EPIGENETICS

TET-mediated active DNA demethylation: mechanism, function and beyond

Xiaoji Wu[1–5] *and Yi Zhang*[1–4]

Abstract | In mammals, DNA methylation in the form of 5-methylcytosine (5mC) can be actively reversed to unmodified cytosine (C) through TET dioxygenase-mediated oxidation of 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), followed by replication-dependent dilution or thymine DNA glycosylase (TDG)-dependent base excision repair. In the past few years, biochemical and structural studies have revealed mechanistic insights into how TET and TDG mediate active DNA demethylation. Additionally, many regulatory mechanisms of this process have been identified. Technological advances in mapping and tracing the oxidized forms of 5mC allow further dissection of their functions. Furthermore, the biological functions of active DNA demethylation in various biological contexts have also been revealed. In this Review, we summarize the recent advances and highlight key unanswered questions.

[1]*Howard Hughes Medical Institute, Boston, Massachusetts 02115, USA.*
[2]*Program in Cellular and Molecular Medicine, Boston Children's Hospital, WAB-149G, 200 Longwood Avenue, Boston, Massachusetts 02115, USA.*
[3]*Harvard Stem Cell Institute, Boston, Massachusetts 02115, USA.*
[4]*Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.*
[5]*Ph.D. Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, Massachusetts 02115, USA.*

Correspondence to Y.Z.
yzhang@genetics.med.harvard.edu

doi:10.1038/nrg.2017.33
Published online 30 May 2017

In mammals, 5-methylcytosine (5mC) is the major form of DNA modification, and it has important roles in development and disease[1,2]. About 60–80% of the CpG sites in the mammalian genome are modified by 5mC[1]. The major functions of 5mC include mediating genomic imprinting and X-chromosome-inactivation, repressing transposable elements and regulating transcription[3].

5mC is both chemically and genetically stable. Chemically, the methyl group is connected to the 5-position of the cytosine base through a stable carbon–carbon bond, creating a barrier for direct removal of the methyl group. Genetically, on its establishment by *de novo* DNA methyltransferase 3A (DNMT3A) and DNMT3B[4], 5mC is maintained by the maintenance methyltransferase DNMT1, which recognizes hemi-methylated CpG dyads through its functional partner UHRF1 (REFS 5–7). This maintenance mechanism is crucial because it ensures faithful re-establishment of 5mC on the newly synthesized strand after DNA replication[8,9].

Despite its stability, mammalian 5mC can still be reversed to its unmodified state in several ways. First, a lack of functional DNA methylation maintenance machinery can result in the dilution of 5mC during DNA replication, a process known as passive DNA demethylation. Second, TET proteins can mediate the iterative oxidation of 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). Replication-dependent dilution of these oxidized forms of 5mC or thymine DNA glycosylase (TDG)-mediated excision of 5fC and 5caC coupled with base excision repair (BER) will also result in demethylation (FIG. 1a). This process, known as active DNA demethylation, is the focus of this Review. Although several other TET–TDG-independent mechanisms have been proposed to mediate active DNA demethylation (reviewed in REFS 10–12), the TET–TDG pathway has gained the most support.

In this Review, we summarize recent advances in understanding the mechanism and function of TET-mediated active DNA demethylation. First, we discuss the biochemical and molecular mechanism of the TET–TDG pathway. Then, we describe how this pathway can be regulated at different levels. Furthermore, we discuss the distribution and dynamics of oxidized 5mC. Finally, we review the roles of TET-mediated demethylation in specific biological contexts.

Mechanism of TET-mediated DNA demethylation

Active erasure of 5mC from the genome can take place through different mechanisms in various organisms. In plants, the REPRESSOR OF SILENCING 1 (ROS1)/DEMETER (DME) family of DNA glycosylases can recognize and directly excise the 5mC base, leading to the restoration of unmodified cytosine through BER[13]. In mammals, in which ROS1/DME-like proteins have not been identified, active DNA demethylation takes place in a TET-dependent manner and can be coupled with TDG-mediated BER (FIG. 1a).

The discovery of the TET–TDG pathway. One of the first pieces of evidence to support the existence of active DNA demethylation in mammals was the discovery of genome-wide loss of 5mC in mouse zygotes, as revealed by immunostaining. Shortly after fertilization, the 5mC signal on the zygotic paternal genome rapidly decreases to an extent that cannot be fully explained by replication-dependent dilution[14]. This observation was independently

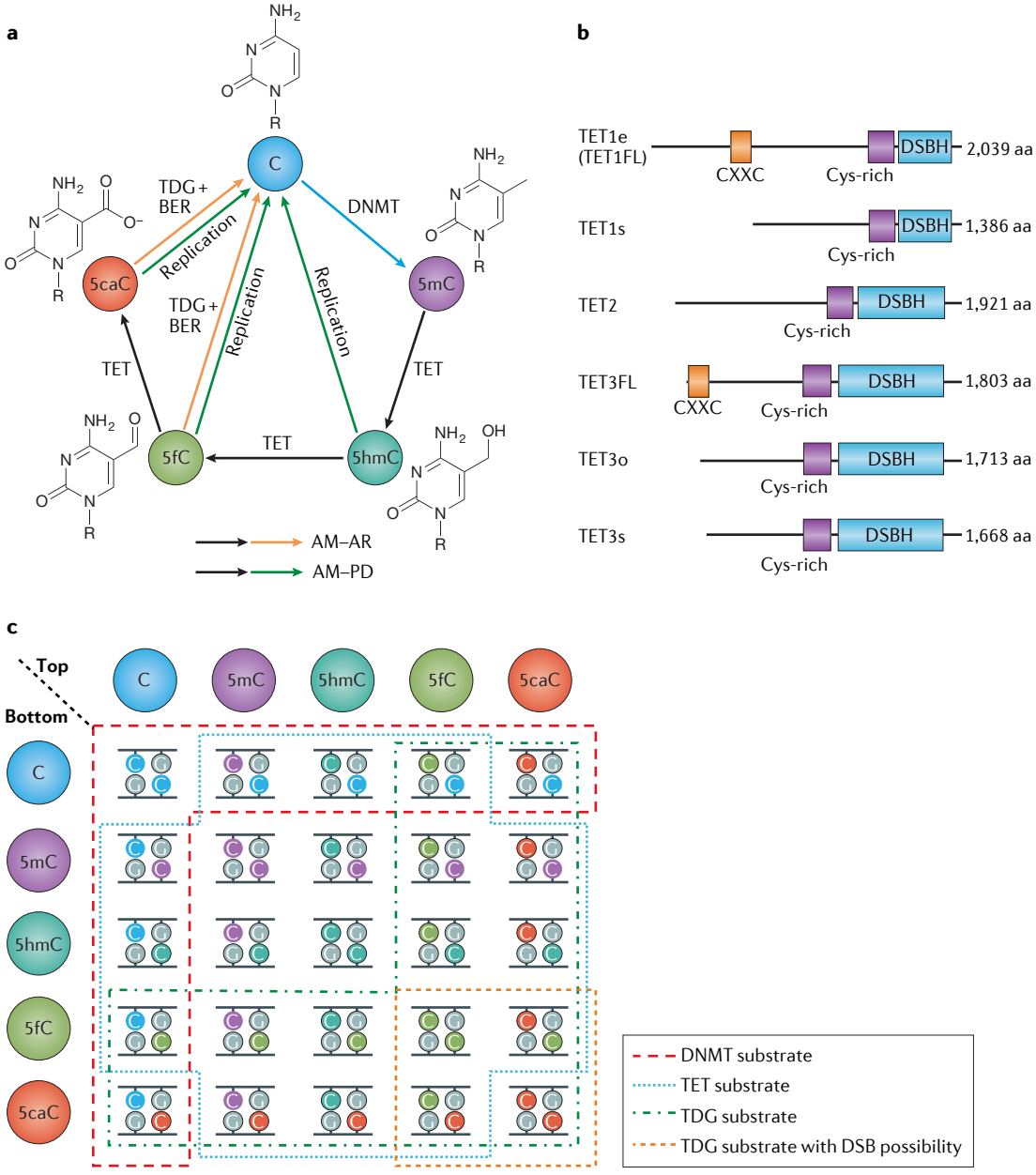


Figure 1 | **TET-mediated active DNA demethylation. a** | The cycle of active DNA demethylation. DNA methyltransferases (DNMTs) convert unmodified cytosine to 5-methylcytosine (5mC). 5mC can be converted back to unmodified cytosine by TET-mediated oxidation to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), followed by excision of 5fC or 5caC mediated by thymine DNA glycosylase (TDG) coupled with base excision repair (BER) (the process of active modification–active removal (AM–AR)) or replication-dependent dilution of 5hmC, 5fC or 5caC (the process of active modification–passive dilution (AM–PD)). **b** | Domain structure of mouse TET proteins. Cysteine-rich and double-stranded β-helix (DSBH) domains at the carboxyl terminus confer catalytic activity. Full-length TET1 (TET1FL) and TET3 have a CXXC domain at the amino terminus, whereas TET2 does not. Multiple splicing isoforms have been reported for TET1 (REF. 33) and TET3 (REFS 34,35). **c** | Possible substrates for DNMT, TET and TDG. CpG dyads with different modification states on the top and bottom strands are shown. In the case of DNMT substrates, DNMT1 prefers hemi-5mC sites (5mC:C), whereas DNMT3A and DNMT3B work in various contexts, including 5hmC:C, 5fC:C and 5caC:C[54,55,57]. TET and TDG both work on their substrates regardless of the modification status of the complementary strand[28,39]. When both strands of a CpG dyad are modified as 5fC or 5caC, it is theoretically possible that a double-strand break (DSB) may arise due to simultaneous TDG–BER on both strands. In reality, the highly coordinated TDG–BER process reduces this possibility[28]. aa, amino acids.

TET
A family of methylcytosine dioxygenase enzymes that are involved in several steps of the oxidative demethylation of 5-methylcytosine. They are named after the cancer-associated 'ten–eleven translocation', which creates a fusion between mixed-lineage leukaemia (*MLL*; also known as *KMT2A*) on chromosome 11 and *TET1* on chromosome 10.

Base excision repair
(BER). The repair of a damaged base through the following steps: excising the base to create an abasic site, generating a single-strand break (SSB) and repairing the SSB through short-patch or long-patch repair.

confirmed by locus-specific bisulfite sequencing (BS-seq)[15]. The molecular mechanism underlying these observations was not revealed until 2009, when two ground-breaking papers showed that 5hmC accumulates to a significant level in certain tissues and that human TET1 is capable of converting 5mC to 5hmC[16,17]. Human TET1 was initially identified as a fusion partner of MLL (also known as KMT2A) in cancer[18,19] and was rediscovered in 2009 as an orthologue of *Trypanosoma brucei* base J-binding protein 1 (JBP1) and JBP2, two enzymes that catalyse oxidation of thymine to 5-hydroxymethyluracil (5hmU) during the biosynthesis of trypanosome base J (β-D-glucosylhydroxymethyluracil)[17,20–22]. The other two TET proteins, TET2 and TET3, were also shown to possess 5mC to 5hmC oxidizing activity[23].

Further studies suggest that TET proteins also catalyse the oxidation of 5hmC to 5fC and 5caC[24,25], a process that is similar to the thymine hydroxylase-mediated oxidation of thymine to 5-hydroxyuracil, 5-formyluracil and 5-carboxyluracil[26]. To complete DNA demethylation, TDG recognizes and excises 5fC and 5caC from the genome, creating abasic sites before unmodified cytosine is restored through BER[24,27,28].

Structure of TET proteins. TET proteins are iron(II)/α-ketoglutarate (Fe(II)/α-KG)-dependent dioxygenases. The core catalytic domain at the carboxyl terminus is comprised of a double-stranded β-helix (DSBH) domain and a cysteine-rich domain[29] (FIG. 1b). The DSBH domain brings Fe(II), α-KG and 5mC together for oxidation, while the cysteine-rich domain wraps around the DSBH core to stabilize the overall structure and TET–DNA interaction. The TET–DNA contact does not involve the methyl group, thus allowing TET to accommodate different forms of modified cytosine[30]. Notably, the C-terminal catalytic domain alone can localize to the nucleus and oxidize 5mC[17,23,31].

Full-length TET1 and TET3 have a CXXC domain at their amino terminus, whereas the putative CXXC domain of TET2 is separated from the protein as the result of a genomic inversion during evolution, forming a gene named *Idax* (also known as *Cxxc4*)[29,32] (FIG. 1b). Interestingly, mouse TET1 preferentially exists in an N-terminus-truncated form (known as TET1s) in somatic tissues but exists in its full-length form (known as TET1e) in early embryos, embryonic stem cells (ESCs) and primordial germ cells (PGCs) (FIG. 1b). TET1s, which does not have a CXXC domain and the other N-terminal sequence, has reduced global chromatin binding compared with TET1e and confers weaker demethylation activity in cells[33]. TET3 also has multiple isoforms, including two without the CXXC domain, TET3s and TET3o[34,35] (FIG. 1b). TET3o is specifically expressed in oocytes, whereas TET3s and TET3 full-length (TET3FL) are upregulated during neuronal differentiation. Unlike TET1, TET3s and TET3o display stronger demethylation activity than TET3FL[35].

Substrate preference of TET enzymes. The substrate preference of TET enzymes can be viewed at three different levels. First, 5mC is predominantly detected in a CpG context but is also observed in non-CpG contexts, and TET may prefer one context over the other. *In vitro* analysis suggests that human TET2 has lower activity towards 5mCpC and 5mCpA than towards 5mCpG, probably due to the impaired base-stacking interaction[30]. Therefore, TET may prefer 5mC in a CpG context.

Second, 5mC, 5hmC and 5fC are all possible substrates of TET-mediated oxidation, but TET may exhibit different binding or catalytic activity towards these three substrates. Indeed, biochemical and structural studies suggest that TET prefers 5mC to 5hmC or 5fC. Enzyme kinetics analyses revealed that 5mC to 5hmC conversion catalysed by human TET1 or TET2 is faster than 5hmC to 5fC and 5fC to 5caC conversions (with a 3–5-fold difference in K_{cat}), an observation that also holds true for mouse TET2 and a *Naegleria gruberi* TET-like protein, NgTET1 (REFS 25,36,37). Mechanistically, although human TET2 has a comparable binding affinity for 5mC, 5hmC and 5fC, the difference of these substrates in hydrogen abstraction makes 5hmC and 5fC less favourable[37]. Interestingly, mutating a conserved Thr1372 residue of human TET2 can make the protein predominantly oxidize 5mC to 5hmC but not to 5fC or 5caC, shifting the substrate preference to an extreme[38].

Third, although in most cases a CpG dyad will be either symmetrically methylated by 5mC or free of any modification, all 25 combinations of modified and unmodified cytosine can happen at a CpG dyad owing to *de novo* methylation, failure of maintenance, TET-mediated oxidation or replication-dependent dilution, resulting in 21 possible substrates for TET proteins (FIG. 1c). *In vitro* studies suggest that TET proteins can use various combinations of substrates, including 5mC, 5hmC or 5fC paired with unmodified cytosine and also 5mC paired with all forms of cytosine[28,39]. The oxidation rate of 5mC at a CpG dyad is largely similar regardless of the modification status of the complementary strand[39]. This lack of substrate preference can be important *in vivo*, as active DNA demethylation is concomitant with *de novo* DNA methylation and DNA replication in many biological contexts.

