Review

Author for correspondence:
M. Azim Surani
e-mail: a.surani@gurdon.cam.ac.uk

Royal Society Publishing
Informing the science of the future

DNA methylation dynamics during the mammalian life cycle

Jamie A. Hackett and M. Azim Surani

Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge CB2 1QN, UK

DNA methylation is dynamically remodelled during the mammalian life cycle through distinct phases of reprogramming and de novo methylation. These events enable the acquisition of cellular potential followed by the maintenance of lineage-restricted cell identity, respectively, a process that defines the life cycle through successive generations. DNA methylation contributes to the epigenetic regulation of many key developmental processes including genomic imprinting, X-inactivation, genome stability and gene regulation. Emerging sequencing technologies have led to recent insights into the dynamic distribution of DNA methylation during development and the role of this epigenetic mark within distinct genomic contexts, such as at promoters, exons or imprinted control regions. Additionally, there is a better understanding of the mechanistic basis of DNA demethylation during epigenetic reprogramming in primordial germ cells and during pre-implantation development. Here, we discuss our current understanding of the developmental roles and dynamics of this key epigenetic system.

1. Introduction

The mammalian life cycle is characterized by extended phases of cellular differentiation, followed by distinct periods of reprogramming in germ cells and the zygote, which resets genomic potential to restart the cycle. Cell differentiation is directed by transcription factor (TF) networks, which establish gene expression patterns in response to developmental cues. Once established, core TF networks maintain robust lineage-restriction and ensure unidirectional development towards defined differentiated cell-types. In parallel, epigenetic mechanisms act to reinforce cell-fate decisions and establish barriers against reversion to preceding cellular states [1]. Multiple epigenetic systems including DNA methylation, histone modifications and small RNAs function both cooperatively and autonomously to establish the epigenome that propagates gene expression profiles through mitosis/meiosis. Disruption of these epigenetic mechanisms can lead to loss of cell identity, cellular transformation or disease, and is generally incompatible with normal development. However, at two unique junctures of mammalian development—in primordial germ cells (PGCs) and pre-implantation embryos—the epigenome is comprehensively reprogrammed towards a basal state to enable relaxation of the epigenetic constraints imposed on cell potential and ultimately, resetting of the life cycle [2]. Of particular recent interest are the mechanisms which erase and establish DNA methylation during development and the roles of these epigenetic events.

In mammals, DNA methylation (5mC) is essential for embryonic development and plays important roles in genomic imprinting, transposon silencing, X-inactivation and gene repression [3]. During development 5mC patterns are established by the de novo methyltransferases DNMT3A and DNMT3B, and subsequently preserved through cell divisions by DNMT1 [4,5]. As such, DNA methylation represents a true epigenetic mechanism because 5mC is faithfully maintained in the absence of the inducing signal. DNA methylation primarily occurs at cytosine residues within CpG dinucleotides, although it is becoming increasingly evident that there may be significant methylation in non-CpG (CpH) contexts, at least in pluripotent cell-types and oocytes [6–8]. Owing to the inherent mutagenicity of methylated cytosine by deamination to

thymidine, the bulk of the genome is depleted of CpG motifs, with the exception of distinct regions known as CpG islands (CGIs), which are often associated with promoters. In general, the vast majority of CGIs remain unmethylated during development, whereas most other CpG sites throughout the genome are methylated irrespective of genomic context, suggesting that methylation is the default state and that CGIs are exceptions to the rule [9,10]. Notably, TET-family proteins can modify CpG methylation to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) [11,12], although it is currently unclear whether these methylation derivatives have specific epigenetic roles or act only as a route towards DNA demethylation [13]. Nonetheless, recent studies have begun to unravel many aspects of DNA methylation biology including its dynamic distribution through development, the mechanisms of demethylation, and the role of CpG methylation in regulating gene transcription in distinct genomic contexts.

Here, we focus on the accumulating knowledge of the dynamics and roles of DNA methylation during the mammalian life cycle. A deeper understanding of this key epigenetic system and the mechanisms which establish and erase it may provide opportunities to manipulate cell identity for therapeutic purposes.

2. Strategies for profiling DNA methylation

A fundamental requirement to study DNA methylation is the ability to detect and track 5mC genomic levels efficiently and at a suitable resolution. However, unlike the chromatin immunoprecipitation (ChIP) method to detect histone modifications, there is no single preferred method for 5mC analysis, and a number of techniques have been developed that each have strengths and weaknesses. Broadly, genome-wide DNA methylation analysis can be divided into three approaches: (i) chemical (bisulphite, BS) modification, (ii) affinity enrichment or (iii) differential enzymatic digestion. These can be combined and/or coupled with single-locus-, microarray- or deep-sequencing- analysis to generate a vast number of different analytical techniques [14].

The BS methods rely upon the deamination of unmethylated cytosine to uracil after exposure to sodium bisulphite, while methylated cytosines are resistant and can thus be distinguished by direct sequencing [15]. This approach can generate a quantitative measure of modified cytosine across the whole genome (WGBS) at single-base resolution (including non-CpG methylation) and can be performed on relatively low input DNA quantities (more than 0.5 ng) [16,17]. In addition, BS sequencing is sensitive to relatively small changes (5–10%) in methylation and has thus been considered the 'gold standard'. However, the method has several drawbacks. First, BS conversion is not specific for precise cytosine modifications, as 5mC is indistinguishable from 5hmC, and unmodified cytosine is indistinguishable from 5caC by current methods [18,19]. A recently developed technique called oxidative bisulphite (oxBS) has partly resolved this issue by specifically identifying 5hmC, although it requires a relatively high amount of input DNA and does not discriminate other cytosine derivatives (5fC and 5caC) [20]. Second, there are potentially significant amplification biases when 5hmC is present, or between differentially methylated amplicons, which could lead to inaccurate quantification

owing to preferential exclusion of certain epigenetic states [21,22]. Third, despite recent improvement in deep-sequencing technologies, a genome-wide BS analysis of sufficient depth remains costly. This last point has been addressed by reduced representation bisulphite sequencing (RRBS), which preferentially interrogates CGIs and therefore reduces sequencing expense, but at the cost of reduced coverage, or through the use of infinium methylation arrays which infer the locus methylation state from dual CpG dinucleotides [9,23,24]. Overall, despite some drawbacks, BS approaches coupled with deep-sequencing offer the highest resolution and reproducibility currently available.

