Establishing, maintaining and modifying DNA methylation patterns in plants and animals

Julie A. Law and Steven E. Jacobsen*‡*

Abstract | Cytosine DNA methylation is a stable epigenetic mark that is crucial for diverse biological processes, including gene and transposon silencing, imprinting and X chromosome inactivation. Recent findings in plants and animals have greatly increased our understanding of the pathways used to accurately target, maintain and modify patterns of DNA methylation and have revealed unanticipated mechanistic similarities between these organisms. Key roles have emerged for small RNAs, proteins with domains that bind methylated DNA and DNA glycosylases in these processes. Drawing on insights from both plants and animals should deepen our understanding of the regulation and biological significance of DNA methylation.

Epigenetic modifications
Chemical additions to DNA and histones that are associated with changes in gene expression and are heritable but do not alter the primary DNA sequence.

The genetic information in a cell is encoded by DNA, which is packaged into chromatin. Epigenetic modifications of DNA and histones, the core components of chromatin, constitute an additional layer of information that influences the expression of the underlying genes. One such epigenetic modification is DNA methylation (the addition of a methyl group to a cytosine base), which is evolutionarily ancient and associated with gene silencing in eukaryotes. DNA methylation defects in mammals are embryonic lethal and in plants they can lead to pleiotropic morphological defects, which attests to the importance of this modification.

In mammals, DNA methylation occurs almost exclusively in the symmetric CG context and is estimated to occur at ~70–80% of CG dinucleotides throughout the genome[1]. However, a small amount of non-CG methylation is observed in embryonic stem (ES) cells[2–4]. The remaining unmethylated CG dinucleotides are mostly found near gene promoters in dense clusters known as CpG islands[5,6]. In plants, DNA methylation commonly occurs at cytosine bases in all sequence contexts: the symmetric CG and CHG contexts (in which H = A, T or C) and the asymmetric CHH context[7]. In *Arabidopsis thaliana*, genome-wide DNA methylation levels of approximately 24%, 6.7% and 1.7% are observed for CG, CHG and CHH contexts, respectively[8]. Unlike in mammals, DNA methylation in plants predominantly occurs on transposons and other repetitive DNA elements[9].

In mammals, DNA methylation patterns are established by the DNA methyltransferase 3 (DNMT3) family of *de novo* methyltransferases and maintained by the maintenance methyltransferase DNMT1[10–12] (FIG. 1). In plants, *de novo* methylation is catalysed by DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), a homologue of the DNMT3 methyltransferases, and maintained by three different pathways: CG methylation is maintained by DNA METHYLTRANSFERASE 1 (MET1, also known as DMT1), the plant homologue of DNMT1; CHG methylation is maintained by CHROMOMETHYLASE 3 (CMT3), a plant-specific DNA methyltransferase; and asymmetric CHH methylation is maintained through persistent *de novo* methylation by DRM2 (REF. 13) (FIG. 1). However, the pathways that control the establishment and maintenance of DNA methylation, as well as those involved in the removal of DNA methylation, are less well characterized.

In this Review, we focus on recent studies in plants and animals that have greatly expanded our understanding of such pathways. We begin with the establishment of DNA methylation, with a separate section focusing on the dynamics of DNA methylation in reproductive cells and the roles of small RNAs at this stage of development. We then discuss mechanisms that govern the maintenance and removal of DNA methylation. In each section, recent advances from plants and animals are presented and similarities and differences are highlighted. As we discuss, small RNAs, methyl-binding domain proteins

*Department of Molecular, Cell and Developmental Biology, University of California-Los Angeles, Los Angeles, California 90095-1606, USA
‡Howard Hughes Medical Institute, University of California-Los Angeles, Los Angeles, California 90095-1662, USA
Correspondence to S.E.J.
e-mail: jacobsen@ucla.edu
doi:10.1038/nrg2719
Published online 9 February 2010