Processivity of TET-mediated oxidation. The processivity of TET enzymes is a topic of growing interest. TET processivity can be viewed at the physical, chemical and genetic levels. Physical processivity refers to the capacity of TET protein to slide along the DNA from one CpG site to another (FIG. 2a). An *in vitro* study showed that DNA-bound TET does not preferentially oxidize other CpG sites on the same DNA molecule, indicating that TET is not physically processive[40]. Chemical processivity refers to the ability of TET to catalyse the oxidation of 5mC iteratively to 5caC without releasing its substrate (FIG. 2b). Two studies examined the chemical processivity *in vitro* but reached conflicting conclusions[39,40]. One explanation for this discrepancy is that TET can work in both chemically processive and non-processive manners depending on reaction conditions. Genetic processivity refers to the genetic outcome of TET-mediated oxidation

Genetic processivity
The genetic outcome of TET-mediated oxidation in cells. Genomic regions or CpG sites modified by 5-hydroxymethylcytosine (5hmC) but not 5-formylcytosine (5fc) or 5-carboxylcytosine (5caC) are regarded as having low genetic processivity, whereas genomic regions or CpG sites modified by 5fC or 5caC are regarded as having higher genetic processivity. Unlike physical and chemical processivity, which describe the biochemical behaviour of the enzyme, genetic processivity describes the outcome of TET-mediated oxidation *in vivo* and is determined by various factors, including enzymatic activity and local chromatin environment.

AP lyase
An enzyme that is capable of cleaving the 3′ side of an abasic (apurinic or apyrimidinic (AP)) site to create a 3′-terminal unsaturated sugar and 5′-deoxyribosephosphate. This activity generates a single-strand break to initiate the downstream steps of base excision repair.

in the genome, as shown by mapping of the oxidized bases (FIG. 2c). In mouse ESCs, many genomic regions or CpG sites are modified by 5hmC but not 5fC or 5caC, whereas many others are modified by 5fC or 5caC but not 5hmC, suggesting that 5mC is processed to different states at different genomic regions or CpG sites. Genetic processivity positively correlates with chromatin accessibility and binding of certain transcription factors, but the causal relationship requires further examination[41–44].

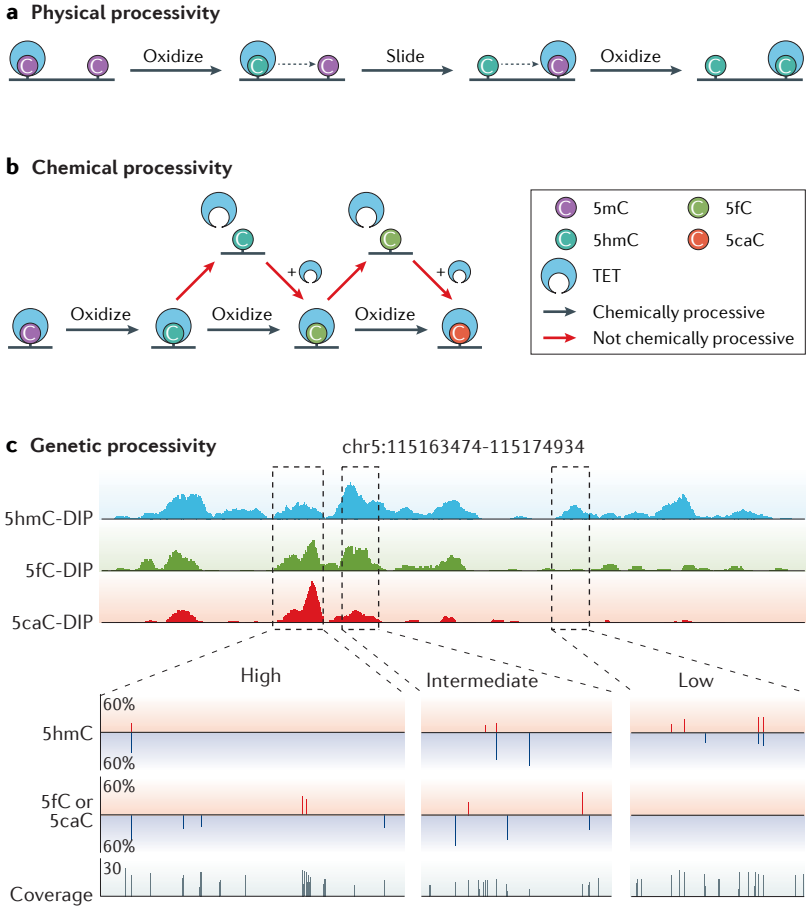


Figure 2 | **Processivity of TET at three different levels. a** | Physical processivity of TET. Physical processivity refers to the ability of TET to slide along DNA from one CpG site to another. An *in vitro* study suggests that TET is not physically processive[40]. **b** | Chemical processivity of TET. Chemical processivity refers to the ability of TET to stay bound to its substrate or oxidation product during the step-wise oxidation from 5-methylcytosine (5mC) to 5-carboxylcytosine (5caC). **c** | Genetic processivity of TET. Genetic processivity refers to the genetic outcome of TET-mediated oxidation, which is determined by various factors, including the intrinsic property of the enzyme, modulation by other factors and the local chromatin environment. The distinct genetic processivity is observed at genomic regions and individual CpG sites, with certain regions or CpG sites being oxidized only to 5-hydroxymethylcytosine (5hmC), whereas some others are further oxidized to 5-formylcytosine (5fC) or 5caC[41–44]. The locus shown here provides three examples of low, intermediate and high genetic processivity revealed by non-base-resolution methods (DNA-immunoprecipitation sequencing (DIP-seq) of 5hmC, 5fC and 5caC of thymine DNA glycosylase (TDG)-depleted mouse embryonic stem cells (ESCs)[43]; upper panel) and base-resolution methods (TET-assisted bisulfite sequencing (TAB-seq) of mouse ESCs[117] and M.SssI methylase-assisted bisulfite sequencing (MAB-seq) of TDG-depleted mouse ESCs[41]; lower panels). Red versus blue lines indicate on which DNA strand the modifications occurred.

Restoration of unmodified cytosine through the TDG–BER pathway. After 5mC is oxidized to 5fC or 5caC, TDG-mediated excision of 5fC or 5caC and BER-dependent repair of the abasic site can restore unmodified cytosine[24,27,28]. This process is defined as active modification–active removal (AM–AR) and is independent of DNA replication[45] (FIG. 1a).

In vitro biochemical assays have shown that TDG specifically excises 5fC and 5caC but not 5hmC[24,27,46,47]. This specificity may be conferred by different mechanisms, including specific TDG–5fC or TDG–5caC interactions, altered C-G base pairing and altered base–sugar bonding[28,46–50]. Like TET, TDG recognizes different forms of substrate at a CpG dyad: for example, 5caC paired with unmodified cytosine, 5mC, 5hmC or 5caC[28] (FIG. 1c).

Recent biochemical reconstruction using purified TDG and BER proteins demonstrated that after TDG-dependent excision of 5fC or 5caC, an abasic site is generated, which can be converted to a single-strand break (SSB) through AP endonuclease 1 (APE1)-mediated incision. Addition of DNA polymerase β (Pol β) inserts a deoxycytidine monophosphate at the break, and further addition of XRCC1 and DNA ligase 3 (LIG3) ligates the nick to restore double-stranded DNA[28]. In addition to APE1, NEIL1 (also known as endonuclease 8-like 1) and NEIL2 (also known as endonuclease 8-like 2), two glycosylases with AP lyase activity, can facilitate the restoration of unmodified cytosine by displacing TDG from the abasic site to create an SSB for downstream processing[51].

In the case when both the top and the bottom strands of a CpG dyad are modified as 5fC or 5caC, it is possible that two SSBs are introduced at the same time, creating a double-strand break (DSB) that might lead to catastrophic outcomes (FIG. 1c). *In vitro* analysis mixing 5caC:5caC substrate with purified TDG and BER proteins (APE1, Pol β, XRCC1 and LIG3) showed that a DSB is generated in less than 1% of this substrate, suggesting that TDG and BER are efficiently coupled to demethylate one strand at a time[28]. In addition to TDG–BER coupling, the protein–protein interaction between TET and TDG may facilitate the quick removal and repair of 5fC or 5caC on generation, minimizing the possibility of DSB generation[28,52,53].

Restoration of unmodified cytosine through replication-dependent dilution of oxidized 5mC. In addition to AM–AR, DNA replication can lead to the dilution of the oxidized 5mC, a process known as active modification–passive dilution (AM–PD)[45] (FIG. 1a). During DNA replication, unmodified cytosine is incorporated into the newly synthesized strand, creating hemi-modified CpG dyads. A 5mC:C dyad is recognized by UHRF1, which helps to recruit DNMT1 to the hemi-5mC site[6,7]. Two studies suggested that UHRF1 has less affinity for 5hmC:C[54,55], but an earlier study provided a conflicting result[56]. Nevertheless, *in vitro* biochemical assays suggest that DNMT1 is much less efficient at 5hmC:C, 5fC:C and 5caC:C dyads than at a 5mC:C dyad[54,55,57]. Through multiple rounds of DNA replication, a 5hmC-, 5fC- or 5caC-modified CpG site can become demethylated.

Regulation of TET-mediated DNA demethylation

Active DNA demethylation can be regulated at various levels. The kinetics of an enzymatic reaction is directly affected by the availability of substrates. An enzymatic reaction can also be modulated by cofactors. Furthermore, all genes involved can be regulated at transcriptional, post-transcriptional and post-translational levels. Finally, factors targeting the demethylation machinery to specific genomic regions can also regulate the process.

Regulation by the availability of substrates and cofactors. TET-mediated oxidation reactions require oxygen and α-KG as substrates and Fe(II) as a cofactor to generate CO_2 and succinate[45,58]. As a result, the availability of the substrates and cofactors can directly affect the reaction kinetics.

α-KG is generated from isocitrate through the activity of the enzymes isocitrate dehydrogenase 1 (IDH1), IDH2 and IDH3 (REF. 59). Overexpression of IDH1 or IDH2 facilitates 5hmC generation in cells[60,61]. By contrast, downregulation of IDH2, as observed in melanoma, is associated with decreased levels of 5hmC[61]. In addition, cancer-associated IDH mutants can inhibit TET activity by producing 2-hydroxyglutarate (2HG), an oncometabolite that competes with α-KG for TET binding[60,62]. In a similar way to 2HG, fumarate and succinate — which can accumulate in cancer owing to deficiency in fumarate hydratase (FH) and succinate dehydrogenase (SDH), respectively — compete with α-KG to inhibit TET activity[63,64]. In the liver of mice, increased α-KG levels resulting from the administration of glucose, glutamate or glutamine correlate with a rapid increase in 5hmC levels[65].

In addition to α-KG, oxygen is another substrate of TET-mediated oxidation. The effect of oxygen on TET-mediated oxidation has been quantified *in vitro*[64], but the outcomes in cells and animals can vary[64,66–68]. In response to hypoxia, certain cell types display an increased 5hmC level, an effect caused by hypoxia-inducible factor (HIF)-mediated upregulation of TET[66–68]. Interestingly, in other cell types, hypoxia reduces 5hmC levels without downregulating TET and is independent of changes in reactive oxygen species (ROS) production, cell proliferation and metabolite concentrations, suggesting a direct regulation by oxygen availability. The association between hypoxia and 5hmC loss is also observed in samples from people with glioblastoma and has been validated in a mouse breast tumour model[68].

As a cofactor of the reaction, Fe(II) availability also influences TET activity. Modulation of cellular iron concentration alters the level of 5hmC[69]. Mutations of critical iron-binding residues of TET reduce its catalytic activity[64].

Vitamin C has been reported to stimulate the enzymatic activity of TET, most likely through acting as a cofactor[70–72]. Mechanistically, vitamin C directly interacts with the catalytic domain of TET proteins to increase their enzymatic activity[70,72]. Additionally, vitamin C may promote TET folding to facilitate the recycling of Fe(II)[72].

Post-transcriptional and post-translational regulation. Following transcription, mRNAs of *TET* and *TDG* can be regulated by microRNAs (miRNAs). Examples include miR-15b[73], miR-22 (REFS 74,75), miR-26 (REFS 76,77), miR-29 (REFS 77–81), miR-125 (REF. 77), miR-494 (REF. 82) and miR-302/367 (REF. 83) for TET, and miR-26a[76] and miR-29 (REFS 78–80) for TDG. The RNA-binding protein deleted in azoospermia-like (DAZL) has also been reported to facilitate *Tet1* translation by binding to *Tet1* mRNA[84].

After translation, the subcellular localization, chromatin binding and enzymatic activity of TET proteins can be regulated by covalent modifications. For example, monoubiquitylation of a conserved lysine residue of TET family members facilitates their chromatin binding[85]. Additionally, acetylation of two conserved lysine residues at the N terminus of human TET2 increases its enzymatic activity, stabilizes the protein and enhances its chromatin targeting during oxidative stress[52]. Furthermore, phosphorylation, GlcNAcylation and PARylation of TET proteins have also been reported[86–90].

The protein levels of TET proteins can also be regulated by protein–protein interaction and proteolysis. Overexpression of IDAX, the interacting partner of TET2, results in caspase-dependent degradation of TET2, whereas its depletion increases TET2 protein levels[32]. Calpain-mediated proteolysis regulates TET protein levels during the maintenance and differentiation of mouse ESCs[91], and the ubiquitin–proteasome pathway regulates TET2 degradation in human cancer cell lines[52].

Regulation of genomic localization. To selectively demethylate a group of CpG sites, the TET and TDG machinery needs to be localized to the corresponding genomic regions, in a process coordinated by intrinsic properties of TET and TDG, the local chromatin environment and extrinsic factors.

Chromatin immunoprecipitation followed by sequencing (ChIP–seq) analysis in mouse ESCs suggests that TET1 is enriched at CpG islands, active promoters and bivalent promoters (marked by both histone H3 lysine 4 trimethylation (H3K4me3) and H3K27me3). TET1 occupancy positively correlates with CpG density and H3K4me3 levels[92–94]. This binding preference may be partially explained by the CXXC domain of TET, which prefers CpG-rich regions[93,95]. In the case of TET2 without a CXXC domain, the chromatin recruitment may be partially mediated by its interacting partner IDAX[32]. Interestingly, TET1s, despite lacking the CXXC domain and the N terminus, distributes across the genome in a manner similar to full-length TET1e, although with less enrichment, suggesting that the N terminus and the CXXC domain facilitate the binding of TET1 but do not fully determine its genomic localization[33]. Thus, the genomic distribution of TET proteins may be regulated by additional factors, such as the local chromatin environment.

Interacting partners of TET proteins may also contribute to their recruitment to specific genomic regions, as shown by the following examples. In mouse ESCs, the pluripotency factor NANOG physically

CpG islands
Genomic regions with a high density of CpG dinucleotides.

