Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells

Mitinori Saitou[1,2,3,*], Saya Kagiwada[1] and Kazuki Kurimoto[1,2]

Summary

Epigenetic modifications are crucial for the identity and stability of cells, and, when aberrant, can lead to disease. During mouse development, the genome-wide epigenetic states of pre-implantation embryos and primordial germ cells (PGCs) undergo extensive reprogramming. An improved understanding of the epigenetic reprogramming mechanisms that occur in these cells should provide important new information about the regulation of the epigenetic state of a cell and the mechanisms of induced pluripotency. Here, we discuss recent findings about the potential mechanisms of epigenetic reprogramming, particularly genome-wide DNA demethylation, in pre-implantation mouse embryos and PGCs.

Key words: DNA demethylation, Epigenetic reprogramming, Pre-implantation embryos, Primordial germ cells, TET proteins

Introduction

The variety of cellular states in multicellular organisms reflects the diversity of the transcriptional programme of cells, despite the fact that nearly all cells in any given organism bear an identical genome sequence. The transcriptional state of a cell is governed by a specific set of transcriptional regulators and also by chemical modifications of the genome, including cytosine methylation (5-methylcytosine; 5mC) and post-translational modifications of histone tails, which regulate the accessibility of transcriptional regulators to, and the on or off states or expression levels of, all genes in the genome (Bonasio et al., 2010). The stability of the phenotype of a cell upon mitosis/meiosis is considered to be underpinned by the stability of these modifications. Modifications that regulate the identity of a cell without changing the DNA sequence are referred to as epigenetic modifications (Bird, 2007; Bonasio et al., 2010), and the genome-wide state of the epigenetic status of a cell is known as the epigenome (Bernstein et al., 2007). Once specified during development or in adult physiology, the epigenome of a cell is stable upon mitosis/meiosis, and cell identities are maintained essentially for a lifetime (Bonasio et al., 2010).

During mammalian development, however, there are two crucial developmental stages and/or cell types in which the epigenome undergoes profound reprogramming: pre-implantation embryos and primordial germ cells (PGCs), the precursors both for oocytes and spermatozoa (Fig. 1) (Surani et al., 2007). Epigenetic

[1]Department of Anatomy and Cell Biology, Graduate School of Medicine, Kyoto University, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan. [2]JST, CREST/ERATO, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan. [3]Institute for Integrated Cell-Material Sciences, Kyoto University, Yoshida-Ushinomiya-cho, Sakyo-ku, Kyoto 606-8501, Japan.

*Author for correspondence (saitou@anat2.med.kyoto-u.ac.jp)

reprogramming in these cells involves genome-wide demethylation of 5mC; 5mC plays a crucial role in genome imprinting, X inactivation, transposon silencing, the stability of centromeric/telomeric structures and gene expression in general. Over the past decade, the technologies that can be used to determine genome-wide 5mC distribution have dramatically evolved, leading to the identification of global 5mC distribution in a number of cultured cell lines and primary tissues (Lister et al., 2009; Suzuki and Bird, 2008).

In this article, we summarize and discuss recent discoveries about the nature and mechanism of epigenetic reprogramming, particularly those relating to genome-wide DNA demethylation in mouse pre-implantation embryos and PGCs. These findings are important not only for understanding the genetic and epigenetic basis of genome inheritance, but also for elucidating the mechanisms of artificially induced epigenetic reprogramming that may be of medical relevance (Hanna et al., 2010; Hayashi and Surani, 2009; Yamanaka and Blau, 2010). [For reviews on other associated epigenetic events such as histone modification changes and X-chromosome inactivation/reactivation, please see other recent reviews (Brockdorf, 2011; Hemberger et al., 2009; Payer and Lee, 2008; Probst and Almouzni, 2011).]

5-methylcytosine: an overview
Cytosine methylation and 5mC distribution

Genome-wide cytosine methylation states, especially those associated with genes, differ among cell types and function as a form of memory of the identity and developmental state of a cell (Lister et al., 2009). The enzymes that methylate cytosine to form 5-methylcytosine (5mC) have been well characterized (see Table 1). DNA methyltransferase (DNMT) 1 preferentially methylates hemi-methylated cytosines in CpG sequences and thus acts as a maintenance methyltransferase to maintain genome-wide methylation patterns during replication (Bestor et al., 1988; Bestor, 1992; Li et al., 1992). DNMT3A and DNMT3B can methylate unmethylated CpG sequences and hence function as de novo methyltransferases (Okano et al., 1998a). DNMT3L has no catalytic activity but recruits DNMT3A and DNMT3B to their targets by recognizing nucleosomes that carry unmethylated histone H3 lysine 4 (H3K4) (Aapola et al., 2000; Bourc'his and Bestor, 2004; Bourc'his et al., 2001; Hata et al., 2002; Ooi et al., 2007).

5mC occurs mostly in CpG sequences and, to a lesser extent, in CpHpG or CpHpH sequences (where H is A, C or T), especially in pluripotent cells (Lister et al., 2009; Ramsahoye et al., 2000; Tomizawa et al., 2011). Mammalian genomes are globally methylated: genes, transposons, repeat sequences and intergenic DNA are all subjected to methylation (Suzuki and Bird, 2008). 5mC can spontaneously deaminate to form thymine (T), creating T:G mismatches, and thus can be a source of point mutations across the genome (see Box 1 and references therein). Unmethylated
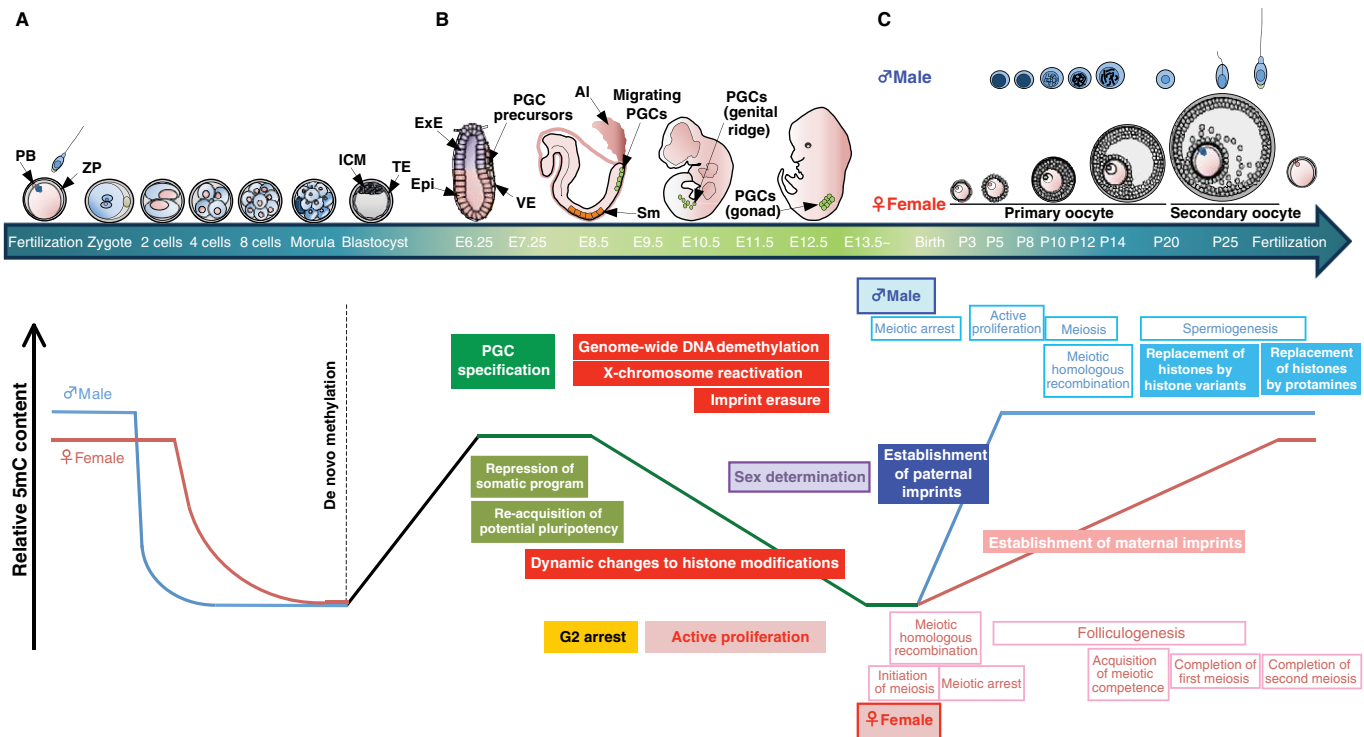
Fig. 1. A schematic of mouse pre-implantation and germ cell development. (Top) A schematic of pre-implantation and germ-cell development in mice. (**A**) Pre-implantation development stages; (**B**) post-implantation embryonic development, following blastocyst implantation at around E4.5; and (**C**) postnatal germ cell development and maturation. Primordial germ cell (PGC) precursors (E6.25) and PGCs are shown as green circles in embryos from E6.25 to E12.5. (Bottom) Key genetic and epigenetic events are shown that are associated with pre-implantation and germ cell development, together with relative levels of 5-methylcytosine (5mC) at different developmental stages. Al, allantois; Epi, epiblast; ExE, extra-embryonic ectoderm; ICM, inner cell mass; PB, polar body; PGCs, primordial germ cells; Sm, somite; TE, trophectoderm; VE, visceral endoderm; ZP, zona pelucida.

sequences are most often found in CpG islands (CGIs) (see Glossary, Box 2), which are typically associated with gene promoters (Suzuki and Bird, 2008) and with ~70% of genes. Based on their CpG ratio, GC content and on the length of the CpG-rich region, promoters are classified as being high-, intermediate- or low-CpG promoters (HCPs, ICPs and LCPs, respectively) (Weber et al., 2007).

5mC and histone modifications

5mC and histone modifications act in concert to form an appropriate epigenome during development and in adult cells (Cedar and Bergman, 2009). Generally, 5mCs are associated with transcriptionally repressive histone modifications, such as histone H3 lysine 9 di- (H3K9me2) or tri-methylation (H3K9me3). This is partly because 5mCs are recognized by methyl-CpG binding proteins, which recruit the histone deacetylase complex (see Glossary, Box 2) (Jones et al., 1998; Nan et al., 1998). The interaction of DNMT1 and G9a, a H3K9 methyltransferase, with the replication complex (see Glossary, Box 2) might also connect 5mC to H3K9me2 (Esteve et al., 2006; Hashimshony et al., 2003). Conversely, methylated H3K9 is bound by heterochromatin protein 1 (HP1), which recruits DNMT1 to confer DNA methylation (Fuks et al., 2003; Smallwood et al., 2007). The interaction of the H3K9 methyltransferases SUV39H1 (suppressor of variegation 3-9 homolog 1) and ESET (also known as SETDB1; SET domain, bifurcated 1) with DNMT3A and DNMT3B can also direct DNA methylation at H3K9me3 (Fuks et al., 2003; Lehnertz et al., 2003;

Li et al., 2006). NP95 (also known as UHRF1, ubiquitin-like, containing PHD and RING finger domains 1) is a multi-domain protein that is essential for recruiting DNMT1 to replication foci by interacting with DNMT1 and binding to hemi-methylated DNA (Bostick et al., 2007; Fujimori et al., 1998; Sharif et al., 2007). NP95 also interacts with DNMT3A and DNMT3B (Meilinger et al., 2009), G9a (Kim et al., 2009) and H3K9me3 (Rottach et al., 2010), integrating the DNA methylation and H3K9 methylation pathways.

Unmethylated CpG sequences, most typically CGIs, by contrast, are generally associated with transcriptionally permissive/active acetylated H3K4, H3K4me2 and H3K4me3 (Guenther et al., 2007). CXXC finger protein 1 (CFP1) binds unmethylated CpG sequences via its CXXC zinc-finger domain and recruits the H3K4 methyltransferase SETD1, thereby inducing a H3K4me2/3-positive transcriptionally permissive/active chromatin state (Thomson et al., 2010). Polycomb repressive complex (PRC) 2 (see Glossary, Box 2), which catalyses H3K27me3 formation and induces a repressive chromatin state, also binds unmethylated CGIs and represses their promoter activity during development and in embryonic stem cells (ESCs) (Barski et al., 2007; Bernstein et al., 2006; Mikkelsen et al., 2007; Mohn et al., 2008; Pan et al., 2007; Zhao et al., 2007). In ESCs, PRC target genes are also often marked by H3K4me3, which creates a bivalent modification state that causes the chromatin of a gene to be configured in an active or inactive state depending on the subsequent signal the cell receives. In some cases, PRC-targeted sequences become DNA methylated (Mohn et

