Current Biology 20, R780–R785, September 14, 2010 ©2010 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2010.07.007

Evolution of Eukaryotic DNA Methylation and the Pursuit of Safer Sex

Minireview

Assaf Zemach and Daniel Zilberman*

Cytosine methylation is an ancient process with conserved enzymology but diverse biological functions that include defense against transposable elements and regulation of gene expression. Here we will discuss the evolution and biological significance of eukaryotic DNA methylation, the likely drivers of that evolution, and major remaining mysteries.

Introduction

The genetic code underpins all of life with almost invariant consistency, but imagine, for a moment, if this were not so. After all, a given codon is not intrinsically better suited to represent leucine than phenylalanine or aspartic acid. What if evolution could break the informational straightjacket, allowing species to tweak the code to their needs? It would be, to say the least, extremely inconvenient. To interpret each genome, the code would have to be cracked anew. We would need to understand how the code has evolved, which features are ancient, which are specific to major lineages, and which commonly fluctuate between species.

DNA methylation may not be as old as the genetic code but is nonetheless exceedingly ancient. Methylation of the fifth carbon of cytosine, the subject of this article, is mediated by the same enzymatic superfamily in bacteria, archaea, and eukaryotes [1]. Like the genetic code, semiconservative inheritance of methylation states of palindromic sites can propagate information through cellular generations [2]. However, the biological meaning of methylated bases is flexible [3]. Considering that mechanistic studies of DNA methylation are confined to a small number of model organisms, uncovering the evolutionary history of this process is required to know which lessons from, for example, the mustard weed *Arabidopsis thaliana* are directly applicable to mammals, which will be useful for distantly related crop plants, and which are esoteric to the genus. Recent advances in sequencing technology have allowed us to read the methylation patterns of entire genomes [4–8]. The quest to decipher the meaning of these patterns is just beginning.

Eukaryotic Methyltransferase Families

Dnmt1 and Dnmt3 are two generally accepted families of functional eukaryotic DNA methyltransferases that predate the divergence of plants and animals [1]. Dnmt1 and the accessory protein UHRF1 mediate methylation of hemimethylated CG dinucleotides following DNA replication, allowing faithful propagation of methylation patterns [1,8]. Because of this functionality, Dnmt1 is generally considered a maintenance methyltransferase. Dnmt1 is the lynchpin of eukaryotic methylation: with the exception of a lineage of ascomycete fungi, all plants, animals and fungi that methylate DNA possess Dnmt1 (Figures 1 and 2).

Dnmt3 enzymes establish methylation of previously unmethylated sequences in plants and animals. Animal Dnmt3s methylate CG sites, while land plant Dnmt3s (called DRMs for Domains Rearranged Methyltransferases because of a rearrangement of the catalytic domain) can methylate cytosine in any context [1]. DRMs are recruited to their sites of action by the RNA interference pathway [2]. Dnmt3 enzymes appear to be more dispensable than Dnmt1. Dnmt3 homologs have not been found in any fungal genome, and Dnmt3 has been lost in some green algae and animal lineages (Figure 2). The green alga *Chlorella* sp. NC64A, the silk moth *Bombyx mori* and zygomycete and basidiomycete fungi have robust Dnmt1-mediated CG methylation without Dnmt3 [7]. In *B. mori* and basidiomycetes, Dnmt1 is the only methyltransferase family, indicating that Dnmt1 can establish as well as maintain DNA methylation, at least in some species [7].

CMT and Dim-2 are Dnmt1-related methyltransferases found in plants and fungi, respectively [9,10]. Both enzymes methylate transposable elements and other repeats, are dependent on methylation of lysine 9 of histone H3, and have acidic carboxy-terminal tails [1,7]. Consistent with the structural and functional similarities, CMT and Dim-2 form a monophyletic group distinct from the Dnmt1 proteins of plants, animals, and fungi (Figure 1), leading us to propose the CMT/Dim-2 enzyme family [7]. Neither CMT-like nor Dim-2-like proteins are present in animals, indicating that this family has been lost early in animal evolution (Figure 2).

