

## REVIEW

# Tet family of 5-methylcytosine dioxygenases in mammalian development

Hongbo Zhao and Taiping Chen

Methylation of cytosines is a major epigenetic modification in mammalian genomes. The levels and patterns of DNA methylation are the results of the opposing actions of methylating and demethylating machineries. Over the past two decades, great progress has been made in elucidating the methylating machinery including the identification and functional characterization of the DNA methyltransferases (Dnmts). However, the mechanisms of demethylation and the major players involved had been elusive. A major breakthrough came in 2009, when the ten-eleven translocation (Tet) family of proteins was discovered as 5-methylcytosine (5mC) dioxygenases that convert 5mC to 5-hydroxymethylcytosine (5hmC). Studies in the past several years have established that 5hmC serves as an intermediate in DNA demethylation and that Tet proteins have important roles in epigenetic reprogramming in early embryos and primordial germ cells. In this review, we discuss recent advances in this exciting field, focusing on the role of Tet proteins in mammalian development.

*Journal of Human Genetics* (2013) 58, 421–427; doi:10.1038/jhg.2013.63; published online 30 May 2013

**Keywords:** DNA methylation; 5-hydroxymethylcytosine; 5-methylcytosine; Tet1; Tet2; Tet3

## INTRODUCTION

Methylation at the 5-position of cytosine (5-methylcytosine, 5mC), which occurs predominantly in the context of cytosine-phosphate-guanine (CpG) dinucleotides, is a common modification present in mammalian genomes. DNA methylation is essential for mammalian development and has crucial roles in a variety of biological processes including regulation of gene expression, genomic imprinting, X chromosome inactivation and retrotransposon silencing.<sup>1,2</sup> Aberrant changes of genomic DNA methylation patterns and genetic mutations of components of the DNA methylation machinery are linked to numerous human diseases including developmental syndromes, neurological diseases, immunological disorders and various types of cancer.<sup>3–6</sup>

In mice, DNA methylation patterns are established and maintained by three active DNA methyltransferases (Dnmts)—Dnmt1, Dnmt3a and Dnmt3b. Dnmt3a and Dnmt3b function primarily as *de novo* methyltransferases that set up DNA methylation patterns during early embryogenesis, whereas Dnmt1 is the major maintenance enzyme that copies the CpG methylation pattern from the parental strand onto the daughter strand during DNA replication.<sup>7</sup> Genetic studies revealed that Dnmt1 and Dnmt3b are essential for embryogenesis and Dnmt3a is required for postnatal survival.<sup>8–10</sup> Dnmt3a also cooperates with its cofactor Dnmt3L, a Dnmt3-like protein with no enzymatic activity, in mediating DNA methylation in developing germ cells, including the establishment of methylation marks at imprinting control regions (ICRs).<sup>11–13</sup>

DNA methylation is considered to be a relatively stable modification. However, waves of global demethylation occur in two developmental stages—preimplantation embryos and developing primordial germ cells (PGCs)—through both DNA replication-independent ‘active’ and DNA replication-dependent ‘passive’ processes.<sup>2,14</sup> Progress in understanding the mechanisms of demethylation and the major players involved had been slow until the recent discovery that 5mC can be converted to 5-hydroxymethylcytosine (5hmC) by the ten-eleven translocation (Tet) family of dioxygenases.<sup>15,16</sup> Studies in the past several years have revealed that 5hmC is an intermediate in the process of demethylation and distinct Tet proteins appear to be involved in methylation erasure in the zygote and PGCs. Evidence has also emerged for the involvement of Tet-mediated 5mC oxidation in other biological processes.

In this review, we discuss recent progress in our understanding of the biological functions of 5hmC and Tet proteins with an emphasis on mammalian development.

## TET PROTEINS AS 5mC DIOXYGENASES

Although 5hmC was identified in mammalian DNA in 1972,<sup>17</sup> its significance and biological function had not been explored, largely because it is present in relatively low levels in most cell types. In 2009, two research groups reported abundant 5hmC in mouse Purkinje neurons and embryonic stem (ES) cells.<sup>15,18</sup> More importantly, Tahiliani *et al.*<sup>15</sup> identified Tet1, Tet2 and Tet3 as mammalian homologs of the trypanosome proteins JBP1 and JBP2, which

Department of Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Smithville, TX, USA

Correspondence: Dr T Chen, Department of Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Science Park—Research Division, 1808 Park Road 1C, PO Box 389, Smithville, TX 78957, USA.

E-mail: tchen2@mdanderson.org

Received 23 April 2013; revised 3 May 2013; accepted 11 May 2013; published online 30 May 2013

oxidize the 5-methyl group of thymine, and experimentally demonstrated that Tet1 has the capacity to catalyze the conversion of 5mC to 5hmC. Shortly afterward, Tet2 and Tet3 were also shown to have 5mC hydroxylase activity.<sup>16,19</sup> Subsequent studies revealed that Tet proteins can further catalyze the oxidation of 5hmC to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), two less abundant bases.<sup>20,21</sup>

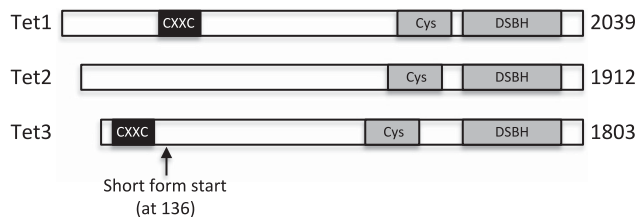
The Tet proteins belong to the 2-oxoglutarate (2OG)- and Fe(II)-dependent dioxygenase (2OGFeDO) superfamily.<sup>15,22</sup> All three Tet proteins contain a C-terminal catalytic domain, which consists of a cysteine-rich region and a double-stranded  $\beta$ -helix fold characteristic of the 2OGFeDO superfamily. Tet1 and Tet3, but not Tet2, also contain an N-terminal CXXC zinc finger domain, a DNA-binding motif (Figure 1).

