

# DNA Demethylation Dynamics

Nidhi Bhutani,<sup>1,2</sup> David M. Burns,<sup>1</sup> and Helen M. Blau<sup>1,\*</sup>

<sup>1</sup>Baxter Laboratory for Stem Cell Biology, Institute for Stem Cell Biology and Regenerative Medicine, Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305-5175, USA

<sup>2</sup>Present address: Department of Orthopaedic Surgery, Stanford University, Stanford, CA 94305-5341, USA

\*Correspondence: [hblau@stanford.edu](mailto:hblau@stanford.edu)

DOI 10.1016/j.cell.2011.08.042

The discovery of cytosine hydroxymethylation (5hmC) suggested a simple means of demethylating DNA and activating genes. Further experiments, however, unearthed an unexpectedly complex process, entailing both passive and active mechanisms of DNA demethylation by the ten-eleven translocation (TET) and AID/APOBEC families of enzymes. The consensus emerging from these studies is that removal of cytosine methylation in mammalian cells can occur by DNA repair. These reports highlight that in certain contexts, DNA methylation is not fixed but dynamic, requiring continuous regulation.

## Introduction

The significant impact of DNA methylation patterns on cell and organismal fate is perhaps most graphically exemplified in honeybees, in which differential DNA methylation determines whether the bee will be a worker or a queen (Kucharski et al., 2008). In mammals, DNA methylation has also long been considered integral to fundamental choices, including the long-term gene silencing that leads to genomic imprinting, X chromosome inactivation, suppression of transposable elements, and the establishment and maintenance of stable cellular identities (Bird, 2002; De Carvalho et al., 2010; Deaton and Bird, 2011; Goll and Bestor, 2005; Jaenisch and Bird, 2003). Yet, studies of cellular reprogramming by three approaches—nuclear transfer, cell fusion, and induced pluripotency by defined factors (i.e., iPSCs)—all demonstrate that “fixed and stable” differentiated cellular states can be radically altered (Jullien et al., 2011; Yamanaka and Blau, 2010). Concurrently, accumulating evidence has suggested that DNA methylation may be reversible in mammalian cells; however, knowledge of the requisite molecules and mechanisms underlying this process has been lacking. In this Perspective, we focus on recent reports that now identify enzymes capable of mediating DNA demethylation in mammalian cells. These findings raise the possibility that regulation by DNA methylation is at times quite dynamic, providing exciting insights into why reprogramming of cell fates is possible.

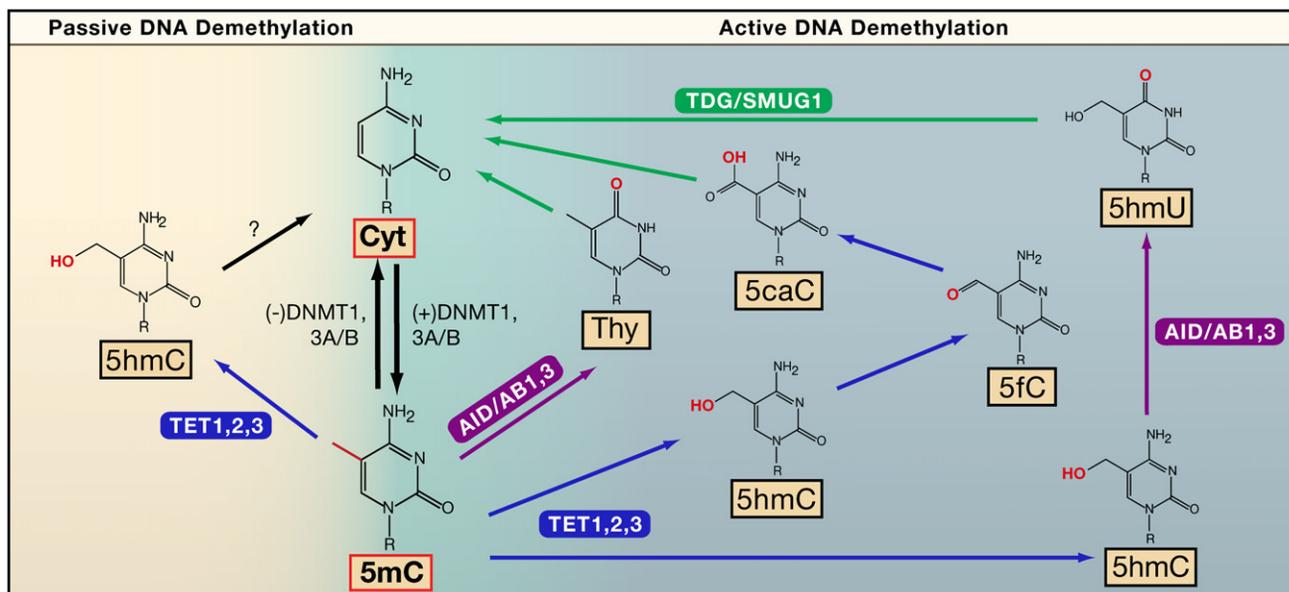
Recent discoveries have generated substantial excitement, as they show that cytosines in mammalian cells can be hydroxymethylated to 5hmC (5-hydroxymethylcytosine) (Figure 1) (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009). 5hmC is especially abundant in tissues such as brain and in pluripotent embryonic stem cells (ESCs), but it is also present at lower levels in blood, lung, kidney, and muscle (Globisch et al., 2010; Kriaucionis and Heintz, 2009; Ruzov et al., 2011; Song et al., 2011; Tahiliani et al., 2009). Initially, hydroxylation seemed like a probable means of activating genes silenced by methylation (Ito et al., 2010; Tahiliani et al., 2009), but recent studies rule out this simple hypothesis (Ficz et al., 2011; Pastor et al., 2011; Williams et al.,

