J Biosci (2020) 45:21
DOI: 10.1007/s12038-019-9973-4

Review

TET methylcytosine oxidases: new insights from a decade of research

Chan-Wang J. Lio[1,†], Xiaojing Yue[1,†], Isaac F. López-Moyado[1,2,3,†], Mamta Tahiliani[4,5], L. Aravind[6] and Anjana Rao[1,2,3,7,8]*

[1]*Division of Signaling and Gene Expression, La Jolla Institute for Immunology, La Jolla, CA 92037, USA*

[2]*Bioinformatics and Systems Biology Graduate Program, University of California San Diego, La Jolla, CA 92093, USA*

[3]*Sanford Consortium for Regenerative Medicine, La Jolla, CA 92093, USA*

[4]*Skirball Institute of Biomolecular Medicine, New York University School of Medicine, New York, NY 10012, USA*

[5]*Department of Biology, New York University, New York, NY 10003, USA*

[6]*National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20892, USA*

[7]*Department of Pharmacology, University of California San Diego, La Jolla, CA 92093, USA*

[8]*Moores Cancer Center, University of California San Diego, La Jolla, CA 92093, USA*

**Corresponding author (Email, arao@lji.org)*

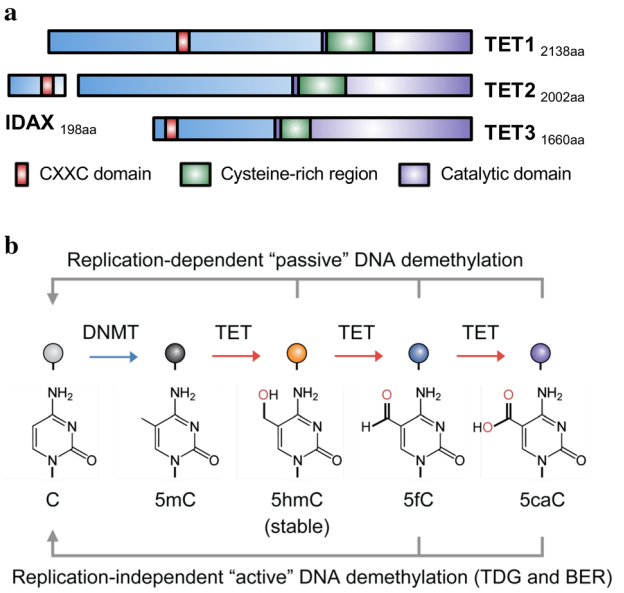
[†]*These authors contributed equally to this work.*

Published online 22 January 2020

In mammals, DNA methyltransferases transfer a methyl group from S-adenosylmethionine to the 5 position of cytosine in DNA. The product of this reaction, 5-methylcytosine (5mC), has many roles, particularly in suppressing transposable and repeat elements in DNA. Moreover, in many cellular systems, cell lineage specification is accompanied by DNA demethylation at the promoters of genes expressed at high levels in the differentiated cells. However, since direct cleavage of the C-C bond connecting the methyl group to the 5 position of cytosine is thermodynamically disfavoured, the question of whether DNA methylation was reversible remained unclear for many decades. This puzzle was solved by our discovery of the TET (Ten-Eleven Translocation) family of 5-methylcytosine oxidases, which use reduced iron, molecular oxygen and the tricarboxylic acid cycle metabolite 2-oxoglutarate (also known as α-ketoglutarate) to oxidise the methyl group of 5mC to 5-hydroxymethylcytosine (5hmC) and beyond. TET-generated oxidised methylcytosines are intermediates in at least two pathways of DNA demethylation, which differ in their dependence on DNA replication. In the decade since their discovery, TET enzymes have been shown to have important roles in embryonic development, cell lineage specification, neuronal function and cancer. We review these findings and discuss their implications here.

Keywords. DNA methylation (5mC); 5-hydroxymethylcytosine (5hmC); DNA cytosine modifications; Ten-Eleven Translocation (TET); epigenetics

This article is part of the Topical Collection: Chromatin Biology and Epigenetics.



◄**Figure 1.** Ten-Eleven Translocation (TET) proteins and DNA modification. (**a**) TET family proteins. Mammalian genomes encode three members of the TET/JBP family: TET1, TET2, and TET3. The diagram depicts the domain structures and the length in amino acids (aa) of human TET proteins. The CXXC domains of TET1 and TET3 (red) bind unmethylated CpG sequences in DNA. Note that during evolution, the CXXC domain of primordial TET2 was separated from the TET2 catalytic domain due to chromosomal inversion and evolved as a different gene (IDAX or CXXC4). All three TET proteins contain cysteine-rich domains (green) followed by a C-terminal catalytic domain (purple). (**b**) TET-mediated DNA modifications and demethylation. DNA methyltransferases (DNMT) methylate unmodified cytosines (C) to yield 5-methylcytosine (5mC). TET proteins can successively oxidize 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). Among these three oxidized methylcytosines (oxi-mC), 5hmC is a stable modification and is the most abundant, accounting for ∼1–10% of 5mC depending on the cell type, while 5fC and 5caC are ∼100-fold and ∼1000-fold less abundant than 5hmC. All three oxi-mCs are intermediates for DNA demethylation. During DNA replication, the 5mC at the CpG motif on the template strand pairs with unmodified CpG on the newly synthesized strand, resulting in the hemi-methylated CpG motif. The maintenance methyltransferase complex, DNMT1/UHRF1 binds to the hemi-methylated CpG and rapidly restores methylation on the CpG on the newly synthesized DNA, to restore symmetrical CpG methylation. In contrast, the presence of oxi-mCs in the template strand inhibits the binding of DNMT1/UHRF1 to hemi-modified CpGs, thus preventing methylation of CpGs in the newly synthesized strand. This process is known as 'passive' DNA demethylation (top arrows). Additionally, 5fC and 5caC can be recognized and removed by thymine DNA glycosylase (TDG). The abasic site will be repaired by the base-excision repair (BER) system and replaced by an unmodified cytosine, a process termed 'active' (replication-independent) DNA demethylation (bottom arrows).

1. Introduction

The biochemical activity of ten-eleven translocation proteins was reported ten years ago in a collaboration between the Aravind and Rao labs (Iyer *et al.* 2009; Tahiliani *et al.* 2009). In this article, we review these and subsequent findings in the field, with a focus on published studies from our labs.

2. Overview of the evolution of TET/JBP proteins

In 2009, the three mammalian members of the TET family were predicted to be members of the large superfamily of 2-oxoglutarate (2OG)- and Fe(II)-dependent (2OGFe) dioxygenases (Iyer *et al.* 2009). TET proteins are the animal homologs of the kinetoplastid JBPs (base J-binding proteins) which oxidise the methyl group of thymine to yield 5-hydroxyuracil (5hU), which is subsequently modified with a sugar moiety to yield Base J (Yu *et al.* 2007; Iyer *et al.* 2009; Iyer *et al.* 2013; Bullard *et al.* 2014). In contrast, TET enzymes were predicted to oxidize 5-methylcytosine (5mC), also a 5-methyl pyrimidine, because two of the three vertebrate TET proteins—TET1 and TET3—contain a CXXC domain, known to bind unmethylated Cytosine-Guanine (CpG) sequences. Thus, TET proteins were predicted to act on the methyl group of 5mC rather than that of thymine (Iyer *et al.* 2009). Although TET2 does not currently possess a CXXC domain, the primordial TET2 did contain such a domain; however, the CXXC and catalytic domains of TET2 were separated during evolution through a chromosomal inversion (Ko *et al.* 2013) (figure 1a).

The prediction that TET proteins were 5-methylcytosine oxidases was experimentally verified in 2009 (Tahiliani *et al.* 2009). The use of recombinant TET proteins confirmed that TET proteins not only oxidized 5mC to 5-hydroxymethylcytosine (5hmC) (Iyer *et al.* 2009; Tahiliani *et al.* 2009) but also carried out two additional oxidations, converting 5hmC to 5-formyl and 5-carboxylcytosine (5fC and 5caC, respectively) (Tahiliani *et al.* 2009; He *et al.* 2011; Ito *et al.* 2011; Crawford *et al.* 2016) (figure 1b). These oxidised methylcytosines (oxi-mC) are intermediates in at least two pathways of DNA demethylation as described below.