© 2010 Macmillan Publishers Limited. All rights reserved

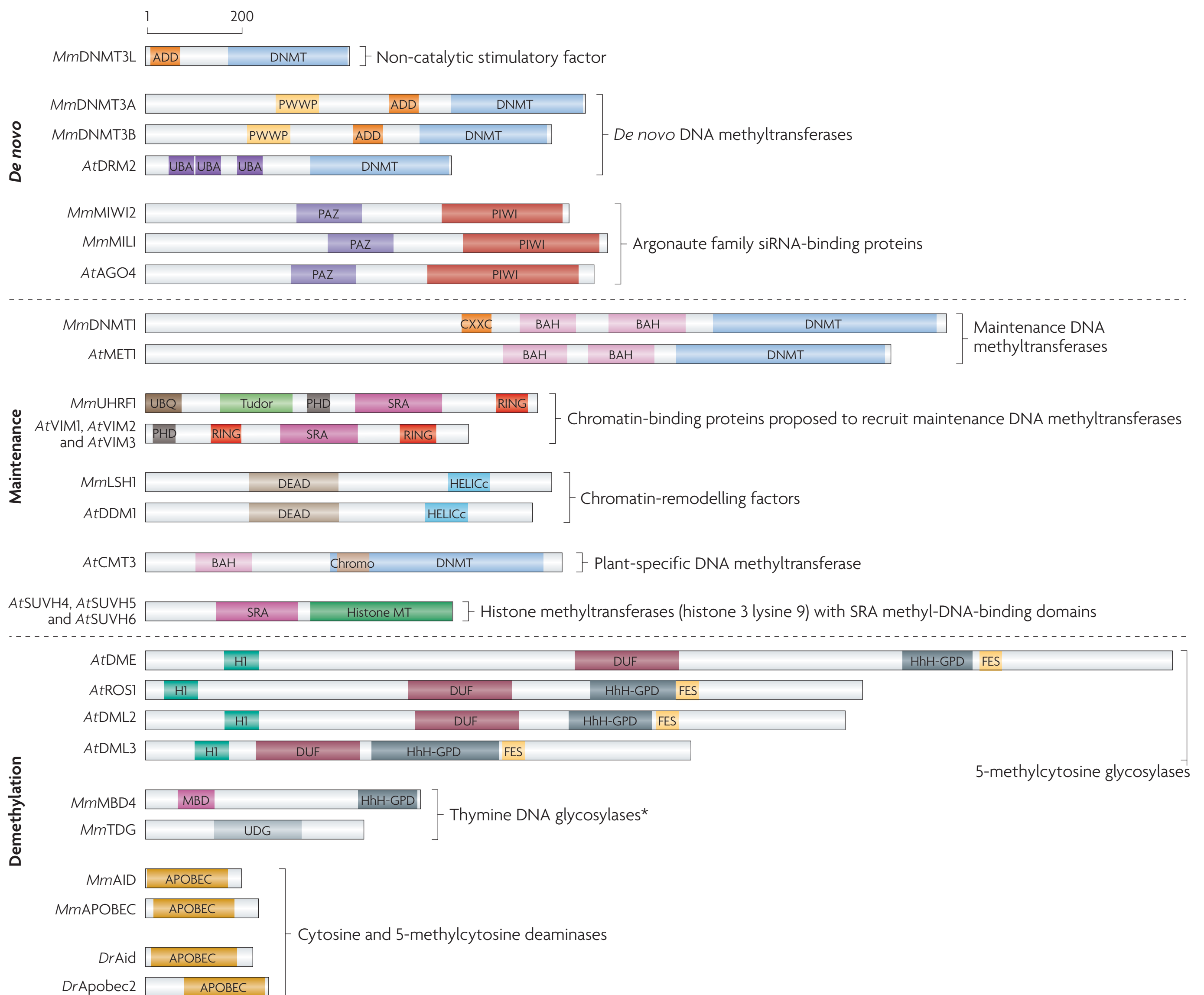

Figure 1 | **Proteins involved in *de novo* DNA methylation, maintenance methylation and demethylation.** Select proteins with established roles in these processes are shown for mice (*Mus musculus* (*Mm*)), *Arabidopsis thaliana* (*At*) and zebrafish (*Danio rerio* (*Dr*)). The *Mm* DNA methyltransferase 3 (DNMT3) family and *At* DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) contain DNMT domains, although in *At*DRM2 the catalytic motifs are rearranged. The *Mm*DNMT3 proteins also possess a cysteine-rich domain that contains a plant homeodomain (PHD) zinc finger motif and is referred to as an ATRX–DNMT3–DNMT3-like (DNMT3L) (ADD) domain. *Mm*DNMT3A and *Mm*DNMT3B possess a proline-tryptophan-tryptophan-proline (PWWP) motif that *Mm*DNMT3L lacks. *At*DRM2 contains ubiquitin-associated (UBA) domains. *Mm*MIWI2 (also known as PIWIL4), *Mm*MILI (also known as PIWIL2) and *At* ARGONAUTE 4 (AGO4) possess a Piwi Ago and Zwille (PAZ) domain and a PIWI domain. *Mm*DNMT1 and *At* DNA METHYLTRANSFERASE 1 (MET1, also known as DMT1) possess bromo-adjacent homology (BAH) domains and a DNMT domain. *Mm*DNMT1 also contains a cysteine-rich (CXXC) domain. *Mm* ubiquitin-like PHD and RING finger domain 1 (UHRF1) and the *At* VARIANT IN METHYLATION (VIM, also known as ORTHRUS) family contain SET- or RING-associated (SRA), RING and PHD domains. *Mm*UHRF1 also has a Tudor domain and a ubiquitin domain (UBQ). *Mm* lymphoid-specific helicase 1 (LSH1, also known as HELLS) and *At* DECREASED IN DNA METHYLATION 1 (DDM1) contain DEAD and HELICc helicase domains. *At* CHROMOMETHYLASE 3 (CMT3) contains a DNMT domain, a chromodomain and a BAH domain. *At* SUPPRESSOR OF VARIEGATION 3-9 HOMOLOGUE 4 (SUVH4), *At*SUVH5 and *At*SUVH6 possess an SRA and a histone methyltransferase (histone MT) domain. The *At* DEMETER (DME)/REPRESSOR OF SILENCING 1 (ROS1) family of glycosylases all possess a helix–hairpin–helix-Gly-Pro-Asp (HhH-GPD) domain, a 4Fe-4S (FES) cluster, a domain with similarity to histone H1 (H1), and a domain of unknown function (DUF). *Mm* methyl-CpG-binding domain 4 (MBD4) contains an HhH-GDP domain and an MBD domain. *Mm* thymine DNA glycosylase (TDG) contains a uracil DNA glycosylase (UDG) domain. *Mm* and *Dr* activation-induced cytosine deaminase (AID) and apolipoprotein B mRNA-editing enzyme (APOBEC) proteins all contain an APOBEC domain. '*' indicates activity on additional substrates[144]. siRNA, small interfering RNA.

Histones
The main protein components of chromatin. The four core histones, H2A, H2B, H3 and H4, form a globular octameric complex called a nucleosome upon which DNA is wrapped. The amino-terminal regions of histone proteins are largely unstructured and are subject to various chemical modifications, including methylation.

CpG island
A sequence of at least 200 bp with a greater number of CpG sites than expected for its GC content. These regions are often GC rich, typically undermethylated, and are found upstream of many mammalian genes.

© 2010 Macmillan Publishers Limited. All rights reserved

RNA-directed DNA methylation
A plant-specific pathway through which small RNAs (24 nucleotides long) target the *de novo* methyltransferase DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) to homologous genomic loci to establish DNA methylation, which leads to transcriptional gene silencing.

and DNA glycosylases are common components of the pathways that define dynamic DNA methylation patterns in the two taxonomic groups.

De novo DNA methylation

De novo *methylation in plants.* Throughout plant development, small RNAs target homologous genomic DNA sequences for cytosine methylation in all sequence contexts through a phenomenon that was initially observed by Wassenegger *et al.*[14] and is known as RNA-directed DNA methylation (RdDM)[7,15]. In addition to the canonical RNA interference (RNAi) machinery (that is, members of the Dicer and Argonaute families) and DRM2, RdDM requires two plant-specific RNA polymerases, Pol IV and Pol V (which have largely non-redundant functions[16,17]), two putative chromatin-remodelling factors and several other recently identified proteins[15]. Through the characterization of these components, an increasingly detailed mechanistic understanding of RdDM is emerging (FIG. 2).