2011; Wu et al., 2011b; Xu et al., 2011). Moreover, although several groups have investigated the genomic distribution of DNA hydroxymethylation, the role and functional significance of this modification are still unclear (Ficz et al., 2011; Pastor et al., 2011; Williams et al., 2011; Wu et al., 2011b; Xu et al., 2011).

We postulate here, based on analyses of recent evidence from our and other laboratories (Bhutani et al., 2010; Cortellino et al., 2011; Guo et al., 2011; He et al., 2011), that DNA methylation and demethylation can be a two-way street, characterized by multiple pathways (Figure 1). Importantly, these findings suggest that 5hmC may serve as an intermediate for the removal of methylated cytosines either by (1) passive dilution via the presence of 5hmC, which impairs remethylation by DNA methyltransferases (DNMTs) when cells divide, or (2) active replacement of modified cytosines via DNA repair in the absence of cell division. The role of DNA repair in DNA demethylation is well established in plants (Gehring et al., 2009). However, this pathway was not thought to operate in mammals, as no mammalian orthologs of the plant enzymes with similar activities were readily apparent (Gehring et al., 2009). Moreover, the mammalian enzymes that have recently been identified as the lead actors in the demethylation plot are well known for their involvement in other processes, for example in leukemia (TETs) and in antibody diversification (AID). Thus, their roles in the saga of DNA demethylation are entirely new. We postulate that the discovery of these regulators and their newly identified roles will provide insights into the *raison d'être* for DNA methylation, its modifications, and its role in gene expression, cell-fate determination, and nuclear reprogramming.

## Active Mechanisms for Loss of DNA Methylation: The Methylation “Editors”

DNMTs are responsible for the establishment and maintenance of DNA methylation as well as passive DNA demethylation in mammalian cells. It has long been thought that an absence or reduction in the DNMT levels gradually and passively removes



**Figure 1. DNA Demethylation Pathways**

Passive DNA demethylation has long been known to occur by a reduction in activity or absence of DNA methyltransferases (DNMTs) (black). DNMT3A and 3B are responsible for de novo DNA methylation, whereas DNMT1 maintains DNA methylation patterns through successive rounds of cell division. Recently, three enzyme families have been implicated in active DNA demethylation via DNA repair. (1) 5-methylcytosine (5mC) can be hydroxylated by the ten-eleven translocation (TET) family of enzymes (blue) to form 5-hydroxymethylcytosine (5hmC) or further oxidized to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). (2) 5mC (or 5hmC) can be deaminated by the AID/APOBEC family members (purple) to form 5-methyluracil (5mU) or 5-hydroxymethyluracil (5hmU). (3) Replacement of these intermediates (i.e., 5mU, 5hmU, or 5caC) is initiated by the UDG family of base excision repair (BER) glycosylases (green) like TDG or SMUG1, culminating in cytosine replacement and DNA demethylation.

DNA methylation in early mammalian development (Monk et al., 1991; Rougier et al., 1998). Specifically, it is well known that de novo methylation in early development is established by DNA methyltransferases 3A (DNMT3A) and 3B (DNMT3B) (reviewed in Law and Jacobsen, 2010). Once established, methylation patterns are faithfully maintained through cell divisions by DNMT1 (Law and Jacobsen, 2010). Thus, to date, inhibition of DNMTs constitutes the primary means of passive DNA demethylation.

A few early studies implicated demethylation of DNA by a rapid and active mechanism, independent of cell division (Mayer et al., 2000; Oswald et al., 2000; Paroush et al., 1990; Zhang et al., 2007), but how this active removal is achieved remained a mystery. Indeed, for decades an enzyme that could cleave the methyl group was sought but not found, suggesting that such a chemical reaction might simply not be possible (Bird, 2002). In the past year or so, a flurry of studies (Bhutani et al., 2010; Cortellino et al., 2011; Ficiz et al., 2011; Guo et al., 2011; He et al., 2011; Ito et al., 2011; Koh et al., 2011; Pastor et al., 2011; Popp et al., 2010; Rai et al., 2008; Song et al., 2011; Williams et al., 2011; Wu et al., 2011a, 2011b; Xu et al., 2011) have identified key players in this process, which may have been overlooked previously because they act indirectly to mediate active DNA demethylation. These enzymes first modify the methylated cytosine (by hydroxylation, deamination, oxidation, or a combination of these modifications), leading to its replacement by DNA repair.