Finally, plants, animals and fungi share the highly conserved Dnmt2 proteins [1]. Dnmt2 contains all catalytic motifs expected of a DNA methyltransferase, but shows no such activity *in vitro* [11]. Instead, Dnmt2 specifically and efficiently methylates cytosine 38 of $tRNA^{Asp}$ *in vitro*, and can reestablish this methylation in *A. thaliana*, mouse and fruit fly *Dnmt2*-deficient cells [12]. The sequence around cytosine 38 is conserved among organisms that have Dnmt2, but is diverged in species lacking Dnmt2 [12]. Several studies have put forth evidence for *in vivo* DNA methylation by Dnmt2, most recently in early *Drosophila* embryos [13]. However, whole-genome analysis of fruit fly embryos at the same stage did not reveal significant methylation [7]. While the possibility that Dnmt2 can function as a DNA methyltransferase remains, the preponderance of evidence so far suggests that Dnmt2 is a very specific RNA methyltransferase with no activity on DNA.

DNA Methylation and Genome Defense

DNA methylation has been long-proposed to function as a genomic immune system [14,15]. Invading transposable elements (TEs), once recognized by the host, would be methylated, suppressing transcription and recombination, thereby preventing further replication and preserving genome integrity [1]. There is very strong evidence that plants and fungi use methylation to defend against TEs. TEs are preferentially methylated in every plant and fungus examined so far [7,8]. CMT enzymes methylate TEs almost exclusively,

Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA.
*E-mail: danielz@berkeley.edu

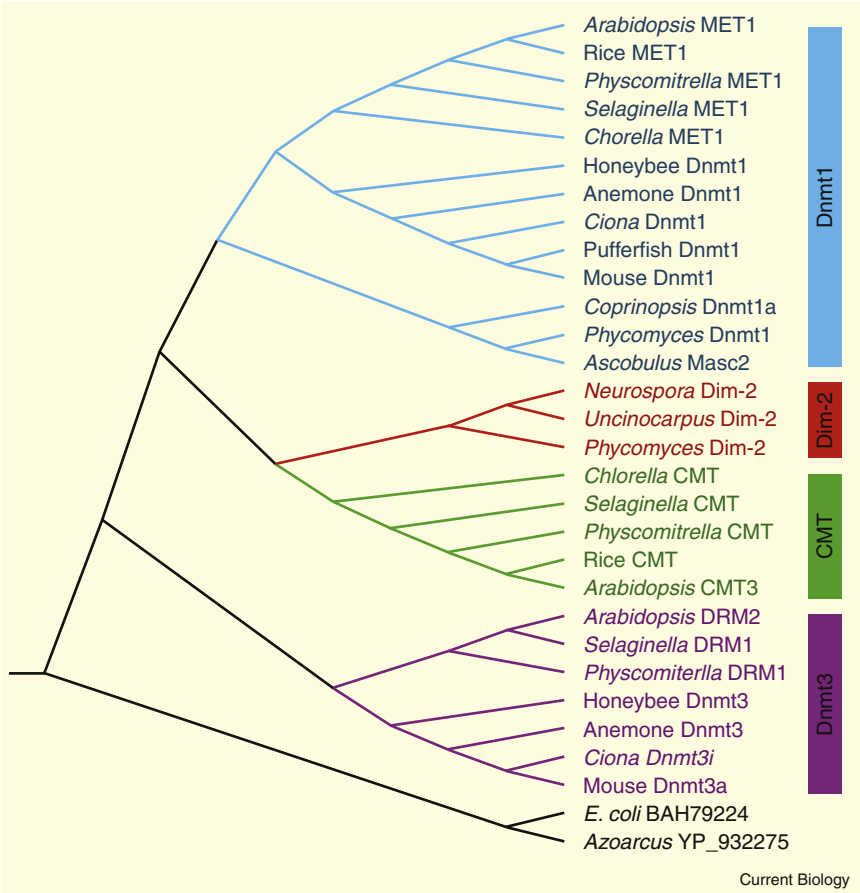

Figure 1. A phylogenetic tree of eukaryotic DNA methyltransferases.

