CellPress
REVIEWS

Review

DNA Methylation: Shared and Divergent Features across Eukaryotes

Robert J. Schmitz,[1] Zachary A. Lewis,[2] and Mary G. Goll[1,*]

Chemical modification of nucleotide bases in DNA provides one mechanism for conveying information in addition to the genetic code. 5-methylcytosine (5mC) represents the most common chemically modified base in eukaryotic genomes. Sometimes referred to simply as DNA methylation, in eukaryotes 5mC is most prevalent at CpG dinucleotides and is frequently associated with transcriptional repression of transposable elements. However, 5mC levels and distributions are variable across phylogenies, and emerging evidence suggests that the functions of DNA methylation may be more diverse and complex than was previously appreciated. We summarize the current understanding of DNA methylation profiles and functions in different eukaryotic lineages.

Highlights

Although DNA methyltransferase content is generally conserved across eukaryotes, there is extensive variation in how this modified base is used for a variety of cellular processes.

The manner in which DNA methyltransferases are recruited to target sequences leads to a diversity of genome-wide DNA methylation patterns across eukaryotes.

Continued exploration of DNA methylation patterns and DNA methyltransferase content in diverse eukaryotic lineages will lead to an expanded understanding of the mechanism by which the modified base functions in genomes.

Expanding Roles of DNA Methylation throughout Eukaryotes

DNA methylation has been the subject of intense investigation for decades. Interest in this modification was stimulated in the 1970s when it was proposed that DNA methylation might be a mechanism for controlling multicellular development, although at the time there was no experimental evidence to support this idea [1,2]. Interest in DNA methylation continued to grow following key findings that 5mC is required in some plant and animal species for proper development, as well as for transposon silencing in plants, animals, and some fungi [3,4]. Additional work uncovered roles for 5mC in mammalian X-chromosome inactivation and mono-allelic expression of imprinted genes in mammals and plants [5–7]. Despite decades of work, however, many key questions about how 5mC is controlled and how this modification functions in eukaryotic genomes remain unanswered. Most early work on DNA methylation was restricted to a handful of model systems, but the emergence of new technologies has facilitated studies of 5mC in diverse organisms and has provided new and surprising insights into the control of DNA methylation and its diverse functions in eukaryotes.

Mechanisms for the Establishment and Maintenance of DNA Methylation

DNA methyltransferase enzymes are responsible for formation of 5mC through the transfer of a methyl group from the cofactor *S*-adenosyl-L-methionine (SAM) to the 5′-position of the cytosine ring in DNA [8]. At the amino acid level, DNA methyltransferases are identified by a series of highly conserved motifs that are associated with catalytic activity [9]. Most DNA methyltransferase enzymes can be categorized into one of two groups. The first group comprises *de novo* DNA methyltransferases, which are primarily responsible for establishing 5mC at previously unmethylated sites. The second group includes proteins that function primarily to maintain already established DNA methylation marks during DNA replication [3,4].

The mechanisms responsible for targeting 5mC establishment to specific sequences are only partially understood. There is evidence to suggest that recruitment by select histone tail modifications, pairing of repetitive sequences, and small RNA pathways can all be involved in guiding the establishment of 5mC [10–12]. In addition, transcription factor binding provides a potent mechanism for shaping the global landscape of 5mC establishment through occlusion of potential DNA methyltransferase target sequences [13–16]. The relative importance of these targeting approaches appears to vary between species, and different mechanisms may be used to direct 5mC to different sequences within the same species.

The process of DNA replication presents a challenge for the propagation of methylation states through mitosis. In some cases this problem may be solved through the *de novo* reestablishment of 5mC at target sites during each round of cell division [4,17]. More commonly, maintenance

[1]Department of Genetics, University of Georgia, Athens, GA 30602, USA

[2]Department of Microbiology, University of Georgia, Athens, GA 30602, USA

*Correspondence: Mary.Goll@uga.edu

 https://doi.org/10.1016/j.tig.2019.07.007

mechanisms promote preservation of 5mC patterns at symmetrical CpG sites following replication (Figure 1). Propagation of methylation at CpGs is achieved by recruitment of maintenance methyltransferases to hemimethylated CpG sites at replication forks [18–20]. This recruitment in turn drives methylation of reciprocal unmethylated cytosines in the newly synthesized DNA. The high fidelity of this maintenance mechanism may explain why methylation of CpGs predominates over other dinucleotide contexts in most species with methylated genomes.

Methods To Detect DNA Methylation

Bisulfite sequencing represents the current gold-standard method for detection of 5mC in DNA. In this approach, treatment with sodium bisulfite preferentially deaminates unmethylated cytosines in DNA [21,22]. Deaminated cytosines are subsequently converted to uracil through desulfonation and are replaced by thymines during PCR amplification. Methylated cytosines are protected from the bisulfite reaction, allowing their detection through sequencing of the converted, amplified DNA. In the past decade the coupling of the bisulfite reaction to high-throughput sequencing has made it possible to map genome-wide cytosine methylation states at single-base resolution for any species that has a publicly available reference genome [23,24], and new innovations are extending the power of these approaches to species that lack reference genomes [25]. Although powerful, these types of whole-genome bisulfite sequencing (WGBS) approaches have some limitations. For example, vertebrate genomes also harbor low levels of the modified base 5-hydroxymethylcytosine, which cannot be distinguished from 5mC using standard bisulfite sequencing approaches. It is also important to recognize that, in genomes with very low levels of 5mC or extensive 5mC in non-CpG contexts, it may be difficult to distinguish background levels of bisulfite nonconversion from true methylation events. Controls including known methylated and unmethylated DNA standards can provide useful context in these cases.