Table 1. Proteins involved in cytosine modification

Protein	Functions	Tissue distribution	Phenotype of knockout/knockdown mice/ESCs	Refs
DNMT1	Methylation of hemimethylated CpGs during DNA replication	Highly expressed in proliferating cells (ubiquitous expression)	Embryonic lethality shortly after gastrulation; extensive demethylation of the genome.	(Bestor et al., 1988; Lei et al., 1996; Li et al., 1992)
DNMT2	Methylation of small RNA	High expression in heart, kidney and testis (ubiquitous expression)	*Dnmt2* knockout mice are viable, fertile and show normal genomic methylation pattern; abolishment of RNA methyltransferase activity.	(Goll et al., 2006; Okano et al., 1998b)
DNMT3A	DNA methylation establishment in early development and germ cells (essential for imprint establishment during gametogenesis, with DNMT3L)	Highly expressed in embryonic tissue and undifferentiated ESCs (ubiquitous expression)	*Dnmt3a* knockout embryos develop to term, but die shortly after birth; loss of imprinting in both male and female germ cells.	(Kaneda et al., 2004; Okano et al., 1999; Okano et al., 1998a)
DNMT3B	DNA methylation establishment in early development and germ cells (crucial for methylation of pericentromeric major and minor satellite repeats)	Highly expressed in embryonic tissue and undifferentiated ESCs (ubiquitous expression)	*Dnmt3b* knockout embryos show embryonic lethality between E13.5 and E16.5, multiple developmental defects and demethylation of major and minor satellite repeats.	(Kaneda et al., 2004; Okano et al., 1999; Okano et al., 1998a; Ueda et al., 2006)
DNMT3L	Non-catalytic activity; essential for the establishment of imprints in oocytes and for silencing of dispersed repeated sequences in male germ cells	Specific expression in germ cells during gametogenesis and embryonic stages	*Dnmt3l* knockout mice are viable; knockout males are sterile, because their germ cells show reactivation of LINE1 and IAP and severe meiosis defects; knockout females fail to deliver viable pups (developing embryos die due to neural tube defects, partly due to bi-allelic expression of maternally imprinted genes owing to loss of imprints).	(Bourc'his and Bestor, 2004; Bourc'his et al., 2001; Hata et al., 2002; Webster et al., 2005)
NP95 (UHRF1)	Recruitment of DNMT1 into the replication foci	Highly expressed in proliferating cells (ubiquitous expression)	Embryonic lethality shortly after gastrulation; extensive demethylation of the genome (similar phenotype to *Dnmt1* knockout embryos).	(Bostick et al., 2007; Fujimori et al., 1998; Sharif et al., 2007)
TET1	Conversion of 5mC to 5hmC; limits accessibility of DNA to DNMTs by binding strongly to unmethylated CpG-rich regions via its CXXC domain	Highly expressed in ESCs	*Tet1* knockdown ESCs show morphological abnormality, decreased AP activity, reduced 5hmC levels and increased DNA methylation in Tet1 binding regions. *Tet1* knockdown embryos fail to form normal blastocysts. *Tet1* knockout ESCs are pluripotent and support full-term mouse development in tetraploid complementation assay. *Tet1* knockout mice are viable and fertile with a reduced litter size; some are small in size.	(Dawlaty et al., 2011; Ito et al., 2010; Koh et al., 2011; Szwagierczak et al., 2010; Tahiliani et al., 2009; Xu et al., 2011)
TET2	Conversion of 5mC to 5hmC	Expressed in ESCs and hematopoietic cells (expression in almost all tissues)	*Tet2* knockdown ESCs show normal morphology, normal AP activity but reduced 5hmC levels; mutations in *Tet2* results in hematopoietic malignancies.	(Figueroa et al., 2010; Ito et al., 2010; Ko et al., 2011; Koh et al., 2011; Langemeijer et al., 2009; Moran-Crusio et al., 2011; Quivoron et al., 2011; Szwagierczak et al., 2010)
TET3	Conversion of 5mC to 5hmC	Highly expressed in oocyte and zygote (expression in NSCs, lung, spleen and pancreas, etc.)	Zygotes injected with Tet3 siRNA show reduced 5hmC and elevated 5mC signals in the paternal pronuclei. A *Tet3* maternal knockout leads to a failure in the elevation of 5hmC and the concomitant reduction of 5mC from the paternal genome, and frequent death of the resulting embryos.	(Gu et al., 2011; Iqbal et al., 2011; Ito et al., 2010; Szwagierczak et al., 2010; Wossidlo et al., 2011)
AID	Deamination of cytosine to uracil in ssDNA, of 5mC to thymine and of 5hmC to 5hmU	Expressed in activated B cells and testis; expressed in oocytes, ESCs and PGCs (low level)	AID knockout mice fail in class switch recombination and somatic hypermutation, and result in hyper-IgM phenotype. AID knockout mice show reduced genome-wide DNA demethylation in PGCs (E13.5). AID knockdown leads to reduced epigenetic reprogramming.	(Muramatsu et al., 1999; Morgan et al., 2004; Muramatsu et al., 2000; Bhutani et al., 2010; Guo et al., 2011; Popp et al., 2010)
APOBEC1	Deamination of cytosine to uracil in RNA and DNA, of 5mC to thymine and of 5hmC to 5hmU	Expressed in small intestine	APOBEC1 knockout mice show no apparent phenotype.	(Guo et al., 2011; Hirano et al., 1996; Morgan et al., 2004; Morrison et al., 1996; Navaratnam et al., 1993; Teng et al., 1993)

5hmC, 5-hydroxymethylcytosine; 5hmU, 5-hydroxymethyluracil; 5mC, 5-methylcytosine; AID, activation-induced deaminase; AP, apurinic/apyrimidinic; APOBEC1, apolipoprotein B mRNA editing enzyme, catalytic polypeptides; DNMT, DNA methyltransferase; ESC, embryonic stem cells; IAP, intracisternal A particle; IgM, immunoglobulin M; LINE1, long interspersed nuclear element 1; NSC, neural stem cell; PGC, primordial germ cell; siRNA, short interfering RNA; ssDNA, single stranded DNA; TET, ten-eleven translocation; UHRF1, ubiquitin-like, containing PHD and RING finger domains 1.

Box 1. 5mC: a major source of point mutations

5mC can be spontaneously deaminated to form thymine (T), creating a T:G mismatch (Duncan and Miller, 1980; Poole et al., 2001). Although T:G mismatches can be repaired by the base-excision repair (BER) system, which uses enzymes that have thymine DNA glycosylase activity, such as MBD4 (methyl CpG binding domain protein 4) and TDG (thymine DNA glycosylase) (Poole et al., 2001), they are often unrecognized and lead to point mutations after DNA replication. Thus, 5mC is a source of point mutations across the genome (Kondrashov, 2003). Indeed, owing to the high mutability of the 5mCpG sequence, the frequency of the CpG sequence is much lower (~0.2-0.25×) than would be expected given the GC content of the genome both in mice and humans (Lander et al., 2001; Rollins et al., 2006; Saxonov et al., 2006; Waterston et al., 2002).

al., 2008; Weber et al., 2007), possibly through the interaction of EZH2 (enhancer of zeste homologue 2), a component of PRC2, with DNMT3A and DNMT3B (Vire et al., 2006).

DNA demethylation in mouse pre-implantation development

Mouse development commences with fertilization – the fusion of an ovulated haploid oocyte with a haploid spermatozoon. Up to blastocyst formation, parental genomes undergo extensive epigenetic reprogramming, most notably genome-wide DNA demethylation. However, some genomic regions escape demethylation at this stage, including centromeric repeats, intracisternal A particle (IAP) retrotransposons (~1000 elements/mouse genome) and the differentially methylated regions (DMRs) (see Glossary, Box 2) that are present in parentally methylated imprinted genes, as well as in some non-imprinted genes (Borgel et al., 2010; Lane et al., 2003; Reik et al., 2001; Rougier et al., 1998).

DNA demethylation in mouse zygotes

A recent genome-wide bisulfite sequence analysis (see Box 3 for more about bisulfite sequence analysis and other techniques for assaying DNA methylation), which covered ~1% of the mouse genome, reported that ~80% of CpG sequences are methylated in sperm (Popp et al., 2010). The maternal genome in mouse oocytes has lower levels of genome-wide DNA methylation than does the paternal genome in sperm, although the precise extent of genome-wide DNA methylation in the maternal genome has yet to be determined (Howlett and Reik, 1991; Monk et al., 1987; Smallwood et al., 2011). Within 1 hour of fertilization, the paternal genome releases protamine and is re-packaged by maternal nucleosomal histones, forming the paternal pronucleus. Either subsequently or concomitantly, the paternal pronucleus enlarges substantially by incorporating further maternal proteins, such as stella (also known as PGC7 and DPPA3, developmental pluripotency associated 3) and nucleoplasmin 2 (NPM2) (Li et al., 2010).

The development of the zygote is defined by the pronuclear stages P0/1 to P5 (Adenot et al., 1997; Santos et al., 2002). P0, P1 and P2 embryos are in the G1 phase, P3 and P4 embryos are largely in the S phase, replicating both the paternal and maternal genomes, and P5 embryos are mostly in the post-replicative G2 phase (Adenot et al., 1997). Several reports have shown that the paternal genome undergoes genome-wide DNA demethylation before replicating its DNA (before or in early P3) via an active

mechanism (Mayer et al., 2000; Oswald et al., 2000; Santos et al., 2002; Wossidlo et al., 2010). By around P3 (7-8 hours post-fertilization), the paternal genome appears to lose a substantial amount of 5mC immunofluorescence (see Box 3 for more about immunofluorescence analysis), whereas the maternal genome retains it at a seemingly constant level. Treatment of zygotes with aphidicolin, which blocks DNA replication, has no effect on the loss of 5mC immunofluorescence from the paternal genome, indicating that a replication-independent, active DNA demethylation mechanism occurs in early zygotes (Mayer et al., 2000; Oswald et al., 2000; Santos et al., 2002; Wossidlo et al., 2010).

In mouse zygotes that lack stella, a maternal-effect protein that is essential for pre-implantation development (Payer et al., 2003), a substantial decrease of 5mC immunofluorescence is also observed in the maternal genome, indicating that the maternal genome is normally protected from demethylation by stella (Nakamura et al., 2007). Bisulfite sequence analysis has also shown that in stella-deficient zygotes, some of the paternally [*H19*, *Rasgrf1* (RAS protein-specific guanine nucleotide-releasing factor 1)] and maternally [*Peg* (paternally expressed gene) *1*, *Peg3*, *Peg10*] methylated imprinted genes, as well as IAPs, are demethylated at the P5 stage: these genes remain methylated at this stage in wild-type zygotes. Therefore, stella also protects some imprinted genes from demethylation. The reason(s) that stella can protect only paternally imprinted genes and the maternal genome remains unknown, particularly because stella localizes to both the paternal and maternal pronucleus (Nakamura et al., 2007; Payer et al., 2003). Underlying differences in chromatin modifications between the paternal and maternal genome may account, at least in part, for this asymmetric action of stella (Nakamura et al., 2007).

Although the immunofluorescence-based observations described above appear to be significant, bisulfite sequencing analyses offer less substantial evidence for DNA demethylation in mouse zygotes before they undergo DNA replication (Lane et al., 2003; Wossidlo et al., 2010). For example, the levels of 5mC in LINE1 (see Glossary, Box 2) elements (~6×10^5 elements/mouse genome, see Glossary, Box 2) in zygotes at P1 is about ~68% and drops to ~53% at early P3 and then to ~27% after DNA replication (Wossidlo et al., 2010). The CpG methylation level of early retrotransposons (ETn) (~300-400 elements/mouse genome) at P1 is about ~77% and drops to ~61% at early P3; after DNA replication, methylation levels somehow then rise to ~73% (Wossidlo et al., 2010). Another study reported on the demethylation of LINE1 and ETn elements by the P4 stage; LINE 1 methylation drops from 87% to 55% and ETn from 89% to 66% (Okada et al., 2010). The demethylation of imprinted genes and IAPs in stella-deficient zygotes is also observed after DNA replication (Nakamura et al., 2007). Thus, although the extent of DNA demethylation in zygotes after DNA replication is substantial, that occurring before DNA replication, by active DNA demethylation, seems less prominent. The discrepancy between the immunofluorescence and bisulfite sequence data thus needs to be resolved. As we discuss below, the recent discovery that 5mC can be oxidized to form 5-hydroxymethylcytosine (5hmC) might offer such a resolution.

5hmC: an intermediate for DNA demethylation?

5hmC, a stable hydroxylated metabolite of 5mC, was first identified in the genome of T-even bacteriophages (Wyatt and Cohen, 1953) and is produced as an oxidation damage product from 5mC (Burdzy et al., 2002; Zuo et al., 1995). However, the

Box 2. Glossary

Base excision repair (BER). A DNA repair pathway that removes mismatched DNA bases, followed by incision of the 5′ phosphodiester bond of the abasic site and gap filling by a DNA polymerase.

CpG island. A genomic region not depleted of CpGs that is typically 200-500 bp in length, has a minimum observed:expected CpG ratio of >0.6 and a minimum GC content of 50-55%.

Differentially methylated region (DMR). A genomic region that is differentially DNA methylated between the two parental chromosomes.

Elongator complex. A protein complex that associates with the RNA polymerase II holoenzyme in transcription elongation and exhibits histone acetyltransferase activity.

Endosperm. A nutritive tissue of flowering plant seeds that is formed by the fertilization of the maternal central cell

HIRA. A histone chaperone for H3/H4 for nucleosome assembly independent of DNA replication.

Histone deacetylase complex. A transcription repressor complex that involves Sin3 and histone deacelylases.

LINE1 (long interspersed nuclear element 1). A retrotranspon-derived genetic element that encodes reverse transcriptase and integrase. Around 6×10^5 LINE1 elements are present in the mouse genome, constituting ~19% of the genome.

NAP1 (nucleosome assembly protein 1). A histone chaperone for H2A/H2B and H1/B4 that removes histones and is implicated in transcription factor binding to DNA.

Nucleotide excision repair (NER). A DNA repair pathway that recognizes DNA lesions that result in conformational distortions, such as a thymine dimer.

Polycomb repressive complex (PRC). Transcriptional repressor complexes. PRC1 contains ubiquitin ligases RING1A and RING1B, which catalyze mono-ubiquitylation of Lys 119 of histone H2A. PRC2 contains EED, SUZ12, RbAp46/48 and EZH1/2, a histone methyltransferase responsible for di-/tri-methylation of H3K27.

Replication complex. A macromolecular structure in which eukaryotic DNA replication occurs.

SIN3A repressor complex. A transcription repressor complex that contains SIN3A and histone deacetylases, together with other proteins.

TET (ten-eleven translocation). Proteins that catalyze oxidization of 5-methylcytosine and produce 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcytosine in a 2-oxoglutarate- and Fe(II)-dependent manner, through conserved catalytic domains (Cys-rich and dioxygenase domains).

Thymine DNA glycosylase. DNA glycosylases that catalyze the base excision of thymine mismatched with guanine (e.g. TDG and MBD4).

5mC glycosylase/lyase. An enzyme that removes 5-methylated cytosine from the backbone sugar of DNA.

discovery that it is a physiologically relevant DNA modification in mammals, such as in mouse neurons and ESCs (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009), is a more recent novel finding.

