zinc finger and BTB domain-containing protein 24 (*ZBTB24*; defining ICF2), cell division cycle-associated protein 7 (*CDCA7*; ICF3) and *HELLS* (which encodes LSH; ICF4)[216,217].

Interestingly, although hypomethylation of pericentromeric satellite repeats is a hallmark of ICF, there are also ICF-type-specific defects that point towards differential genomic-context-dependent interactions of DNMT3B with ZBTB14, CDCA7 and LSH[218]. Notably, individuals with ICF1 uniquely show hypomethylation of CGI promoters of germline genes and X-linked genes, in agreement with DNMT3B targeting these sequences during mouse development[91,116] and indicating that this targeting may occur independently of ZBTB14, CDCA7 and LSH. By contrast, hypomethylation of centromeric α-satellite repeats and neural gene clusters distinguishes ICF2, ICF3 and ICF4 from ICF1. This may implicate ZBTB14, CDCA7 and LSH in the recruitment of a DNMT other than DNMT3B to these sequences. Finally, how whole-body hypomethylation of pericentromeric repeats specifically results in a defective immune system remains unclear. Modelling the ICF syndrome in mice is poorly informative owing to embryonic lethality and/or lack of immunoglobulin deficiency[219,220]. However, a recent study of *zbtb24*-mutant zebrafish reported that pericentromeric hypomethylation during embryonic development primarily activates the innate immune system through the derepression of pericentromeric transcripts[221]. Such an animal model may help refine our mechanistic understanding of ICF pathology in humans.

DNMT3A, cell growth and malignancies

Heterozygous, germline mutations in *DNMT3A* are associated with growth disorders of early prenatal onset, with contrasting growth phenotypes depending on the nature of the mutations (TABLE 1). Missense, gain-of-function mutations in the PWWP domain are found in individuals with microcephalic dwarfism[222] — a group of pathologies that manifest as a profound yet proportionate reduction in body size and head size. These are gain-of-function mutations, which disrupt interactions of DNMT3A with H3K36me3 moieties and alter its chromatin binding specificity: in fibroblasts and blood cells of individuals carrying the mutations, ectopic gain of DNA methylation is observed in large regions known as DNA methylation valleys or canyons[223,224], at the expense of the bivalent H3K27me3 and H3K4me3 modifications that normally characterize these regions and regulate important developmental genes, such as Hox transcription factors and morphogen genes. The introduction of a PWWP gain-of-function mutation in mice remarkably recapitulated the reductions in body size and brain weight in microcephalic dwarfism and the ectopic methylation phenotypes[222,225].

By contrast, heterozygous *DNMT3A* haploinsufficiency mutations characterize the Tatton-Brown–Rahman syndrome (TBRS; also known as DNMT3A-overgrowth syndrome; OMIM 602729), which combines

Neomorphic mutations
Typically, dominant mutations, which confer altered expression or novel function for the protein product.

macrocephalic overgrowth with moderate intellectual disability[226] (TABLE 1). TBRS mutations occur across the *DNMT3A* gene, including in the PWWP domain (upstream from the gain-of-function mutations), ADD domain and MTase domain. The effects of these mutations on DNA methylation patterns have not yet been determined, although the large genomic domains that are hypermethylated in the DNMT3A gain-of-function mutants are seemingly not affected in individuals with TBRS[222]. Given that DNMT3A loss of function results in stem cell hyperproliferation in various tissues in mice[227,228], *DNMT3A* haploinsufficiency may promote overgrowth in TBRS by increasing the number of cells in tissues and, consequently, body size. Inversely, DNMT3A gain of function could favour differentiation over stem cell proliferation, and, in turn, reduce organ and body size. The *DNMT3A* locus is strongly linked with natural height variation in genome-wide association studies[229], pointing towards an important role of DNMT3A in human size physiology.

Finally, somatic *DNMT3A* mutations are linked to enhanced proliferation of immature myeloid cells and the development of adult haematological malignancies. The mutations are found in ~15–35% of cases of acute myeloid leukaemia (AML; OMIM 601626), the vast majority being missense mutations in arginine 882 (R882), which is in the MTase domain. R882 mutations have dominant negative effects on the formation of wild-type DNMT3A tetramers, thereby reducing the methylation activity of the enzyme by 80% in vitro[230]. Consequently, very focal yet genome-wide hypomethylation is observed at specific sites in CpG-dense regions, which precedes the onset of AML[231]. Hypermethylation of bivalent CGI promoters was also reported, but is a consequence rather than a cause of AML progression, being secondary to rapid cellular proliferation of malignant myeloid cells. Interestingly, germline R882 mutations also exist and cause early-onset AML in individuals with TBRS[232].

Hydroxymethylation by TET2 and cancer

Whereas somatic loss of DNMT3A activity is fairly common in AML, *TET2* mutations are even more frequent causes of haematological malignancies, including AML, chronic myelomonocytic leukaemia, lymphomas and myeloproliferative neoplasms (TABLE 1). Acquisition of (mostly heterozygous) *TET2* mutations confers a proliferative advantage to haematopoietic stem cells and likely is an early driver of leukaemogenesis[233]. Moreover, restoration of TET2 function can reverse leukaemic progression, further supporting a causative role for the enzyme in the aetiology of the disease[234]. In contrast to the *DNMT3A* mutations that primarily decrease DNA methylation levels, *TET2* mutations promote hypermethylation in myeloid malignancies, mostly at enhancer regions[235,236]. How haploinsufficiency mutations in TET2 and DNMT3A — two enzymes with opposing functions — can cause similar malignancy phenotypes is puzzling and compels explanation. *TET2*-mutant cancer cells exhibit the expected reduction in 5hmC content[235,237]; in *DNMT3A*-mutant cancer cells, decreased 5hmC is also expected owing to the lower availability of 5mC for TET2. The 5hmC modification may therefore be the culprit of malignancy, by affecting gene regulation and/or other cellular processes through the recruitment of 5hmC readers[238]. A recent study of mice with mutations in both *Dnmt3A* and *Tet2* illustrates the complex epistasis between the two enzymes, which involves a combination of cooperative, competitive and independent activities[239].

