

REVIEW

Open Access



Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond

Achim Breiling and Frank Lyko*

Abstract

The chemical modification of DNA bases plays a key role in epigenetic gene regulation. While much attention has been focused on the classical epigenetic mark, 5-methylcytosine, the field garnered increased interest through the recent discovery of additional modifications. In this review, we focus on the epigenetic regulatory roles of DNA modifications in animals. We present the symmetric modification of 5-methylcytosine on CpG dinucleotide as a key feature, because it permits the inheritance of methylation patterns through DNA replication. However, the distribution patterns of cytosine methylation are not conserved in animals and independent molecular functions will likely be identified. Furthermore, the discovery of enzymes that catalyse the hydroxylation of 5-methylcytosine to 5-hydroxymethylcytosine not only identified an active demethylation pathway, but also a candidate for a new epigenetic mark associated with activated transcription. Most recently, N6-methyladenine was described as an additional eukaryotic DNA modification with epigenetic regulatory potential. Interestingly, this modification is also present in genomes that lack canonical cytosine methylation patterns, suggesting independent functions. This newfound diversity of DNA modifications and their potential for combinatorial interactions indicates that the epigenetic DNA code is substantially more complex than previously thought.

Keywords: DNA modification, Modified DNA bases, DNA methylation, Epigenetic marks, Gene regulation, Chromatin, Inheritance, 5-Methylcytosine, 5-Hydroxymethylcytosine, N6-methyladenine, DNA demethylation

Background

To establish and maintain cellular identity during development, specific memory mechanisms have evolved that regulate gene expression patterns epigenetically. Once determined, these lineage-specific expression profiles have to be maintained through cell divisions. Active or inactive states of gene expression are defined by specific epigenetic modification patterns that are either accessible to transcription factors and activators, or result in a closed chromatin structure that prevents activated transcription [1–3]. Central to this is the concept of epigenetic marks, specific DNA or chromatin modifications that can be inherited through cell divisions. These marks maintain the epigenetic information and serve as interaction sites for specific binder or reader proteins, which include epigenetic modifier enzymes, repressors,

chromatin remodeling complexes and the transcription machinery. The most prominent of these marks is the methylation of the carbon-5 of cytosine (5mC), which is traditionally considered incompatible with activated transcription when present near gene regulatory regions. At these regions, 5mC can modulate the binding of transcription factors [4, 5] or induce the binding of specific 5mC-binding proteins that can lead to the recruitment of co-repressor complexes to methylated target promoters [6].

While there is an enormous number of published studies on epigenetic modifications, most of them are correlative in nature. This is exemplified by the increasing use of powerful genome-wide mapping technologies that have revealed numerous associations between changes in epigenetic modification patterns and cell fate transitions [7–9]. However, functional insight remains relatively limited. Furthermore, the field has broadened significantly through the discovery of two additional DNA modifications with epigenetic regulatory functions,

*Correspondence: f.lyko@dkfz.de

Division of Epigenetics, DKFZ-ZMBH Alliance, German Cancer Research Center, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany

5-hydroxymethylcytosine (5hmC) and N6-methyladenine (6mA), as well as the identification of the corresponding modifying enzymes (Figure 1). Our review aims to illustrate the epigenetic regulatory functions of these DNA modifications, with a predominant focus on animal models. Epigenetic regulation in plants has recently been reviewed elsewhere [10–12].

5-Methylcytosine: the fifth base

5-Methylcytosine has been termed the “fifth base” of the human genome. This reflects the relatively high abundance of this modification, as about 4% of the cytosine residues in the human genome have been found to be methylated. However, cytosine methylation levels can differ greatly among animal genomes (see below), and it would therefore be misleading to define the significance of 5mC by its abundance. Rather, the key feature of cytosine methylation is its enrichment or even specificity for “symmetric” CpG dinucleotides [13]. Symmetric methylation means that methylation marks are present on both strands of DNA and that methylation patterns can be faithfully propagated through DNA replication by copying from the parental strand to the unmethylated newly synthesized strand. This methylation maintenance is carried out by the Dnmt1 DNA methyltransferase that has a strong preference for hemimethylated DNA and provides a key paradigm for the stability and heritability of epigenetic information [14]. Dnmt1 is complemented by the Dnmt3 DNA methyltransferases that do not show any selectivity for hemimethylated DNA and have therefore been termed “de novo methyltransferases” [14]. Together,

both enzymes catalyze the establishment and maintenance of cytosine DNA methylation patterns during animal development and cell fate specification.

While the overall specificity of animal methylation patterns for CpG dinucleotides has been confirmed in numerous studies, several notable exceptions have also been described. A prominent example is non-CpG methylation in mouse embryonic stem cells (ESCs), which was verified in the first genome-wide methylation analysis of ESCs [15]. While levels of non-CpG methylation are very low in most somatic tissues, extensive postnatal accumulation of this modification has been observed in the mouse and human brain [16–18]. Targeted depletion of Dnmt3a in specific brain regions resulted in significant reduction of non-CpG methylation [18, 19]. In contrast to ESCs where non-CpG methylation seems to correlate with gene expression [15], the modification exhibited an inverse correlation with transcription in neurons, which could partly be explained through the recruitment of the methyl-CpG binding protein 2 (MeCP2) [18, 19]. Context dependent non-CpG methylation might therefore have an impact on specific readers of DNA methylation, thus influencing tissue-specific gene expression.