The CXXC domain, present in multiple chromatin-interacting proteins such as CFP1, MBD1, MLL and Dnmt1, has been shown to selectively bind unmethylated CpG dinucleotides.<sup>23</sup> Sequence alignment revealed that the CXXC domains of Tet1 and Tet3 lack a KFGG motif that is present in many other CXXC domains.<sup>23,24</sup> The DNA-binding property of the Tet1 CXXC domain is controversial: one report showed no specific DNA-binding activity,<sup>24</sup> whereas another report showed binding to unmodified, 5mC-modified and 5hmC-modified CpG-rich DNA.<sup>25</sup> The discrepancy could be because of the different DNA-binding assays and substrates used in these studies. Recently, it was shown that the CXXC domains of *Xenopus* and human TET3 bind only unmodified cytosines, regardless of the sequence contexts (with a slight preference for CpG content), and that in *Xenopus*, the CXXC domain of Tet3 works cooperatively with the catalytic domain in targeting Tet3 to its binding sites during development.<sup>26</sup> Notably, mouse Tet3 has a shorter isoform that lacks the N-terminal 135 amino acids including the CXXC domain (GenBank accession no. NM\_183138). It would be interesting to determine the expression patterns and functional differences of the two isoforms.

Tet2 lost its CXXC domain during evolution due to a chromosomal inversion event that split the ancestral *Tet2* gene into two distinct genes, *Idax*, which encodes a protein containing the ancestral CXXC domain of Tet2, and *Tet2*, which encodes the current Tet2 protein. Interestingly, the two proteins physically interact.<sup>27</sup> IDAX, which preferentially binds DNA sequences containing unmethylated CpG via its CXXC domain, could have a role in recruiting Tet2 to its genomic targets. Strikingly, IDAX induces Tet2 degradation by caspase-mediated cleavage.<sup>27</sup>

### 5hmC, 5fC AND 5caC AS INTERMEDIATES OF DNA DEMETHYLATION PATHWAYS

The finding that 5mC can be converted to 5hmC by Tet proteins immediately raised the possibility that this conversion could be



**Figure 1** Schematic diagrams of mouse Tet proteins. There are three Tet proteins in mice: Tet1, Tet2 and Tet3. They all have a C-terminal catalytic domain, consisting of a cysteine-rich region and the double-stranded  $\beta$ -helix (DSBH) fold characteristic of the 2-oxoglutarate- and Fe(II)-dependent dioxygenase family. Tet1 and Tet3, but not Tet2, have an N-terminal CXXC zinc finger domain, a DNA-binding domain.

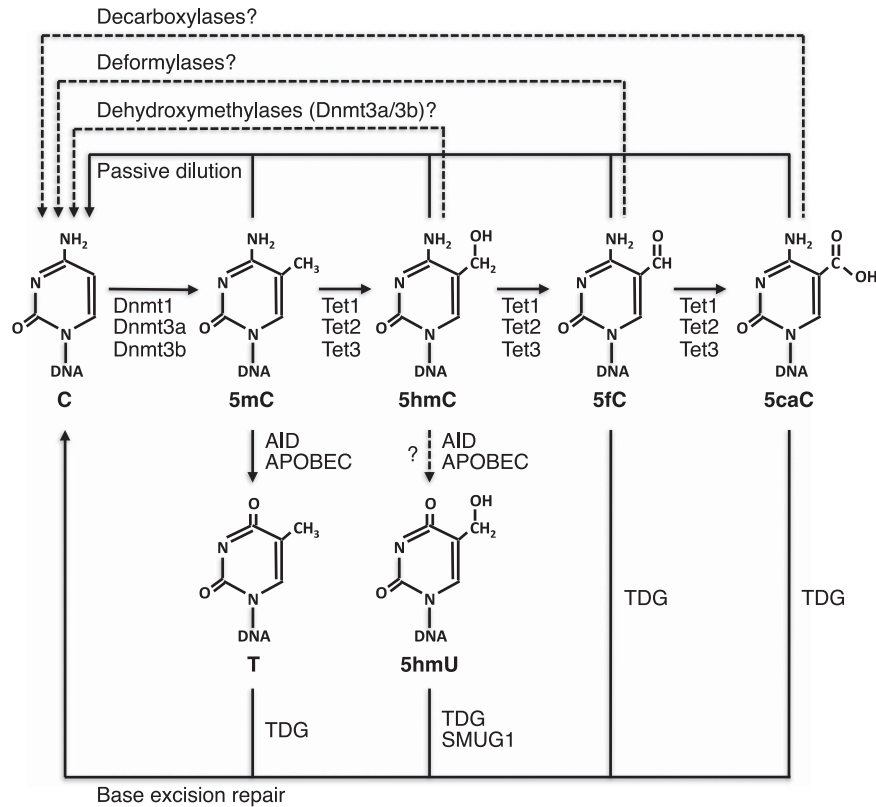
involved in DNA demethylation.<sup>15,16</sup> Indeed, Tet-mediated oxidation of 5mC appears to be the only source of 5hmC, as Dnmt1/Dnmt3a/Dnmt3b triple knockout ES cells lack both 5mC and 5hmC<sup>28–30</sup> and depletion of Tet proteins substantially reduces or abolishes the production of 5hmC in various cells and tissues.<sup>15,16,29,31–37</sup> However, Tet proteins alone do not seem to be sufficient to complete DNA demethylation by converting 5hmC, 5fC or 5caC to unmodified cytosine. Thus, it is generally accepted that 5hmC, 5fC and 5caC are intermediates in the process of DNA demethylation.