The tree is based on the conserved catalytic domains of Dnmt1, CMT, Dim-2 and Dnmt3 proteins, with bacterial methyltransferases included as an outgroup. Conserved domains were aligned using MUSCLE v3.7 and the phylogenetic tree was inferred using MrBayes v3.1.2 as described [7]. Dnmt1, Dim-2, CMT, Dnmt3 and bacterial proteins are colored blue, red, green, purple and black, respectively.

and loss of DNA methylation causes transcriptional reactivation and transposition of plant TEs [16]. With the exception of a small amount of methylation in *Uncinocarpus reesii* genes, all fungal methylation is located in transcriptionally silent, primarily repetitive loci [7,17]. Ascomycete and basidiomycete fungi methylate and silence repeated sequences as part of the sexual cycle [18]. In the ascomycete *Ascobolus immersus* this methylation is initiated by the methyltransferase Masc1 [19]. Other ascomycete fungi have taken a step further, riddling repeats with C/G > T/A point mutations. This process, termed repeat-induced point mutation (RIP), was first identified and is best studied in *Neurospora crassa,* where it requires the RID protein, a Masc1 homolog [20]. There is much evidence that DNA methylation is mutagenic: deamination of unmethylated cytosine produces uracil, which is efficiently repaired, while deamination of 5-methylcytosine produces thymine. This, and the involvement of RID, has led to the hypothesis that RIP works by methylation-coupled deamination [18].

Masc1/RID methyltransferases have so far been found only in the Pezizomycotina branch of ascomycete fungi [21]. Although two RID-like proteins were suggested to exist in the basidiomycete *Coprinopsis cinerea* [21], a phylogenetic reconstruction shows that these are Dnmt1 proteins [7]. Ascomycete Masc1/RID proteins, though quite diverged, likely evolved from Dnmt1. Some proteins in this family contain a bromo-adjacent homology (BAH) domain, a feature of Dnmt1 [21]. We recently found that mutation of cytosines within repeats of *U. reesii* is restricted (or at least has a very high preference for) CG sites, suggesting that *U. reesii* RID has retained the ancestral Dnmt1 sequence preference [7].

The relationship between DNA methylation and TE defense is much less straightforward in animals. The genomes of vertebrates, even those with very few TEs, exhibit almost blanket methylation [6–8]. This pattern is consistent with methylation being the default state, so that unmethylated sequences are protected from methylation [22]. Nevertheless, puffer fish, zebrafish and mouse TEs are significantly enriched in methylation compared to adjacent regions, indicating that TEs are at least not protected from methylation [7,8], and DNA methylation may be directed to vertebrate TEs by small RNAs [23]. Vertebrate methylation is strongly associated with transcriptional repression, including of TEs [1,22], so DNA methylation appears to be a key mechanism of vertebrate TE defense. Unlike vertebrates, there is no evidence that DNA methylation silences TEs, or for that matter silences anything, in invertebrates [7,8,22]. TEs are not preferentially methylated in the chordate *Ciona intestinalis,* insects, or the sea anemone *Nematostella vectensis* [7,8]. Methylation of invertebrate transcriptional start sites does not negatively correlate with transcription [7]. The decoupling of DNA methylation from TE silencing in invertebrates suggests that the ancestral methylation-dependent TE silencing pathway was lost in early animal evolution, and that the use of methylation for TE defense has evolved independently in the vertebrate lineage (Figure 2).

Gene Bodies Are Ancient Methylation Targets

Although DNA methylation of active genes has long been a known feature of mammalian genomes, the blanket character of vertebrate methylation and the abundance of TEs within introns made it difficult to ascertain the significance of genic methylation [22]. The discovery of gene body methylation in the chordate *C. intestinalis* and the subsequent identification of such methylation in *Arabidopsis* led to the recognition that gene body methylation may be a significant phenomenon [24–27]. By now, extensive methylation of active genes has been found in many animals (humans, mouse, zebrafish, pufferfish, *C. intestinalis*, honeybee, silk moth, and sea anemone) and plants (*Arabidopsis*, poplar, rice, maize, and three species of green algae) [5–8,28,29]. Gene body methylation is almost