Beyond mammalian methylomes, the comparative analysis of single-base resolution methylation maps has shown a substantial degree of variation between animal species [15, 20, 21]. The available information can be used to define three major categories (Figure 2): the first group is defined by mammalian methylomes and is characterized by pervasive methylation. In the human genome, more than 80% of the CpG dinucleotides are methylated,

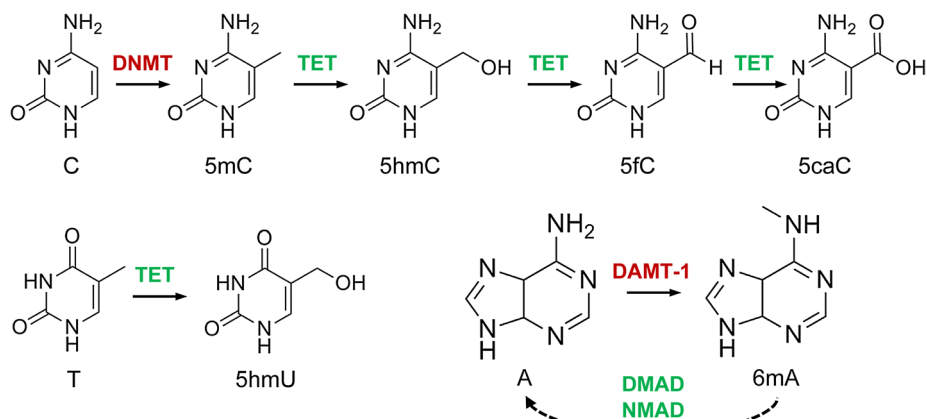
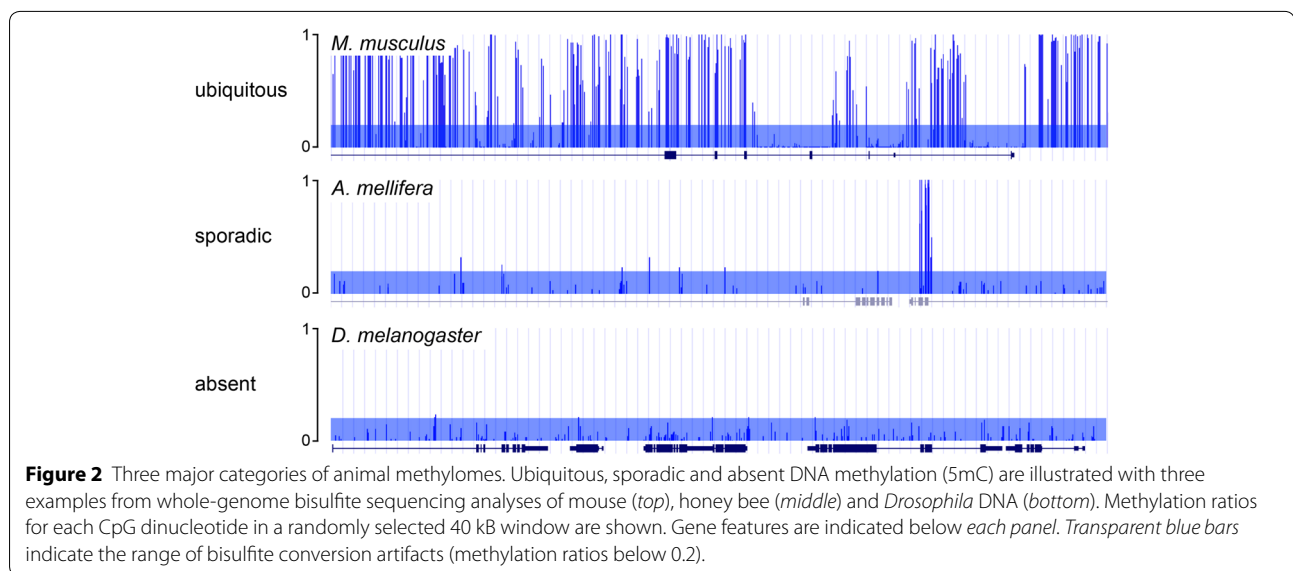


Figure 1 DNA modifications with epigenetic regulatory functions and their interdependencies. Cytosine (C) is methylated to 5-methylcytosine (5mC) by DNA methyltransferases (DNMT) and then further oxidised to 5hmC, 5fC and 5caC by Tet dioxygenases. 5-Hydroxymethyluracil (5hmU) is produced by Tet-catalysed oxidation of thymine (T). N6-methyladenine (6mA) is likely catalysed by DNA N6 adenine methyltransferases (DAMT-1 in *C. elegans*), even though the biochemical activity of these enzymes remains to be characterized. The Tet-like ALKB enzymes NMAD (N6-methyl adenine demethylase 1) and DMAD (DNA 6mA demethylase) have been shown to be involved in 6mA demethylation in *C. elegans* and in *Drosophila*, respectively, possibly by using a conserved dioxygenase mechanism.



creating a landscape of ubiquitous methylation, but with local gaps that are often found at active regulatory elements, such as promoters and enhancers (Figure 2). It seems plausible to assume that the default state of these methylomes is “methylated” and that active mechanisms (see below) are required to keep specific regions free of methylation.