Various mechanisms of DNA demethylation involving Tet-mediated oxidation have been proposed (Figure 2). The simplest mechanism is ‘passive’ dilution of 5hmC, 5fC and 5caC owing to the lack of maintenance during DNA replication. In support of this mechanism, *in vitro* assays have revealed that Dnmt1 methylates hemi-hydroxymethylated CpG sites much more poorly than hemi-methylated CpG sites.<sup>38,39</sup>

Several DNA replication-independent ‘active’ pathways have also been suggested. First, 5hmC can be further oxidized to 5fC and 5caC, which can be recognized and excised from DNA by thymine DNA glycosylase (TDG).<sup>20,40,41</sup> The resulting abasic site could then be repaired by the base excision repair (BER) pathway, thus generating an unmodified cytosine. Another possibility is that deformylases or decarboxylases could convert 5fC and 5caC directly to unmodified cytosine, although whether such enzymes exist remains an open question. Second, the AID/APOBEC family of deaminases has been shown to deaminate 5hmC to 5-hydroxymethyluracil, which can then be excised by TDG and SMUG1, another DNA glycosylase, and replaced by cytosine through BER.<sup>42,43</sup> Deamination of 5mC by AID/APOBEC enzymes, resulting in a T:G mismatch leading to subsequent repair by TDG and BER, has also been implicated in DNA demethylation.<sup>43–46</sup> Third, a recent study provided *in vitro* evidence that Dnmt3a and Dnmt3b, in addition to their methyltransferase activity, function as dehydroxymethylases that convert 5hmC directly to cytosine. The methyltransferase and dehydroxymethylase activities seem to be regulated by the redox state of the enzymes. Reduction conditions (for example, the presence of DTT or  $\beta$ -mercaptoethanol) inhibit their dehydroxymethylase activity, whereas oxidation conditions (for example, presence of  $H_2O_2$ ) inhibit their methyltransferase activity.<sup>47</sup> The bacterial HhaI methyltransferase has also been shown to have dehydroxymethylase activity *in vitro*.<sup>48</sup> Interestingly, a previous study suggests that Dnmt3a and Dnmt3b exhibit dual actions in mammalian cells, being involved in both CpG methylation and active demethylation at some loci (for example, *pS2* gene promoter), although the mechanism of demethylation was suspected to involve Dnmt3a/3b-mediated deamination of 5mC, TDG and BER.<sup>49</sup> It would be interesting to revisit the mechanism of demethylation and determine whether the dehydroxymethylase activities of Dnmt3a and Dnmt3b are partly responsible.

### TET PROTEINS IN DEMETHYLATION IN PREIMPLANTATION EMBRYOS

Genome-wide analysis reveals that the male and female gametes have different levels of CpG methylation, with  $\sim 90\%$  in sperm and  $\sim 40\%$  in oocytes.<sup>50,51</sup> After fertilization, most DNA methylation marks inherited from gametes are erased during preimplantation development, exceptions include those associated with ICRs and intracisternal-A particles that resist this wave of global demethylation.<sup>50–53</sup> The mechanisms by which the paternal and maternal genomes undergo demethylation are distinct. In the zygote, the male pronucleus, but not its female counterpart, undergoes rapid global loss of 5mC before the onset of DNA

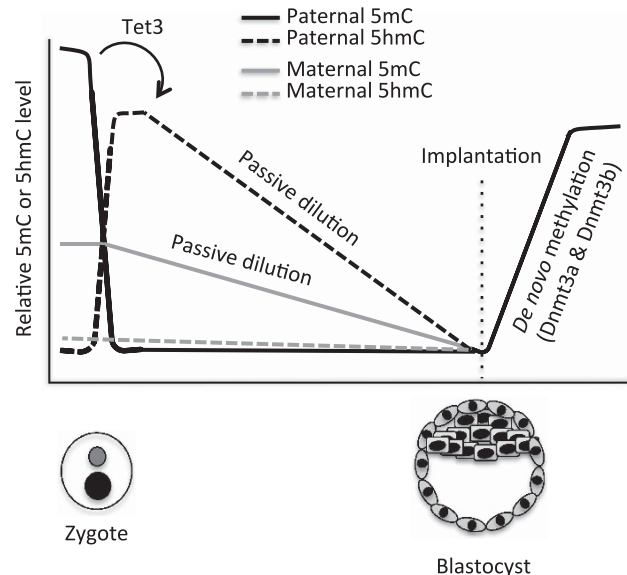


**Figure 2** Proposed DNA demethylation pathways involving Tet proteins. Tet proteins catalyze 5mC oxidation to 5hmC, which can be further converted to 5fC and 5caC by Tet proteins or to 5-hydroxymethyluracil (5hmU) by AID/APOBEC deaminases (recent evidence suggests that 5hmC is an unlikely substrate for AID/APOBEC). 5mC can also be deaminated to thymine (T) by AID/APOBEC. 5fC, 5caC, 5hmU and T can then be excised by glycosylases (TDG and SMUG1) and replaced by unmodified cytosine (C) following base excision repair. As 5hmC, 5fC, and 5caC are poorly recognized by Dnmt1, demethylation can also be achieved by passive dilution with DNA replication. In addition, Dnmt3a and Dnmt3b have been shown to function as dehydroxymethylases that directly convert 5hmC to C *in vitro*, and putative deformylases and decarboxylases could directly convert 5fC and 5caC, respectively, to C. Solid lines represent processes with relatively strong evidence and dashed lines represent processes that need to be further confirmed.