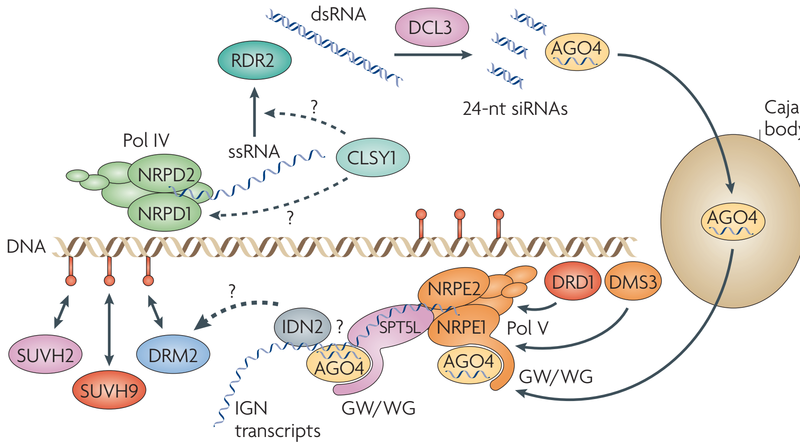

Figure 2 | **Model for RNA-directed DNA methylation.** Single-stranded RNA transcripts corresponding to transposons and repeat elements are thought to be generated by RNA polymerase IV (Pol IV). CLASSY 1 (CLSY1, also known as CHR38), a putative chromatin-remodelling factor, is likely to function early in RNA-directed DNA methylation (RdDM), possibly recruiting Pol IV to chromatin or aiding in ssRNA transcript processing. RNA-DEPENDENT RNA POLYMERASE 2 (RDR2) is proposed to generate dsRNA from the ssRNA transcripts. DICER-LIKE 3 (DCL3) is thought to process the dsRNAs into 24-nucleotide (nt) small interfering RNAs (siRNAs), which are bound by an Argonaute protein, AGO4. AGO4 localizes to Cajal bodies, and although the function of this association remains unknown, it seems to be necessary for wild-type levels of RdDM[33]. AGO4 also colocalizes with two Pol V subunits — NUCLEAR RNA POLYMERASE E1 (NRPE1) and NRPE2 — and DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) at a distinct nuclear focus, the AGO4–NRPE1 body (not depicted), which may represent a site of active RdDM[33]. Pol V is thought to transcribe intergenic non-coding (IGN) regions throughout the genome. NRPE1 association with chromatin requires another putative chromatin-remodelling factor, DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1), and a structural maintenance of chromosome (SMC) domain protein, DEFECTIVE IN MERISTEM SILENCING 3 (DMS3). IGN transcripts may serve as a scaffold for recruiting AGO4, which interacts with the GW/WG motifs of NRPE1 and SUPPRESSOR OF TY INSERTION 5-LIKE (SPT5L, also known as KTF1), possibly through interactions between AGO4-bound siRNAs and the nascent transcript. An RNA-binding protein, INVOLVED IN DE NOVO 2 (IDN2), is proposed to recognize the siRNA–nascent transcript duplex. These associations may aid in targeting DRM2 to genomic loci that produce both 24-nt siRNAs and IGN transcripts. Recruitment or retention of DRM2 at such loci may be aided by SUPPRESSOR OF VARIEGATION 3-9 HOMOLOGUE 9 (SUVH9) and SUVH2, two proteins that bind methylated DNA and are likely to act late in RdDM. '?' indicates a putative function. The red circles represent DNA methylation.

The biogenesis of the 24-nucleotide (nt) small interfering RNAs (siRNAs) that are required for targeting DNA methylation depends on Pol IV, RNA-DEPENDENT RNA POLYMERASE 2 (RDR2) and DICER-LIKE 3 (DCL3). Other RdDM components, including DRM2, ARGONAUTE 4 (AGO4) and Pol V, are needed for siRNA accumulation for a subset of loci; however, these proteins do not seem to be involved in the initial production of siRNAs and are proposed to reinforce siRNA biogenesis by an unknown mechanism[7,18]. Additional subunits or interacting partners of Pol IV and Pol V have recently been identified[19–23]. Whereas some subunits are shared with Pol II, others are unique to Pol IV, Pol V or both[20]. Although no polymerase activity has been shown for Pol IV, mutations in the largest subunit, NUCLEAR RNA POLYMERASE D1 (NRPD1), including mutations in the conserved metal-binding motif, greatly reduce the abundance of siRNAs[18,24–30], which suggests that Pol IV may be an active polymerase. Pol IV is suggested to initiate siRNA biogenesis by producing long ssRNA transcripts. These transcripts are then thought to be acted upon by RDR2, which generates dsRNAs that are processed into 24-nt siRNAs by DCL3 and loaded into AGO4 (REFS 7,15). AGO4 interacts with the Pol V subunit NUCLEAR RNA POLYMERASE E1 (NRPE1)[31,32], and this interaction is required for RdDM[31], leading to the hypothesis that this complex functions as a downstream effector of DNA methylation. *In vivo*, AGO4 colocalizes with Cajal bodies or with NRPE1, NRPE2 and DRM2 at a separate discrete nuclear body known as the AGO4–NRPE1 (AB) body (note that NRPE1 was previously known as NRPD1b)[32,33]. The AB body is adjacent to 45S ribosomal DNA and may be a site of active RdDM[33].

A recent study further clarified the role of Pol V in RdDM by identifying low-abundance intergenic non-coding (IGN) transcripts from several loci that depend on Pol V for their accumulation[34]. NRPE1 is present at these transcribed DNA regions and is associated with the RNA transcripts, which suggests that Pol V is an active polymerase[34]. These Pol V-dependent transcripts are required for DNA methylation and silencing of surrounding loci, but their accumulation does not depend on NRPD1, DCL3 or RDR2 (REF. 34), which suggests that Pol V acts in RdDM through a pathway that is independent of siRNAs. These IGN transcripts are proposed to function as scaffolds for the recruitment of the silencing machinery, possibly facilitated by base-pairing interactions between AGO4-bound siRNAs and nascent Pol V transcripts[34]. A requirement of transcription for silencing is also observed in fission yeast, in which transcription of heterochromatic DNA by Pol II is required for siRNA-mediated heterochromatin formation[35].

Current models of RdDM posit that Pol V-dependent transcripts and siRNAs are both required for silencing a particular locus. Several studies support the hypothesis that AGO4 and/or SUPPRESSOR OF TY INSERTION 5-LIKE (SPT5L, also known as KTF1) may bridge the siRNA and IGN transcript-generating pathways. SPT5L, a protein with homology to the yeast transcription elongation factor Spt5, was recently identified as a

© 2010 Macmillan Publishers Limited. All rights reserved

RNA interference
A process of post-transcriptional gene silencing in which small RNAs, often generated by the activity of an RNA-dependent RNA polymerase and a Dicer endoribonuclease, are bound by Argonaute proteins and target cleavage of homologous mRNA transcripts.

Dicer
An RNase III family endonuclease that processes dsRNAs into small interfering RNAs.

Argonautes
Effector proteins of small RNA-directed silencing. Small RNAs guide Argonautes to their RNA targets. Argonaute proteins are characterized by two domains — PIWI (a ribonuclease domain) and Piwi Argonaute and Zwille (PAZ; an ssRNA-binding module).

Chromatin-remodelling factors
Proteins that have the capacity to remodel chromatin, often using the energy of ATP, so that gene transcription can be activated or silenced.

Small interfering RNAs
20–25 nucleotide-long RNAs that are generated from dsRNAs and serve as guides for the cleavage of homologous mRNAs in RNA interference or for the addition of chromatin modifications, including histone and DNA methylation at homologous genomic sequences in transcriptional gene silencing.

Cajal bodies
Nuclear bodies that are associated with the maturation of ribonucleoprotein complexes.

Heterochromatin
A densely packaged form of chromatin that is associated with repressive histone modifications, DNA methylation and gene silencing.

Primordial germ cells
The population of embryonic cells from which germ cells are formed.

Imprinted genes
Genes in which one allele is expressed in a parent-of-origin-specific manner.

downstream effector of RdDM[21,23,36]. SPT5L and NRPE1 can both interact with AGO4 through a conserved GW/WG motif (also known as an Ago hook motif) that is present in their carboxy-terminal regions[23,31,32,36]. *In vivo*, both SPT5L and AGO4 interact with Pol V-dependent transcripts[36,37], prompting speculation that SPT5L serves as an adaptor protein that binds AGO4 and nascent Pol V transcripts, aiding in the recruitment of AGO4 to Pol V-transcribed loci. This interaction may also be required to recruit the silencing machinery, including DRM2, to establish DNA methylation.