Bivalent promoters
Promoters that are enriched for both the active mark histone H3 lysine 4 trimethylation (H3K4me3) and the repressive mark H3K27me3.

Box 1 | **Methods for mapping the genomic distribution of oxidized 5mC**

Various methods have been developed to map the genomic distribution of oxidized 5-methylcytosine (5mC) (TABLE 1; also reviewed in REFS 12,247,248). These methods can be separated into base-resolution and non-base-resolution categories. TABLE 1 lists the principles and features of these mapping techniques.

Non-base-resolution methods generally involve enrichment of oxidized 5mC followed by high-throughput sequencing. In the case of DNA immunoprecipitation (DIP), 5-hydroxymethylcytosine (5hmC)[94,118,122,124,137], 5-formylcytosine (5fC)[43] and 5-carboxylcytosine (5caC)[43] can be enriched using antibodies directly recognizing these modifications. In other cases — for example, cytosine 5-methylenesulfonate sequencing (CMS-seq; for 5hmC)[121], glucosylation, periodate oxidation and biotinylation sequencing (GLIB-seq; for 5hmC)[121], selective chemical labelling (hMe-Seal for 5hmC[132], and fC-Seal for 5fC[128]) and 5fC DNA pulldown (5fC-DP; for 5fC)[129,142] — oxidized 5mC can be converted to other forms that allow enrichment. These methods capture DNA fragments containing the oxidized 5mC but do not determine where in the fragment the modification is; hence, they provide low spatial resolution. Other caveats and disadvantages include potential contaminants of pulldown, CpG-density-related bias and lack of absolute quantification.

Most base-resolution methods are based on bisulfite sequencing (BS-seq). After sodium bisulfite treatment, 5mC and 5hmC are read as cytosine, whereas unmodified cytosine, 5fC and 5caC are read as thymine[24,119,249]. Altering the behaviour of these different forms of cytosine during bisulfite conversion allows the decoding of individual modifications, either directly, as in the case of TET-assisted BS-seq (TAB-seq; for 5hmC)[117], M.SssI methylase-assisted BS-seq (MAB-seq; for 5fC + 5caC)[41,107,159,207] and a sodium-borohydride-reduced derivate of MAB-seq (caMAB-seq; for 5caC)[41], or indirectly after comparison with regular BS-seq data, as in the case of oxidized BS-seq (oxBS-seq; for 5mC and 5hmC)[119], reduced BS-seq (redBS-seq; for 5fC)[250], 5fC chemical-modification-assisted bisulfite sequencing (fCAB-seq; for 5fC)[128,130] and caCAB-seq (for 5caC)[130,251]. Cyclization-enabled C-to-T transition of 5fC (fC-CET) and chemical-labelling-enabled C-to-T conversion sequencing (CLEVER-seq), two bisulfite-free methods that directly map 5fC, have also been established, in which 5fC, after its reaction with 1,3-indandione or malononitrile, respectively, behaves as thymine instead of cytosine during PCR amplification[42,163]. A similar strategy may also allow bisulfite-free mapping of 5hmC[252]. These methods can quantify the absolute levels of the modifications and avoid some of the biases of non-base-resolution methods. One disadvantage of base-resolution methods is their relatively high sequencing cost when the whole genome is analysed without pre-enrichment.

Other than the base-resolution methods mentioned above, enzymes whose cutting activity depends on cytosine modification status have also been applied to map oxidized 5mC at base resolution[44,138,253–255]. Depending on the methods, the limitations may include sequence dependency or lack of quantification.

Recently, some of the methods have been developed to be applicable to small numbers of cells or even single cells. Examples include single-cell application of the DNA-modification-dependent restriction endonuclease AbaSI coupled with sequencing (scAba-seq)[256] and Nano-hmC-Seal[257] for mapping 5hmC, single-cell MAB-seq (scMAB-seq) and low-input MAB-seq (liMAB-seq)[168] for mapping 5fC and 5caC, and CLEVER-seq[163] for mapping 5fC. Given that active DNA demethylation occurs in many biological contexts that involve limited cell numbers (for example, in zygotes and primordial germ cells) and that cell-to-cell heterogeneity of active DNA demethylation exists, these low-input or single-cell methods will broaden the scope of genomic analysis of oxidized 5mC.

interacts with TET1 and TET2, and NANOG depletion results in reduced TET1 binding at NANOG-bound regions[96]. Similarly, PR domain zinc finger protein 14 (PRDM14)[97], Polycomb repressive complex 2 (PRC2)[98] and LIN28A[99] have also been reported to interact with and recruit TET proteins in mouse ESCs. In acute myeloid leukaemia (AML) cells, the transcription factor WT1 recruits TET2 to WT1 target genes through their physical interaction[100,101]. During monocyte-to-osteoclast differentiation, PU.1 interacts with TET2 and facilitates TET2 recruitment to PU.1 target genes for demethylation[102]. In the mouse retina, RE1-silencing transcription factor (REST) has been identified as an interaction partner of TET3 and may facilitate the recruitment of TET3 to REST target genes[103]. During 3T3-L1 fibroblast-to-adipocyte *trans*-differentiation, TET-mediated DNA demethylation takes place around binding sites of peroxisome proliferator-activated receptor-γ (PPARγ) and CCCTC-binding factor (CTCF)[104–106], which might be mediated through the interaction between TET and PPARγ[105] or TET and CTCF[104]. Collectively, these results show or imply that the interacting partners of TET, in many cases key transcription factors of the cells studied, contribute to TET recruitment. In some of the studies, further analysis is needed to determine whether the interaction per se mediates the recruitment or instead the interacting partner helps to establish a favourable chromatin environment for TET binding.

In mouse ESCs, TDG is distributed in a similar way to TET1 and is enriched at active promoters and enhancers[107]. This distribution pattern can be explained by the physical interaction between TET and TDG[28,52,53], or the recruitment of TDG by the oxidation products 5fC and 5caC[108,109], or both. Other interacting factors such as growth arrest and DNA damage-inducible protein GADD45A and oestrogen receptor-β (ERβ) may also contribute to the genomic localization of TDG[110,111].

Distribution and dynamics of oxidized 5mC

To identify the biological contexts of active DNA demethylation, various studies based on mass spectrometry have quantified the amount of 5hmC, 5fC and 5caC in different tissues. Moving one step further, various sequencing techniques, including base-resolution and non-base-resolution methods, have been developed to examine the genomic distribution of oxidized 5mC (BOX 1; TABLE 1). Finally, efforts based on genetic, chemical and biochemical approaches have helped to determine whether the observed oxidized 5mC represents ongoing demethylation dynamics or products accumulated from previous oxidation reactions. In this section, we summarize the tissue distribution, genomic distribution and dynamics of oxidized 5mC with an emphasis on ESCs and neurons. We also discuss the potential biological functions of oxidized 5mC.

Tissue distribution of oxidized 5mC. Mass spectrometry analyses suggest that unlike 5mC, the levels of oxidized 5mC are highly variable in different tissues[25,112–116]. In adult mice, 5hmC is present at high levels in the central nervous system (CNS)[16,25,115,116]. In mouse cerebellar Purkinje neurons, 5hmC abundance is nearly 40% of that of 5mC[16]. Some somatic tissues such as the kidney and heart have medium levels of 5hmC (25–50% of that of CNS tissues), whereas some others, such as the spleen and thymus, have low levels of 5hmC (5–15% of that of CNS tissues)[25,114,115]. 5hmC abundance seems to be anti-correlated with cell proliferation[114]. 5hmC is also present in embryonic tissues. For example, in mouse ESCs, the amount of 5hmC is about 1.3×10^3 in every 10^6 cytosines, a level comparable to that of non-CNS somatic tissues[25].

Table 1 | **Sequencing techniques for genome-scale mapping of oxidized 5mC**

Name	Principle	Base resolution	Absolute level	Refs
5hmC mapping				
5hmC-DIP	Enrich 5hmC-modified DNA using a 5hmC-targeted antibody	No	No	94,118,122, 124,137
hMe-Seal	Convert 5hmC to N_3-5gmC with β-GT. Label N_3-5gmC with biotin. Enrichment	No	No	132,257
CMS-seq	Convert 5hmC to CMS by bisulfite conversion. Enrich CMS-containing DNA using a CMS-targeted antibody	No	No	121
GLIB-seq	Convert 5hmC to 5gmC with β-GT. Label 5gmC with biotin through sodium periodate and ARP treatment. Enrichment	No	No	121
TAB-seq	Convert 5hmC to 5gmC with β-GT. Oxidize 5mC and 5fC to 5caC by TET treatment. Directly read 5hmC out as C after bisulfite conversion	Yes	Yes	117
oxBS-seq	Convert 5hmC to 5fC with potassium perruthenate. Sequence C + 5hmC + 5fC + 5caC as T after bisulfite conversion. Quantify 5hmC by subtracting traditional BS-seq signal (C + 5fC + 5caC)	Yes	Yes	119
RRHP	Digest DNA with MspI (which cuts both 5mC and 5hmC in a CCGG context). Add adaptors. Convert 5hmC to 5gmC with β-GT. Digest again with MspI (DNA fragments with 5gmC at adaptor–DNA junctions will not be digested). Amplify adaptor-tagged fragments to detect 5hmC in a CCGG context	Yes	No	253
Aba-seq	Convert 5hmC to 5gmC with β-GT. Treat DNA with AbaSI, which preferentially recognizes and cuts 5gmC in a sequence context where two cytosines are positioned symmetrically around the cleavage site. Adaptor ligation and amplification. Identify 5hmC-modified sites	Yes	No	126,256
Pvu-Seal-seq	Digest DNA with PvuRts1I, which cuts 5hmC with high preference. Add adaptors and further enrich 5hmC-modified fragments with hMe-Seal. Identify 5hmC-modified sites	Yes	No	44
SCL-exo	Follow the procedures of hMe-Seal to convert 5hmC to biotin-conjugated 5gmC for enrichment. Digest the enriched DNA with exonuclease, which preferentially stops at biotin-conjugated 5gmC. Identify 5hmC-modified sites	Yes	No	255
5fC or 5caC mapping				
5fC-DIP and 5caC-DIP	Enrich 5fC- or 5caC-modified DNA through 5fC- and 5caC-targeted antibodies	No	No	43
fC-Seal	Convert 5hmC to 5gmC with β-GT. Convert 5fC to 5hmC by sodium borohydride-mediated reduction. Follow the procedures of hMe-Seal to enrich DNA originally modified by 5fC	No	No	128
5fC-DP	Label 5fC with biotin through reaction with ARP. Enrichment	No	No	129,142
MAB-seq	Convert unmodified C in a CpG context to 5mC by M.SssI treatment. Directly read 5fC + 5caC out as T after bisulfite conversion	Yes	Yes	41,107,159, 168,207
caMAB-seq	Convert 5fC to 5hmC with sodium borohydride. Convert unmodified C in CpG context to 5mC with M.SssI treatment. Directly read 5caC out as T after bisulfite conversion	Yes	Yes	41
redBS-seq	Convert 5fC to 5hmC by sodium borohydride. Sequence 5mC + 5hmC + 5fC as C after bisulfite conversion. Quantify 5fC after subtracting traditional BS-seq signal (5mC + 5hmC)	Yes	Yes	250
fCAB-seq	Treat DNA with $EtONH_2$, which reacts with 5fC. Sequence 5mC + 5hmC + 5fC as C after bisulfite conversion. Quantify 5fC after subtracting traditional BS-seq signal (5mC + 5hmC)	Yes	Yes	128,130
caCAB-seq	Treat DNA with EDC, which reacts with 5caC. Sequence 5mC + 5hmC + 5caC as C after bisulfite conversion. Quantify 5caC after subtracting traditional BS-seq signal (5mC + 5hmC)	Yes	Yes	130,251
fC-CET	Treat DNA with an azido derivative of 1,3-indandione to label 5fC. Conjugate with biotin. Enrich 5fC-modified DNA. Directly read 5fC out as C-to-T transitions after PCR amplification	Yes	No	42
CLEVER-seq	Treat DNA with malononitrile to label 5fC. MALBAC-based amplification. Directly read 5fC out as C-to-T transition after PCR amplification	Yes	Yes	163
Modified Pvu-Seal-seq	Convert 5hmC to 5gmC with β-GT. Convert 5fC to 5hmC with sodium borohydride-mediated reduction. Follow the procedures of Pvu-Seal-seq to identify 5fC-modified sites	Yes	No	44

Techniques are categorized based on the modifications profiled. For each method, we describe its principle, whether it is base resolution and whether it allows quantification of the absolute level of the modification. 5caC, 5-carboxylcytosine; 5fC, 5-formylcytosine; 5gmC, β-glycosyl-5-hydroxymethylcytosine; 5hmC, 5-hydroxymethylcytosine; 5mC, 5-methylcytosine; β-GT, β-glucosyltransferase; Aba-seq, DNA-modification-dependent restriction endonuclease AbaSI coupled with sequencing; ARP, aldehyde-reactive probe; BS-seq, bisulfite sequencing; CAB-seq, chemical-modification-assisted bisulfite sequencing; caMAB-seq, a sodium-borohydride-reduced derivate of MAB-seq; CET, cyclization-enabled C-to-T transition of 5fC; CLEVER-seq, chemical-labelling-enabled C-to-T conversion sequencing; CMS, cytosine 5-methylenesulfonate; DIP, DNA immunoprecipitation; DP, DNA pulldown; EDC, 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride; fC-Seal, 5fC selective chemical labelling; GLIB, glucosylation, periodate oxidation and biotinylation; hMe-Seal, 5hmC selective chemical labelling; MAB-seq, M.SssI methylase-assisted bisulfite sequencing; MALBAC, multiple annealing and looping-based amplification cycles; N3-5gmC, N_3-β-glycosyl-5-hydroxymethylcytosine; oxBS-seq, oxidized BS-seq; redBS-seq, reduced BS-seq; RRHP, reduced-representation 5hmC profiling; SCL, selective chemical labelling; TAB-seq, TET-assisted bisulfite sequencing.

5fC and 5caC are much less abundant than 5hmC, either because TDG is highly efficient at removing these two bases or because the conversion of 5hmC to 5fC and 5caC is less efficient, or both. In wild-type mouse ESCs, the amount of 5fC and 5caC is about 20 and 3 in every 10^6 cytosines, respectively[25]. Consistent with efficient removal by TDG, depletion of TDG in mouse ESCs results in a 5.6-fold increase in 5fC levels and an 8.4-fold increase in 5caC levels[43]. Beyond ESCs, 5fC can be readily detected in various somatic tissues in postnatal mice[25,113].

For individual tissues, the levels of 5hmC, 5fC and 5caC are not obviously correlated[25,113]. For example, although 5hmC is more abundant in the mouse brain cortex than in ESCs, the levels of 5fC and 5caC are reversed in the two cell types[25]. This complex pattern suggests that different steps of the demethylation cycle are differentially regulated in different tissues.

Genomic distribution of oxidized 5mC in ESCs. Various genome-wide profiling studies have been performed in ESCs because all of the components of active DNA demethylation (TET, TDG and BER) are functional, and all three oxidized forms of 5mC are detectable in this cell type[12].