Diversity of Methylation Profiles and Functions across Eukaryotic Genomes

Traditional views of 5mC distribution and function in eukaryotes have been heavily influenced by early analysis of a few species including humans, mouse, the flowering plant *Arabidopsis thaliana*, and the filamentous fungus *Neurospora crassa*. Today, WGBS data are available for more than 150 eukaryotic genomes. This wealth of new sequencing data has revealed more extensive taxonomic diversity among methylomes and methyltransferase enzymes than was previously appreciated. Below we summarize our current understanding of 5mC distribution and function in different eukaryotic lineages and discuss similarities and differences across species.

Vertebrates

Vertebrate genomes are extensively methylated, and 5mC is detected at >70% of CpGs in somatic tissues [26]. Low levels of non-CpG methylation have also been reported in some cellular contexts, most prominently in neurons and embryonic stem cells [27–29]. At the sequence level, transposable elements and satellite repeats near centromeres and telomeres are commonly cited as being highly enriched in 5mC in vertebrate genomes. However, it may be more accurate to describe vertebrate genomes as being methylated at CpGs in all sequences, with two types of exceptions. The first exception is nonmethylated islands (NMIs). Located near gene promoters, NMIs represent the only sequence class that consistently escapes DNA methylation in vertebrate species [30]. In mammals, NMIs are often referred to as CpG islands owing to their high CpG density, whereas in other vertebrates NMIs may exhibit significantly lower CpG densities [30,31]. In general, a relatively small fraction of CpG dinucleotides in vertebrate genomes exhibit dynamic changes in 5mC levels between different tissues and developmental stages [32]. These differentially methylated regions (DMRs) often include binding sites for transcription factors, and their hypomethylation correlates with active transcription of nearby genes.

Cytosine methylation in vertebrates is achieved through a combination of *de novo* and maintenance methyltransferases (Figure 2). Most vertebrate genomes encode a single maintenance DNA methyltransferase of the Dnmt1 family and a variable number of *de novo* DNA methyltransferases of the Dnmt3 family. In mammals, two Dnmt3 proteins, Dnmt3a and Dnmt3b, cooperate to establish the bulk of 5mC [33]. An additional Dnmt3 gene, *Dnmt3c*, is specifically found in rodents, and is important for the methylation of young retrotransposons in the male germ line [34,35]. In addition to these

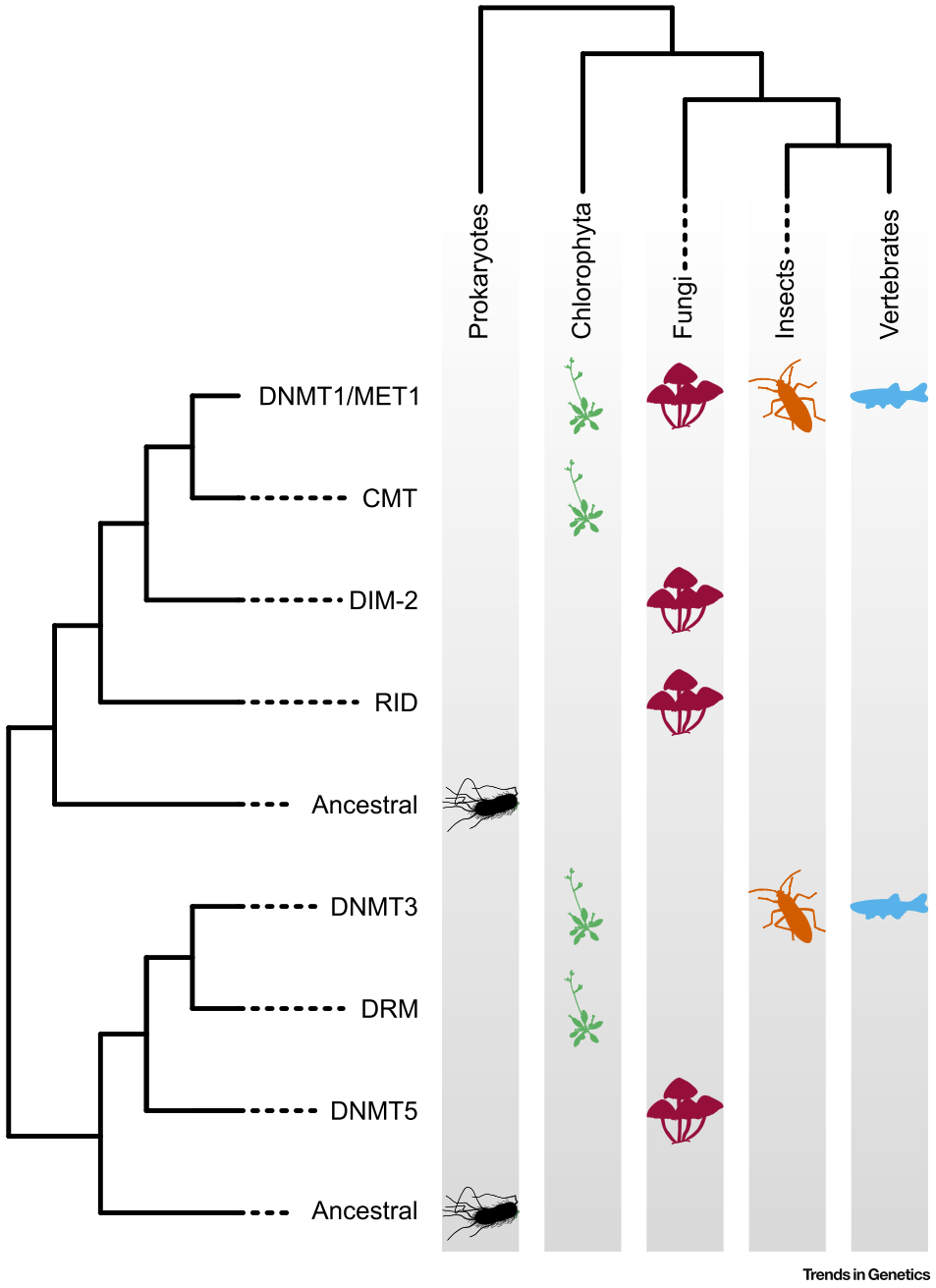

Figure 1. Evolutionary Relationship of Eukaryotic DNA Methyltransferases.
DNMT1 homologs are found in essentially all eukaryotes that utilize 5-methylcytosine (5mC), whereas lineage-specific losses and gains of DNA methyltransferases (DNMTs) are found in specific taxa. This phylogeny is a representation and is not applicable to all species within each lineage owing to recurrent loss of the DNA methylation machinery. Figure courtesy of Adam Bewick.