Affinity-based methods use specific antibodies or protein interactions to enrich methylated DNA fragments, with the number of subsequent sequencing reads directly related to the underlying methylation level at a locus. Affinity approaches can also be used to determine the overall global level of DNA modification through immunofluorescence, dot-blot or ELISA. Recently, genome-wide methylated DNA immunoprecipitation (meDIP-seq) has become popular in part due to the ability to distinguish 5mC- and 5hmC- profiles with specific antibodies (meDIP & hmeDIP) [25–28]. This is particularly important in studies investigating dynamic DNA demethylation or underlying epigenetic states, as 5hmC is a likely demethylation intermediate [13], and may have a specific epigenetic role [29]. Additionally, (h)meDIP-seq can interrogate both CpG and non-CpG methylation patterns and has recently been optimized for use with low input DNA quantities from rare cell populations [30,31]. Other enrichment methods such as the methylated CGI recovery assay (MIRA) or methyl-binding domain-seq (MBD-seq) can be used to selectively purify only CpG-methylated DNA (rather than CpH methylation) [32,33]. Alternatively, genomic regions that are selectively unmethylated can be assayed using CXXC-affinity purification (CAP-seq) [34]. The resolution of affinity-based methods is dependent on input fragment size, with approximately 150 bp generally the optimal trade-off between resolution and immunoprecipitation efficiency [31]. Notably fragment sizes of more than 500 bp, as used in some studies, can lead to false enrichments over promoter regions, where the first internal exon is often heavily methylated, leading to an apparent upstream promoter peak. A significant drawback of affinity approaches is that they preferentially interrogate CpG-dense regions and that enrichment represents a distribution of DNA modification rather than a quantitative measure [35]. This second issue is compounded by the difficulty in comparing genomic regions of differing CpG density, where the underlying substrate is unevenly distributed, although analytical algorithms have been developed to account for this [36]. Overall, while affinity enrichment does not enable the base-resolution of BS sequencing, the capacity to distinguish cytosine derivatives and the markedly reduced cost currently make it an important complementary technique for genome-wide DNA modification analysis.

The final broad grouping of global DNA methylation analysis methods entail differential enzymatic digestion of methylated and unmethylated regions. Several approaches have been used, such as comprehensive high-throughput analysis for relative methylation (CHARM) and *HpaII* tiny fragment enrichment by ligation-mediated PCR (HELP-seq) [37,38]. In general, enzymatic approaches suffer from the resolution limit imposed by the necessity of specific restriction sites but have found specific niches. In particular, the recent development of the glucosyltransferase assay has permitted
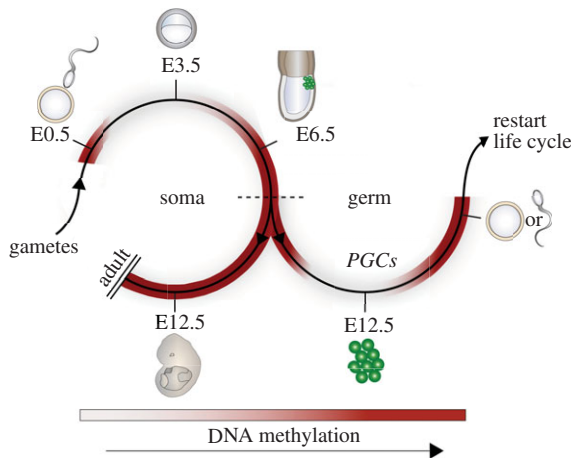
Figure 1. Global DNA methylation dynamics during the life cycle. Upon fertilization genome-wide DNA demethylation occurs in the zygote by conversion to 5hmC on the paternally derived genome and direct passive 5mC depletion of the maternally derived genome. The low-point of global methylation occurs at approximately E3.5 in the ICM, after which global re-methylation begins and reaches near completion by E6.5. At approximately E6.5, cells either continue to develop towards a somatic fate (left) or are specified as primordial germ cells (PGCs—right). Somatic fated cells acquire distinct methylomes according to their lineage but maintain high global levels of DNA methylation. PGCs initiate a phase of comprehensive DNA demethylation, which is complete by approximately E12.5, and which enables subsequent establishment of a unique gamete-specific methylome during gametogenesis. Mature gametes can then fuse to form the zygote and initiate a new life cycle. ICM, inner cell mass.

quantitative analysis of 5hmC levels at specific loci by differential digestion between *MspI* and *HpaII* [26].

3. DNA methylation during the mammalian life cycle

DNA methylation undergoes dynamic remodelling during early embryogenesis to initially establish a globally demethylated state and then subsequently, a progressively lineage-specific methylome that maintains cellular identity and genomic stability (figure 1). The process of extensive 5mC erasure begins in the zygote following fertilization and involves both conversion to 5hmC and direct passive depletion of 5mC [39]. Reprogramming of CpG methylation culminates in a globally demethylated genome in the inner cell mass (ICM) of the pre-implantation embryo (by approx. E3.5 in mice), and correlates with the establishment of pluripotent cells, which can form embryonic stem (ES) cells *in vitro* [17,40]. Indeed, ES cells completely lacking DNA methylation are viable and competent for self-renewal, indicating DNA methylation is dispensable for the naive ground state [41]. However, DNA methylation-deficient ES cells undergo apoptosis upon *in vitro* differentiation and *Dnmt1*-null embryos die around approximately E8.5, suggesting that 5mC is a critical requirement to direct and maintain terminal cell differentiation in embryonic lineages [5]. Thus, the reduced genomic 5mC levels in the pre-implantation embryo generates an epigenetic state that is conducive for subsequent embryonic lineage-specification to proceed in parallel with de novo methylation to progressively lock in cellular identity.