The TETs and the JBPs define a distinct family within the double-stranded β-helix fold 2OGFe-dioxygenase superfamily. Within the superfamily they are more closely related to the AlkB family, members of which specialize in the oxidative repair of N^6-alkyl adducts to adenine and the resetting of N^6-methyladenosine marks in eukaryotic DNA and RNA (Aravind and Koonin 2001; Iyer *et al.* 2016). This suggests that both the TET/JBP and AlkB families diversified as part of an ancient radiation of nucleic acid-modifying 2OGFe-dioxygenases. Indeed, both share certain common features in their nucleic acid binding interface (Pastor *et al.* 2013).

The TET/JBP family additionally includes members from several bacteriophages, certain bacteria and diverse eukaryotes such as the filamentous fungi, the chlorophyte algae and basal land plants, and the heterolobosean such as *Naegleria* (Iyer *et al.* 2009; Iyer *et al.* 2013). Phylogenetic analysis indicates that the origin of TET/JBP family lies in the bacteriophages, where they are part of the highly diverse DNA-modification systems typical of DNA phages. In the phages, DNA modifications by TET/JBP enzymes are likely to help in evading host restriction and marking the genome for packaging into the phage-head. Notably, in several phages, the 5-hydroxymethylpyrimidine is further modified by phosphorylation by a P-loop kinase and is used as an intermediate for the generation of hypermodified bases (Iyer *et al.* 2009; Iyer *et al.* 2013; Lee *et al.* 2018).

The bacterial and phage versions already show a divergence into two types which might respectively act on 5mC and T. These appear to have been laterally transferred on more than one occasion to eukaryotes to give rise to their TET/JBP proteins (Iyer *et al.* 2009; Iyer *et al.* 2013). Interestingly, TET/JBP proteins are also encoded by intracellular pathogenic bacteria such as *Legionella* and related genera. These proteins are predicted to function as effectors that are delivered into the eukaryotic host cell to modify its DNA. Thus, other than direct transfer from phages, such endo-parasitic bacteria might have also served as a conduit for the transfer of TET/JBP genes to eukaryotes (Iyer *et al.* 2009; Iyer *et al.* 2013). A comparable scenario has been proposed for the origin of the histone methylase H3K79 methylase Dot1 from a *Legionella* effector secreted into eukaryotic host cells (Aravind *et al.* 2011).

3. Enzymatic activities of TET proteins

Like all 2OGFe-dioxygenases, TET enzymes utilize 2OG, reduced iron (Fe(II)) and both atoms of molecular oxygen, to generate their oxidised substrates, with CO_2 and succinate as byproducts (Hausinger 2004). Succinate, which structurally resembles 2OG, is an inhibitor of many 2OGFe-dioxygenases, including the TET enzymes (Xiao *et al.* 2012), whereas Vitamin C, which likely facilitates the reduction of Fe(III) at the active site back to Fe(II), is an activator of these enzymes (Blaschke *et al.* 2013; Yue *et al.* 2016).

A major function of mammalian TET proteins is to facilitate DNA demethylation through the production of oxi-mC through both passive (replication-dependent) and active (replication-independent) mechanisms (figure 1b). The first pathway relies on the fact that the maintenance DNA methyltransferase, DNMT1, efficiently methylates hemi-methylated CpGs, in which 5mC is present across from the unmethylated cytosine on the newly-replicated strand. However, DNMT1 is much less efficient at methylating the unmodified CpGs on newly replicated DNA strands if an oxi-mC (rather than 5mC) is present on the template strand (Hashimoto *et al.* 2012; Otani *et al.* 2013). This process of TET-dependent 'passive' DNA demethylation displays an absolute requirement for replication and for TET catalytic activity, and may be the major process that operates to demethylate the promoters and enhancers of genes that characterize specific cellular lineages during the process of cell lineage specification (Inoue and Zhang 2011; Lio *et al.* 2019). A second, replication-independent, mechanism of DNA demethylation relies on the ability of the DNA repair enzyme thymine DNA glycosylase (TDG) to excise 5fC and 5caC from hemi-modified DNA strands, a process that requires base excision repair to replace the original 5fC or 5caC with an unmodified cytosine (He *et al.* 2011; Maiti and Drohat 2011). This mechanism appears to make only a minor contribution, if any, to DNA demethylation in replicating cells. Moreover, TET-dependent active demethylation in the zygote is unaffected by TDG deletion, suggesting the existence of additional active demethylation pathways downstream of TET-mediated oxidation (Guo *et al.* 2014). Intriguingly, a 5caC decarboxylase activity has been claimed in mouse embryonic stem cells (mESC) (Schiesser *et al.* 2012), although currently, there are no likely candidates in the human genome for such an activity.

Notably, most eukaryotes that possess one or more genes encoding a TET-like member of the TET/JBP-family proteins also code for a DNA methyltransferase (DNMT) gene (Iyer *et al.* 2009; Iyer *et al.* 2011), suggesting a strong functional link between these TET-like enzymes and DNA methylation in eukaryotes. In eukaryotes, other than animals and kinetoplastids, the TET/JBP enzymes of the amoeba *Naegleria*, the

mushroom *Coprinopsis cinerea* and the chlorophyte alga *Chlamydomonas reinhardtii* have been biochemically characterized. Both *Naegleria* and *C. cinerea* TETs produce 5hmC, 5fC and 5caC in differing proportions using 5mC as a substrate (Chavez *et al.* 2014; Zhang *et al.* 2014). In fungi like *C. cinerea*, *TET* genes are genomically linked to novel transposon families, which belong to the so-called Kyajuka-Dileera-Zisupton class of transposons. *TET* genes have probably been widely disseminated across the chromosomes of the fungi by these transposons (Iyer *et al.* 2014), and their protein products appear to have a role in regulating the activity of the linked transposons in addition to marking certain regions of the chromatin. The TET from the yeast *Schizosaccharomyces pombe* is catalytically inactive but might have a role in inducing certain epigenetic states via a non-enzymatic mechanism (Iyer *et al.* 2014).

An interesting recent finding was that one of the TET enzymes from the green alga *C. reinhardtii* utilizes ascorbate instead of 2OG as its essential co-substrate *in vitro*. The enzyme, CMD1, produces a mixture of stereoisomers of 5-glyceryl-methylcytosine (5gmC), in which the glyceryl moiety is linked to the -CH_2 group at the 5 position of cytosine (Xue *et al.* 2019). Like the oxi-mCs generated by mammalian and fungal TET/JBP enzymes, 5gmC antagonized the repressive effects of DNA cytosine methylation; mutants lacking CMD1 showed increased cytosine methylation and decreased expression of two genes encoding LHCSR3 (light-harvesting complex stress-related protein 3), a complex that is required for growth under conditions of high light intensity (Aravind *et al.* 2019; Xue *et al.* 2019).

4. 5hmC is present in euchromatin and is enriched at expressed genes and active enhancers

The genomes of most mammalian cell types can be roughly divided into euchromatic and heterochromatic compartments (Dekker *et al.* 2013), which correspond to actively transcribed and transcriptionally silent regions of the genome. These compartments were originally defined by cytology and then later by immunocytochemistry, but have recently become amenable to definition using an unbiased genome-wide chromosome conformation capture method known as Hi-C. Briefly, principal component analysis of the interaction matrix obtained from Hi-C data can be used to partition the genome into A and B compartments that correspond, respectively, to euchromatin and heterochromatin (Lieberman-Aiden *et al.* 2009). Euchromatin is defined by positive PC1 values and high gene density; it contains expressed genes whose promoters bear the 'active' histone modification H3K4 trimethylation (H3K4me3), and replicates early during S phase (van Steensel and Belmont 2017). In contrast the heterochromatic compartment is gene-poor and transcriptionally silent, replicates during late S phase, is enriched for histone 3 lysine 9 di- and tri- methylation (H3K9me2 and me3), and is associated with the nuclear lamina (van Steensel and Belmont 2017).

In all cell types examined, 5hmC is most highly enriched in gene bodies of the most highly expressed genes, and also at the most active enhancers defined by the highest levels of histone 3 lysine 4 mono-methylation (H3K4me1) and histone 3 lysine 27 acetylation (H3K27Ac) (Tsagaratou *et al.* 2014; Lio *et al.* 2019) (figure 2a). In contrast, the TET substrate 5mC is present throughout the genome, in both euchromatin and heterochromatin. To determine the extent to which 5hmC was present in the heterochromatic, transcriptionally silent Hi-C B compartment, we integrated 5hmC mapping data from TAB-seq (Tet-Assisted Bisulfite Sequencing) (Hon *et al.* 2014) and CMS-IP (Cytosine-5-Methylene-Sulfonate Immunoprecipitation) (Huang *et al.* 2014) with Hi-C data from the same cell type—mouse embryonic stem cells (mESC). The data showed, unambiguously, that the bulk of 5hmC was in the euchromatic Hi-C A compartment (with similar observations in haematopoietic stem/precursor cells, pro-B cells and natural-killer-T/NKT cells), as expected from the known overlap of 5hmC-containing regions with transcribed genes and active enhancers (Lopez-Moyado *et al.* 2019).