In ESCs, both enrichment-based analyses and base-resolution analyses (BOX 1; TABLE 1) suggest that 5hmC levels are low at promoters with high CpG density, which are typically marked by H3K4me3 (REFS 93,94,117–119). This observation, in contrast to the high TET binding at these promoters, as discussed above, is not unexpected given that 5mC, the substrate of oxidation, is largely absent. Comparatively, promoters with low-to-intermediate CpG density have higher levels of 5hmC[93,117–120]. 5hmC is low at promoters of highly expressed genes, consistent with the high CpG density of these promoters[93,94,118,121,122]. Interestingly, a high level of 5hmC is observed at bivalent promoters, which correspond to genes repressed in ESCs but activated on differentiation[93,117,118,121].

In gene bodies, 5hmC displays an increase from transcription start sites (TSSs) to transcription termination sites (TTSs), implying a potential coupling between 5mC oxidation and transcription elongation[44,93,118]. 5hmC is also enriched at distal regulatory elements such as enhancers, H3K4me1-marked regions, DNase hypersensitivity sites and transcription-factor-bound regions[44,117,118,121–126]. These regions typically have low-to-intermediate CpG density and intermediate levels of DNA methylation[127]. Notably, base-resolution analysis further shows that 5hmC levels are relatively low at the transcription factor binding sites (±100 bp) but high at flanking regions[117]. Targeting demethylation activity to distal regulatory elements, which usually have fewer TET ChIP–seq signals than promoters and CpG-rich regions[33,125], may result from direct recruitment by transcription factors, as discussed above, or from enhancer–promoter looping.

TDG depletion in mouse ESCs through knockdown or knockout results in accumulation of 5fC and 5caC[24,43,128,129]. 5fC and 5caC accumulation in TDG-depleted cells occurs preferentially at distal regulatory elements and bivalent promoters, similar to the enrichment of 5hmC at these regions[41,43,128,130]. Certain CpG sites and genomic regions, such as regions bound by the pluripotency transcription factors OCT4 (also known as POU5F1) and SOX2, have relatively higher 5fC and 5caC levels but lower 5hmC levels, reflecting a higher genetic processivity at these regions[41–44].

Genomic distribution of oxidized 5mC in neurons. Multiple studies have profiled 5hmC in neuronal samples or purified neurons. Unlike ESCs, replication-dependent dilution does not take place in neurons, as these cells are postmitotic[12].

In adult neuronal samples or purified neurons, 5hmC is generally depleted from TSSs regardless of gene expression level or CpG content, a situation that differs from that in ESCs, in which 5hmC can accumulate at the promoters of poised or repressed genes[131–134]. Comparatively, enrichment is observed at around 875 bp upstream of TSSs and around 160–200 bp downstream of TTSs[131–133]. In gene bodies, 5hmC is highly enriched and positively correlates with gene expression[131–137]. 5hmC also peaks at exon–intron boundaries, suggesting a potential connection with splicing[134,138]. Interestingly, 5hmC is slightly but significantly higher on the sense strand, implying an association with transcription[134]. Finally, in a pattern similar to that in ESCs, 5hmC is observed around enhancer regions, suggesting that active DNA demethylation may have regulatory roles[134,136].

When neuronal samples of different ages are compared, the overall 5hmC level is positively correlated with age[131,132,134,139]. The changes at individual genomic regions are more dynamic: some regions gain 5hmC on maturing or ageing, whereas others lose the mark[131,139], although the functional importance of 5hmC dynamics during ageing is not fully understood.

Dynamics of oxidized 5mC. Mapping the genomic location of oxidized 5mC may not fully reflect the demethylation dynamics owing to the following two reasons. First, the absence of oxidized 5mC does not preclude the possibility that the demethylation process is highly efficient, leaving no trace of intermediates. Second, for postmitotic cells such as neurons, the presence of oxidized 5mC may reflect oxidation events in the past (for example, during neuronal differentiation) instead of real-time dynamics in the cells. In certain cases — for example, development — changes in 5mC and oxidized 5mC can be examined by comparing cells of different developmental stages. In other cases, when the cells are not undergoing cell fate transition, a comparison might be difficult. In these cases, genetic perturbation of the key factors of the pathway or tracing the turnover of oxidized 5mC through isotope labelling is needed.

Genetic perturbation of the TET–TDG–BER pathway can be achieved through depleting TET or TDG. An increase in 5mC or decrease in 5hmC levels on TET depletion through knockout or knockdown, as shown in ESCs, can indicate that active turnover is taking place[92,93,124,140,141]. A caveat of this interpretation is that

an increase in 5mC levels may not be solely due to the loss of TET catalytic activity. For example, it is possible that the physical presence of TET proteins, especially at methylation-free promoters, prevents 5mC deposition by DNMTs. This possibility, along with other catalytic-independent roles of TET, is discussed in BOX 2. Alternatively, the accumulation of 5fC or 5caC on TDG depletion can support the existence of ongoing TET-mediated oxidation and TDG-mediated excision[24,43,128,129,142,143]. The limitation is that in mitotic cells, accumulation of 5fC or 5caC on TDG depletion reflects AM–AR but does not fully reflect the degree of AM–PD, in which oxidation to 5hmC coupled with replication can complete the demethylation cycle independently of 5fC and 5caC generation and excision.

In addition to genetic perturbation, isotope labelling has been applied to trace the turnover of oxidized 5mC. In two studies, mice were fed with food containing isotope-labelled L-methionine, which is converted to the methyl donor *S*-adenosyl methionine (SAM) in cells, leading to the labelling of 5mC by DNMTs. For adult mice fed by labelled food, the proportion of 5mC being labelled is greater than that of 5hmC and 5fC in all tissues examined. If all 5hmC and 5fC undergo active turnover during the feeding period, these two modifications should have been labelled in the same proportions as 5mC. Therefore, these observations imply that a proportion of the existing 5hmC and 5fC in adult tissues is stable, without active turnover[113,114].

Potential functions of oxidized 5mC. Because 5mC oxidation derivatives can accumulate to a high level (for example, 5hmC in the brain) or exist in a relatively stable state in certain biological contexts, they have been proposed to carry out additional functions besides serving as demethylation intermediates. Several studies have reported the identification of proteins capable of binding to 5hmC, 5fC and 5caC[108,109,133,144–146]. SALL4A has been shown to bind 5hmC and facilitate further oxidation of 5hmC, promoting genetic processivity[144]. In addition to serving as docking sites for protein binding, 5fC and 5caC have been shown to impede transcriptional elongation by RNA polymerase II[147]. 5fC may also affect the structure of the DNA double helix[148]. Given the relatively low abundance of 5fC and 5caC, additional studies are needed to demonstrate their functional importance under biologically relevant conditions.

Functions of active DNA demethylation

Active DNA demethylation occurs in various biological contexts, including pre-implantation and PGC development, ESC maintenance and differentiation, and neuronal functions. Aberrant demethylation is observed in cancer. Recent studies have also revealed the involvement of TET and active DNA demethylation in genomic instability and DNA damage repair, which is discussed in BOX 3.

Active DNA demethylation in pre-implantation development. Shortly after fertilization, mouse and human zygotes undergo extensive epigenetic reprogramming, including global DNA demethylation of both the paternal and maternal genomes[12,149,150]. Current evidence indicates that the demethylation of the paternal genome mainly occurs through a combination of passive dilution of 5mC and TET3-mediated AM–PD of 5hmC, 5fC and 5caC (FIG. 3a), whereas the demethylation of the maternal genome mainly occurs through passive dilution.

Immunostaining reveals that the 5mC signal from the highly methylated paternal genome decreases rapidly after fertilization[14]. This rapid decrease in 5mC is accompanied by the generation of 5hmC, 5fC and 5caC[151–153] and is mediated by TET3 (REFS 151,154,155). Immunostaining further showed that DNA replication results in the dilution of all oxidized forms of 5mC and also 5mC[153,156,157], which is not efficiently maintained in this context, possibly owing to the exclusion of the maintenance machinery from the nucleus[158]. TET3-mediated oxidation and DNA replication overlap in timing, but these two processes seem to be independent from each other[151,153,157]. Comparison of the sperm and paternal pronuclei post-replication by BS-seq reveals that the global BS-seq signal (5mC + 5hmC) drops by 40–50% from sperm to post-replication paternal pronuclei. DNA

Box 2 | Catalytic-activity-independent functions of TET proteins

The preferential binding of TET proteins at 5-methylcytosine (5mC)-free promoters and the ability of TET proteins to interact with various proteins both suggest that TET may function independently of its catalytic activity by repelling or recruiting other factors[12]. For example, TET1 can regulate transcription through its association with the SIN3A histone deacetylase complex[94], the MOF histone acetyltransferase (also known as KAT8)[258] and possibly Polycomb repressive complex 2 (PRC2)[92,98], and these associations might not require the catalytic activity of TET1. TET2 can regulate transcription by interacting and recruiting histone deacetylase 2 (HDAC2)[259]. In addition, TET proteins can regulate transcription by recruiting *O*-linked *N*-acetylglucosamine transferase (OGT) in a catalytic-activity-independent manner[120,260,261].

Similar to their complex roles in transcriptional regulation, TET proteins might also affect DNA methylation dynamics in a catalytic-activity-independent manner. Triple knockout (TKO) of TET1, TET2 and TET3 in mouse ESCs results in 5mC accumulation at active promoters where TET binds strongly while oxidized forms of 5mC are present at low levels[140]. It is thus possible that 5mC accumulation at these regions on TET TKO is caused by increased accessibility to DNA methyltransferases rather than decreased TET activity.

Further evidence supporting catalytic-activity-independent functions of TET proteins comes from the demonstration that catalytically dead TET mutants are able to confer biological functions or rescue the phenotype of TET knockout or knockdown. In the mouse hippocampus, overexpression of a catalytically dead TET1 mutant affects learning and memory in a way similar to that of wild-type TET1 (REF. 228). In *Xenopus laevis*, a TET3 catalytically dead mutant can partially rescue the developmental defects caused by TET3 knockout[95]. Various catalytic-activity-independent transcriptional roles of TET proteins have been identified in mouse and human cells: during the inflammation resolution of innate myeloid cells, catalytically dead TET2 represses interleukin-6 transcription by recruiting HDAC2 (REF. 259); in H1299 cells, catalytically dead TET1 regulates the hypoxia response by acting as a transcriptional co-activator[67]; and in 293T cells, overexpression of a catalytically dead TET1 mutant results in transcription changes that are highly similar to wild-type TET1 (REF. 262). Finally, in mast cells, hyperproliferation resulting from TET2 knockout can be rescued by a catalytically dead mutant[213]. In many of these scenarios, the catalytic-activity-dependent and -independent functions probably coordinate to reinforce the functional outcome.

In the future, it will be interesting to determine whether the reported TET knockdown or knockout phenotypes can be rescued by catalytically dead mutants. Given the convenience of current genome editing techniques, it will also be interesting to introduce catalytic mutations at endogenous *Tet* loci to examine the catalytic-activity-dependent effects.

Box 3 | **TET proteins and active DNA demethylation in DNA repair and genomic instability**

Active DNA demethylation is intrinsically linked to DNA repair, as thymine DNA glycosylase (TDG)-mediated base excision repair (BER) is part of the active modification–active removal (AM–AR) process. In addition, emerging evidence suggests that TETs and active DNA demethylation participate in DNA repair and genomic instability in various ways.

In several biological contexts, TET1 deficiency can lead to defective DNA repair, increased DNA damage and genomic instability. For example, in mouse pachytene and early diplotene stage oocytes, TET1 depletion leads to an accumulation of DNA breaks, marked by γH2AX, and delayed removal of the double-strand break (DSB) repair recombinase RAD51, possibly due to dysregulation of meiosis[186]. In mouse embryonic fibroblasts, TET1 knockout results in increased DSBs and genomic instability, possibly due to the downregulation of DNA-repair-associated genes. Consistent with *in vitro* data, TET1 heterozygous knockout mice are more sensitive to X-ray exposure than are wild-type control mice[258]. In N2a cells treated with the topoisomerase inhibitor etoposide, TET1 deficiency results in increased levels of γH2AX and S15-phosphorylated p53, whereas TET1 overexpression leads to a quicker extinction of these DNA damage markers[90]. Finally, pro-B cells of TET1 knockout mice have increased levels of γH2AX, probably due to the downregulation of DNA repair genes[243].

In addition to TET1, deficiency in TET2 and TET3 can affect DNA repair. In mouse bone marrow and spleen, TET2 and TET3 double knockout (DKO) results in a progressive γH2AX increase. In response to irradiation, TET2 and TET3 DKO cells from these tissues resolve DSBs less efficiently, probably due to a downregulation of DNA repair genes[263].

A loss of all three TET proteins can result in chromosomal abnormalities. In mouse embryonic stem cells (ESCs), TET1, TET2 and TET3 triple knockout (TKO) resulted, in one study, in telomere elongation and increased telomere sister chromatid exchange, probably through ZSCAN4 upregulation[140], whereas in another study it resulted in chromosomal fusion and telomere loss but not telomere elongation[264]. The difference between the two studies may be due to different ZSCAN4 levels in the TKO ESCs generated by different laboratories. In response to replication stress induced by aphidicolin treatment, TET TKO mouse ESCs display more severe chromosomal abnormalities, although the expression of DNA repair genes is largely unchanged[265].

In many of the examples above, TET regulates DNA repair and genomic instability indirectly by modulating the expression of genes that are relevant to these processes. However, a direct involvement of active DNA demethylation in the DNA damage response has been reported by several recent studies. One study showed that 5-hydroxymethylcytosine (5hmC) colocalizes with γH2AX and p53-binding protein 1 (53BP1), either in unchallenged cells or in cells treated with aphidicolin or micro-irradiation[265]. Furthermore, DNA derived from γH2AX-targeted chromatin immunoprecipitation (ChIP) has a higher 5hmC level than that from H2A-targeted ChIP and input DNA, indicating that 5hmC is enriched in the DNA damage sites[265]. In addition, mouse cortical neurons exhibit an increase in 5hmC in response to irradiation, in a process that is dependent on ataxia telangiectasia mutated (ATM) and TET1 (REF. 90). Furthermore, the global 5hmC level was found to increase in the A2780 ovarian cancer cell line in response to oxidative stress, in a process that seems to be mediated by TET2 (REF. 52). In the future, it will be interesting to further explore the biological importance of these observations.

replication inhibition by aphidicolin treatment largely abrogates this drop, whereas depletion of zygotic TET3 through maternal knockout has a much less pronounced effect, indicating that replication-dependent dilution is the driving force for paternal genome demethylation, while TET3-mediated oxidation contributes to some degree[159,160].