Following 5mC erasure, de novo remethylation commences at implantation and is directed by DNMT3A and DNMT3B. Several mechanisms have been proposed to account for how DNA methylation patterns are established during development. These can be broadly divided into (i) interactions with chromatin modifications, (ii) protection by DNA-binding factors at autonomous DNA sequences (including CpG-density), (iii) transcription read-through, and (iv) small RNA guides [42–48]. Notably, these mechanisms probably function synergistically, at least to some extent, but also in a locus and context-dependent manner. In general, DNA methylation is the default state and variation in 5mC patterns is generated through protection of genomic regions, such as CGIs, from de novo activity. Recent studies using meDIP and RRBS have demonstrated that the bulk of de novo methylation initiates in the early ICM and is largely complete by approximately E6.5 [17,30]. Consistent with this, there is an accumulation of DNA methylation at a significant number of promoters during ES cell differentiation towards neuronal progenitors [49]. However, upon subsequent terminal differentiation, neuronal progenitors only acquire limited additional 5mC, indicating that the majority of promoter CpG methylation occurs as an early developmental step, possibly to define initial lineage-restriction rather than the terminal fate. Indeed, differential DNA methylation defines lineage-restriction to either myeloid or lymphoid progenitors during haematopoiesis [50] and DNMT3A-dependent activity is essential for haematopoietic differentiation [51]. Moreover, methylation of the *Elf5* promoter is thought to be an early epigenetic barrier between trophectoderm and embryonic lineages [52]. Thus, the progressive acquisition of developmental DNA methylation at key promoters underpins Waddington's image of canalization of cell fate, particularly, the early developmental decisions.

While the majority of studies on DNA methylation patterns during development have focused primarily on promoters, it is likely that dynamic 5mC at genic, intergenic or repeat elements also contributes to lineage restriction and the observed phenotypes in DNMT-deficient systems. Indeed, only a relatively modest number of promoters, mainly of intermediate CpG-density, exhibit lineage-dependent de novo methylation during development, implying that DNA methylation is also developmentally relevant in other genomic contexts [9,30]. The recent identification of dynamically regulated low-methylated regions (LMR) at cell-type specific enhancers supports this possibility [44], as does the observation of differential methylation at intragenic CGIs [53]. Moreover, the accumulation of DNA methylation at transposable and repeat elements likely plays a key role in maintaining genomic stability during development [54,55].

An exception to the generally unidirectional process of coupled differentiation and de novo methylation during development occurs in PGCs, where a second comprehensive reprogramming event during the life cycle erases global 5mC (discussed below). This process is essential for erasure of parent-of-origin dependent genomic imprints and underpins the acquisition of genomic plasticity [56]. Following reprogramming, PGCs enter a phase of global remethylation which results in mature gametes that exhibit either a highly methylated (sperm) or a partially methylated (oocyte) genome [17,57] (reviewed in [58]). This second phase of 5mC reprogramming results in the establishment of a unique epigenome in PGCs and enables differentiation to mature gametes, which ultimately fuse to restart the mammalian life cycle.
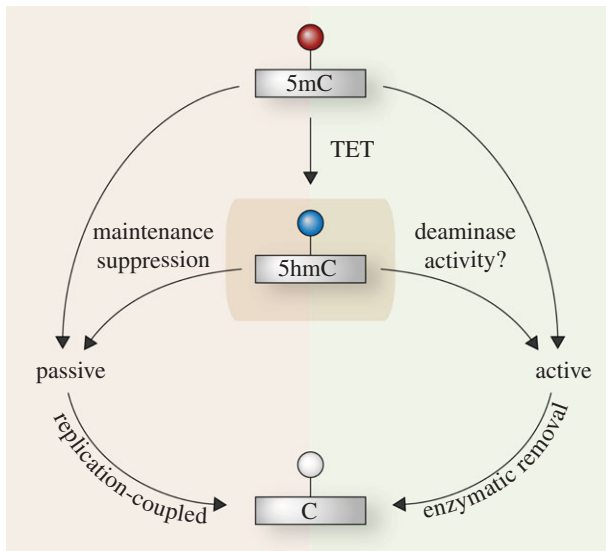
Figure 2. Mechanisms of DNA demethylation. Cytosine methylation (5mC) can be demethylated towards unmodified cytosine (C) through several potential mechanisms that can be broadly described as passive (left) or active (right). Passive demethylation occurs via replication-coupled dilution owing to a lack of re-establishment of DNA modification on the new daughter strand after DNA synthesis. Active erasure occurs through direct enzymatic removal of a DNA modification independent of DNA replication, and can involve a deamination reaction. Conversion to 5-hydroxymethylcytosine (5hmC) by TET proteins is likely a hub for demethylation (central box) and can lead to either passive or active erasure of DNA methylation. Additionally, 5hmC can be further modified to 5fC and 5caC (not shown), which can subsequently be actively or passively erased. 5mC can also be directly depleted passively, and potentially actively following deamination without conversion to 5hmC (upper arrows).

4. Reprogramming DNA methylation

DNA methylation acts as a lineage-restricting barrier during development and it is therefore essential to reprogramme the stable 5mC mark to reset the life cycle for each new generation. The mechanisms that direct this process are of great interest and have recently begun to be unravelled (figure 2). In the zygote, DNA demethylation is mechanistically compartmentalized, with the maternally and paternally derived genomes undergoing distinct processes of 5mC erasure. Ultimately, this leads to a highly demethylated epigenome, with the exception of imprinted loci, rare maternally derived promoters and some transposable elements (TEs) including ETn and intracisternal A particle (IAP) [30,59]. Original studies demonstrated that the paternal genome becomes globally demethylated prior to DNA replication, while the maternal genome apparently retains 5mC, with a subsequent progressive depletion over cell divisions [60,61]. This led to the suggestion that the paternal and maternal genomes undergo active and passive DNA demethylation, respectively. However, recent studies have demonstrated that at a global level paternal 5mC is converted to 5hmC, which is subsequently removed via passive replication-coupled dilution [39,62,63]. The paternal conversion to 5hmC is driven by the hydroxylase TET3, which can also further modify 5hmC to 5fC and 5caC [39,64]. Notably, maternal knockout of *Tet3* leads to embryonic death in a subset of embryos suggesting the conversion to 5hmC is important, if not essential, to overcome the epigenetic barriers imposed by 5mC [65]. Thus, it appears that the bulk of zygotic demethylation is induced through