The second group is exemplified by the honeybee methylome, that can be defined by only 60,000 CpG-specific methylation marks that are highly enriched in exons [22]. In this case, the default state of the genome appears to be “unmethylated” and the selective targeting of DNA methyltransferases to specific CpGs would be a key step for shaping the methylation landscape (Figure 2). Such sporadic methylation patterns have been described in several animals, particularly in insects. However, the functional significance of sparse methylation remains to be fully understood, which is largely due to the limited potential of the corresponding organisms for genetic manipulation. Importantly, it has been shown that queen-like phenotypes can be enhanced in honeybees following siRNA-mediated knockdown of the Dnmt3 orthologue [23]. While the mechanisms underlying this phenomenon remain to be elucidated, these results strongly suggest a functional role of this enzyme in caste specification, possibly through the modulation of caste-specific methylation patterns.

Finally, several animal genomes have failed to reveal canonical cytosine methylation patterns (Figure 2), which implies that 5mC is not essential for development and cell fate specification of well-known laboratory models such as *S. cerevisiae*, *S. pombe*, *C. elegans* and *D. melanogaster* [24]. The absence of conserved cytosine

methylation patterns in these organisms was instrumental for the identification and characterization of other epigenetic mechanisms, including covalent histone modifications and small noncoding RNAs [25–27]. Moreover, it also played an important role in the recent discovery of N6-methyladenine as an epigenetic DNA modification in eukaryotes (see below).

The functional analysis of cytosine methylation has proven to be surprisingly complex and difficult, even in well-characterized mammalian organisms. While knockout models demonstrated a role of Dnmt1 and Dnmt3 in mouse development [28, 29] and in general epigenetic phenomena, such as genomic imprinting [30], X-chromosome inactivation [31] and transposon control [32], the specific function of cytosine methylation in epigenetic gene regulation remains to be fully understood. However, recent integrative studies that combine the targeted disruption of Dnmt genes with genome-wide mapping approaches have provided interesting insight into the functional specificities of individual Dnmts. For example, Dnmt3a-mediated gene body methylation at transcriptionally active genes was shown to be prevalent in postnatal neuronal stem cells and is required for postnatal neurogenesis [33]. In addition, other Dnmts were found to interact with actively transcribed gene bodies, suggesting that gene body methylation promotes transcription [34]. Most recently, Dnmt3b-mediated gene body methylation in mouse ESCs was shown to depend on the presence of histone H3 lysine 36 methylation in the same regions [35]. This represents a novel and unexpected feature of de novo methyltransferases, as it suggests the recruitment of cytosine methyltransferases by the co-transcriptional modification of histones.

In another study, it was shown that human embryonic stem cells lacking both DNMT3A and DNMT3B progressively lose cytosine methylation marks, thus illustrating an imperfect maintenance activity of DNMT1 and a supporting role of DNMT3 enzymes in maintenance methylation [36]. Similar results were obtained with Dnmt-deficient mouse ESCs, which also revealed differential specificities of Dnmt1 and Dnmt3a/b for distinct subclasses of retrotransposons [37]. Further analyses of human ESCs revealed a novel role of DNMT3A in the hypermethylation of genes associated with endoderm differentiation and a rapid, replication-dependent loss of global DNA methylation in DNMT1-deficient cells [36]. It will be important to use similar approaches for the characterization of additional cell types and model systems in order to fully understand the epigenetic regulatory function of 5mC.

5-Hydroxymethylcytosine: oxidation creates a new modification

With the discovery of the catalytic dioxygenase activity of Ten eleven translocation (Tet) proteins, novel epigenetic DNA modifications started to emerge [38, 39]. 5-Hydroxymethylcytosine (5hmC, Figure 1) was originally discovered in mammalian DNA in 1972 [40], but its biological significance was investigated only almost 40 years later [41]. Cytosine hydroxymethylation levels are often around 0.1% in mammalian tissues, but can vary greatly [42], with highest values in the brain, where up to 1% of the cytosines can be hydroxymethylated [41]. The three mammalian Tet homologues generate 5hmC from existing 5mC, which they can further process to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC, Figure 1) [43, 44]. About 30,000 molecules of 5mC, 1,300 of 5hmC, 20 of 5fC, and 3 of 5caC were found per million Cs in mouse embryonic stem cells [44, 45], indicating a very low abundance of 5fC and 5caC. As both modifications are targeted by base excision repair mechanisms mediated by thymine-DNA-glycosylases, they are mainly interpreted as intermediates of an active demethylation pathway via Tet-dependent 5mC oxidation [43, 44].

We are only beginning to understand the functional significance of 5hmC as an epigenetic mark and the specific roles of the three Tet enzymes. Tet1 and Tet2 are highly expressed in mouse ESCs, but their single depletion does not affect pluripotency or development [46–49]. Tet3 homozygous mutant mice develop properly, but die at birth [50], suggesting that Tet3 is also dispensable for embryonic development. ESCs deficient for both Tet1 and Tet2 show insignificant levels of 5hmC, but retain pluripotency. However, the majority of mice lacking both proteins showed developmental defects, which was found to be associated with ectopic hypermethylation

[51]. Combined deficiency of all three Tet proteins in ESCs depleted 5hmC completely, but did not affect ESC viability and pluripotency [52–54]. Nevertheless, triple knockout ESCs and embryoid bodies showed impaired differentiation potential, promoter hypermethylation and correlated deregulation of genes implicated in embryonic development and differentiation [52]. In agreement, severe defects in somatic cell reprogramming and mesenchymal-epithelial transition have been described in double and triple Tet knockout mouse embryonic fibroblasts [53].