The zygotic maternal genome also has detectable immunostaining signals of oxidized 5mC, although the signals are much weaker than those of the paternal genome[151–153]. BS-seq, TET-assisted bisulfite sequencing (TAB-seq) and chemical-labelling-enabled C-to-T conversion sequencing (CLEVER-seq) analyses (BOX 1; TABLE 1) have confirmed that the maternal genome does undergo TET3-mediated oxidation, although to a much lesser degree[159–163]. This difference in oxidation between the two parental genomes is probably due to the protection of the maternal genome by developmental pluripotency-associated protein 3 (DPPA3; also known as PGC7), which is recruited to the maternal genome through H3K9me2 (REFS 151,164,165). One possible molecular mechanism might be that DPPA3 interacts with TET3 to inhibit its enzymatic activity[166].

An unsolved question is whether unmodified cytosine can be actively generated through AM–AR in zygotes. Unlike ESCs and most other cell types, oocytes and zygotes have very low levels of *Tdg* mRNA, implying that restoration of unmodified cytosine through TDG may not occur[167]. In addition, 5fC and 5caC immunostaining signals persist to at least the four-cell stage, indicating that TDG-mediated excision is absent or inefficient[153]. Surprisingly, comparative analyses of changes in 5mC + 5hmC (quantified by BS-seq) and 5fC + 5caC (quantified by M.SssI methylase-assisted bisulfite sequencing (MAB-seq)) revealed a lower-than-expected gain of 5fC + 5caC, indicating that a proportion of 5mC is processed to unmodified cytosine[159,168]. Further supporting the existence of AM–AR, BER activation and SSB generation accompanying TET3-mediated oxidation have been reported and seem to be enriched on the paternal genome[169–171]. Given that *Tdg* mRNA is expressed at a negligible level[167] and that maternal TDG knockout did not lead to 5fC + 5caC accumulation at selected genomic loci[159], there might be a TDG-independent and BER-coupled AM–AR mechanism for restoring unmodified cytosine. In addition to a potential TDG-independent mechanism, two immunostaining-based studies suggest that the paternal genome might have undergone a first wave of 5mC loss before DNA replication and TET3-mediated oxidation, implying a possible TET- and replication-independent mechanism for removing 5mC[172,173].

If all modified forms of cytosine including 5mC are diluted by subsequent rounds of DNA replication, why is there the need to oxidize 5mC in zygotes? One explanation is that TET3-mediated oxidation ensures the success

γH2AX
Phosphorylated histone H2AX, a marker of a DNA strand break. At the genomic site of a DNA double-strand break (DSB), histone variant H2AX becomes phosphorylated (at S139 for human H2AX). Although γH2AX is most commonly used as a marker for DSBs, single-strand breaks (SSBs) can also induce γH2AX.

of replication-dependent dilution at its target regions. In fact, certain genomic loci, including imprinted genes, maintain their methylated status, indicating the existence of methylation maintenance machinery[159–162,174]. However, there is no evidence supporting an essential role of TET3-mediated 5mC oxidation in zygotes. Although maternal TET3 knockout results in partial embryonic lethality[155] or neonatal sublethality[175,176], TET3 haploinsufficiency, but not defective 5mC oxidation, seems to be the real cause for these phenotypes[175]. It is possible that the oxidation process may have some non-essential roles, such as facilitating *Oct4* activation[155] or reducing transcriptome variability[177]. Consistent with the notion that 5mC oxidation is not essential, active erasure of 5mC is observed in some but not all mammals[10].

Active DNA demethylation in PGCs. Like preimplantation development, mammalian PGC development also involves global epigenomic reconfiguration, including two stages of DNA demethylation[12,149,150,178]. In mice, PGCs arise at embryonic day 7.25 (E7.25) in the extra-embryonic mesoderm[178]. The first stage of demethylation takes place between E7.25 and E9.5 during PGC proliferation and migration, whereby global 5mC levels decrease through passive dilution. The second stage occurs between E9.5 and E13.5 and involves TET1-mediated 5mC oxidation and passive dilution (FIG. 3b). This process is also largely conserved in humans[179–182].

The global DNA methylation level of mouse E6.5 epiblast cells, from which PGCs originate, is comparable to that of J1 ESCs and E6.5 embryos (~70%)[183,184]. Founder PGCs (E7.25) are thought to have a similar methylation level[183]. At E9.5, the methylation level drops to ~30%, with most genomic regions demethylated[183]. This stage is largely TET-independent because 5hmC is low between E8.5 and E9.5 (REF. 185) and TET deficiency does not markedly affect global demethylation[186,187]. The lack of both *de novo* and maintenance methylation, due to downregulation of UHRF1, DNMT3A and DNMT3B, is probably the major driving force for the loss of DNA methylation at this stage[183,188–191]. Despite this wave of global demethylation, some genomic regions are resistant to demethylation, including imprinting control regions, meiotic gene promoters and particular types of repetitive element[183,186,192,193].

Immunostaining showed that 5hmC signals in PGCs display an increase from E9.5 to E10.5, indicating that TET-mediated oxidation takes place between E9.5 and E10.5 (REF. 185). Gene expression profiling suggests that TET1 and possibly TET2 are responsible for the oxidation[186,189,194]. 5hmC levels peak at around E10.5–E11.5 and gradually decrease in the following days[185,194]. BS-seq and 5mC DNA-immunoprecipitation sequencing (DIP-seq) confirmed the demethylation of germline-specific genes (such as meiotic genes), imprinting control regions and other regions between E9.5 and E13.5 (REFS 183,186,187,189,192–195). These changes in methylation also correlate with changes in gene expression[183,185,186,194]. However, certain genomic elements, such as intracisternal A-particles (IAPs), are still resistant to this wave of demethylation (BS-seq signals >40%)[183,192,193]. The presence of BER machinery and SSB formation at this stage suggests that TDG, which is expressed at this stage[189], may still mediate the active restoration of unmodified cytosine to some degree[194].

For female mice, TET1 deficiency results in meiotic defects of PGCs, probably due to insufficient demethylation and failed activation of meiotic genes[186]. For male mice, TET1 deficiency results in aberrant methylation patterns at imprinted loci in PGCs and sperm cells. When TET1-deficient male mice are crossed with wild-type females, the progeny exhibit a number of phenotypes associated with abnormal imprinting erasure[187]. These observations definitively prove that TET1 mediates imprinting erasure in PGCs and explain the aberrant methylation pattern observed at imprinted loci in TET1 and TET2 double knockout mice[196]. Notably, global methylation in germ cells is not greatly affected by TET1 or combined TET1 and TET2 deficiency, suggesting that passive dilution is the driving force for global demethylation, whereas TET-mediated oxidation has a locus-specific effect[186,187,196].

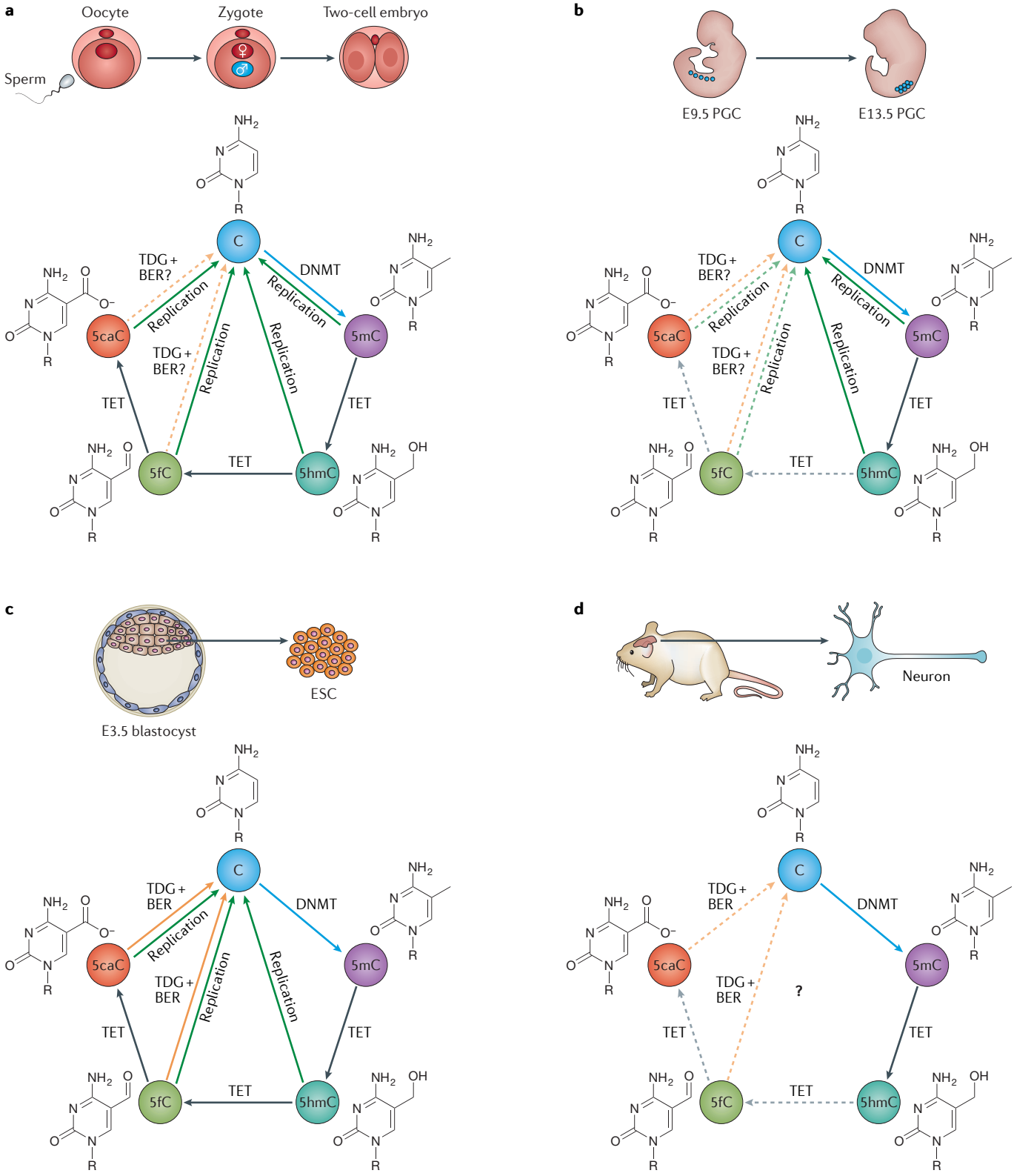
An interesting question is how the imprinted genes and meiotic genes are protected from the first stage of demethylation. DNMT1 conditional knockout in PGCs results in hypomethylation of imprinted loci, meiotic genes and IAPs at E10.5, suggesting that DNMT1 is responsible for keeping these regions methylated in the first stage. The purpose of this mechanism is probably for preventing early activation of germline-specific genes, as DNMT1 deficiency in PGCs results in precocious germline differentiation[197].

Active DNA demethylation in pluripotency and differentiation. Mouse ESCs are pluripotent cells derived from the inner cell mass of the E3.5 blastocyst[198]. When cultured in serum conditions, mouse ESCs express all components of the methylation and demethylation pathways (FIG. 3c) and have detectable levels of all oxidized forms of 5mC[12]. Both TET1 and TET2 are expressed in mouse ESCs. Single knockdown or knockout of TET1 (REFS 23,92–94,118,124,141,199–201) or TET2 (REFS 23,141,196,200,201) in mouse ESCs decreases 5hmC levels and induces transcriptional changes. Comparisons between TET1-depleted and TET2-depleted cells revealed distinct target preferences of these two enzymes, with TET1-preferring promoters and TET2-preferring gene bodies of highly expressed genes and enhancers[141,201]. TET1 deficiency may skew the differentiation of ESCs towards specific lineages[124,199,200], whereas TET2 deficiency represses enhancer activity and delays transcriptional changes during differentiation[141]. Despite these observations, TET1 or TET2 single knockout mice are viable, suggesting that deficiency in either one does not compromise pluripotency and development[199,202–205].

TET1 and TET2 double knockdown or knockout abrogates most 5hmC generation in mouse ESCs, but the cells are still pluripotent, with the ability to form all three germ layers[124,140,196,200]. Although most DKO embryos die perinatally, some develop normally[196]. TET1, TET2 and TET3 triple knockout (TKO) completely abrogates 5hmC in ESCs[140,206,207]. TKO ESCs have

largely normal morphology and express pluripotency markers, but their differentiation and developmental potentials are compromised[206]. Consistent with these developmental impairments, TET TKO embryos display gastrulation defects[208].

Many genomic regions in mouse TET TKO ESCs are hypermethylated in comparison with control cells. Most of the hypermethylated regions are those harbouring oxidized 5mC in TET wild-type cells, supporting the role of TET-mediated oxidation in DNA demethylation.



Penetrance
The proportion of individuals or animals (with a particular genotype) manifesting the phenotype of interest.

Enhancer regions gain the most 5mC in TET TKO cells[140], indicating that they are the major targets of demethylation. At promoter regions, bivalent promoters gain a substantial amount of 5mC, while active and initiated promoters also gain 5mC to a lesser extent. For active or initiated genes, promoter or enhancer hypermethylation correlates with reduced expression, whereas the opposite is true for bivalent genes[140].

TDG deficiency in mouse ESCs does not affect ESC maintenance[24,43,128,129]. TDG knockout is embryonic lethal at around E11.5 for unknown reasons[209,210]. Therefore, AM–AR is not required for ESC maintenance but may be important for differentiation.

Overall, these data suggest that active DNA demethylation is not required for ESC maintenance, but is required for proper differentiation through modulating the activity of enhancers and possibly bivalent promoters. In addition to its role in ESC differentiation, active DNA demethylation regulates somatic cell reprogramming[96,207,211] and other biological processes involving cell fate transitions[105,143,212–218]. In these cases, active DNA demethylation may function similarly by demethylating enhancers and other key regulatory elements, thus allowing crucial transcriptional changes to occur.

Active DNA demethylation in neuronal functions. How active DNA demethylation affects neuronal functions has been a topic of great interest owing to the high abundance of 5hmC in various subtypes of neuron[16,25,114–116] and as a result of earlier research reporting neuronal-activity-induced DNA demethylation[219–225] (FIG. 3d).

Various studies have reported roles for TET proteins in neural functions. For TET1, its deficiency in mice results in impaired adult hippocampal neurogenesis and abnormalities in learning and memory, and synaptic plasticity[226,227], as well as the downregulation of multiple neuronal-activity-induced genes in the hippocampus, possibly owing to promoter hypermethylation[227]. TET1 overexpression in the hippocampus results in transcriptional and functional changes, but the effects may be independent of its enzymatic activity[228]. Repeated cocaine administration downregulates TET1 in the nucleus accumbens and alters 5hmC distribution and gene expression. Interestingly, 5hmC changes at splicing sites positively correlate with the expression changes of the splicing isoforms[229]. For TET3, its overexpression in mouse olfactory sensory neurons alters the levels of gene-body 5hmC and gene expression, and disrupts the axonal targeting of these cells[135]. TET3 depletion in the mouse infralimbic prefrontal cortex impairs fear extinction, whereas fear extinction results in TET3-mediated redistribution of 5hmC[230]. Hippocampal neurons subjected to different types of activity induction can have bidirectional changes in TET3, and altering TET3 or inhibiting BER affects synaptic functions[231]. Furthermore, depletion of both TET1 and TET3 in cerebella slices from postnatal day 7 (P7) mice inhibits dendritic arborization of granule cells[232]. Finally, P25 mice subjected to 3-day monocular deprivation show altered 5mC and 5hmC levels and gene expression in the visual cortex[233].