Another factor that is thought to act as a downstream RdDM effector, INVOLVED IN DE NOVO 2 (IDN2), was recently identified[38]. IDN2 has homology with SUPPRESSOR OF GENE SILENCING 3 (SGS3), a protein that is involved in post-transcriptional gene silencing, and like SGS3, IDN2 contains an XS domain that can recognize dsRNAs with 5′ overhangs[38]. A possible RNA substrate for IDN2 is the duplex that is formed between AGO4-bound siRNAs and Pol V non-coding transcripts[38], which could also be a signal that aids in recruitment of DRM2 to establish DNA methylation.

In addition to Pol V-dependent transcripts, Pol II-dependent non-coding transcripts that are required for transcriptional gene silencing at some loci have recently been identified in a weak *nuclear RNA polymerase B2* (*nrpb2*) mutant, and these transcripts are also proposed to act as scaffolds for the recruitment of RdDM factors, including AGO4 and possibly Pol IV and Pol V[39]. Two genetic screens for RdDM factors provided further support for a role of Pol II in RdDM by identifying a conserved protein, DEFECTIVE IN MERISTEM SILENCING 4 (DMS4, also known as RDM4), that has similarity to the yeast protein known as interacts with Pol II (Iwr1)[40,41]. The precise relationship among Pol II, Pol V and Pol IV remains elusive, but these studies suggest they may be more intimately connected than previously thought.

Although the mechanisms through which Pol IV and Pol V are targeted to specific loci are poorly understood, several recent findings are beginning to shed light on these aspects of RdDM. DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1), a putative chromatin-remodelling factor[42], and DMS3, an RdDM component with similarity to structural maintenance of chromosome (SMC) proteins[38,43], are needed for NRPE1 chromatin association and for accumulation of IGN transcripts[34,37], although how these components are targeted is unknown. In addition, the NRPB2 subunit of Pol II aids in the association of both NRPE1 and NRPD1 with chromatin, which suggests that Pol II-dependent transcripts, or the act of transcription, may recruit Pol IV and Pol V to specific loci[39]. Finally, a putative chromatin-remodelling factor, CLASSY 1 (CLSY1, also known as CHR38), may be involved at an early stage of siRNA production, possibly at the level of Pol IV or RDR2 activity[44].

De novo *methylation in mammals.* Unlike in plants, DNA methylation in mammals covers most of the genome, with the main exception being CpG islands. This DNA methylation pattern is largely established during early embryogenesis, at around the time of implantation[45,46], through the activity of DNMT3A and DNMT3B[5,47,48]. However, during post-implantation development, further epigenetic reprogramming occurs in primordial germ cells (PGCs). Following a wave of demethylation, which is required to erase DNA methylation imprints established in the previous generation, DNA methylation patterns are re-established at imprinted loci and transposable elements (TEs) during gametogenesis by DNMT3A and a non-catalytic paralogue, DNMT3-like (DNMT3L)[5,12,47,48]. Recent studies suggest that DNA methylation may be targeted to TEs and imprinted genes during germ cell development through different mechanisms, with targeting to TEs involving Piwi-interacting RNAs (piRNAs) (discussed later) and targeting to imprinted genes involving interactions between DNMT3L and unmethylated histone 3 lysine 4 (H3K4) tails.

Biochemical purification of DNMT3L led to the discovery that DNMT3L interacts with unmethylated H3K4 tails through its cysteine-rich ATRX–DNMT3–DNMT3L (ADD) domain[49,50]. As DNMT3L also interacts with DNMT3A[12,49], a model was proposed in which DNMT3L binds unmethylated H3K4 tails and recruits the DNMT3A2 isoform to specific loci, including imprinted loci[49,51] (FIG. 3). Indeed, an inverse relationship between H3K4 methylation and allele-specific DNA methylation has been reported at several imprinted loci[52–55]. Further supporting this model, an oocyte-specific H3K4 demethylase[56], lysine demethylase 1B (KDM1B), was recently shown to be required for the establishment of DNA methylation at several differentially methylated regions (DMRs) that are associated with imprinted genes during oogenesis[56], and defects in DNA methylation at such loci resulted in a loss of imprinting in developing embryos[56]. This model is also consistent with studies that show that H3K4 methylation seems to be anticorrelated with DNA methylation in multiple mammalian cell types[57–61] and with findings that H3K4 dimethylation and trimethylation are anticorrelated with DNA methylation in plants[62].

There is also evidence that, in addition to DNMT3A and DNMT3L, transcription across DMRs is required for imprinting[63]. Chotalia *et al.*[63] showed that such transcription occurs during oocyte growth (before or around the time when *de novo* methylation occurs) and is required for the establishment of DNA methylation, at least at the imprinted guanine nucleotide-binding protein, α-stimulating (*Gnas*) locus[63]. These findings led to the proposal that the act of transcription, or the transcripts themselves, may alter the chromatin structure of imprinted loci and/or recruit the histone-modifying enzymes and DNA methyltransferases that are required for establishing DNA methylation imprints[63].

Mechanistic insights into how *de novo* methyltransferases function once they are targeted to a particular locus have been provided by several biophysical studies that focused on the interaction between DNMT3L and DNMT3A. Co-crystallization of the C-terminal regions of DNMT3A and DNMT3L revealed a tetrameric

© 2010 Macmillan Publishers Limited. All rights reserved

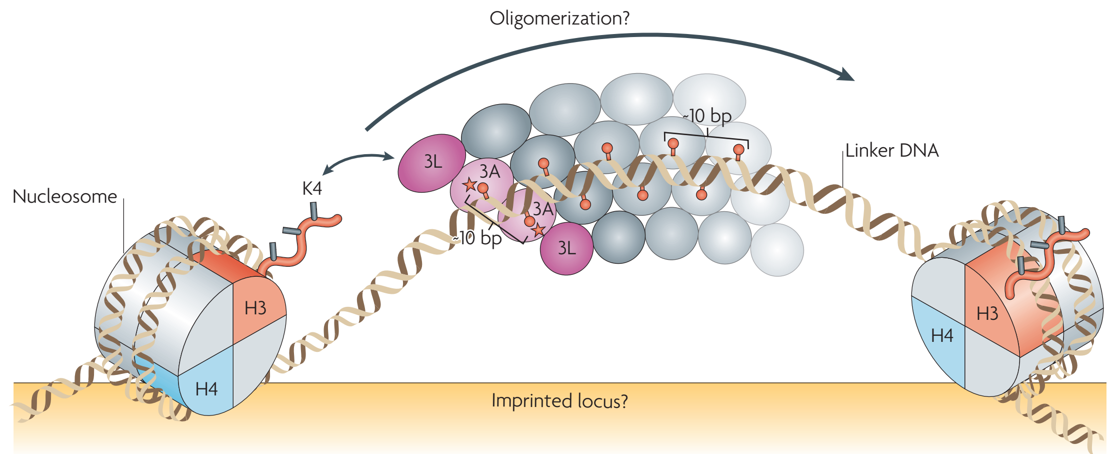

Figure 3 | **Model of recruitment of the *de novo* methylation machinery by unmethylated histone 3 lysine 4 tails.** The amino-terminal domain of DNA methyltransferase 3-like (DNMT3L, shown as 3L) possesses a cysteine-rich domain that interacts with unmethylated histone 3 lysine 4 (H3K4) tails, and this interaction is proposed to recruit or activate the DNMT3A2 isoform. The carboxy-terminal domains of DNMT3L and DNMT3A (shown as 3A) form a tetrameric complex in which two DNMT3A proteins interact with each other and are flanked by two DNMT3L proteins. The DNMT3A active sites (red stars) are thought to be separated by approximately one helical turn and therefore could catalyse methylation (red circles) on opposite DNA strands ~10 bp apart. After being recruited to a specific locus, the DNMT3L–DNMT3A tetramer might be able to oligomerize, which could result in an ~10 bp periodic pattern of DNA methylation along the same DNA strand.