These data point to a major role of Tet-mediated oxidation in DNA demethylation, most likely by keeping regulatory genomic regions free of 5mC. Particularly important are enhancers, that have been shown to be hypermethylated in Tet-deficient mouse ESCs, resulting in a reduced activity of associated differentiation genes [54, 55]. Tet-dependent oxidation of 5mC as a first step of active demethylation is therefore an early event of enhancer activation [54–56], but might also more generally allow functional interactions with regulatory DNA elements and counteract aberrant spreading of DNA methylation into CpG islands [57].

Nevertheless, 5hmC was also found as a relative stable base at a subset of mammalian promoters, at gene bodies of actively transcribed genes and at poised and active enhancers [58, 59]. 5fC was also mapped to a subset of these 5hmC-marked regions [60–62], suggesting a role as an independent epigenetic mark. Indeed, several “reader” proteins for oxidised 5mC-derivatives have been identified, which might mediate epigenetic regulation [63, 64]. Among these were, in addition to DNA damage- and repair-related proteins, chromatin modifiers and transcriptional regulators like e.g. MBD3, MeCP2, UHRF2 and FOX transcription factors [64–66]. While the functional relevance and specificity of the interactions remains to be fully understood (e.g. many 5hmC interacting proteins also have significant affinities for 5mC) these readers might recruit chromatin regulatory complexes to their targets and support activated transcription.

A role of 5hmC as active mark is supported by mass spectrometric analyses of isotope labelled DNA from mammalian cell culture and mice showing that 5hmC is mostly a stable modification and not a transient intermediate [67]. The high abundance in post-mitotic brain tissues [41, 42] also suggests a direct epigenetic function of 5hmC. Indeed, 5hmC levels increase during neuronal differentiation and a very stable intragenic enrichment of 5hmC was observed at many active neuron-specific genes [66, 68–70]. These findings suggest that 5hmC functions as epigenetic mark in mammalian neuronal development. This is further supported by the observations that the activated human HOXA cluster becomes stably enriched

in 5hmC upon retinoic acid stimulated neuronal differentiation [71] and that increased 5hmC levels at neuronal marker genes in Sirtuin-6-deficient mice induce skewed differentiation versus neuroectoderm [72].

While there is evidence for a direct epigenetic function for 5hmC at least in some tissues, a similar role for its oxidation derivatives appears less likely. The levels of 5fC and 5caC have been found to increase at 5fC sites in thymine-DNA-glycosylase-deficient mouse ESCs, suggesting that 5caC sites primarily represent sites of active demethylation [60–62]. It remains possible that, due to the chemical differences between the oxidised 5mC-derivatives, each modification might attract specific readers. However, considering the relatively strong DNA-damage response triggered by 5fC and 5caC (in contrast to 5hmC) and their very low abundances, it seems more likely that these modifications transiently accumulate at the regions of the hydroxymethylome that undergoes demethylation. In contrast, a subset of 5hmC sites appears to be stable and might act as an independent epigenetic mark. Very recently, it has been shown that Tet proteins can also oxidize thymine to 5-hydroxymethyluracil (5hmU, Figure 1) [73]. Tet-dependent 5hmU is present at levels similar to 5caC in mESCs, increases during early ESC differentiation and recruits specific interacting proteins [73], suggesting an epigenetic function for Tet-dependent 5hmU. Nevertheless, 5hmU paired with adenine is a target for the Smug1 DNA glycosylase [74] and might therefore trigger base excision repair mechanisms. Indeed knock down of Smug1 in mESCs led to increased 5hmU levels [73], indicating that 5hmU might also serve to promote active demethylation by recruiting repair factors to Tet targets.

N6-methyladenine: revival of an old acquaintance

In bacterial genomes 5mC is outshined by a second base modification, N6-methyladenine (6mA, Figure 1). Adenine methylation has been shown to be essential for the viability of several bacteria, as methylation of GATC sequences by the Dam methylase creates specific marks that are important for DNA replication, chromosome segregation, mismatch repair and the regulation of gene expression [75, 76]. However, several older studies also suggested the presence of 6mA in eukaryotic genomes, even though detection was often indirect and modification levels appeared close to the detection limit [76]. Several unicellular eukaryotes, including the green alga *Chlamydomonas reinhardtii*, had consistently shown comparably high levels of DNA adenine methylation [76], which established this organism as an attractive model to investigate 6mA further.