catalytically active *de novo* methyltransferases, mammalian genomes also encode a stimulatory cofactor, DNMT3L, which shares homology to other Dnmt3 proteins but lacks key catalytic motifs. This cofactor is critical for *de novo* methylation in the germline [36,37]. Other non-mammalian vertebrate lineages appear to lack Dnmt3L orthologs, but exhibit larger expansions of the Dnmt3 *de novo*

methyltransferase family. For example, the zebrafish genome contains six *dnmt3* genes [38]. The reasons for this expansion remain unclear. The mechanisms that target vertebrate Dnmt3 enzymes to particular sequences are not completely understood. However, pathways associated with small RNAs called piRNAs have been implicated in establishing 5mC at transposons in germ cells [39]. Histone tail modifications have also been identified as potential modulators of 5mC deposition. For example, Dnmt3 proteins can directly bind to histone H3 trimethylated on lysine 36 (H3K36me3) to facilitate *de novo* methylation [10,40]. Recent evidence also suggests that methylation may actually be broadly targeted across vertebrate genomes, with protection by bound transcription factors serving as a dominant force driving hypomethylation at DMRs [13–16].

Mutation or inhibition of the maintenance DNA methyltransferase Dnmt1 leads to global loss of 5mC and embryonic lethality in all vertebrates tested to date [41–43]. Global loss of 5mC is associated with significant derepression of transposable elements, suggesting that a major function of the vertebrate DNA methyltransferase machinery is to suppress transcription from these parasitic elements [44]. In addition to roles in controlling transposon expression, 5mC has long been hypothesized to be a key regulator of tissue-specific gene expression. However, although altered gene expression profiles at autosomal, biallelically expressed genes have been noted in vertebrate genomes following 5mC depletion, expression of most genes remains unaffected and not all observed expression changes can be easily attributed to methylation changes at corresponding DMRs [45–47]. Where DNA methylation changes impact on transcription, these changes are likely mediated by sequence-specific DNA-binding proteins that preferentially bind to either methylated or unmethylated recognition sites [13]. In mammals, 5mC is also important for monoallelic expression from imprinted genes, where high levels of 5mC typically accumulate on the silenced allele [5,6]. Similarly, the silent X chromosome of mammalian females is associated with high levels of DNA methylation [7]. Global reprogramming of 5mC patterns is observed in the mammalian germ line and the early embryo. By contrast, the zebrafish methylome does not appear to undergo similar widespread global demethylation during embryogenesis [48,49].

Insects

To date, methylomes of more than 40 insect species have been reported, including representatives from at least six different orders [26,50–61]. 5mC is not detected in all insect genomes, and 5mC and the DNA methyltransferases that mediate this mark seem to have been independently lost multiple times in the insect lineage. Most notably, 5mC appears to be absent from the genomes of dipteran insects [50] including the popular laboratory model *Drosophila melanogaster*. When methylation is

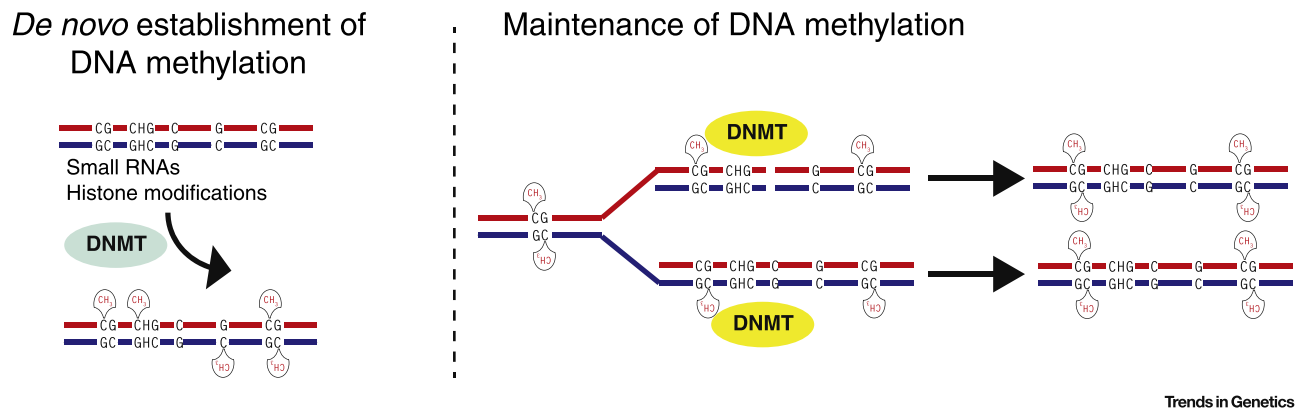

Figure 2. Schematic Diagram of *De Novo* and Maintenance DNA Methylation.
Methylation of cytosines *de novo* is independent of existing 5-methylcytosine (5mC) and can be targeted to cytosines in CpG and non-CpG contexts. Recruitment of *de novo* cytosine methyltransferases is regulated by small RNAs and/or specific chromatin modifications. Maintenance methylation involves recognition of hemi-methylated CpG sites generated during DNA replication. Maintenance methyltransferases target CpG sequences on the newly synthesized strand to generate a fully methylated site.