replication, suggesting an active mechanism.<sup>54–56</sup> In contrast, the maternal genome is passively demethylated during cleavage divisions, presumably owing to the exclusion of Dnmt1, the maintenance DNA methyltransferase, from the nucleus.<sup>57</sup>

Recent studies using immunofluorescence revealed that, in the zygote, concomitant with the loss of 5mC signal in the male pronucleus, there is a dramatic increase in 5hmC, as well as 5fC and 5caC, thus suggesting Tet-mediated 5mC oxidation.<sup>58–61</sup> Tet3, but not Tet1 and Tet2, is highly expressed in oocytes and zygotes.<sup>36,58,59</sup> Indeed, depletion of maternal Tet3 blocks the conversion of 5mC to 5hmC in the male pronucleus in the zygote.<sup>36,58</sup> 5mC oxidation seems to be a key step in the erasure of paternal methylation marks, as Tet3 deficiency inhibits demethylation of paternal genes.<sup>36</sup> Although BER has been proposed to be involved in active demethylation in preimplantation embryos,<sup>62,63</sup> 5hmC, 5fC and 5caC do not appear to be rapidly replaced by unmodified cytosine. Instead, they persist in the paternal genome and gradually decline during cleavage divisions.<sup>36,59–61</sup> These results suggest that, although 5hmC, 5fC and 5caC are generated in the zygote by an enzyme-catalyzed process, their loss during preimplantation development is primarily through a DNA replication-coupled passive process (Figure 3).

Although the maternal and paternal genomes are exposed to an identical environment in the zygote, the maternal genome is protected from Tet3-mediated 5mC oxidation. PGC7 (also known as Stella and Dppa3), a maternal factor, has recently been shown to be required for this protection. Depletion of maternal PGC7 results in conversion of



**Figure 3** DNA demethylation during preimplantation development. Shortly after fertilization, paternal 5mC is rapidly oxidized by Tet3. The resulting 5hmC, as well as maternal 5mC, gradually declines during subsequent cleavage divisions primarily through passive dilution. After implantation, Dnmt3a/3b-mediated *de novo* methylation occurs to establish lineage-specific methylation patterns.

5mC to 5hmC in both the male and female pronuclei.<sup>58</sup> Consistent with this finding, a previous study showed that PGC7 protects the maternal genome from demethylation in early embryos.<sup>64</sup> In normal zygotes, Tet3 is enriched and preferentially associated with the male pronucleus.<sup>36,65</sup> PGC7 seems to bind histone H3K9me2, which is abundant in the maternal chromatin but absent in the paternal chromatin with the exception of some imprinted loci, and inhibit Tet3 binding to the maternal chromatin and paternally imprinted loci.<sup>65</sup>

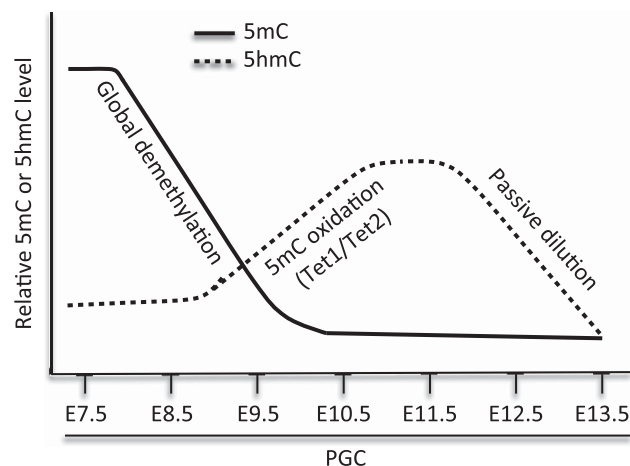
Reprogramming of the parental genomes in early embryos is believed to be important for the establishment of totipotency, but its biological significance remains largely unknown. Embryos conceived from Tet3-depleted oocytes implant normally but show high frequency of degeneration and morphological abnormalities, starting from mid-gestation, with only ~20% surviving to term,<sup>36</sup> whereas embryos without PGC7 show preimplantation defects and rarely reach the blastocyst stage.<sup>66</sup> These findings support the notion that epigenetic reprogramming is crucial for embryonic development, although the developmental phenotypes observed may not be entirely attributable to defects in epigenetic reprogramming.

### TET PROTEINS IN DEMETHYLATION IN PGCs

In mice, PGCs are specified around embryonic day (E) 7.25 in the epiblast of the developing embryo, with the involvement of bone morphogenetic protein (BMP) signaling and the transcription factors BLIMP1 and PRDM14.<sup>67</sup> Shortly afterward, PGCs begin migrating along the embryonic–extraembryonic interface and eventually arrive at the genital ridge, mostly by E11.5.<sup>68</sup> PGCs initially have similar epigenetic marks as other epiblast cells, including significant levels of DNA methylation,<sup>69,70</sup> and thus need to be reprogrammed to generate an epigenome for the development of germ cells.

Previous studies indicated that, during their migration, PGCs undergo genome-wide demethylation including the erasure of DNA methylation marks at ICRs.<sup>71,72</sup> Exceptions include intracisternal-A particles and other active retrotransposons, which appear to be resistant to complete demethylation.<sup>71–73</sup> In the past several years, a number of groups have generated genome-wide high-resolution DNA methylation profiles of PGCs at different stages of the reprogramming process.<sup>45,73–77</sup> These studies reveal that PGCs undergo demethylation in two phases. The first phase occurs during PGC expansion and migration from ~E8.5, and involves global demethylation affecting sequences of almost all genomic features. Passive demethylation may have a major role in this phase of genome-wide loss of methylation, as Dnmt3a and Dnmt3b, as well as Uhrf1 (also known as NP95), an essential factor for Dnmt1 function, are repressed in PGCs.<sup>78,79</sup> The second phase occurs from E9.5 to E13.5 and affects specific loci including ICRs, CpG islands (CGIs) on the X chromosome and germline-specific genes.<sup>73–76</sup>