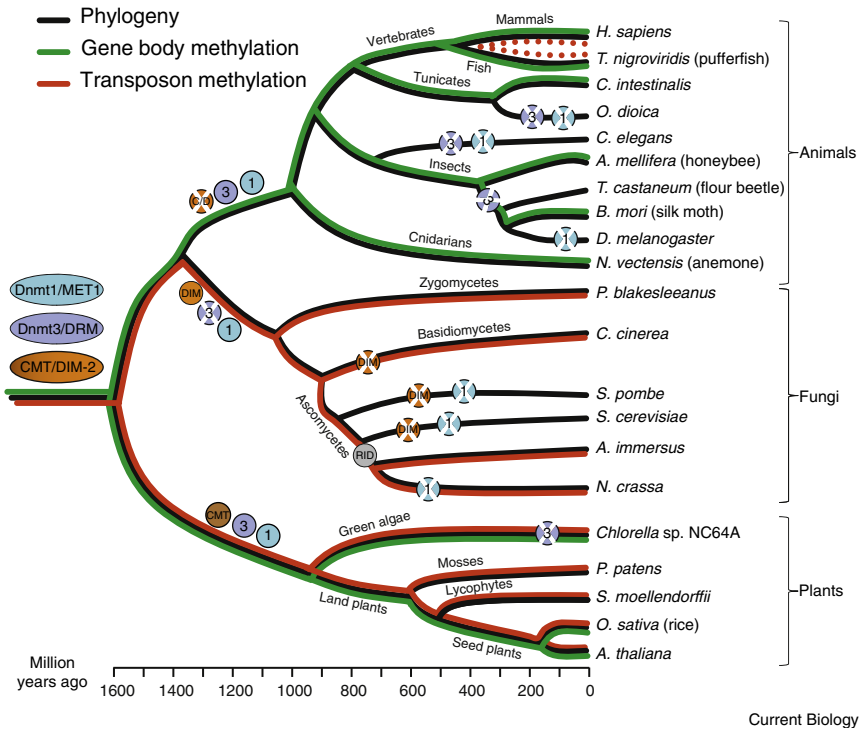

Figure 2. The evolution of eukaryotic DNA methylation.

Circles with 1, 3, CMT, DIM and C/D represent Dnmt1, Dnmt3, CMT, Dim-2 and CMT/Dim-2, respectively. Circles with white Xs represent loss of the indicated gene family. Red and green lines indicate the evolutionary trajectory of TE and gene body methylation, respectively. The dotted red lines represent TE methylation in vertebrates.

exclusively found in the CG context, even in plants that also exhibit methylation in other contexts, a phenomenon explained in part by exclusion of non-CG methylation by the putative IBM1 histone demethylase in *Arabidopsis*, and likely in other plants [30].

Gene body methylation shows a similar pattern in all examined genomes, with methylation excluded near the transcription start and termination sites, and a preference for exons over introns [7,8]. Several mechanisms likely account for this pattern. ROS1-related *Arabidopsis* DNA glycosylases that specifically remove methylated cytosine have a preference for gene ends [5,31]. This mechanism is unlikely to be general because the ROS1 family is found only in plants [32]. A more universal factor may be the histone variant H2A.Z, an ancient eukaryotic protein that preferentially localizes near the transcription start site in all plants, animals and fungi examined so far [33]. DNA methylation in *Arabidopsis* and pufferfish shows a very strong anticorrelation with H2A.Z, and disruption of an *Arabidopsis* chromatin factor that deposits H2A.Z leads to hypermethylation [7,34]. Methylation of lysine 4 of histone H3 (H3K4me), which antagonizes DNA methylation, at least in mammals, and is concentrated around the transcription start site, may also contribute to the pattern of gene body methylation [35,36]. A distinctive feature of flowering plants is that methylation is depleted mostly downstream of the transcription start site, whereas in green algae and animal genes methylation is lowest upstream of this region [7,8]. Depletion of methylation correlates with enrichment of H2A.Z and H3K4me, suggesting that sequences required for gene activity may differ in relation to the transcription start site between flowering plants and other eukaryotes.

Plant and animal species show a characteristic relationship between gene body methylation and transcription: Modestly transcribed genes are most likely to be methylated, while genes at either transcriptional extreme are least likely [7,26,27]. In *Arabidopsis*, there is an inverse correlation between variability of gene expression and methylation, with most methylation found at constitutively expressed genes, and least at tissue-specific or inducible ones [26,37]. It is probable that this is the more relevant correlation, and that it also holds true in other species, but this remains to be tested. Although gene body methylation, as characterized in plants and animals, has not been found in fungi (Figure 2), the ascomycete *U. reesii* does methylate its genes. This methylation occurs in all sequence contexts, has a very strong preference for exons and is directly correlated with gene transcription [7]. The last two features suggest that methylation is either targeted by spliced mRNA or the splicing process itself.