The above studies connect three enzymatic families to active DNA demethylation: the ten-eleven translocation (TET) family,

which modifies methylated cytosines first by hydroxylation and then by further oxidation; the AID/APOBEC family, which deaminates the base (5mC or 5hmC); and finally, a family of base excision repair (BER) glycosylases, which mediate DNA repair (Figure 1). Here, we synthesize recent data that link these enzyme families. We suggest a role for them in active DNA demethylation in mammals in response to cell signaling and in early development, differentiation, and nuclear reprogramming to new cell states. Furthermore, we provide a speculative scheme for the circuitry by which DNA can be demethylated and remethylated, and how these states may be rapidly reversed by these methylation “editors.”

#### **The TET Family: Mediators of 5mC to 5hmC Conversion**

The existence of 5hmC was reported in the 1950s (Wyatt and Cohen, 1952), but its significance was unknown, and it remained largely ignored for the next half century. The discovery of the TET proteins, TET1, 2, and 3, that catalyze the conversion of 5-methylcytosine (5mC) to 5hmC heralded a revival of interest in this modified base. TET1 was initially identified as a fusion partner of the MLL protein in acute myeloid leukemia and named leukemia-associated protein, LCX, although its functional role in this type of cancer remained unknown (Ono et al., 2002). More recently, Rao and colleagues rekindled interest in the TETs when they identified them as potential modifiers of 5mC (Tahiliani et al., 2009). Based on knowledge that in *Trypanosoma*, the J base binding proteins 1 (JBP1) and 2 (JBP2) oxidize 5-methylthymine, Rao and colleagues cloned the mammalian homolog, TET1, cleverly reasoning that it might serve a similar

function in higher metazoa and mediate 5mC hydroxylation (Tahiliani et al., 2009). Recombinant human TET1 was found to be capable of converting 5mC to 5hmC in mammalian DNA, providing evidence for its putative role in mediating DNA demethylation (Tahiliani et al., 2009). Like their human counterparts, mouse TET1, 2, and 3 catalyze the conversion of 5mC to 5hmC (Ito et al., 2010).

Much of the initial excitement regarding the discovery of 5hmC was the prediction that it could readily lift the repression of gene expression imposed by 5mC at many gene promoters (Ito et al., 2010; Tahiliani et al., 2009). However, like methylation, a high concentration of 5hmC correlates with transcriptionally nonproductive or altogether inactive gene promoters (Ficz et al., 2011; Pastor et al., 2011; Williams et al., 2011; Wu et al., 2011a; Xu et al., 2011). Thus, contrary to expectations, the 5mC to 5hmC modification is clearly not the functional equivalent of 5mC to cytosine, which is associated with derepression of certain gene promoters (Pastor et al., 2011).

The hypothesis that cytosine hydroxylation might play a functional role in maintaining the pluripotent state was first suggested by Zhang and colleagues (Ito et al., 2010). They reported that TET1 results in a loss of ESC self-renewal by reducing the expression of the pluripotency regulator NANOG. This finding, however, has been challenged by others who suggest that, although levels of TET1 and TET2 (and therefore 5hmC) are high in ESCs, these proteins largely mediate regulation of lineage-specific genes, not the pluripotency regulator NANOG (Ficz et al., 2011; Koh et al., 2011).

In parallel with studies in ES cells, other experiments have been performed to determine whether the TET proteins are involved in active DNA demethylation in early development. In this case, the third family member, TET3, is most abundant and plays a role in the rapid and active loss of 5mC in the male pronucleus upon zygote formation prior to cell division. Recent experiments have shed light on the mechanism underlying this process. An increase in 5hmC is concomitant with a decrease in 5mC in zygotes (as determined by immunohistochemistry), suggesting that 5mC is converted to 5hmC (Iqbal et al., 2011; Wossidlo et al., 2011). In addition, knocking down TET3 by RNA interference (RNAi) led to an increase in 5mC. Thus, TET3 is responsible for 5hmC generation post-fertilization in mouse zygotes, suggesting a potential role for TET proteins in DNA demethylation early in development.

An unexpected complication of interpretations of experiments regarding the effects of TET proteins on gene expression is that TET1 plays a repressive role, independent of its enzymatic activity as a hydroxylase (Williams et al., 2011; Wu et al., 2011b). TET1 has been found to associate with two different transcriptional repressor complexes containing PRC2 (polycomb repressive complex 2) or SIN3A (Smi-independent 3A) (Williams et al., 2011). SIN3A and TET1 directly interact with one another (as shown by coimmunoprecipitation), whereas PRC2 and TET1 may act indirectly (Williams et al., 2011; Wu et al., 2011b). Importantly, a high degree of target overlap is observed between TET1 and PRC2, as well as between TET1 and SIN3A, in global chromatin immunoprecipitation (ChIP) analyses (Wu et al., 2011b; Williams et al., 2011). Furthermore, a subset of TET1 target genes is upregulated upon loss of SIN3A

function by siRNA knockdown, further substantiating that the SIN3A corepressor complex is required for TET1-mediated repression of these genes, independent of its catalytic role in generating 5hmC (Williams et al., 2011).