Although the above observations suggest a role for TET proteins in neuronal function, the interpretation of these data may not be straightforward. First, some observed phenotypes may result from catalytic-activity-independent functions of TET proteins (BOX 2). Second, some phenotypes observed in TET-deficient mice may result from the developmental roles of TET proteins rather than altered functions of mature neurons. Third, the experimental perturbation and methylation analysis may or may not be neuron-specific or neuronal-subtype-specific, depending on the experimental design. In the future, catalytic-activity-specific, developmental-stage-specific and cell-type-specific analyses will help to resolve some of these ambiguities.

Aberrant active DNA demethylation in cancer. Around the time when TET-mediated oxidation was discovered, multiple studies reported TET2 mutations in myeloid disorders, including AML[234–240]. Many of these mutations compromise the enzymatic activity of TET2, suggesting that defects in active DNA demethylation may be a cause of haematopoietic malignancies[241]. Further supporting this notion, TET2 deficiency in mice results in increased self-renewal of haematopoietic stem or progenitor cells, and may eventually lead to malignancy[202–205]. Notably, the low penetrance and long latency for these mice to develop malignancy suggest that additional mutations may be needed (reviewed in REF. 242). In addition to TET2 mutations, TET1 and TET3 mutations were observed in haematopoietic malignancies[236]. TET1 deficiency or TET1 and TET2 double deficiency in mice promotes B cell malignancy[243,244]. In other cases of haematological cancers, mutations in regulators of TET — for example, IDH1, IDH2 and WT1 — also inhibit active DNA demethylation and are thus observed in a mutually exclusive manner with TET mutations[62,100,101]. In AML, mutations in IDH1 or IDH2 produce oncometabolites

◀ Figure 3 | **Mechanism of DNA demethylation in different biological contexts.** **a** | DNA demethylation in the paternal genome of the zygote. In all figure parts, dashed lines denote steps that require more supporting evidence. 5-Methylcytosine (5mC) is oxidized by TET3 to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). DNA replication occurs at around the same time as oxidation and results in the dilution of all modified forms of cytosine, including 5mC. Evidence of single-strand break (SSB) generation, base excision repair (BER) activation and replication-independent restoration of unmodified cytosine was shown despite the negligible level of thymine DNA glycosylase (*Tdg*) transcripts, implying the potential existence of a TDG-independent and BER-coupled mechanism for restoring unmodified cytosine[159,168–171]. **b** | The second stage of DNA demethylation during primordial germ cell (PGC) development. At this stage, TET1 and possibly TET2 oxidize 5mC to 5hmC, and both modifications are diluted by replication. No major change in 5fC or 5caC was detected at this stage, either because further oxidation is limited or because TDG removes the two modifications efficiently[185]. **c** | DNA demethylation in embryonic stem cells (ESCs). In ESCs, all components of the methylation and demethylation machinery are present, which allows both active modification–active removal (AM–AR) and active modification–passive dilution (AM–PD). **d** | Active DNA demethylation in neurons. Neurons express all factors that are needed for methylation and demethylation and they do not undergo DNA replication. Despite high levels of 5hmC, neurons have very low levels of 5fC and 5caC, either because further oxidation is limited or because TDG has efficiently removed these two modifications[25,113]. For mature neurons, further analysis is required to show whether the methylation and demethylation dynamics take place at all (denoted by the question mark in the centre), as an isotope-tracing study suggests that 5hmC is predominantly stable in the brain[114]. DNMT, DNA methyltransferase; E3.5, embryonic day 3.5.

that compete with α-KG, thus inhibiting TET activity[62], whereas mutations in WT1 reduce the genomic recruitment of TET2 (REF. 100).

Aberrant active DNA demethylation, as shown by the reduction of 5hmC levels, is also observed in various solid tumours[61,68,245,246]. In many cases, IDH downregulation, IDH mutations or TET downregulation is the cause of reduced 5hmC levels[61,245]. Interestingly, tumour hypoxia, as discussed above, also inhibits TET-mediated oxidation and contributes to the DNA hypermethylation that is observed in solid tumours without mutations in TET or IDH[68].

Conclusions and perspectives

Despite the great progress in the past few years in understanding the mechanism and function of active DNA demethylation, some questions still remain to be addressed. First, the catalytic-dependent and -independent functions of TET have not been well defined in various biological processes (BOX 2). Second, further investigation is required to determine whether oxidized 5mC, particularly 5hmC, can serve as a docking site for reader proteins to mediate biological functions. Third, certain genomic regions are resistant to TET-mediated oxidation — for example, the imprinted loci during zygotic reprogramming and IAPs during PGC development — and the mechanism for selectively protecting these regions from oxidation is not fully understood. Fourth, given that replication-dependent dilution is the driving force in global demethylation of both pre-implantation and PGC development, the exact roles of TET-mediated oxidation during these processes remain to be determined. Fifth, most published studies have focused on TET but not TDG. However, for cells that are slowly proliferating or postmitotic, AM–AR mediated by TDG is the major or only way to restore unmodified cytosine (assuming no other demethylation pathways exist) and thus requires further examination. Last, despite the high abundance of 5hmC in mature neurons, it is not clear whether, to what extent and where active DNA demethylation takes place. Answering the above questions will deepen our understanding of the mechanism and functions of active DNA demethylation.

1. Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220 (2013).
2. Li, E. & Zhang, Y. DNA methylation in mammals. *Cold Spring Harb. Perspect. Biol.* **6**, a019133 (2014).
3. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
4. 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).
5. Hermann, A., Goyal, R. & Jeltsch, A. The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J. Biol. Chem.* **279**, 48350–48359 (2004).
6. Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760–1764 (2007).
7. Sharif, J. *et al.* The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908–912 (2007).
8. Holliday, R. & Pugh, J. E. DNA modification mechanisms and gene activity during development. *Science* **187**, 226–232 (1975).
9. Riggs, A. D. X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* **14**, 9–25 (1975).
10. Wu, S. C. & Zhang, Y. Active DNA demethylation: many roads lead to Rome. *Nat. Rev. Mol. Cell Biol.* **11**, 607–620 (2010).
11. Bochtler, M., Kolano, A. & Xu, G. L. DNA demethylation pathways: additional players and regulators. *Bioessays* **39**, 1–13 (2017).
12. Wu, H. & Zhang, Y. Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell* **156**, 45–68 (2014).
13. Zhu, J. K. Active DNA demethylation mediated by DNA glycosylases. *Annu. Rev. Genet.* **43**, 143–166 (2009).
14. Mayer, W., Niveleau, A., Walter, J., Fundele, R. & Haaf, T. Demethylation of the zygotic paternal genome. *Nature* **403**, 501–502 (2000).
15. Oswald, J. *et al.* Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10**, 475–478 (2000).
16. Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929–930 (2009).
This paper shows convincingly for the first time that 5hmC is highly abundant in the brain.
17. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
This study reveals that TET1 is capable of converting 5mC to 5hmC.
18. Lorsbach, R. B. *et al.* TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia* **17**, 637–641 (2003).
19. Ono, R. *et al.* LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer Res.* **62**, 4075–4080 (2002).
20. Yu, Z. *et al.* The protein that binds to DNA base J in trypanosomatids has features of a thymidine hydroxylase. *Nucleic Acids Res.* **35**, 2107–2115 (2007).
21. Borst, P. & Sabatini, R. Base J: discovery, biosynthesis, and possible functions. *Annu. Rev. Microbiol.* **62**, 235–251 (2008).
22. Cliffe, L. J. *et al.* JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes. *Nucleic Acids Res.* **37**, 1452–1462 (2009).
23. Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
This paper demonstrates that all three mouse TET proteins are capable of oxidizing 5mC to 5hmC.
24. He, Y. F. *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011).
This work shows that TET converts 5mC and 5hmC to 5caC, which can be excised by TDG. The study also shows that TDG depletion results in 5caC accumulation in mouse ESCs.
25. Ito, S. *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303 (2011).
This paper shows that TET can convert 5mC to 5fC and 5caC, and that 5fC and 5caC are present in mouse ESCs and organs.
26. Smiley, J. A., Kundracik, M., Landfried, D. A., Barnes, V. R. Sr & Axhemi, A. A. Genes of the thymidine salvage pathway: thymine-7-hydroxylase from a Rhodotorula glutinis cDNA library and isoorotate decarboxylase from *Neurospora crassa*. *Biochim. Biophys. Acta* **1723**, 256–264 (2005).
27. 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).
These authors show that TDG is capable of excising both 5fC and 5caC.
28. Weber, A. R. *et al.* Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. *Nat. Commun.* **7**, 10806 (2016).
Using *in vitro* biochemical reconstruction, this paper demonstrates that TET–TDG–BER is sufficient to convert 5mC to unmodified cytosine. It also shows that TET, TDG and BER are tightly coupled to minimize the possibility of DSB formation during demethylation.
29. Pastor, W. A., Aravind, L. & Rao, A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat. Rev. Mol. Cell Biol.* **14**, 341–356 (2013).
30. Hu, L. *et al.* Crystal structure of TET2–DNA complex: insight into TET-mediated 5mC oxidation. *Cell* **155**, 1545–1555 (2013).
This work describes the crystal structure of TET2–5mC, providing insights into the interaction between TET2 and methylated DNA.
31. Zhang, H. *et al.* TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res.* **20**, 1390–1393 (2010).
32. Ko, M. *et al.* Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* **497**, 122–126 (2013).
33. Zhang, W. *et al.* Isoform switch of TET1 regulates DNA demethylation and mouse development. *Mol. Cell* **64**, 1062–1073 (2016).
This paper describes two different isoforms of mouse TET1 and their distinct expression pattern and activity.
34. Liu, N. *et al.* Intrinsic and extrinsic connections of Tet3 dioxygenase with CXXC zinc finger modules. *PLoS ONE* **8**, e62755 (2013).
35. Jin, S. G. *et al.* Tet3 reads 5-carboxylcytosine through its CXXC domain and is a potential guardian against neurodegeneration. *Cell Rep.* **14**, 493–505 (2016).
36. Hashimoto, H. *et al.* Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. *Nature* **506**, 391–395 (2014).
37. Hu, L. *et al.* Structural insight into substrate preference for TET-mediated oxidation. *Nature* **527**, 118–122 (2015).
38. Liu, M. Y. *et al.* Mutations along a TET2 active site scaffold stall oxidation at 5-hydroxymethylcytosine. *Nat. Chem. Biol.* **13**, 181–187 (2016).
39. Crawford, D. J. *et al.* Tet2 catalyzes stepwise 5-methylcytosine oxidation by an iterative and *de novo* mechanism. *J. Am. Chem. Soc.* **138**, 730–733 (2016).
40. Tamanaha, E., Guan, S., Marks, K. & Saleh, L. Distributive processing by the iron(II)/alpha-ketoglutarate-dependent catalytic domains of the TET enzymes is consistent with epigenetic roles for oxidized 5-methylcytosine bases. *J. Am. Chem. Soc.* **138**, 9345–9348 (2016).