5. Dynamic changes in 5mC and 5hmC at *de novo* enhancers during signal-dependent cell activation and differentiation

Studies in many different systems have established the general principle that transcription factors recruit TET enzymes to enhancers, where they deposit 5hmC and facilitate DNA demethylation (figure 2b). In the following sections, we describe the roles of TET proteins at three different types of immune cell enhancers examined in the Rao lab, which control a developmental switch in immature B cells and two signal-dependent processes in mature B cells and in T 'regulatory' cells respectively.

TET proteins mediate 5hmC deposition and DNA demethylation at the Igκ locus during B cell development: Rearrangement of the immunoglobulin light

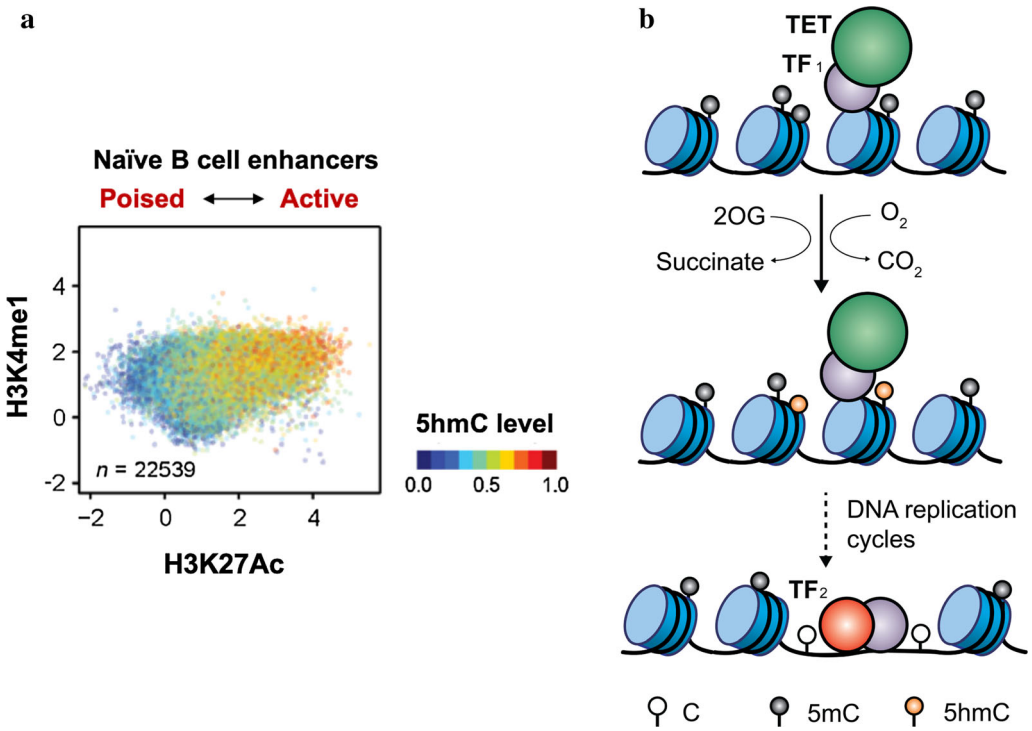


Figure 2. Regulation of enhancers by TET proteins. (**a**) 5hmC levels at enhancers show a strong positive correlation with enhancer activity. The diagram depicts all enhancers in naïve mouse B cells (n=22,539), ranked according to their relative levels of H3-lysine 4-monomethylation (H3K4me1; a mark for most enhancers) and H3-lysine 27-acetylation (H3K27Ac; a mark for enhancer activity). The color indicates the relative enrichment of 5hmC. In general, active enhancers bearing both marks (right) are enriched in 5hmC relative to poised enhancers bearing only the K3K4me1 mark (left). The figure was adapted from Lio *et al.* (2019) with permission. (**b**) Working model for TET-mediated enhancer regulation. Pioneer transcription factors (TF_1, purple circle) are able to bind to nucleosomes at enhancers and recruit TET proteins. Using 2-oxoglutarate (2OG; also known as alpha-ketoglutarate), reduced iron (Fe(II)) and O_2, TET proteins oxidize 5mC into 5hmC at CpG motifs around the enhancer, releasing succinate and CO_2. After rounds of DNA replication, the CpG motifs become demethylated and the enhancer becomes more accessible for binding of additional transcription factors (TF_2, orange circle).

chain including the kappa chain (Igκ) occurs during the pro-B to pre-B switch in early B cell development, and is required for the expression of immunoglobulin M (IgM) on the surface of mature B cells (Hamel *et al.* 2014). At least 3 Eκ enhancers are known to be important for germline Igκ locus transcription, a prerequisite for Igκ chain rearrangement: an intronic enhancer (iEκ), a 3′ enhancer (3′Eκ) and a distal enhancer (dEκ) (Hamel *et al.* 2014). TET proteins are recruited to and deposit 5hmC at the 3′ and distal Eκ enhancers, which contain 2 and 3 CpG sequences respectively; the intronic κ enhancer, which is essential for κ chain rearrangement, does not contain any CpGs and so is unlikely to be a target of regulation by TET proteins, given that the vast majority of 5mC occurs symmetrically on CpGs (Lio *et al.* 2016).

We generated mice in which deletion of the *Tet2* and *Tet3* genes was induced with *Mb1-Cre*, which is expressed at the early pro-B stage. Using pro-B cells from these mice, we showed that TET proteins are required for Igκ germline transcription and rearrangement by rendering the Igκ enhancers accessible. Mechanistically, TET2 co-immunoprecipitates under stringent conditions with PU.1 and E2A, two transcription factors essential for Igκ rearrangement and the pro-B to pre-B cell transition. The evidence supports a mechanism whereby TET proteins are recruited by PU.1 to the Igκ enhancers, and the associated increase in chromatin accessibility at the enhancers permits E2A and TET-induced IRF4 to bind the enhancers and facilitate subsequent germline transcription of the Igκ locus (Lio *et al.* 2016).

TET proteins act at activation-dependent 'de novo' enhancers to facilitate AID expression and class switch recombination (CSR) in mature B cells: To avoid complications arising from TET deletion during development, we deleted the *Tet2* and *Tet3* genes in mature B cells using Cre-ERT2, a tamoxifen-inducible fusion of Cre recombinase with the estrogen receptor ligand-binding domain. This inducible system permits

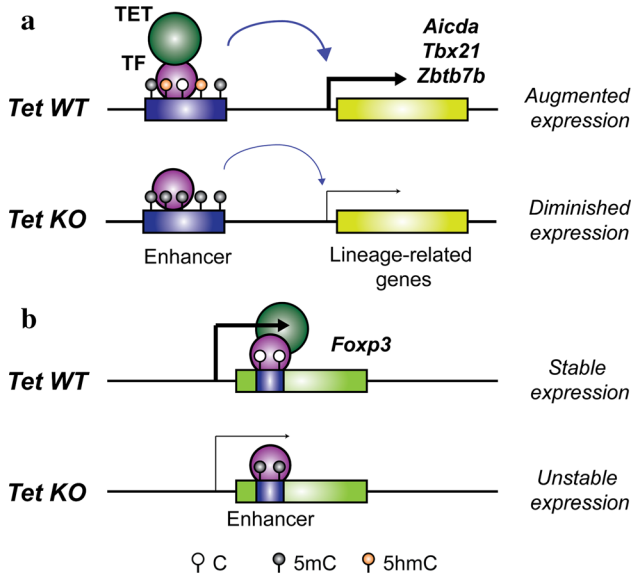


Figure 3. Function of TET proteins in immune system. (**a**) TET proteins are required for the full potential of enhancers. During T cell development and B cell activation, transcription factors (TFs) recruit TET proteins to the key enhancers that promote the expression of lineage-related genes (*Tbx21* and *Zbtb7b* in T cells; *Aicda* in B cells) (Tsagaratou *et al.* 2014; Lio *et al.* 2019). TET proteins oxidize and demethylate enhancers, augmenting gene expression. In the absence of TET proteins, the inability to demethylate enhancers results in decreased gene expression, potentially by affecting chromatin conformation and the binding of additional transcription factors. (**b**) TET proteins are required for stable gene expression. A variety of transcription factors recruit TET proteins and assemble at the intronic enhancer (*CNS2*) of *Foxp3*, the lineage-defining transcription factor for regulatory T (Treg) cells. This results in the demethylation of ~12 CpGs located in the *CNS2* enhancer, a process central to establishing and maintaining the stable expression of *Foxp3* (Yue *et al.* 2016).