Recent studies provide evidence for the involvement of Tet-mediated 5mC oxidation in demethylation in PGCs.<sup>37,75,76,80–82</sup> Tet1 and Tet2 are expressed in PGCs between E9.25 and E11.5, but Tet3 is undetectable in PGCs.<sup>75,80,81</sup> Hackett *et al.*<sup>75</sup> and Yamaguchi *et al.*<sup>76</sup> used immunofluorescence to analyze PGCs at various time points and showed that both 5mC and 5hmC levels are low at E8.5, 5hmC levels begin to increase between E9.5 and E10.5, peak at ~E11.5 and then gradually decline from E11.5 to E13.5. Genetic studies reveal that deficiency for Tet1 or both Tet1 and Tet2 has no effect on global demethylation in PGCs, but results in defective demethylation and altered expression of specific genes including meiotic genes and imprinted genes.<sup>37,80</sup> These results suggest that Tet1 and Tet2 are responsible for the production of 5hmC in PGCs and that 5hmC enrichment is followed by replication-coupled dilution. Using PGCs



**Figure 4** DNA demethylation in primordial germ cells (PGCs). DNA demethylation in PGCs occurs in two phases. The first phase involves global demethylation. The second phase affects specific loci including imprinting control regions. Tet1/Tet2-mediated 5mC oxidation occurs mainly in the second phase, and 5hmC enrichment is followed by gradual decline at a rate consistent with passive dilution.

differentiated from wild-type or Tet1- and Tet2-depleted ES cells *in vitro*, Vincent *et al.*<sup>81</sup> showed that, in the absence of Tet1 and Tet2, the first phase of global demethylation is unaffected but numerous promoters and gene bodies become hypermethylated. Taken together, these findings suggest that Tet-mediated conversion of 5mC to 5hmC is mainly involved in the second phase of demethylation in PGCs including the erasure of imprints at ICRs (Figure 4).

Previous studies have shown that embryonic germ cells (EGCs) have the capacity to reprogram somatic genomes, including the erasure of imprints at ICRs, in hybrid cells.<sup>83</sup> Recently, Piccolo *et al.*<sup>82</sup> used this system to address the requirement of Tet1 and Tet2 in EGC-induced pluripotent reprogramming. Intriguingly, Tet2 induces 5mC oxidation at pluripotent genes (for example, *Oct4*), as well as expression of these genes, and is thus required for the efficient reprogramming capacity of EGCs, whereas Tet1 is necessary to induce 5mC oxidation specifically at ICRs. These results suggest that Tet1 and Tet2 may have distinct genomic targets.

Despite the participation of Tet1 and Tet2 in epigenetic reprogramming in PGCs, these enzymes do not seem to be essential for germ cell development and fertility. Mice deficient for either Tet1 or Tet2 or both are viable and fertile, although Tet1-null and Tet1/Tet2-double-null female mice show reduced fertility owing to meiotic defects.<sup>32,35,37,80,84</sup>

### TET PROTEINS IN EMBRYONIC AND POSTNATAL DEVELOPMENT

Tet1, Tet2 and Tet3 show different expression patterns. Tet1 and Tet2 are highly expressed in the inner cell mass of mouse blastocysts, as well as in ES cells (which are derived from inner cell mass), whereas Tet3 is highly expressed in mouse oocytes and zygotes.<sup>16,58,59</sup> Upon differentiation of mouse ES cells, Tet1 and Tet2 are rapidly downregulated and Tet3 is upregulated.<sup>15,16,31</sup> Tet2 and Tet3 also appear to be widely expressed, at various levels, in adult tissues.<sup>16</sup> Consistent with the distinct expression patterns of Tet proteins, recent genetic studies indicate that these enzymes have different functions in mammalian development.

Multiple groups have reported that depletion of Tet1 in mouse ES cells results in 5hmC reduction, alterations in gene expression and

defects in self-renewal or differentiation.<sup>16,29–32</sup> However, Tet1-null mice show no overt developmental abnormalities, although some mutant mice are slightly smaller at birth.<sup>32,80</sup> Several Tet2-mutant alleles have been generated. Tet2-null mice develop normally and are fertile.<sup>35,84</sup> However, Tet2 deletion, either systemically or in the hematopoietic compartment, results in hematological phenotypes in adult animals characterized by progressive enlargement of the hematopoietic stem cell pool and eventual myeloid malignancies.<sup>33–35,84</sup> Tet2 expression is ubiquitous in the hematopoietic compartment, including in the stem and progenitor subsets and in mature myeloid and lymphoid cells.<sup>33,35</sup> Although the molecular mechanisms by which Tet2 deficiency leads to the hematological phenotypes remain to be elucidated, Tet2-null mice show decreased levels of 5hmC and concurrent increased levels of 5mC in bone marrow and spleen.<sup>33–35</sup> Tet2 could regulate genes important for hematopoiesis by modulating DNA methylation. Consistent with the observed phenotypes in Tet2-deficient mice, *TET2* is frequently mutated in patients with various myeloid malignancies, such as myelodysplastic syndromes, myeloproliferative neoplasms, chronic myelomonocytic leukemia, acute myeloid leukemia and secondary acute myeloid leukemia.<sup>85</sup> The fact that Tet2 deficiency in mice recapitulates the major phenotypes in human patients suggests that *TET2* mutations are driver mutations in hematological malignancies.