Over the past few years, several powerful technologies were developed to analyze 6mA in RNA, where this

modification plays an important regulatory role. When these methods were adapted to characterize the distribution of 6mA in the *Chlamydomonas* genome, some key characteristics of this modification could be defined [77]. For example, the results showed that the algal adenine methylome consists of about 85,000 fully methylated 6mA sites, corresponding to a global adenine methylation level of approximately 0.4%. Methylation was often found in symmetric ApT target sequences, but there was no evidence for symmetric 6mA methylation. The modification was enriched at promoter regions, and particularly in linker regions between adjacent nucleosomes. The authors propose a model in which the DNA 6mA modification either restricts or marks the positions of nucleosomes near transcriptional start sites in *Chlamydomonas*. As such, the presence of 6mA may position nucleosomes to facilitate initiation of transcription. While these findings are highly interesting, they are difficult to generalize because of a highly specific periodic pattern of nucleosome occupancy around transcriptional start sites in *Chlamydomonas*. Furthermore, the *Chlamydomonas* genome has an unusual pattern of 5mC: it is characterised by low levels of CpG methylation but also contains CHG and CHH methylation in gene bodies, which corresponds to known plant methylation patterns [20].

A parallel study also revealed novel details of adenine DNA methylation in *Caenorhabditis elegans* [78]. Similar to *Chlamydomonas*, adenine methylation was found to be variable, and maximum levels were rather low (0.3%). Mapping of 6mA residues by SMRT sequencing revealed that methylation was targeted to GAGG and AGAA consensus sequences, indicating strand-specific adenine methylation. Interestingly, 6mA accumulated in worms deficient for *spr-5* (coding for a H3K4me2 demethylase), an important paradigm of trans-generational epigenetic inheritance [78]. Further work led to the identification of a *C. elegans* DNA adenine demethylase (Nmad-1), belonging to the ALKB family of dioxygenases that also contains the Tet proteins. In addition, the authors identified a candidate DNA adenine methyltransferase (Damt-1) related to bacterial 6mA DNA methyltransferases. This enzyme belongs to a highly conserved family of proteins that is characterized by a C-terminal circularly permuted methyltransferase domain fused to a distinctive N-terminal domain [79]. While the biochemical activity of the enzyme remains to be characterized, deletion of Damt-1 suppressed the trans-generational phenotypes of *spr-5* mutant worms, suggesting that 6mA might be a transgenerationally inheritable epigenetic mark.

Additional insight into the function of adenine methylation came from a recent analysis in *Drosophila*. Flies represent a particularly interesting model for DNA

modifications, because of the longstanding controversial discussions surrounding the cytosine methylation status of the *Drosophila* genome. In addition, the fly genome encodes an unusual DNA methylation machinery, with no canonical Dnmt1/3 homologue, but with a clear Tet homologue. The former is consistent with the reported absence of Dnmt-dependent cytosine methylation patterns in *Drosophila* [24, 80], but the latter seemed to indicate that methylation may have been overlooked so far. By using highly sensitive mass spectrometry approaches, Zhang et al. have now demonstrated the presence of low (0.07%) but significant levels of adenine methylation during the earliest stages of *Drosophila* embryogenesis [81]. Most interestingly, the authors showed 6mA demethylation by the *Drosophila* Tet homologue DMAD in vitro and a specific increase of 6mA levels in the genomic DNA of DMAD mutants suggesting that DMAD is a 6mA-specific enzyme [81]. Furthermore, both deletion and overexpression of DMAD resulted in lethality, thus demonstrating an important developmental function of 6mA in *Drosophila*. One such function could be the regulation of transposons, as 6mA appeared enriched in transposon regions and transposons marked with 6mA were derepressed in DMAD mutants. Taken together, if 6mA will also be found in significant quantities in the genome of other eukaryotes, it might turn out to be an important carrier of epigenetic information, involved in the regulation of gene expression and possibly playing a complementary role to 5mC at certain loci or during specific stages of development.

Conclusions

Epigenetic DNA modifications generally affect the accessibility of genomic regions for regulatory proteins or protein complexes, for example by preventing interactions or by recruiting specific readers. Consequently, this can influence the chromatin structure and/or directly regulate enhancer and promoter activity or transcriptional processivity. Cytosine methylation is so far the only known symmetric modification with an established maintenance mechanism, which represents a unique feature that currently distinguishes 5mC from all other epigenetic modifications. 5mC has mostly been related to gene repression, in particular at enhancer and promoter regions of genes (Figure 3), but might also play an important role in positively influencing transcription, either by recruiting methylation-specific transcription factors [82, 83] or by a yet to be understood mechanism when present in the body of active genes [35].

Dynamic epigenetic processes also require the active removal of a mark. With the discovery of the enzymatic functions of the Tet proteins, the main enzymes for the removal of DNA methylation were identified. 5hmC and its Tet-dependent oxidation products are demethylation intermediates, but might also have significant roles as independent epigenetic marks (Figure 3). Specific readers for 5hmC, 5fC and 5caC have been identified that function in transcription regulation and chromatin remodeling, mostly promoting the active state. In addition, 5fC, 5caC and 5hmU might primarily function in

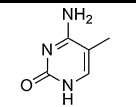
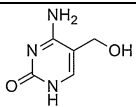
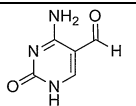
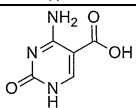
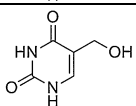
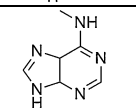
	5mC	5-methylcytosine. Repressive mark at enhancers and promoters, enriched in active gene bodies, recruits specific binders.
	5hmC	5-hydroxymethylcytosine. Oxidised 5mC, demethylation intermediate. Active mark at enhancers and gene bodies, recruits specific binders.
	5fC	5-formylcytosine. Oxidised 5hmC, demethylation intermediate, recruits DNA repair machinery.
	5caC	5-carboxycytosine. Oxidised 5fC, demethylation intermediate, recruits DNA repair machinery.
	5hmU	5-hydroxymethyluracil. Oxidised form of thymine, recruits DNA repair machinery.
	6mA	N6-methyladenine: active mark, enriched at TSS, influences nucleosome positioning, demethylation by Tet homologues, correlation with H3K4-methylation.