41. Wu, H., Wu, X., Shen, L. & Zhang, Y. Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing. *Nat. Biotechnol.* **32**, 1231–1240 (2014).
This study is the first to assess the genetic processivity of TET by comparing 5hmC profiles with 5fC/5caC profiles at base resolution.
42. Xia, B. *et al.* Bisulfite-free, base-resolution analysis of 5-formylcytosine at the genome scale. *Nat. Methods* **12**, 1047–1050 (2015).
43. Shen, L. *et al.* Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell* **153**, 692–706 (2013).
44. Sun, Z. *et al.* A sensitive approach to map genome-wide 5-hydroxymethylcytosine and 5-formylcytosine at single-base resolution. *Mol. Cell* **57**, 750–761 (2015).
45. Kohli, R. M. & Zhang, Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502**, 472–479 (2013).
46. Hashimoto, H., Hong, S., Bhagwat, A. S., Zhang, X. & Cheng, X. Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. *Nucleic Acids Res.* **40**, 10203–10214 (2012).
47. Zhang, L. *et al.* Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nat. Chem. Biol.* **8**, 328–330 (2012).
48. Bennett, M. T. *et al.* Specificity of human thymine DNA glycosylase depends on N-glycosidic bond stability. *J. Am. Chem. Soc.* **128**, 12510–12519 (2006).
49. Malik, S. S., Coey, C. T., Varney, K. M., Pozharski, E. & Drohat, A. C. Thymine DNA glycosylase exhibits negligible affinity for nucleobases that it removes from DNA. *Nucleic Acids Res.* **43**, 9541–9552 (2015).
50. Maiti, A., Michelson, A. Z., Armwood, C. J., Lee, J. K. & Drohat, A. C. Divergent mechanisms for enzymatic excision of 5-formylcytosine and 5-carboxylcytosine from DNA. *J. Am. Chem. Soc.* **135**, 15813–15822 (2013).
51. Schomacher, L. *et al.* Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation. *Nat. Struct. Mol. Biol.* **23**, 116–124 (2016).
52. Zhang, Y. W. *et al.* Acetylation enhances TET2 function in protecting against abnormal DNA methylation during oxidative stress. *Mol. Cell* **65**, 323–335 (2017).
53. Muller, U., Bauer, C., Siegl, M., Rottach, A. & Leonhardt, H. TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation. *Nucleic Acids Res.* **42**, 8592–8604 (2014).
54. Hashimoto, H. *et al.* Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res.* **40**, 4841–4849 (2012).
55. Otani, J. *et al.* Cell cycle-dependent turnover of 5-hydroxymethyl cytosine in mouse embryonic stem cells. *PLoS ONE* **8**, e82961 (2013).
56. Frauer, C. *et al.* Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS ONE* **6**, e21306 (2011).
57. Ji, D., Lin, K., Song, J. & Wang, Y. Effects of Tet-induced oxidation products of 5-methylcytosine on Dnmt1- and DNMT3a-mediated cytosine methylation. *Mol. Biosyst.* **10**, 1749–1752 (2014).
58. Lu, X., Zhao, B. S. & He, C. TET family proteins: oxidation activity, interacting molecules, and functions in diseases. *Chem. Rev.* **115**, 2225–2239 (2015).
59. Losman, J. A. & Kaelin, W. G. Jr. What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. *Genes Dev.* **27**, 836–852 (2013).
60. Xu, W. *et al.* Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell* **19**, 17–30 (2011).
61. Lian, C. G. *et al.* Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell* **150**, 1135–1146 (2012).
62. Figueroa, M. E. *et al.* Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* **18**, 553–567 (2010).
63. Xiao, M. *et al.* Inhibition of alpha-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev.* **26**, 1326–1338 (2012).
64. Laukka, T. *et al.* Fumarate and succinate regulate expression of hypoxia-inducible genes via TET enzymes. *J. Biol. Chem.* **291**, 4256–4265 (2016).
65. Yang, H. *et al.* TET-catalyzed 5-methylcytosine hydroxylation is dynamically regulated by metabolites. *Cell Res.* **24**, 1017–1020 (2014).
66. Mariani, C. J. *et al.* TET1-mediated hydroxymethylation facilitates hypoxic gene induction in neuroblastoma. *Cell Rep.* **7**, 1343–1352 (2014).
67. Tsai, Y. P. *et al.* TET1 regulates hypoxia-induced epithelial–mesenchymal transition by acting as a co-activator. *Genome Biol.* **15**, 513 (2014).
68. Thienpont, B. *et al.* Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature* **537**, 63–68 (2016).
By analysing cell lines, mouse models and patient samples, this paper shows that hypoxia can cause DNA hypermethylation by reducing TET-mediated oxidation.
69. Zhao, B. *et al.* Redox-active quinones induces genome-wide DNA methylation changes by an iron-mediated and Tet-dependent mechanism. *Nucleic Acids Res.* **42**, 1593–1605 (2014).
70. Blaschke, K. *et al.* Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**, 222–226 (2013).
71. Minor, E. A., Court, B. L., Young, J. I. & Wang, G. Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. *J. Biol. Chem.* **288**, 13669–13674 (2013).
72. Yin, R. *et al.* Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J. Am. Chem. Soc.* **135**, 10396–10403 (2013).
73. Lv, X., Jiang, H., Liu, Y., Lei, X. & Jiao, J. MicroRNA-15b promotes neurogenesis and inhibits neural progenitor proliferation by directly repressing TET3 during early neocortical development. *EMBO Rep.* **15**, 1305–1314 (2014).
74. Song, S. J. *et al.* The oncogenic microRNA miR-22 targets the TET2 tumor suppressor to promote hematopoietic stem cell self-renewal and transformation. *Cell Stem Cell* **13**, 87–101 (2013).
75. Song, S. J. *et al.* MicroRNA-antagonism regulates breast cancer stemness and metastasis via TET-family-dependent chromatin remodeling. *Cell* **154**, 311–324 (2013).
76. Fu, X. *et al.* MicroRNA-26a targets ten eleven translocation enzymes and is regulated during pancreatic cell differentiation. *Proc. Natl Acad. Sci. USA* **110**, 17892–17897 (2013).
77. Cheng, J. *et al.* An extensive network of TET2-targeting MicroRNAs regulates malignant hematopoiesis. *Cell Rep.* **5**, 471–481 (2013).
78. Morita, S. *et al.* miR-29 represses the activities of DNA methyltransferases and DNA demethylases. *Int. J. Mol. Sci.* **14**, 14647–14658 (2013).
79. Sengupta, S. *et al.* MicroRNA 29c is down-regulated in nasopharyngeal carcinomas, up-regulating mRNAs encoding extracellular matrix proteins. *Proc. Natl Acad. Sci. USA* **105**, 5874–5878 (2008).
80. Zhang, P., Huang, B., Xu, X. & Sessa, W. C. Ten-eleven translocation (Tet) and thymine DNA glycosylase (TDG), components of the demethylation pathway, are direct targets of miRNA-29a. *Biochem. Biophys. Res. Commun.* **437**, 368–373 (2013).
81. Takayama, K. *et al.* TET2 repression by androgen hormone regulates global hydroxymethylation status and prostate cancer progression. *Nat. Commun.* **6**, 8219 (2015).
82. Chuang, K. H. *et al.* MicroRNA-494 is a master epigenetic regulator of multiple invasion-suppressor microRNAs by targeting ten eleven translocation 1 in invasive human hepatocellular carcinoma tumors. *Hepatology* **62**, 466–480 (2015).
83. Fidalgo, M. *et al.* Zfp281 coordinates opposing functions of Tet1 and Tet2 in pluripotent states. *Cell Stem Cell* **19**, 355–369 (2016).
84. Welling, M. *et al.* DAZL regulates Tet1 translation in murine embryonic stem cells. *EMBO Rep.* **16**, 791–802 (2015).
85. Nakagawa, T. *et al.* CRL4(VprBP) E3 ligase promotes monoubiquitylation and chromatin binding of TET dioxygenases. *Mol. Cell* **57**, 247–260 (2015).
86. Shi, F. T. *et al.* Ten-eleven translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells. *J. Biol. Chem.* **288**, 20776–20784 (2013).
87. Ciccarone, F., Valentini, E., Zampieri, M. & Caiafa, P. 5mC-hydroxylase activity is influenced by the PARylation of TET1 enzyme. *Oncotarget* **6**, 24333–24347 (2015).
88. Zhang, Q. *et al.* Differential regulation of the ten-eleven translocation (TET) family of dioxygenases by O-linked beta-N-acetylglucosamine transferase (OGT). *J. Biol. Chem.* **289**, 5986–5996 (2014).
89. Bauer, C. *et al.* Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT). *J. Biol. Chem.* **290**, 4801–4812 (2015).
90. Jiang, D. *et al.* Alteration in 5-hydroxymethylcytosine-mediated epigenetic regulation leads to Purkinje cell vulnerability in ATM deficiency. *Brain* **138**, 3520–3536 (2015).
91. Wang, Y. & Zhang, Y. Regulation of TET protein stability by calpains. *Cell Rep.* **6**, 278–284 (2014).
92. Wu, H. *et al.* Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473**, 389–393 (2011).
These authors show that TET1 functions at both active and repressed (bivalent) genes.
93. Xu, Y. *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).
94. Williams, K. *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343–348 (2011).
This paper shows that TET1 can contribute to transcriptional repression by recruiting the SIN3A co-repressor complex.
95. Xu, Y. *et al.* Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for *Xenopus* eye and neural development. *Cell* **151**, 1200–1213 (2012).
96. Costa, Y. *et al.* NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* **495**, 370–374 (2013).
97. Okashita, N. *et al.* PRDM14 promotes active DNA demethylation through the ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. *Development* **141**, 269–280 (2014).
98. Neri, F. *et al.* Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. *Genome Biol.* **14**, R91 (2013).
99. Zeng, Y. *et al.* Lin28A binds active promoters and recruits Tet1 to regulate gene expression. *Mol. Cell* **61**, 153–160 (2016).
100. Wang, Y. *et al.* WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. *Mol. Cell* **57**, 662–673 (2015).
101. Rampal, R. *et al.* DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. *Cell Rep.* **9**, 1841–1855 (2014).
102. de la Rica, L. *et al.* PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. *Genome Biol.* **14**, R99 (2013).
103. Perera, A. *et al.* TET3 is recruited by REST for context-specific hydroxymethylation and induction of gene expression. *Cell Rep.* **11**, 283–294 (2015).
104. Dubois-Chevalier, J. *et al.* A dynamic CTCF chromatin binding landscape promotes DNA hydroxymethylation and transcriptional induction of adipocyte differentiation. *Nucleic Acids Res.* **42**, 10943–10959 (2014).
105. Fujiki, K. *et al.* PPARgamma-induced PARylation promotes local DNA demethylation by production of 5-hydroxymethylcytosine. *Nat. Commun.* **4**, 2262 (2013).
106. Serandour, A. A. *et al.* Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers. *Nucleic Acids Res.* **40**, 8255–8265 (2012).
107. Neri, F. *et al.* Single-base resolution analysis of 5-formyl and 5-carboxyl cytosine reveals promoter DNA methylation dynamics. *Cell Rep.* **10**, 674–683 (2015).
108. Iurlaro, M. *et al.* A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. *Genome Biol.* **14**, R119 (2013).
109. Spruijt, C. G. *et al.* Dynamic readers for 5-(hydroxy) methylcytosine and its oxidized derivatives. *Cell* **152**, 1146–1159 (2013).
This study identifies potential reader proteins for 5hmC in mouse ESCs, neuronal progenitor cells and the adult mouse brain.
110. Arab, K. *et al.* Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A. *Mol. Cell* **55**, 604–614 (2014).

111. Liu, Y. *et al.* Oestrogen receptor beta regulates epigenetic patterns at specific genomic loci through interaction with thymine DNA glycosylase. *Epigenetics Chromatin* **9**, 7 (2016).
112. Pfaffeneder, T. *et al.* The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angew. Chem. Int. Ed.* **50**, 7008–7012 (2011).
113. Bachman, M. *et al.* 5-Formylcytosine can be a stable DNA modification in mammals. *Nat. Chem. Biol.* **11**, 555–557 (2015).
114. Bachman, M. *et al.* 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat. Chem.* **6**, 1049–1055 (2014).
115. Globisch, D. *et al.* Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS ONE* **5**, e15367 (2010).
116. Munzel, M. *et al.* Quantification of the sixth DNA base hydroxymethylcytosine in the brain. *Angew. Chem. Int. Ed.* **49**, 5375–5377 (2010).
117. Yu, M. *et al.* Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* **149**, 1368–1380 (2012).
This paper describes TAB-seq, the first method capable of profiling 5hmC at genome-scale and base-resolution.
118. Wu, H. *et al.* Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev.* **25**, 679–684 (2011).
119. Booth, M. J. *et al.* Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science* **336**, 934–937 (2012).
120. Vella, P. *et al.* Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. *Mol. Cell* **49**, 645–656 (2013).
121. Pastor, W. A. *et al.* Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* **473**, 394–397 (2011).
122. Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S. & Jacobsen, S. E. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol.* **12**, R54 (2011).
123. Szulwach, K. E. *et al.* Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet.* **7**, e1002154 (2011).
124. Ficz, G. *et al.* Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473**, 398–402 (2011).
125. Feldmann, A. *et al.* Transcription factor occupancy can mediate active turnover of DNA methylation at regulatory regions. *PLoS Genet.* **9**, e1003994 (2013).
126. Sun, Z. *et al.* High-resolution enzymatic mapping of genomic 5-hydroxymethylcytosine in mouse embryonic stem cells. *Cell Rep.* **3**, 567–576 (2013).
127. Stadler, M. B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495 (2011).
128. Song, C. X. *et al.* Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell* **153**, 678–691 (2013).
129. Raiber, E. A. *et al.* Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. *Genome Biol.* **13**, R69 (2012).
130. Lu, X. *et al.* Base-resolution maps of 5-formylcytosine and 5-carboxylcytosine reveal genome-wide DNA demethylation dynamics. *Cell Res.* **25**, 386–389 (2015).
131. Szulwach, K. E. *et al.* 5-HmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat. Neurosci.* **14**, 1607–1616 (2011).
132. Song, C. X. *et al.* Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat. Biotechnol.* **29**, 68–72 (2011).
133. 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).
134. Wen, L. *et al.* Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. *Genome Biol.* **15**, R49 (2014).
135. Colquitt, B. M., Allen, W. E., Barnea, G. & Lomvardas, S. Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. *Proc. Natl Acad. Sci. USA* **110**, 14682–14687 (2013).
136. Lister, R. *et al.* Global epigenomic reconfiguration during mammalian brain development. *Science* **341**, 1237905 (2013).
137. Jin, S. G., Wu, X., Li, A. X. & Pfeifer, G. P. Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Res.* **39**, 5015–5024 (2011).
138. Khare, T. *et al.* 5-HmC in the brain is abundant in synaptic genes and shows differences at the exon–intron boundary. *Nat. Struct. Mol. Biol.* **19**, 1037–1043 (2012).
139. Wang, T. *et al.* Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum. *Hum. Mol. Genet.* **21**, 5500–5510 (2012).
140. Lu, F., Liu, Y., Jiang, L., Yamaguchi, S. & Zhang, Y. Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev.* **28**, 2103–2119 (2014).
By analysing the DNA methylome of *Tet1*;*Tet2*;*Tet3* TKO mouse ESCs, this paper shows that TET regulates enhancer activity.
141. Hon, G. C. *et al.* 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol. Cell* **56**, 286–297 (2014).
By profiling DNA methylation and hydroxymethylation of *Tet2* knockout mouse ESCs, this paper shows that TET2 regulates enhancer activity.
142. Iurlaro, M. *et al. In vivo* genome-wide profiling reveals a tissue-specific role for 5-formylcytosine. *Genome Biol.* **17**, 141 (2016).
143. Wheldon, L. M. *et al.* Transient accumulation of 5-carboxylcytosine indicates involvement of active demethylation in lineage specification of neural stem cells. *Cell Rep.* **7**, 1353–1361 (2014).
144. Xiong, J. *et al.* Cooperative action between SALL4A and TET proteins in stepwise oxidation of 5-methylcytosine. *Mol. Cell.* **64**, 913–925 (2016).
This study demonstrates that SALL4A can regulate the genetic processivity of TET.
145. Hashimoto, H. *et al.* Wilms tumor protein recognizes 5-carboxylcytosine within a specific DNA sequence. *Genes Dev.* **28**, 2304–2313 (2014).
146. Wang, D. *et al.* MAX is an epigenetic sensor of 5-carboxylcytosine and is altered in multiple myeloma. *Nucleic Acids Res.* **45**, 2396–2407 (2016).
147. Wang, L. *et al.* Molecular basis for 5-carboxycytosine recognition by RNA polymerase II elongation complex. *Nature* **523**, 621–625 (2015).
148. Raiber, E. A. *et al.* 5-Formylcytosine alters the structure of the DNA double helix. *Nat. Struct. Mol. Biol.* **22**, 44–49 (2015).
149. Saitou, M., Kagiwada, S. & Kurimoto, K. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development* **139**, 15–31 (2012).
150. Lee, H. J., Hore, T. A. & Reik, W. Reprogramming the methylome: erasing memory and creating diversity. *Cell Stem Cell* **14**, 710–719 (2014).
151. Wossidlo, M. *et al.* 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun.* **2**, 241 (2011).
152. 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).
References 151 and 152 show that active erasure of 5mC in the zygotic paternal genome is accompanied by the generation of 5hmC. Reference 151 also shows that TET3 mediates this process.
153. 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).
154. Inoue, A., Matoba, S. & Zhang, Y. Transcriptional activation of transposable elements in mouse zygotes is independent of Tet3-mediated 5-methylcytosine oxidation. *Cell Res.* **22**, 1640–1649 (2012).
155. Gu, T. P. *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606–610 (2011).
By establishing *Tet3* conditional knockout mice, this paper examines the role of maternal TET3 in oxidizing 5mC and epigenetic reprogramming.
156. Rougier, N. *et al.* Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev.* **12**, 2108–2113 (1998).
157. Inoue, A. & Zhang, Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* **334**, 194 (2011).
This study shows by immunostaining that 5hmC on the zygotic paternal genome is passively diluted by replication.
158. Howell, C. Y. *et al.* Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell* **104**, 829–838 (2001).
159. Guo, F. *et al.* Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell* **15**, 447–458 (2014).
160. Shen, L. *et al.* Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell Stem Cell* **15**, 459–470 (2014).
References 159 and 160 show that DNA replication is the driving force of demethylation in the zygote, although TET3-mediated oxidation facilitates this process to some extent.
161. Peat, J. R. *et al.* Genome-wide bisulfite sequencing in zygotes identifies demethylation targets and maps the contribution of TET3 oxidation. *Cell Rep.* **9**, 1990–2000 (2014).
162. Wang, L. *et al.* Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**, 979–991 (2014).
163. Zhu, C. *et al.* Single-cell 5-formylcytosine landscapes of mammalian early embryos and ESCs at single-base resolution. *Cell Stem Cell* **5**, 720–731.e5 (2017).
164. Nakamura, T. *et al.* PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**, 415–419 (2012).
165. Nakamura, T. *et al.* PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat. Cell Biol.* **9**, 64–71 (2007).
166. Bian, C. & Yu, X. PGC7 suppresses TET3 for protecting DNA methylation. *Nucleic Acids Res.* **42**, 2893–2905 (2014).
167. Tang, F. *et al.* Deterministic and stochastic allele specific gene expression in single mouse blastomeres. *PLoS ONE* **6**, e21208 (2011).
168. Wu, X., Inoue, A., Suzuki, T. & Zhang, Y. Simultaneous mapping of active DNA demethylation and sister chromatid exchange in single cells. *Genes Dev.* **31**, 511–523 (2017).
169. Wossidlo, M. *et al.* Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. *EMBO J.* **29**, 1877–1888 (2010).
170. Ladstatter, S. & Tachibana-Konwalski, K. A surveillance mechanism ensures repair of DNA lesions during zygotic reprogramming. *Cell* **15**, 1774–1787.e13 (2016).
171. Hajkova, P. *et al.* Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* **329**, 78–82 (2010).
172. Santos, F. *et al.* Active demethylation in mouse zygotes involves cytosine deamination and base excision repair. *Epigenetics Chromatin* **6**, 39 (2013).
173. Amouroux, R. *et al. De novo* DNA methylation drives 5hmC accumulation in mouse zygotes. *Nat. Cell Biol.* **18**, 225–233 (2016).
174. Arand, J. *et al.* Selective impairment of methylation maintenance is the major cause of DNA methylation reprogramming in the early embryo. *Epigenetics Chromatin* **8**, 1 (2015).
175. Inoue, A., Shen, L., Matoba, S. & Zhang, Y. Haploinsufficiency, but not defective paternal 5mC oxidation, accounts for the developmental defects of maternal Tet3 knockouts. *Cell Rep.* **10**, 463–470 (2015).
176. Tsukada, Y., Akiyama, T. & Nakayama, K. I. Maternal TET3 is dispensable for embryonic development but is required for neonatal growth. *Sci. Rep.* **5**, 15876 (2015).
177. Kang, J. *et al.* Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. *Proc. Natl Acad. Sci. USA* **112**, E4236–E4245 (2015).
178. Sasaki, H. & Matsui, Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat. Rev. Genet.* **9**, 129–140 (2008).
179. Gkountela, S. *et al.* DNA demethylation dynamics in the human prenatal germline. *Cell* **161**, 1425–1436 (2015).
180. Guo, F. *et al.* The transcriptome and DNA methylome landscapes of human primordial germ cells. *Cell* **161**, 1437–1452 (2015).
181. Tang, W. W. *et al.* A unique gene regulatory network resets the human germline epigenome for development. *Cell* **161**, 1453–1467 (2015).