present in insects, it predominates in CpG contexts. However, the levels and localization of CpG methylation are distinct from vertebrates. Methylation levels in insects are typically much lower than in vertebrate genomes because most insect genomes exhibit methylation at <15% of CpGs [50]. 5mC is consistently enriched in exon sequences of expressed insect genes, and genes that possess housekeeping functions are most likely to be methylated [56,57,62]. Methylation of repetitive sequences is highly variable in the insect lineage. Transposons and other repetitive sequences are not primary targets of DNA methyltransferases in holometabolous insects such as the honey bee, *Apis mellifera*, and the jewel wasp, *Nasonia vitripennis*, whereas similar repeats are often methylated in hemimetabolous insects such as the milkweed bug, *Oncopeltus fasciatus*, and the termite, *Zootermopsis nevadensis* [51,57]. The mechanistic basis for this distinction is unknown, but it likely involves how DNA methyltransferases are recruited to their targets.

Although 5mC distributions vary significantly between vertebrates and insects, the methyltransferase enzymes that mediate this modification are closely related in both phylogenies. Insect species with 5mC typically encode one or more maintenance DNA methyltransferases of the Dnmt1 family and at least one *de novo* methyltransferase with high similarity to vertebrate Dnmt3 proteins (Figure 2) [50]. Duplication of Dnmt1 has occurred in some families within the Hymenoptera and other lineage-specific duplications are reported [50]. Curious exceptions have been noted in which only the *Dnmt1* or *Dnmt3* methyltransferase genes can be detected in a given insect genome. This observation raises the possibility that, in some insects, DNA methyltransferase enzymes of the Dnmt1 or Dnmt3 family may have developed the capacity to efficiently perform both *de novo* and maintenance functions. Alternatively, in some cases the loss of Dnmt1 or Dnmt3 orthologs may reflect an intermediate stage in the lineage-specific decay or adaptation of the DNA methylation machinery.

To date there have been few studies addressing the functional roles of DNA methylation and DNA methyltransferases in insects. The lack of tractable insect species for reverse genetics that also harbor 5mC has represented one challenge to such exploration. However, new studies have begun to fill this void. Knockdown of *Dnmt1* in the milkweed bug, *O. fasciatus*, was recently used to successfully reduce genome-wide 5mC levels in ovary tissues, providing an experimental framework for assessing function in an insect genome [51]. Affected females produced only limited numbers of poor-quality eggs, which developed abnormally when fertilized. However, loss of methylation within transposable elements or genes did not appear to significantly affect their expression [51]. Intriguingly, a similar developmental arrest of progeny was also noted following maternal depletion of Dnmt1 in the red flour beetle *Tribolium castaneum*, even though the *Tribolium* genome appears to lack 5mC [63]. Together, these two experiments raise the possibility that Dnmt1 may have DNA methylation-independent functions in at least some insect species. Future reverse genetic studies of 5mC pathway components will enable testing of the hypothesized effects of DNA methylation on gene regulation in insect genomes.

Plants

5mC has been found in all plant species examined. In plants, DNA methylation is primarily found at transposons and other repetitive sequences. In some angiosperm species, DNA methylation at CpG sites is also common in exons of genes that are moderately and broadly expressed [64–66]. This type of methylation is referred to as gene body methylation (gbM) and is not associated with silencing of gene expression [67,68]. Although its function, if any, is unknown, angiosperm gene body methylation shares similarities with the exonic DNA methylation of highly conserved housekeeping genes noted in insects [69,70]. Curiously, gbM has been lost from the genomes of some angiosperm species, further adding to the mysterious nature of this subclass of methylated loci [71].

The composition of DNA methyltransferases in plants is somewhat similar to that of animals, in that there are maintenance and *de novo* enzymes (Figure 2). However, there are also features of the plant DNA methylation machinery that are unique. DNA methyltransferases of the MET1 family, which are orthologous with Dnmt1 methyltransferases found in animals, are responsible for maintaining methylation at CpG sites in all plant species examined. *De novo* methyltransferase proteins with strong

similarity to animal Dnmt3 enzymes can be detected in the genomes of some early land plants such as mosses [72], but clear Dnmt3 orthologs are absent from the majority of plant species studied to date. Instead, most plants encode domain-rearranged methyltransferases of the DRM2 family. Although related to Dnmt3 proteins, the motifs important for methyltransferase activity are rearranged in DRM2 proteins compared with Dnmt3 enzymes [73,74]. In addition, unlike Dnmt3 proteins, DRM2 methyltransferases appear similarly effective in catalyzing *de novo* methylation at CpHpH sites (H = A, C, or T) with increased activity at CpG, CpT, and CpA sites [74]. In contrast to mammals, targeting of this unique family of *de novo* DNA methyltransferases appears to be mediated almost exclusively through small RNAs via the RNA-directed DNA methylation pathway [4]. These small RNAs are typically 24 nt in length and are generated by the repetitive sequences that are ultimately methylated by these enzymes. Plant genomes also possess DNA methyltransferases, called chromomethylases (CMTs), which are mostly responsible for methylating CpHpG sites in repetitive DNA [75]. To date, these methyltransferases have not been found outside the plant kingdom. CMTs in angiosperm species are recruited to sequences by H3 histone tail methylation at lysine 9 (H3K9me2) [76,77]. Following recruitment, CMT-dependent methylation is subsequently able to recruit the enzymes that mediate H3K9 methylation at these same sequences [78]. In this way, CMT enzymes participate in a feedforward loop that allows long-term propagation of both non-CpG and histone H3K9 methylation at target loci.