The “rediscovery” of 5hmC and TET proteins has led to a rapid succession of reports regarding the location and putative function of both 5hmC and TET in regulating DNA demethylation (Cortellino et al., 2011; Ficz et al., 2011; Guo et al., 2011; He et al., 2011; Ito et al., 2010, 2011; Koh et al., 2011; Kriaucionis and Heintz, 2009; Pastor et al., 2011; Tahiliani et al., 2009; Williams et al., 2011; Wossidlo et al., 2011; Wu et al., 2011a, 2011b; Xu et al., 2011). These studies have answered some fundamental questions, but they also raise others, underscoring the need for additional experiments that probe the mechanisms of DNA methylation and demethylation, how they are regulated, and how they affect gene expression.

In particular, the discovery of 5hmC raises a new technical conundrum. Many of the past methylation studies have relied on two techniques that cannot distinguish between 5mC and 5hmC: bisulfite conversion sequencing and methylation-sensitive restriction digests (Huang et al., 2010; Pastor et al., 2011). New tools have been developed based on specific modifications of 5hmC coupled with DNA immunoprecipitation and sequencing, chromatographic separation techniques, and improved immunohistochemical visualization using specific antibodies to distinguish 5mC and 5hmC (He et al., 2011; Pastor et al., 2011; Song et al., 2011; Wossidlo et al., 2011). Experiments performed using these new methodologies will undoubtedly enhance our understanding of the complex relationship between cytosine methylation and demethylation, as well as the new roles of cytosine hydroxymethylation and deamination.

#### **The AID/APOBEC Family: Mediators of 5mC or 5hmC Deamination**

Activation-induced cytidine deaminase (AID) has only recently been implicated in DNA demethylation (Bhutani et al., 2010; Cortellino et al., 2011; Guo et al., 2011; Popp et al., 2010). However, AID has been a focus of intense study by numerous groups over the past 10 years because of its critical role in generating antibody diversity in lymphocytes (reviewed in Chaudhuri et al., 2007; Delker et al., 2009). In B lymphocytes, AID participates in somatic hypermutation and class-switch recombination (Muramatsu et al., 2000), both of which entail error-prone DNA repair and, therefore, are mutagenic. AID is a member of a family of proteins, the APOBECs, which unlike AID were originally identified as RNA editors—hence their name “apolipoprotein B mRNA-editing catalytic polypeptides,” or APOBECs (reviewed in Conticello et al., 2007). AID mediates deamination of cytosine residues to uracils, which are then repaired by either BER or mismatch repair (MMR). This repair is error prone, leading to mutations essential to generating the vast repertoire of diverse antibodies seen in mammals (Liu and Schatz, 2009; Maul and Gearhart, 2010).

AID was thought to preferentially target the immunoglobulin (Ig) locus in B lymphocytes by unknown mechanisms, as the frequency of AID-generated mutations at non-Ig loci is very low. However, recent studies in B lymphocytes deficient in BER and MMR repair (i.e., *Ung*<sup>-/-</sup>*Msh2*<sup>-/-</sup>) revealed that AID acts extensively on non-Ig loci as well. These regions are protected from mutations, presumably by high-fidelity error-free

repair mechanisms (Liu et al., 2008). Clearly, an increased understanding of how error-prone and error-free DNA repair pathways are targeted to Ig versus non-Ig loci warrants further investigation, as AID is key to both antibody generation and DNA demethylation.

A role for AID in “global” DNA demethylation was first shown in zebrafish embryos by Cairns and colleagues (Rai et al., 2008). Upon overexpression of AID or zebrafish APOBEC deaminases and the glycosylase MBD4, active DNA demethylation was observed in zebrafish embryos injected with a methylated linearized nonreplicating DNA. Reik and colleagues suggested a similar role for AID in global DNA demethylation at a later stage of embryogenesis in mice (Popp et al., 2010). Mice completely lacking AID (AID<sup>-/-</sup>) (Muramatsu et al., 2000) exhibited an increase in genome-wide hypermethylation in their primordial germ cells (PGCs), suggesting that AID is involved in DNA demethylation. However, if AID-mediated global DNA demethylation plays a crucial role in early development, a more profound phenotype would be expected in AID null mice, which are both viable and fertile, albeit with somewhat smaller litter sizes (Popp et al., 2010). These findings raise the possibility that, in the absence of AID, other family members may play compensatory roles in DNA demethylation.

Studies of nuclear reprogramming provided the first evidence that AID plays a role in active DNA demethylation in mammals and in somatic cells (Bhutani, et al., 2010). Upon fusion of an excess of mouse ESCs with human somatic cells (fibroblasts) in nondividing heterokaryons, rapid demethylation was detected at the promoters of the human pluripotency genes *OCT4* and *NANOG*, accompanied by their transcriptional induction. This effect on reprogramming was dependent on AID, as a reduction of AID by four different siRNAs completely blocked pluripotency gene promoter demethylation and gene expression. In somatic cells, AID does not appear to act globally as in zebrafish embryos and mouse PGCs (Popp et al., 2010; Rai et al., 2008), but instead it is targeted to specific loci (Bhutani et al., 2010). However, if specific loci are involved, the question arises as to how targeting of this enzyme is mediated. Unlike TET, which has a DNA-binding motif (Ono et al., 2002), AID does not. Future studies to decipher the mechanism by which AID is targeted will not only provide insights into its sites of action but also illuminate its role in active DNA demethylation in different cell types and in response to different stimuli.