complex that positions two DNMT3A proteins such that their active sites are adjacent to one another[51]. Two DNMT3L proteins are located on either side of the DNMT3A dimer and residues of DNMT3L may stabilize the active site loop in DNMT3A[51], which could account for the observed stimulatory effect of DNMT3L on DNMT3A and DNMT3B activity[64,65]. Superimposition of the DNMT3A C-terminal structure with that of the bacterial M.*Hha*I methyltransferase complexed with DNA[66] provided a model in which the two DNMT3A active sites are separated by approximately one helical DNA turn, which suggests that each tetrameric complex could simultaneously methylate two cytosine residues at a defined spacing of 8–10 bp[51]. This tetrameric complex was subsequently shown to oligomerize on DNA substrates, forming a filamentous nucleoprotein complex[67] (FIG. 3).

Consistent with the determined structural parameters, a periodicity for DNA methylation on opposite strands of a DNA duplex as well as along the same strand of DNA was observed *in vitro* through bisulphite sequencing analyses[51,67]. *In vivo*, the spacing of CG dinucleotides at many DMRs is also consistent with an ~8–10 bp periodicity[51], as is the finding that CG dinucleotides at an 8 bp spacing are overrepresented across the human genome[68,69] and, to a lesser extent, across the mouse genome[68]. As DNMT3A seems to be a non-processive DNA methyltransferase[70], the formation of an oligomer could help to explain the observed periodic pattern of DNA methylation. Whether oligomerization occurs *in vivo* remains unknown, but it is tempting to propose a model in which interactions between DNMT3L and unmethylated H3K4 tails, or possibly between DNMT3A and other histone modifications or histone methyltransferases, might target and set the register for oligomerization of tetramers consisting of DNMT3A and DNMT3L, resulting in an ~8–10 bp periodicity.

In *A. thaliana*, nucleotide-resolution DNA methylation mapping revealed an element of periodicity for DNA methylation. For CHH methylation (which is mostly controlled by DRM2), a period of ~10 bp was observed genome wide[8], suggesting that the periodicity observed for DNMT3A may be a common feature of *de novo* methyltransferases and that it may also occur on a genome-wide scale in mammals. For CHG methylation (which is mostly controlled by CMT3), a period of approximately the size of a nucleosome, 167 nt, was found[8], which is consistent with the chromodomain in CMT3 interacting with methylated H3 tails[71].

In addition to interactions with unmethylated H3K4 tails, other mechanisms for targeting DNA methylation to specific loci throughout the genome shape the overall methylation landscape during mammalian development. These include interactions between DNMT3A and DNMT3B and the histone methyltransferases G9a, enhancer of zeste homologue 2 (EZH2), suppressor of variegation 3-9 homologue 1 (SUV39H1) and SET domain bifurcated 1 (SETDB1)[5]. More recently, Zhao *et al.*[72] showed that symmetric methylation of histone 4 arginine 3 (H4R3me2s) by protein arginine *N*-methyltransferase 5 (PRMT5) can recruit DNMT3A to the human β-globin locus, which is required for the DNA methylation and silencing of this gene. *In vitro* characterization of the interaction between the H4R3me2s modification and DNMT3A showed that the ADD domain of DNMT3A is sufficient to mediate this interaction[72]. Although histone arginine methylation has been implicated in gene silencing[73], this finding is the first direct link between arginine methylation and DNA methylation.

Bisulphite sequencing
A technique in which the treatment of DNA with bisulphite, which converts cytosines into uracils but does not modify methylated cytosines, is used to determine the DNA methylation pattern.

© 2010 Macmillan Publishers Limited. All rights reserved

DNA methylation in reproductive cells

Transposons and other repetitive DNA elements are highly abundant in plant and mammal genomes. Owing to the high risk TEs pose to genome integrity, their expression must be tightly regulated. Such control is particularly important in cells that transmit genetic information to the subsequent generation. In plants and mammals, such elements are targeted by the *de novo* methylation machinery and are maintained in a methylated and silenced state. Recent evidence suggests that in mammals, like in plants, small RNAs have an important role in targeting transposons for methylation.

RdDM in plant reproductive cells. In plants, DNA methylation patterns seem to be maintained in a multigenerational manner, which has led to the view that DNA methylation in plants is quite static. However, several complementary studies have shown that transposon reactivation and genome-wide losses of DNA methylation occur during male and female gametogenesis, respectively, indicating that DNA methylation patterns in plants are dynamic during development. This hypomethylation is similar to the global demethylation observed in PGCs and on the paternal genome during mammalian development[47,48]. These recent studies suggest that in *A. thaliana*, these changes may reinforce transposon silencing in the sperm and egg cells[74–76] (FIG. 4).

During male gametogenesis, tricellular pollen grains that contain a vegetative nucleus and two sperm cells are produced[77] (FIG. 4Ba). Analysis of transposon expression in different plant tissues revealed that transposons, which are methylated and silenced in most tissues, are expressed and mobile in pollen[74]. In the pollen grain, transposon reactivation seems to be restricted to the vegetative nucleus. This is a key distinction, as the sperm cells, but not the vegetative nucleus, provide genetic information to subsequent generations[77] and therefore their genome integrity must be protected. Consistent with decreased DNA methylation and transposon activation, several RdDM components are downregulated in pollen[74,78], and DECREASE IN DNA METHYLATION 1 (DDM1), a chromatin-remodelling factor that is required for maintenance of CG methylation[79], seems to be excluded from the vegetative nucleus[74]. Sequencing of siRNA populations from pollen and isolated sperm cells showed an increase in 21-nt siRNAs in sperm cells[74]. As these siRNAs correspond to transposons that do not seem to be expressed in the sperm cells, it was postulated that siRNAs generated in the vegetative nucleus might travel to the sperm cells and reinforce silencing by an unknown mechanism[74] (FIG. 4Ba).

The two sperm cells fertilize the central cell and the egg cell of the multicellular female gametophyte in a double fertilization event that generates the embryo and the endosperm, respectively[77] (FIG. 4B,C). Although previous studies have documented decreased DNA methylation at discrete imprinted loci in endosperm[77], two recent studies show that endosperm DNA methylation is reduced genome-wide, and this reduction is likely to originate from demethylation in the central cell nucleus of the female gametophyte[75,76]. These findings are in line with observations that chromatin seems to be less condensed in endosperm nuclei[80]. Despite this global decrease in DNA methylation, Hsieh *et al.*[75] found increased CHH methylation in both the endosperm and embryo tissues relative to adult shoot tissue and suggest that this hypermethylation could result from enhanced RdDM. Consistent with these findings, profiling of Pol IV-dependent siRNA levels in different plant tissues shows that maternal-derived siRNAs accumulate to high levels in the endosperm[81]. Analogous with the model of reinforced silencing in the male gametophyte[74], these findings led to the suggestion that siRNAs potentially generated in the central cell may reinforce silencing in the egg cell and possibly in the developing embryo[75] (FIG. 4B,C).