TET2 is also indirectly impaired in multiple cancers due to genetic mutations in enzymes that regulate the production of metabolites necessary for its oxidative activity. TET2 requires 2-oxoglutarate as a co-substrate for catalysing the three-step 5mC hydroxylation reaction. Isocitrate dehydrogenase 1 (IDH1) and IDH2 promote TET2 activity by producing 2-oxoglutarate. In the vast majority of lower-grade gliomas, neomorphic mutations in IDH1 and IDH2 render them capable of producing 2-hydroxyglutarate. By competing with 2-oxoglutarate as a TET2 co-substrate, 2-hydroxyglutarate induces extensive gain of 5mC and transcriptomic changes that are relevant to the oncogenic phenotype, thereby classifying it as a potential oncometabolite[240]. In addition, as part of the Krebs cycle in mitochondria, succinate dehydrogenases (SDHs) convert the TET2 inhibitor succinate. Somatic mutations in *SDH* genes are found in gastrointestinal stromal tumours and neuroendocrine tumours (paragangliomas and pheochromocytomas), in association with increased intracellular levels of succinate, inhibition of TET2 activity and — similarly to *IDH* mutations — widespread hypermethylation within and outside CGIs[241–243].

In summary, the diseases associated with mutations in DNA methylation and demethylation factors offer a nuanced understanding of the functions of these proteins, beyond what can be achieved in loss-of-function studies. Continued research in these areas will hopefully lead to therapeutic breakthroughs for both cancer and genetic diseases.

Conclusions and future perspective

As Timothy Bestor — one of the fathers of the mammalian DNA methylation field — likes to note, the number of studies mentioning *DNMT1* has increased linearly since its discovery in the 1980s, and the scientific interest in DNA methylation has continued to grow. However, despite the major inroads made into understanding DNA methylation mechanisms and patterns, much remains unknown. For example, why do DNA methylation-deficient embryos die at such an early stage of development? It seems like a foregone conclusion that lack of DNA methylation would be lethal, but it is not clear what confers the embryonic failure: aberrant expression of protein-coding genes; massive derepression of transposable elements; genomic instability; a combination or all of the above or something less obvious? Furthermore, in cancer cells, DNA methylation is commonly deregulated and *DNMT3* and *TET* genes are frequently mutated. But which methylation changes are relevant for oncogenesis and which are only incidental?

We are now entering an era of unprecedented genetic tools, sensitive and highly quantitative sequencing technologies, and the ability to alter DNA methylation

with surgical precision (BOX 2). Major questions that have been traditionally difficult to address now seem possible to answer. Yet challenges remain. For example, when assessing DNA methylation patterns associated with a specific cell type or developmental stage, we can only infer steady-state levels of methylation that result from the opposing activities of the enzymes that generate and remove DNA methylation. There is a need for sequence-specific quantification of DNA methylation turnover. Combining experimental and theoretical approaches would allow modelling of the principles of local DNA methylation dynamics and help us to understand where and why imbalances may occur during ageing and tumorigenesis. Furthermore, a major conundrum in mammalian DNA methylation biology is the function of the extensive demethylation that occurs in the early embryo and germline, in particular whether it is required for the establishment of pluripotency. At present, it appears impossible to stymie this process to assess the developmental effects of DNA methylation.

As a final note, although the focus of this review is mice and humans, our understanding of mammalian DNA methylation has greatly benefited from studies in distant species such as flowering plants and filamentous fungi (reviewed elsewhere[244]). Comparative analyses of non-traditional model organisms provide valuable insights into conserved mechanisms and functions, and also surprises: in the milkweed bug, DNA methylation is important for fertility, but this may occur independently of the traditional role of DNA methylation in transcription control[245]. With the torrid pace of technological advancement and knowledge gleaned from studies in diverse systems, the future of mammalian DNA methylation research promises exciting and impactful discoveries ahead.