The hydroxylation of 5mC to 5hmC is catalyzed by a family of dioxygenases, the TET (ten-eleven translocation) 1/2/3 proteins (see Glossary, Box 2), which have different tissue distributions (Cimmino et al., 2011; Szwagierczak et al., 2010; Tahiliani et al., 2009) (Table 1). 5hmC is abundant in the brain (~40% and ~13% as abundant as 5mC in Purkinje and granule cells, respectively), but is present at lower levels in other mouse tissues (Kriaucionis and Heintz, 2009). 5hmC is detected in mouse ESCs (~7-10% as abundant as 5mC) but is undetectable in human T cells and in mouse dendritic cells (Tahiliani et al., 2009). These findings raise the possibility that demethylation of 5mC to cytosine occurs via the

generation of a 5hmC intermediate, which is in turn converted into unmethylated cytosine. Furthermore, recent studies show that TET proteins further convert 5hmC into 5-formylcytosine (5fC) and then into 5-carboxylcytosine (5caC) (He et al., 2011; Ito et al., 2011; Pfaffeneder et al., 2011). The genomic contents of these cytosine derivatives in mouse ESCs are, however, very low, e.g. 20 5fC and three 5caC in every 10^6 Cs (5hmC is about 1.3×10^3 in every 10^6 Cs), and the significance of these modifications needs to be clarified (Ito et al., 2011).

5hmC may also be a biological end-product of demethylation, as methyl-CpG binding proteins have a significantly lower affinity for 5hmC (Valinluck et al., 2004). DNMT1 also recognizes 5hmC very poorly in vitro (Valinluck and Sowers, 2007); thus, 5hmC might passively convert into cytosine during replication. A more recent report, however, shows that Np95, which recruits DNMT1 to replication foci, recognizes 5hmC as efficiently as it does 5mC (Frauer et al., 2011), raising the possibility that 5hmC may have the same capacity as 5mC for 5mC propagation during replication. The biological significance of 5hmC thus requires further clarification.

As we discuss in more detail below, recent studies of the TET proteins have revealed more about their genome-wide binding sites, their functions in epigenetic reprogramming and about the genome-wide distribution of 5hmC.

5hmC and TET proteins in mouse ESCs
Mouse ESCs highly express *Tet1*, express *Tet2* to a lesser extent and do not express *Tet3* (Ito et al., 2010; Koh et al., 2011). Upon ESC differentiation, both *Tet1* and *Tet2* are downregulated (Ito et al., 2010; Koh et al., 2011). When knocked down by RNAi, *Tet1* and *Tet2* were found to be involved in regulating the expression of pluripotency transcription factors, such as *Nanog*, *Esrrb* (estrogen-related receptor β) and *Prdm14* (PR domain containing 14) (Ito et al., 2010; Koh et al., 2011; Ficz et al., 2011; Williams et al., 2011). The depletion of *Tet1* and *Tet2* skews ESC differentiation towards the extra-embryonic lineages (Ito et al., 2010; Koh et al., 2011; Ficz et al., 2011; Williams et al., 2011). Surprisingly, however, *Tet1* knockout ESCs, which show a ~35% reduction in 5hmC levels, exhibit only subtle changes in gene expression, are pluripotent and support full-term mouse development in the tetraploid complementation assay (Dawlaty et al., 2011). To examine the possibility that *Tet1* and *Tet2* are functionally redundant and to investigate 5hmC functions further in ESCs, *Tet1* and *Tet2* double knockout ESCs will need to be generated in the near future.

Chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq) for TET1 has shown that most TET1-binding sites are in the transcribed regions of genes, with the highest density around transcription start sites (TSSs) of HCPs and ICPs (Williams et al., 2011; Wu et al., 2011; Xu et al., 2011); Williams et al. (Williams et al., 2011), for example, reported ~6500 TSSs with TET1-binding sites. The CXXC zinc-finger domain of TET1 is required to recruit TET1 onto CpG-rich sequences (Xu et al., 2011). TET1 binding is positively correlated with H3K4me3 and also with bivalent chromatin modifications. 5hmC immunoprecipitation followed by DNA sequencing (hMeDIP-seq, see Box 3) shows that 5hmC is also enriched within gene bodies and at TSSs of HCPs and ICPs (Ficz et al., 2011; Pastor et al., 2011; Williams et al., 2011; Xu et al., 2011). Williams et al. (Williams et al., 2011), for example, identified ~2400 TSSs enriched for 5hmC. In DNMT1/DNMT3A/DNMT3B triple-knockout (TKO) cells that lack all 5mC (Tsumura et al., 2006), nearly all 5hmC signals were found to be absent (Ficz et al., 2011; Szwagierczak et al., 2010;

Box 3. Methods for genome-wide quantitation of DNA methylation

Methylation-sensitive/dependent enzyme digestion

Genomic DNA samples digested with methylation-sensitive and -insensitive enzymes (e.g. *Hpa*II and *Msp*I, respectively) are compared by microarrays (Tompa et al., 2002) or by next-generation sequencing (Oda et al., 2009). Genomic DNA can also be digested by McrBC, which digests nearly all methylated CpG islands (Sutherland et al., 1992), and then compared with non-digested DNA (Irizarry et al., 2008; Lippman et al., 2004). These methods depend on enzyme restriction sites.

Methylated/hydroxymethylated DNA immunoprecipitation (MeDIP/hMeDIP)

5-methylcytosine (5mC) or 5-hydroxymethylcytosine (5hmC) in fragmented genomic DNA is enriched by immunoprecipitation using specific antibodies and analyzed using microarray (MeDIP/hMeDP-chip) (Weber et al., 2005) or next-generation sequencing (MeDIP/hMeDIP-seq) (Down et al., 2008). These methods do not quantify absolute 5mC/5hmC levels. The densities of 5mC and 5hmC influence the efficiencies of immunoprecipitation (Pastor et al., 2011; Weber et al., 2007). The GLIB (glucosylation, periodate oxidation, biotinylation) method enables the efficient pulldown of 5hmC, even when it is at a low density (Pastor et al., 2011). MeDIP-chip has been combined with DNA amplification and applied to ~10^4 cells from early mouse embryos (Borgel et al., 2010).

Bisulfite sequencing

Genome DNA is treated with sodium bisulfite, which converts cytosine, but not 5mC/5hmC, to uracil, and analyzed using next-generation sequencing (Cokus et al., 2008; Lister et al., 2008). This method determines methylation sites at single-base resolution, and, with sufficient read depths, absolutely quantifies methylation levels. However, 5mC and 5hmC are indistinguishable by this method. By reducing genome representation with *Msp*I digestion (Meissner et al., 2005), this method has been applied to ~10^3 cells from oocytes and pre-implantation embryos (Smallwood et al., 2011).

Immunofluorescence analysis

5mC/5hmC are marked by specific antibodies and by fluorophore-conjugated secondary antibodies in situ, followed by fluorescent microscopic analysis. This method cannot distinguish the methylation state of specific sequences, but detects genome-wide methylation levels in single cells. 5mC/5hmC antibodies detect densely methylated sequences efficiently but single CpG methylation less efficiently (Pastor et al., 2011; Suzuki and Bird, 2008). As transposon-related elements occupy ~40% of the genome and genes only ~2-3% (Lander et al., 2001; Waterston et al., 2002), most 5mC/5hmC signals should be derived from the methylation of transposon-related elements.

Williams et al., 2011). These findings indicate that in ESCs, 5hmC is generated from pre-existing 5mC by the action of TET1, and that a significant fraction of 5mC is converted to 5hmC at the TSSs of HCPs and ICPs, where 5mC becomes depleted. One of the functions of TET1 would therefore be to remove aberrant stochastic DNA methylation from HCPs and ICPs, thereby regulating DNA methylation fidelity in ESCs.

TET1 contributes to the transcriptional repression of a fraction [~7-8%, as reported previously (Williams et al., 2011)] of its target genes and to a lesser extent their transcriptional activation [~3% as reported previously (Williams et al., 2011)]. Most of the transcriptional effects of TET1 are independent of the conversion of 5mC to 5hmC, as TET1 has similar transcriptional activity in DNMT TKO ESCs (Williams et al., 2011). Instead, TET1 contributes to transcriptional repression by forming a complex with

the SIN3A repressor complex (see Glossary, Box 2) or indirectly by recruiting PRC2 (Williams et al., 2011; Wu et al., 2011). Whether the catalytic activity of TET1 is required for the functions of TET1 in ESCs thus remains to be explored.

5hmC and DNA demethylation in mouse zygotes

Immunofluorescence analysis has shown that 5hmC levels elevate on the paternal genome at around postnatal day (P) 3, concomitant with the reduction of 5mC (Gu et al., 2011; Iqbal et al., 2011; Wossidlo et al., 2011). This elevation occurs independently of DNA replication, and 5hmC persists at least until the two-cell stage (Iqbal et al., 2011; Wossidlo et al., 2011). In stella-deficient zygotes, 5hmC increases and 5mC decreases on both the paternal and maternal genomes (Wossidlo et al., 2011), and when *Tet3* is knocked down (*Tet3* is highly expressed in oocytes and zygotes), 5mC levels increase whereas 5hmC levels reduce, compared with wild type, on the paternal genome (Wossidlo et al., 2011). Importantly, a maternal knockout of *Tet3* leads to a failure in the elevation of 5hmC and in the reduction of 5mC from the paternal genome, impaired promoter demethylation of *Oct4* (*Pou5f1*, POU domain, class 5, transcription factor 1) and *Nanog*, delay in the activation of a paternally derived *Oct4* transgene, and frequent death of the resulting embryos (Gu et al., 2011). These findings suggest that, in normal development, the paternal genome is targeted by TET3, which converts 5mC to 5hmC from around the P3 stage onwards and that the TET3-mediated hydroxylation of 5mC accounts, at least in part, for the active DNA demethylation of the paternal genome.

Crucially, both 5mC and 5hmC are resistant to deamination by bisulfite treatment and are indistinguishable in bisulfite sequence analysis (Hayatsu and Shiragami, 1979). This may explain why the paternal genome seems to retain persistent levels of methylation by bisulfite sequence analysis, despite the fact that it shows highly reduced 5mC immunofluorescence. The development of a technology that can discriminate between cytosine, 5mC and 5hmC by quantitative sequencing analysis is crucial for obtaining more detailed information on DNA demethylation of the paternal genome in zygotes.

As discussed earlier, the mechanism by which 5hmC is converted into cytosine in zygotes remains unclear. In the plant *Arabidopsis thaliana*, it is already well established that DNA demethylation involves 5mC glycosylases/lyases (see Glossary, Box 2; Box 4) and the base excision repair (BER) pathway (see Glossary, Box 2) (Zhu, 2009). This pathway contributes to the genome-wide DNA demethylation that occurs in the endosperm (see Glossary, Box 2) (Gehring et al., 2009; Hsieh et al., 2009). Although there are as yet no known mammalian homologues of plant 5mC glycosylases/lyases, there is evidence that the BER, but not the nucleotide excision repair (NER) (see Glossary, Box 2), pathway is involved in the DNA demethylation of the mammalian paternal genome (Hajkova et al., 2010; Wossidlo et al., 2010; Ziegler-Birling et al., 2009). Accordingly, γH2A.X, the Serine[139] phosphorylated form of the histone H2 protein H2AX, which marks DNA strand breaks, and PARP1 [poly(ADP-ribose) polymerase family, member 1], a sensor of single-stranded DNA (ssDNA) breaks and a component of the BER pathway, are recognized specifically on the paternal genome at early P3 (Hajkova et al., 2010; Wossidlo et al., 2010; Ziegler-Birling et al., 2009). At this stage, XRCC1 (x-ray repair complementing defective repair in Chinese hamster cells 1), a core BER component, is tightly bound only to the paternal genome. In stella-deficient zygotes, XRCC1 binds to both the paternal and maternal

> **Box 4. DNA demethylation in plants**
> Compelling genetic and biochemical evidence exists in *Arabidopsis thaliana* that active DNA demethylation is carried out by 5-methylcytosine (5mC)-specific glycosylases/lyases of the DEMETER family, which consists of four members: DME; repressor of silencing 1 (ROS1, also known as DML1); DML2; and DML3 (Zhu, 2009). These proteins remove 5mC by a glycosylation reaction and cleave one of the phosphodiester bonds. Subsequently, the remaining sugar and phosphate group is removed by an apurinic/apyrimidinic endonuclease and a phosphodiesterase, a proper nucleotide is then inserted by a DNA repair polymerase, and the nick is sealed by a DNA ligase. DME functions to demethylate transposable elements and imprinted genes globally in the endosperm of plants, thereby allowing their parent-of-origin specific (maternal) expression (Gehring et al., 2009; Hsieh et al., 2009).

genome and inhibition of PARP and APE1 (apurinic/apyrimidinic endonuclease 1) activity results in the reduced demethylation of the paternal genome (Hajkova et al., 2010). It is possible that the TET3-mediated hydroxylation of 5mC on the paternal genome directly or indirectly triggers the BER pathway. This possibility needs to be verified experimentally.

Other mechanism of DNA demethylation in mouse zygotes
It has been reported that components of the elongator complex (see Glossary, Box 2), including ELP1, ELP3 and ELP4, are involved in the pre-replicative DNA demethylation of the paternal genome (Okada et al., 2010). The radical SAM (S-adenosylmethionine) domain but not the HAT (histone acetyltransferase) domain of ELP3 appears to be required for this activity. The mechanism by which the elongator complex is involved in demethylating the paternal genome remains to be explored.

DNA demethylation in pre-implantation embryos
There is evidence that DNA methylation is passively removed both from the paternal and maternal genomes from the first S-phase (one-cell stage) up to the morula/early blastocyst stage (Howlett and Reik, 1991; Kafri et al., 1992; Lane et al., 2003; Monk et al., 1987; Oda et al., 2006; Okano et al., 1999; Rougier et al., 1998). Embryos at the morula/early blastocyst stage are therefore considered to bear substantially lower levels of genome-wide DNA methylation than do zygotes. A study that used reduced representation bisulfite sequencing (RRBS) (see Box 3) has recently shown that CGIs that are methylated in mature oocytes are indeed demethylated in blastocysts, but not to the extent that would be expected if passive demethylation occurs at every cleavage division, indicating that mechanisms of DNA demethylation in pre-implantation embryos need to be further investigated (Smallwood et al., 2011). Given that ~10 primitive ectoderm (PEct) cells constitute the inner cell mass (ICM) of ~E4.0-4.5 blastocysts and give rise to all somatic and germ cells, it remains an important challenge to elucidate the epigenome of the primitive ectoderm.