passive replicative-dilution, but through different mechanisms on the paternal and maternal genomes. This difference may be related to the necessity to maintain genomic imprints through reprogramming, which are primarily located on the maternal genome [66]. Paternal imprints may be maintained via an alternative mechanism based on their association with repeat elements [67]. Interestingly, both the paternal and maternal genomes exhibit a dramatic decrease in 5mC, including at imprints, in mouse zygotes deficient for STELLA. This implies that the maternal genome is protected from 5hmC conversion, rather than the paternal genome being specifically targeted in normal zygotes [68]. However, it is unclear precisely how the direct passive loss of 5mC on the maternal genome occurs, as at least some DNMT1 is present and sufficient to maintain imprints [69].

The most comprehensive DNA methylation reprogramming event in the mammalian life cycle occurs in PGCs between approximately E8.5–E12.5 [70,71]; (reviewed in [56]). This process leads to almost complete genome-wide DNA demethylation by E13.5 [72]. Importantly, 5mC erasure in PGCs includes genomic imprints, which enables their subsequent re-establishment according to the sex of the cell, and ultimately the acquisition of totipotency upon fertilization [71]. Additionally, most repeat elements are demethylated to some extent in PGCs, although IAP elements are partially resistant, which may enable transgenerational epigenetic inheritance [72–74]. The mechanistic basis of DNA demethylation in PGCs has yet to be fully determined, but could potentially occur through a number of overlapping routes including: conversion to 5hmC coupled with passive depletion, direct replication-dependent loss of 5mC, or active erasure associated with DNA repair pathways (figure 2) [56]. Given the high expression of *Tet1* and *Tet2* in PGCs, a demethylation mechanism involving 5hmC seems probable [75]. Moreover, both TET1 and TET2 may contribute to this process redundantly, as supported by only a modest reduction of 5hmC levels in *Tet1*-null ES cells (which also express *Tet2*) and the fertility, albeit subfertility, of *Tet1*-null mice [76]. Other mechanisms may also operate in conjunction with 5hmC conversion to drive DNA demethylation in PGCs, possibly including base excision repair (BER) systems because loss of the deaminase AID leads to a partial reduction of 5mC erasure [72,75,77]. Further studies are required to elucidate the mechanisms fully.

While the two primary phases of global 5mC erasure during the life cycle occur in pre-implantation embryos and in PGCs, significant DNA demethylation has also been reported in other contexts. A significant reduction at all genomic loci examined, except imprinted loci, has been shown using an RRBS approach during mouse erythropoiesis [78]. It is unclear how this global demethylation occurs mechanistically, but it may be linked to the requirement to activate lineage-specific genes. Indeed, despite the common view that once established, DNA methylation is stably maintained throughout development, several examples of locus-specific demethylation have emerged outwith of global reprogramming phases, particularly in neuronal and haematopoietic lineages. During terminal differentiation, the CpG-dense promoters of the eye-specific genes *Cplx4* and *Cryaa* are demethylated, possibly to enable their subsequent expression. Similarly, promoter methylation at *Tlr6* and *Cytip*, which is established by E6.5, is subsequently erased during further embryonic differentiation towards haematopoietic lineages [30]. At low CpG-density promoters, the presence of activating signals alone may be sufficient to initiate dynamic
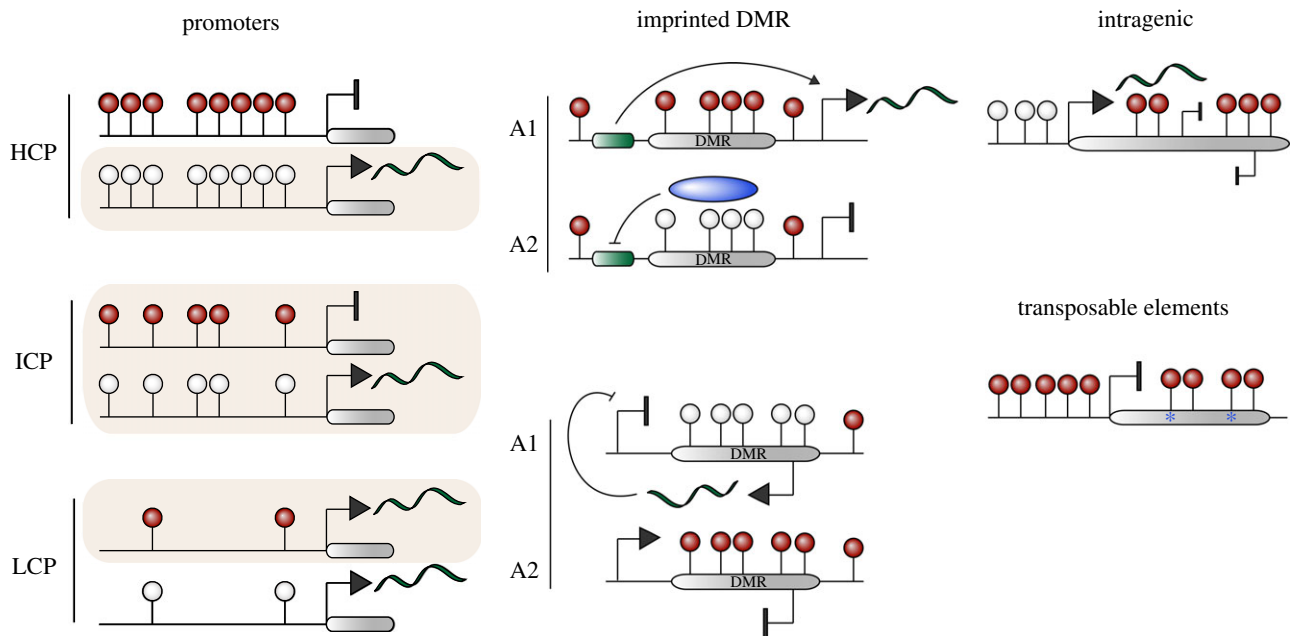
Figure 3. Distribution and roles of DNA methylation. The distribution of DNA methylation varies according to genomic landmarks. High-CpG density promoters (HCP) are usually hypomethylated, low CpG-density promoters (LCP) are usually methylated, and intermediate CpG-density promoters (ICP) can be either methylated or unmethylated (shown as shaded). In general, methylation only has a significant transcriptional effect at HCPs and ICPs, whereas methylation at LCPs does not correlate with repression. Note also that the absence of methylation only generates a permissive state for transcription and does not necessarily result in gene activity. Imprinted loci are methylated on one allele and hypomethylated on the other. This can have allele-specific effects by either modulating interactions between enhancers (green) and promoters (upper imprinted gene) or regulating expression of an antisense ncRNA, which silences genes in *cis* (lower imprinted gene). Gene bodies are generally hypermethylated, which may function to repress cryptic internal promoters. Transposable elements are usually highly methylated in the promoter and coding regions, which silences their expression and can lead to their mutation and inactivation through cytosine deamination (asterisks), respectively. A1/A2: Allele 1 and Allele 2. DMR, differentially methylated region.