a detailed kinetic analysis of 5hmC deposition, DNA demethylation and transcriptional and chromatin changes occurring over a four-day time period (Lio *et al.* 2019). The experiments showed that TET enzymes regulate CSR in mature B cells activated with lipopolysaccharide (LPS) and Interleukin-4 (IL-4). Briefly, B cell activation results in rapid upregulation of the basic region/leucine zipper (bZIP) transcription factor BATF, and later upregulation of the activation-induced cytidine deaminase (AID). Both BATF and AID are essential for CSR, a process in which B cells replace the IgM-encoding exons with those encoding other antibody isotypes such as IgG1 or IgA. In the absence of TET proteins, CSR was reduced by 50%, but reconstitution with catalytically active, but not inactive, AID fully reconstituted CSR. The mechanism involves recruitment of TET proteins to at least two activation-dependent ('*de novo*') enhancers in the *Aicda* locus by BATF; in the absence of BATF, TET proteins were unable to mediate the progressive 5hmC deposition and DNA demethylation seen at these loci in wildtype cells (figure 3a).

TET proteins regulate an intronic enhancer required for the stable expression of FOXP3 in T regulatory cells: Regulatory T (Treg) cells are a minor subpopulation of T cells that are critical for immune homeostasis and prevention of autoimmune disease (Sakaguchi *et al.* 2008; Lio and Hsieh 2011; Josefowicz *et al.* 2012). The lineage-determining transcription factor for Treg cells is FOXP3; germline mutations in FOXP3 in either mice or humans, as well as induced deletion of the *Foxp3* gene in healthy adult mice, leads to fulminant autoimmune disease (Sakaguchi *et al.* 2008; Josefowicz *et al.* 2012). Loss of TET function does not impair the development of thymic Treg cells, but greatly impairs the stability of *Foxp3* expression through cell division (Yue *et al.* 2016). The stability of *Foxp3* expression is controlled by an intronic enhancer, *CNS2*, within the *Foxp3* locus, in a manner linked to its DNA methylation status (Floess *et al.* 2007; Zheng *et al.* 2010; Feng *et al.* 2014; Li *et al.* 2014). *CNS2* is fully methylated in naïve T cells but mainly demethylated in Treg cells (Floess *et al.* 2007). The demethylation is controlled by TET proteins, since Treg cells from mice lacking *Tet2* and *Tet3* (or *Tet1* and *Tet2*) show DNA hypermethylation at *CNS2* and consequent loss of *Foxp3* expression as a function of cell division (Yang *et al.* 2015; Yue *et al.* 2016) (figure 3b).

Tregs can be generated *in vitro* from naïve T cells by culturing them in the presence of TGFβ (Chen *et al.* 2003) and/or retinoic acid (RA) (Benson *et al.* 2007); these cells have been termed 'induced' Tregs (iTregs). However, iTregs generated under these conditions do not show demethylation of *CNS2* (Floess *et al.* 2007; Yue *et al.* 2016). Rather, addition of the TET activator Vitamin C (Blaschke *et al.* 2013) to cultures of naïve T cells with TGFβ and/or RA results in full demethylation of *CNS2* and a substantial increase in the stability of FOXP3 expression, compared to iTregs cultured with TGFβ or TGFβ + RA alone, in both mouse and human (Sasidharan Nair *et al.* 2016; Yue *et al.* 2016). Moreover, inhibition of the Vitamin C transporter reverses the demethylation status of *CNS2*, both in Vitamin C-treated iTregs *in vitro* and in peripheral Tregs generated *in vivo* (Sasidharan Nair *et al.* 2016).

6. Association of TET loss-of-function with cancer

In mouse models developed in the Rao lab, deletion of the *Tet2* and *Tet3* genes in developing T cells using *CD4Cre* resulted in the rapid oligoclonal expansion of a normally minor T cell population known as NKT cells, which recognize lipid antigens presented on a non-classical major histocompatibility complex protein (CD1d) and undergo controlled proliferation rather than being deleted in the thymus due to self-reactivity (Tsagaratou *et al.* 2017). The expansion is quickly followed by the development of aggressive transmissible T cell lymphomas in 100% of mice, which show various hallmarks of cancer, including DNA damage, and the mice succumb within 5–8 weeks (Tsagaratou *et al.* 2017; Lopez-Moyado *et al.* 2019). Similarly, deletion of both *Tet2* and *Tet3* in B cells using *Mb1-Cre* results in a fully-penetrant B cell lymphoma that arises from a few surviving B cells in these mice, and is fatal within 5 months (Lio *et al.* 2016). In both cases, deletion of either the *Tet2* or *Tet3* genes alone resulted in a less dramatic phenotype, suggesting that profound TET deficiency was necessary. We proved this point in a different model system in which the *Tet2* gene was disrupted in the germline and the *Tet3* gene was inducibly deleted (i.e. adult *Tet2* $^{-/-}$ *Tet3* $^{fl/fl}$ *Mx1-Cre* and *Tet2*$^{-/-}$ *Tet3* $^{fl/fl}$ *Cre-ERT2* mice, in which Cre recombinase is induced by injection of polyI:polyC and tamoxifen respectively (An *et al.* 2015)). In this system, tamoxifen-treated (but not mock-treated) mice almost immediately showed massive myeloid expansion with concomitant loss of T, B and erythroid cells, and rapidly developed an aggressive acute myeloid leukemia that caused them to succumb within 4–5 weeks of injection (An *et al.* 2015). Together these data indicate that profound TET loss-of-function predisposes cells to rapid, signal-dependent expansion that quickly progresses to frank malignancy.

Even in the absence of TET coding region mutations, TET loss-of-function and low 5hmC levels are frequently observed in many different types of cancers (Ko *et al.* 2010; Huang and Rao 2014; Ko *et al.* 2015b; Marcais *et al.* 2017; Lemonnier *et al.* 2018), including both blood malignancies and solid tumours. This may occur as a result of silencing or degradation of TET proteins at different stages of gene expression, including transcriptional silencing as a result of TET promoter hypermethylation, post-transcriptional processes including microRNA-mediated silencing, and increased degradation, as posttranslational modifications differentially impact TET proteins stability (Cimmino *et al.* 2015; Ko *et al.* 2015a; Raffel *et al.* 2017; Wu *et al.* 2018). Additionally, hypoxia and metabolic alterations could lead to TET loss-of-function by impairing its enzymatic activity (along with other dioxygenases), by decreasing the levels of the substrates 2-oxoglutarate and molecular oxygen or by increasing the levels of the competitive inhibitor 2-hydroxyglutarate (2-HG) (Kaelin and McKnight 2013; Losman and Kaelin 2013; Huang and Rao 2014; Ko *et al.* 2015b; Raffel *et al.* 2017). For example, gain-of-function mutations in the isocitrate dehydrogenases, IDH1 and IDH2, lead to accumulation of 2-HG, and mutations in these genes are frequently observed in patients with acute myeloid leukemia (AML) and glioblastoma (Dang *et al.* 2010; Cairns and Mak 2013; Losman and Kaelin 2013). Similarly, overexpression of the Branched chain amino acid transaminase 1 (BCAT1) gene, as reported in AML, leads to decreased levels of 2-oxoglutarate and therefore low TET function (Raffel *et al.* 2017).