DNA demethylation in pre-implantation embryos could partly be due to a reduction in DNMT1, as DNMT1, but not DNMT3A or DNMT3B, immunofluorescence is excluded from the nucleus during pre-implantation development (Branco et al., 2008; Hirasawa et al., 2008). The DNA methylation of imprinted genes, IAPs and centromeric repeats is, however, maintained during this period (Borgel et al., 2010; Lane et al., 2003; Reik et al., 2001; Rougier et al., 1998). Interestingly, the conditional knockout of both maternal and zygotic *Dnmt1* leads to a complete erasure of

DNA methylation at imprinted genes in the blastocyst, demonstrating that DNMT1, which is present in the nuclei of pre-implantation embryos at a low level that is undetectable by immunofluorescence analysis, is sufficient to maintain the DNA methylation of imprinted genes (Hirasawa et al., 2008). The conditional knockout of both maternal and zygotic *Dnmt3a* and *Dnmt3b* leads to a partial demethylation of a paternally imprinted gene, *Rasgrf1*, at E9.5 (Hirasawa et al., 2008), indicating that imprinting maintenance requires the presence of all three DNMTs. The maintenance of DNA methylation at IAPs in pre-implantation embryos also depends on DNMT1 function (Gaudet et al., 2004).

Recent studies have re-examined the long-held view that the DMRs of imprinted genes are resistant to genome-wide DNA demethylation during pre-implantation development (Kobayashi et al., 2006; Tomizawa et al., 2011). Accordingly, the DMRs of imprinted genes, particularly of paternally imprinted genes, are partly demethylated during pre-implantation development, especially at their peripheral regions, and are subsequently remethylated, exhibiting an unexpectedly dynamic regulation (Tomizawa et al., 2011). The mechanism that targets DNMTs to demethylation-resistant sequences remains to be clarified.

Functional significance of DNA demethylation
What is the functional significance of DNA demethylation in pre-implantation embryos? It is most likely to be in the creation of the pluripotent epigenome of the primitive ectoderm (PEct) (Surani et al., 2007).

Genome-wide promoter methylation in sperm seems to be generally similar to that in ESCs and embryonic germ cells (EGCs), except at certain loci that encode pluripotency factors, such as *Nanog* and *Brd1* (bromodomain containing 1) (Farthing et al., 2008). Therefore, the genome-wide DNA demethylation of the paternal genome in the zygote may occur preferentially at transposons, such as at LINE1 elements, and perhaps in intergenic and intragenic regions (Farthing et al., 2008). Given that most imprints conferred during germ cell development are maternally derived, demethylation of the paternal genome may be a consequence of a need for the maternal cytoplasm to erase paternal imprints (Reik and Walter, 2001).

One report has shown that when round spermatids, the DNA of which is still associated with histones, are injected into oocytes by round spermatid injection (ROSI), paternal genome demethylation is not observed, whereas when mature sperm, the DNA of which is associated mainly with protamines, are injected into oocytes by intracytoplasmic sperm injection (ICSI), the paternal genome is demethylated, indicating that the protamine-histone exchange that occurs in the paternal genome once it is in the oocyte may cause paternal DNA demethylation (Polanski et al., 2008). Notably, both ROSI- and ICSI-derived embryos develop to term at the same ratio, indicating that paternal genome demethylation may have no functional significance in development (Polanski et al., 2008). As this study used only immunofluorescence analysis to detect 5mC, it is possible that functionally important DNA demethylation of the ROSI-derived paternal genome escaped detection.

Epigenetic reprogramming in primordial germ cells
The blastocyst at implantation (~E4.0-E4.5) consists of three cell types, the trophectoderm (TE), PEct and the primitive endoderm (PE) (Rossant and Tam, 2009) (Figs 1, 2). After implantation, the PEct gives rise to the epiblast, the source of all somatic cells, including the PGCs (Figs 1, 2). Genome-wide DNA methylation levels increase in PEct-derived tissues in response to the

activities of DNMT3A and DNMT3B (Borgel et al., 2010; Kafri et al., 1992; Oda et al., 2006; Okano et al., 1999). For example, many gene-specific CpG sequences that are demethylated by the blastocyst stage, are remethylated by E6.5 (Kafri et al., 1992). A more recent study has shown that during implantation, de novo DNA methylation conferred mainly by DNMT3B is primarily targeted to the CGIs of many germline genes, as well as to lineage-specific genes, to repress their expression (Borgel et al., 2010). Moreover, the methylation levels of the major satellite sequences, of LINE1 elements and of IAP elements increase from the blastocyst to the E8.5 stage from 15 to 80% (major satellites), 30 to 80% (LINE1 elements) and from 60 to 95% (IAP elements) (Oda et al., 2006). Compared with the PEct-derived embryonic tissues, the TE-derived placenta remains hypomethylated, with 30% of major satellites, 40% of LINE1 and 65% of IAPs being methylated at E9.5 (Oda et al., 2006). At E13.5, the genome-wide methylation level of placenta is 43.2% (Popp et al., 2010).

PGC specification takes place in the most proximal epiblast in response to bone morphogenetic protein (BMP) signalling from the extra-embryonic ectoderm at ~E6.0 (Lawson et al., 1999) (Fig. 2A,B). At this stage, epiblast cells are still pluripotent but are being propelled towards somatic fates and are in the process of losing their pluripotency (Kurimoto et al., 2008). BMP signalling induces the expression of the transcriptional regulators BLIMP1 (also known as PRDM1, PR domain containing 1, with ZNF domain) and PRDM14, in the most proximal epiblasts at ~E6.25 and E6.5, respectively; the BLIMP1- and PRDM14-positive cells go on to form a cluster of ~40 alkaline phosphatase (AP)-positive PGCs at the base of the incipient allantois at ~E7.25 (Ginsburg et al., 1990; Ohinata et al., 2009; Ohinata et al., 2005; Vincent et al., 2005; Yamaji et al., 2008) (see Fig. 2). These established PGCs shut down the somatic transcriptional programme (for example, by turning off Hox gene expression), re-acquire the expression of pluripotency factors (such as Sox2) and prepare for the epigenetic reprogramming that manifests after E7.75. From ~E7.5, PGCs start
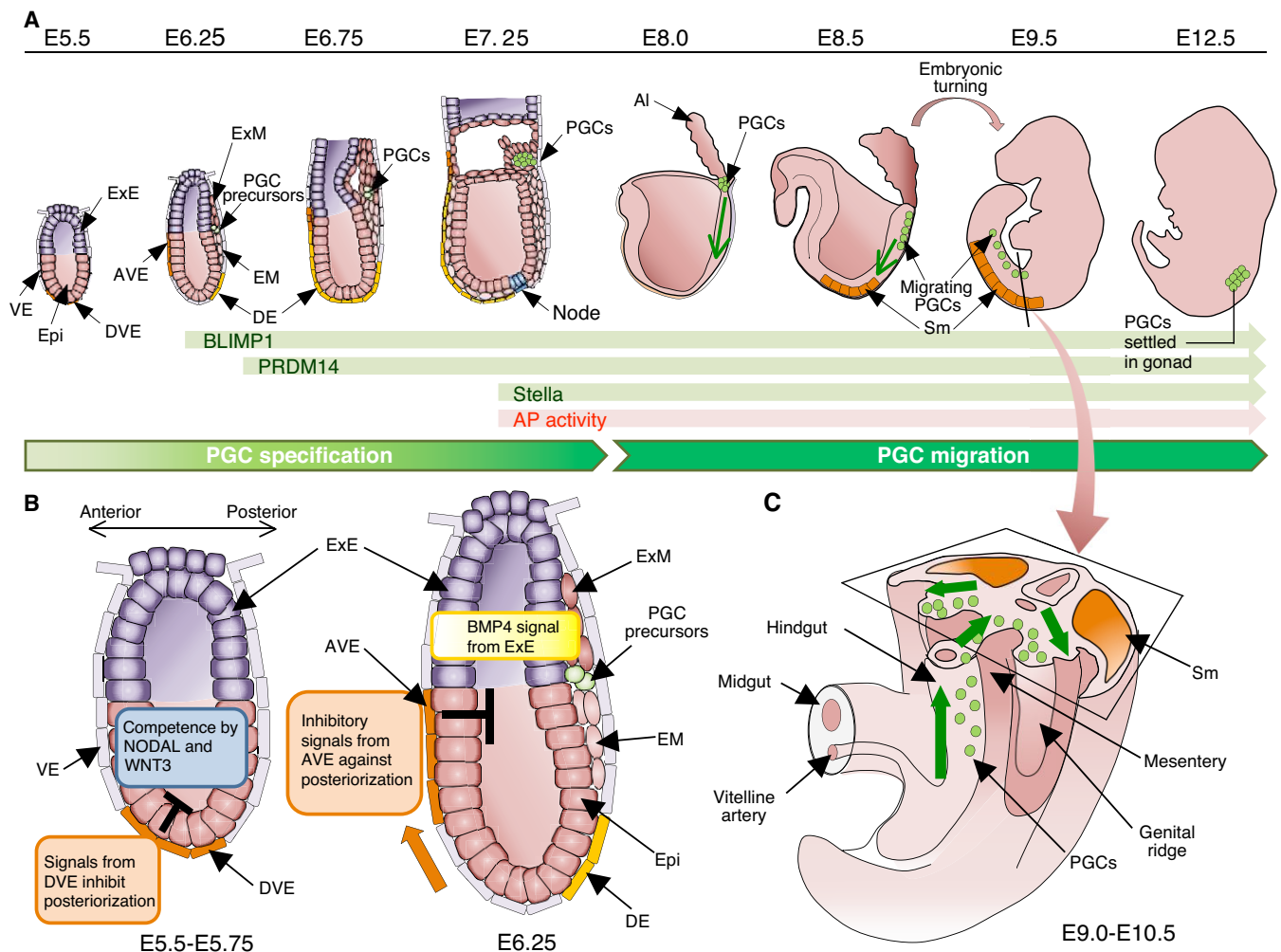

Fig. 2. Specification and migration of mouse primordial germ cells. (**A**) A schematic of germ cell specification and migration in developing mouse embryos (prospective anterior is towards the left). Primordial germ cell (PGC) precursors (E6.25) and PGCs are shown as green circles in embryos from E6.25 to E12.5, and the direction of PGC migration is denoted by a green arrow. The timing of expression of key genes (*Blimp1*, *Prdm14* and stella) and alkaline-phosphastase activity is shown below. (**B**) Signalling activities for PGC specification at E5.5-E5.75 and at E6.25. (**C**) A detailed view of PGC migration from the hindgut through the mesentery to genital ridges at E9.0-E10.5. The direction of PGC migration is denoted by a green arrow and anterior is towards the top. Al, allantois; AVE, anterior visceral endoderm; DE, distal endoderm; DVE, distal visceral endoderm; EM, embryonic mesoderm; Epi, epiblast; ExE, extra-embryonic ectoderm; ExM, extra-embryonic mesoderm; PGCs, primordial germ cells; Sm, somite; VE, visceral endoderm.

to migrate to the hindgut, from where they migrate to the mesentery and finally to the genital ridges, which they colonize by E10.5 to initiate sexually dimorphic development (Kurimoto et al., 2008; Saitou et al., 2002; Seki et al., 2005; Seki et al., 2007) (Fig. 2A,C). In PGCs, BLIMP1 is required for the repression of the somatic programme, and both BLIMP1 and PRDM14 are involved in the re-expression of pluripotency factors and in epigenetic reprogramming (Kurimoto et al., 2008; Yamaji et al., 2008).

Although the precise nature of the epigenome of the pre-gastrulating epiblast and of established PGCs at E7.25 is unknown and requires further investigation, we do know that, in early PGCs, methylation at imprinted loci is maintained (Hajkova et al., 2002; Lee et al., 2002), that one X chromosome in females is inactivated (Sugimoto and Abe, 2007; Tam et al., 1994) and that transposable elements, such as LINE1 and IAP, are relatively highly methylated (both are ~70% methylated at E11.5) (Hajkova et al., 2002). It is therefore likely that PGCs at their outset bear a genome-wide DNA methylation pattern that is comparable with that of somatic cells at the same stage.

DNA demethylation in PGCs

The most striking epigenetic event in PGCs is the genome-wide DNA demethylation that encompasses genic, intergenic and transposon sequences, which is completed in both sexes by E13.5 (see Fig. 3). As a consequence of this demethyation, one inactivated X-chromosome in females is reactivated, imprinted loci are fully demethylated and methylation at most transposable elements is erased (Hayashi and Surani, 2009). Notably, long terminal repeat (LTR) retrotransposon sequences, including IAPs, are more resistant to demethylation (Hajkova et al., 2002; Lane et al., 2003; Popp et al., 2010) and can cause transgenerational epigenetic inheritance (Whitelaw and Whitelaw, 2008).