DNA demethylation [79,80]. Moreover, physiological factors such as increased muscle activity may drive dynamic 5mC erasure at key metabolic promoters in skeletal muscle [81]. Recent studies have also suggested that cyclical demethylation may contribute to the generally unmethylated state of CGIs during all stages of development. In this model TET1, which preferentially binds CpG-rich sequences, is targeted to CGIs where any aberrant 5mC could be hydroxylated and removed via subsequent passive or active mechanisms [13,26,82]. This continuous turnover could thus be considered as reiterative demethylation to maintain CGIs in a permissive state for transcription. Taken together, 5mC reprogramming appears to occur both locus-specifically and globally at least to some extent, during normal cell differentiation, although this may be specifically linked to lineage gene-activation.

5. The genomic role of DNA methylation

At the cellular level DNA methylation is essential to propagate lineage restriction and to maintain genomic stability. However, at the molecular level the role of 5mC is dictated by multiple parameters, including its genomic context—such as promoter or exon, the local CpG density and local chromatin modifications [10,83]. While DNA methylation is generally considered to impede and repress transcription, genome-wide studies have demonstrated that the actual transcriptional effect of DNA methylation at promoter regions is related to the local CpG-density (figure 3) [9,84]. For example, DNA methylation at low CpG-density promoters (LCP) is not correlated with transcriptional silencing, and indeed most LCPs are methylated irrespective of their expression state. In contrast,

significant 5mC at high CpG-density promoters (HCP) is strongly associated with gene silencing, although HCPs are very rarely methylated during normal development, excepting imprinted genes and the inactive X-chromosome. Thus, concerning transcriptional regulation, DNA methylation primarily functions at intermediate CpG-density promoters (ICP), which have a greater capacity to acquire lineage-dependent methylation and which are also correlated with strong gene silencing when methylated [9,30]. Of particular prominence is a role for DNA methylation in stably silencing ICPs associated with germline-specific genes [9,85,86], which may serve to stringently prevent their expression in somatic cells owing to their potential for contributing to carcinogenesis [87,88]. Further studies have shown that promoter methylation at germline-specific genes is indeed uniquely associated with their transcriptional regulation [89]. At other promoters, principally lineage- or pluripotency-specific genes, CpG methylation may function as a secondary 'silencing lock' that is acquired after other chromatin modifications have established a repressive state [90]. However, CpG methylation is generally not a regulatory mechanism that causally directs repression at most promoters during normal development. Indeed, it has even been reported to mediate gene activation in some rare contexts [91].

In contrast to its secondary role at most promoters, CpG methylation has a key role in modulating genomic imprinting and in transcriptional silencing of CpG-dense promoters on the inactive X-chromosome in females. Imprinted genes are expressed monoallelically according to parental origin, a process that in most cases is directly linked to differential methylation of a CpG-dense region (DMR) [92,93]. Some DMRs regulate the monoallelic expression of multiple linked genes, as demonstrated by genetic interventions, in

which case the DMR is known as an imprinted control region (ICR). In general, DNA methylation at DMRs functions through two mechanisms: (i) impeding binding of a chromatin insulator protein, thereby enabling differential interactions between enhancers and promoters according to methylation state or (ii) acting as a promoter for an antisense ncRNA, which represses *cis*-linked genes when expressed from the unmethylated allele (figure 3) [94]. The deregulation of imprinted genes may significantly contribute to the phenotype of DNA methylation-deficient systems, as imprinted genes play an essential role in development, growth, postnatal behaviour and cell metabolism.

The role of intragenic DNA methylation in mammals has recently emerged as a key point of interest, since both intronic and particularly exonic sequences tend to be highly methylated. One possible role of intragenic 5mC is to prevent spurious transcriptional activation from cryptic internal promoters [10]. Alternatively, genic methylation may simply be derived from the relatively open chromatin state of active genes owing to transcription read-through, which could enable DNMTs or chromatin-modifying enzymes enhanced access. This is supported by observations that the active X-chromosome (X_a) in females is more highly methylated in gene-bodies than the inactive X-chromosome (X_i) [95]. Another point of interest is DNA methylation at regions adjacent to promoter CGIs, termed CpG-shores. Here, 5mC has been correlated with gene expression and correspondingly, CpG-shores are reported to be methylated tissue-specifically [50,96]. However, using an affinity-based approach, it has also been suggested that CpG-shore methylation shows no association with gene expression and indeed, that they do not exhibit significant levels of differential DNA methylation [53]. Thus, the role of CpG-shore methylation in genome regulation will require further investigation.