Tet1 and Tet2 seem to have partially redundant functions in embryonic development. Although a fraction of Tet1/Tet2 double knockout (DKO) mice are viable and fertile, some DKO embryos exhibit mid-gestation abnormalities and most DKO animals die perinatally with a variety of malformations, such as exencephaly, hemorrhage in the head and profound growth retardation. Tet3 is upregulated in DKO mice, suggesting that compensation by Tet3 may contribute to the viability of DKO mice.<sup>37</sup> Systemic deletion of Tet3 leads to neonatal lethality, and maternal deletion impairs reprogramming in the zygote.<sup>36</sup> Given the wide expression of Tet3 in somatic tissues,<sup>16</sup> it would be interesting to determine the function of Tet3 in adult animals.

## CONCLUDING REMARKS

Since the discovery that Tet proteins can convert 5mC to 5hmC in 2009,<sup>15</sup> tremendous progress has been made in understanding the functions of these enzymes and their products (5hmC, 5fC and 5caC). It is now widely accepted that 5hmC, 5fC and 5caC serve as intermediates in the process of DNA demethylation. Genetic studies in mice have confirmed the involvement of distinct Tet proteins in demethylation in the zygote and PGCs. Despite the progress, complete models of DNA demethylation in various physiological contexts remain to be assembled. One of the challenges is that multiple mechanisms seem to work cooperatively to achieve demethylation, and the relative contribution of these mechanisms and how they are orchestrated need to be clarified. Some of the proposed mechanisms may not be relevant or significant. For instance, recent biochemical studies suggest that 5hmC is an unlikely substrate for the AID/APOBEC family of deaminases.<sup>86,87</sup> Emerging evidence suggests that 5hmC, in addition to its role in DNA demethylation, may function as a stable epigenetic mark. Indeed, recent studies have identified several 5hmC-specific 'readers' including MBD3, MeCP2, Uhrf1 and Uhrf2.<sup>88–91</sup> Further work needs to be done to determine the significance of 5hmC in regulating chromatin structure and function, including gene expression, and the mechanism by which 5hmC is maintained. Tet proteins are large molecules and may have other functions, some of which may be independent of their enzymatic

activities. For example, several recent reports show that Tet proteins interact with O-GlcNAc transferase and promote histone O-GlcNAcylation.<sup>92–94</sup> Another area of intense research is the role of Tet proteins and 5hmC in cancer. We expect to see exciting discoveries addressing these issues in the coming years.

## ACKNOWLEDGEMENTS

We thank A.K. Singh for help with formatting the references. Work in our laboratory is supported by the Cancer Prevention and Research Institute of Texas (CPRIT) (R1108). T.C. is a CPRIT Scholar in Cancer Research.

- 1 Li, E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat. Rev. Genet.* **3**, 662–673 (2002).
- 2 Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220 (2013).
- 3 Tomizawa, S. & Sasaki, H. Genomic imprinting and its relevance to congenital disease, infertility, molar pregnancy and induced pluripotent stem cell. *J. Hum. Genet.* **57**, 84–91 (2012).
- 4 Jakovcevski, M. & Akbarian, S. Epigenetic mechanisms in neurological disease. *Nat. Med.* **18**, 1194–1204 (2012).
- 5 Ballestar, E. Epigenetic alterations in autoimmune rheumatic diseases. *Nat. Rev. Rheumatol.* **7**, 263–271 (2011).
- 6 Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome - biological and translational implications. *Nat. Rev. Cancer* **11**, 726–734 (2011).
- 7 Chen, T. & Li, E. Structure and function of eukaryotic DNA methyltransferases. *Curr. Top. Dev. Biol.* **60**, 55–89 (2004).
- 8 Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
- 9 Lei, H., Oh, S. P., Okano, M., Juttermann, R., Goss, K. A., Jaenisch, R. *et al.* De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122**, 3195–3205 (1996).
- 10 Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
- 11 Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B. & Bestor, T. H. Dnmt3L and the establishment of maternal genomic imprints. *Science* **294**, 2536–2539 (2001).
- 12 Hata, K., Okano, M., Lei, H. & Li, E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129**, 1983–1993 (2002).
- 13 Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E. *et al.* Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **429**, 900–903 (2004).
- 14 Seisenberger, S., Peat, J. R. & Reik, W. Conceptual links between DNA methylation reprogramming in the early embryo and primordial germ cells. *Curr. Opin. Cell Biol.* **S0955-0674**, 00036–00037 (2013).
- 15 Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
- 16 Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C. & Zhang, Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
- 17 Penn, N. W., Suwalski, R., O'Riley, C., Bojanowski, K. & Yura, R. The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem. J.* **126**, 781–790 (1972).
- 18 Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929–930 (2009).
- 19 Ko, M., Huang, Y., Jankowska, A. M., Pape, U. J., Tahiliani, M., Bandukwala, H. S. *et al.* Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839–843 (2010).
- 20 He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q. *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011).
- 21 Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A. *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303 (2011).
- 22 Iyer, L. M., Tahiliani, M., Rao, A. & Aravind, L. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle* **8**, 1698–1710 (2009).
- 23 Hashimoto, H., Vertino, P. M. & Cheng, X. Molecular coupling of DNA methylation and histone methylation. *Epigenomics* **2**, 657–669 (2010).
- 24 Frauer, C., Rottach, A., Meilinger, D., Bultmann, S., Fellinger, K., Hasenoder, S. *et al.* Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1. *PLoS ONE* **6**, e16627 (2011).
- 25 Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J. *et al.* Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol. Cell* **42**, 451–464 (2011).