182. Tang, W. W., Kobayashi, T., Irie, N., Dietmann, S. & Surani, M. A. Specification and epigenetic programming of the human germ line. *Nat. Rev. Genet.* **17**, 585–600 (2016).
183. Seisenberger, S. *et al.* The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol. Cell* **48**, 849–862 (2012).
184. Smith, Z. D. *et al.* A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).
185. Yamaguchi, S. *et al.* Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. *Cell Res.* **23**, 329–339 (2013).
186. Yamaguchi, S. *et al.* Tet1 controls meiosis by regulating meiotic gene expression. *Nature* **492**, 443–447 (2012).
187. Yamaguchi, S., Shen, L., Liu, Y., Sendler, D. & Zhang, Y. Role of Tet1 in erasure of genomic imprinting. *Nature* **504**, 460–464 (2013).
This paper demonstrates that TET1 mediates the erasure of genomic imprinting during PGC development.
188. Kurimoto, K. *et al.* Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. *Genes Dev.* **22**, 1617–1635 (2008).
189. 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).
190. Ohno, R. *et al.* A replication-dependent passive mechanism modulates DNA demethylation in mouse primordial germ cells. *Development* **140**, 2892–2903 (2013).
191. Seki, Y. *et al.* 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).
192. Guibert, S., Forne, T. & Weber, M. Global profiling of DNA methylation erasure in mouse primordial germ cells. *Genome Res.* **22**, 633–641 (2012).
193. Hajkova, P. *et al.* Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117**, 15–23 (2002).
194. Hackett, J. A. *et al.* Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* **339**, 448–452 (2013).
This study presents the dynamics of 5mC, 5hmC and TET during PGC development.
195. Kobayashi, H. *et al.* High-resolution DNA methylome analysis of primordial germ cells identifies gender-specific reprogramming in mice. *Genome Res.* **23**, 616–627 (2013).
196. Dawlaty, M. M. *et al.* Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev. Cell* **24**, 310–323 (2013).
197. Hargan-Calvopina, J. *et al.* Stage-specific demethylation in primordial germ cells safeguards against precocious differentiation. *Dev. Cell* **39**, 75–86 (2016).
198. Hanna, J. H., Saha, K. & Jaenisch, R. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. *Cell* **143**, 508–525 (2010).
199. Dawlaty, M. M. *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).
200. Koh, K. P. *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).
201. Huang, Y. *et al.* Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA* **111**, 1361–1366 (2014).
202. Moran-Crusio, K. *et al.* Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**, 11–24 (2011).
203. Quivoron, C. *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).
204. Ko, M. *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).
205. Li, Z. *et al.* Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood* **118**, 4509–4518 (2011).
206. Dawlaty, M. M. *et al.* Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. *Dev. Cell* **29**, 102–111 (2014).
207. Hu, X. *et al.* Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. *Cell Stem Cell* **14**, 512–522 (2014).
208. Dai, H. Q. *et al.* TET-mediated DNA demethylation controls gastrulation by regulating Lefty–Nodal signalling. *Nature* **538**, 528–532 (2016).
209. Cortazar, D. *et al.* Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* **470**, 419–423 (2011).
210. Cortellino, S. *et al.* Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146**, 67–79 (2011).
211. Doege, C. A. *et al.* Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* **488**, 652–655 (2012).
212. Klug, M., Schmidhofer, S., Gebhard, C., Andreesen, R. & Rehli, M. 5-Hydroxymethylcytosine is an essential intermediate of active DNA demethylation processes in primary human monocytes. *Genome Biol.* **14**, R46 (2013).
213. Montagner, S. *et al.* TET2 regulates mast cell differentiation and proliferation through catalytic and non-catalytic activities. *Cell Rep.* **15**, 1566–1579 (2016).
214. Yue, X. *et al.* Control of Foxp3 stability through modulation of TET activity. *J. Exp. Med.* **213**, 377–397 (2016).
215. Lio, C. J. *et al.* Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *eLife* **5**, e18290 (2016).
216. Tsagaratou, A. *et al.* TET proteins regulate the lineage specification and TCR-mediated expansion of iNKT cells. *Nat. Immunol.* **18**, 45–53 (2016).
217. Orlanski, S. *et al.* Tissue-specific DNA demethylation is required for proper B-cell differentiation and function. *Proc. Natl Acad. Sci. USA* **113**, 5018–5023 (2016).
218. Ichiyama, K. *et al.* The methylcytosine dioxygenase Tet2 promotes DNA demethylation and activation of cytokine gene expression in T cells. *Immunity* **42**, 613–626 (2015).
219. Martinowich, K. *et al.* DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* **302**, 890–893 (2003).
220. Ma, D. K. *et al.* Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* **323**, 1074–1077 (2009).
221. Guo, J. U. *et al.* Neuronal activity modifies the DNA methylation landscape in the adult brain. *Nat. Neurosci.* **14**, 1345–1351 (2011).
222. Miller, C. A. & Sweatt, J. D. Covalent modification of DNA regulates memory formation. *Neuron* **53**, 857–869 (2007).
223. Lubin, F. D., Roth, T. L. & Sweatt, J. D. Epigenetic regulation of BDNF gene transcription in the consolidation of fear memory. *J. Neurosci.* **28**, 10576–10586 (2008).
224. 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).
225. Chen, W. G. *et al.* Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science* **302**, 885–889 (2003).
226. Zhang, R. R. *et al.* Tet1 regulates adult hippocampal neurogenesis and cognition. *Cell Stem Cell* **13**, 237–245 (2013).
227. Rudenko, A. *et al.* Tet1 is critical for neuronal activity-regulated gene expression and memory extinction. *Neuron* **79**, 1109–1122 (2013).
228. Kaas, G. A. *et al.* TET1 controls CNS 5-methylcytosine hydroxylation, active DNA demethylation, gene transcription, and memory formation. *Neuron* **79**, 1086–1093 (2013).
229. Feng, J. *et al.* Role of Tet1 and 5-hydroxymethylcytosine in cocaine action. *Nat. Neurosci.* **18**, 536–544 (2015).
230. Li, X. *et al.* Neocortical Tet3-mediated accumulation of 5-hydroxymethylcytosine promotes rapid behavioral adaptation. *Proc. Natl Acad. Sci. USA* **111**, 7120–7125 (2014).
231. Yu, H. *et al.* Tet3 regulates synaptic transmission and homeostatic plasticity via DNA oxidation and repair. *Nat.Neurosci.* **18**, 836–843 (2015).
232. Zhu, X. *et al.* Role of Tet1/3 genes and chromatin remodeling genes in cerebellar circuit formation. *Neuron* **89**, 100–112 (2016).
233. Tognini, P. *et al.* Experience-dependent DNA methylation regulates plasticity in the developing visual cortex. *Nat. Neurosci.* **18**, 956–958 (2015).
234. Delhommeau, F. *et al.* Mutation in TET2 in myeloid cancers. *N. Engl. J. Med.* **360**, 2289–2301 (2009).
235. Langemeijer, S. M. *et al.* Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat. Genet.* **41**, 838–842 (2009).
236. Abdel-Wahab, O. *et al.* Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood* **114**, 144–147 (2009).
237. Jankowska, A. M. *et al.* Loss of heterozygosity 4q24 and TET2 mutations associated with myelodysplastic/myeloproliferative neoplasms. *Blood* **113**, 6403–6410 (2009).
238. Tefferi, A. *et al.* TET2 mutations and their clinical correlates in polycythemia vera, essential thrombocythemia and myelofibrosis. *Leukemia* **23**, 905–911 (2009).
239. Tefferi, A. *et al.* Frequent TET2 mutations in systemic mastocytosis: clinical, KITD816V and FIP1L1-PDGFRA correlates. *Leukemia* **23**, 900–904 (2009).
240. Tefferi, A. *et al.* Detection of mutant TET2 in myeloid malignancies other than myeloproliferative neoplasms: CMML, MDS, MDS/MPN and AML. *Leukemia* **23**, 1343–1345 (2009).
241. Ko, M. *et al.* Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839–843 (2010).
242. Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* **30**, 733–750 (2016).
243. Cimmino, L. *et al.* TET1 is a tumor suppressor of hematopoietic malignancy. *Nat. Immunol.* **16**, 653–662 (2015).
244. Zhao, Z. *et al.* Combined loss of Tet1 and Tet2 promotes B cell, but not myeloid malignancies, in mice. *Cell Rep.* **13**, 1692–1704 (2015).
245. Yang, H. *et al.* Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation. *Oncogene* **32**, 663–669 (2013).
246. Haffner, M. C. *et al.* Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell compartments and in human cancers. *Oncotarget* **2**, 627–637 (2011).
247. Wu, H. & Zhang, Y. Charting oxidized methylcytosines at base resolution. *Nat. Struct. Mol. Biol.* **22**, 656–661 (2015).
248. Shen, L., Song, C. X., He, C. & Zhang, Y. Mechanism and function of oxidative reversal of DNA and RNA methylation. *Annu. Rev. Biochem.* **83**, 585–614 (2014).
249. Huang, Y. *et al.* The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS ONE* **5**, e8888 (2010).
250. Booth, M. J., Marsico, G., Bachman, M., Beraldi, D. & Balasubramanian, S. Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. *Nat. Chem.* **6**, 435–440 (2014).
251. Lu, X. *et al.* Chemical modification-assisted bisulfite sequencing (CAB-Seq) for 5-carboxylcytosine detection in DNA. *J. Am. Chem. Soc.* **135**, 9315–9317 (2013).
252. Hayashi, G. *et al.* Base-resolution analysis of 5-hydroxymethylcytosine by one-pot bisulfite-free chemical conversion with peroxotungstate. *J. Am. Chem. Soc.* **138**, 14178–14181 (2016).
253. Petterson, A., Chung, T. H., Tan, D., Sun, X. & Jia, X. Y. RRHP: a tag-based approach for 5-hydroxymethylcytosine mapping at single-site resolution. *Genome Biol.* **15**, 456 (2014).
254. Kinney, S. M. *et al.* Tissue-specific distribution and dynamic changes of 5-hydroxymethylcytosine in mammalian genomes. *J. Biol. Chem.* **286**, 24685–24693 (2011).
255. Serandour, A. A. *et al.* Single-CpG resolution mapping of 5-hydroxymethylcytosine by chemical labeling and exonuclease digestion identifies evolutionarily unconserved CpGs as TET targets. *Genome Biol.* **17**, 56 (2016).
256. Mooijman, D., Dey, S. S., Boisset, J. C., Crosetto, N. & van Oudenaarden, A. Single-cell 5hmC sequencing reveals chromosome-wide cell-to-cell variability and enables lineage reconstruction. *Nat. Biotechnol.* **34**, 852–856 (2016).

257. Han, D. *et al.* A highly sensitive and robust method for genome-wide 5hmC profiling of rare cell populations. *Mol. Cell* **63**, 711–719 (2016).
258. Zhong, J. *et al.* TET1 modulates H4K16 acetylation by controlling auto-acetylation of hMOF to affect gene regulation and DNA repair function. *Nucleic Acids Res.* **45**, 672–684 (2016).
259. Zhang, Q. *et al.* Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. *Nature* **525**, 389–393 (2015).
260. Chen, Q., Chen, Y., Bian, C., Fujiki, R. & Yu, X. TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature* **493**, 561–564 (2013).
261. Deplus, R. *et al.* TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J.* **32**, 645–655 (2013).
262. Jin, C. *et al.* TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells. *Nucleic Acids Res.* **42**, 6956–6971 (2014).
263. An, J. *et al.* Acute loss of TET function results in aggressive myeloid cancer in mice. *Nat. Commun.* **6**, 10071 (2015).
264. Yang, J. *et al.* Tet enzymes regulate telomere maintenance and chromosomal stability of mouse ESCs. *Cell Rep.* **15**, 1809–1821 (2016).
265. Kafer, G. R. *et al.* 5-Hydroxymethylcytosine marks sites of DNA damage and promotes genome stability. *Cell Rep.* **14**, 1283–1292 (2016).

Acknowledgements
The authors thank members of the Zhang laboratory and also H. Wu for their comments on the manuscript. X.W. was supported by a fellowship from the China Scholarship Council. Y.Z. is an investigator of the Howard Hughes Medical Institute. The authors apologize to colleagues whose work cannot be cited owing to space limitations.

Competing interests statement
The authors declare no competing interests.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.