A genome-wide bisulfite sequence analysis (covering ~1% of the genome) has quantified levels of 5mC in PGCs at E13.5, as well as in various cell types, and has shown that both male and female PGCs are extremely hypomethylated relative to other tissues (Popp et al., 2010). For example, whereas median methylation levels at CpGs in sperm are 85%, in ESCs they are 75%, in E13.5 embryos they are 73.2%, and in placenta they are 43.2%, those in E13.5
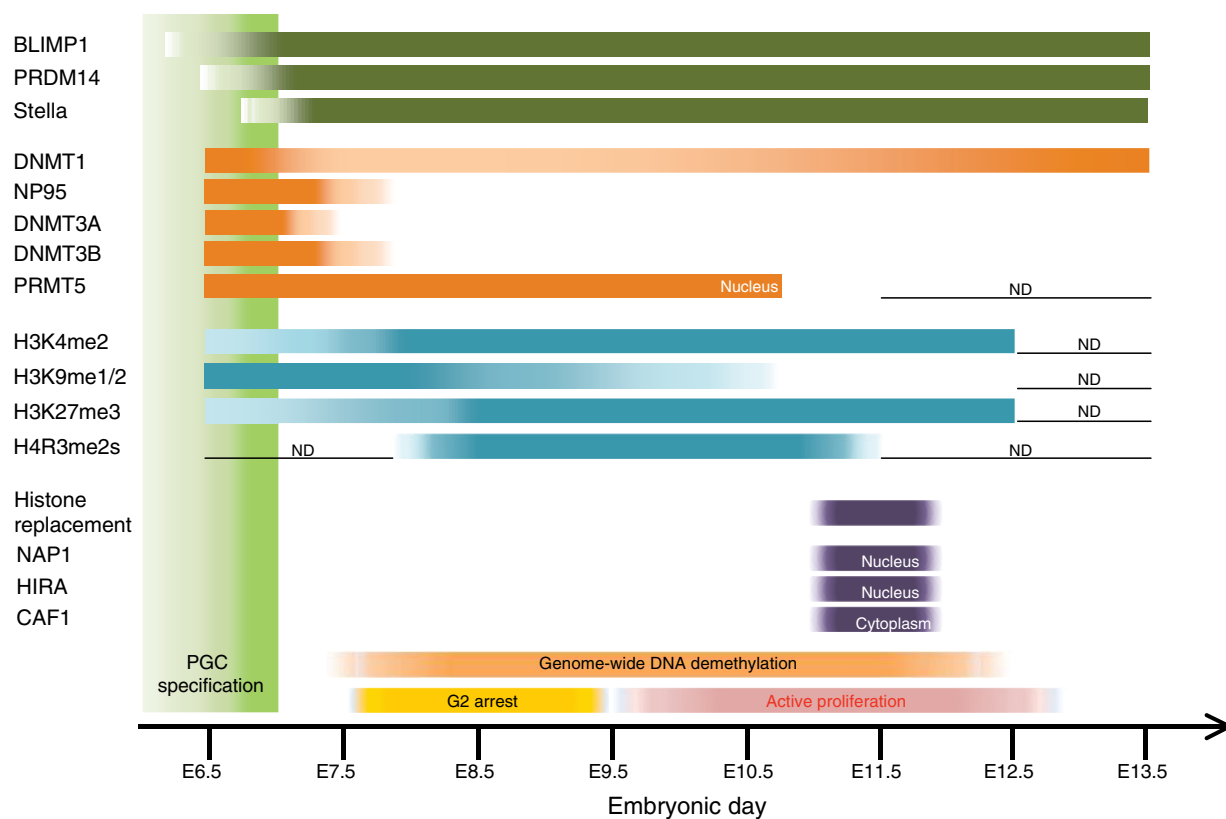

Fig. 3. A summary of epigenetic reprogramming in mouse primordial germ cells. A summary of the temporal expression patterns of key proteins involved in epigenetic reprogramming of mouse primordial germ cells (PGCs), and of the presence of key histone modifications, as revealed by immunohistochemistry and other methods (Hajkova et al., 2008; Popp et al., 2010; Seki et al., 2005; Seki et al., 2007). Migrating PGCs reduce H3K9me2 and upregulate H3K27me3 when they are arrested at the G2 phase of the cell cycle and become transcriptionally quiescent. Extensive remodelling of histone modifications occurs in the genital ridges at around E11.5, which perhaps reflects extensive histone replacement triggered by DNA repair (Hajkova et al., 2008; Hajkova et al., 2010). The level of genome-wide DNA methylation in PGCs becomes as low as or even lower than that of methylation-defective $Np95^{-/-}$ ES cells (Popp et al., 2010). The light-green box represents the period for PGC specification. The dark-green bars represent the expression of indicated genes associated with PGC specification. The orange bars represent the expression of the chromatin modifiers indicated. The light-blue bars represent the relative levels of the histone modifications indicated. The purple bars represent the occurrence of histone replacement and the expression of histone chaperones indicated (Hajkova et al., 2008; Hajkova et al., 2010). N.D., not determined. BLIMP1 (PRDM1), PR domain containing 1, with ZNF domain; CAF1, chromatin assembly factor 1; DNMT, DNA methyltransferase; HIRA, histone cell cycle regulation defective homolog A; NP95 (UHRF1), ubiquitin-like, containing PHD and RING finger domains 1; PRDM, PR domain containing; PRMT, protein arginine methyltransferase.

male and female PGCs are only 16.3% and 7.8%, respectively. Methylation levels in PGCs are substantially lower than the level of 22% that has been recorded in methylation-deficient $Np95^{-/-}$ ESCs. The fact that female PGCs have considerably lower methylation levels than do male PGCs may be because, owing to X-reactivation, female PGCs bear two active X-chromosomes, which may encode a modifier locus to lower genome-wide methylation levels (Zvetkova et al., 2005).

Dynamics of DNA demethylation in PGCs

The genome-wide DNA methylation state of PGCs before E13.5 has not yet been reported. Nonetheless, several studies have examined the timing of the demethylation of imprinted genes, some single-copy genes and transposable elements in PGCs from E10.5 to E12.5/E13.5 (Hajkova et al., 2002; Lane et al., 2003; Lee et al., 2002). In general, the timing of demethylation depends on the genes being analyzed and is thus heterogeneous. This finding might also reflect heterogeneity in the timing of demethylation in each PGC. One study has identified the rapid demethylation of imprinted genes between E11.5 and E12.5, proposing the involvement of active demethylation (Hajkova et al., 2002) when considering the doubling time of PGCs of ~16 hours (Tam and Snow, 1981); the DMRs of maternally methylated genes *Snrpn* (small nuclear ribonucleoprotein N), *Peg3* and *Lit1* [also known as *Kcnq1ot1* (potassium voltage-gated channel, subfamily Q, member 1, overlapping transcript 1)] are all nearly fully methylated at E11.5, but become almost fully demethylated at E12.5.

Conversely, other studies have demonstrated a gradual erasure of methylation at several imprinted genes and retrotransposons, such as LINE1 and IAP (Lane et al., 2003; Lee et al., 2002). Notably, imprinted genes such as *Nnat* (neuronatin), *H19* and *Peg10* are already partly (~50%) demethylated at E10.5 (Lee et al., 2002). These observations are compatible with the occurrence of replication-dependent passive demethylation. More comprehensive measurements of DNA methylation states during PGC development should provide further insights into the dynamics, and hence the mechanism, of DNA demethylation.

Active DNA demethylation in PGCs?

There is evidence that the cytosine deaminases AID (activation-induced deaminase) and APOBEC1 (apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1) (Box 5, Fig. 4, Table 1) can convert 5mCs to thymines by deamination, creating T:G mismatches that might then become targets of thymine glycosylases (see Glossary, Box 2), such as MBD4 (methyl CpG binding domain protein 4) or TDG (thymine DNA glycosylase; see Glossary, Box 2), which then trigger the BER pathway (Morgan et al., 2004). AID-deficient male and female PGCs at E13.5 show median methylation levels of ~22% and 20%, respectively, which are higher than the methylation levels of wild-type male and female PGCs (16.3% and 7.8%, respectively), indicating that AID functions in genome-wide DNA demethylation in PGCs (Popp et al., 2010). Importantly, AID deficiency does not impact genome-wide methylation levels in cells/tissues other than PGCs (Popp et al., 2010). As the methylation levels of AID-deficient PGCs are still lower compared with those of earlier wild-type PGCs, the demethylation events occur even without AID, possibly owing to compensation by other deaminases, including APOBEC1/2/3.

However, it is important to note that both AID-deficient male and female mice are relatively healthy (except for their B-cell-derived phenotype) and fertile, although some abnormalities in

litter size and progeny birth weights have been reported (Popp et al., 2010). Considering that the deregulated dose of even a single imprinted gene profoundly affects development and adult physiology, the genome-wide DNA demethylation that occurs in PGCs, which erases imprints and contributes to the creation of proper imprinted gene dose, should be a crucial event. Therefore, the finding that DNA demethylation deficiencies in AID-mutant PGCs does not lead to profound reproductive defects, such as infertility, subfertility or marked adult phenotypes, appears to be counter-intuitive. It has also been shown that AID is unable to act on double-stranded DNA and that 5mC is a much more inefficient target for AID-mediated deamination than is unmethylated cytosine in vitro (Bransteitter et al., 2003; Di Noia and Neuberger, 2007; Larijani et al., 2005), raising the issue of whether AID can directly deaminate 5mC in vivo.

In support of the active removal of 5mC during DNA demethylation in PGCs, the BER, but not the NER, pathway has been reported to operate in PGCs during their genome-wide DNA demethylation (Hajkova et al., 2010). As in the paternal pronucleus of the zygote, in ~E11.5 PGCs, components of this pathway, including XRCC1, APE1 and PARP1 are found to be enriched in their nuclei, and XRCC1 is found bound to PGC chromatin, suggesting that ssDNA breaks are present in PGCs. Hajkova et al. have shown that, as a potential consequence of the DNA demethylation mediated by the DNA repair mechanisms, PGCs at ~E11.5 show dramatic changes in their chromatin states, including rapid loss of linker histone H1, loss of detectable chromophores, significant enlargement of nuclei, and a concomitant loss of H3K9me3, H3K27me3, H4/H2AR3me2s and H3K9ac (Fig. 3). The loss of H3K9me3 and H3K27me3 seems transient, with these modifications being recovered after E12.5, whereas the loss of H4/H2AR3me2s and H3K9ac seems persistent. These dynamic
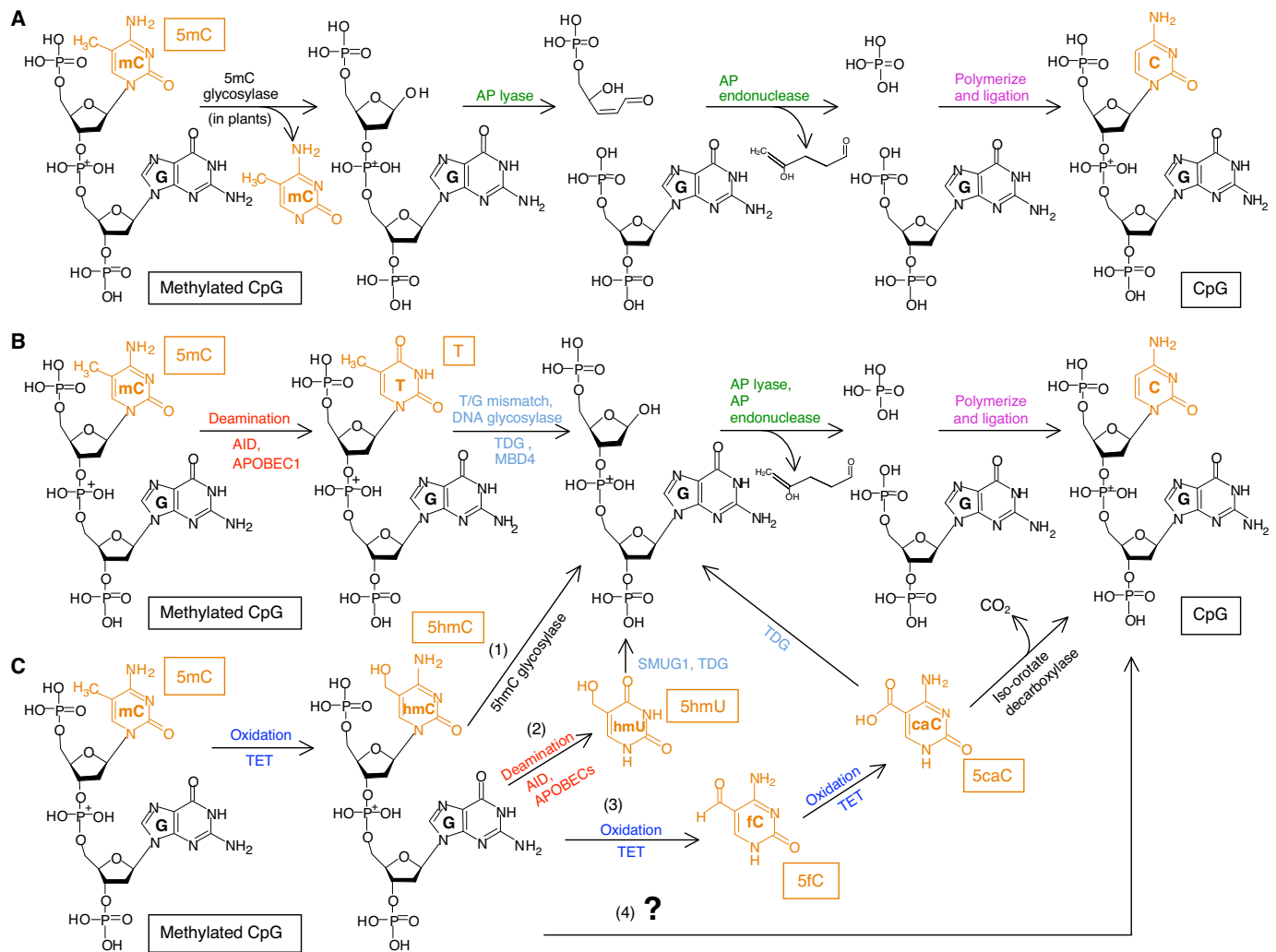

Box 5. AID and APOBECs: cytidine deaminases

AID (activation-induced deaminase) and APOBECs (apolipoprotein B mRNA editing enzyme, catalytic polypeptides) are a group of cytidine deaminases in vertebrates that can introduce mutations in DNA and RNA by deaminating cytidine to uridine. AID is involved in class switch recombination (CSR) and in somatic hypermutation (SHM) of the immunoglobulin (Ig) genes (Muramatsu et al., 2000; Revy et al., 2000). In one model, in activated B-cells, AID deaminates cytosines into uracils on the Ig loci, which creates U:G mismatches, triggering the error-prone DNA repair system (Di Noia and Neuberger, 2007). Consequently, U:G mismatches occurring in the V, D and J genes lead to affinity maturation or gene conversion, whereas U:G mismatches occurring in the switch regions lead to CSR (Di Noia and Neuberger, 2007). Another model posits that AID deaminates unidentified mRNA, leading to the production of a potential endonuclease that cleaves DNA during the immune response (Honjo et al., 2005). In the absence of AID, none of these events occurs (Muramatsu et al., 2000; Revy et al., 2000). APOBEC1 is an RNA deaminase that converts cytidine to uridine. Most typically, it edits apolipoprotein B RNA, generating a truncated apolipoprotein B in a tissue-specific manner (Conticello, 2008). APOBEC3 functions to restrict the activity of viruses and retrotransposons in primates by editing their DNAs (Conticello, 2008). Interestingly, both *Aid* and *Apobec1* are located in close proximity to *Nanog* and *stella/Pgc7* (*Dppa3*, developmental pluripotency-associated 3) on mouse chromosome 6. This may account for the expression of *Aid* and *Apobec1* in pluripotent cell lineages, such as oocytes, embryonic stem cells and primordial germ cells (Morgan et al., 2004).

