DNA demethylation pathways: Additional players and regulators

Matthias Bochtler[1)2)*], Agnieszka Kolano[1)] and Guo-Liang Xu[3)]

DNA demethylation can occur passively by ''dilution'' of methylation marks by DNA replication, or actively and independently of DNA replication. Direct conversion of 5-methylcytosine (5mC) to cytosine (C), as originally proposed, does not occur. Instead, active DNA methylation involves oxidation of the methylated base by ten-eleven trans-locations (TETs), or deamination of the methylated or a nearby base by activation induced deaminase (AID). The modified nucleotide, possibly together with surrounding nucleotides, is then replaced by the BER pathway. Recent data clarify the roles and the regulation of well-known enzymes in this process. They identify base excision repair (BER) glycosylases that may cooperate with or replace thymine DNA glycosylase (TDG) in the base excision step, and suggest possible involvement of DNA damage repair pathways other than BER in active DNA demethylation. Here, we review these new developments.

Keywords:
◼ 5-carboxylcytosine; 5-formylcytosine;
5-hydroxymethylcytosine; 5-hydroxymethyluracil; active
DNA demethylation; AID; deamination; oxidation; TDG; TET

[1)] International Institute of Molecular and Cell Biology, Warsaw, Poland
[2)] Institute of Biochemistry and Biophysics Polish Academy of Sciences, Warsaw, Poland
[3)] Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai, China

*Corresponding author:
Matthias Bochtler
E-mail: mbochtler@iimcb.gov.pl

Abbreviations:
5caC, 5-carboxylcytosine; **5fC,** 5-formylcytosine; **5hmC,** 5-hydroxymethyl-cytosine; **5mC,** 5-methylcytosine; **AID,** activation induced deaminase; **BER,** base excision repair; **MBD,** methyl binding domain; **NEIL,** Nei-Like; **NER,** nucleotide excision repair; **SMUG1,** single-strand selective monofunctional uracil DNA glycosylase 1; **TCA,** tricarboxylic acid; **TDG,** thymine DNA glycosylase; **TET,** ten-eleven translocation; **UNG,** uracil DNA glycosylase.

Introduction

DNA methylation in vertebrates is involved in multiple processes, including the control of gene expression, X-chromosome-inactivation, imprinting, and the silencing of mobile genetic elements [1]. The importance of DNA methylation is evident from the severe consequences of its loss. Lack of the DNA maintenance methyltransferase Dnmt1 leads to death of differentiated cells within a few generations in culture, in both murine and human cells [2]. At the organismal level, *Dnmt1* deficiency causes a delay in development, and death in mid-gestation [3]. The de novo methyltransferases DNMT3A and DNMT3B are not required for the survival of cells in culture (at least not separately), but *Dnmt3b* deletion causes embryonic lethality in the mouse. *Dnmt3a* knockout mice are born, but die at about 4 weeks of age [4].

DNA demethylation does not have such a clear role, in part because many pathways may contribute and act redundantly [5]. Due to the resilience of carbon-carbon bonds, methylation was originally considered as an irrevers-ible modification, which could only be altered through dilution or de novo synthesis of DNA [6, 7]. However, it is now clear that DNA methylation need not be passively "diluted" through DNA replication, but can also be actively erased [8]. Active genome wide DNA demethylation is well documented in the sperm-derived paternal pronucleus of the zygote for rapid activation of zygotic transcription [9–11]. Locus-specific active DNA demethylation is well documented in post-mitotic cells in the adult brain [12, 13], and also during cell fate changes [14].

One step conversion of 5-methylcytosine (5mC) to cytosine (C) is unknown, and also unlikely because of the difficulty of breaking non-activated carbon-carbon bonds. Pathways that convert 5mC back to C in several steps without the formation of nicks or double strand breaks in DNA have been described, but are improbable in vivo [15–17]. All known in vivo conversions of 5mC to C involve replacement of the methylated nucleotide, at the cost of the creation of single strand nicks. Unlike plants, animals do not seem to be able to excise 5mC paired with G directly (except in the context of excision of larger patches of DNA). Instead, the base (or a nearby base) is always first modified, either by deamination,

oxidation, or – according to models that now appear less likely – a combination of both, followed by replacement of the modified nucleotide (possibly together with surrounding nucleotides [18, 19]). Most evidence points to the base excision repair (BER) pathway [20–24], but its role in global DNA demethylation has been challenged [25]. Alternative pathways include nucleotide excision [26, 27] and non-canonical mismatch repair [28]. Modified DNA bases can also contribute to DNA demethylation by "dilution" of methylation during DNA replication. 5-Hydroxymethylcytosine (5hmC) in the template strand in the CpG context blocks methylation of the newly synthesized strand by the maintenance methyl-transferase DNMT1, and thus has the same effect as an unmodified C in the parental strand [29]. In the following, we first review modifications to the 5mC base. We then discuss the fates of the altered DNA bases.

DNA base modification is the first step of active demethylation

Genetic data implicate AID in demethylation, but the substrate is not clear

Activation induced deaminase (AID) deaminates cytosine to uracil and to a lesser extent 5mC to thymine, by simple hydrolysis. AID has been named for its role in class switch recombination, hypermutation, and gene conversion in activated B-cells [30, 31], mediated by its deaminase activity against cytosines in DNA [32]. AID is mainly expressed in lymphoid cells [33], and only scarcely and transiently in other tissues [34], except in the context of malignancy [35–37]. Consistent with the expression pattern of Aid in germinal centers, lack of Aid abrogates class switch recombination and somatic hypermutation in the mouse [30]. In humans, AID deficiency causes the autosomal recessive form of hyper-IgM syndrome (HIGM2) [38]. Both the class switch recombination and hypermutation defects can be explained without reference to DNA demethylation, but additional experiments show that Aid deficiency may nevertheless impair DNA demethylation.

Expression of Aid in embryonic stem (ES) cells [39] and the role of Aid in demethylation during reprogramming experiments involving ES or iPSC cells are controversial. In heterokaryon experiments (involving the fusion of ES cells and differentiated cells), two independent groups found that Aid knockdown reduced the efficiency of reprogramming, and impaired demethylation of promoters of pluripotency associated genes [39, 40], whereas a third group reported no expression of Aid in ES cells and concluded that Aid was not required for reprogramming [41]. Reports about the involvement of AID in the generation of induced pluripotent stem cells (iPSCs) are equally contradictory. The original report concluded that AID was required early in reprogramming for the demethylation of selected promoters of pluripotency genes [42]. A second study corroborated AID involvement, but found that the enzyme regulated the induction of pluripotency first negatively and then positively [43]. Two further studies concluded that Aid did not play any role in reprogram-ming [44, 45] (Fig. 1).

Studies addressing the role of deamination in DNA demethylation in animals have also not yet led to consensus conclusions. A genetic study of methylation dynamics of germinal center B-cells from mice convincingly supports a role of the enzyme in DNA demethylation [46]. However, the same authors found no effect of Aid on DNA methylation levels when using cells in culture, in agreement with the results of at least two other groups [47, 48]. A pathway for DNA demethylation in zebrafish embryos involving Gadd45a, Aid/Apobec, and Mbd4 (to excise G:T mispairs) was proposed [49] based on a correlation between Aid/Apobec levels and DNA demethylation activity. However, the experimental findings have since been challenged [50]. Slightly higher levels of DNA methylation were found in primordial germ cells of *Aid* null mice compared to controls, but it was not demonstrated that the effect was direct [51], and results may be influenced by different genetic backgrounds of *Aid* deficient and control mice [47]. Demethylation in the mouse brain by linked deamination and BER has been suggested based on the detection of a complex of Tdg, Gadd45a, and Aid, but Aid activity on 5mC containing DNA could not be directly demonstrated [23] (Fig. 1).

Many reports on a role of AID in demethylation are hard to reconcile with prevailing views about AID expression patterns. Moreover, they conflict with biochemical results, particularly with the tacit assumption that AID acts directly on the 5mC base. Several reports agree that AID deaminates C much more efficiently than 5mC [52–54]. Zebrafish Aid is a better 5mC deaminase than other orthologs, but still prefers a C over a 5mC [55]. Moreover, AID works only on single stranded DNA (after treatment with RNase to clear RNA-DNA hybrids), but not on double-stranded DNA or RNA-DNA heteroduplexes [56].

Several investigators have sought to resolve the difficulty of poor AID activity toward 5mC by postulating the involvement of alternative deaminases. At least one human APOBEC3 paralog (3A, but not 3G) indeed deaminates 5mC in DNA faster than AID [57], but the single murine APOBEC3 also discriminates against 5mC [52]. Some prokaryotic methyltransferases can catalyze deamination of Cs to uracils [58] and of 5mCs to thymines [59] in the absence of S-adenosylmethionine (SAM), but mammalian DNA methyltransferases are not known to catalyze this reaction, either in vitro or in vivo. Moreover, the implication of alternative deaminases does not explain the genetic data that implicate specifically Aid and not deaminases in general in DNA demethylation.

The conflicting genetic and biochemical data suggest that AID acts indirectly. AID may deaminate C, not 5mC, and "regional" repair may then lead to the replacement of methylated cytidines by unmethylated ones in the vicinity [28, 60, 61] (Fig. 2). The model is appealing, because it is consistent with both genetic and biochemical data. Moreover, deamination of C rather than 5mC avoids the formation of a legitimate, but miscoding (and therefore highly mutagenic) DNA base during DNA demethylation. For more detail, and additional arguments for and against a deamination-based DNA demethylation pathway, the reader is directed to other reviews in the literature [62–64].