The majority of our knowledge regarding the function of DNA methylation in plants derives from studies on two angiosperm species, *Arabidopsis thaliana* and *Zea mays*. Somewhat remarkably, viable *A. thaliana* plants with significant depletion of methylation in CpG or CpHpG contexts can be isolated, although they do not develop normally. By contrast, significant depletion of 5mC is lethal in *Z. mays* [79,80]. In both species, decreases in 5mC levels are associated with increased expression of transposable elements and aberrant expression of some gene loci. However, there is currently only limited understanding of how these changes drive the development abnormalities observed in 5mC-depleted plants. DNA methylation is also involved in genome imprinting in the endosperm of flowering plants [81], although it should be noted that imprinting evolved independently in flowering plants and vertebrates despite mechanistic similarities in both groups.

Fungi

DNA methylation is absent from the genomes of several extensively studied fungi including *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Aspergillus nidulans*. However, a recent broad survey of DNA methyltransferases and DNA methylomes from representative fungal species indicates that 5mC is more widespread than was previously appreciated in this kingdom, and highest levels of 5mC are detected among basidiomycetes [82]. Although gene body methylation is a common feature of animal and plant genomes, it is notably absent in the vast majority of fungal genomes, and instead methylation is localized primarily to repeat sequences [82]. In most fungal species where 5mC is observed, methylation of cytosines in all dinucleotide contexts can be detected. Four classes of fungal DNA methyltransferases have been observed: DNMT1, DNMT5, DIM-2, and RID (Figure 2). Fungal Dnmt1 orthologs share strong homology to Dnmt1/MET1 enzymes from plants and animals, suggesting extensive conservation of this maintenance machinery across eukaryotes. Phylogenetic evidence would suggest that DNMT5 proteins are likely maintenance enzymes. There is evidence to suggest that DIM-2 and RID can function as *de novo* methyltransferases, although the methylation capacities of RID homologs are not well defined [83,84]. As a general rule, ascomycete fungi encode homologs of DIM-2, whereas basidiomycetes encode DNMT1 and DNMT5 homologs, but all possible combinations of DNMTs are observed, presumably due to horizontal transfer of DNMT genes between fungi [82].

To date, most functional studies of DNA methylation in fungi have been performed using the ascomycete fungus *N. crassa*. Most 5mC in *N. crassa* is associated with repeat sequences that have been altered by a homology-based genome defense system called repeat-induced point mutation (RIP) [85]. RIP was first identified in *N. crassa* and involves DNA methylation and subsequent mutation of duplicated sequences in DNA [86,87]. RIP has been reported to occur in other ascomycete fungi

[88–93], and a RIP-like process called MIP (methylation induced premeiotically) methylates but does not mutate repeats during sexual development in the ascomycete *Ascobolus immersus* [84,94,95]. Interestingly, homologs of RID are required for sexual development in *A. immersus* and *Aspergillus nidulans*, but the precise role of these proteins in meiotic tissues is not understood [84,96].

In *N. crassa*, 5mC in is not limited to symmetrical sites, and a mechanism for maintenance methylation does not appear to exist, instead it is likely that 5mC is reestablished *de novo* at every round of cell division [97–102]. This *de novo* targeting is mediated primarily by H3K9 methylation at AT-rich targets [99,101,103], although antisense transcripts have been implicated in targeting 5mC to specific genes that are GC-rich and present as single-copies [104]. In *N. crassa*, loss of 5mC is not associated with significant gene expression changes, but 5mC is necessary to restrict the mobility of functional transposons [105]. 5mC was shown to inhibit transcriptional elongation in *N. crassa*, but the mechanism is not understood [106].

Concluding Remarks and Future Perspectives

The ever-increasing availability of eukaryotic methylomes allows a more holistic analysis of DNA methylation in eukaryotes than could be performed only a decade ago. Nevertheless, there are many lineages that remain under- or unexplored. For example, analysis of 5mC in the parasitic nematode *Trichinella spiralis* has recently challenged the longstanding assumption that the nematode lineage lacks this modification [107]. In the ciliate *Oxytricha trifallx*, DNA methylation is likely required for DNA elimination independently of gene regulation [108]. Many other protostome lineages have been similarly neglected, and almost nothing is known regarding methylation in cnidaria, porifera, and entire groups of protists. These omissions suggest that there is still much to learn regarding the diversity of methylomes and DNA methyltransferase systems in eukaryotes.

There are numerous unresolved issues in this field (see Outstanding Questions). Why there is such outstanding natural diversity in eukaryotic DNA methyltransferases and 5mC distributions remains an open question. The loss of a subset of methylation pathways in addition to the complete loss of methylation in some lineages emphasizes that 5mC is not absolutely required for gene regulation, transposon control, or survival in eukaryotes, suggesting that compensatory pathways can support these functions in some contexts. Nevertheless, the essential nature of 5mC in vertebrates and *Z. mays* argues that in other organisms such redundancies are lacking or are insufficient to compensate for newly evolved 5mC requirements. Specialist functions for 5mC, such as imprinted gene regulation in mammals and in the endosperm of flowering plants, repeat-induced point mutation in *N. crassa*, and gene body methylation in insects and angiosperms, suggest that the base functions of 5mC can readily be usurped to support unique requirements in different lineages. Indeed, the lack of 5mC in repetitive sequences of some insect species suggests that common functions for 5mC in transposon control may not always be retained as the predominant 5mC function in species with 5mC. There is clearly much more to learn about 5mC in eukaryotes. However, new methylome data underscore the risk of broadly extrapolating findings regarding DNA methylation and DNA methyltransferase function from one species to the next.