Published online: 09 August 2019

1. Zemach, A., McDaniel, I. E., Silva, P. & Zilberman, D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**, 916–919 (2010).
2. Feng, S. et al. Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl Acad. Sci. USA* **107**, 8689–8694 (2010).
3. Doskočil, J. & Šorm, F. Distribution of 5-methylcytosine in pyrimidine sequences of deoxyribonucleic acids. *Biochim. Biophys. Acta* **55**, 953–959 (1962).
4. Riggs, A. D. X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* **14**, 9–25 (1975).
5. Bird, A. P. Use of restriction enzymes to study eukaryotic DNA methylation. II. The symmetry of methylated sites supports semi-conservative copying of the methylation pattern. *J. Mol. Biol.* **118**, 49–60 (1978).
6. Ben-Hattar, J. & Jiricny, J. Methylation of single CpG dinucleotides within a promoter element of the Herpes simplex virus tk gene reduces its transcription in vivo. *Gene* **65**, 219–227 (1988).
7. Watt, F. & Molloy, P. L. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev.* **2**, 1136–1143 (1988).
8. Iguchi-Ariga, S. M. & Schaffner, W. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes Dev.* **3**, 612–619 (1989).
9. Ferguson-Smith, A. C., Sasaki, H., Cattanach, B. M. & Surani, M. A. Parental-origin-specific epigenetic modification of the mouse H19 gene. *Nature* **362**, 751–755 (1993).
10. Li, E., Beard, C. & Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature* **366**, 362–365 (1993).
11. Bartolomei, M. S., Webber, A. L., Brunkow, M. E. & Tilghman, S. M. Epigenetic mechanisms underlying the imprinting of the mouse H19 gene. *Genes Dev.* **7**, 1663–1673 (1993).
12. Stöger, R. et al. Maternal-specific methylation of the imprinted mouse Igf2r locus identifies the expressed locus as carrying the imprinting signal. *Cell* **73**, 61–71 (1993).
13. Mohandas, T., Sparkes, R. S. & Shapiro, L. J. Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science* **211**, 393–396 (1981).
14. Lock, L. F., Takagi, N. & Martin, G. R. Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. *Cell* **48**, 39–46 (1987).
15. Zemach, A. & Zilberman, D. Evolution of eukaryotic DNA methylation and the pursuit of safer sex. *Curr. Biol.* **20**, R780–R785 (2010).
16. Raddatz, G. et al. Dnmt2-dependent methylomes lack defined DNA methylation patterns. *Proc. Natl Acad. Sci. USA* **110**, 8627–8631 (2013).
17. Holliday, R. & Grigg, G. W. DNA methylation and mutation. *Mutat. Res.* **285**, 61–67 (1993).
18. Bird, A. P. & Taggart, M. H. Variable patterns of total DNA and rDNA methylation in animals. *Nucleic Acids Res.* **8**, 1485–1497 (1980).
19. Cooper, D. N. & Krawczak, M. Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. *Hum. Genet.* **83**, 181–188 (1989).
20. Rošić, S. et al. Evolutionary analysis indicates that DNA alkylation damage is a byproduct of cytosine DNA methyltransferase activity. *Nat. Genet.* **50**, 452–459 (2018).
21. Li, E. & Zhang, Y. DNA methylation in mammals. *Cold Spring Harb. Perspect. Biol.* **6**, a019133 (2014).
22. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
23. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
24. Baylin, S. B. & Jones, P. A. Epigenetic determinants of cancer. *Cold Spring Harb. Perspect. Biol.* **8**, a019505 (2016).
25. Arand, J. et al. In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. *PLOS Genet.* **8**, e1002750 (2012).
26. Walsh, C. P., Chaillet, J. R. & Bestor, T. H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20**, 116–117 (1998).
27. Borgel, J. et al. Targets and dynamics of promoter DNA methylation during early mouse development. *Nat. Genet.* **42**, 1093–1100 (2010).
28. 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).
29. Lister, R. et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
30. Monk, M., Boubelik, M. & Lehnert, S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**, 371–382 (1987).
31. Sanford, J. P., Clark, H. J., Chapman, V. M. & Rossant, J. Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes Dev.* **1**, 1039–1046 (1987).
32. Okano, M., Xie, S. & Li, E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.* **19**, 219–220 (1998).
33. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
34. Ooi, S. K. et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714–717 (2007).
35. Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B. & Bestor, T. H. Dnmt3L and the establishment of maternal genomic imprints. *Science* **294**, 2536–2539 (2001).
36. Piunti, A. & Shilatifard, A. Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science* **352**, aad9780 (2016).
37. Otani, J. et al. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep.* **10**, 1235–1241 (2009).
38. Zhang, Y. et al. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res.* **38**, 4246–4253 (2010).
39. Guo, X. et al. Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature* **517**, 640–644 (2015).
40. Krogan, N. J. et al. Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol.* **23**, 4207–4218 (2003).
41. Sun, X. J. et al. Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *J. Biol. Chem.* **280**, 35261–35271 (2005).
42. Dhayalan, A. et al. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J. Biol. Chem.* **285**, 26114–26120 (2010).
43. Baubec, T. et al. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520**, 243–247 (2015).
44. Xu, Q. et al. SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. *Nat. Genet.* **51**, 844–856 (2019).
45. Bostick, M. et al. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760–1764 (2007).
46. Sharif, J. et al. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908–912 (2007).
47. Nady, N. et al. Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. *J. Biol. Chem.* **286**, 24300–24311 (2011).
48. Arita, K. et al. Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1. *Proc. Natl Acad. Sci. USA* **109**, 12950–12955 (2012).
49. Rothbart, S. B. et al. Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat. Struct. Mol. Biol.* **19**, 1155–1160 (2012).
50. Rothbart, S. B. et al. Multivalent histone engagement by the linked tandem tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation. *Genes Dev.* **27**, 1288–1298 (2013).
51. Song, J., Rechkoblit, O., Bestor, T. H. & Patel, D. J. Structure of DNMT1–DNA complex reveals a role for autoinhibition in maintenance DNA methylation. *Science* **331**, 1036–1040 (2011).
52. Takeshita, K. et al. Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1). *Proc. Natl Acad. Sci. USA* **108**, 9055–9059 (2011).