7. DNA hypermethylation in TET2-mutant cancers

Most studies of TET2-mutant cancers have focused on the fact that loss-of-TET function results in increased methylation at genomic regions where TET proteins play a transcriptional regulatory role. This focal DNA hypermethylation, which occurs primarily at promoters and enhancers, can result in transcriptional silencing of tumour suppressor genes and genes involved in DNA damage repair, thus promoting oncogenesis (Jones and Baylin 2002; Baylin and Jones 2016). Indeed, focal DNA hypermethylation is commonly observed in tumours with impaired expression or activity of TET proteins. The presence of hypermethylation at these active genomic regions is consistent with the finding that 5hmC in wildtype cells is primarily present in euchromatin, at active enhancers and in the gene bodies of highly transcribed genes (Tsagaratou *et al.* 2014). DNA hypermethylation signature have been defined for many cancers, and some of these are characteristic of either *TET2* mutations or TET deficiency resulting from metabolic and other aberrations. For instance, both IDH-mutant and BCAT-overexpressing cancers have been shown to have a DNA hypermethylation signature that resembles that of TET-deficient cancers (Sasaki *et al.* 2012; Raffel *et al.* 2017). However, whether IDH-mutant and BCAT-overexpressing cancers show a second feature observed in TET-mutant cancers—DNA hypomethylation in heterochromatin—has not yet been resolved (see the following section).

8. TET deficiency is associated with a paradoxical loss of DNA methylation in heterochromatin

As detailed above, TET mutation or deficiency—which could result from TET coding region mutations, changes in mRNA or protein expression or stability, or metabolic alterations that result in inhibition of TET enzymatic activity—results in increased DNA methylation at genomic regions, including enhancers and promoters active in the cell type being examined, as well as certain CTCF sites (Cimmino *et al.* 2015; Rasmussen *et al.* 2015; Flavahan *et al.* 2016; Rasmussen and Helin 2016; Yue *et al.* 2016; Tsagaratou *et al.* 2017). This feature is expected from TET biochemical activity (figure 1b). Surprisingly, however, several studies that mapped DNA methylation, genome-wide, in TET-deficient cells noted unexpected and widespread decreases of DNA methylation (Hon *et al.* 2014; Lu *et al.* 2014; An *et al.* 2015; Tsagaratou *et al.* 2017). These hypomethylated regions did not overlap with active or regulatory regions of the genome, and so were largely ignored.

We recently reported a comprehensive analysis of DNA methylation in many different wildtype and TET-deficient cell types, including embryonic stem (ES) cells, neuronal precursor cells, haematopoietic stem cells, B cells and T cells (Lopez-Moyado *et al.* 2019). The TET-deficient cells bore individual deletions of the *Tet1* or *Tet2* genes, *Tet2/3* double deletions, or triple deletions of all three *TET* genes, *Tet1*, *Tet2* and *Tet3*. Our study revealed that in each of these distinct cell types, the widespread DNA hypomethylation observed in TET-deficient cells was confined to the heterochromatin compartment (Lopez-Moyado *et al.* 2019). Notably, the heterochromatin hypomethylation cannot be explained simply by increased proliferation, since TET triple-deficient ES cells do not proliferate faster than their wildtype counterparts (Li *et al.* 2016). Rather, in *Tet1*-deficient mESC, we observed a relocalization of DNMT3A from the heterochromatic to the euchromatic compartment, to the sites where TET1 would bind in wildtype conditions. These data provide a potential mechanism for the heterochromatic DNA hypomethylation associated with TET mutations, independent of proliferation rate (figure 4).

In addition to focal hypermethylation, cancer genomes have long been known to have widespread DNA hypomethylation (Feinberg and Vogelstein 1983; Jones and Baylin 2002; Ehrlich 2009; Baylin and Jones 2016). In these cases, as well as in TET-deficient genomes, DNA hypomethylation is primarily present in the heterochromatic compartment (Lopez-Moyado *et al.* 2019). Although we currently have a reasonable understanding of the biochemical mechanisms underlying focal hypermethylation and their consequences for gene transcription, the causes and consequences of DNA hypomethylation in cancer remain unclear.

9. Unexpected synergy between TET2 and DNMT3A mutations

Despite their opposing catalytic activities (TET removes DNA methylation whereas DNMT3A deposits this modification), *TET2* and *DNMT3A* mutations are frequently observed, individually and together, in diverse blood malignancies including myelodysplastic syndromes (MDS), acute myeloid leukemias (AML) and peripheral T cell lymphomas (PTCL) (Couronne *et al.* 2012; Ley *et al.* 2013; Odejide *et al.* 2014; Palomero *et al.* 2014; Sakata-Yanagimoto *et al.* 2014; Papaemmanuil *et al.* 2016). A previous study (Zhang *et al.* 2016) comparing the phenotypes of *Dnmt3a*, *Tet2*, and double *Dnmt3a/Tet2* loss-of-function mutations in the mouse hematopoietic precursors found that the *Dnmt3a/Tet2* double mutation resulted in decreased survival and increased number of hematopoietic precursor cells and white cells (monocytes) in the peripheral blood, compared to that of the mice singly deficient for *Dnmt3a* or *Tet2*. Additionally, we recently found that *Dnmt3a*/*Tet2* doubly-deficient cells displayed more profound losses of DNA methylation than *Dnmt3a* or *Tet2* mutations alone, even though both mutations resulted in heterochromatic DNA hypomethylation to different extents (Lopez-Moyado *et al.* 2019). Potentially, the similar phenotypes of DNMT3A and TET2 mutations could be a result of the loss of oxi-mC (TET deletions will decrease the amount of 5hmC, 5fC and 5caC, whereas DNMT mutations will decrease the amount of 5mC, which is the substrate for the TET-mediated cytosine oxidations). Thus, our study (Lopez-Moyado *et al.* 2019) opens up the possibility that some of the similarities between DNMT3A and TET2 mutations are a result of a shared loss of DNA methylation in heterochromatin.

DNA hypomethylation has been associated with increased mutational load and genome instability (Chen *et al.* 1998; Eden *et al.* 2003; Gaudet *et al.* 2003). It is well known that cancer genomes display DNA hypomethylation which covers long regions of the genome and overlaps with lamina-associated domains, H3K9me2/3-marked, late-replicating regions of the genome (Berman *et al.* 2011; Hon *et al.* 2012; Zhou *et al.* 2018). Furthermore, it has been previously

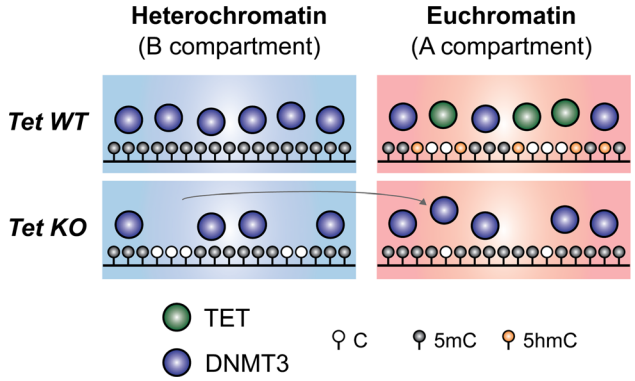


Figure 4. Proposed model for loss of DNA methylation in heterochromatin of TET-deficient embryonic stem cells. Loss of TET proteins results in relocalization of the *de novo* methyltransferase DNMT3 proteins, from the heterochromatic compartment to euchromatin regions previously occupied by TET proteins. Potentially, this relocalization contributes both to the heterochromatic DNA hypomethylation and the euchromatin DNA hypermethylation observed in TET-deficient cells. Whether this relocalization also occurs in other systems with TET loss-of-function is still an open question. Adapted from López-Moyado *et al.* (2019).

reported that cancer genomes display increased mutation rates in H3K9me3-marked regions of the genome (Schuster-Bockler and Lehner 2012). An interesting open question is if the heterochromatic DNA hypomethylation observed in TET deficient genomes could account for their increased levels of DNA damage, genome instability, and ultimately their role in oncogenesis. For instance, in the case of a NKT cell lymphoma that arises as a result of double *Tet2/Tet3* deletion (Tsagaratou *et al.* 2017), there was an association between progressive loss of methylation, increased levels of DNA damage, pronounced enrichment for single-nucleotide variations (SNVs) in the heterochromatin, and genome instability (Lopez-Moyado *et al.* 2019). However, the relationship between oncogenic transformation and DNA hypomethylation in heterochromatin and cancer is only just beginning to be elucidated.