Outstanding Questions

Is there a specialized role for DNA methylation during meiosis? In mammals, disruption of DNA methylation leads to failed meiosis. Similarly, knockdown of *Dnmt1* in the milkweed bug, *Oncopeltus fasciatus*, caused females to cease laying eggs, and eggs that were produced failed to develop. In fungi, homologs of RID are required for sexual development in two ascomycetes, *Ascobolus immersus* and *Aspergillus nidulans*.

Is there a function of gene body DNA methylation in plants and insects? Gene body DNA methylation is found within moderately and constitutively expressed housekeeping genes in flowering plants and insects. The modification is often evolutionary conserved, but a well-defined function remains to be determined.

Why are DNA methyltransferases and DNA methylation lost in numerous independent insect taxa? Curiously, some insect species have only lost Dnmt1 or Dnmt3, but still retain DNA methylation. These findings suggest that additional undiscovered enzymatic activities of DNA methyltransferases might exist or that compensatory mechanisms have evolved to allow loss of DNA methylation.

Does DNA methylation regulate gene expression beyond transposon silencing in fungi? In *N. crassa*, a small number of protein-coding genes are associated with DNA methylation directed by a class of small RNAs, and, in several basidiomycetes, large methylated domains found on chromosome arms span multiple protein-coding genes. These observations raise the possibility that 5mC could regulate gene expression during development or in response to changing environmental conditions.

Acknowledgments

This work was supported by the National Institutes of Health (R01GM110092) to M.G.G and (R01GM132644) to Z.A.L., as well as by the National Science Foundation (MCB-1856143), the Technical University of Munich–Institute for Advanced Study funded by the Excellence Initiative, and the EU Seventh Framework Programme (grant 291763) to R.J.S. R.J.S. is a Pew Scholar in the Biomedical Sciences, supported by The Pew Charitable Trust.

References

1. Holliday, R. and Pugh, J.E. (1975) DNA modification mechanisms and gene activity during development. *Science* 187, 226–232
2. Riggs, A.D. (1975) X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* 14, 9–25

3. Goll, M.G. and Bestor, T.H. (2005) Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* 74, 481–514
4. Law, J.A. and Jacobsen, S.E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11, 204–220
5. Elhamamsy, A.R. (2017) Role of DNA methylation in imprinting disorders: an updated review. *J. Assist. Reprod. Genet.* 34, 549–562
6. Satyaki, P.R. and Gehring, M. (2017) DNA methylation and imprinting in plants: machinery and mechanisms. *Crit. Rev. Biochem. Mol. Biol.* 52, 163–175
7. Disteche, C.M. and Berletch, J.B. (2015) X-chromosome inactivation and escape. *J. Genet.* 94, 591–599
8. Chen, L. *et al.* (1991) Direct identification of the active-site nucleophile in a DNA (cytosine-5)-methyltransferase. *Biochemistry* 30, 11018–11025
9. Posfai, J. *et al.* (1989) Predictive motifs derived from cytosine methyltransferases. *Nucleic Acids Res.* 17, 2421–2435
10. Dhayalan, A. *et al.* (2010) The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J. Biol. Chem.* 285, 26114–26120
11. Gladyshev, E. and Kleckner, N. (2017) DNA sequence homology induces cytosine-to-thymine mutation by a heterochromatin-related pathway in *Neurospora*. *Nat. Genet.* 49, 887–894
12. Matzke, M.A. and Mosher, R.A. (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 15, 394–408
13. Zhu, H. *et al.* (2016) Transcription factors as readers and effectors of DNA methylation. *Nat. Rev. Genet.* 17, 551–565
14. Domcke, S. *et al.* (2015) Competition between DNA methylation and transcription factors determines binding of NRF1. *Nature* 528, 575–579
15. Stadler, M.B. *et al.* (2011) DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* 480, 490–495
16. Krebs, A.R. *et al.* (2014) High-throughput engineering of a mammalian genome reveals building principles of methylation states at CG rich regions. *eLife* 3, e04094
17. Chen, T. *et al.* (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell. Biol.* 23, 5594–5605
18. Bostick, M. *et al.* (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317, 1760–1764
19. Woo, H.R. *et al.* (2007) VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. *Genes Dev.* 21, 267–277
20. Sharif, J. *et al.* (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 450, 908–912
21. Frommer, M. *et al.* (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. U. S. A.* 89, 1827–1831
22. Clark, S.J. *et al.* (1994) High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* 22, 2990–2997
23. Cokus, S.J. *et al.* (2008) Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 452, 215–219
24. Lister, R. *et al.* (2008) Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* 133, 523–536
25. Bewick, A.J. *et al.* (2015) FASTmC: a suite of predictive models for nonreference-based estimations of DNA methylation. *G3* 6, 447–452
26. Feng, S. *et al.* (2010) Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. U. S. A.* 107, 8689–8694
27. Ziller, M.J. *et al.* (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet.* 7, e1002389
28. Guo, J.U. *et al.* (2014) Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat. Neurosci.* 17, 215–222
29. He, Y. and Ecker, J.R. (2015) Non-CG methylation in the human genome. *Annu. Rev. Genomics Hum. Genet.* 16, 55–77
30. Long, H.K. *et al.* (2013) Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. *Elife* 2, e00348
31. Bird, A. (1987) CpG islands as gene markers in the vertebrate nucleus. *Trends Genet.* 3, 342–347
32. Luo, C. *et al.* (2018) Dynamic DNA methylation: In the right place at the right time. *Science* 361, 1336–1340
33. Okano, M. *et al.* (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247–257
34. Jain, D. *et al.* (2017) *rahu* is a mutant allele of *Dnmt3c*, encoding a DNA methyltransferase homolog required for meiosis and transposon repression in the mouse male germline. *PLoS Genet.* 13, e1006964
35. Barau, J. *et al.* (2016) The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science* 354, 909–912
36. Bourc'his, D. *et al.* (2001) Dnmt3L and the establishment of maternal genomic imprints. *Science* 294, 2536–2539
37. Bourc'his, D. and Bestor, T.H. (2004) Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* 431, 96–99
38. Goll, M.G. and Halpern, M.E. (2011) DNA methylation in zebrafish. *Prog. Mol. Biol. Transl. Sci.* 101, 193–218
39. Aravin, A.A. *et al.* (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell* 31, 785–799
40. Baubec, T. *et al.* (2015) Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* 520, 243–247
41. Anderson, R.M. *et al.* (2009) Loss of Dnmt1 catalytic activity reveals multiple roles for DNA methylation during pancreas development and regeneration. *Dev. Biol.* 334, 213–223
42. Stancheva, I. and Meehan, R.R. (2000) Transient depletion of xDnmt1 leads to premature gene activation in *Xenopus* embryos. *Genes Dev.* 14, 313–327
43. Li, E. *et al.* (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915–926
44. Yoder, J.A. *et al.* (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* 13, 335–340
45. Jackson-Grusby, L. *et al.* (2001) Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat. Genet.* 27, 31–39
46. Chernyavskaya, Y. *et al.* (2017) Loss of DNA methylation in zebrafish embryos activates retrotransposons to trigger antiviral signaling. *Development* 144, 2925–2939
47. Gutierrez-Arcelus, M. *et al.* (2013) Passive and active DNA methylation and the interplay with genetic variation in gene regulation. *eLife* 2, e00523