As potentially deleterious transposition events occurring in the sperm or egg cells would be inherited in subsequent generations, demethylation during gametogenesis may function to reveal TEs in the genome that have the potential to be expressed and may arm siRNA-based pathways to ensure that these elements are efficiently silenced. Such a mechanism would be inherently adaptable, as newly integrated transposons would also be expressed, leading to siRNA production and the establishment of silencing. Interestingly, Teixeira *et al.*[82] recently showed that siRNA-producing loci, unlike other regions of the *A. thaliana* genome, can be remethylated in all sequence contexts when methylation is lost in previous generations, suggesting a dynamic role for RdDM in correcting DNA methylation defects. However, remethylation to approximately wild-type levels was only observed after multiple generations, as is the case when newly inserted transgenes become silenced. It is tempting to speculate that decreased methylation and transposon reactivation during gametogenesis might be required to generate siRNA signals and allow the observed re-establishment of silencing.

piRNAs in mammalian germ cells. A small RNA pathway is also required for silencing some transposons in mammals during male gametogenesis (FIG. 5). Whereas RdDM in plants uses 24-nt siRNAs, transposon control in mammals uses 25–30-nt piRNAs, which were initially identified in *Drosophila melanogaster*[83]. In *D. melanogaster*, piRNAs bound by the Piwi clade of Agos guide cleavage of transposon transcripts, which results in post-transcriptional gene silencing[84]. This clade of Agos is highly conserved in animals[85] and initial genetic analysis in mammals and flies suggested that roles for piRNAs in germ cell development and transposon silencing were also conserved[84]. But early studies of mammalian piRNA populations revealed that, unlike in *D. melanogaster*, mammalian piRNAs were not enriched for repetitive regions of the genome[86,87], leaving it unclear whether mammalian piRNAs function to silence transposons. However, subsequent analyses of piRNA populations isolated at earlier stages of mouse development revealed an enrichment in repetitive DNA sequences[88–90]. These piRNA populations possess the characteristic sequence properties of primary piRNAs[88–90] and secondary piRNAs[88,90], which suggests that they are generated by a mechanism

Vegetative nucleus
The nucleus of a terminally differentiated vegetative cell. It does not contribute genetic information to subsequent generations.

Gametophyte
A multicellular structure that is generated from a haploid spore through mitotic cell divisions and contains the male or female gamete.

Endosperm
The product of fertilization of the central cell of the female gametophyte. It is present in the seeds of most flowering plants and provides nutrition to the developing embryo.

Primary piRNAs
(Primary Piwi-interacting RNAs.) The products of piRNA precursor transcript processing. These piRNAs have a preference for a 5′ uridine.

Secondary piRNAs
(Secondary Piwi-interacting RNAs.) The products of a ping-pong amplification cycle. These piRNAs are antisense to primary piRNAs and have a preference for an adenine at position 10.