Fig. 4. Potential chemical pathways for active DNA demethylation. (A) Direct excision of 5mC (orange) by a 5mC glycosylase followed by repair via the base excision repair (BER) pathway (green and pink), as occurs in plants. **(B)** Cytosine deamination by AID/APOBEC1 (red), followed by base excision mismatch repair, involving the TDG/MBD4 (pale blue) and BER pathways. **(C)** Hydroxylation by TET (blue) initiates four potential pathways leading to demethylated cytosine: (1) removal of 5hmC by an unidentified 5hmC glycosylase, followed by BER; (2) deamination of 5hmC by AID or APOBECs creates 5hmU, which is removed by SMUG1 (single-strand selective monofunctional uracil DNA glycosylase) or TDG, followed by BER; (3) further oxidization of 5hmC to 5fC and then to 5caC, which then may be converted to C by a decarboxylase or by TDG followed by BER; and (4) direct conversion of 5hmC to 5mC by an unidentified enzyme (?). 5caC, 5-carboxylcytosine; 5fC, 5-formylcytosine; 5hmC, 5-hydroxymethylcytosine; 5hmU, 5-hydroxymethyluracil;5mC, 5-methylcytosine; AP, apurinic/apyrimidinic; AID, activation-induced deaminase; APOBEC1, apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1; C, cytosine; G, guanine; MBD4, methyl CpG binding domain protein 4; SMUG1, single-strand selective monofunctional uracil DNA glycosylase; T, thymidine; TDG, thymine DNA glycosylase; TET, ten-eleven translocation.

changes are possibly mediated through histone replacement, perhaps by the histone chaperone HIRA (histone cell cycle regulation defective homolog A) or NAP1 (nucleosome assembly protein 1) (see Glossary, Box 2) (Hajkova et al., 2008).

Although the expression of *Aid*, *Apobec1* (apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1), *Mbd4* and *Tdg* (thymine DNA glycosylase) is low in PGCs from E10.5 to E12.5, significant expression of *Tet1* has been found in these cells (Hajkova et al., 2010), indicating that, in PGCs, TET1 may convert 5mC into 5hmC, which could be cleaved by an as yet unidentified 5hmC glycosylase, leading to the activation of BER. However, recently generated *Tet1* knockout mice are viable and fertile, and mating between homozygous mutant males and females produces viable progeny, although with a reduced average litter size (three

to six pups compared with wild-type litters of five to nine pups) (Dawlaty et al., 2011). The role of TET proteins in gametogenesis and fertility thus requires further investigation.

Passive DNA demethylation in PGCs?
The extent of genome-wide DNA demethylation in PGCs is extraordinary and more global compared even with that in pre-implantation embryos, because, in PGCs, genome imprints are erased and the demethylation of transposable elements is more extensive (Hajkova et al., 2002; Lane et al., 2003; Lee et al., 2002; Popp et al., 2010). This suggests the presence of mechanisms unique to PGCs, allowing nearly complete DNA demethylation. Although AID and TET1 are suggested to be part of this process, these molecules are expressed elsewhere (e.g. B-cells and ESCs),

in which genome-wide DNA demethylation is not reported. To understand more fully the mechanism of this extensive genome-wide DNA demethylation, it is important to investigate the events that are unique to PGCs.

Upon PGC specification, *Dnmt3b*, *Dnmt3a* and *Np95* are transcriptionally repressed (Kurimoto et al., 2008; Seki et al., 2005), although PGCs continue to express *Dnmt1*. GLP and G9a, the histone methyltransferases that confer the H3K9me2 mark to chromatin during development and in ESCs (Tachibana et al., 2002; Tachibana et al., 2005), are also repressed in PGCs at around E7.5 and E9.5, respectively (Kurimoto et al., 2008; Seki et al., 2005). PGCs continue to repress these molecules at least until E12.5. Thus, PGCs have little to no DNA methyltransferase and H3K9 di-methylase activity from soon after their specification (~E7.5) to E12.5.

Immunofluorescence studies indicate that, during the migration period, PGCs show reduced genome-wide DNA methylation, and exhibit decreases in H3K9me2 and increases in H3K27me3 in a progressive, cell-by-cell manner. By E9.5, when PGCs emigrate out into the mesentery, nearly all of them bear low H3K9me2 and high H3K27me3 levels (Seki et al., 2005; Seki et al., 2007). Western blot analysis confirms that PGCs at E12.5 have highly reduced H3K9me2 and significantly elevated H3K27me3 levels compared with E6.5 epiblasts and with somatic cells in the gonads at E12.5 (Seki et al., 2005).

These findings, together with the cell cycle dynamics of migrating PGCs, indicate that their genome-wide DNA demethylation might occur partly through a passive mechanism (Fig. 3). The low H3K9me2 state of PGCs may be of relevance to their DNA demethylation, because in *G9a*/*Glp*-knockout ES cells, which show highly reduced H3K9me1/2, some single-copy genes and retrotransposable elements are DNA demethylated even in the presence of the three DNMTs (Dong et al., 2008; Tachibana et al., 2008). The timing of DNA demethylation in PGCs might depend on the target preference of the residual DNMTs. Indeed, as discussed earlier, in pre-implantation embryos, in which genome-wide DNA methylation levels substantially decrease by a presumably passive mechanism, the methylation of DMRs at imprinted genes is maintained by the activity of DNMT1, which is expressed at a very low level (Hirasawa et al., 2008). Moreover, the maintenance of DNA methylation at some sequences, including at retrotransposons, requires cooperation between DNMT1, DNMT3A and DNMT3B (Chen et al., 2003; Liang et al., 2002), indicating that DNA methylation patterns can be altered by the absence of even one of these three enzymes.

The conversion of 5mC into 5hmC and its subsequent passive demethylation may also be a potential DNA demethylation pathway in PGCs. The fact that TET1 binding (and hence the presence of 5hmC) is enriched in the promoters of LINE1 elements but is absent at repetitive elements, such as at IAP and minor satellite repeats in ESCs (Ficz et al., 2011; Williams et al., 2011; Wu et al., 2011), may account for the preferential demethylation at LINE1 elements but the relatively persistent presence of 5mC at IAP and minor satellite repeats in PGCs.

Active DNA demethylation in other contexts

Active DNA demethylation is reported to occur in a highly locus-specific fashion and to control gene expression in various contexts. There is evidence that GADD45A (growth arrest and DNA-damage-inducible 45a), a protein involved in the maintenance of genomic stability, DNA repair and suppression of cell growth, has a role in active DNA demethylation through the NER pathway in cultured fibroblasts (Barreto et al., 2007). It has also been reported that overexpression of AID and MBD4 in zebrafish embryos leads to active DNA demethylation through a combined pathway of 5mC deamination by AID followed by thymine base excision by MBD4, which is promoted by GADD45 (Rai et al., 2008). In somatic cell-ESC fusion experiments, AID has also been shown to facilitate epigenetic reprogramming towards pluripotency, which requires DNA demethylation (Bhutani et al., 2010). RNF4 (RING finger protein 4), a SUMO-dependent ubiquitin E3-ligase implicated in the maintenance of genome stability, has also been shown to have a role in active DNA demethylation both in mouse embryonic development and in cultured cells (Hu et al., 2010): in RNF4-deficient embryonic fibroblasts, DNA methylation at imprinted genes, such as at *Peg1* and *Peg3*, is elevated from ~50% to ~75%, indicating that maintenance of the unmethylated state of the DMRs of the paternal alleles of these genes requires protection (demethylation) from erroneous methylation (Hu et al., 2010). RNF4 interacts with and requires TDG and APE1 for active demethylation, indicating the involvement of the BER pathway in this process (Hu et al., 2010).

In adult neurons, activity induced GADD45B has been shown to demethylate DNA actively at promoters of key genes involved in adult neurogenesis and to induce their expression (Ma et al., 2009). Furthermore, another study shows that TET1 and APOBEC1 are involved in neuronal activity-induced region-specific active DNA demethylation and subsequent gene expression in the dentate gyrus of the adult mouse brain (Guo et al., 2011). This study shows that TET1 promotes DNA demethylation in human cultured cell lines, and this requires the BER pathway. In this system, 12 known human DNA glycosylases have been shown to not act directly on 5hmC. However, AID and APOBECs can efficiently deaminate 5hmC into 5hm uracil (5hmU) (AID cannot deaminate 5mC efficiently), which is then a preferable target for DNA glycosylases such as SMUG1 (single-strand selective monofunctional uracil DNA glycosylase) and TDG in their activation of the BER pathway. AID-mediated 5hmC deamination recapitulated the properties of the AID-mediated cytosine deamination observed in B cells, such as processivity, sequence selectivity, transcription dependence and strand preference. Thus, this study proposes a TET1-induced oxidation-deamination mechanism for active DNA demethylation (Guo et al., 2011).

Gene-knockout studies have revealed that the DNA glycosylase TDG has important functions in mouse embryogenesis (Cortazar et al., 2011; Cortellino et al., 2011). It has been shown to maintain the unmethylated state of CGIs at the promoters of developmentally regulated genes, such as *Hoxa10*, *Hoxd13*, *Sfrp2* (secreted frizzled-related protein 2), *Twist2* (twist homolog 2) and *Rarb* (retinoic acid receptor β) (Cortazar et al., 2011). In wild-type mouse embryonic fibroblasts, the CGI at the promoters of these genes are free of 5mC and are associated with H3K4me2, but in TDG-deficient cells, they are aberrantly methylated and are associated with H3K27me3. On the promoters of wild-type cells, TDG forms a complex with BER pathway components, including XRCC1, APE and PARP1, and with the transcription-activating histone acetyltransferase CBP/p300 and the H3K4-specific methyltransferase MLL1. Interestingly, although TDG also associates with the promoters of such genes in ESCs, the epigenetic aberrations only manifest upon their differentiation, indicating that TDG contributes to the maintenance of active chromatin during cell differentiation, facilitating a proper assembly of the chromatin modifying complex and undergoing BER to counter aberrant de novo methylation (Cortazar et al., 2011). Another study supports

the conclusion of the above-mentioned study and furthermore shows that, in TDG mutants, imprinted genes such as *H19* and *Igf2* show hypermethylation and that the developmentally regulated demethylation of the albumin gene enhancer fails to occur (Cortellino et al., 2011). Moreover, TDG forms a complex with AID and GADD45A, and shows a strong glycosylase activity towards 5hmU (Cortellino et al., 2011). Thus, the authors propose a two-step mechanism for DNA demethylation in mammals, in which 5mC or 5hmC is first deaminated by AID to thymine or 5hmU, respectively, which is then excised and repaired by the TDG-mediated BER pathway.

However, some of these pathways may not have a role in DNA demethylation in pre-implantation embryos and in PGCs. For example, *Mbd4*-deficient mice are fertile, and genome-wide DNA demethylation appears to occur normally in *Mbd4*-deficient zygotes (Millar et al., 2002; Santos and Dean, 2004). One study has shown that GADD45A has no DNA demethylation activity (Jin et al., 2008), and another that *Gadd45a*-deficient mice have neither loci-specific nor global defects in DNA methylation levels (Engel et al., 2009). In addition, the knocking out of *Gadd45b* does not affect the paternal DNA demethylation in zygotes (Okada et al., 2010). Furthermore, *Apobec1* knockout mice are fully fertile (Hirano et al., 1996; Morrison et al., 1996). There remains a possibility that the negative results of these knockout experiments are due to functional redundancy with other proteins. As such, we should await the results of compound mutants, such as *Gadd45a/Gadd45b*-double knockout mice, *Aid/Apobec1*-double knockout mice or *Aid/Tet1*-double knockout mice. The conditional deletion of TDG in PGCs and in oocytes should also provide important new insights into the role of TDG in genome-wide DNA demethylation in pre-implantation embryos and in PGCs.

Conclusion

Despite recent considerable progress, much remains to be learned about the mechanisms and the consequences of the epigenetic reprogramming in pre-implantation embryos and in PGCs. As multiple and compound pathways for active DNA demethylation have been reported, future genetic analyses of candidate components of these DNA-demethylation pathways are required to substantiate their proposed mechanisms. This will require the creation of compound mutants, i.e. double or triple knockouts, as the candidate enzymes for DNA demethylation belong to families with similar activities. In PGCs, in addition to active DNA demethylation, replication-dependent passive DNA demethylation may also be involved, and this possibility should be examined experimentally by overexpressing key repressed genes in PGCs. At the same time, a more comprehensive determination of the mode of DNA demethylation, including the analysis of hemi-methylation states during critical developmental periods is crucially required to obtain new insights into the mechanisms of DNA demethylation. Genome-wide quantification of underlying histone modifications would also provide key information about how epigenetic reprogramming proceeds. The development of new technologies to quantify genome-wide epigenetic modifications from small amounts of starting materials would also help to advance research in this field, as would new procedures to reconstitute epiblast and PGC development from pluripotent stem cells in vitro, in order to provide greater quantities of experimental material for such experiments (Hayashi et al., 2011). Concerted efforts along these lines will help to clarify the mechanisms of epigenetic reprogramming and may lead to the development of a strategy that will ultimately allow us to control the epigenetic state of a cell in vitro.

References
Aapola, U., Kawasaki, K., Scott, H. S., Ollila, J., Vihinen, M., Heino, M., Shintani, A., Minoshima, S., Krohn, K., Antonarakis, S. E. et al. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. *Genomics* **65**, 293-298.

Adenot, P. G., Mercier, Y., Renard, J. P. and Thompson, E. M. (1997). Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. *Development* **124**, 4615-4625.

Barreto, G., Schafer, A., Marhold, J., Stach, D., Swaminathan, S. K., Handa, V., Doderlein, G., Maltry, N., Wu, W., Lyko, F. et al. (2007). Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* **445**, 671-675.

Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K. et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-326.