#### **The BER Glycosylase Family: Mediators of DNA Repair**

As described above, the accumulating data from the TET and AID/APOBEC studies suggest that active demethylation involves cytosine replacement via DNA repair. In principle, 5hmC or 5mC can be removed passively in the course of cell division. However, the rapid loss of methylation that occurs independent of DNA replication (Bhutani et al., 2010; Frank et al., 1990; Oswald et al., 2000; Paroush et al., 1990) must be mediated by an active mechanism. In plants, active DNA demethylation is a well-characterized process in which the accepted mechanism involves the BER pathway. BER glycosylases mediate the first step in the repair pathway by removing the methylated cytosine and creating an abasic site, which is then further acted upon by other enzymes (Gehring et al., 2009). In mammals, the process is more complex, as no glycosylases have been identified that act

directly on 5mC or 5hmC. Recently, an intermediate step, deamination, has been suggested to precede BER in mammalian DNA demethylation (Cortellino et al., 2011; Guo et al., 2011). The family of glycosylases implicated in the BER pathway are members of the uracil DNA glycosylase (UDG) family that includes thymine-DNA glycosylase (TDG) and single-strand-selective monofunctional uracil-DNA glycosylase 1 (SMUG1) (Cortellino et al., 2011; Guo et al., 2011). Therefore, like plants, mammalian DNA demethylation involves DNA repair pathways.

The DNA glycosylases TDG and SMUG1 are capable of converting 5hmU to cytosine, suggesting that they act in a partnership with TET and AID/APOBEC (Cortellino et al., 2011; Guo et al., 2011). Notably, mice lacking TDG (TDG<sup>-/-</sup>) are embryonic lethal, underscoring the significant role that BER glycosylases play during development and DNA demethylation. A direct physical interaction has been demonstrated between AID and TDG by coimmunoprecipitation experiments (Cortellino et al., 2011). In addition, recent reports suggest that 5hmC can be further oxidized by TET proteins to form 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (He et al., 2011; Ito et al., 2011). These are unstable intermediates that can be detected in ESCs, but they are 100-fold less abundant than 5hmC. Importantly, 5caC can be repaired by TDG, providing further support for a DNA repair pathway (He et al., 2011). However, another study also raises the possibility that a putative decarboxylase could directly convert 5caC to cytosine independent of DNA repair (Ito et al., 2011). The extent to which DNA repair pathways are involved in the removal of 5caC and its relative importance to DNA demethylation and gene expression remain to be determined.

In summary, the BER glycosylases, along with the TET and AID/APOBEC families of enzymes, mediate DNA demethylation via DNA repair. It remains to be tested whether other DNA repair pathways besides BER participate in DNA demethylation.

#### **Examples of Active DNA Demethylation in Mammals**

As described below, a body of evidence is accumulating in this nascent and rapidly evolving field, which supports the thesis that active DNA demethylation is more common than previously anticipated. Examples are presented below that indicate the role of DNA demethylation in rapid responses to changes in extrinsic signals, in early stages of development, and in highly specialized postmitotic cells. This is merely the tip of the iceberg, and more studies are needed to ascertain the extent to which the methylation editors are involved in the spatial and temporal regulation of DNA demethylation.

#### **Rapid Active DNA Demethylation in Response to Signals**

Perhaps the most striking example of “active” DNA demethylation in adult cells to date is the activity-dependent DNA demethylation at the promoters of brain-derived neurotrophic factor (BDNF) and fibroblast growth factor 1 (FGF1) in postmitotic neurons (Martinowich et al., 2003). Recent studies have elegantly elucidated the molecular mechanisms underlying this active DNA demethylation process and revealed a partnership between the TET, AID/APOBEC enzymes, and the BER glycosylases (Guo et al., 2011). Reconstitution in HEK293 cells and knockdown experiments demonstrate that TET-induced conversion of 5mC to 5hmC is followed by AID/APOBEC-mediated

deamination of 5hmC to 5hmU and its further replacement by an unmethylated cytosine through the BER pathway. Several laboratories have implicated both SMUG1 and TDG as the BER glycosylases in this process (Cortellino et al., 2011; Guo et al., 2011; He et al., 2011).

Other examples of active DNA demethylation as a rapid response to signal transduction include interleukin-2 (IL-2) stimulation of T lymphocytes and estrogen stimulation of breast cancer cells (Bruniquel and Schwartz, 2003). IL-2 activates T lymphocytes to mount an immune response, and this process is accompanied by a rapid demethylation of 5mCs within 20 min. This signal-dependent DNA demethylation is active because it is unaffected by the presence of an inhibitor of DNA synthesis. In breast cancer cells, the promoter of the pS2/trefoil factor 1 (TFF1) gene undergoes cycles of methylation and demethylation within 20–40 min in response to estrogen, demonstrating a strikingly dynamic interplay of demethylation and de novo methylation (Kangaspeska et al., 2008; Métivier et al., 2008). The putative roles of hydroxylation, deamination, and DNA repair remain to be explored in these scenarios.