53. Ishiyama, S. et al. Structure of the Dnmt1 reader module complexed with a unique two-mono-ubiquitin mark on Histone H3 reveals the basis for DNA methylation maintenance. *Mol. Cell* **68**, 350–360 (2017).
54. Nishiyama, A. et al. Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature* **502**, 249–253 (2013).
55. Qin, W. et al. DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination. *Cell Res.* **25**, 911–929 (2015).
56. Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in purkinje neurons and the brain. *Science* **324**, 929–930 (2009).
57. Tahiliani, M. et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
58. Ito, S. et al. Role of tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
59. Ito, S. et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303 (2011).
60. He, Y. F. et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011).
61. Hashimoto, H. et al. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res.* **40**, 4841–4849 (2012).
62. Otani, J. et al. Cell cycle-dependent turnover of 5-hydroxymethyl cytosine in mouse embryonic stem cells. *PLOS ONE* **8**, e82961 (2013).
63. Maiti, A. & Drohat, A. C. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: Potential implications for active demethylation of CpG sites. *J. Biol. Chem.* **286**, 35334–35338 (2011).
64. Weber, A. R. et al. Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. *Nat. Commun.* **7**, 10806 (2016).
65. Weber, M. et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **39**, 457–466 (2007).
66. Stadler, M. B. et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495 (2011).
67. Yin, Y. et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* **356**, eaaj2239 (2017).
68. Dennis, K., Fan, T., Geiman, T., Yan, Q. & Muegge, K. Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes Dev.* **15**, 2940–2944 (2001).
69. Myant, K. et al. LSH and G9a/GLP complex are required for developmentally programmed DNA methylation. *Genome Res.* **21**, 83–94 (2011).
70. Tao, Y. et al. Lsh, chromatin remodeling family member, modulates genome-wide cytosine methylation patterns at nonrepeat sequences. *Proc. Natl Acad. Sci. USA* **108**, 5626–5631 (2011).
71. Esteve, P. O. et al. Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev.* **20**, 3089–3103 (2006).
72. Epsztejn-Litman, S. et al. De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat. Struct. Mol. Biol.* **15**, 1176–1183 (2008).
73. Fuks, F., Burgers, W. A., Brehm, A., Hughes-Davies, L. & Kouzarides, T. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat. Genet.* **24**, 88–91 (2000).
74. Fuks, F., Burgers, W. A., Godin, N., Kasai, M. & Kouzarides, T. Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J.* **20**, 2536–2544 (2001).
75. Deplus, R. et al. Dnmt3L is a transcriptional repressor that recruits histone deacetylase. *Nucleic Acids Res.* **30**, 3831–3838 (2002).
76. Meehan, R. R., Lewis, J. D., McKay, S., Kleiner, E. L. & Bird, A. P. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **58**, 499–507 (1989).
77. Hendrich, B. & Bird, A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol* **18**, 6538–6547 (1998).
78. Baubec, T., Ivánek, R., Lienert, F. & Schübeler, D. methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell* **153**, 480–492 (2013).
79. Saito, M. & Ishikawa, F. The mCpG-binding domain of human MBD3 does not bind to mCpG but interacts with NuRD/Mi2 components HDAC1 and MTA2. *J. Biol. Chem.* **277**, 35434–35439 (2002).
80. Nan, X. et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386–389 (1998).
81. Ng, H. H. et al. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat. Genet.* **23**, 58–61 (1999).
82. Ren, R., Horton, J. R., Zhang, X., Blumenthal, R. M. & Cheng, X. Detecting and interpreting DNA methylation marks. *Curr. Opin. Struct. Biol.* **53**, 88–99 (2018).
83. Gardiner-Garden, M. & Frommer, M. CpG Islands in vertebrate genomes. *J. Mol. Biol.* **196**, 261–282 (1987).
84. Larsen, F., Gundersen, G., Lopez, R. & Prydz, H. CpG islands as gene markers in the human genome. *Genomics* **13**, 1095–1107 (1992).
85. Ku, M. et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLOS Genet.* **4**, e1000242 (2008).
86. 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* **40**, 91–99 (1985).
87. Marasca, F., Bodega, B. & Orlando, V. How polycomb-mediated cell memory deals with a changing environment. *BioEssays* **40**, e1700137 (2018).
88. Singer-Sam, J. et al. Use of a HpaII-polymerase chain reaction assay to study DNA methylation in the Pgk-1 CpG island of mouse embryos at the time of X-chromosome inactivation. *Mol. Cell. Biol.* **10**, 4987–4989 (1990).
89. Grant, M., Zuccotti, M. & Monk, M. Methylation of CpG sites of two X-linked genes coincides with X-inactivation in the female mouse embryo but not in the germ line. *Nat. Genet.* **2**, 161–166 (1992).
90. Keohane, A. M., O'Neill, L. P., Belyaev, N. D., Lavender, J. S. & Turner, B. M. X-inactivation and histone H4 acetylation in embryonic stem cells. *Dev. Biol.* **180**, 618–630 (1996).
91. Gendrel, A.-V. et al. Smchd1-dependent and -independent pathways determine developmental dynamics of CpG Island methylation on the inactive X chromosome. *Dev. Cell* **23**, 265–279 (2012).
92. Blewitt, M. E. et al. SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat. Genet.* **40**, 663–669 (2008).
93. Gdula, M. R. et al. The non-canonical SMC protein SmcHD1 antagonises TAD formation and compartmentalisation on the inactive X chromosome. *Nat. Commun.* **10**, 30 (2019).