Bernstein, B. E., Meissner, A. and Lander, E. S. (2007). The mammalian epigenome. *Cell* **128**, 669-681.

Bestor, T., Laudano, A., Mattaliano, R. and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J. Mol. Biol.* **203**, 971-983.

Bestor, T. H. (1992). Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *EMBO J.* **11**, 2611-2617.

Bhutani, N., Brady, J. J., Damian, M., Sacco, A., Corbel, S. Y. and Blau, H. M. (2010). Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463**, 1042-1047.

Bird, A. (2007). Perceptions of epigenetics. *Nature* **447**, 396-398.

Bonasio, R., Tu, S. and Reinberg, D. (2010). Molecular signals of epigenetic states. *Science* **330**, 612-616.

Borgel, J., Guibert, S., Li, Y., Chiba, H., Schubeler, D., Sasaki, H., Forne, T. and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. *Nat. Genet.* **42**, 1093-1100.

Bostick, M., Kim, J. K., Esteve, P. O., Clark, A., Pradhan, S. and Jacobsen, S. E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760-1764.

Bourc'his, D. and Bestor, T. H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431**, 96-99.

Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B. and Bestor, T. H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science* **294**, 2536-2539.

Branco, M. R., Oda, M. and Reik, W. (2008). Safeguarding parental identity: Dnmt1 maintains imprints during epigenetic reprogramming in early embryogenesis. *Genes Dev.* **22**, 1567-1571.

Bransteitter, R., Pham, P., Scharff, M. D. and Goodman, M. F. (2003). Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl. Acad. Sci. USA* **100**, 4102-4107.

Brockdorff, N. (2011). Chromosome silencing mechanisms in X chromosome inactivation: unknown unknowns. *Development* **138**, (in press)

Burdzy, A., Noyes, K. T., Valinluck, V. and Sowers, L. C. (2002). Synthesis of stable-isotope enriched 5-methylpyrimidines and their use as probes of base reactivity in DNA. *Nucleic Acids Res.* **30**, 4068-4074.

Cedar, H. and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nat. Rev. Genet.* **10**, 295-304.

Chen, T., Ueda, Y., Dodge, J. E., Wang, Z. and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell. Biol.* **23**, 5594-5605.

Cimmino, L., Abdel-Wahab, O., Levine, R. L. and Aifantis, I. (2011). TET family proteins and their role in stem cell differentiation and transformation. *Cell Stem Cell* **9**, 193-204.

Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., Pradhan, S., Nelson, S. F., Pellegrini, M. and Jacobsen, S. E. (2008). Shotgun

28   REVIEW

Development 139 (1)

bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* **452**, 215-219.

Conticello, S. G. (2008). The AID/APOBEC family of nucleic acid mutators. *Genome Biol.* **9**, 229.

Cortazar, D., Kunz, C., Selfridge, J., Lettieri, T., Saito, Y., MacDougall, E., Wirz, A., Schuermann, D., Jacobs, A. L., Siegrist, F. et al. (2011). Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* **470**, 419-423.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D. et al. (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146**, 67-79.

Dawlaty, M. M., Ganz, K., Powell, B. E., Hu, Y. C., Markoulaki, S., Cheng, A. W., Gao, Q., Kim, J., Choi, S. W., Page, D. C. 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.

Di Noia, J. M. and Neuberger, M. S. (2007). Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.* **76**, 1-22.

Dong, K. B., Maksakova, I. A., Mohn, F., Leung, D., Appanah, R., Lee, S., Yang, H. W., Lam, L. L., Mager, D. L., Schubeler, D. et al. (2008). DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J.* **27**, 2691-2701.

Down, T. A., Rakyan, V. K., Turner, D. J., Flicek, P., Li, H., Kulesha, E., Graf, S., Johnson, N., Herrero, J., Tomazou, E. M. et al. (2008). A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat. Biotechnol.* **26**, 779-785.

Duncan, B. K. and Miller, J. H. (1980). Mutagenic deamination of cytosine residues in DNA. *Nature* **287**, 560-561.

Engel, N., Tront, J. S., Erinle, T., Nguyen, N., Latham, K. E., Sapienza, C., Hoffman, B. and Liebermann, D. A. (2009). Conserved DNA methylation in Gadd45a(–/–) mice. *Epigenetics* **4**, 98-99.

Esteve, P. O., Chin, H. G., Smallwood, A., Feehery, G. R., Gangisetty, O., Karpf, A. R., Carey, M. F. and Pradhan, S. (2006). Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev.* **20**, 3089-3103.

Farthing, C. R., Ficz, G., Ng, R. K., Chan, C. F., Andrews, S., Dean, W., Hemberger, M. and Reik, W. (2008). Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS Genet.* **4**, e1000116.

Ficz, G., Branco, M. R., Seisenberger, S., Santos, F., Krueger, F., Hore, T. A., Marques, C. J., Andrews, S. and Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473**, 398-402.

Figueroa, M. E., Abdel-Wahab, O., Lu, C., Ward, P. S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H. F. et al. (2010). Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* **18**, 553-567.

Frauer, C., Hoffmann, T., Bultmann, S., Casa, V., Cardoso, M. C., Antes, I. and Leonhardt, H. (2011). Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS ONE* **6**, e21306.

Fujimori, A., Matsuda, Y., Takemoto, Y., Hashimoto, Y., Kubo, E., Araki, R., Fukumura, R., Mita, K., Tatsumi, K. and Muto, M. (1998). Cloning and mapping of Np95 gene which encodes a novel nuclear protein associated with cell proliferation. *Mamm. Genome* **9**, 1032-1035.

Fuks, F., Hurd, P. J., Deplus, R. and Kouzarides, T. (2003). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res.* **31**, 2305-2312.

Gaudet, F., Rideout, W. M., 3rd, Meissner, A., Dausman, J., Leonhardt, H. and Jaenisch, R. (2004). Dnmt1 expression in pre- and postimplantation embryogenesis and the maintenance of IAP silencing. *Mol. Cell. Biol.* **24**, 1640-1648.

Gehring, M., Bubb, K. L. and Henikoff, S. (2009). Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science* **324**, 1447-1451.

Ginsburg, M., Snow, M. H. and McLaren, A. (1990). Primordial germ cells in the mouse embryo during gastrulation. *Development* **110**, 521-528.

Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C. L., Zhang, X., Golic, K. G., Jacobsen, S. E. and Bestor, T. H. (2006). Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science* **311**, 395-398.

Gu, T. P., Guo, F., Yang, H., Wu, H. P., Xu, G. F., Liu, W., Xie, Z. G., Shi, L., He, X., Jin, S. G. et al. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606-610.

Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. and Young, R. A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77-88.

Guo, J. U., Su, Y., Zhong, C., Ming, G. L. and Song, H. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145**, 423-434.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J. and Surani, M. A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117**, 15-23.

Hajkova, P., Ancelin, K., Waldmann, T., Lacoste, N., Lange, U. C., Cesari, F., Lee, C., Almouzni, G., Schneider, R. and Surani, M. A. (2008). Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* **452**, 877-881.

Hajkova, P., Jeffries, S. J., Lee, C., Miller, N., Jackson, S. P. and Surani, M. A. (2010). Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* **329**, 78-82.

Hanna, J. H., Saha, K. and Jaenisch, R. (2010). Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. *Cell* **143**, 508-525.

Hashimshony, T., Zhang, J., Keshet, I., Bustin, M. and Cedar, H. (2003). The role of DNA methylation in setting up chromatin structure during development. *Nat. Genet.* **34**, 187-192.

Hata, K., Okano, M., Lei, H. and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129**, 1983-1993.

Hayashi, K. and Surani, M. A. (2009). Resetting the epigenome beyond pluripotency in the germline. *Cell Stem Cell* **4**, 493-498.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519-532.

Hayatsu, H. and Shiragami, M. (1979). Reaction of bisulfite with the 5-hydroxymethyl group in pyrimidines and in phage DNAs. *Biochemistry* **18**, 632-637.

He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L. et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303-1307.

Hemberger, M., Dean, W. and Reik, W. (2009). Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nat. Rev. Mol. Cell. Biol.* **10**, 526-537.

Hirano, K., Young, S. G., Farese, R. V., Jr, Ng, J., Sande, E., Warburton, C., Powell-Braxton, L. M. and Davidson, N. O. (1996). Targeted disruption of the mouse apobec-1 gene abolishes apolipoprotein B mRNA editing and eliminates apolipoprotein B48. *J. Biol. Chem.* **271**, 9887-9890.

Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R. and 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.

Honjo, T., Nagaoka, H., Shinkura, R. and Muramatsu, M. (2005). AID to overcome the limitations of genomic information. *Nat. Immunol.* **6**, 655-661.

Howlett, S. K. and Reik, W. (1991). Methylation levels of maternal and paternal genomes during preimplantation development. *Development* **113**, 119-127.

Hsieh, T. F., Ibarra, C. A., Silva, P., Zemach, A., Eshed-Williams, L., Fischer, R. L. and Zilberman, D. (2009). Genome-wide demethylation of Arabidopsis endosperm. *Science* **324**, 1451-1454.

Hu, X. V., Rodrigues, T. M., Tao, H., Baker, R. K., Miraglia, L., Orth, A. P., Lyons, G. E., Schultz, P. G. and Wu, X. (2010). Identification of RING finger protein 4 (RNF4) as a modulator of DNA demethylation through a functional genomics screen. *Proc. Natl. Acad. Sci. USA* **107**, 15087-15092.

Iqbal, K., Jin, S. G., Pfeifer, G. P. and Szabo, P. E. (2011). Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl. Acad. Sci. USA* **108**, 3642-3647.

Irizarry, R. A., Ladd-Acosta, C., Carvalho, B., Wu, H., Brandenburg, S. A., Jeddeloh, J. A., Wen, B. and Feinberg, A. P. (2008). Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res.* **18**, 780-790.

Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C. and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129-1133.

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., He, C. and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300-1303.

Jin, S. G., Guo, C. and Pfeifer, G. P. (2008). GADD45A does not promote DNA demethylation. *PLoS Genet.* **4**, e1000013.

Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J. and Wolffe, A. P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* **19**, 187-191.

Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H. and Razin, A. (1992). Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev.* **6**, 705-714.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E. and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **429**, 900-903.

Kim, J. K., Esteve, P. O., Jacobsen, S. E. and Pradhan, S. (2009). UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells. *Nucleic Acids Res.* **37**, 493-505.

Ko, M., Huang, Y., Jankowska, A. M., Pape, U. J., Tahiliani, M., Bandukwala, H. S., An, J., Lamperti, E. D., Koh, K. P., Ganetzky, R. et al. (2011). Impaired

hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839-843.

Kobayashi, H., Suda, C., Abe, T., Kohara, Y., Ikemura, T. and Sasaki, H. (2006). Bisulfite sequencing and dinucleotide content analysis of 15 imprinted mouse differentially methylated regions (DMRs): paternally methylated DMRs contain less CpGs than maternally methylated DMRs. *Cytogenet. Genome Res.* **113**, 130-137.

Koh, K. P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C. A., Mostoslavsky, G. et al. (2011). Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**, 200-213.

Kondrashov, A. S. (2003). Direct estimates of human per nucleotide mutation rates at 20 loci causing Mendelian diseases. *Hum. Mutat.* **21**, 12-27.

Kriaucionis, S. and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929-930.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K. and Saitou, M. (2008). Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. *Genes Dev.* **22**, 1617-1635.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. et al. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J. and Reik, W. (2003). Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* **35**, 88-93.

Langemeijer, S. M., Kuiper, R. P., Berends, M., Knops, R., Aslanyan, M. G., Massop, M., Stevens-Linders, E., van Hoogen, P., van Kessel, A. G., Raymakers, R. A. et al. (2009). Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat. Genet.* **41**, 838-842.

Larijani, M., Frieder, D., Sonbuchner, T. M., Bransteitter, R., Goodman, M. F., Bouhassira, E. E., Scharff, M. D. and Martin, A. (2005). Methylation protects cytidines from AID-mediated deamination. *Mol. Immunol.* **42**, 599-604.

Lawson, K. A., Dunn, N. R., Roelen, B. A., Zeinstra, L. M., Davis, A. M., Wright, C. V., Korving, J. P. and Hogan, B. L. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* **13**, 424-436.

Lee, J., Inoue, K., Ono, R., Ogonuki, N., Kohda, T., Kaneko-Ishino, T., Ogura, A. and Ishino, F. (2002). Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. *Development* **129**, 1807-1817.

Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. and Peters, A. H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* **13**, 1192-1200.

Lei, H., Oh, S. P., Okano, M., Juttermann, R., Goss, K. A., Jaenisch, R. and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122**, 3195-3205.

Li, E., Bestor, T. H. and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915-926.

Li, H., Rauch, T., Chen, Z. X., Szabo, P. E., Riggs, A. D. and Pfeifer, G. P. (2006). The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. *J. Biol. Chem.* **281**, 19489-19500.

Li, L., Zheng, P. and Dean, J. (2010). Maternal control of early mouse development. *Development* **137**, 859-870.

Liang, G., Chan, M. F., Tomigahara, Y., Tsai, Y. C., Gonzales, F. A., Li, E., Laird, P. W. and Jones, P. A. (2002). Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Mol. Cell. Biol.* **22**, 480-491.

Lippman, Z., Gendrel, A. V., Black, M., Vaughn, M. W., Dedhia, N., McCombie, W. R., Lavine, K., Mittal, V., May, B., Kasschau, K. D. et al. (2004). Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430**, 471-476.

Lister, R., O'Malley, R. C., Tonti-Filippini, J., Gregory, B. D., Berry, C. C., Millar, A. H. and Ecker, J. R. (2008). Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* **133**, 523-336.

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q. M. et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315-322.

Ma, D. K., Jang, M. H., Guo, J. U., Kitabatake, Y., Chang, M. L., Pow-Anpongkul, N., Flavell, R. A., Lu, B., Ming, G. L. and Song, H. (2009). Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* **323**, 1074-1077.

Mayer, W., Niveleau, A., Walter, J., Fundele, R. and Haaf, T. (2000). Demethylation of the zygotic paternal genome. *Nature* **403**, 501-502.

Meilinger, D., Fellinger, K., Bultmann, S., Rothbauer, U., Bonapace, I. M., Klinkert, W. E., Spada, F. and Leonhardt, H. (2009). Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic

silencing of the viral CMV promoter in embryonic stem cells. *EMBO Rep.* **10**, 1259-1264.