Figure 1. Evidence for and against deamination-based DNA demethylation. The experiments in favor of and against AID involvement in DNA demethylation in germinal center B-cells are not directly comparable, because both Hogenbirk et al. [47] and Fritz et al. [48] only assessed the effect of the presence or absence of Aid in cell culture, whereas Dominguez et al. [46] used ex vivo cells. In culture, Dominguez et al. also do not observe AID dependent demethylation [46].

Genetic and biochemical data implicate TETs in DNA oxidative demethylation

Ten-eleven-translocation (TET) proteins are Fe(II)-dependent dioxygenases that oxidize DNA using molecular oxygen and α-ketoglutarate as co-substrates, generating oxidized DNA, succinate, and carbon dioxide (CO_2) as the co-products. TETs act on modified cytosines. They convert 5mC to 5hmC [65, 66], 5hmC to 5-formylcytosine (5fC), and 5fC to 5-carboxylcytosine (5caC) [67]. Simplistically speaking, TET enzymes split

Figure 2. Indirect replacement of 5mC nucleotides. The reaction is initiated by deamination of C (as shown) or 5mC (not shown). The base is then excised by UDG or TDG, and a nick is created at the abasic site. Long patch DNA resynthesis leads to replacement of 5mC nucleotides by unmodified nucleotides (shown with tilted bases). Flap trimming and DNA ligation complete the demethylation reaction. In principle, a similar resynthesis reaction could also occur after excision of oxidation based demethylation intermediates. The NER [27, 130] and ncMMR [28] pathways may take over the role of LP-BER in local DNA resynthesis in some circumstances.

molecular oxygen into oxygen atoms (with two unpaired electrons and a carbene-like tendency to insert into other bonds). One oxygen atom inserts into succinic semi-aldehyde, the formal product of α-ketoglutarate decarboxylation, to form succinate, the other into the 5-substituent of the C-base, converting 5mC to 5hmC, 5hmC to 5fC (as a result of the equilibrium between gem-diol and aldehyde by dehydration/hydration), and 5fC to 5caC (Fig. 3A). More realistic models of the reaction mechanism of α-ketoglutarate-dependent dioxygenases in general [68] and of TETs in particular [69] agree that oxygen is split while bound to the iron in the TET active site. The iron resting state is Fe(II). Reaction with oxygen leads to a Fe(IV) = O species involved in oxidizing the DNA base (Fig. 3B).

TETs were originally identified in the context of hematologic malignancies. The first paralog, TET1, is named for the karyotype aberration t(10;11)(q22;q23), that creates a TET1-MLL fusion protein [70]. The same study also showed that there are three TET paralogs in humans, as was later found to be true also for mouse and zebrafish. Xenopus has only Tet2 and Tet3, presumably due to lineage specific loss of Tet1. Functional Tets have also been found in invertebrates [71], and even in the protozoan *Naegleria gruberi* [72, 73], but not in plants [74]. Tets are large multidomain proteins (typically in excess of 1,000 amino acids). Tet1 consists of a CXXC domain, a Cys-rich domain (distinct from the CXXC domain, which is also cysteine rich), and a double stranded β-helix (DSBH) domain, which is interrupted by a spacer of unknown function. The Cys-rich and DSBH domains form the catalytic do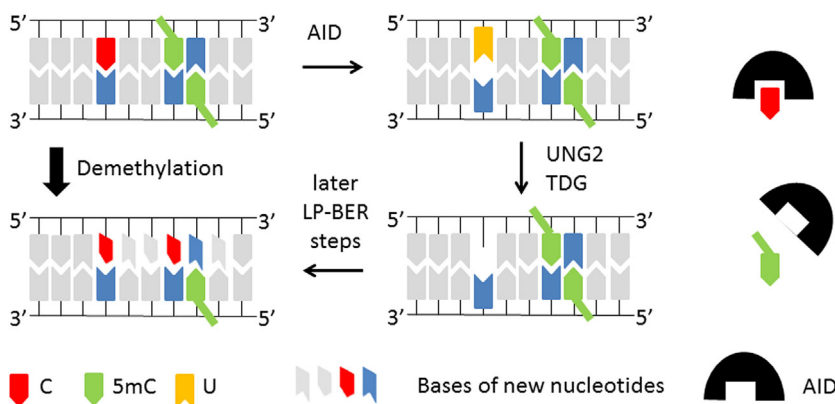main harboring the dioxygenase activity. The CXXC domain is involved in Tet targeting. In the case of Tet2, the CXXC domain is provided on a separate poly-peptide chain termed IDAX, due to an ancestral chromosomal gene inversion [75]. Tet3 occurs in several isoforms with and without CXXC domain [76].

The link between TET activity and demethylation is supported by both genetic and biochemical data. As TETs oxidize 5mC, they obviously reduce the level of 5mC, and conversely, a loss of TETs causes hypermethylation [77]. This does not necessarily imply that all oxidized nucleobases are ultimately converted back to the unmodified ones. In the early embryo, 5hmC has been shown to persist through the early divisions [11]. Particularly in the brain, 5hmC levels are relatively high [78], and a plethora of 5hmC reader proteins, with no obvious connection to DNA repair, suggest a role for 5hmC as an epigenomic mark in its
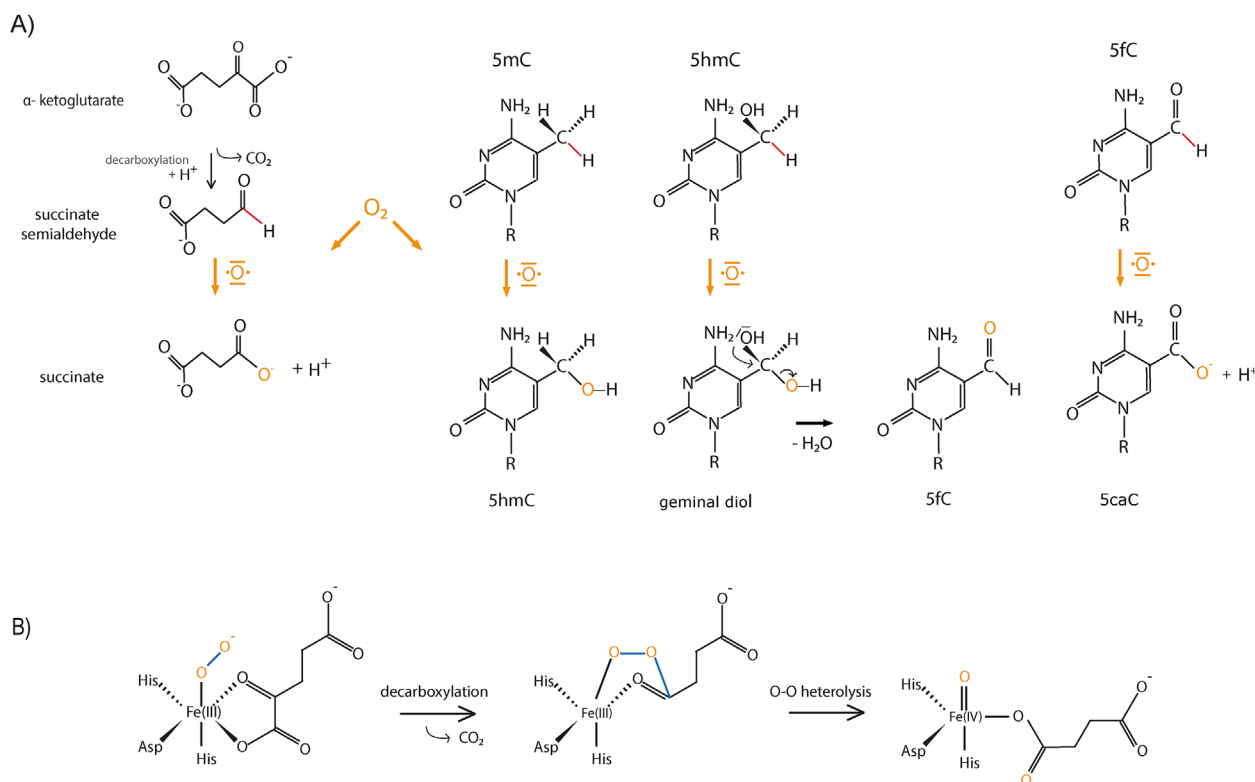
A)

B)

Figure 3. TET-catalyzed oxidation reactions. **A:** Incorporation of atomic oxygen into 5mC, 5hmC, and 5fC. C—H bonds for insertion are marked red, and inserted oxygen atoms orange. Note that the description is a gross simplification to keep track of reaction substrates and products. **B:** Activation of molecular oxygen and formation of the Fe(IV)=O intermediate according to the currently accepted model for the reaction.

own right [79]. However, there is solid evidence for the entire pathway from 5mC ultimately back to C [22, 80]. Recently, the entire pathway has been reconstituted in vitro from defined components [81]. TET knockout phenotypes are informative and consistent with a demethylation defect.

Triple *Tet* knockout ES cells and murine embryonic fibroblasts (MEFs) are viable, but have differentiation and de-differentiation defects [77, 82]. Triple *Tet* KO ES cells have problems with teratoma formation, contribute poorly to chimeric embryos, and cannot support embryonic development [77]. MEFs have problems with de-differentiation under the influence of defined reprogramming factors (related to the original Yamanaka factors [83]). The defect was traced to an impaired mesenchymal-to-epithelial transition (MET), and could be rescued by overexpression of miRNAs that were suppressed by the TET deficiency [82]. Conversely, TET overexpression facilitates reprogramming. Tet1 can replace Oct4 in the standard cocktail for iPSC induction [84], and together with Oct4 suffices for iPSC generation in the absence of other exogenous reprogramming factors [85]. The positive effects of Tets on reprogramming and differentiation are likely mediated by the control of the methylation state of regulatory loci, particularly enhancers that need to be activated to bring about cell-type changes [14, 86, 87].