94. Nozawa, R. S. et al. Human inactive X chromosome is compacted through a PRC2-independent SMCHD1-HBiX1 pathway. *Nat. Struct. Mol. Biol.* **20**, 566–573 (2013).
95. Wutz, A. et al. Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature* **389**, 745–749 (1997).
96. Thorvaldsen, J. L., Duran, K. L. & Bartolomei, M. S. Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. *Genes Dev.* **12**, 3693–3702 (1998).
97. Yang, T. et al. A mouse model for Prader–Willi syndrome imprinting-centre mutations. *Nat. Genet.* **19**, 25–31 (1998).
98. Fitzpatrick, G. V., Soloway, P. D. & Higgins, M. J. Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nat. Genet.* **32**, 426–431 (2002).
99. Lin, S. P. et al. Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the Dlk1-Gtl2 imprinted cluster on mouse chromosome 12. *Nat. Genet.* **35**, 97–102 (2003).
100. Williamson, C. M. et al. Identification of an imprinting control region affecting the expression of all transcripts in the Gnas cluster. *Nat. Genet.* **38**, 350–355 (2006).
101. Proudhon, C. et al. Protection against de novo methylation is instrumental in maintaining parent-of-origin methylation inherited from the gametes. *Mol. Cell* **47**, 909–920 (2012).
102. Kaneda, M. et al. Genetic evidence for Dnmt3a-dependent imprinting during oocyte growth obtained by conditional knockout with Zp3-Cre and complete exclusion of Dnmt3b by chimera formation. *Genes Cells* **15**, 169–179 (2010).
103. Chotalia, M. et al. Transcription is required for establishment of germline methylation marks at imprinted genes. *Genes Dev.* **23**, 105–117 (2009).
104. Smith, E. Y., Futtner, C. R., Chamberlain, S. J., Johnstone, K. A. & Resnick, J. L. Transcription is required to establish maternal imprinting at the Prader–Willi syndrome and Angelman syndrome locus. *PLOS Genet.* **7**, e1002422 (2011).
105. Smallwood, S. A. et al. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat. Genet.* **43**, 811–814 (2011).
106. Veselovska, L. et al. Deep sequencing and de novo assembly of the mouse oocyte transcriptome define the contribution of transcription to the DNA methylation landscape. *Genome Biol.* **16**, 209 (2015).
107. Peaston, A. E. et al. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell* **7**, 597–606 (2004).
108. Franke, V. et al. Long terminal repeats power evolution of genes and gene expression programs in mammalian oocytes and zygotes. *Genome Res.* **27**, 1384–1394 (2017).
109. Brind'Amour, J. et al. LTR retrotransposons transcribed in oocytes drive species-specific and heritable changes in DNA methylation. *Nat. Commun.* **9**, 3331 (2018).
110. Li, X. et al. A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev. Cell* **15**, 547–557 (2008).
111. Quenneville, S. et al. In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol. Cell* **44**, 361–372 (2011).
112. Strogantsev, R. et al. Allele-specific binding of ZFP57 in the epigenetic regulation of imprinted and non-imprinted monoallelic expression. *Genome Biol.* **16**, 112 (2015).
113. Messerschmidt, D. M. et al. Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science* **335**, 1499–1502 (2012).
114. Galonska, C., Ziller, M. J., Karnik, R. & Meissner, A. Ground State Conditions Induce Rapid Reorganization of Core Pluripotency Factor Binding before Global Epigenetic Reprogramming. *Cell Stem Cell* **17**, 462–470 (2015).
115. Takahashi, N. et al. ZNF445 is a primary regulator of genomic imprinting. *Genes Dev.* **33**, 49–54 (2019).
116. Auclair, G., Guibert, S., Bender, A. & Weber, M. Ontogeny of CpG island methylation and specificity of DNMT3 methyltransferases during embryonic development in the mouse. *Genome Biol.* **15**, 545 (2014).
117. Karimi, M. M. et al. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* **8**, 676–687 (2011).
118. O'Neill, K. M. et al. Depletion of DNMT1 in differentiated human cells highlights key classes of sensitive genes and an interplay with polycomb repression. *Epigenetics Chromatin* **11**, 12 (2018).
119. Auclair, G. et al. EHMT2 directs DNA methylation for efficient gene silencing in mouse embryos. *Genome Res.* **26**, 192–202 (2016).
120. Velasco, G. & Francastel, C. Genetics meets DNA methylation in rare diseases. *Clin. Genet.* **95**, 210–220 (2019).
121. Endoh, M. et al. PCGF6-PRC1 suppresses premature differentiation of mouse embryonic stem cells by regulating germ cell-related genes. *eLife* **6**, e21064 (2017).
122. Stielow, B., Finkernagel, F., Stiewe, T., Nist, A. & Suske, G. MGA, L3MBTL2 and E2F6 determine genomic binding of the non-canonical Polycomb repressive complex PRC1.6. *PLOS Genet.* **14**, e1007193 (2018).
123. Gao, Z. et al. PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol. Cell* **45**, 344–356 (2012).
124. Maeda, I. et al. Max is a repressor of germ cell-related gene expression in mouse embryonic stem cells. *Nat. Commun.* **4**, 1754 (2013).
125. Velasco, G. et al. Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues. *Proc. Natl Acad. Sci. USA* **107**, 9281–9286 (2010).
126. Qin, J. et al. The polycomb group protein L3mbtl2 assembles an atypical PRC1-family complex that is essential in pluripotent stem cells and early development. *Cell Stem Cell* **11**, 319–332 (2012).
127. Trojer, P. et al. L3MBTL2 protein acts in concert with PcG protein-mediated monoubiquitination of H2A to establish a repressive chromatin structure. *Mol. Cell* **42**, 438–450 (2011).
128. Yoder, J. A., Walsh, C. P. & Bestor, T. H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet* **13**, 335–340 (1997).