Meissner, A., Gnirke, A., Bell, G. W., Ramsahoye, B., Lander, E. S. and Jaenisch, R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868-5877.

Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P. et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-560.

Millar, C. B., Guy, J., Sansom, O. J., Selfridge, J., MacDougall, E., Hendrich, B., Keightley, P. D., Bishop, S. M., Clarke, A. R. and Bird, A. (2002). Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science* **297**, 403-405.

Mohn, F., Weber, M., Rebhan, M., Roloff, T. C., Richter, J., Stadler, M. B., Bibel, M. and Schubeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* **30**, 755-766.

Monk, M., Boubelik, M. and Lehnert, S. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**, 371-382.

Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., Figueroa, M. E., Vasanthakumar, A., Patel, J., Zhao, X. et al. (2011). Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**, 11-24.

Morgan, H. D., Dean, W., Coker, H. A., Reik, W. and Petersen-Mahrt, S. K. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J. Biol. Chem.* **279**, 52353-52360.

Morrison, J. R., Paszty, C., Stevens, M. E., Hughes, S. D., Forte, T., Scott, J. and Rubin, E. M. (1996). Apolipoprotein B RNA editing enzyme-deficient mice are viable despite alterations in lipoprotein metabolism. *Proc. Natl. Acad. Sci. USA* **93**, 7154-7159.

Muramatsu, M., Sankaranand, V. S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N. O. and Honjo, T. (1999). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J. Biol. Chem.* **274**, 18470-18476.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y. and Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553-563.

Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M. et al. (2007). PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat. Cell Biol.* **9**, 64-71.

Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N. and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386-389.

Navaratnam, N., Morrison, J. R., Bhattacharya, S., Patel, D., Funahashi, T., Giannoni, F., Teng, B. B., Davidson, N. O. and Scott, J. (1993). The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *J. Biol. Chem.* **268**, 20709-20712.

Oda, M., Yamagiwa, A., Yamamoto, S., Nakayama, T., Tsumura, A., Sasaki, H., Nakao, K., Li, E. and Okano, M. (2006). DNA methylation regulates long-range gene silencing of an X-linked homeobox gene cluster in a lineage-specific manner. *Genes Dev.* **20**, 3382-3394.

Oda, M., Glass, J. L., Thompson, R. F., Mo, Y., Olivier, E. N., Figueroa, M. E., Selzer, R. R., Richmond, T. A., Zhang, X., Dannenberg, L. et al. (2009). High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers. *Nucleic Acids Res.* **37**, 3829-3839.

Ohinata, Y., Payer, B., O'Carroll, D., Ancelin, K., Ono, Y., Sano, M., Barton, S. C., Obukhanych, T., Nussenzweig, M., Tarakhovsky, A. et al. (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* **436**, 207-213.

Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T. and Saitou, M. (2009). A signaling principle for the specification of the germ cell lineage in mice. *Cell* **137**, 571-584.

Okada, Y., Yamagata, K., Hong, K., Wakayama, T. and Zhang, Y. (2010). A role for the elongator complex in zygotic paternal genome demethylation. *Nature* **463**, 554-558.

Okano, M., Xie, S. and Li, E. (1998a). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.* **19**, 219-220.

Okano, M., Xie, S. and Li, E. (1998b). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res.* **26**, 2536-2540.

Okano, M., Bell, D. W., Haber, D. A. and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257.

Ooi, S. K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S. P., Allis, C. D. et al. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714-717.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W. and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10**, 475-478.

Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G. A., Stewart, R. and Thomson, J. A. (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* **1**, 299-312.

Pastor, W. A., Pape, U. J., Huang, Y., Henderson, H. R., Lister, R., Ko, M., McLoughlin, E. M., Brudno, Y., Mahapatra, S., Kapranov, P. et al. (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* **473**, 394-397.

Payer, B. and Lee, J. T. (2008). X chromosome dosage compensation: how mammals keep the balance. *Annu. Rev. Genet.* **42**, 733-772.

Payer, B., Saitou, M., Barton, S. C., Thresher, R., Dixon, J. P., Zahn, D., Colledge, W. H., Carlton, M. B., Nakano, T. and Surani, M. A. (2003). Stella is a maternal effect gene required for normal early development in mice. *Curr. Biol.* **13**, 2110-2117.

Pfaffeneder, T., Hackner, B., Truss, M., Munzel, M., Muller, M., Deiml, C. A., Hagemeier, C. and Carell, T. (2011). The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angew. Chem. Int. Ed. Engl.* **50**, 7008-7012.

Polanski, Z., Motosugi, N., Tsurumi, C., Hiiragi, T. and Hoffmann, S. (2008). Hypomethylation of paternal DNA in the late mouse zygote is not essential for development. *Int. J. Dev. Biol.* **52**, 295-298.

Poole, A., Penny, D. and Sjoberg, B. M. (2001). Confounded cytosine! Tinkering and the evolution of DNA. *Nat. Rev. Mol. Cell. Biol.* **2**, 147-151.

Popp, C., Dean, W., Feng, S., Cokus, S. J., Andrews, S., Pellegrini, M., Jacobsen, S. E. and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**, 1101-1105.

Probst, A. V. and Almouzni, G. (2011). Heterochromatin establishment in the context of genome-wide epigenetic reprogramming. *Trends Genet.* **27**, 177-185.

Quivoron, C., Couronne, L., Della Valle, V., Lopez, C. K., Plo, I., Wagner-Ballon, O., Do Cruzeiro, M., Delhommeau, F., Arnulf, B., Stern, M. H. et al. (2011). TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell* **20**, 25-38.

Rai, K., Huggins, I. J., James, S. R., Karpf, A. R., Jones, D. A. and Cairns, B. R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* **135**, 1201-1212.

Ramsahoye, B. H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A. P. and 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.

Reik, W. and Walter, J. (2001). Evolution of imprinting mechanisms: the battle of the sexes begins in the zygote. *Nat. Genet.* **27**, 255-256.

Reik, W., Dean, W. and Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science* **293**, 1089-1093.

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Labelouse, R., Gennery, A. et al. (2000). Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* **102**, 565-575.

Rollins, R. A., Haghighi, F., Edwards, J. R., Das, R., Zhang, M. Q., Ju, J. and Bestor, T. H. (2006). Large-scale structure of genomic methylation patterns. *Genome Res.* **16**, 157-163.

Rossant, J. and Tam, P. P. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* **136**, 701-713.

Rottach, A., Frauer, C., Pichler, G., Bonapace, I. M., Spada, F. and Leonhardt, H. (2010). The multi-domain protein Np95 connects DNA methylation and histone modification. *Nucleic Acids Res.* **38**, 1796-1804.

Rougier, N., Bourc'his, D., Gomes, D. M., Niveleau, A., Plachot, M., Paldi, A. and Viegas-Pequignot, E. (1998). Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev.* **12**, 2108-2113.

Saitou, M., Barton, S. C. and Surani, M. A. (2002). A molecular programme for the specification of germ cell fate in mice. *Nature* **418**, 293-300.

Santos, F. and Dean, W. (2004). Epigenetic reprogramming during early development in mammals. *Reproduction* **127**, 643-651.

Santos, F., Hendrich, B., Reik, W. and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev. Biol.* **241**, 172-182.

Saxonov, S., Berg, P. and Brutlag, D. L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl. Acad. Sci. USA* **103**, 1412-1417.

Seki, Y., Hayashi, K., Itoh, K., Mizugaki, M., Saitou, M. and 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.

Seki, Y., Yamaji, M., Yabuta, Y., Sano, M., Shigeta, M., Matsui, Y., Saga, Y., Tachibana, M., Shinkai, Y. and Saitou, M. (2007). Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. *Development* **134**, 2627-2638.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T. A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K. et al. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908-912.

Smallwood, A., Esteve, P. O., Pradhan, S. and Carey, M. (2007). Functional cooperation between HP1 and DNMT1 mediates gene silencing. *Genes Dev.* **21**, 1169-1178.

Smallwood, S. A., Tomizawa, S., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., Sato, S., Hata, K., Andrews, S. R. and Kelsey, G. (2011). Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat. Genet.* **43**, 811-814.

Sugimoto, M. and Abe, K. (2007). X chromosome reactivation initiates in nascent primordial germ cells in mice. *PLoS Genet.* **3**, e116.

Surani, M. A., Hayashi, K. and Hajkova, P. (2007). Genetic and epigenetic regulators of pluripotency. *Cell* **128**, 747-762.

Sutherland, E., Coe, L. and Raleigh, E. A. (1992). McrBC: a multisubunit GTP-dependent restriction endonuclease. *J. Mol. Biol.* **225**, 327-348.

Suzuki, M. M. and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* **9**, 465-476.

Szwagierczak, A., Bultmann, S., Schmidt, C. S., Spada, F. and Leonhardt, H. (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res.* **38**, e181.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H. et al. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* **16**, 1779-1791.

Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T. and Shinkai, Y. (2005). Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev.* **19**, 815-826.

Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H. and Shinkai, Y. (2008). G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO J.* **27**, 2681-2690.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L. et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-935.

Tam, P. P. and Snow, M. H. (1981). Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. *J. Embryol. Exp. Morphol.* **64**, 133-147.

Tam, P. P., Zhou, S. X. and Tan, S. S. (1994). X-chromosome activity of the mouse primordial germ cells revealed by the expression of an X-linked lacZ transgene. *Development* **120**, 2925-2932.

Teng, B., Burant, C. F. and Davidson, N. O. (1993). Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science* **260**, 1816-1819.

Thomson, J. P., Skene, P. J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A. R., Deaton, A., Andrews, R., James, K. D. et al. (2010). CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* **464**, 1082-1086.

Tomizawa, S., Kobayashi, H., Watanabe, T., Andrews, S., Hata, K., Kelsey, G. and 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.

Tompa, R., McCallum, C. M., Delrow, J., Henikoff, J. G., van Steensel, B. and Henikoff, S. (2002). Genome-wide profiling of DNA methylation reveals transposon targets of CHROMOMETHYLASE3. *Curr. Biol.* **12**, 65-68.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H. R. 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.

Ueda, Y., Okano, M., Williams, C., Chen, T., Georgopoulos, K. and Li, E. (2006). Roles for Dnmt3b in mammalian development: a mouse model for the ICF syndrome. *Development* **133**, 1183-1192.

Valinluck, V. and Sowers, L. C. (2007). Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res.* **67**, 946-950.

Valinluck, V., Tsai, H. H., Rogstad, D. K., Burdzy, A., Bird, A. and Sowers, L. C. (2004). Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res.* **32**, 4100-4108.

Vincent, S. D., Dunn, N. R., Sciammas, R., Shapiro-Shalef, M., Davis, M. M., Calame, K., Bikoff, E. K. and Robertson, E. J. (2005). The zinc finger

transcriptional repressor Blimp1/Prdm1 is dispensable for early axis formation but is required for specification of primordial germ cells in the mouse. *Development* **132**, 1315-1325.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J. M. et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**, 871-874.

Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P. et al. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520-562.

Weber, M., Davies, J. J., Wittig, D., Oakeley, E. J., Haase, M., Lam, W. L. and Schubeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **37**, 853-862.

Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Paabo, S., Rebhan, M. and Schubeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **39**, 457-466.

Webster, K. E., O'Bryan, M. K., Fletcher, S., Crewther, P. E., Aapola, U., Craig, J., Harrison, D. K., Aung, H., Phutikanit, N., Lyle, R. et al. (2005). Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. *Proc. Natl. Acad. Sci. USA* **102**, 4068-4073.

Whitelaw, N. C. and Whitelaw, E. (2008). Transgenerational epigenetic inheritance in health and disease. *Curr. Opin. Genet. Dev.* **18**, 273-279.

Williams, K., Christensen, J., Pedersen, M. T., Johansen, J. V., Cloos, P. A., Rappsilber, J. and Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343-348.

Wossidlo, M., Arand, J., Sebastiano, V., Lepikhov, K., Boiani, M., Reinhardt, R., Scholer, H. and Walter, J. (2010). Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. *EMBO J.* **29**, 1877-1888.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C. J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W. and Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun.* **2**, 241.

Wu, H., D'Alessio, A. C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Eve Sun, Y. and Zhang, Y. (2011). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473**, 389-393.

Wyatt, G. R. and Cohen, S. S. (1953). The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. *Biochem. J.* **55**, 774-782.

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A. J., Zheng, L., Zhang, H., Huang, S. et al. (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 Hydroxylase in mouse embryonic stem cells. *Mol. Cell* **42**, 451-464.

Yamaji, M., Seki, Y., Kurimoto, K., Yabuta, Y., Yuasa, M., Shigeta, M., Yamanaka, K., Ohinata, Y. and Saitou, M. (2008). Critical function of Prdm14 for the establishment of the germ cell lineage in mice. *Nat. Genet.* **40**, 1016-1022.

Yamanaka, S. and Blau, H. M. (2010). Nuclear reprogramming to a pluripotent state by three approaches. *Nature* **465**, 704-712.

Zhao, X. D., Han, X., Chew, J. L., Liu, J., Chiu, K. P., Choo, A., Orlov, Y. L., Sung, W. K., Shahab, A., Kuznetsov, V. A. et al. (2007). Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Cell Stem Cell* **1**, 286-298.

Zhu, J. K. (2009). Active DNA demethylation mediated by DNA glycosylases. *Annu. Rev. Genet.* **43**, 143-166.

Ziegler-Birling, C., Helmrich, A., Tora, L. and Torres-Padilla, M. E. (2009). Distribution of p53 binding protein 1 (53BP1) and phosphorylated H2A.X during mouse preimplantation development in the absence of DNA damage. *Int. J. Dev. Biol.* **53**, 1003-1011.

Zuo, S., Boorstein, R. J. and Teebor, G. W. (1995). Oxidative damage to 5-methylcytosine in DNA. *Nucleic Acids Res.* **23**, 3239-3243.

Zvetkova, I., Apedaile, A., Ramsahoye, B., Mermoud, J. E., Crompton, L. A., John, R., Feil, R. and Brockdorff, N. (2005). Global hypomethylation of the genome in XX embryonic stem cells. *Nat. Genet.* **37**, 1274-1279.