Figure 3 Modified DNA bases and their functions.

the recruitment of DNA repair-associated complexes and thus enhance demethylation (Figure 3). Finally, these marks might also directly contribute to gene regulation by triggering “scheduled” DNA repair, which has been suggested to be coupled with activated transcription [84].

The discovery of 6mA in eukaryotes recently identified an additional methylation mark (Figure 3). With *C. elegans* and *D. melanogaster*, two species with negligible 5mC/5hmC levels were shown to contain low, but significant genomic 6mA levels. In both species, this novel modification can be cautiously interpreted as an active epigenetic mark, as data from *C. elegans* suggests a functional interplay with an established active histone mark (H3K4me2) [78], whereas in *Drosophila* mutations in the 6mA-demethylase DMAD (a Tet-homologue) caused increased transposon expression [81]. In both organisms mutations in the 6mA-specific enzymes resulted in significant phenotypes (developmental defects, infertility), suggesting important roles in development. Also in *Chlamydomonas*, 6mA marks actively transcribed genes near the transcriptional start site (TSS).

Future research needs to address the conservation of 6mA and the enzymes that can set and remove this modification. Interestingly, the candidate *C. elegans* 6mA methyltransferase Damt-1 belongs to a widely conserved family of enzymes [78] that also includes a human homologue (METTL4). Nevertheless, reports on 6mA in higher eukaryotes have been sparse and the results were often inconclusive [76]. Highly sensitive mass spectrometry detected less than one molecule of 6mA per million nucleotides in DNA from selected mouse tissues [85], suggesting that 6mA is not a constitutive modification, or is rapidly turned over by demethylation processes. It might be possible to enrich 6mA by depleting the 6mA-demethylase, as shown for *Drosophila* [81]. Furthermore, additional enzymes potentially involved in adenine methylation and demethylation in mammals can be identified using genome editing tools. Finally, the observation that 6mA demethylation in *Drosophila* can be mediated by a Tet-like enzyme [81], raises the fascinating possibility that cytosine and adenine (de)methylation are coordinated. It will be most interesting to investigate the potential interplay between specific DNA modifications and to explore the full complexity of this epigenetic code.

Abbreviations

5mC: 5-methylcytosine; 5hmC: 5-hydroxymethylcytosine; 5hmU: 5-hydroxymethyluracil; 5caC: 5-carboxylcytosine; 5fC: 5-formylcytosine; 6mA: N6-methyladenine; ALKB: alpha-ketoglutarate-dependent dioxygenase; CpG: a cytosine followed by a guanine nucleotide; CHG: a cytosine followed by a non-guanine nucleotide followed by a guanine; CHH: a cytosine followed by two non-guanine nucleotides; Dam: deoxyadenosine methylase; Damt: 6mA DNA methyltransferase; Dmad: *Drosophila* DNA 6mA demethylase;

DNA: deoxyribonucleic acid; DNMT: DNA methyltransferase; ESCs: embryonic stem cells; Fox: forkhead box; H3K4me2: dimethylation of lysine 4 at histone H3; MBD: methyl-CpG-binding domain; MECP: methyl-CpG-binding protein; Nmad: N6-methyladenine demethylase; RNA: ribonucleic acid; siRNA: small interfering RNA; SMRT: single molecule real time; Smug: single-strand-selective monofunctional uracil-DNA glycosylase; Spr: suppressor of presenilin; Tet: ten eleven translocation; TSS: transcriptional start site; Uhrf: ubiquitin-like with PHD and ring finger domains.

Authors' contributions

AB and FL designed and wrote the manuscript. Both authors read and approved the final manuscript.

Acknowledgements

We would like to thank Günter Raddatz for help with Figure 2. Our work is supported by grants from the Deutsche Forschungsgemeinschaft.

Compliance with ethical guidelines

Competing interests

The authors declare that they have no competing interests.