#### **Active DNA Demethylation in Early Mammalian Development**

DNA demethylation of paternal and maternal genomes differs in the zygote, indicating that there is specificity and targeting of the DNA demethylation machinery even at this early stage of development. Studies report that, following fertilization, the paternal pronuclei undergo an extensive loss of 5mC (Oswald et al., 2000), whereas the maternal pronuclei are resistant to this loss due to the presence of the protein Stella (Nakamura et al., 2007). The loss of 5mC in the paternal genome is rapid and independent of cell division, and it serves as a classic example of active DNA demethylation. More recent studies have revealed that the active loss of 5mC is actually a conversion of 5mC to 5hmC in a TET3-dependent manner (Iqbal et al., 2011; Wossidlo et al., 2011). The fate of the resulting 5hmC and how it may be converted back to unmethylated cytosines remain to be elucidated.

A similar active DNA demethylation process has been reported in mouse PGCs (Popp et al., 2010). It is still unknown, however, whether 5hmC plays a role in this process. The genome-wide hypermethylation observed in AID null PGCs has suggested a role for AID-mediated deamination in active DNA demethylation (Popp et al., 2010). However, this study used bisulfite conversion and sequencing, which cannot distinguish between 5mC and 5hmC (Huang et al., 2010). Thus, it remains possible that 5hmC generated by TET activity is an intermediate for AID-mediated deamination and subsequent DNA demethylation. Independently, DNA repair by the BER pathway has been reported by other investigators to occur in PGCs (Hajkova et al., 2010). Taken together, it is reasonable to speculate that the TET–AID/APOBEC–BER pathway plays a role in DNA demethylation in early development both in zygotes and in PGCs.

#### **Active DNA Demethylation in Tissue-Specific Differentiation**

Skeletal muscle constitutes an example of somatic cells in mammals in which active DNA demethylation has been reported (Blau et al., 1983; Zhang et al., 2007). In these reprogramming studies, when human fibroblasts and mouse muscle cells were

fused to form nondividing heterokaryons, active DNA demethylation was observed at the human MyoD promoter, which accompanied its activation and expression in the fibroblasts. Remarkably, the Cedar laboratory postulated more than 20 years ago that DNA demethylation occurs by an active mechanism in muscle cells (Paroush et al., 1990; Weiss et al., 1996), but the factors responsible were unknown. These studies suggest that a dynamic interplay of methylation and demethylation may also function during differentiation.

#### **Is DNA Methylation-Demethylation Bidirectional?**

Although indirect evidence has been accumulating for decades, recent advances discussed here now support the hypothesis that DNA demethylation and methylation may be bidirectional and dynamically regulated throughout early and late development and in certain adult tissues, especially the brain (Guo et al., 2011; Miller and Sweatt, 2007). Much remains to be learned, including which loci are targeted for demethylation and how the process is spatially and temporally regulated in diverse cell types and stages of development. The long-held notion that DNA methylation patterns are generally maintained in stable differentiated states is likely true. Nonetheless, nuclear reprogramming shows that perturbations are possible (Jullien et al., 2011; Yamanaka and Blau, 2010). As is the case for the regulation of gene expression by transcription factors (Blau and Baltimore, 1991; Jacob and Monod, 1961; Ptashne, 2009; Blau, 1992), the regulation of DNA methylation may also be continuous and dictated by a balance of enzymes and targeting factors.

As shown in Figure 1, our current understanding of the DNA methylation and demethylation circuitry entails members of the following enzyme families with roles in either passive or active DNA demethylation: (1) the DNMT family of three methyltransferases responsible for the de novo generation and maintenance of 5mC. DNA demethylation can occur passively by a dilution or inactivation of DNMTs; (2) the TET family of three 5mC hydroxylases, which generate 5hmC (and further oxidized intermediates) from 5mC; (3) the AID/APOBEC family of deaminases, which initiate an active process of demethylation by deaminating either 5mC or 5hmC generated by the TET family; (4) the family of BER glycosylases that initiate DNA repair culminating in the replacement of methylated cytosines with unmethylated cytosines. We have designated these enzymes as the DNA methylation editors that are responsible for the regulation of the DNA methylome associated with a particular cell fate. It remains to be determined whether active DNA demethylation in different scenarios always requires a representative member from each of these families. In other words, does the entire TET–AID/APOBEC–BER pathway operate broadly, or is only a subset thereof required to achieve active DNA demethylation in different cell contexts.