- 26 Xu, Y., Xu, C., Kato, A., Tempel, W., Abreu, J. G., Bian, C. *et al.* Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for *Xenopus* eye and neural development. *Cell* **151**, 1200–1213 (2012).
- 27 Ko, M., An, J., Bandukwala, H. S., Chavez, L., Aljo, T., Pastor, W. A. *et al.* Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* **497**, 122–126 (2013).
- 28 Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C. *et al.* 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 (2006).
- 29 Ficiz, G., Branco, M. R., Seisenberger, S., Santos, F., Krueger, F., Hore, T. A. *et al.* Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473**, 398–402 (2011).
- 30 Williams, K., Christensen, J., Pedersen, M. T., Johansen, J. V., Cloos, P. A., Rappasber, J. *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343–348 (2011).
- 31 Koh, K. P., Yabuuchi, A., Rao, S., Huang, Y., Cunliff, K., Nardone, J. *et al.* Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**, 200–213 (2011).
- 32 Dawlaty, M. M., Ganz, K., Powell, B. E., Hu, Y. C., Markoulaki, S., Cheng, A. W. *et al.* Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* **9**, 166–175 (2011).
- 33 Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Loby, D., Lohby, C. *et al.* Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**, 11–24 (2011).
- 34 Ko, M., Bandukwala, H. S., An, J., Lamperti, E. D., Thompson, E. C., Hastie, R. *et al.* Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc. Natl Acad. Sci. USA* **108**, 14566–14571 (2011).
- 35 Li, Z., Cai, X., Cai, C. L., Wang, J., Zhang, W., Petersen, B. E. *et al.* Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood* **118**, 4509–4518 (2011).
- 36 Gu, T. P., Guo, F., Yang, H., Wu, H. P., Xu, G. F., Liu, W. *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606–610 (2011).
- 37 Dawlaty, M. M., Breiling, A., Le, T., Raddatz, G., Barrasa, M. I., Cheng, A. W. *et al.* Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev. Cell* **24**, 310–323 (2013).
- 38 Valinluck, V. & Sowers, L. C. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res.* **67**, 946–950 (2007).
- 39 Hashimoto, H., Liu, Y., Upadhyay, A. K., Chang, Y., Howerton, S. B., Vertino, P. M. *et al.* Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res.* **40**, 4841–4849 (2012).
- 40 Maiti, A. & Drohat, A. C. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J. Biol. Chem.* **286**, 35334–35338 (2011).
- 41 Zhang, L., Lu, X., Lu, J., Liang, H., Dai, Q., Xu, G. L. *et al.* Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nat. Chem. Biol.* **8**, 328–330 (2012).
- 42 Guo, J. U., Su, Y., Zhong, C., Ming, G. L. & Song, H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145**, 423–434 (2011).
- 43 Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A. *et al.* Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146**, 67–79 (2011).
- 44 Rai, K., Huggins, I. J., James, S. R., Karpf, A. R., Jones, D. A. & Cairns, B. R. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* **135**, 1201–1212 (2008).
- 45 Popp, C., Dean, W., Feng, S., Cokus, S. J., Andrews, S., Pellegrini, M. *et al.* Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**, 1101–1105 (2010).
- 46 Bhutani, N., Brady, J. J., Damian, M., Sacco, A., Corbel, S. Y. & Blau, H. M. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463**, 1042–1047 (2010).
- 47 Chen, C. C., Wang, K. Y. & Shen, C. K. The mammalian de novo DNA methyltransferases DNMT3A and DNMT3B are also DNA 5-hydroxymethylcytosine dehydroxymethylases. *J. Biol. Chem.* **287**, 33116–33121 (2012).
- 48 Liutkeviciute, Z., Lukinavicius, G., Masevicius, V., Daujotyte, D. & Klimasauskas, S. Cytosine-5-methyltransferases add aldehydes to DNA. *Nat. Chem. Biol.* **5**, 400–402 (2009).
- 49 Metivier, R., Gallais, R., Tiffocche, C., Le Peron, C., Jurkowska, R. Z., Carmouche, R. P. *et al.* Cyclical DNA methylation of a transcriptionally active promoter. *Nature* **452**, 45–50 (2008).
- 50 Kobayashi, H., Sakurai, T., Imai, M., Takahashi, N., Fukuda, A., Yayoi, O. *et al.* Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. *PLoS Genet.* **8**, e1002440 (2012).
- 51 Smith, Z. D., Chan, M. M., Mikkelsen, T. S., Gu, H., Gnirke, A., Regev, A. *et al.* A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).
- 52 Borgel, J., Guibert, S., Li, Y., Chiba, H., Schubeler, D., Sasaki, H. *et al.* Targets and dynamics of promoter DNA methylation during early mouse development. *Nat. Genet.* **42**, 1093–1100 (2010).
- 53 Smallwood, S. A., Tomizawa, S., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A. *et al.* Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat. Genet.* **43**, 811–814 (2011).
- 54 Mayer, W., Niveleau, A., Walter, J., Fundele, R. & Haaf, T. Demethylation of the zygotic paternal genome. *Nature* **403**, 501–502 (2000).
- 55 Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R. *et al.* Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10**, 475–478 (2000).
- 56 Santos, F., Hendrich, B., Reik, W. & Dean, W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev. Biol.* **241**, 172–182 (2002).
- 57 Howell, C. Y., Bestor, T. H., Ding, F., Latham, K. E., Mertineit, C., Trasler, J. M. *et al.* Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell* **104**, 829–838 (2001).
- 58 Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C. J., Zakhartchenko, V., Boiani, M. *et al.* 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun.* **2**, 241 (2011).
- 59 Iqbal, K., Jin, S. G., Pfeifer, G. P. & Szabo, P. E. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl Acad. Sci. USA* **108**, 3642–3647 (2011).
- 60 Inoue, A. & Zhang, Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* **334**, 194 (2011).
- 61 Inoue, A., Shen, L., Dai, Q., He, C. & Zhang, Y. Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. *Cell Res.* **21**, 1670–1676 (2011).
- 62 Gehring, M., Reik, W. & Henikoff, S. DNA demethylation by DNA repair. *Trends Genet.* **25**, 82–90 (2009).
- 63 Wu, S. C. & Zhang, Y. Active DNA demethylation: many roads lead to Rome. *Nat. Rev. Mol. Cell Biol.* **11**, 607–620 (2010).
- 64 Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H. *et al.* PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat. Cell Biol.* **9**, 64–71 (2007).
- 65 Nakamura, T., Liu, Y. J., Nakashima, H., Umehara, H., Inoue, K., Matoba, S. *et al.* PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**, 415–419 (2012).
- 66 Payer, B., Saitou, M., Barton, S. C., Thresher, R., Dixon, J. P., Zahn, D. *et al.* Stella is a maternal effect gene required for normal early development in mice. *Curr. Biol.* **13**, 2110–2117 (2003).
- 67 Saitou, M. Germ cell specification in mice. *Curr. Opin. Genet. Dev.* **19**, 386–395 (2009).
- 68 Molyneux, K. A., Stallock, J., Schaible, K. & Wylie, C. Time-lapse analysis of living mouse germ cell migration. *Dev. Biol.* **240**, 488–498 (2001).
- 69 Seki, Y., Hayashi, K., Itoh, K., Mizugaki, M., Saitou, M. & Matsui, Y. 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 (2005).
- 70 Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T. & Saitou, M. A signaling principle for the specification of the germ cell lineage in mice. *Cell* **137**, 571–584 (2009).
- 71 Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maari, O., Reik, W. *et al.* Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117**, 15–23 (2002).
- 72 Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J. *et al.* Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* **35**, 88–93 (2003).
- 73 Guibert, S., Forne, T. & Weber, M. Global profiling of DNA methylation erasure in mouse primordial germ cells. *Genome Res.* **22**, 633–641 (2012).
- 74 Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F. *et al.* The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol. Cell* **48**, 849–862 (2012).
- 75 Hackett, J. A., Sengupta, R., Zyllicz, J. J., Murakami, K., Lee, C., Down, T. A. *et al.* Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* **339**, 448–452 (2013).
- 76 Yamaguchi, S., Hong, K., Liu, R., Inoue, A., Shen, L., Zhang, K. *et al.* Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. *Cell Res.* **23**, 329–339 (2013).
- 77 Kobayashi, H., Sakurai, T., Miura, F., Imai, M., Mochiduki, K., Yanagisawa, E. *et al.* High-resolution DNA methylome analysis of primordial germ cells identifies gender-specific reprogramming in mice. *Genome Res.* **23**, 616–627 (2013).
- 78 Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K. & Saitou, M. Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. *Genes Dev.* **22**, 1617–1635 (2008).
- 79 Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M. & Saitou, M. Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J.* **32**, 340–353 (2013).
- 80 Yamaguchi, S., Hong, K., Liu, R., Shen, L., Inoue, A., Diep, D. *et al.* Tet1 controls meiosis by regulating meiotic gene expression. *Nature* **492**, 443–447 (2012).
- 81 Vincent, J. J., Huang, Y., Chen, P. Y., Feng, S., Calvopina, J. H., Nee, K. *et al.* Stage-specific roles for tet1 and tet2 in DNA demethylation in primordial germ cells. *Cell Stem Cell* **12**, 470–478 (2013).
- 82 Piccolo, F. M., Bagci, H., Brown, K. E., Landeira, D., Soza-Ried, J., Feytout, A. *et al.* Different roles for Tet1 and Tet2 proteins in reprogramming-mediated erasure of imprints induced by EGC fusion. *Mol. Cell* **49**, 1023–1033 (2013).