10. Hypomethylation of human heterochromatin is associated with increased replication fork stalling, DNA damage and chromosomal abnormalities

An interesting recent finding is that hypomethylation of heterochromatin causes DNA damage and chromosomal abnormalities through the induction of replication stress (Delpu *et al.* 2019). Cells deficient in DNA methylation struggle to complete S phase, suggesting an essential and unexplored role for DNA methylation in regulating DNA replication (Jacob *et al.* 2015; Haruta *et al.* 2016). One of the most striking examples of genomic instability triggered by hypomethylation involves the repetitive sequence, Satellite 2 (SAT2), which occurs in megabase-long tracts in the pericentromeric heterochromatin of human chromosomes 1 and 16 (Ehrlich 2009; Altemose *et al.* 2014). Gains and losses of the long arms of these two chromosomes are overrepresented across many types of cancers, as well as in aging cells, and correlate strongly with SAT2 hypomethylation (Qu *et al.* 1999; Suzuki *et al.* 2002; Tsuda *et al.* 2002; Neve *et al.* 2006). Striking chromosomal rearrangements involving SAT2 have also been reported in lymphocytes and fibroblasts from patients with the fatal genetic disease Immunodeficiency, Centromeric instability, and Facial anomalies (ICF) syndrome, which is caused by germline mutations (including mutations in the *de novo* methyltransferase *DNMT3B*), that result in a dramatic loss of methylation at SAT2 (Ehrlich *et al.* 2001; Thijssen *et al.* 2015). Structural studies have demonstrated that sequences contained in SAT2 have the potential to fold into highly stable non-B DNA structures (Catasti *et al.* 1994). Such non-canonical DNA structures are known to stall replication forks leading to the formation of breaks and chromosomal rearrangements (Leon-Ortiz *et al.* 2014), suggesting that hypomethylation of SAT2 may lead to chromosomal abnormalities through the dysregulation of genomic secondary structures and the induction of replication stress.

The Tahiliani Lab developed a single-molecule approach that combined DNA combing with fluorescence in situ hybridization (FISH) to directly visualize the impact of hypomethylation of SAT2 on replication. Replication dynamics in well-characterized ICF patient cell lines were compared to those in normally methylated cells. This approach revealed that SAT2 hypomethylation results in increased DNA damage specifically at SAT2 and strongly impairs the efficiency of replicating these sequences (Delpu *et al.* 2019). Consistent with increased fork stalling at these sequences, they found increased levels of the single-stranded DNA (ssDNA) binding protein, RPA2, as well as asymmetric progression of sister replication forks within hypomethylated SAT2 sequences (Delpu *et al.* 2019). Together these findings indicate that impaired replication triggers the formation of chromosomal aberrations observed at hypomethylated SAT2 sequences and also suggests a mechanistic basis for

how the loss of DNA methylation may contribute to genomic instability in diverse pathological conditions.

11. Conclusion and perspectives

The studies of TET protein function over the last decade have focused on its ability to facilitate DNA demethylation through the production of oxi-mC. It has only recently been recognized that loss of TET function can also compromise heterochromatin integrity, and that this process could be deleterious for genome stability and start cells on the road to oncogenic transformation. We anticipate that studies over the coming decade will elucidate the mechanisms involved.

Acknowledgments

The authors thank all current and previous Rao lab members for their work on published papers cited in this review, and for their inputs and many stimulating discussions over the years. We also thank the Flow Cytometry and Functional Genomics cores at the La Jolla Institute for Immunology for cell sorting and next-generation sequencing. This work was supported by National Institutes of Health (NIH) Grants R01 AI128589 and R35 CA210043 (to A.R). C.W.L. was supported by the Independent Investigator Fund (La Jolla Institute/Kyowa Kirin) and the Irvington Postdoctoral Fellowship from the Cancer Research Institute. I.F.L.-M. was supported by a University of California Institute for Mexico and the United States–Consejo Nacional de Ciencia y Tecnología Fellowship. L.A. is supported by the intramural funds of the National Library of Medicine, NIH, USA. M.T. was supported by the Packard Fellowship for Science and Engineering. Funding for Illumina HiSeq 2500 and BD FACSAria II is supported by NIH (NIH S10OD016262 and NIH S10RR027366).

Conflict of interest AR is on the scientific advisory board of Cambridge Epigenetix (Cambridge, UK). The other authors declare no competing interests.

References

Altemose N, Miga KH, Maggioni M and Willard HF 2014 Genomic characterization of large heterochromatic gaps in the human genome assembly. *PLoS Comput. Biol.* **10** e1003628

An J, González-Avalos E, Chawla A, Jeong M, López-Moyado IF, Li W, Goodell MA, Chavez L, Ko M and Rao A 2015 Acute loss of TET function results in aggressive myeloid cancer in mice. *Nat. Commun.* **6** 10071

Aravind, L, Abhiman, S, and Iyer LM 2011 Natural history of the eukaryotic chromatin protein methylation system. *Prog Mol Biol Transl Sci* **101** 105–176

Aravind L, Balasubramanian S and Rao A 2019 Unusual Activity of a Chlamydomonas TET/JBP Family Enzyme. *Biochemistry* **(in press)**

Aravind L and Koonin EV 2001 The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. *Genome Biol.* **2** RESEARCH0007

Baylin SB and Jones PA 2016 Epigenetic Determinants of Cancer. *Cold Spring Harb. Perspect. Biol.* **8**

Benson MJ, Pino-Lagos K, Rosemblatt M and Noelle RJ 2007 All-trans retinoic acid mediates enhanced T reg cell growth, differentiation, and gut homing in the face of high levels of co-stimulation. *The Journal of experimental medicine* **204** 1765–1774

Berman BP, Weisenberger DJ, Aman JF, Hinoue T, Ramjan Z, Liu Y, Noushmehr H, Lange CP, van Dijk CM, Tollenaar RA, *et al.* 2011 Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nat. Genet.* **44** 40–46

Blaschke K, Ebata KT, Karimi MM, Zepeda-Martinez JA, Goyal P, Mahapatra S, Tam A, Laird DJ, Hirst M, Rao A, *et al.* 2013 Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500** 222–226

Bullard W, Lopes da Rosa-Spiegler J, Liu S, Wang Y and Sabatini R 2014 Identification of the glucosyltransferase that converts hydroxymethyluracil to base J in the trypanosomatid genome. *J. Biol. Chem.* **289** 20273–20282

Cairns RA and Mak TW 2013 Oncogenic isocitrate dehydrogenase mutations: mechanisms, models and clinical opportunities. *Cancer Discov.* **3** 730–741

Catasti P, Gupta G, Garcia AE, Ratliff R, Hong L, Yau P, Moyzis RK and Bradbury EM 1994 Unusual structures of the tandem repetitive DNA sequences located at human centromeres. *Biochemistry* **33** 3819–3830

Chavez L, Huang Y, Luong K, Agarwal S, Iyer LM, Pastor WA, Hench VK, Frazier-Bowers SA, Korol E, Liu S, *et al.* 2014 Simultaneous sequencing of oxidized methylcytosines produced by TET/JBP dioxygenases in *Coprinopsis cinerea*. *Proc. Natl. Acad. Sci. USA* **111** E5149–5158

Chen RZ, Pettersson U, Beard C, Jackson-Grusby L and Jaenisch R 1998 DNA hypomethylation leads to elevated mutation rates. *Nature* **395** 89–93

Chen W, Jin W, Hardegen N, Lei KJ, Li L, Marinos N, McGrady G and Wahl SM 2003 Conversion of peripheral CD4+CD25− naive T cells to CD4+CD25+ regulatory T cells by TGF-beta induction of transcription factor

Foxp3. *The Journal of experimental medicine* **198** 1875–1886

Cimmino L, Dawlaty MM, Ndiaye-Lobry D, Yap YS, Bakogianni S, Yu Y, Bhattacharyya S, Shaknovich R, Geng H, Lobry C*, et al.* 2015 TET1 is a tumor suppressor of hematopoietic malignancy. *Nat. Immunol.* **16** 653–662

Couronne L, Bastard C and Bernard OA 2012 *TET2* and *DNMT3A* mutations in human T-cell lymphoma. *The New England journal of medicine* **366** 95–96

Crawford DJ, Liu MY, Nabel CS, Cao X-J, Garcia BA and Kohli RM 2016 Tet2 Catalyzes Stepwise 5-Methylcytosine oxidation by an iterative and de novo mechanism. *J. Am. Chem. Soc.* **138** 730–733

Dang L, Jin S and Su SM 2010 IDH mutations in glioma and acute myeloid leukemia. *Trends Mol. Med.* **16** 387–397

Dekker J, Marti-Renom MA and Mirny LA 2013 Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat. Rev. Genet.* **14** 390–403

Delpu Y, McNamara T, Griffin P, Kaleem S, Narayan S, Schildkraut C, Miga K and Tahiliani M 2019 Chromosomal rearrangements at hypomethylated Satellite 2 sequences are associated with impaired replication efficiency and increased fork stalling. *BioRxiv*