The concept of a dynamic interplay of regulators has emerged in parallel with the demonstration of the remarkable plasticity of cellular fates illustrated by nuclear reprogramming. When the balance of transcription factors that recognize DNA sequence is perturbed by either nuclear transfer, cell fusion, or defined factors (i.e., in generating iPSCs), it leads to a dramatic shift in cell fate. A provocative, yet perhaps overly simplistic view of how cell fate is controlled and maintained is provided by an

analogy with a sailboat. Transcription factors comprise the rudder that determines the direction of the differentiated state (i.e., whether it is muscle or liver). Threshold concentrations of transcription factors, achieved by feedback loops, continuously regulate the differentiated states. Similarly, the editors of DNA methylation described in this Review can be regulated actively and continuously serving as the keel and preventing the cell from responding to minor changes in wind or current. A blast of ectopic transcription factors can overwhelm the rudder and reset it as well as the DNA methylation regulators. This occurs in cellular reprogramming, either following nuclear transfer into oocytes, upon cell fusion in heterokaryons, or in induced pluripotency (iPSCs). In the first two cases, the somatic nucleus encounters an overwhelming abundance of pre-existing proteins, whereas in iPSCs, this protein abundance is progressive, as it derives from the overexpression of four genes (reviewed in [Yamanaka and Blau, 2010](#)).

The recent discoveries that TET and AID/APOBEC enzymes are active regulators of DNA demethylation support the hypothesis that even apparently stable states are continuously regulated ([Blau, 1992](#); [Blau and Baltimore, 1991](#)). Thus, the stable differentiated state is governed by regulatory pathways that are surprisingly perturbable. This raises the intriguing question of how cellular plasticity is kept in check to maintain specific cell fates. A future goal and major challenge is to understand how cell plasticity can be first enlisted to reprogram cells and then regulated to derive stable differentiated cell types. Indeed, understanding the mechanisms that govern this dichotomy is critical for successfully applying cellular reprogramming to regenerative medicine.

## ACKNOWLEDGMENTS

We thank Steve Hennikoff, Mary Goll, Mark Ptashne, Peter Jones, Jason Pomerantz, and members of our laboratory for critical reading of the manuscript and constructive suggestions and Stephane Corbel for expert artwork. This work was funded by a PHS grant 5T32CA09151 awarded by NCI to D.M.B. and funding from NIH grants HL096113, HL100397, AG020961, and AG009521; JDRF grant 34-2008-623; CIRM grants RT1-01001 and RB1-01292; and the Baxter Foundation to H.M.B.

## REFERENCES

- Bhutani, N., Brady, J.J., Damian, M., Sacco, A., Corbel, S.Y., and Blau, H.M. (2010). *Nature* **463**, 1042–1047.
- Bird, A. (2002). *Genes Dev.* **16**, 6–21.
- Blau, H.M. (1992). *Annu. Rev. Biochem.* **61**, 1213–1230.
- Blau, H.M., and Baltimore, D. (1991). *J. Cell Biol.* **112**, 781–783.
- Blau, H.M., Chiu, C.P., and Webster, C. (1983). *Cell* **32**, 1171–1180.
- Bruniquel, D., and Schwartz, R.H. (2003). *Nat. Immunol.* **4**, 235–240.
- Chaudhuri, J., Basu, U., Zarrin, A., Yan, C., Franco, S., Perlot, T., Vuong, B., Wang, J., Phan, R.T., Datta, A., et al. (2007). *Adv. Immunol.* **94**, 157–214.
- Coticello, S.G., Langlois, M.A., Yang, Z., and Neuberger, M.S. (2007). *Adv. Immunol.* **94**, 37–73.
- Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., et al. (2011). *Cell* **146**, 67–79.
- De Carvalho, D.D., You, J.S., and Jones, P.A. (2010). *Trends Cell Biol.* **20**, 609–617.
- Deaton, A.M., and Bird, A. (2011). *Genes Dev.* **25**, 1010–1022.
- Delker, R.K., Fugmann, S.D., and Papavasiliou, F.N. (2009). *Nat. Immunol.* **10**, 1147–1153.
- Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S., and Reik, W. (2011). *Nature* **473**, 398–402.
- Frank, D., Lichtenstein, M., Paroush, Z., Bergman, Y., Shani, M., Razin, A., and Cedar, H. (1990). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **326**, 241–251.
- Gehring, M., Reik, W., and Henikoff, S. (2009). *Trends Genet.* **25**, 82–90.
- Globisch, D., Münzel, M., Müller, M., Michalakis, S., Wagner, M., Koch, S., Brückl, T., Biel, M., and Carell, T. (2010). *PLoS ONE* **5**, e15367.
- Goll, M.G., and Bestor, T.H. (2005). *Annu. Rev. Biochem.* **74**, 481–514.
- Guo, J.U., Su, Y., Zhong, C., Ming, G.L., and Song, H. (2011). *Cell* **145**, 423–434.
- Hajkova, P., Jeffries, S.J., Lee, C., Miller, N., Jackson, S.P., and Surani, M.A. (2010). *Science* **329**, 78–82.
- He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). *Science*. Published online August 4, 2011.
- Huang, Y., Pastor, W.A., Shen, Y., Tahiliani, M., Liu, D.R., and Rao, A. (2010). *PLoS ONE* **5**, e8888.
- Iqbal, K., Jin, S.G., Pfeifer, G.P., and Szabó, P.E. (2011). *Proc. Natl. Acad. Sci. USA* **108**, 3642–3647.
- Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). *Nature* **466**, 1129–1133.
- Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). *Science*. Published online July 21, 2011.
- Jacob, F., and Monod, J. (1961). *J. Mol. Biol.* **3**, 318–356.
- Jaenisch, R., and Bird, A. (2003). *Nat. Genet. Suppl.* **33**, 245–254.
- Jullien, J., Pasque, V., Halley-Stott, R.P., Miyamoto, K., and Gurdon, J.B. (2011). *Nat. Rev. Mol. Cell Biol.* **12**, 453–459.
- Kangaspekka, S., Stride, B., Métivier, R., Polycarpou-Schwarz, M., Ibberson, D., Carmouche, R.P., Benes, V., Gannon, F., and Reid, G. (2008). *Nature* **452**, 112–115.
- Koh, K.P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C.A., Mostoslavsky, G., et al. (2011). *Cell Stem Cell* **8**, 200–213.
- Kriaucionis, S., and Heintz, N. (2009). *Science* **324**, 929–930.
- Kucharski, R., Maleszka, J., Foret, S., and Maleszka, R. (2008). *Science* **319**, 1827–1830.
- Law, J.A., and Jacobsen, S.E. (2010). *Nat. Rev. Genet.* **11**, 204–220.
- Liu, M., and Schatz, D.G. (2009). *Trends Immunol.* **30**, 173–181.
- Liu, M., Duke, J.L., Richter, D.J., Vinuesa, C.G., Goodnow, C.C., Kleinstein, S.H., and Schatz, D.G. (2008). *Nature* **451**, 841–845.
- Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., Fan, G., and Sun, Y.E. (2003). *Science* **302**, 890–893.
- Maul, R.W., and Gearhart, P.J. (2010). *Adv. Immunol.* **105**, 159–191.
- Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). *Nature* **403**, 501–502.
- Métivier, R., Gallais, R., Tiffoche, C., Le Péron, C., Jurkowska, R.Z., Carmouche, R.P., Ibberson, D., Barath, P., Demay, F., Reid, G., et al. (2008). *Nature* **452**, 45–50.
- Miller, C.A., and Sweatt, J.D. (2007). *Neuron* **53**, 857–869.
- Monk, M., Adams, R.L., and Rinaldi, A. (1991). *Development* **112**, 189–192.
- Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). *Cell* **102**, 553–563.
- Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M., et al. (2007). *Nat. Cell Biol.* **9**, 64–71.
- Ono, R., Taki, T., Taketani, T., Taniwaki, M., Kobayashi, H., and Hayashi, Y. (2002). *Cancer Res.* **62**, 4075–4080.
- Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). *Curr. Biol.* **10**, 475–478.