© 2010 Macmillan Publishers Limited. All rights reserved

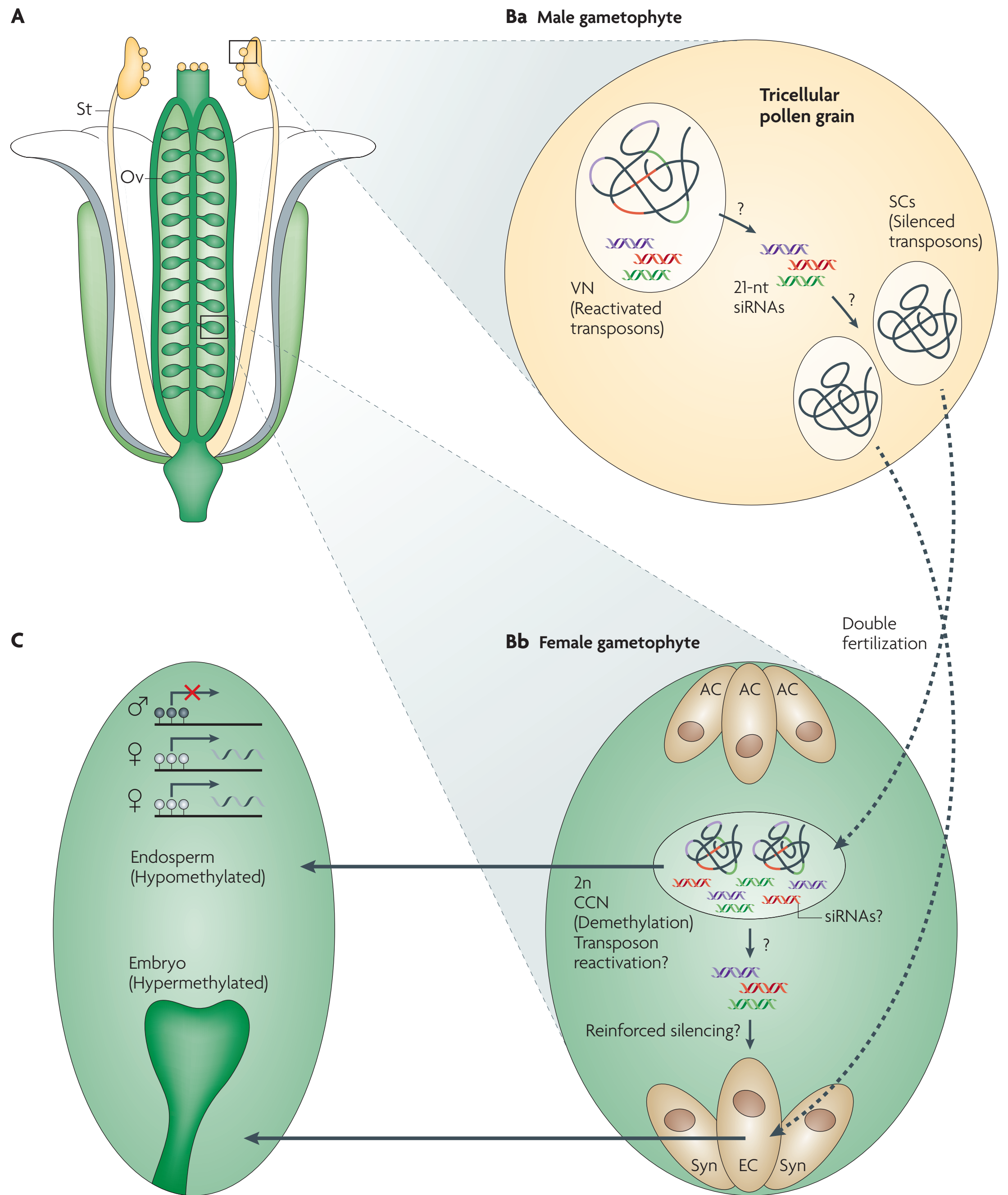

Figure 4 | **DNA methylation changes during plant development. A** | Diagram of an *Arabidopsis thaliana* flower. A stamen (St) and ovule (Ov) are labelled. **Ba** | Male gametogenesis occurs in stamens and generates tricellular pollen grains that contain a vegetative nucleus (VN) and two sperm cells (SCs). The model shows that transposon reactivation and small interfering RNA (siRNA) production take place specifically in the VN. These siRNAs may travel to the SCs to reinforce transposon silencing. **Bb** | Female gametogenesis occurs in ovules and produces a multicellular gametophyte with three antipodal cells (ACs), two synergid cells (labelled with 'Syn'), one egg cell (EC) and a diploid (2n) central cell nucleus (CCN)[77]. The model shows that siRNAs are found in the CCN, possibly as a consequence of global demethylation. These siRNAs may travel to the EC to reinforce silencing. Reinforced silencing in the sperm and egg cells could account for the observed hypermethylation of the embryo. **C** | Fertilization of the EC and CCN generate the embryo and endosperm, respectively. The embryo will give rise to the mature *A. thaliana* plant, whereas the endosperm is a terminally differentiated tissue. Imprinting is observed in the endosperm, which nourishes the embryo and is therefore analogous to the placenta in mammals, in which imprinting also occurs. In plants, maternal imprinting results from demethylation in the CCN by the DEMETER (DME) glycosylase, which is likely to account for the observed hypomethylation in this tissue. After fertilization, the unmethylated (white circles) maternal alleles are expressed in the endosperm, whereas the paternal allele is methylated (black circles) and silent.

© 2010 Macmillan Publishers Limited. All rights reserved

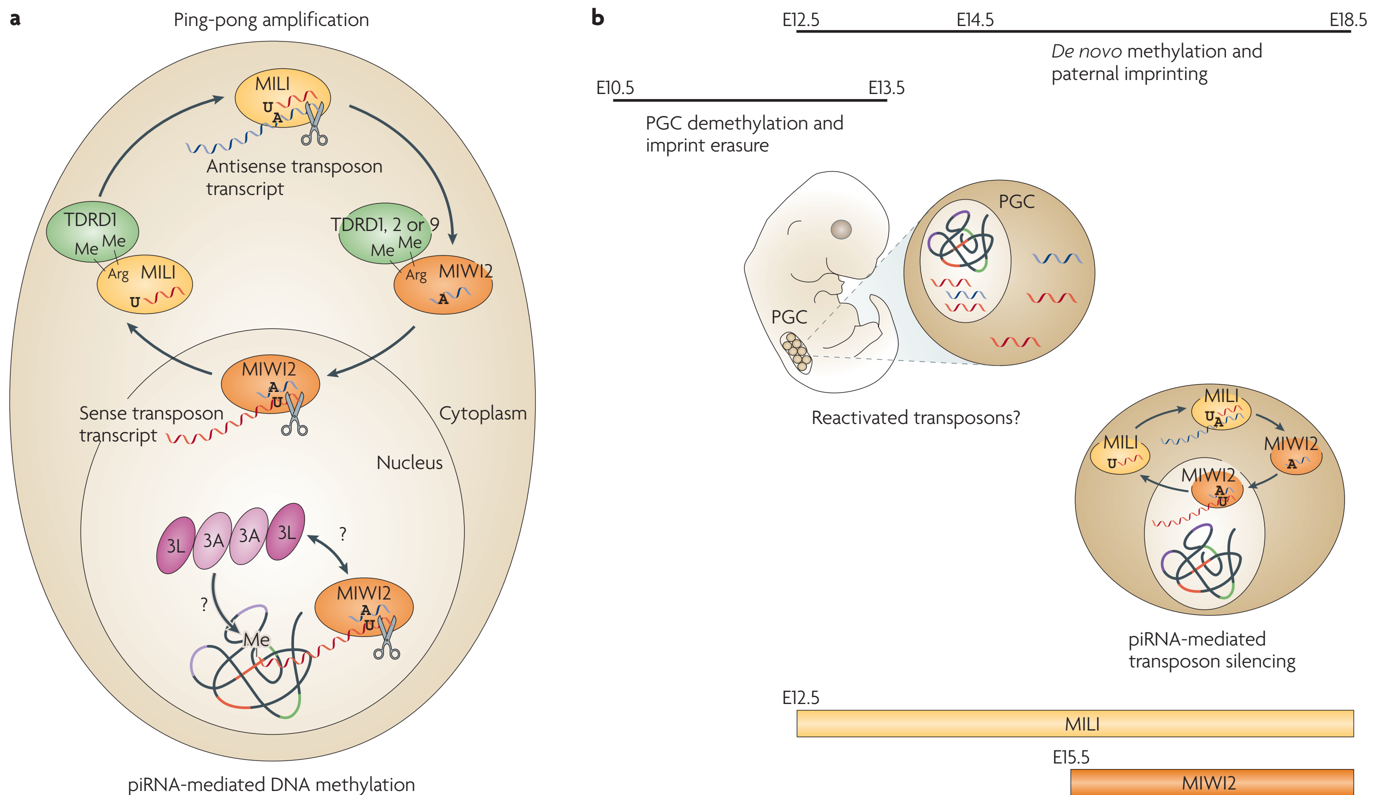

Figure 5 | **Piwi-interacting RNAs and male gametogenesis. a** | The ping-pong amplification model. In mammals, as in flies, Piwi-interacting RNAs (piRNAs) are proposed to arise through a ping-pong amplification cycle that produces primary piRNAs with a 5′ uridine (U) and secondary piRNAs with an adenine (A) at position 10. In mammals, transposon transcripts, mainly sense oriented, are the presumed substrates for primary piRNA production[90]. Cleavage of these transcripts produces primary piRNAs that are proposed to preferentially associate with cytoplasmic MILI (also known as PIWIL2)[90]. MILI, which is bound with sense piRNAs, cleaves antisense transcripts, producing secondary piRNAs that preferentially associate with cytoplasmic and nuclear MIWI2 (also known as PIWIL4)[90]. Nuclear MIWI2, which is bound with antisense piRNAs, cleaves sense transposon transcripts, producing more primary piRNAs. piRNA complexes are also proposed to guide DNA methylation to homologous genomic loci by potentially interacting with nascent transposon transcripts and directly or indirectly recruiting *de novo* methyltransferases, possibly complexes of DNA methyltransferase 3A (DNMT3A, shown as 3A) and DNMT3-like (DNMT3L, shown as 3L). MILI and MIWI2 contain symmetrical dimethylarginines and interact with Tudor domain-containing (TDRD) proteins. **b** | Model for transposon silencing during male gametogenesis. Genome-wide demethylation (embryonic day (E)10.5–E13.5) in primordial germ cells (PGCs) erases DNA imprints and could briefly reactivate transposons. *De novo* methylation and paternal imprinting are observed in testes from E14.5 until birth. Consistent with prior transposon expression, piRNAs bound to MILI and MIWI2 (expressed by E12.5 and E15.5, respectively) are enriched for transposon sequences during this time period and are proposed to facilitate targeted re-establishment of DNA methylation at transposons. Although less well-studied then in male germ cells, Piwi Argonautes are expressed in the female germ cells of flies, mammals[84], frogs[180] and silkworms[181], and piRNAs are present in the ovaries of flies[182] and silkworms[181,183] and in the oocytes of frogs[184]. Arg, arginine; Me, methyl group.

that is similar to the ping-pong amplification model initially proposed in *D. melanogaster*[91,92] and that they are indeed involved in the post-transcriptional silencing of transposons (FIG. 5a).