Eden A, Gaudet F, Waghmare A and Jaenisch R 2003 Chromosomal instability and tumors promoted by DNA hypomethylation. *Science* **300** 455

Ehrlich M 2009 DNA hypomethylation in cancer cells. *Epigenomics* **1** 239–259

Ehrlich M, Buchanan KL, Tsien F, Jiang G, Sun B, Uicker W, Weemaes CM, Smeets D, Sperling K, Belohradsky BH*, et al.* 2001 DNA methyltransferase 3B mutations linked to the ICF syndrome cause dysregulation of lymphogenesis genes. *Hum. Mol. Genet.* **10** 2917–2931

Feinberg AP and Vogelstein B 1983 Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301** 89–92

Feng Y, Arvey A, Chinen T, van der Veeken J, Gasteiger G and Rudensky AY 2014 Control of the inheritance of regulatory T cell identity by a cis element in the *Foxp3* locus. *Cell* **158** 749–763

Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, Suva ML and Bernstein BE 2016 Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **529** 110–114

Floess S, Freyer J, Siewert C, Baron U, Olek S, Polansky J, Schlawe K, Chang HD, Bopp T, Schmitt E*, et al.* 2007 Epigenetic control of the foxp3 locus in regulatory T cells. *PLoS Biol* **5** e38

Gaudet F, Hodgson JG, Eden A, Jackson-Grusby L, Dausman J, Gray JW, Leonhardt H and Jaenisch R 2003 Induction of tumors in mice by genomic hypomethylation. *Science* **300** 489–492

Guo F, Li X, Liang D, Li T, Zhu P, Guo H, Wu X, Wen L, Gu TP, Hu B*, et al.* 2014 Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell* **15** 447–459

Hamel KM, Mandal M, Karki S and Clark MR 2014 Balancing proliferation with Igkappa recombination during B-lymphopoiesis. *Front. Immunol.* **5** 139

Haruta M, Shimada M, Nishiyama A, Johmura Y, Le Tallec B, Debatisse M and Nakanishi M 2016 Loss of maintenance DNA methylation results in abnormal DNA origin firing during DNA replication. *Biochem. Biophys. Res. Commun.* **469** 960–966

Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, Zhang X and Cheng X 2012 Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res.* **40** 4841–4849

Hausinger RP 2004 FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol.* **39** 21–68

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

Hon GC, Hawkins RD, Caballero OL, Lo C, Lister R, Pelizzola M, Valsesia A, Ye Z, Kuan S, Edsall LE*, et al.* 2012 Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Res.* **22** 246–258

Hon GC, Song CX, Du T, Jin F, Selvaraj S, Lee AY, Yen CA, Ye Z, Mao SQ, Wang BA*, et al.* 2014 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol. Cell* **56** 286–297

Huang, Y, Chavez, L, Chang X, Wang X, Pastor WA, Kang J, Zepeda-Martinez JA, Pape UJ, Jacobsen SE, Peters B and Rao A 2014 Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *P. Natl. Acad. Sci. USA* **111** 1361–1366

Huang Y and Rao A 2014 Connections between TET proteins and aberrant DNA modification in cancer. *Trends Genet.* **30** 464–474

Inoue A and Zhang Y 2011 Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* **334** 194

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

Iyer LM, Abhiman S and Aravind L 2011 Natural history of eukaryotic DNA methylation systems. *Prog. Mol. Biol. Transl. Sci.* **101** 25–104

Iyer LM, Tahiliani M, Rao A and Aravind L 2009 Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle* **8** 1698–1710

Iyer LM, Zhang D and Aravind L 2016 Adenine methylation in eukaryotes: Apprehending the complex evolutionary

history and functional potential of an epigenetic modification. *Bioessays* **38** 27–40

Iyer LM, Zhang D, Burroughs AM and Aravind L 2013 Computational identification of novel biochemical systems involved in oxidation, glycosylation and other complex modifications of bases in DNA. *Nucleic Acids Res.* **41** 7635–7655

Iyer LM, Zhang D, de Souza RF, Pukkila PJ, Rao A and Aravind L 2014 Lineage-specific expansions of TET/JBP genes and a new class of DNA transposons shape fungal genomic and epigenetic landscapes. *P. Natl. Acad. Sci. USA* **111** 1676–1683

Jacob V, Chernyavskaya Y, Chen X, Tan PS, Kent B, Hoshida Y and Sadler KC 2015 DNA hypomethylation induces a DNA replication-associated cell cycle arrest to block hepatic outgrowth in uhrf1 mutant zebrafish embryos. *Development* **142** 510–521

Jones PA and Baylin SB 2002 The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* **3** 415–428

Josefowicz SZ, Lu LF and Rudensky AY 2012 Regulatory T cells: mechanisms of differentiation and function. *Annu. Rev. Immunol.* **30** 531–564

Kaelin WG Jr. and McKnight SL 2013 Influence of metabolism on epigenetics and disease. *Cell* **153** 56–69

Ko M, An J, Bandukwala HS, Chavez L, Aijo T, Pastor WA, Segal MF, Li H, Koh KP, Lahdesmaki H*, et al.* 2013 Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* **497** 122–126

Ko M, An J, Pastor WA, Koralov SB, Rajewsky K and Rao A 2015a TET proteins and 5-methylcytosine oxidation in hematological cancers. *Immunol. Rev.* **263** 6–21

Ko M, An J and Rao A 2015b DNA methylation and hydroxymethylation in hematologic differentiation and transformation. *Curr. Opin. Cell Biol.* **37** 91–101

Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, An J, Lamperti ED, Koh KP, Ganetzky R*, et al.* 2010 Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant *TET2*. *Nature* **468** 839–843

Lee YJ, Dai N, Walsh SE, Muller S, Fraser ME, Kauffman KM, Guan C, Correa IR Jr. and Weigele PR 2018 Identification and biosynthesis of thymidine hypermodifications in the genomic DNA of widespread bacterial viruses. *P. Natl. Acad. Sci. USA* **115** E3116–E3125

Lemonnier F, Poullot E, Dupuy A, Couronne L, Martin N, Scourzic L, Fataccioli V, Bruneau J, Cairns RA, Mak TW*, et al.* 2018 Loss of 5-hydroxymethylcytosine is a frequent event in peripheral T-cell lymphomas. *Haematologica* **103** e115–e118

Leon-Ortiz AM, Svendsen J and Boulton SJ 2014 Metabolism of DNA secondary structures at the eukaryotic replication fork. *DNA Repair* **19** 152–162

Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, Robertson A, Hoadley K, Triche TJ Jr., Laird PW, Baty JD*, et al.* 2013 Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *The New England Journal of Medicine* **368** 2059–2074

Li X, Liang Y, LeBlanc M, Benner C and Zheng Y 2014 Function of a Foxp3 cis-element in protecting regulatory T cell identity. *Cell* **158** 734–748

Li X, Yue X, Pastor WA, Lin L, Georges R, Chavez L, Evans SM and Rao A 2016 Tet proteins influence the balance between neuroectodermal and mesodermal fate choice by inhibiting Wnt signaling. *P. Natl. Acad. Sci. USA* **113** E8267–E8276

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO*, et al.* 2009 Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326** 289–293

Lio CJ, Shukla V, Samaniego-Castruita D, Gonzalez-Avalos E, Chakraborty A, Yue X, Schatz DG, Ay F and Rao A 2019 TET enzymes augment activation-induced deaminase (AID) expression via 5-hydroxymethylcytosine modifications at the Aicda superenhancer. *Sci. Immunol.* 4

Lio CW and Hsieh CS 2011 Becoming self-aware: the thymic education of regulatory T cells. *Curr. Opin. Immunol.* **23** 213–219

Lio CW, Zhang J, Gonzalez-Avalos E, Hogan PG, Chang X and Rao A 2016 Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *eLife*

López-Moyado IF, Tsagaratou A, Yuita H, Seo H, Delatte B, Heinz S, Benner C and Rao A 2019 Paradoxical association of TET loss of function with genome-wide DNA hypomethylation. *P. Natl. Acad. Sci. USA* **34** 16933–16942

Losman JA and Kaelin WG Jr. 2013 What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate and cancer. *Gene. Dev.* **27** 836–852

Lu F, Liu Y, Jiang L, Yamaguchi S and Zhang Y 2014 Role of Tet proteins in enhancer activity and telomere elongation. *Gene. Dev.* **28** 2103–2119