- Paroush, Z., Keshet, I., Yisraeli, J., and Cedar, H. (1990). *Cell* 63, 1229–1237.
- Pastor, W.A., Pape, U.J., Huang, Y., Henderson, H.R., Lister, R., Ko, M., McLoughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P., et al. (2011). *Nature* 473, 394–397.
- Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). *Nature* 463, 1101–1105.
- Ptashne, M. (2009). *Curr. Biol.* 19, R234–R241.
- Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). *Cell* 135, 1201–1212.
- Rougier, N., Bourc'his, D., Gomes, D.M., Niveleau, A., Plachot, M., Paldi, A., and Viegas-Péquignot, E. (1998). *Genes Dev.* 12, 2108–2113.
- Ruzov, A., Tsenkina, Y., Serio, A., Dudnakova, T., Fletcher, J., Bai, Y., Chebotareva, T., Pells, S., Hannoun, Z., Sullivan, G., et al. (2011). *Cell Res.* Published online July 12, 2011. 10.1038/cr.2011.113.
- Song, C.X., Szulwach, K.E., Fu, Y., Dai, Q., Yi, C., Li, X., Li, Y., Chen, C.H., Zhang, W., Jian, X., et al. (2011). *Nat. Biotechnol.* 29, 68–72.
- Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., and Rao, A. (2009). *Science* 324, 930–935.
- Weiss, A., Keshet, I., Razin, A., and Cedar, H. (1996). *Cell* 86, 709–718.
- Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P.A., Rappsilber, J., and Helin, K. (2011). *Nature* 473, 343–348.
- Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). *Nat Commun* 2, 241.
- Wu, H., D'Alessio, A.C., Ito, S., Wang, Z., Cui, K., Zhao, K., Sun, Y.E., and Zhang, Y. (2011a). *Genes Dev.* 25, 679–684.
- Wu, H., D'Alessio, A.C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Sun, Y.E., and Zhang, Y. (2011b). *Nature* 473, 389–393.
- Wyatt, G.R., and Cohen, S.S. (1952). *Nature* 170, 1072–1073.
- Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., Zhang, H., Huang, S., et al. (2011). *Mol. Cell* 42, 451–464.
- Yamanaka, S., and Blau, H.M. (2010). *Nature* 465, 704–712.
- Zhang, F., Pomerantz, J.H., Sen, G., Palermo, A.T., and Blau, H.M. (2007). *Proc. Natl. Acad. Sci. USA* 104, 4395–4400.