The function of gene body methylation remains a mystery. We have proposed that methylation may repress cryptic intragenic promoters by preventing RNA Polymerase II recruitment, possibly through exclusion of H2A.Z [7,27]. Some organisms, like vertebrates and Chlorella, methylate the vast majority of their genes, so it is likely that genic methylation plays an additional role in these species. Vertebrates have immune receptors that recognize DNA with unmethylated CG sites as foreign, so extensive CG methylation serves as a 'self' signal [22,38]. A similar function might exist in Chlorella, which is infected by a wide variety of DNA viruses, and thus has a strong incentive to differentiate its DNA from that of the pathogen [39]. The preference of methylation for exons [8] is consistent with a role in splicing, as has been recently reported for histone modifications [40]. The best evidence for functional importance of gene body methylation comes from honeybees, where knockdown of Dnmt3 by RNA interference (RNAi) mimics the effect of royal jelly, causing larvae to develop into queens [41]. Given the wide conservation of gene body methylation, understanding its functions is a major outstanding challenge.

DNA Methylation, Transposon Defense, and Development

Proper DNA methylation patterns are crucial for plant and animal development [2]. In part, developmental disruption is the expected consequence of derepressing TEs that are near or within genes, but methylation has also been recruited for regulatory purposes. In *Arabidopsis*, a glycosylase homologous to ROS1, called DEMETER, is expressed in the central cell, which is adjacent to the egg [42]. Plant pollen carries two sperm nuclei, one of which fertilizes the central cell to give rise to endosperm, a terminal placenta-like tissue

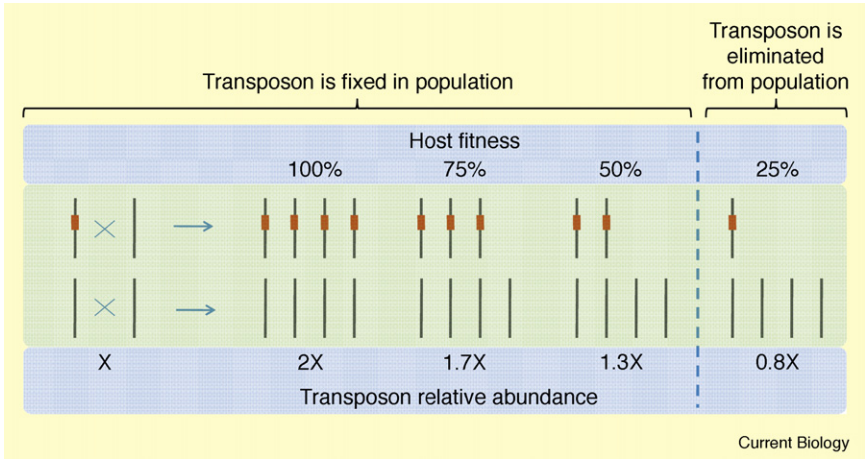

Figure 3. The relationship between TE success and host fitness in sexual outcrossers.

In the middle panel (green), a population starts with four individuals, each represented by a vertical line. TEs are illustrated as red boxes. A cross between individuals lacking the TE in question yields four offspring. Reduction of host fitness caused by the TE manifests in reduced number of offspring. Fitness of TE hosts versus non-carriers (upper blue panel) is shown compared to TE abundance in the filial population relative to the parental population (lower blue panel). A fitness reduction of up to 50% still allows the TE to increase in abundance, leading to eventual fixation.

that nourishes the embryo and makes up the bulk of grain seeds. DEMETER is expressed in the central cell prior to fertilization, causing global demethylation of the maternal genome [43,44]. This results in genomic imprinting — epigenetic differentiation between maternal and paternal genomes — that causes some genes to be preferentially or exclusively expressed either from the maternal or paternal alleles. The imprinted genes themselves do not appear to be targeted for demethylation with any specificity [44]; instead, the likely purpose of this process is reactivation of TEs, a hypothesis supported by a massive burst of small RNAs arising specifically from maternal chromosomes in the endosperm [45]. The small RNAs, which can trigger DNA methylation of cognate sequences, would immunize the egg cell and embryo by reinforcing TE silencing. This process has also been proposed to operate in pollen, where TEs are derepressed in the vegetative nurse cell, and a small RNA expressed in the vegetative cell has been shown to silence a gene in sperm [46]. Removal of methylation from maternal TEs and repeats that are near or within genes can influence gene expression, so that parent-of-origin specific expression is a consequence of TE defense, and new imprinted genes can arise through transposition events or sequence duplications that bring a gene under control of the methylation pathway [43,44].