In animals, *Tet* deletion has species-dependent effects. In the mouse, deletion of *Tet3* causes neonatal lethality. Germline-specific ablation of *Tet3* abolishes the appearance of oxidized 5mC derivatives in the paternal pronucleus [10]. However, the consequences of *Tet3* ablation from the oocyte are mild. *Tet3* ablation reduces fertility, but surprisingly embryos from mating of *Tet3* null mothers and wild-type fathers show no overt defects in pre-implantation development and are viable, suggesting that demethylation of the paternal pronucleus is not essential [88]. In the mouse, loss of either *Tet1* or *Tet2* is compatible with embryonic and postnatal development [89, 90]. Double knockout of *Tet1* and *Tet2* leads to partly penetrant mid-gestation abnormalities. Nevertheless, some overtly normal mice are born, albeit with reduced fertility of females [91]. Triple *Tet* KO mice exhibit defects already at the gastrulation stage. The phenotype is akin to the Nodal gain of function phenotype and probably due to hypermethylation of the promoters and consequent decreased expression of the *Lefty* genes that antagonize Nodal signaling [92]. In Xenopus, *Tet3* deficiency causes defects in eye and neural development [93]. In zebrafish, *Tet3* deficiency alone is inconsequential, but combined loss of *Tet2* and *Tet3* genes, or of all three *Tet* genes, is incompatible with survival beyond the larval period [94].

Impaired TET activity alone does not seem to cause malignancies, but it clearly contributes, particularly in the hematopoietic system. TETs were named for a translocation affecting the TET1 gene, but the case for the involvement of TETs in malignant transformation is now clearest for a genetic loss of TET2 [95–97]. Interestingly, impaired TET activity can play a role in malignancies also in the absence of genetic alterations of TET genes. Perturbations of the tricarboxylic acid cycle (TCA) lead to accumulation of genuine
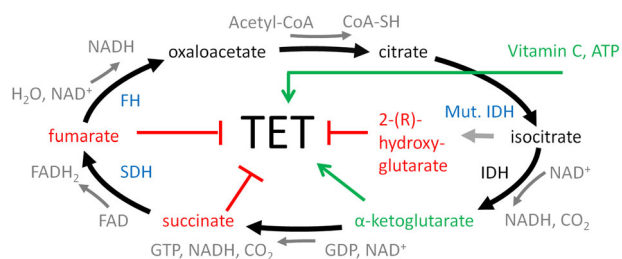
Figure 4. Influence of small molecule metabolites on TET activity. The TCA cycle is shown in highly simplified form, omitting enzymes and metabolites without clear link to TET activity. The co-substrate α-ketoglutarate and the positive regulators of TET activity are in green, negative regulators are in red. Enzymes that have been (in their mutant variants) implicated in TET inhibition are shown in blue. Isocitrate dehydrogenase (IDH), succinate dehydrogenase (SDH), and fumarate hydratase (FH) have mitochondrial and cytoplasmic isoenzymes, encoded by separate genes in the case of IDH. Mutations in both the cytosolic and mitochondrial forms of IDH have been implicated in TET dysregulation. Note that redox equivalents generated in the TCA cycle may influence vitamin C status, and are used in oxidative phosphorylation for the production of ATP, which also affects TET activity.

or aberrant TCA metabolites, which inhibit TETs, but also other α-ketoglutarate-dependent dioxygenases [98, 99] (Fig. 4 and Box 1).

AID and TETs may cooperate, but are unlikely to act on the same base

The clear evidence for the involvement of both AID and TETs in DNA demethylation has led to suggestions that they may act sequentially on the same DNA base [23]. In principle, AID could act upstream or downstream of TETs. Models that AID acts upstream require the formation of the legitimate DNA base thymine as a demethylation intermediate and imply a pathway that should be highly mutagenic. Moreover, they have to be reconciled with the weak activity of AID on 5mC (compared to C) [52]. Models that AID acts downstream require the action of AID on 5hmC, 5fC, or 5caC, and conflict directly with biochemical data on AID [52]. Moreover, they are hard to reconcile with the observation that most 5-hydroxymethyluracil (5hmU) in the genome appears to be formed by oxidation of T, not deamination of 5hmC (Fig. 5 and Box 2).

Nucleotide replacement is the second step of active demethylation

Direct reversion of modified DNA bases to C does not seem to occur in vivo

The products of C or 5mC deamination by AID can only be resolved by repair (because deamination decreases similarity to C and may affect base close to, but different from 5mC). In contrast, at least some of the products 5mC oxidation could in principle revert back to C. The search for enzymes that

could catalyze such reversion reactions has paradoxically led to the DNA methyltransferases. In vitro, thiol reagents and the de novo mammalian DNA methyltransferases DNMT3a and DNMT3b can convert 5hmC (with loss of formaldehyde) and 5caC (with loss of CO_2) to C in the absence of the methyl donor S-adenosylmethionine [15–17, 100]. De novo DNA methyltransferases are present both in the germline and in the oocyte (judging from defects in germline development in their absence [101–104]), but a role of these enzymes in demethylation would require the improbable absence of SAM. 5caC decarboxylation activity has been reported for stem cell nuclear extracts [105]. Orotidine 5-phosphate decarboxylase (ODC) decarboxylates orotidine 5-phosphate (6-carboxy-uridine 5-phosphate) very efficiently [106], but lacks activity on 5caC containing DNA. The stem cell decarboxylase activity therefore remains unconfirmed and its source unknown. Thus, DNA repair appears to be the only in vivo pathway to resolve DNA demethylation intermediates.

Various BER glycosylases are involved in the nucleotide replacement step

Uracil, generated from C by deamination and mispaired with guanine, can be excised by UNG2, SMUG1, TDG, and MBD4 [107]. Thymine, generated by deamination of 5mC and mispaired with guanine, is excised by TDG and MBD4. The reactions are well documented in the context of somatic hypermutation. In the case of Ung2 [61] and Tdg [28], there is also evidence for the involvement of these reactions in the removal of nearby 5mC bases, by either long patch base excision or non-canonical mismatch repair.

In contrast to deamination products, oxidized methyl-cytosine derivatives are not mispaired, at least when 5fC and 5caC are present in their dominant tautomeric forms [108, 109]. However 5caC:G pairs may resemble T:G mispairs when 5caC when 5caC adopts a minor and still debated [109, 110] alternative tautomer [111]. 5fC and 5caC nucleotides also resemble BER substrates in many other ways. As formyl and carboxyl groups are electron-withdrawing, both 5fC and 5caC nucleotides have weakened glycosidic bonds [110, 112] like many other oxidatively and otherwise damaged DNA nucleotides [113]. Base pairing is weaker for 5caC:G than for C:G [112], presumably leading to a higher rate of spontaneous flipping, as for many damaged DNA bases [114]. Weakened glycosidic bonds and increased flipping rates suggest that 5fC and 5caC may be excised by BER glycosylases, and indeed, this is the case (Fig. 6).

Thymine DNA glycosylase (TDG), a monofunctional DNA glycosylase, has been named for its activity against T bases arising from 5mC deamination mispaired to G. TDG was also the first enzyme found to be able to excise 5fC and 5caC [20, 80]. Favorable interactions between TDG and the flipped 5caC base [115], and faster excision of 5fC than T from pairs with G [20] have prompted speculation that the "main" role of TDG may be in DNA demethylation rather than in deamination repair. Tdg is barely expressed in oocytes or zygotes (judging from the RNA levels), and not required for demethylation of the paternal pronucleus [116]. It is essential for MET [82], which plays a role in somitogenesis and organogenesis (e.g.

Box 1

Impaired 5mC oxidation contributes to malignancies

Loss or reduction of TET function and reduced 5hmC levels are clearly correlated with various malignancies, particularly of the hematopoietic system. The *TET1* gene was first identified as a breakpoint of the t(10;11)(q22;q23) translocation leading to a TET1-MLL fusion protein in a group of patients with acute myelogenic leukemia (AML) [70]. *TET2* mutations are frequent in myeloid malignancies, including myelodysplastic syndrome, acute myelogenic leukemias, and particularly chronic myelo-monocytic leukemias [95–97]. The link between the loss of TET function and hematopoietic malignancies is at least in part causal.

Mice with conditional loss of *Tet2* (not only of both, but also of a single allele) in hemotopoietic cells exhibit progressive enlargement of the hematopoietic stem cell compartment and eventual myeloproliferation in vivo [137]. Overall loss of *Tet2* leads to a tendency to develop myeloid malignancies over the time-course of a year [90]. Acute loss of *Tet3* in a *Tet2* deletion background strongly enhances this tendency, and results in aggressive myeloid cancer [138]. Unexpectedly, combined loss of *Tet1* and *Tet2* promotes B-cell, but not myeloid malignancies in mice [139]. In zebrafish, loss of *Tet2* causes erythroid dysplasia and anemia [140], and myelodyspastic syndrome [141]. Reports on DNA methylation levels in TET deficiency are contradictory. Some of the early studies reported paradoxical hypomethylation [75, 142], while other data suggested mechanistically more plausible hyper-methylation [143], particularly at enhancers. The common theme of the TET deficiency related malignancies is a failure of progenitor cells to mature. The defect is most parsimoniously explained by silencing of genes required for differentiation due to hypermethylation.