129. Barau, J. et al. The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science* **354**, 909–912 (2016).
130. Sanchez-Delgado, M. et al. Human oocyte-derived methylation differences persist in the placenta revealing widespread transient imprinting. *PLOS Genet.* **12**, e1006427 (2016).
131. Walsh, C. P., Chaillet, J. R. & Bestor, T. H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20**, 116–117 (1998).
132. Jain, D. et al. rahu is a mutant allele of Dnmt3c, encoding a DNA methyltransferase homolog required for meiosis and transposon repression in the mouse male germline. *PLOS Genet.* **13**, e1006964 (2017).
133. Kuramochi-Miyagawa, S. et al. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev.* **22**, 908–917 (2008).
134. Aravin, A. A. et al. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol Cell* **31**, 785–799 (2008).
135. Aravin, A. A., Hannon, G. J. & Brennecke, J. The Piwi–piRNA pathway provides an adaptive defense in the transposon arms race. *Science* **318**, 761–764 (2007).
136. Carmell, M. A. et al. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev. Cell* **12**, 503–514 (2007).
137. Molaro, A. et al. Two waves of de novo methylation during mouse germ cell development. *Genes Dev.* **28**, 1544–1549 (2014).
138. Manakov, S. A. et al. MIWI2 and MILI have differential effects on piRNA biogenesis and DNA methylation. *Cell Rep.* **12**, 1234–1243 (2015).
139. Barau, J. *et al*. The novel DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science* **354**, 909–912 (2016).
140. Jordà, M. et al. The epigenetic landscape of Alu repeats delineates the structural and functional genomic architecture of colon cancer cells. *Genome Res.* **27**, 118–132 (2017).
141. Zamudio, N. et al. DNA methylation restrains transposons from adopting a chromatin signature permissive for meiotic recombination. *Genes Dev.* **29**, 1256–1270 (2015).
142. Nishibuchi, G. & Déjardin, J. The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. *Chromosom. Res.* **25**, 77–87 (2017).
143. Hutnick, L. K., Huang, X., Loo, T.-C., Ma, Z. & Fan, G. Repression of retrotransposal elements in mouse embryonic stem cells is primarily mediated by a DNA methylation-independent mechanism. *J. Biol. Chem.* **285**, 21082–21091 (2010).
144. Matsui, T. et al. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **464**, 927 (2010).
145. Walter, M. et al. An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. *eLife* **5**, e11418 (2016).
146. Sharif, J. et al. Activation of endogenous retroviruses in Dnmt1−/− ESCs involves disruption of SETDB1-mediated repression by NP95 binding to hemimethylated DNA. *Cell Stem Cell* **19**, 81–94 (2016).
147. Berrens, R. V. et al. An endosiRNA-based repression mechanism counteracts transposon activation during global DNA demethylation in embryonic stem cells. *Cell Stem Cell* **21**, 694–703 (2017).
148. Bender, C. M. et al. Roles of cell division and gene transcription in the methylation of CpG islands. *Mol. Cell. Biol.* **19**, 6690–6698 (1999).
149. Varley, K. E. et al. Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res.* **23**, 555–567 (2013).
150. Laurent, L. et al. Dynamic changes in the human methylome during differentiation. *Genome Res.* **20**, 320–331 (2010).
151. Gelfman, S., Cohen, N., Yearim, A. & Ast, G. DNA-methylation effect on cotranscriptional splicing is dependent on GC architecture of the exon-intron structure. *Genome Res.* **23**, 789–799 (2013).
152. Shayevitch, R., Askayo, D., Keydar, I. & Ast, G. The importance of DNA methylation of exons on alternative splicing. *RNA* **24**, 1351–1362 (2018).
153. Yang, X. et al. Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer Cell* **26**, 577–590 (2014).
154. Shukla, S. et al. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* **479**, 74–79 (2011).
155. Maunakea, A. K., Chepelev, I., Cui, K. & Zhao, K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res.* **23**, 1256–1269 (2013).
156. Yearim, A. et al. HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. *Cell Rep.* **10**, 1122–1134 (2015).
157. Carrozza, M. J. et al. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**, 581–592 (2005).
158. Maunakea, A. K. et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**, 253–257 (2010).
159. Neri, F. et al. Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**, 72–77 (2017).
160. Teissandier, A. & Bourc'his, D. Gene body DNA methylation conspires with H3K36me3 to preclude aberrant transcription. *EMBO J.* **36**, 1471–1473 (2017).
161. Guibert, S., Forné, T. & Weber, M. Global profiling of DNA methylation erasure in mouse primordial germ cells. *Genome Res.* **22**, 633–641 (2012).
162. Vincent, J. J. et al. Stage-specific roles for Tet1 and Tet2 in DNA demethylation in primordial germ cells. *Cell Stem Cell* **12**, 470–478 (2013).
163. Yamaguchi, S. et al. Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. *Cell Res.* **23**, 329–339 (2013).
164. Yamaguchi, S. et al. Tet1 controls meiosis by regulating meiotic gene expression. *Nature* **492**, 443–447 (2012).
165. Yamaguchi, S., Shen, L., Liu, Y., Sendler, D. & Zhang, Y. Role of Tet1 in erasure of genomic imprinting. *Nature* **504**, 460–464 (2013).
166. Hackett, J. A. et al. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* **339**, 448–452 (2013).
167. Gu, T.-P. et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606–610 (2011).