Finally, DNA methylation is essential for stable silencing of some TEs. In mammals retrotransposons and repeat-derived elements are usually highly methylated, and loss of 5mC can lead to significant upregulation of at least IAP and LINE elements, which can potentially lead to insertional mutagenesis [54,55]. The potential significance of 5mC in maintaining genomic integrity by silencing TEs has led to the 'genome defence' hypothesis that TE repression is *the* central role of DNA methylation in mammals [97], although this has been disputed [10]. Notably, the function of DNA methylation in TE silencing has been proposed to be directly coupled to its unique role in regulating a subset of germline-specific genes. Here, it has been noted that genes involved in post-transcriptional genome-defence against TEs in the germline, such as *Tex19.1* [98], are preferentially regulated by promoter CpG methylation [89]. Because these genes are activated by erasure of promoter 5mC during epigenetic reprogramming in PGCs, they become expressed at precisely the time when TEs might become transcriptionally active, owing to loss of global 5mC. Thus, by linking DNA methylation to both TE silencing and regulation of genome defence genes, a failsafe system is created whereby the potential for TE expression during 5mC loss in PGCs is coupled to upregulation of mechanisms that suppress TE activity.

6. Concluding remarks

As an epigenetic mechanism, DNA methylation is both elegant and enigmatic. While we understand both the mechanism that faithfully propagates CpG methylation through cell divisions and the genomic distribution of this modification, the dynamics and the mechanisms of reprogramming this mark remain to be fully elucidated. Likewise, while DNA methylation appears to carry out distinct roles in different genomic and developmental contexts, we have not precisely defined how CpG methylation contributes to genome regulation and development. Emerging technologies that can comprehensively map cell-type specific DNA methylation patterns could potentially enable a fuller understanding of 5mC in normal development, cancer, ageing and inter-individual variations. The ability to profile methylation in rare or inaccessible cell populations such as in PGCs is integral to this process. Downstream analysis of key mediators of DNA methylation patterns such as TET proteins, and integration of chromatin and transcription datasets may shed further light on the roles and dynamics of this key epigenetic modification.

References

1. Reik W. 2007 Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**, 425–432. (doi:10.1038/nature05918)

2. Surani MA, Hayashi K, Hajkova P. 2007 Genetic and epigenetic regulators of pluripotency. *Cell* **128**, 747–762. (doi:10.1016/j.cell.2007.02.010)

3. Bird A. 2002 DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21. (doi:10.1101/gad.947102)

4. Okano M, Bell DW, Haber DA, Li E. 1999 DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257. (doi:10.1016/S0092-8674(00)81656-6)

5. Li E, Bestor TH, Jaenisch R. 1992 Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926. (doi:10.1016/0092-8674(92)90611-F)

6. Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. 2000 Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl Acad. Sci. USA* **97**, 5237–5242. (doi:10.1073/pnas.97.10.5237)

7. Tomizawa S, Kobayashi H, Watanabe T, Andrews S, Hata K, Kelsey G, Sasaki H. 2011 Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development* **138**, 811–820. (doi:10.1242/dev.061416)

8. Ziller MJ *et al.* 2011 Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet.* **7**, e1002389. (doi:10.1371/journal.pgen.1002389)

9. Meissner A *et al.* 2008 Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770. (doi:10.1038/nature07107)

10. Suzuki MM, Bird A. 2008 DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* **9**, 465–476. (doi:10.1038/nrg2341)

11. Tahiliani M *et al.* 2009 Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935. (doi:10.1126/science.1170116)

12. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. 2011 Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303. (doi:10.1126/science.1210597)

13. Branco MR, Ficz G, Reik W. 2012 Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat. Rev. Genet.* **13**, 7–13.

14. Laird PW. 2010 Principles and challenges of genomewide DNA methylation analysis. *Nat. Rev. Genet.* **11**, 191–203. (doi:10.1038/nrg2732)

15. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL. 1992 A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl Acad. Sci. USA* **89**, 1827–1831. (doi:10.1073/pnas.89.5.1827)

16. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. 2008 Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* **133**, 523–536. (doi:10.1016/j.cell.2008.03.029)

17. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A. 2012 A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344. (doi:10.1038/nature10960)

18. Nestor C, Ruzov A, Meehan R, Dunican D. 2010 Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA. *BioTechniques* **48**, 317–319. (doi:10.2144/000113403)

19. He YF et al. 2011 Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307. (doi:10.1126/science.1210944)

20. Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F, Reik W, Balasubramanian S. 2012 Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science* **336**, 934–937. (doi:10.1126/science.1220671)

21. Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A, Liu J. 2010 The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS ONE* **5**, e8888. (doi:10.1371/journal.pone.0008888)

22. Warnecke PM, Stirzaker C, Melki JR, Millar DS, Paul CL, Clark SJ. 1997 Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res.* **25**, 4422–4426. (doi:10.1093/nar/25.21.4422)

23. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. 2005 Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868–5877. (doi:10.1093/nar/gki901)

24. Bibikova M et al. 2006 High-throughput DNA methylation profiling using universal bead arrays. *Genome Res.* **16**, 383–393. (doi:10.1101/gr.4410706)

25. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PAC, Rappsilber J, Helin K. 2011 TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343–348. (doi:10.1038/nature10066)

26. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, Marques CJ, Andrews S, Reik W. 2011 Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473**, 398–402. (doi:10.1038/nature10008)

27. Pastor WA et al. 2011 Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* **473**, 394–397. (doi:10.1038/nature10102)

28. Nestor CE et al. 2012 Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res.* **22**, 467–477. (doi:10.1101/gr.126417.111)

29. Yildirim O, Li R, Hung J-H, Chen P, Dong X, Ee L-S, Weng Z, Rando OJ, Fazzio TAG. 2011 Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* **147**, 1498–1510. (doi:10.1016/j.cell.2011.11.054)

30. Borgel J, Guibert S, Li Y, Chiba H, Schübeler D, Sasaki H, Forné T, Weber M. 2010 Targets and dynamics of promoter DNA methylation during early mouse development. *Nat. Genet.* **42**, 1093–1100. (doi:10.1038/ng.708)