48. Potok, M.E. *et al.* (2013) Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* 153, 759–772
49. Jiang, Y. *et al.* (2013) Genome-wide distribution of DNA methylation and DNA demethylation and related chromatin regulators in cancer. *Biochim. Biophys. Acta* 1835, 155–163
50. Bewick, A.J. *et al.* (2017) Evolution of DNA methylation across insects. *Mol. Biol. Evol.* 34, 654–665
51. Bewick, A.J. *et al.* (2019) Dnmt1 is essential for egg production and embryo viability in the large milkweed bug, *Oncopeltus fasciatus*. *Epigenetics Chromatin* 12, 6
52. Bonasio, R. *et al.* (2012) Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Curr. Biol.* 22, 1755–1764
53. Libbrecht, R. *et al.* (2016) Robust DNA methylation in the clonal raider ant brain. *Curr. Biol.* 26, 391–395
54. Glastad, K.M. *et al.* (2017) Variation in DNA methylation is not consistently reflected by sociality in hymenoptera. *Genome Biol. Evol.* 9, 1687–1698
55. Standage, D.S. *et al.* (2016) Genome, transcriptome and methylome sequencing of a primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social insect. *Mol. Ecol.* 25, 1769–1784
56. Rehan, S.M. *et al.* (2016) The genome and methylome of a subsocial small carpenter bee, *Ceratina calcarata*. *Genome Biol. Evol.* 8, 1401–1410
57. Glastad, K.M. *et al.* (2016) The caste- and sex-specific DNA methylome of the termite *Zootermopsis nevadensis*. *Sci. Rep.* 6, 37110
58. Xiang, H. *et al.* (2010) Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. *Nat. Biotechnol.* 28, 516–520
59. Wang, X. *et al.* (2013) Function and evolution of DNA methylation in *Nasonia vitripennis*. *PLoS Genet.* 9, e1003872
60. Cunningham, C.B. *et al.* (2015) The genome and methylome of a beetle with complex social behavior, *Nicrophorus vespilloides* (Coleoptera: Silphidae). *Genome Biol. Evol.* 7, 3383–3396
61. Zemach, A. *et al.* (2010) Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* 328, 916–919
62. Hunt, B.G. *et al.* (2013) The function of intragenic DNA methylation: insights from insect epigenomes. *Integr. Comp. Biol.* 53, 319–328
63. Schulz, N.K.E. *et al.* (2018) Dnmt1 has an essential function despite the absence of CpG DNA methylation in the red flour beetle *Tribolium castaneum*. *Sci. Rep.* 8, 16462
64. Niederhuth, C.E. *et al.* (2016) Widespread natural variation of DNA methylation within angiosperms. *Genome Biol.* 17, 194
65. Bewick, A.J. *et al.* (2017) The evolution of CHROMOMETHYLASES and gene body DNA methylation in plants. *Genome Biol.* 18, 65
66. Bewick, A.J. and Schmitz, R.J. (2017) Gene body DNA methylation in plants. *Curr. Opin. Plant Biol.* 36, 103–110
67. Tran, R.K. *et al.* (2005) DNA methylation profiling identifies CG methylation clusters in *Arabidopsis* genes. *Curr. Biol.* 15, 154–159
68. Zhang, X. *et al.* (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* 126, 1189–1201
69. Takuno, S. and Gaut, B.S. (2012) Body-methylated genes in *Arabidopsis thaliana* are functionally important and evolve slowly. *Mol. Biol. Evol.* 29, 219–227
70. Takuno, S. and Gaut, B.S. (2013) Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. *Proc. Natl. Acad. Sci. U. S. A.* 110, 1797–1802
71. Bewick, A.J. *et al.* (2016) On the origin and evolutionary consequences of gene body DNA methylation. *Proc. Natl. Acad. Sci. U. S. A.* 113, 9111–9116
72. Yaari, R. *et al.* (2019) RdDM-independent de novo and heterochromatin DNA methylation by plant CMT and DNMT3 orthologs. *Nat. Commun.* 10, 1613
73. Cao, X. and Jacobsen, S.E. (2002) Role of the *Arabidopsis* DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr. Biol.* 12, 1138–1144
74. Zhong, X. *et al.* (2014) Molecular mechanism of action of plant DRM de novo DNA methyltransferases. *Cell* 157, 1050–1060
75. Lindroth, A.M. *et al.* (2001) Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science* 292, 2077–2080
76. Lindroth, A.M. *et al.* (2004) Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. *EMBO J.* 23, 4286–4296
77. Du, J. *et al.* (2012) Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. *Cell* 151, 167–180
78. Jackson, J.P. *et al.* (2002) Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416, 556–560
79. Li, Q. *et al.* (2014) Genetic perturbation of the maize methylome. *Plant Cell* 26, 4602–4616
80. Fu, F.F. *et al.* (2018) Loss of RNA-directed DNA methylation in maize vhromomethylase and DDM1-type nucleosome remodeler mutants. *Plant Cell* 30, 1617–1627
81. Gehring, M. and Satyaki, P.R. (2017) Endosperm and Imprinting, Inextricably Linked. *Plant Physiol.* 173, 143–154
82. Bewick, A.J. *et al.* (2019) Diversity of cytosine methylation across the fungal tree of life. *Nat. Ecol. Evol.* 3, 479–490
83. Kouzminova, E.A. and Selker, E.U. (2001) dim-2 encodes a DNA-methyltransferase responsible for all known cytosine methylation in *Neurospora*. *EMBO J.* 20, 4309–4323
84. Malagnac, F. *et al.* (1997) A gene essential for de novo methylation and development in *Ascobolus* reveals a novel type of eukaryotic DNA methyltransferase structure. *Cell* 91, 281–290
85. Lewis, Z.A. *et al.* (2009) Relics of repeat-induced point mutation direct heterochromatin formation in *Neurospora crassa*. *Genome Res.* 19, 427–437
86. Aramayo, R. and Selker, E.U. (2013) *Neurospora crassa*, a model system for epigenetics research. *Cold Spring Harb. Perspect. Biol.* 5, a017921
87. Selker, E.U. *et al.* (1987) Rearrangement of duplicated DNA in specialized cells of *Neurospora*. *Cell* 51, 741–752
88. Goldfarb, M. *et al.* (2016) Evidence of ectopic recombination and a repeat-induced point (RIP) mutation in the genome of *Sclerotinia sclerotiorum*, the agent responsible for white mold. *Genet. Mol. Biol.* 39, 426–430
89. Rouxel, T. *et al.* (2011) Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. *Nat. Commun.* 2, 202