168. Wossidlo, M. et al. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun.* **2**, 241 (2011).
169. Iqbal, K., Jin, S.-G., Pfeifer, G. P. & Szabo, P. E. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl Acad. Sci. USA* **108**, 3642–3647 (2011).
170. Howell, C. Y. et al. Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell* **104**, 829–838 (2001).
171. Wang, L. et al. Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**, 979–991 (2014).
172. Guo, F. et al. Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell* **15**, 447–459 (2014).
173. Shen, L. et al. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell Stem Cell* **15**, 459–470 (2014).
174. Amouroux, R. et al. De novo DNA methylation drives 5hmC accumulation in mouse zygotes. *Nat. Cell Biol.* **18**, 225–233 (2016).
175. Santos, F. et al. Active demethylation in mouse zygotes involves cytosine deamination and base excision repair. *Epigenetics Chromatin* **6**, 39 (2013).
176. Smith, Z. D. et al. DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611–615 (2014).
177. Okae, H. et al. Genome-wide analysis of DNA methylation dynamics during early human development. *PLOS Genet.* **10**, e1004868 (2014).
178. Guo, H. et al. The DNA methylation landscape of human early embryos. *Nature* **511**, 606 (2014).
179. Zhu, P. et al. Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat. Genet.* **50**, 12–19 (2018).
180. Hill, S. et al. Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte. *Nature* **555**, 392–396 (2018).
181. Skvortsova, K., Iovino, N. & Bogdanović, O. Functions and mechanisms of epigenetic inheritance in animals. *Nat. Rev. Mol. Cell Biol.* **19**, 774–790 (2018).
182. Duffié, R. et al. The Gpr1/Zdbf2 locus provides new paradigms for transient and dynamic genomic imprinting in mammals. *Genes Dev.* **28**, 463–478 (2014).
183. Greenberg, M. V. C. et al. Transient transcription in the early embryo sets an epigenetic state that programs postnatal growth. *Nat. Genet.* **49**, 110–118 (2017).
184. Lane, N. et al. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* **35**, 88–93 (2003).
185. Smith, Z. D. et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).
186. Rowe, H. M. et al. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**, 237–240 (2010).
187. Ecco, G., Imbeault, M. & Trono, D. KRAB zinc finger proteins. *Development* **144**, 2719–2729 (2017).
188. Kazachenka, A. et al. Identification, characterization, and heritability of murine metastable epialleles: implications for non-genetic inheritance. *Cell* **175**, 1259–1271 (2018).
189. Gkountela, S. et al. DNA demethylation dynamics in the human prenatal germline. *Cell* **161**, 1425–1436 (2015).
190. Tang, W. W. C. et al. A unique gene regulatory network resets the human germline epigenome for development. *Cell* **161**, 1453–1467 (2015).
191. Hajkova, P. et al. Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117**, 15–23 (2002).
192. Kobayashi, H. et al. Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. *PLOS Genet.* **8**, e1002440 (2012).
193. Li, Y. et al. Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. *Nature* **564**, 136–140 (2018).
194. Smith, Z. D. et al. Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. *Nature* **549**, 543–547 (2017).
195. Zhang, Y. et al. Dynamic epigenomic landscapes during early lineage specification in mouse embryos. *Nat. Genet.* **50**, 96–105 (2018).
196. Gama-Sosa, M. A. et al. Tissue-specific differences in DNA methylation in various mammals. *Biochim. Biophys. Acta* **740**, 212–219 (1983).
197. Hon, G. C. et al. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nat. Genet.* **45**, 1198–1206 (2013).
198. Seisenberger, S. et al. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol. Cell* **48**, 849–862 (2012).
199. Schultz, M. D. et al. Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* **523**, 212–216 (2015).
200. Ziller, M. J. et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**, 477–481 (2013).
201. Sardina, J. L. et al. Transcription factors drive Tet2-mediated enhancer demethylation to reprogram cell fate. *Cell Stem Cell* **23**, 727–741 (2018).
202. Rasmussen, K. D. et al. TET2 binding to enhancers facilitates transcription factor recruitment in hematopoietic cells. *Genome Res.* **29**, 564–575 (2019).
203. Lister, R. et al. Global epigenomic reconfiguration during mammalian brain development. *Science* **341**, 6146 (2013).
204. Guo, J. U. et al. Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat. Neurosci.* **17**, 215–222 (2014).
205. Gabel, H. W. et al. Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. *Nature* **522**, 89–93 (2015).
206. Lagger, S. et al. MeCP2 recognizes cytosine methylated tri-nucleotide and di-nucleotide sequences to tune transcription in the mammalian brain. *PLOS Genet.* **13**, e1006793 (2017).
207. Stroud, H. et al. Early-life gene expression in neurons modulates lasting epigenetic states. *Cell* **171**, 1151–1164 (2017).
208. Ziller, M. J. et al. Dissecting the functional consequences of *de novo* DNA methylation dynamics in human motor neuron differentiation and physiology. *Cell Stem Cell* **22**, 559–574 (2018).
209. Klein, C. J. et al. Mutations in DNMT1 cause hereditary sensory neuropathy with dementia and hearing loss. *Nat. Genet.* **43**, 595 (2011).
210. Winkelmann, J. et al. Mutations in DNMT1 cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. *Hum. Mol. Genet.* **21**, 2205–2210 (2012).