Regulation of vertebrate gene expression by DNA methylation appears to be substantially more dynamic than in plants. Imprinted expression of mammalian genes is caused by differential DNA methylation established during gametogenesis [47]. DNA methylation patterns differ between cell types, with gamete and embryo-specific genes methylated in adult tissues [48]. Aberrant methylation is a common hallmark of cancer cells, including hypermethylation and silencing of tumor suppressor genes [49]. Like in plants, methylation near the transcription start site of vertebrate genes causes silencing, so the regulatory aspects of vertebrate DNA methylation are also likely an outgrowth of TE defense. As long as methylation has the capacity to repress transcription, it seems evolution will recruit it to do so in genes as well as TEs.

In both plants and animals, methylation primarily represses genes by impeding transcriptional initiation [50]. In *N. crassa*, however, initiation is unaffected, but transcriptional elongation is disrupted [51]. Why the dichotomy? A possibility is that the loss of gene body methylation freed fungi to use methylation to inhibit elongation, while plants and animals that methylate bodies of active genes simply do not have this option.

All About Sex?

The patchy evolutionary distribution of DNA methylation has been a long-standing mystery. Methylation has been lost several times in the course of animal evolution, including in lineages leading to *Drosophila* and *C. elegans*, as well as the chordate *O. dioica* (Figure 2). Methylation is uncommon in fungi, with saccharomycete and schizosaccharomyces ascomycetes, and most basidiomycete and zygomycete species lacking methyltransferase genes [22]. Most species of green algae with sequenced genomes also lack DNA methyltransferases. On the other hand, all examined vertebrates and land plants, lineages that date back at least 500 million years, have extensive DNA methylation. Why should this be?

We recently proposed that the major forces driving the evolution of methylation are TEs and sex [7]. Because TEs, with rare exceptions, cannot move from host to host, they spread in a population through genetic drift, some selective advantage to the host, or through mating of carriers with non-carriers. The first two strategies depend on no or very little harm to the host, but the last one does not. When a carrier mates with a non-carrier, all of the offspring inherit the TE, so a TE that does not harm its host doubles its prevalence in a population in a single generation (Figure 3). The corollary is that a TE can decrease the fitness of its host by as much as half and still become fixed (Figure 3), logic that has even led to the proposal that TEs invented sexual reproduction [52]. TE aggressiveness correlates very well with the extent of sexual outcrossing of the host [53], so while TEs may not have invented sex, they are an evolutionary price we pay for having it [54].

Extant unicellular eukaryotes are primarily asexual, so losing the ability to silence TEs via DNA methylation likely imposes a modest enough evolutionary cost that some lineages can acquire compensatory mutations before going extinct. This would explain the dearth of methylation among fungi and green algae. If today's unicellular eukaryotes are a reliable guide, early unicellular animals also likely relied primarily on asexual reproduction, and thus could retain gene body methylation while losing the ability to use methylation to silence TEs. Extant invertebrate lineages, though primarily sexual, evolved from a primarily asexual state without a link between DNA methylation and TE silencing,

instead using other pathways, such as repressive chromatin mediated by RNAi, to silence TEs [55]. Sexually outcrossing invertebrates like fruit flies could thus lose DNA methylation without compromising TE defense.

Land plants and vertebrates reproduce primarily by sexual outcrossing and use DNA methylation to silence TEs, so they are stuck with methylation: Loss of the system would disrupt gene transcription by activating dormant TEs that frequently reside near or within genes, and lead to a burst of mutagenic transposition, driving the host lineage to extinction. The known differences between plant and vertebrate methylation dynamics may be a consequence of the differing evolutionary trajectories of the methylation system. Plants, which appear to have inherited the use of methylation for TE defense from ancestral eukaryotes [7], do not undergo major fluctuations in methylation except in terminal nurse cells that do not contribute to the next generation [43,44]. Vertebrates, which appear to have reinvented the use of methylation for TE silencing, and thus were building upon invertebrate systems that do not rely on methylation, would have more flexibility, potentially relying on chromatin-based silencing to prevent TE reactivation during the demethylation events that occur in early development of some species [56].

DNA methylation shows remarkable diversity in extent and function across eukaryotic evolution. In our view, the need to mitigate the damage that TEs inflict on sexually reproducing species — a search, if you will, for safer sex — is the main evolutionary force maintaining DNA methylation in eukaryotic genomes. Methylation of gene bodies, whatever its function, appears to be more dispensable, as it has been independently lost in multiple plant and animal lineages. In plants and animals, methylation has also been co-opted to regulate development. Understanding the full scope of DNA methylation functionality, and how the various functions interact in different species, is a challenge for the future.