A reduction of TET activity may play a role in malignancies in the absence of genetic alterations of the TET genes. The TET co-substrate α-ketoglutarate is an intermediate of the tricarboxylic acid (TCA) cycle and an also a metabolite in amino acid synthesis (as the product of transamination). Two genuine TCA metabolites, succinate and fumarate, and the oncometabolite (R)-2-hydroxy-glutarate inhibit TETs [144–146] and accumulate as result of succinate dehydrogenase (SDH) [98], fumarate hydra-tase (FH) [98], and isocitrate dehydrogenase (IDH1/IDH2) deficiency [99]. Succinate, fumarate, and (R)-2-hydroxy-glutarate all broadly inhibit α-ketoglutarate-dependent dioxygenases, including the Jumonji histone demethy-lases. Among other effects, they also cause histone hypermethylation [147], demonstrating that TETs are not the only mediators of the oncogenic properties of the di- or tri-carboxylic acids (Fig. 4).

The antioxidant vitamin C, controversially hailed as an anti-cancer dietary supplement [148], enhances TET activity in vitro and in vivo [149–151], presumably because it affects the redox state of the iron co-factor [152]. Vitamin C also facilitates reprogramming, which depends in part on TET mediated DNA demethylation [82, 153, 154], and enhances the effect of the hypomethylating drug 5-azacytidine in hepatocellular cancer cell lines [155]. Many investigators use ATP for TET dependent reac-tions [149, 156], because it appears to enhance oxidation. However, the mechanistic role of ATP in the dioxygenase reaction is not understood, and it is not clear whether the dependence reflects a link between the metabolic state of cells and TET activity.

nephrogenesis, cardiogenesis, or hepatogenesis). The timing of MET fits nicely with the appearance of a *Tdg* deficiency phenotype around embryonic day 11.5 in the mouse [23, 117], but causality remains to be established.

Nei-like 1 (NEIL1) and Nei-like 2 (NEIL2), both bifunctional DNA glycosylases, excise oxidatively damaged pyrimidines
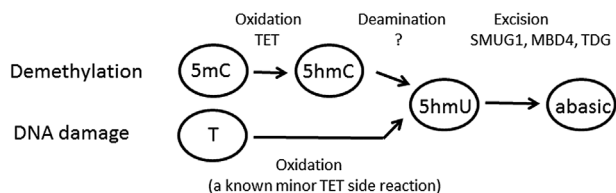

Figure 5. Possible pathways for generation and excision of 5hmU. Suggestions that 5hmU may be a demethylation intermediate suffer from the lack of a convincing candidate for the 5hmC deaminase. TETs discriminate only imperfectly against T. Therefore, and because of the vast excess of T over 5mC in the genome, enzymatic oxidation of T to 5hmU can occur. Isotope tracer studies indicate that this pathway is the main source of 5hmU in the genome [161], which is subsequently excised by several BER pathway enzymes.

(5-hydroxycytosine, 5-hydroxyuracil, thymine glycol) and purines chemically degraded to bases structurally resembling damaged pyrimidines (hydantoins, formamidopyrimidines) [118]. In several screens, NEIL1 or NEIL2 contributed to reactivation of epigenetically silenced reporter plasmids, although to a lesser extent than TDG [22, 119]. Leonhardt and co-workers concluded that NEIL1 and NEIL2 could act redundantly with TDG [22]. Niehrs and co-workers dispute glycosylase activity on 5fC and 5caC containing DNA, and suggest instead that TDG recruits NEIL1 or NEIL2 (instead of APEX) to excise the (deoxy)ribonucleosides [119]. A physio-logical demethylation process requiring either NEIL1 or NEIL2 remains to be discovered. *Neil1* knockout (and heterozygotic) mice develop metabolic syndrome, a combination of severe obesity, dyslipidemia, fatty liver disease, and a tendency for hyperinsulinemia [120], but none of the phenotypes has been linked to a demethylation defect. Neil2 knockdown leads to a neural crest phenotype in Xenopus [119] not seen in *Neil2* deficient mice [121].

NEIL3 (unlike NEIL1 and NEIL2) is a monofunctional glycosylase lacking β, δ-lyase activity. Reports are either inconclusive [21] or conflict on whether [22] or not [119] NEIL3

Box 2

Could 5fC and 5caC be off the main demethylation pathway?

5fC and 5caC share many features with damaged DNA bases (Fig. 6). They are rare, they inhibit basic trans-actions on DNA such as transcription [132, 133], and they stimulate repair (the BER pathway, the exonuclease activity of replicative DNA polymerases [111]). Could oxidation of 5hmC to 5fC and 5caC be a form of DNA damage, the result of overshooting TET activity? Thymine hydroxylase (thymine dioxygenase), another α-ketoglutarate-dependent dioxygenase, also overshoots [157], suggesting that dioxygenases in general may have a problem to "cleanly" stop oxidation at the level of the hydroxymethyl group. Could there be a pathway from 5hmC back to C, not involving 5fC or 5caC? In other words, what could happen to the 5hmC base other than further oxidation?

5hmC is chemically stable and resistant to all DNA glycosylases that have been tested, but it may be channeled to base excision repair by deamination to 5-hydroxymethyluracil (5hmU). A demethylation pathway involving 5mC oxidation prior to deamination has indeed been claimed for the mouse, based on the detection of a complex of Gadd45a, Aid, and Tdg, and the finding that Tdg has glycosylase activity on 5hmU [23]. It is also supported by a requirement for both Aid and Tet1 for demethylation of certain reporter constructs in ES cells. Finally, 5hmU is excised efficiently by the SMUG1 [158], MBD4 or TDG [23, 159] glycosylases and can also be removed by non-canonical mismatch repair, at least when the base is mispaired with G [28]. Glycosylase excision is 60-fold more efficient when 5hmU is paired with G than when it is paired with A [160], suggesting that the pathway may be dedicated to the removal of deamination products of bases derived from C and rather than to the removal of oxidation products of T (Fig. 5).

Nevertheless, a mixed oxidation/deamination pathway remains doubtful unless a convincing deaminase is discovered. AID activity on 5hmC nucleobases in DNA has never been demonstrated. Instead, biochemical studies have made a convincing case that 5hmC has too much steric bulk in the C5 position of the pyrimidine ring to be a substrate for AID [52, 53]. Moreover, isotope tracer studies indicate that most 5hmU in the genome is formed by Tet mediated oxidation of T and not by the deamination of 5hmC [161]. The efficient excision of 5hmU by several BER glycosylases therefore seems primarily required to repair oxidative damage to T. Thus, on balance current data speak against a mixed oxidation/deamination pathway for active DNA demethylation, and support the role of 5fC and 5caC as bona fide intermediates of oxidative DNA demethylation.

can excise 5caC from DNA. *Neil3* is highly expressed in the mouse oocyte, the unfertilized ovum, and the zygote, but then the expression falls strikingly after the zygote stage [122], suggesting a possible involvement in zygotic DNA demethyl-ation (together with UNG2, see below). *Neil3* is also expressed in the developing mouse brain, particularly in regions where neurogenesis takes place [122]. *Neil3* deficient mice are viable and fertile, but exhibit a loss of neural progenitors [123], pointing to a role of the enzyme in rapidly dividing cells, rather than in active DNA demethylation.

Uracil DNA glycosylases occur in mitochondrial (UNG1) and nuclear (UNG2) isoforms that are generated by alternative splicing. The enzymes excise uracils arising from cytosine deamination. Ung2 was recently identified in an unbiased screen for glycosylases involved in TET-dependent gene reactivation. In cultured cells, Ung2, but not an inactive Ung2 variant, prevented the accumulation of 5caC in genomic DNA resulting from Tet2 (catalytic domain) overexpression [21]. The finding was unexpected, because in vitro UNG2 had been reported to lack activity against 5caC containing substrates [80]. Whether weak UNG2 activity in vitro against 5-carboxyuracil (5caU) containing DNA [21] explains the cell culture data is not clear, because catalyzed conversion of 5caC to 5caU has not been demonstrated yet. As *Ung* transcript levels are high in the zygote and early embryo, a role of the enzyme in DNA demethylation at this stage was tested. *Ung* deficiency in the zygote impaired demethylation at some loci (*Nanog* and *Line1* elements, some maternally hypermethy-lated regions), but did not perturb global levels of oxidized 5mC bases [21]. As Ung2 has also been reported to cooperate with Aid in active demethylation in the zygote [61], it is not clear whether loss of *Ung* impairs oxidation or deamination based demethylation. *Ung* deficient mice have elevated levels of uracil in DNA, but develop normally into adulthood with no overt phenotype [124].

It is generally assumed that standard steps of BER [125] operate downstream of glycosylase action in DNA demethylation. However, little attention has been paid to whether a single or more nucleotides are replaced. For BER downstream of AID, only long patch repair is productive for DNA demethylation, if indeed AID deaminates Cs and not 5mCs. For BER downstream of TETs, short and long patch repair could be productive for demethylation. Short patch repair with Polβ alone has been reconstituted [81], but whether this is the dominant pathway downstream of TETs in vivo is not clear.

Other damage repair factors and pathways may contribute

In addition to the BER pathway, other DNA damage repair pathways may be involved in active DNA demethylation.