211. Baets, J. et al. Defects of mutant DNMT1 are linked to a spectrum of neurological disorders. *Brain* **138**, 845–861 (2015).
212. Sun, Z. et al. Aberrant signature methylome by DNMT1 hot spot mutation in hereditary sensory and autonomic neuropathy 1E AU. *Epigenetics* **9**, 1184–1193 (2014).
213. Kernohan, K. D. et al. Identification of a methylation profile for DNMT1-associated autosomal dominant cerebellar ataxia, deafness, and narcolepsy. *Clin. Epigenetics* **8**, 91 (2016).
214. 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).
215. Xu, G. L. et al. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402**, 187–191 (1999).
216. Thijssen, P. E. et al. Mutations in CDCA7 and HELLS cause immunodeficiency–centromeric instability–facial anomalies syndrome. *Nat. Commun.* **6**, 7870 (2015).
217. de Greef, J. C. et al. Mutations in ZBTB24 are associated with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *Am. J. Hum. Genet.* **88**, 796–804 (2011).
218. Velasco, G. et al. Comparative methylome analysis of ICF patients identifies heterochromatin loci that require ZBTB24, CDCA7 and HELLS for their methylated state. *Hum. Mol. Genet.* **27**, 2409–2424 (2018).
219. Ueda, Y. et al. Roles for Dnmt3b in mammalian development: a mouse model for the ICF syndrome. *Development* **133**, 1183–1192 (2006).
220. Wu, H. et al. Converging disease genes in ICF syndrome: ZBTB24 controls expression of CDCA7 in mammals. *Hum. Mol. Genet.* **25**, 4041–4051 (2016).
221. Rajshekar, S. et al. Pericentromeric hypomethylation elicits an interferon response in an animal model of ICF syndrome. *eLife* **7**, e39658 (2018).
222. Heyn, P. et al. Gain-of-function DNMT3A mutations cause microcephalic dwarfism and hypermethylation of Polycomb-regulated regions. *Nat. Genet.* **51**, 96–105 (2019).
223. Li, Y. et al. Genome-wide analyses reveal a role of Polycomb in promoting hypomethylation of DNA methylation valleys. *Genome Biol.* **19**, 18 (2018).
224. Jeong, M. et al. Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nat. Genet.* **46**, 17–23 (2014).
225. Sendžikaitė, G., Hanna, C. W., Stewart-Morgan, K. R., Ivanova, E. & Kelsey, G. A DNMT3A PWWP mutation leads to methylation of bivalent chromatin and growth retardation in mice. *Nat. Commun.* **10**, 1884 (2019).
226. Tatton-Brown, K. et al. Mutations in the DNA methyltransferase gene DNMT3A cause an overgrowth syndrome with intellectual disability. *Nat. Genet.* **46**, 385–388 (2014).
227. Challen, G. A. et al. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat. Genet.* **44**, 23–31 (2012).
228. Wu, H. et al. Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* **329**, 444–448 (2010).
229. Tatton-Brown, K. et al. Mutations in epigenetic regulation genes are a major cause of overgrowth with intellectual disability. *Am. J. Hum. Genet.* **100**, 725–736 (2017).
230. Russler-Germain, D. A. et al. The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. *Cancer Cell* **25**, 442–454 (2014).
231. Spencer, D. H. et al. CpG island hypermethylation mediated by DNMT3A is a consequence of AML progression. *Cell* **168**, 801–816 (2017).
232. Kosaki, R., Terashima, H., Kubota, M. & Kosaki, K. Acute myeloid leukemia-associated DNMT3A p. Arg882His mutation in a patient with Tatton-Brown–Rahman overgrowth syndrome as a constitutional mutation. *Am. J. Med. Genet. Part A* **173**, 250–253 (2017).
233. Langemeijer, S. M. C. et al. Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat. Genet.* **41**, 838–842 (2009).
234. Cimmino, L. et al. Restoration of TET2 function blocks aberrant self-renewal and leukemia progression. *Cell* **170**, 1079–1095 (2017).
235. Ko, M. et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839–843 (2010).
236. Rasmussen, K. D. et al. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. *Genes Dev.* **29**, 910–922 (2015).
237. Madzo, J. et al. Hydroxymethylation at gene regulatory regions directs stem/early progenitor cell commitment during erythropoiesis. *Cell Rep.* **6**, 231–244 (2014).
238. Spruijt, C. G. et al. Dynamic readers for 5-(hydroxy) methylcytosine and its oxidized derivatives. *Cell* **152**, 1146–1159 (2013).
239. Zhang, X. et al. DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors in hematopoietic stem cells. *Nat. Genet.* **48**, 1014–1023 (2016).
240. Turcan, S. et al. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* **483**, 479–483 (2012).
241. Xiao, M. et al. Inhibition of α-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev.* **26**, 1326–1338 (2012).
242. Letouzé, E. et al. SDH mutations establish a hypermethylator phenotype in paraganglioma. *Cancer Cell* **23**, 739–752 (2013).
243. Killian, J. K. et al. Succinate dehydrogenase mutation underlies global epigenomic divergence in gastrointestinal stromal tumor. *Cancer Discov.* **3**, 648–657 (2013).
244. Zhang, H., Lang, Z. & Zhu, J.-K. Dynamics and function of DNA methylation in plants. *Nat. Rev. Mol. Cell Biol.* **19**, 489–506 (2018).
245. Bewick, A. J. et al. Dnmt1 is essential for egg production and embryo viability in the large milkweed bug, *Oncopeltus fasciatus*. *Epigenetics Chromatin* **12**, 6 (2019).
246. Hu, S. et al. DNA methylation presents distinct binding sites for human transcription factors. *eLife* **2**, e00726 (2013).
247. Liu, Y. et al. Structural basis for Klf4 recognition of methylated DNA. *Nucleic Acids Res.* **42**, 4859–4867 (2014).
248. Hashimoto, H. et al. Distinctive Klf4 mutants determine preference for DNA methylation status. *Nucleic Acids Res.* **44**, 10177–10185 (2016).
249. Rishi, V. et al. CpG methylation of half-CRE sequences creates C/EBPα binding sites that activate some tissue-specific genes. *Proc. Natl Acad. Sci. USA* **107**, 20311–20316 (2010).
250. Tanay, A., O'Donnell, A. H., Damelin, M. & Bestor, T. H. Hyperconserved CpG domains underlie Polycomb-binding sites. *Proc. Natl Acad. Sci. USA* **104**, 5521–5526 (2007).
251. Brinkman, A. B. et al. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res.* **22**, 1128–1138 (2012).
252. Statham, A. L. et al. Bisulfite sequencing of chromatin immunoprecipitated DNA (BisChIP-seq) directly informs methylation status of histone-modified DNA. *Genome Res.* **22**, 1120–1127 (2012).
253. Jermann, P., Hoerner, L., Burger, L. & Schubeler, D. Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. *Proc. Natl Acad. Sci. USA* **111**, E3415–E3421 (2014).
254. Hon, G. C. et al. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Res.* **22**, 246–258 (2012).
255. Bahar Halpern, K., Vana, T. & Walker, M. D. Paradoxical role of DNA methylation in activation of FoxA2 gene expression during endoderm development. *J. Biol. Chem.* **289**, 23882–23892 (2014).
256. Takahashi, Y. et al. Integration of CpG-free DNA induces de novo methylation of CpG islands in pluripotent stem cells. *Science* **356**, 503–508 (2017).
257. Li, F. et al. Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes. *Nucleic Acids Res.* **35**, 100–112 (2007).
258. Bernstein, D. L., Le Lay, J. E., Ruano, E. G. & Kaestner, K. H. TALE-mediated epigenetic suppression of CDKN2A increases replication in human fibroblasts. *J. Clin. Invest.* **125**, 1998–2006 (2015).
259. Maeder, M. L. et al. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat. Biotechnol.* **31**, 1137–1142 (2013).
260. Valton, J. et al. Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *J. Biol. Chem.* **287**, 38427–38432 (2012).
261. Vojta, A. et al. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res* **44**, 5615–28 (2016).
262. Liu, X. S. et al. Editing DNA methylation in the mammalian genome. *Cell* **167**, 233–247 (2016).
263. Lei, Y. et al. Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein. *Nat. Commun.* **8**, 16026 (2017).
264. Stepper, P. et al. Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. *Nucleic Acids Res.* **45**, 1703–1713 (2017).
265. Huang, Y.-H. et al. DNA epigenome editing using CRISPR-Cas SunTag-directed DNMT3A. *Genome Biol.* **18**, 176 (2017).
266. Pflueger, C. et al. A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs. *Genome Res.* **28**, 1193–1206 (2018).
267. Morita, S. et al. Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. *Nat. Biotechnol.* **34**, 1060–1065 (2016).
268. Amabile, A. et al. Inheritable silencing of endogenous genes by hit-and-run targeted epigenetic editing. *Cell* **167**, 219–232 (2016).
269. Saunderson, E. A. et al. Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors. *Nat. Commun.* **8**, 1450 (2017).

Acknowledgements
The laboratory of D.B. is part of the Laboratoire d'Excellence (LABEX) entitled DEEP (11-LBX0044) and is supported by the European Research Council (ERC) (grant ERC-Cog EpiRepro).

Author contributions
The authors contributed equally to all aspects of the article.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Molecular Cell Biology thanks G. Kelsey, R. Lister and T. Nakano for their contribution to the peer review of this work.

Supplementary information
Supplementary information is available for this paper at https://doi.org/10.1038/s41580-019-0159-6.