References

1. Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. Annu. Rev. Biochem. *74*, 481–514.
2. 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.
3. Zilberman, D. (2008). The evolving functions of DNA methylation. Curr. Opin. Plant Biol. *11*, 554–559.
4. Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature *452*, 215–219.
5. Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H., and Ecker, J.R. (2008). Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell *133*, 523–536.
6. Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M., *et al*. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.
7. Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science *328*, 916–919.
8. Feng, S., Cokus, S.J., Zhang, X., Chen, P.Y., Bostick, M., Goll, M.G., Hetzel, J., Jain, J., Strauss, S.H., Halpern, M.E., *et al*. (2010). Conservation and divergence of methylation patterning in plants and animals. Proc. Natl. Acad. Sci. USA *107*, 8689–8694.
9. Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S., and Jacobsen, S.E. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science *292*, 2077–2080.
10. Kouzminova, E., and Selker, E.U. (2001). dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in Neurospora. EMBO J. *20*, 4309–4323.
11. Schaefer, M., and Lyko, F. (2009). Solving the Dnmt2 enigma. Chromosoma *119*, 35–40.
12. Goll, M.G., Kirpekar, F., Maggert, K.A., Yoder, J.A., Hsieh, C.L., Zhang, X., Golic, K.G., Jacobsen, S.E., and Bestor, T.H. (2006). Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science *311*, 395–398.
13. Phalke, S., Nickel, O., Walluscheck, D., Hortig, F., Onorati, M.C., and Reuter, G. (2009). Retrotransposon silencing and telomere integrity in somatic cells of Drosophila depends on the cytosine-5 methyltransferase DNMT2. Nat. Genet. *41*, 696–702.
14. Bestor, T.H. (1990). DNA methylation: evolution of a bacterial immune function into a regulator of gene expression and genome structure in higher eukaryotes. Philos. Trans. R. Soc. Lond. B Biol. Sci. *326*, 179–187.
15. Bird, A.P. (1995). Gene number, noise reduction and biological complexity. Trends Genet. *11*, 94–100.
16. Kato, M., Miura, A., Bender, J., Jacobsen, S.E., and Kakutani, T. (2003). Role of CG and non-CG methylation in immobilization of transposons in Arabidopsis. Curr. Biol. *13*, 421–426.
17. Lewis, Z.A., Honda, S., Khlafallah, T.K., Jeffress, J.K., Freitag, M., Mohn, F., Schubeler, D., and Selker, E.U. (2009). Relics of repeat-induced point mutation direct heterochromatin formation in Neurospora crassa. Genome Res. *19*, 427–437.
18. Galagan, J.E., and Selker, E.U. (2004). RIP: the evolutionary cost of genome defense. Trends Genet. *20*, 417–423.
19. Malagnac, F., Wendel, B., Goyon, C., Faugeron, G., Zickler, D., Rossignol, J.L., Noyer-Weidner, M., Vollmayr, P., Trautner, T.A., and Walter, J. (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.
20. Freitag, M., Williams, R.L., Kothe, G.O., and Selker, E.U. (2002). A cytosine methyltransferase homologue is essential for repeat-induced point mutation in Neurospora crassa. Proc. Natl. Acad. Sci. USA *99*, 8802–8807.
21. Lee, D.W., Freitag, M., Selker, E.U., and Aramayo, R. (2008). A cytosine methyltransferase homologue is essential for sexual development in Aspergillus nidulans. PLoS One *3*, e2531.
22. Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. Nat. Rev. Genet. *9*, 465–476.
23. Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol. Cell *31*, 785–799.
24. Simmen, M.W., Leitgeb, S., Charlton, J., Jones, S.J., Harris, B.R., Clark, V.H., and Bird, A. (1999). Nonmethylated transposable elements and methylated genes in a chordate genome. Science *283*, 1164–1167.
25. Tran, R.K., Henikoff, J.G., Zilberman, D., Ditt, R.F., Jacobsen, S.E., and Henikoff, S. (2005). DNA methylation profiling identifies CG methylation clusters in Arabidopsis genes. Curr. Biol. *15*, 154–159.
26. Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E., *et al*. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. Cell *126*, 1189–1201.
27. Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat. Genet. *39*, 61–69.
28. Wang, X., Elling, A.A., Li, X., Li, N., Peng, Z., He, G., Sun, H., Qi, Y., Liu, X.S., and Deng, X.W. (2009). Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize. Plant Cell *21*, 1053–1069.
29. Xiang, H., Zhu, J., Chen, Q., Dai, F., Li, X., Li, M., Zhang, H., Zhang, G., Li, D., Dong, Y., *et al*. Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. Nat. Biotechnol. *28*, 516–520.
30. Miura, A., Nakamura, M., Inagaki, S., Kobayashi, A., Saze, H., and Kakutani, T. (2009). An Arabidopsis jmjC domain protein protects transcribed genes from DNA methylation at CHG sites. EMBO J. *28*, 1078–1086.
31. Penterman, J., Zilberman, D., Huh, J.H., Ballinger, T., Henikoff, S., and Fischer, R.L. (2007). DNA demethylation in the Arabidopsis genome. Proc. Natl. Acad. Sci. USA *104*, 6752–6757.
32. Zhu, J.K. (2009). Active DNA demethylation mediated by DNA glycosylases. Annu. Rev. Genet. *43*, 143–166.
33. Zlatanova, J., and Thakar, A. (2008). H2A.Z: view from the top. Structure *16*, 166–179.
34. Zilberman, D., Coleman-Derr, D., Ballinger, T., and Henikoff, S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. Nature *456*, 125–130.
35. Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D., *et al*. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714–717.
36. Thomson, J.P., Skene, P.J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A.R., Deaton, A., Andrews, R., James, K.D., *et al*. (2010). CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature *464*, 1082–1086.
37. Aceituno, F.F., Moseyko, N., Rhee, S.Y., and Gutierrez, R.A. (2008). The rules of gene expression in plants: organ identity and gene body methylation are key factors for regulation of gene expression in Arabidopsis thaliana. BMC Genomics *9*, 438.