Gadd45a (growth arrest and DNA damage inducible) participates in various pathways that affect genome integ-rity [126, 127]. The gene was also identified in a screen for factors involved in DNA demethylation, and its overexpres-sion was reported to cause activation of methylation silenced
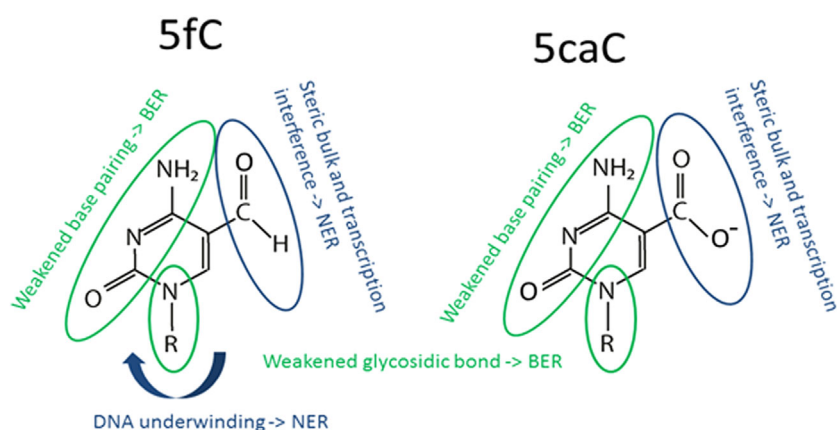
Figure 6. Features of 5fC and 5caC nucleotides shared with damaged DNA nucleotides, and pathways involved in their replacement. The statements on weakened glycosidic bonds and N3 hydrogen bonding for 5fC:G and 5caC:G pairs are based on the work of Maiti et al. [110] and Dai et al. [112]. The figure shows only the major tautomers. 5caC has to adopt a minor tautomeric form for 5caC:G pairs to resemble T:G mispairs [111].

reporter constructs and global DNA demethylation [128]. Moreover, based on experiments in zebrafish, it was suggested that Gadd45 acted in deamination based DNA demethylation [49]. Both studies have been challenged [50, 129], but the involvement of Gadd45 proteins in active DNA demethylation has been confirmed also in several other systems [13, 130]. It is not entirely clear which DNA demethylation pathways are stimulated by Gadd45 proteins. Initial data pointed to an involvement of Gadd45 in AID based DNA demethylation [49]. More recently, Gadd45a was shown enhance oxidation based DNA demethylation [128]. Several DNA repair pathways may be coopted by Gadd45 proteins. Gadd45a interacts directly with the BER glycosylase TDG and enhances oxidation based demethylation. The protein also binds the NER 3′-endonuclease XPG and cooperates with it in demethylation [27].

Nucleotide excision repair (NER) repairs primarily bulkier lesions than BER, or lesions that interfere with transcription (TC-NER). At first sight, the formyl and carboxyl groups of 5fC and 5caC appear too small to trigger NER. However, 5fC causes DNA under-winding [108], like many DNA lesions that are repaired by the nucleotide excision repair pathway [131]. 5fC and 5caC in the template strand interfere with transcription [132, 133], suggesting that they may initiate transcription coupled nucleotide excision repair (TC-NER) [134] (Fig. 6). Indeed, the NER 3-endonuclease XPG was shown to be required for demethylation of a reporter in Xenopus oocytes. Moreover, the NER factors XPA, XPG and XPF were shown to be required for demethylation of an rRNA promoter, and it was additionally shown that the catalytic activity of XPG (and not just the XPG protein itself) was required for demethylation [130].

Non-canonical mismatch repair (ncMMR) has also been considered as a possible pathway for the repair step of DNA demethylation. The process is normally geared towards the removal of mismatches starting from a nick elsewhere in the DNA, and leads to replacement of nucleotide patches from the nick or DNA end to ∼150 nucleotides past the

mismatch site [135]. In a recent study, it was demonstrated that uracils in DNA could trigger the replacement of 5mC nucleotides in DNA in a nick-dependent manner [28]. Although the process competes with BER for U:G mismatches, it is appears to be enhanced by BER, because UDG or TDG create the nicks at U sites that are required to prime new DNA synthesis in ncMMR [28]. Due to the role of U bases in the process, ncMMR would appear to be involved in AID dependent demethylation. On the other hand, a report that MutSα preferentially binds DNA with 5caC:G pairs over unmodified DNA suggests possible involvement of ncMMR in oxidation based DNA demethylation [111], which has not been corroborated in a functional assay [28]. At present, the involvement of ncMMR in DNA demethylation has not yet been demonstrated in physiological circumstances, and it remains unclear how the genome integrity hazards associated with the combined action of AID and ncMMR could be avoided (Box 3).

Conclusions and outlook

All known pathways for active DNA demethylation that have been shown to operate in vivo in animals (but not plants) begin with DNA base modification and end with nucleotide replacement.

The base modification step involves either AID catalyzed deamination, or TET catalyzed oxidation. TET participation in active DNA demethylation is supported by genetic and biochemical data. The genetic evidence for a role of AID in active DNA demethylation has been challenged, but remains strong on balance, especially for B-cells. However, biochemical data argue clearly against AID acting directly on 5mC or its oxidized derivatives. The apparent contradiction between genetic and biochemical data for AID is likely caused by the tacit assumption that AID acts on the modified base directly. Instead, it now appears more likely that AID acts on an unmodified C in the vicinity to trigger repair that involves exchange of one or several 5mC nucleotides with unmodified C nucleotides.

Nucleotide exchange downstream of base modification seems to occur predominantly by the BER pathway in most circumstances. If indeed AID deaminates in the vicinity of the methylated nucleotide, but not at the methylated nucleotide, then only LP-BER can be effective in active DNA demethylation. In contrast, oxidation based DNA demethylation could be completed by either SP-BER or LP-BER. The main BER glycosylases involved in the base excision step are UNG (for U) and TDG (for U, T, 5fC and 5caC). For the excision of 5fC and 5caC, several BER glycosylases have been identified that may cooperate with or replace TDG in some circumstances. In addition to BER, NER and perhaps also ncMMR may have roles in the nucleotide exchange step of active DNA demethylation. At present, neither the choice between BER and alternative

Box 3

Avoiding DNA damage from DNA demethylation

Active DNA demethylation is hazardous. Some intermediates may cause mutations (U and T from C and 5mC deamination) or interfere with transcription (5fC, 5caC). Moreover, uncoordinated introduction of nicks or gaps in both DNA strands may even lead to DNA double strand breaks. How these risks are minimized in the context of active DNA demethylation is only partly understood.

The U and T intermediates of deamination based DNA demethylation are mutagenic in the context of DNA replication. Active DNA demethylation occurs (by definition) in the absence of DNA replication. In these circumstances, U bases can always be identified as illegitimate, and T bases (and not G bases in the opposite strand) are excised by TDG, so that the original and not a mutated sequence is restored.

The 5fC and 5caC intermediates of oxidation based demethylation interfere with transcription [132, 133], and resemble damaged DNA bases (Fig. 6). Their levels are therefore kept low [78, 162]. Rapid elimination of 5fC and 5caC from the genome appears to be aided by physical association between TET enzymes and BER glycosylases [22]. Moreover, formation of 5fC and 5caC is slow enough (rate constant about 1/h, five times slower than the rate of 5hmC formation [69], in the case of TET2 [69]) that TDG can keep up with this rate (Tdg excises 5fC or 5caC at rates of 5/h and 1/h in the absence and presence of product inhibition, respectively [119]). Biochemical reconstitution experiments support this conclusion [81].

Simultaneous base excision repair of sites that are close together could potentially lead to double strand breaks, which are much more deleterious to cells than single strand nicks. Given the low levels of intermediates generated by deamination or oxidation, repair in both strands in close proximity is unlikely. However, fully methylated CpG sites represent a special hazard, and extra safeguards seem to be in place to prevent that both DNA strands are processed simultaneously. 5hmC in a CpG dinucleotide site blocks oxidation of 5mC in the other strand [159]. Should 5mC bases at a CpG site be oxidized simultaneously nonetheless, repair proceeds in a highly coordinated manner to avoid double strand break formation [81]. Possible nucleotide replacement by NER or ncMMR appears to be more hazardous than replacement by BER, because nucleotide replacement tracts are longer and therefore more likely to overlap unless such overlap is actively prevented.

The combined action of AID and ncMMR (that is implied by recent data [28]) appears to pose particularly severe threats to genome integrity. Cooperation between ncMMR and AID occurs physiologically in somatic hypermutation (SHM) and class switch recombination (CSR) [163]. In SHM, mutations are generated not only by deamination, but also by the recruitment of a low fidelity polymerase for the DNA resynthesis step of ncMMR [163]. In CSR, double strand breaks are generated by a combination of AID and UNG2 dependent nicks in one strand and DNA resynthesis in the other strand [163]. How AID and ncMMR could cooperate in active DNA demethylation without these side effects is not clear.

---

repair pathways nor the choice of glycosylases operating within the BER pathway are well understood, and is not fully clear how the hazards of nucleotide replacement are minimized in active DNA demethylation.

The likely involvement of LP-BER and NER and the possible involvement of ncMMR blur the distinction between active DNA demethylation (in the absence of DNA replication) and passive DNA demethylation (in the presence of DNA replication). The indirect active and passive DNA demethylation pathways are mechanistically similar, except that new DNA synthesis is either local or global. Moreover, there are now clear examples for simultaneous active and passive DNA demethylation, for example in the pronuclei of the mammalian zygote [116, 136]. Cooperation between active and passive demethylation pathways may also be expected on theoretical grounds. Parental DNA strands remain methylated after DNA replication, and hemi-methylated DNA is in danger of being remethylated by the ubiquitous maintenance methyltransferase Dnmt1. Therefore, active DNA demethylation may "secure" the results of passive DNA demethyation.

Many interesting questions remain open: how are loci targeted for demethylation? How is the choice between different DNA demethylation pathways made? When is the replacement of nucleotides initiated by AID, when by TETs? To what extent do pathways other than BER (such as NER and ncMMR) contribute to the DNA repair step of DNA demethylation? Which BER glycosylases are required for which physiological demethylation events? How are detrimental effects of active DNA demethylation on DNA integrity minimized? Do some demethylation intermediates, particularly 5hmC, function as epigenetic marks in their own right? To what extent are the phenotypes of TET deficient animals attributable to impaired demethylation, as opposed to a lack of 5hmC? What are the respective contributions of impaired DNA repair and demethylation to the phenotypes of BER glycosylase knockouts? How do defects in genes involved in DNA demethylation lead to the (mostly known, but diverse) phenotypes in development and disease? Why do knockouts of orthologous genes (or combinations of genes) cause seemingly dissimilar phenotypes in different vertebrates, at least in some cases? To what extent can defects be attributed to the dysregulation of specific pathways? Will additional links between metabolism (such as the tricarboxylic acid cycle) and gene regulation by DNA demethylation emerge? The field will stay fascinating.