31. Taiwo O, Wilson GA, Morris T, Seisenberger S, Reik W, Pearce D, Beck S, Butcher LM. 2012 Methylome analysis using MeDIP-seq with low DNA concentrations. *Nat. Protoc.* **7**, 617–636. (doi:10.1038/nprot.2012.012)

32. Rauch T, Li H, Wu X, Pfeifer GP. 2006 MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells. *Cancer Res.* **66**, 7939–7947. (doi:10.1158/0008-5472.CAN-06-1888)

33. Serre D, Lee BH, Ting AH. 2010 MBD-isolated genome sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Res.* **38**, 391–399. (doi:10.1093/nar/gkp992)

34. Illingworth R et al. 2008 A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol.* **6**, e22. (doi:10.1371/journal.pbio.0060022)

35. Harris RA et al. 2010 Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat. Biotech.* **28**, 1097–1105. (doi:10.1038/nbt.1682)

36. Down TA et al. 2008 A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat. Biotech.* **26**, 779–785. (doi:10.1038/nbt1414)

37. Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddeloh JA, Wen B, Feinberg AP. 2008 Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res.* **18**, 780–790. (doi:10.1101/gr.7301508)

38. Khulan B et al. 2006 Comparative isoschizomer profiling of cytosine methylation: the HELP assay. *Genome Res.* **16**, 1046–1055. (doi:10.1101/gr.5273806)

39. Wossidlo M et al. 2011 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun.* **2**, 241. (doi:10.1038/ncomms1240)

40. Evans MJ, Kaufman MH. 1981 Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156. (doi:10.1038/292154a0)

41. Tsumura A et al. 2006 Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11**, 805–814. (doi:10.1111/j.1365-2443.2006.00984.x)

42. Ooi SK et al. 2007 DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714–717. (doi:10.1038/nature05987)

43. Ciccone DN, Su H, Hevi S, Gay F, Lei H, Bajko J, Xu G, Li E, Chen T. 2009 KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* **461**, 415–418. (doi:10.1038/nature08315)

44. Stadler MB et al. 2011 DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495.

45. Thomson JP et al. 2010 CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* **464**, 1082–1086. (doi:10.1038/nature08924)

46. Chotalia M, Smallwood SA, Ruf N, Dawson C, Lucifero D, Frontera M, James K, Dean W, Kelsey G. 2009 Transcription is required for establishment of germline methylation marks at imprinted genes. *Genes Dev.* **23**, 105–117. (doi:10.1101/gad.495809)

47. Lienert F, Wirbelauer C, Som I, Dean A, Mohn F, Schabeler D. 2011 Identification of genetic elements that autonomously determine DNA methylation states. *Nat. Genet.* **43**, 1091–1097. (doi:10.1038/ng.946)

48. Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ. 2008 A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell* **31**, 785–799. (doi:10.1016/j.molcel.2008.09.003)

49. Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schabeler D. 2008 Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* **30**, 755–766. (doi:10.1016/j.molcel.2008.05.007)

50. Ji H et al. 2010 Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* **467**, 338–342. (doi:10.1038/nature09367)

51. Challen GA et al. 2012 Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat. Genet.* **44**, 23–31. (doi:10.1038/ng.1009)

52. Ng RK, Dean W, Dawson C, Lucifero D, Madeja Z, Reik W, Hemberger M. 2008 Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. *Nat. Cell Biol.* **10**, 1280–1290. (doi:10.1038/ncb1786)

53. Deaton AM, Webb S, Kerr ARW, Illingworth RS, Guy J, Andrews R, Bird A. 2011 Cell type-specific DNA methylation at intragenic CpG islands in the immune system. *Genome Res.* **21**, 1074–1086. (doi:10.1101/gr.118703.110)

54. Walsh CP, Chaillet JR, Bestor TH. 1998 Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20**, 116–117. (doi:10.1038/2413)

55. Bourc'his D, Bestor TH. 2004 Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431**, 96–99. (doi:10.1038/nature02886)

56. Hackett JA, Zylicz JJ, Surani MA. 2012 Parallel mechanisms of epigenetic reprogramming in the germline. *Trends Genet.* **28**, 164–174. (doi:10.1016/j.tig.2012.01.005)

57. Smallwood SA *et al*. 2011 Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat. Genet.* **43**, 811–814. (doi:10.1038/ng.864)

58. Smallwood SA, Kelsey G. 2012 De novo DNA methylation: a germ cell perspective. *Trends Genet.* **28**, 33–42. (doi:10.1016/j.tig.2011.09.004)

59. Wossidlo M, Arand J, Sebastiano V, Lepikhov K, Boiani M, Reinhardt R, Schaler H, Walter J. 2010 Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. *EMBO J.* **29**, 1877–1888. (doi:10.1038/emboj.2010.80)

60. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. 2000 Demethylation of the zygotic paternal genome. *Nature* **403**, 501–502. (doi:10.1038/35000656)

61. Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, Dean W, Reik W, Walter J. 2000 Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10**, 475–478. (doi:10.1016/S0960-9822(00)00448-6)

62. Iqbal K, Jin S-G, Pfeifer GP, Szabo PE. 2011 Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl Acad. Sci. USA* **108**, 3642–3647. (doi:10.1073/pnas.1014033108)

63. Inoue AY, Zhang Y. 2011 Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* **334**, 194. (doi:10.1126/science.1212483)

64. Inoue A, Shen L, Dai Q, He C, Zhang Y. 2011 Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. *Cell Res.* **21**, 1670–1676. (doi:10.1038/cr.2011.189)

65. Gu TP *et al*. 2011 The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606–610. (doi:10.1038/nature10443)

66. Reik W, Walter J. 2001 Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet.* **2**, 21–32. (doi:10.1038/35047554)

67. Watanabe T *et al*. 2011 Role for piRNAs and noncoding RNA in de Novo DNA methylation of the imprinted mouse Rasgrf1 locus. *Science* **332**, 848–852. (doi:10.1126/science.1203919)