Received: 28 May 2015 Accepted: 7 July 2015

Published online: 21 July 2015

References

- Bird A. Perceptions of epigenetics. *Nature*. 2007;447:396–8.
- Kouzarides T. Chromatin modifications and their function. *Cell*. 2007;128(4):693–705.
- Sexton T, Cavalli G. The role of chromosome domains in shaping the functional genome. *Cell*. 2015;160(6):1049–59.
- Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet*. 2013;14(3):204–20.
- Baubec T, Schubeler D. Genomic patterns and context specific interpretation of DNA methylation. *Curr Opin Genet Dev*. 2014;25C:85–92.
- Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci*. 2006;31(2):89–97.
- Xie W, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, et al. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell*. 2013;153(5):1134–48.
- Gifford CA, Ziller MJ, Gu H, Trapnell C, Donaghey J, Tsankov A, et al. Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell*. 2013;153(5):1149–63.
- Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518(7539):317–30.
- Feng S, Jacobsen SE. Epigenetic modifications in plants: an evolutionary perspective. *Curr Opin Plant Biol*. 2011;14(2):179–86.
- Bemer M, Grossniklaus U. Dynamic regulation of Polycomb group activity during plant development. *Curr Opin Plant Biol*. 2012;15(5):523–9.
- Pikaard CS, Mittelsten Scheid O. Epigenetic regulation in plants. *Cold Spring Harb Perspect Biol*. 2014;6(12):a019315.
- Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13(7):484–92.
- Goll MG, Bestor TH. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem*. 2005;74:481–514.
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009;462(7271):315–22.
- Xie W, Barr CL, Kim A, Yue F, Lee AY, Eubanks J, et al. Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell*. 2012;148(4):816–31.
- Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, et al. Global epigenomic reconfiguration during mammalian brain development. *Science*. 2013;341(6146):1237905.
- Guo JU, Su Y, Shin JH, Shin J, Li H, Xie B, et al. Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat Neurosci*. 2014;17(2):215–22.

19. Gabel HW, Kinde B, Stroud H, Gilbert CS, Harmin DA, Kastan NR, et al. Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. *Nature*. 2015;522(7554):89–93.
20. Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, et al. Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci USA*. 2010;107(19):8689–94.
21. Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science*. 2010;328(5980):916–9.
22. Lyko F, Foret S, Kucharski R, Wolf S, Falckenhayn C, Maleszka R. The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol*. 2010;8(11):e1000506.
23. Kucharski R, Maleszka J, Foret S, Maleszka R. Nutritional control of reproductive status in honeybees via DNA methylation. *Science*. 2008;319(5871):1827–30.
24. Raddatz G, Guzzardo PM, Olova N, Fantappie MR, Rampp M, Schaefer M, et al. Dnmt2-dependent methylomes lack defined DNA methylation patterns. *Proc Natl Acad Sci USA*. 2013;110(21):8627–31.
25. Beisel C, Paro R. Silencing chromatin: comparing modes and mechanisms. *Nat Rev Genet*. 2011;12(2):123–35.
26. Lejeune E, Allshire RC. Common ground: small RNA programming and chromatin modifications. *Curr Opin Cell Biol*. 2011;23(3):258–65.
27. Reyes-Turcu FE, Grewal SI. Different means, same end-heterochromatin formation by RNAi and RNAi-independent RNA processing factors in fission yeast. *Curr Opin Genet Dev*. 2012;22(2):156–63.
28. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*. 1992;69:915–26.
29. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*. 1999;99:247–57.
30. Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. *Nature*. 1993;366:362–5.
31. Panning B, Jaenisch R. DNA hypomethylation can activate *Xist* expression and silence X-linked genes. *Genes Dev*. 1996;10:1991–2002.
32. Walsh CP, Chaillet JR, Bestor TH. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet*. 1998;20:116–7.
33. Wu H, Coskun V, Tao J, Xie W, Ge W, Yoshikawa K, et al. Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science*. 2010;329(5990):444–8.
34. Jin B, Ernst J, Tiedemann RL, Xu H, Sureshchandra S, Kellis M, et al. Linking DNA methyltransferases to epigenetic marks and nucleosome structure genome-wide in human tumor cells. *Cell Rep*. 2012;2(5):1411–24.
35. Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, et al. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature*. 2015;520(7546):243–7.
36. Liao J, Karnik R, Gu H, Ziller MJ, Clement K, Tsankov AM, et al. Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat Genet*. 2015;47(5):469–78.
37. Li Z, Dai H, Martos SN, Xu B, Gao Y, Li T, et al. Distinct roles of DNMT1-dependent and DNMT1-independent methylation patterns in the genome of mouse embryonic stem cells. *Genome Biol*. 2015;16(1):115.
38. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009;324(5929):930–5.
39. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*. 2010;466(7310):1129–33.
40. Penn NW, Suwalski R, O'Riley C, Bojanowski K, Yura R. The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem J*. 1972;126(4):781–90.
41. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*. 2009;324(5929):929–30.
42. Globisch D, Munzel M, Muller M, Michalakis S, Wagner M, Koch S, et al. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One*. 2010;5(12):e15367.
43. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*. 2011;333(6047):1303–7.
44. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*. 2011;333:1300–3.
45. Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, et al. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev*. 2011;25(7):679–84.
46. Dawlaty MM, Ganz K, Powell BE, Hu YC, Markoulaki S, Cheng AW, et al. Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell*. 2011;9(2):166–75.
47. Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, et al. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell*. 2011;8(2):200–13.
48. Ko M, Bandukwala HS, An J, Lamperti ED, Thompson EC, Hastie R, et al. Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc Natl Acad Sci USA*. 2011;108(35):14566–71.
49. Li Z, Cai X, Cai CL, Wang J, Zhang W, Petersen BE, et al. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood*. 2011;118(17):4509–18.
50. Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*. 2011;477(7366):606–10.
51. Dawlaty MM, Breiling A, Le T, Raddatz G, Barrasa MI, Cheng AW, et al. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev Cell*. 2013;24(3):310–23.
52. Dawlaty MM, Breiling A, Le T, Barrasa MI, Raddatz G, Gao Q, et al. Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. *Dev Cell*. 2014;29(1):102–11.
53. Hu X, Zhang L, Mao SQ, Li Z, Chen J, Zhang RR, et al. Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. *Cell Stem Cell*. 2014;14(4):512–22.
54. Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y. Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev*. 2014;28(19):2103–19.
55. Hon GC, Song CX, Du T, Jin F, Selvaraj S, Lee AY, et al. 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol Cell*. 2014;56:286–97.
56. Serandour AA, Avner S, Oger F, Bizot M, Percevault F, Lucchetti-Miganeh C, et al. Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers. *Nucleic Acids Res*. 2012;40(17):8255–65.
57. Jin C, Lu Y, Jelinek J, Liang S, Estecio MR, Barton MC, et al. TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells. *Nucleic Acids Res*. 2014;42(11):6956–71.
58. Pastor WA, Aravind L, Rao A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol*. 2013;14(6):341–56.
59. Wu H, Zhang Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev*. 2011;25(23):2436–52.
60. Raiber EA, Beraldi D, Ficiz G, Burgess HE, Branco MR, Murat P, et al. Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. *Genome Biol*. 2012;13(8):R69.
61. Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL, et al. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell*. 2013;153(3):692–706.
62. Song CX, Szulwach KE, Dai Q, Fu Y, Mao SQ, Lin L, et al. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell*. 2013;153(3):678–91.
63. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PW, Bauer C, et al. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell*. 2013;152(5):1146–59.
64. Iurlaro M, Ficiz G, Oxley D, Raiber EA, Bachman M, Booth MJ, et al. A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. *Genome Biol*. 2013;14(10):R119.
65. Yildirim O, Li R, Hung JH, Chen PB, Dong X, Ee LS, et al. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell*. 2011;147(7):1498–510.
66. Mellen M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell*. 2012;151(7):1417–30.