- 83 Tada, M., Tada, T., Lefebvre, L., Barton, S. C. & Surani, M. A. Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *EMBO J.* **16**, 6510–6520 (1997).
- 84 Quivoron, C., Couronne, L., Della Valle, V., Lopez, C. K., Plo, I., Wagner-Ballon, O. *et al.* TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell* **20**, 25–38 (2011).
- 85 Abdel-Wahab, O. Genetics of the myeloproliferative neoplasms. *Curr. Opin. Hematol.* **18**, 117–123 (2011).
- 86 Nabel, C. S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H. L., Stivers, J. T. *et al.* AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat. Chem. Biol.* **8**, 751–758 (2012).
- 87 Rangam, G., Schmitz, K. M., Cobb, A. J. & Petersen-Mahrt, S. K. AID enzymatic activity is inversely proportional to the size of cytosine C5 orbital cloud. *PLoS ONE* **7**, e43279 (2012).
- 88 Yildirim, O., Li, R., Hung, J. H., Chen, P. B., Dong, X., Ee, L. S. *et al.* Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* **147**, 1498–1510 (2011).
- 89 Frauer, C., Hoffmann, T., Bultmann, S., Casa, V., Cardoso, M. C., Antes, I. *et al.* Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS ONE* **6**, e21306 (2011).
- 90 Mellen, M., Ayata, P., Dewell, S., Kriaucionis, S. & Heintz, N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151**, 1417–1430 (2012).
- 91 Spruijt, C. G., Gnerlich, F., Smits, A. H., Pfaffeneder, T., Jansen, P. W., Bauer, C. *et al.* Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* **152**, 1146–1159 (2013).
- 92 Chen, Q., Chen, Y., Bian, C., Fujiki, R. & Yu, X. TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature* **493**, 561–564 (2013).
- 93 Vella, P., Scelfo, A., Jammula, S., Chiacchiera, F., Williams, K., Cuomo, A. *et al.* Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. *Mol. Cell* **49**, 645–656 (2013).
- 94 Deplus, R., Delatte, B., Schwinn, M. K., Defrance, M., Mendez, J., Murphy, N. *et al.* TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J.* **32**, 645–655 (2013).