68. Nakamura T *et al*. 2007 PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat. Cell Biol.* **9**, 64–71. (doi:10.1038/ncb1519)

69. Hirasawa R, Chiba H, Kaneda M, Tajima S, Li E, Jaenisch R, Sasaki H. 2008 Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev.* **22**, 1607–1616. (doi:10.1101/gad.1667008)

70. Seki Y, Hayashi K, Itoh K, Mizugaki M, Saitou M, Matsui Y. 2005 Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev. Biol.* **278**, 440–458. (doi:10.1016/j.ydbio.2004.11.025)

71. Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J, Surani MA. 2002 Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117**, 15–23. (doi:10.1016/S0925-4773(02)00181-8)

72. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE, Reik W. 2010 Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**, 1101–1105. (doi:10.1038/nature08829)

73. Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J, Reik W. 2003 Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* **35**, 88–93. (doi:10.1002/gene.10168)

74. Morgan HD, Whitelaw E, Sutherland HGE, Martin DIK. 1999 Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* **23**, 314–318. (doi:10.1038/15490)

75. Hajkova P, Jeffries J, Lee C, Miller N, Jackson SP, Surani MA. 2010 Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* **329**, 78–82. (doi:10.1126/science.1187945)

76. Dawlaty MM *et al*. 2011 Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* **9**, 166–175. (doi:10.1016/j.stem.2011.07.010)

77. Cortellino S *et al*. 2011 Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146**, 67–79. (doi:10.1016/j.cell.2011.06.020)

78. Shearstone JR, Pop R, Bock C, Boyle P, Meissner A, Socolovsky M. 2011 Global DNA demethylation during mouse erythropoiesis in vivo. *Science* **334**, 799–802. (doi:10.1126/science.1207306)

79. Métivier R *et al*. 2008 Cyclical DNA methylation of a transcriptionally active promoter. *Nature* **452**, 45–50. (doi:10.1038/nature06544)

80. Kangaspeska S, Stride B, Metivier R, Polycarpou-Schwarz M, Ibberson D, Carmouche RP, Benes V, Gannon F, Reid G. 2008 Transient cyclical methylation of promoter DNA. *Nature* **452**, 112–115. (doi:10.1038/nature06640)

81. Barres R *et al*. 2012 Acute exercise remodels promoter methylation in human skeletal muscle. *Cell Metab.* **15**, 405–411. (doi:10.1016/j.cmet.2012.01.001)

82. Williams K, Christensen J, Helin K. 2012 DNA methylation: TET proteins-guardians of CpG islands? *EMBO Rep.* **13**, 28–35. (doi:10.1038/embor.2011.233)

83. Cedar H, Bergman Y. 2009 Linking DNA methylation and histone modification: patterns and paradigms. *Nat. Rev. Genet.* **10**, 295–304. (doi:10.1038/nrg2540)

84. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. 2007 Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **39**, 457–466. (doi:10.1038/ng1990)

85. Velasco G, Hube F, Rollin J, Neuillet D, Philippe C, Bouzinba-Segard H, Galvani A, Viegas-Pequignot E, Francastel C. 2010 Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues. *Proc. Natl Acad. Sci. USA* **107**, 9281–9286. (doi:10.1073/pnas.1000473107)

86. Maatouk DM, Kellam LD, Mann MR, Lei H, Li E, Bartolomei MS, Resnick JL. 2006 DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development* **133**, 3411–3418. (doi:10.1242/dev.02500)

87. Janic A, Mendizabal L, Llamazares S, Rossell D, Gonzalez C. 2010 Ectopic expression of germline genes drives malignant brain tumor growth in Drosophila. *Science* **330**, 1824–1827. (doi:10.1126/science.1195481)

88. Simpson AJ, Caballero OL, Jungbluth A, Chen Y-T, Old LJ. 2005 Cancer/testis antigens, gametogenesis and cancer. *Nat. Rev. Cancer* **5**, 615–625. (doi:10.1038/nrc1669)

89. Hackett JA *et al*. 2012 Promoter DNA methylation couples genome-defence mechanisms to epigenetic reprogramming in the mouse germline. *Development* **139**, 3623–3632. (doi:10.1242/dev.081661)

90. Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y, Cedar H, Bergman Y. 2006 G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat. Cell Biol.* **8**, 188–194. (doi:10.1038/ncb1353)

91. Rishi V, Bhattacharya P, Chatterjee R, Rozenberg J, Zhao J, Glass K, Fitzgerald P, Vinson C. 2010 CpG methylation of half-CRE sequences creates C/EBPalpha binding sites that activate some tissue-specific genes. *Proc. Natl Acad. Sci. USA* **107**, 20 311–20 316. (doi:10.1073/pnas.1008688107)

92. Bartolomei MS. 2009 Genomic imprinting: employing and avoiding epigenetic processes. *Genes Dev.* **23**, 2124–2133. (doi:10.1101/gad.1841409)

93. Ferguson-Smith AC. 2011 Genomic imprinting: the emergence of an epigenetic paradigm. *Nat. Rev. Genet.* **12**, 565–575. (doi:10.1038/nrg3032)

94. Abramowitz LK, Bartolomei MS. 2012 Genomic imprinting: recognition and marking of imprinted loci. *Curr. Opin. Genet. Dev.* **22**, 72–78. (doi:10.1016/j.gde.2011.12.001)

95. Hellman A, Chess A. 2007 Gene body-specific methylation on the active X chromosome. *Science* **315**, 1141–1143. (doi:10.1126/science.1136352)

96. Irizarry RA *et al*. 2009 The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat. Genet.* **41**, 178–186. (doi:10.1038/ng.298)

97. Yoder JA, Walsh CP, Bestor TH. 1997 Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**, 335–340. (doi:10.1016/S0168-9525(97)01181-5)

98. Ollinger R *et al*. 2008 Deletion of the pluripotency-associated Tex19.1 gene causes activation of endogenous retroviruses and defective spermatogenesis in mice. *PLoS Genet.* **4**, e1000199. (doi:10.1371/journal.pgen.1000199)