67. Bachman M, Uribe-Lewis S, Yang X, Williams M, Murrell A, Balasubramanian S. 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nature chemistry*. 2014;6(12):1049–55.
68. Song CX, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol*. 2011;29(1):68–72.
69. Szulwach KE, Li X, Li Y, Song CX, Wu H, Dai Q, et al. 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat Neurosci*. 2011;14(12):1607–16.
70. Hahn MA, Qiu R, Wu X, Li AX, Zhang H, Wang J, et al. Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. *Cell Rep*. 2013;3(2):291–300.
71. Bocker MT, Tuorto F, Raddatz G, Musch T, Yang FC, Xu M, et al. Hydroxylation of 5-methylcytosine by TET2 maintains the active state of the mammalian HOXA cluster. *Nat Commun*. 2012;3:818.
72. Etchegaray JP, Chavez L, Huang Y, Ross KN, Choi J, Martinez-Pastor B, et al. The histone deacetylase SIRT6 controls embryonic stem cell fate via TET-mediated production of 5-hydroxymethylcytosine. *Nat Cell Biol*. 2015;17(5):545–57.
73. Pfaffeneder T, Spada F, Wagner M, Brandmayr C, Laube SK, Eisen D, et al. Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat Chem Biol*. 2014;10(7):574–81.
74. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell*. 2011;146(1):67–79.
75. Casadesus J, Low D. Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev MMBR*. 2006;70(3):830–56.
76. Ratel D, Ravanat JL, Berger F, Wion D. N6-methyladenine: the other methylated base of DNA. *BioEssays*. 2006;28(3):309–15.
77. Fu Y, Luo GZ, Chen K, Deng X, Yu M, Han D, et al. N-Methyldeoxyadenosine marks active transcription start sites in *Chlamydomonas*. *Cell*. 2015;161:879–92.
78. Greer EL, Blanco MA, Gu L, Sendinc E, Liu J, Aristizabal-Corralles D, et al. DNA Methylation on N-Adenine in *C. elegans*. *Cell*. 2015;161:868–78.
79. Iyer LM, Abhiman S, Aravind L. Natural history of eukaryotic DNA methylation systems. *Prog Mol Biol Transl Sci*. 2011;101:25–104.
80. Takayama S, Dhahbi J, Roberts A, Mao G, Heo SJ, Pachter L, et al. Genome methylation in *D. melanogaster* is found at specific short motifs and is independent of DNMT2 activity. *Genome Res*. 2014;24(5):821–30.
81. Zhang G, Huang H, Liu D, Cheng Y, Liu X, Zhang W, et al. N-methyladenine DNA modification in *Drosophila*. *Cell*. 2015;161:893–906.
82. Buck-Koehn BA, Defossez PA. On how mammalian transcription factors recognize methylated DNA. *Epigenetics*. 2013;8(2):131–7.
83. Spruijt CG, Vermeulen M. DNA methylation: old dog, new tricks? *Nat Struct Mol Biol*. 2014;21(11):949–54.
84. Fong YW, Cattoglio C, Tjian R. The intertwined roles of transcription and repair proteins. *Mol Cell*. 2013;52(3):291–302.
85. Ratel D, Ravanat JL, Charles MP, Platet N, Breuillaud L, Lunardi J, et al. Undetectable levels of N6-methyl adenine in mouse DNA: cloning and analysis of PRED28, a gene coding for a putative mammalian DNA adenine methyltransferase. *FEBS Lett*. 2006;580(13):3179–84.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