Review essays

References

1. **Jones PA**. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* **13**: 484–92.

2. **Brown KD, Robertson KD**. 2007. DNMT1 knockout delivers a strong blow to genome stability and cell viability. *Nat Genet* **39**: 289–90.

3. **Li E, Bestor TH, Jaenisch R**. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**: 915–26.

4. **Okano M, Bell DW, Haber DA, Li E**. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**: 247–57.

5. **Ooi SKT, Bestor TH**. 2008. The colorful history of active DNA demethylation. *Cell* **133**: 1145–8.

6. **Ohno R, Nakayama M, Naruse C, Okashita N**, et al. 2013. A replication-dependent passive mechanism modulates DNA demethylation in mouse primordial germ cells. *Development* **140**: 2892–903.

7. **Kagiwada S, Kurimoto K, Hirota T, Yamaji M**, et al. 2013. Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J* **32**: 340–53.

8. **Wu SC, Zhang Y**. 2010. Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol* **11**: 607–20.

9. **Mayer W, Niveleau A, Walter J, Fundele R**, et al. 2000. Demethylation of the zygotic paternal genome. *Nature* **403**: 501–2.

10. **Gu T-P, Guo F, Yang H, Wu H-P**, et al. 2011. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**: 606–10.

11. **Iqbal K, Jin S-G, Pfeifer GP, Szabo PE**. 2011. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci USA* **108**: 3642–7.

12. **Guo JU, Su Y, Zhong C, Ming GL**, et al. 2011. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145**: 423–34.

13. **Ma DK, Jang M-H, Guo JU, Kitabatake Y**, et al. 2009. Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* **323**: 1074–7.

14. **Pastor WA, Aravind L, Rao A**. 2013. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol* **14**: 341–56.

15. **Chen CC, Wang KY, Shen CK**. 2012. The mammalian de novo DNA methyltransferases DNMT3A and DNMT3B are also DNA 5-hydroxymethylcytosine dehydroxymethylases. *J Biol Chem* **287**: 33116–21.

16. **Chen CC, Wang KY, Shen CK**. 2013. DNA 5-methylcytosine demethylation activities of the mammalian DNA methyltransferases. *J Biol Chem* **288**: 9084–91.

17. **Liutkeviciute Z, Kriukiene E, Licyte J, Rudyte M**, et al. 2014. Direct decarboxylation of 5-carboxylcytosine by DNA C5-methyltransferases. *J Am Chem Soc* **136**: 5884–7.

18. **Zhu J-K**. 2009. Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet* **43**: 143–66.

19. **Gehring M, Reik W, Henikoff S**. 2009. DNA demethylation by DNA repair. *Trends Genet* **25**: 82–90.

20. **Maiti A, Drohat AC**. 2011. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem* **286**: 35334–8.

21. **Xue JH, Xu GF, Gu TP, Chen GD**, et al. 2016. Uracil-DNA glycosylase UNG promotes Tet-mediated DNA demethylation. *J Biol Chem* **291**: 731–8.

22. **Muller U, Bauer C, Siegl M, Rottach A**, et al. 2014. TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation. *Nucleic Acids Res* **42**: 8592–604.

23. **Cortellino S, Xu J, Sannai M, Moore R**, et al. 2011. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146**: 67–79.

24. **Hajkova P, Jeffries SJ, Lee C, Miller N**, et al. 2010. Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* **329**: 78–82.

25. **Jin C, Qin T, Barton MC, Jelinek J**, et al. 2015. Minimal role of base excision repair in TET-induced global DNA demethylation in HEK293T cells. *Epigenetics* **10**: 1006–13.

26. **Niehrs C, Schafer A**. 2012. Active DNA demethylation by Gadd45 and DNA repair. *Trends Cell Biol* **22**: 220–7.

27. **Barreto G, Schafer A, Marhold J, Stach D**, et al. 2007. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* **445**: 671–5.

28. **Grin I, Ishchenko AA**. 2016. An interplay of the base excision repair and mismatch repair pathways in active DNA demethylation. *Nucleic Acids Res* **44**: 3713–27.

29. **Hashimoto H, Liu Y, Upadhyay AK, Chang Y**, et al. 2012. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res* **40**: 4841–9.

30. **Muramatsu M, Kinoshita K, Fagarasan S, Yamada S**, et al. 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**: 553–63.

31. **Arakawa H, Hauschild J, Buerstedde JM**. 2002. Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. *Science* **295**: 1301–6.

32. **Petersen-Mahrt SK, Harris RS, Neuberger MS**. 2002. AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature* **418**: 99–103.

33. **Muramatsu M, Sankaranand VS, Anant S, Sugai M**, et al. 1999. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem* **274**: 18470–6.

34. **Rommel PC, Bosque D, Gitlin AD, Croft GF**, et al. 2013. Fate mapping for activation-induced cytidine deaminase (AID) marks non-lymphoid cells during mouse development. *PLoS ONE* **8**: e69208.

35. **Nakanishi Y, Kondo S, Wakisaka N, Tsuji A**, et al. 2013. Role of activation-induced cytidine deaminase in the development of oral squamous cell carcinoma. *PLoS ONE* **8**: e62066.

36. **Kou T, Marusawa H, Kinoshita K, Endo Y**, et al. 2007. Expression of activation-induced cytidine deaminase in human hepatocytes during hepatocarcinogenesis. *Int J Cancer* **120**: 469–76.

37. **Marusawa H**. 2008. Aberrant AID expression and human cancer development. *Int J Biochem Cell Biol* **40**: 1399–402.

38. **Revy P, Muto T, Levy Y, Geissmann F**, et al. 2000. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* **102**: 565–75.

39. **Bhutani N, Brady JJ, Damian M, Sacco A**, et al. 2010. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463**: 1042–7.

40. **Sabag O, Zamir G, Keshet I, Hecht M**, et al. 2014. Establishment of methylation patterns in ES cells. *Nat Struct Mol Biol* **21**: 110–2.

41. **Foshay KM, Looney TJ, Chari S, Mao FF**, et al. 2012. Embryonic stem cells induce pluripotency in somatic cell fusion through biphasic reprogramming. *Mol Cell* **46**: 159–70.

42. **Bhutani N, Decker MN, Brady JJ, Bussat RT**, et al. 2013. A critical role for AID in the initiation of reprogramming to induced pluripotent stem cells. *FASEB J* **27**: 1107–13.

43. **Kumar R, DiMenna L, Schrode N, Liu TC**, et al. 2013. AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes. *Nature* **500**: 89–92.

44. **Habib O, Habib G, Do JT, Moon SH**, et al. 2014. Activation-induced deaminase-coupled DNA demethylation is not crucial for the generation of induced pluripotent stem cells. *Stem Cells Dev* **23**: 209–18.

45. **Shimamoto R, Amano N, Ichisaka T, Watanabe A**, et al. 2014. Generation and characterization of induced pluripotent stem cells from Aid-deficient mice. *PLoS ONE* **9**: e94735.

46. **Dominguez PM, Teater M, Chambwe N, Kormaksson M**, et al. 2015. DNA methylation dynamics of germinal center B cells are mediated by AID. *Cell Rep* **12**: 2086–98.

47. **Hogenbirk MA, Heideman MR, Velds A, van den Berk PC**, et al. 2013. Differential programming of B cells in AID deficient mice. *PLoS ONE* **8**: e69815.

48. **Fritz EL, Rosenberg BR, Lay K, Mihailovic A**, et al. 2013. A comprehensive analysis of the effects of the deaminase AID on the transcriptome and methylome of activated B cells. *Nat Immunol* **14**: 749–55.

49. **Rai K, Huggins IJ, James SR, Karpf AR**, et al. 2008. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* **135**: 1201–12.

50. **Shimoda N, Hirose K, Kaneto R, Izawa T**, et al. 2014. No evidence for AID/MBD4-coupled DNA demethylation in zebrafish embryos. *PLoS ONE* **9**: e114816.

51. **Popp C, Dean W, Feng S, Cokus SJ**, et al. 2010. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**: 1101–5.

52. **Nabel CS, Jia H, Ye Y, Shen L**, et al. 2012. AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat Chem Biol* **8**: 751–8.

53. **Rangam G, Schmitz KM, Cobb AJ, Petersen-Mahrt SK**. 2012. AID enzymatic activity is inversely proportional to the size of cytosine C5 orbital cloud. *PLoS ONE* **7**: e43279.

54. **Larijani M, Frieder D, Sonbuchner TM, Bransteitter R**, et al. 2005. Methylation protects cytidines from AID-mediated deamination. *Mol Immunol* **42**: 599–604.

55. **Abdouni H, King JJ, Suliman M, Quinlan M**, et al. 2013. Zebrafish AID is capable of deaminating methylated deoxycytidines. *Nucleic Acids Res* **41**: 5457–68.

56. **Bransteitter R, Pham P, Scharff MD, Goodman MF**. 2003. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc Natl Acad Sci USA* **100**: 4102–7.

57. **Wijesinghe P, Bhagwat AS**. 2012. Efficient deamination of 5-methylcytosines in DNA by human APOBEC3A, but not by AID or APOBEC3G. *Nucleic Acids Res* **40**: 9206–17.