Maiti A and 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** 35334–35338

Marcais A, Waast L, Bruneau J, Hanssens K, Asnafi V, Gaulard P, Suarez F, Dubreuil P, Gessain A, Hermine O and Pique C 2017 Adult T cell leukemia aggressivenness correlates with loss of both 5-hydroxymethylcytosine and TET2 expression. *Oncotarget* **8** 52256–52268

Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F*, et al.* 2006 A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer cell* **10** 515–527

Odejide O, Weigert O, Lane AA, Toscano D, Lunning MA, Kopp N, Kim S, van Bodegom D, Bolla S, Schatz JH*, et al.* 2014 A targeted mutational landscape of angioimmunoblastic T-cell lymphoma. *Blood* **123** 1293–1296

Otani J, Kimura H, Sharif J, Endo TA, Mishima Y, Kawakami T, Koseki H, Shirakawa M, Suetake I and

Tajima S 2013 Cell cycle-dependent turnover of 5-hydroxymethyl cytosine in mouse embryonic stem cells. *PloS one* **8** e82961

Palomero T, Couronne L, Khiabanian H, Kim MY, Ambesi-Impiombato A, Perez-Garcia A, Carpenter Z, Abate F, Allegretta M, Haydu JE, *et al.* 2014 Recurrent mutations in epigenetic regulators RHOA and FYN kinase in peripheral T cell lymphomas. *Nat. Genet.* **46** 166–170

Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, Potter NE, Heuser M, Thol F, Bolli N, *et al.* 2016 Genomic classification and prognosis in acute myeloid leukemia. *The New England Journal of Medicine* **374** 2209–2221

Pastor WA, Aravind L and Rao A 2013 TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat. Rev. Mol. Cell. Biol.* **14** 341–356

Qu GZ, Grundy PE, Narayan A and Ehrlich M 1999 Frequent hypomethylation in Wilms tumors of pericentromeric DNA in chromosomes 1 and 16. *Cancer Genet. Cytogen.* **109** 34–39

Raffel S, Falcone M, Kneisel N, Hansson J, Wang W, Lutz C, Bullinger L, Poschet G, Nonnenmacher Y, Barnert A, *et al.* 2017 BCAT1 restricts alphaKG levels in AML stem cells leading to IDHmut-like DNA hypermethylation. *Nature* **551** 384–388

Rasmussen KD and Helin K 2016 Role of TET enzymes in DNA methylation development and cancer. *Gene. Dev.* **30** 733–750

Rasmussen KD, Jia G, Johansen JV, Pedersen MT, Rapin N, Bagger FO, Porse BT, Bernard OA, Christensen J and Helin K 2015 Loss of *TET2* in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. *Gene. Dev.* **29** 910–922

Sakaguchi S, Yamaguchi T, Nomura T and Ono M 2008 Regulatory T cells and immune tolerance. *Cell* **133** 775–787

Sakata-Yanagimoto M, Enami T, Yoshida K, Shiraishi Y, Ishii R, Miyake Y, Muto H, Tsuyama N, Sato-Otsubo A, Okuno Y, *et al.* 2014 Somatic RHOA mutation in angioimmunoblastic T cell lymphoma. *Nat. Genet.* **46** 171–175

Sasaki M, Knobbe CB, Munger JC, Lind EF, Brenner D, Brustle A, Harris IS, Holmes R, Wakeham A, Haight J, *et al.* 2012 IDH1(R132H) mutation increases murine haematopoietic progenitors and alters epigenetics. *Nature* **488** 656–659

Sasidharan Nair V, Song MH and Oh KI 2016 Vitamin C facilitates demethylation of the Foxp3 enhancer in a Tet-dependent manner. *J. Immunol.* **196** 2119–2131

Schiesser S, Hackner B, Pfaffeneder T, Muller M, Hagemeier C, Truss M and Carell T 2012 Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. *Angew. Chem. Int. Ed. Engl.* **51** 6516–6520

Schuster-Bockler B and Lehner B 2012 Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature* **488** 504–507

Suzuki T, Fujii M and Ayusawa D 2002 Demethylation of classical satellite 2 and 3 DNA with chromosomal instability in senescent human fibroblasts. *Exp. Gerontol.* **37** 1005–1014

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L and Rao A 2009 Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324** 930–935

Thijssen PE, Ito Y, Grillo G, Wang J, Velasco G, Nitta H, Unoki M, Yoshihara M, Suyama M, Sun Y, *et al.* 2015 Mutations in CDCA7 and HELLS cause immunodeficiency-centromeric instability-facial anomalies syndrome. *Nat. Commun.* **6** 7870

Tsagaratou A, Aijo T, Lio CW, Yue X, Huang Y, Jacobsen SE, Lahdesmaki H and Rao A 2014 Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation. *P. Natl. Acad. Sci. USA* **111** E3306–3315

Tsagaratou A, Gonzalez-Avalos E, Rautio S, Scott-Browne JP, Togher S, Pastor WA, Rothenberg EV, Chavez L, Lahdesmaki H and Rao A 2017 TET proteins regulate the lineage specification and TCR-mediated expansion of iNKT cells. *Nat. Immunol.* **18** 45–53

Tsuda H, Takarabe T, Kanai Y, Fukutomi T and Hirohashi S 2002 Correlation of DNA hypomethylation at pericentromeric heterochromatin regions of chromosomes 16 and 1 with histological features and chromosomal abnormalities of human breast carcinomas. *The American Journal of Pathology* **161** 859–866

van Steensel B and Belmont AS 2017 Lamina-associated domains: Links with chromosome architecture heterochromatin and gene Repression. *Cell* **169** 780–791

Wu D, Hu D, Chen H, Shi G, Fetahu IS, Wu F, Rabidou K, Fang R, Tan L Xu S, *et al.* 2018 Glucose-regulated phosphorylation of TET2 by AMPK reveals a pathway linking diabetes to cancer. *Nature* **559** 637–641

Xiao M, Yang H, Xu W, Ma S, Lin H, Zhu H, Liu L, Liu Y, Yang C, Xu Y, *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. *Gene. Dev.* **26** 1326–1338

Xue J-H, Chen G-D, Hao F, Chen H, Fang Z, Chen F-F, Pang B, Yang Q-L, Wei X, Fan Q-Q, *et al.* 2019 A vitamin-C-derived DNA modification catalysed by an algal TET homologue. *Nature* **569** 581–585

Yang R, Qu C, Zhou Y, Konkel JE, Shi S, Liu Y, Chen C, Liu S, Liu D, Chen Y, *et al.* 2015 Hydrogen sulfide promotes Tet1- and Tet2-mediated Foxp3 demethylation to drive regulatory T cell differentiation and maintain immune homeostasis. *Immunity* **43** 251–263

Yu Z, Genest PA, ter Riet B, Sweeney K, DiPaolo C, Kieft R, Christodoulou E, Perrakis A, Simmons JM, Hausinger RP, *et al.* 2007 The protein that binds to DNA base J in

trypanosomatids has features of a thymidine hydroxylase. *Nucleic Acids Res.* **35** 2107–2115

Yue X, Trifari S, Aijo T, Tsagaratou A, Pastor WA, Zepeda-Martinez JA, Lio CW, Li X, Huang Y, Vijayanand P, *et al.* 2016 Control of Foxp3 stability through modulation of TET activity. *The Journal of Experimental Medicine* **213** 377–397

Zhang L, Chen W, Iyer LM, Hu J, Wang G, Fu Y, Yu M, Dai Q, Aravind L and He C 2014 A TET homologue protein from *Coprinopsis cinerea* (CcTET) that biochemically converts 5-methylcytosine to 5-hydroxymethylcytosine 5-formylcytosine and 5-carboxylcytosine. *J. Am. Chem. Soc.* **136** 4801–4804

Zhang X, Su J, Jeong M, Ko M, Huang Y, Park HJ, Guzman A, Lei Y, Huang YH, Rao A, *et al.* 2016 DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors in hematopoietic stem cells. *Nat. Genet.* **48** 1014–1023

Zheng Y, Josefowicz S, Chaudhry A, Peng XP, Forbush K and Rudensky AY 2010 Role of conserved non-coding DNA elements in the *Foxp3* gene in regulatory T-cell fate. *Nature* **463** 808–812

Zhou W, Dinh HQ, Ramjan Z, Weisenberger DJ, Nicolet CM, Shen H, Laird PW and Berman BP 2018 DNA methylation loss in late-replicating domains is linked to mitotic cell division. *Nat. Genet.* **50** 591–602