38. Krieg, A.M. (2006). Therapeutic potential of Toll-like receptor 9 activation. Nat. Rev. Drug Discov. *5*, 471–484.
39. Van Etten, J.L. (2003). Unusual life style of giant chlorella viruses. Annu. Rev. Genet. *37*, 153–195.
40. Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. Science *327*, 996–1000.
41. Kucharski, R., Maleszka, J., Foret, S., and Maleszka, R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. Science *319*, 1827–1830.
42. Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E., and Fischer, R.L. (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. Cell *110*, 33–42.
43. Gehring, M., Bubb, K.L., and Henikoff, S. (2009). Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science *324*, 1447–1451.
44. Hsieh, T.F., Ibarra, C.A., Silva, P., Zemach, A., Eshed-Williams, L., Fischer, R.L., and Zilberman, D. (2009). Genome-wide demethylation of Arabidopsis endosperm. Science *324*, 1451–1454.
45. Mosher, R.A., Meinyk, C.W., Kelly, K.A., Dunn, R.L., Studholme, D.J., and Baulcombe, D.C. (2009). Uniparental expression of PolIV-dependent siRNAs in developing endosperm of Arabidopsis. Nature *460*, 283–287.
46. Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzic, M., Becker, J.D., Feijo, J.A., and Martienssen, R.A. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell *136*, 461–472.
47. Edwards, C.A., and Ferguson-Smith, A.C. (2007). Mechanisms regulating imprinted genes in clusters. Curr. Opin. Cell Biol. *19*, 281–289.
48. Illingworth, R.S., and Bird, A.P. (2009). CpG islands – 'a rough guide'. FEBS Lett. *583*, 1713–1720.
49. Jones, P.A., and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. Nat. Rev. Genet. *3*, 415–428.
50. Klose, R.J., and Bird, A.P. (2006). Genomic DNA methylation: the mark and its mediators. Trends Biochem. Sci. *31*, 89–97.
51. 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.
52. Hickey, D.A. (1982). Selfish DNA: a sexually-transmitted nuclear parasite. Genetics *101*, 519–531.
53. Bestor, T.H. (1999). Sex brings transposons and genomes into conflict. Genetica *107*, 289–295.
54. Orgel, L.E., and Crick, F.H. (1980). Selfish DNA: the ultimate parasite. Nature *284*, 604–607.
55. Aravin, A.A., Hannon, G.J., and Brennecke, J. (2007). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science *318*, 761–764.
56. Dean, W., Santos, F., Stojkovic, M., Zakhartchenko, V., Walter, J., Wolf, E., and Reik, W. (2001). Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proc. Natl. Acad. Sci. USA *98*, 13734–13738.