58. **Shen JC, Rideout WM, 3rd, Jones PA**. 1992. High frequency mutagenesis by a DNA methyltransferase. *Cell* **71**: 1073–80.

59. **Stier I, Kiss A**. 2013. Cytosine-to-uracil deamination by SssI DNA methyltransferase. *PLoS ONE* **8**: e79003.

60. **Franchini DM, Chan CF, Morgan H, Incorvaia E**, et al. 2014. Processive DNA demethylation via DNA deaminase-induced lesion resolution. *PLoS ONE* **9**: e97754.

61. **Santos F, Peat J, Burgess H, Rada C**, et al. 2013. Active demethylation in mouse zygotes involves cytosine deamination and base excision repair. *Epigenetics Chromatin* **6**: 39.

62. **Ramiro AR, Barreto VM**. 2015. Activation-induced cytidine deaminase and active cytidine demethylation. *Trends Biochem Sci* **40**: 172–81.

63. **Fritz EL, Papavasiliou FN**. 2010. Cytidine deaminases: AIDing DNA demethylation? *Genes Dev* **24**: 2107–14.

64. **Xu G-L, Wong, J**. 2015. Oxidative DNA demethylation mediated by Tet enzymes. *Nat Sci Rev* **2**: 318–28.

65. **Tahiliani M, Koh KP, Shen Y, Pastor WA**, et al. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**: 930–5.

66. **Kriaucionis S, Heintz N**. 2009. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**: 929–30.

67. **Ito S, Shen L, Dai Q, Wu SC**, et al. 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**: 1300–3.

68. **Silverman RB**. 2000. *The Organic Chemistry of Enzyme-Catalyzed Reactions*. San Diego, CA: Academic Press.

69. **Hu L, Lu J, Cheng J, Rao Q**, et al. 2015. Structural insight into substrate preference for TET-mediated oxidation. *Nature* **527**: 118–22.

70. **Lorsbach RB, Moore J, Mathew S, Raimondi SC**, et al. 2003. 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**: 637–41.

71. **Wojciechowski M, Rafalski D, Kucharski R, Misztal K**, et al. 2014. Insights into DNA hydroxymethylation in the honeybee from in-depth analyses of TET dioxygenase. *Open Biol* **4**: 140110.

72. **Hashimoto H, Pais JE, Zhang X, Saleh L**, et al. 2014. Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. *Nature* **506**: 391–5.

73. **Pais JE, Dai N, Tamanaha E, Vaisvila R**, et al. 2015. Biochemical characterization of a Naegleria TET-like oxygenase and its application in single molecule sequencing of 5-methylcytosine. *Proc Natl Acad Sci USA* **112**: 4316–21.

74. **Iyer LM, Tahiliani M, Rao A, Aravind L**. 2009. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle* **8**: 1698–710.

75. **Ko M, An J, Bandukwala HS, Chavez L**, et al. 2013. Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* **497**: 122–6.

76. **Jin SG, Zhang ZM, Dunwell TL, Harter MR**, et al. 2016. Tet3 reads 5-carboxylcytosine through its CXXC domain and is a potential guardian against neurodegeneration. *Cell Rep* **14**: 493–505.

77. **Dawlaty MM, Breiling A, Le T, Barrasa MI**, et al. 2014. Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. *Dev Cell* **29**: 102–11.

78. **Wagner M, Steinbacher J, Kraus TF, Michalakis S**, et al. 2015. Age-dependent levels of 5-methyl-, 5-hydroxymethyl-, and 5-formylcytosine in human and mouse brain tissues. *Angew Chem Int Ed Engl* **54**: 12511–4.

79. **Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T**, et al. 2013. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* **152**: 1146–59.

80. **He Y-F, Li B-Z, Li Z, Liu P**, et al. 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**: 1303–7.

81. **Weber AR, Krawczyk C, Robertson AB, Kusnierczyk A**, et al. 2016. Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. *Nat Commun* **7**: 10806.

82. **Hu X, Zhang L, Mao SQ, Li Z**, et al. 2014. Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. *Cell Stem Cell* **14**: 512–22.

83. **Buganim Y, Faddah DA, Jaenisch R**. 2013. Mechanisms and models of somatic cell reprogramming. *Nat Rev Genet* **14**: 427–39.

84. **Gao Y, Chen J, Li K, Wu T**, et al. 2013. Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell* **12**: 453–69.

85. **Chen J, Gao Y, Huang H, Xu K**, et al. 2015. The combination of Tet1 with Oct4 generates high-quality mouse-induced pluripotent stem cells. *Stem Cells* **33**: 686–98.

86. **Lu F, Liu Y, Jiang L, Yamaguchi S**, et al. 2014. Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev* **28**: 2103–19.

87. **Hon GC, Song C-X, Du T, Jin F**, et al. 2014. 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol Cell* **56**: 286–97.

88. **Inoue A, Shen L, Matoba S, Zhang Y**. 2015. Haploinsufficiency, but not defective paternal 5mC oxidation, accounts for the developmental defects of maternal Tet3 knockouts. *Cell Rep* **10**: 463–70.

89. **Dawlaty MM, Ganz K, Powell BE, Hu Y-C**, et al. 2011. Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* **9**: 166–75.

90. **Li Z, Cai X, Cai C-L, Wang J**, et al. 2011. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood* **118**: 4509–18.

91. **Dawlaty MM, Breiling A, Le T, Raddatz G**, et al. 2013. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev Cell* **24**: 310–23.

92. **Dai HQ, Wang BA, Yang L, Chen JJ**, et al. 2016. DNA demethylation by TET dioxygenases controls gastrula patterning by regulating Lefty-1 Nodal signaling. *Nature* **538**: 528–532.

93. **Xu Y, Xu C, Kato A, Tempel W**, et al. 2012. Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. *Cell* **151**: 1200–13.

94. **Li C, Lan Y, Schwartz-Orbach L, Korol E**, et al. 2015. Overlapping requirements for Tet2 and Tet3 in normal development and hematopoietic stem cell emergence. *Cell Rep* **12**: 1133–43.

95. **Langemeijer SMC, Kuiper RP, Berends M, Knops R**, et al. 2009. Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat Genet* **41**: 838–42.

96. **Delhommeau F, Dupont S, Della Valle V, James C**, et al. 2009. Mutation in TET2 in myeloid cancers. *N Engl J Med* **360**: 2289–301.

97. **Kosmider O, Gelsi-Boyer V, Ciudad M, Racoeur C**, et al. 2009. TET2 gene mutation is a frequent and adverse event in chronic myelomonocytic leukemia. *Haematologica* **94**: 1676–81.

98. **Pollard PJ, Briere JJ, Alam NA, Barwell J**, et al. 2005. Accumulation of Krebs cycle intermediates and over-expression of HIF1alpha in tumours which result from germline FH and SDH mutations. *Hum Mol Genet* **14**: 2231–9.

99. **Dang L, White DW, Gross S, Bennett BD**, et al. 2010. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **465**: 966.

100. **Schiesser S, Pfaffeneder T, Sadeghian K, Hackner B**, et al. 2013. Deamination, oxidation, and C-C bond cleavage reactivity of 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxycytosine. *J Am Chem Soc* **135**: 14593–9.

101. **Kaneda M, Okano M, Hata K, Sado T**, et al. 2004. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **429**: 900–3.

102. **Bourc'his D, Xu GL, Lin CS, Bollman B**, et al. 2001. Dnmt3L and the establishment of maternal genomic imprints. *Science* **294**: 2536–9.

Review essays

Review essays

103. **Hata K, Kusumi M, Yokomine T, Li E**, et al. 2006. Meiotic and epigenetic aberrations in Dnmt3L-deficient male germ cells. *Mol Reprod Dev* **73**: 116–22.

104. **Hata K, Okano M, Lei H, Li E**. 2002. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129**: 1983–93.

105. **Schiesser S, Hackner B, Pfaffeneder T, Muller M**, et al. 2012. Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. *Angew Chem Int Ed Engl* **51**: 6516–20.

106. **Iiams V, Desai BJ, Fedorov AA, Fedorov EV**, et al. 2011. Mechanism of the orotidine 5′-monophosphate decarboxylase-catalyzed reaction: importance of residues in the orotate binding site. *Biochemistry* **50**: 8497–507.

107. **Krokan HE, Drablos F, Slupphaug G**. 2002. Uracil in DNA-occurrence, consequences and repair. *Oncogene* **21**: 8935–48.

108. **Raiber EA, Murat P, Chirgadze DY, Beraldi D**, et al. 2015. 5-Formylcytosine alters the structure of the DNA double helix. *Nat Struct Mol Biol* **22**: 44–9.

109. **Szulik MW, Pallan PS, Nocek B, Voehler M**, et al. 2015. Differential stabilities and sequence-dependent base pair opening dynamics of Watson-Crick base pairs with 5-hydroxymethylcytosine, 5-formylcytosine, or 5-carboxylcytosine. *Biochemistry* **54**: 1294–305.

110. **Maiti A, Michelson AZ, Armwood CJ, Lee JK**, et al. 2013. Divergent mechanisms for enzymatic excision of 5-formylcytosine and 5-carboxylcytosine from DNA. *J Am Chem Soc* **135**: 15813–22.

111. **Shibutani T, Ito S, Toda M, Kanao R**, et al. 2014. Guanine-5-carboxylcytosine base pairs mimic mismatches during DNA replication. *Sci Rep* **4**: 5220.

112. **Dai Q, Sanstead PJ, Peng CS, Han D**, et al. 2016. Weakened N3 hydrogen bonding by 5-formylcytosine and 5-carboxylcytosine reduces their base-pairing stability. *ACS Chem Biol* **11**: 470–7.