90. Clutterbuck, A.J. (2011) Genomic evidence of repeat-induced point mutation (RIP) in filamentous ascomycetes. *Fungal Genet. Biol.* 48, 306–326
91. Hamann, A. *et al.* (2000) The degenerate DNA transposon Pat and repeat-induced point mutation (RIP) in *Podospora anserina*. *Mol. Gen. Genet.* 263, 1061–1069
92. Nielsen, M.L. *et al.* (2001) A family of DNA repeats in *Aspergillus nidulans* has assimilated degenerated retrotransposons. *Mol. Gen. Genomics.* 265, 883–887
93. Ikeda, K. *et al.* (2002) Repeat-induced point mutation (RIP) in *Magnaporthe grisea*: implications for its sexual cycle in the natural field context. *Mol. Microbiol.* 45, 1355–1364
94. Goyon, C. and Faugeron, G. (1989) Targeted transformation of *Ascobolus* immersus and de novo methylation of the resulting duplicated DNA sequences. *Mol. Cell. Biol.* 9, 2818–2827
95. Rhounim, L. *et al.* (1992) Epimutation of repeated genes in *Ascobolus immersus*. *EMBO J.* 11, 4451–4457
96. Lee, D.W. *et al.* (2008) A cytosine methyltransferase homologue is essential for sexual development in *Aspergillus nidulans*. *PLoS One* 3, e2531
97. Selker, E.U. *et al.* (1987) A portable signal causing faithful DNA methylation de novo in *Neurospora crassa*. *Science* 238, 48–53
98. Miao, V.P. *et al.* (2000) Short TpA-rich segments of the zeta-eta region induce DNA methylation in *Neurospora crassa*. *J. Mol. Biol.* 300, 249–273
99. Tamaru, H. and Selker, E.U. (2001) A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* 414, 277–283
100. Tamaru, H. and Selker, E.U. (2003) Synthesis of signals for de novo DNA methylation in *Neurospora crassa*. *Mol. Cell. Biol.* 23, 2379–2394
101. Freitag, M. *et al.* (2004) HP1 is essential for DNA methylation in *Neurospora*. *Mol. Cell* 13, 427–434
102. Honda, S. and Selker, E.U. (2008) Direct interaction between DNA methyltransferase DIM-2 and HP1 is required for DNA methylation in *Neurospora crassa*. *Mol. Cell. Biol.* 28, 6044–6055
103. Tamaru, H. *et al.* (2003) Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in *Neurospora crassa*. *Nat. Genet.* 34, 75–79
104. Dang, Y. *et al.* (2013) Convergent transcription induces dynamic DNA methylation at disiRNA loci. *PLoS Genet.* 9, e1003761
105. Zhou, Y. *et al.* (2001) DNA methylation inhibits expression and transposition of the *Neurospora* Tad retrotransposon. *Mol. Gen. Genomics.* 265, 748–754
106. Rountree, M.R. and Selker, E.U. (1997) DNA methylation inhibits elongation but not initiation of transcription in *Neurospora crassa*. *Genes Dev.* 11, 2383–2395
107. Gao, F. *et al.* (2012) Differential DNA methylation in discrete developmental stages of the parasitic nematode *Trichinella spiralis*. *Genome Biol.* 13, R100
108. Bracht, J.R. *et al.* (2012) Cytosine methylation and hydroxymethylation mark DNA for elimination in *Oxytricha trifallax*. *Genome Biol.* 13, R99