113. **Lenz SAP, Kellie JL, Wetmore SD**. 2015. Glycosidic bond cleavage in DNA nucleosides: effect of nucleobase damage and activation on the mechanism and barrier. *J Phys Chem B* **119**: 15601–12.

114. **Verdine GL, Bruner SD**. 1997. How do DNA repair proteins locate damaged bases in the genome? *Chem Biol* **4**: 329–34.

115. **Hashimoto H, Zhang X, Cheng X**. 2013. Activity and crystal structure of human thymine DNA glycosylase mutant N140A with 5-carboxylcytosine DNA at low pH. *DNA Repair (Amst)* **12**: 535–40.

116. **Guo F, Li X, Liang D, Li T**, et al. 2014. Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell* **15**: 447–58.

117. **Cortazar D, Kunz C, Selfridge J, Lettieri T**, et al. 2011. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* **470**: 419–23.

118. **Nemec AA, Wallace SS, Sweasy JB**. 2010. Variant base excision repair proteins: contributors to genomic instability. *Semin Cancer Biol* **20**: 320–8.

119. **Schomacher L, Han D, Musheev MU, Arab K**, et al. 2016. Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation. *Nat Struct Mol Biol* **23**: 116–24.

120. **Vartanian V, Lowell B, Minko IG, Wood TG**, et al. 2006. The metabolic syndrome resulting from a knockout of the NEIL1 DNA glycosylase. *Proc Natl Acad Sci USA* **103**: 1864–9.

121. **Chakraborty A, Wakamiya M, Venkova-Canova T, Pandita RK**, et al. 2015. Neil2-null mice accumulate oxidized DNA bases in the transcriptionally active sequences of the genome and are susceptible to innate inflammation. *J Biol Chem* **290**: 24636–48.

122. **Hildrestrand GA, Neurauter CG, Diep DB, Castellanos CG**, et al. 2009. Expression patterns of Neil3 during embryonic brain development and neoplasia. *BMC Neurosci* **10**: 45.

123. **Sejersted Y, Hildrestrand GA, Kunke D, Rolseth V**, et al. 2011. Endonuclease VIII-like 3 (Neil3) DNA glycosylase promotes neurogenesis induced by hypoxia-ischemia. *Proc Natl Acad Sci USA* **108**: 18802–7.

124. **Nilsen H, Rosewell I, Robins P, Skjelbred CF**, et al. 2000. Uracil-DNA glycosylase (UNG)-deficient mice reveal a primary role of the enzyme during DNA replication. *Mol Cell* **5**: 1059–65.

125. **Fortini P, Dogliotti E**. 2007. Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. *DNA Repair* **6**: 398–409.

126. **Hollander MC, Sheikh MS, Bulavin DV, Lundgren K**, et al. 1999. Genomic instability in Gadd45a-deficient mice. *Nat Genet* **23**: 176–84.

127. **Tamura RE, de Vasconcellos JF, Sarkar D, Libermann TA**, et al. 2012. GADD45 proteins: central players in tumorigenesis. *Curr Mol Med* **12**: 634–51.

128. **Li Z, Gu TP, Weber AR, Shen JZ**, et al. 2015. Gadd45a promotes DNA demethylation through TDG. *Nucleic Acids Res* **43**: 3986–97.

129. **Engel N, Tront JS, Erinle T, Nguyen N**, et al. 2009. Conserved DNA methylation in Gadd45a(-/-) mice. *Epigenetics* **4**: 98–9.

130. **Schmitz K-M, Schmitt N, Hoffmann-Rohrer U, Schafer A**, et al. 2009. TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. *Mol Cell* **33**: 344–53.

131. **Bellon SF, Coleman JH, Lippard SJ**. 1991. DNA unwinding produced by site-specific intrastrand cross-links of the antitumor drug cis-diamminedichloroplatinum(II). *Biochemistry* **30**: 8026–35.

132. **Kellinger MW, Song CX, Chong J, Lu XY**, et al. 2012. 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat Struct Mol Biol* **19**: 831–3.

133. **Wang L, Zhou Y, Xu L, Xiao R**, et al. 2015. Molecular basis for 5-carboxycytosine recognition by RNA polymerase II elongation complex. *Nature* **523**: 621–5.

134. **Fousteri M, Mullenders LHF**. 2008. Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. *Cell Res* **18**: 73–84.

135. **Jiricny J**. 2006. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* **7**: 335–46.

136. **Shen L, Inoue A, He J, Liu Y**, et al. 2014. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell Stem Cell* **15**: 459–70.

137. **Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O**, et al. 2011. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**: 11–24.

138. **An J, Gonzalez-Avalos E, Chawla A, Jeong M**, et al. 2015. Acute loss of TET function results in aggressive myeloid cancer in mice. *Nat Commun* **6**: 10071.

139. **Zhao Z, Chen L, Dawlaty MM, Pan F**, et al. 2015. Combined Loss of Tet1 and Tet2 promotes B cell, but not myeloid malignancies, in mice. *Cell Rep* **13**: 1692–704.

140. **Ge L, Zhang RP, Wan F, Guo DY**, et al. 2014. TET2 plays an essential role in erythropoiesis by regulating lineage-specific genes via DNA oxidative demethylation in a zebrafish model. *Mol Cell Biol* **34**: 989–1002.

141. **Gjini E, Mansour MR, Sander JD, Moritz N**, et al. 2015. A zebrafish model of myelodysplastic syndrome produced through tet2 genomic editing. *Mol Cell Biol* **35**: 789–804.

142. **Perez C, Martinez-Calle N, Martin-Subero JI, Segura V**, et al. 2012. TET2 mutations are associated with specific 5-methylcytosine and 5-hydroxymethylcytosine profiles in patients with chronic myelomonocytic leukemia. *PLoS One* **7**: e31605.

143. **Yamazaki J, Jelinek J, Lu Y, Cesaroni M**, et al. 2015. TET2 mutations affect non-CpG island DNA methylation at enhancers and transcription factor-binding sites in chronic myelomonocytic leukemia. *Cancer Res* **75**: 2833–43.

144. **Laukka T, Mariani CJ, Ihantola T, Cao JZ**, et al. 2016. Fumarate and succinate regulate expression of hypoxia-inducible genes via TET enzymes. *J Biol Chem* **291**: 4256–65.

145. **Sciacovelli M, Goncalves E, Johnson TI, Zecchini VR**, et al. 2016. Fumarate is an epigenetic modifier that elicits epithelial-to-mesenchymal transition. *Nature* **537**: 544–547.

146. **Xu W, Yang H, Liu Y, Yang Y**, et al. 2011. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell* **19**: 17–30.

147. **Xiao M, Yang H, Xu W, Ma S**, et al. 2012. 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**: 1326–38.

148. **Henson DE, Block G, Levine M**. 1991. Ascorbic acid: biologic functions and relation to cancer. *J Natl Cancer Inst* **83**: 547–50.

149. **Blaschke K, Ebata KT, Karimi MM, Zepeda-Martinez JA**, et al. 2013. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**: 222–6.

150. **Chung T-L, Brena RM, Kolle G, Grimmond SM**, et al. 2010. Vitamin C promotes widespread yet specific DNA demethylation of the epigenome in human embryonic stem cells. *Stem Cells* **28**: 1848–55.

151. **Yin R, Mao SQ, Zhao B, Chong Z**, et al. 2013. Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J Am Chem Soc* **135**: 10396–403.

152. **Young JI, Zuchner S, Wang G**. 2015. Regulation of the epigenome by Vitamin C. *Annu Rev Nutr* **35**: 545–64.

153. **Costa Y, Ding J, Theunissen TW, Faiola F**, et al. 2013. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* **495**: 370–4.

154. **Doege CA, Inoue K, Yamashita T, Rhee DB**, et al. 2012. Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* **488**: 652–5.

155. **Sajadian SO, Tripura C, Samani FS, Ruoss M**, et al. 2016. Vitamin C enhances epigenetic modifications induced by 5-azacytidine and cell cycle arrest in the hepatocellular carcinoma cell lines HLE and Huh7. *Clin Epigenetics* **8**: 46.

156. **Tan L, Shi YG**. 2012. Tet family proteins and 5-hydroxymethylcytosine in development and disease. *Development* **139**: 1895–902.

157. **Liu CK, Hsu CA, Abbott MT**. 1973. Catalysis of three sequential dioxygenase reactions by thymine 7-hydroxylase. *Arch Biochem Biophys* **159**: 180–7.

158. **Boorstein RJ, Cummings A, Jr., Marenstein DR, Chan MK**, et al. 2001. Definitive identification of mammalian 5-hydroxy-methyluracil DNA N-glycosylase activity as SMUG1. *J Biol Chem* **276**: 41991–7.

159. **Hashimoto H, Liu Y, Upadhyay AK, Chang Y**, et al. 2012. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res* **40**: 4841–9.

160. **Rusmintratip V, Sowers LC**. 2000. An unexpectedly high excision capacity for mispaired 5-hydroxymethyluracil in human cell extracts. *Proc Natl Acad Sci USA* **97**: 14183–7.

161. **Pfaffeneder T, Spada F, Wagner M, Brandmayr C**, et al. 2014. Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat Chem Biol* **10**: 574–81.

162. **Chowdhury B, Cho IH, Hahn N, Irudayaraj J**. 2014. Quantification of 5-methylcytosine, 5-hydroxymethylcytosine and 5-carboxylcytosine from the blood of cancer patients by an enzyme-based immunoassay. *Anal Chim Acta* **852**: 212–7.

163. **Zanotti KJ, Gearhart PJ**. 2016. Antibody diversification caused by disrupted mismatch repair and promiscuous DNA polymerases. *DNA Repair* **38**: 110–6.

Review essays