In mammals, decreases in DNA methylation and increases in expression were observed at several TEs in two Piwi clade mutants, *Mili* (also known as *Piwil2*) and *Miwi2* (also known as *Piwil4*)[89,93], which suggests that piRNAs silence transposons at both the transcriptional and post-transcriptional levels. However, these initial methylation studies were carried out at a developmental stage that was many mitoses after the establishment of DNA methylation, which occurs after a wave of genome-wide demethylation in PGCs (FIG. 5b). Therefore it could not be determined whether methylation defects were occurring at the level of maintenance or *de novo* methylation. Two recent studies provide compelling evidence that piRNAs are indeed involved in *de novo* methylation by showing that DNA methylation defects in *Mili* mutants occur at the stage in development at which *de novo* methylation in male germ cells is observed[88] and that piRNA populations from this time period are highly enriched in transposon sequences[88,90]. Aravin *et al.*[90] further show that piRNAs are present in DNMT3L mutants, which suggests that the piRNA pathway acts upstream of *de novo* methylation.

© 2010 Macmillan Publishers Limited. All rights reserved

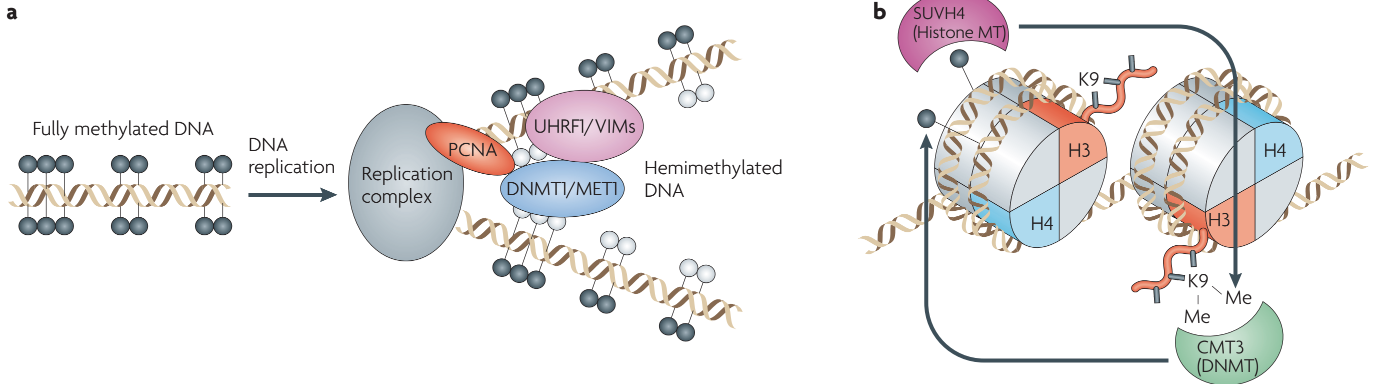

Figure 6 | **Maintenance of DNA methylation in plants and mammals. a** | Model depicting the maintenance of CG methylation during replication. DNA methyltransferase 1 (DNMT1) is proposed to be recruited to replication foci through interactions with ubiquitin-like plant homeodomain and RING finger domain 1 (UHRF1) — a SET- or RING-associated (SRA) domain protein that specifically interacts with hemimethylated DNA — and with proliferating cell nuclear antigen (PCNA). After being recruited, DNMT1 functions to maintain methylation patterns by restoring the hemimethylated DNA to a fully methylated state. In plants, DNA METHYLTRANSFERASE 1 (MET1, also known as DMT1) and the VARIANT IN METHYLATION (VIM, also known as ORTHRUS) family of SRA domain proteins, which are homologues of DNMT1 and UHRF1, respectively, are likely to function in a similar manner to maintain CG methylation patterns. Black and white circles represent methylated and unmethylated cytosines, respectively. **b** | Model depicting the maintenance of CHG methylation in plants. A reinforcing loop of DNA and histone methylation is proposed to maintain CHG methylation in plants. The CHROMOMETHYLASE 3 (CMT3) DNA methyltransferase maintains methylation in the CHG context, which is recognized by the SRA domain of the SUPPRESSOR OF VARIEGATION 3-9 HOMOLOGUE 4 (SUVH4, also known as KYP) histone methyltransferase (histone MT). SUVH4 catalyses histone 3 lysine 9 dimethylation (H3K9me2), a modification that is required for the maintenance of CHG methylation, and the chromodomain of CMT3 binds methylated H3 tails.

As in models for siRNA- and piRNA-mediated transposon control in *A. thaliana* and *D. melanogaster*, respectively, the demethylation in PGCs may reveal TEs with the potential to be expressed when hypomethylated. Such expression may lead to the production of piRNAs and the targeting of DNA methylation to homologous sequences throughout the genome (FIG. 5). It has been suggested that Piwi–piRNA complexes could interact with nascent transposon transcripts and directly recruit *de novo* methyltransferases (FIG. 5a). However, preliminary studies failed to show an interaction between Piwi Agos and DNMT3 proteins[90]. Alternatively, this recruitment could be indirect, first involving the recruitment of chromatin modifiers, which catalyse modifications that subsequently recruit the DNA methyltransferases[94].

Recently, it was found that Piwi family members in mice[95–98], *D. melanogaster* and *Xenopus laevis* contain symmetrical dimethylarginine modifications[96]. Methylated arginines can be recognized by Tudor domains, and purification of MILI-, MIWI (also known as PIWIL1)- or MIWI2-containing complexes showed interactions with various Tudor domain-containing (TDRD) proteins[95,97–100]: TDRD1 was found to interact with MILI[95,97–100], and TDRD1, TDRD2 (also known as TDRKH) and TDRD9 interacted with MIWI2 (REF. 97) (FIG. 5a). Like MILI, TDRD1 is required for DNA methylation and transposon silencing in mouse germ cells[97,98]. In *Tdrd1* mutants, there are changes in the profiles of MILI-bound piRNAs (which contain a higher proportion of non-transposon sequences[97,98]) and of MIWI2-bound piRNAs (which contain a lower proportion of antisense piRNAs[97]), which may explain the observed transposon reactivation.

Tudor domain
A conserved protein motif that is able to recognize symmetrically dimethylated arginines.

Maintenance of DNA methylation patterns

Once established, global DNA methylation patterns must be stably maintained to ensure that transposons remain in a silenced state and to preserve cell type identity.

Maintenance of CG methylation. In mammals, DNA methylation is maintained by DNMT1 (FIG. 6a). This methyltransferase is associated with replication foci and functions to restore hemimethylated DNA generated during DNA replication to the fully methylated state[10]. Early studies showed that DNMT1 is recruited to replication foci by an interaction with the proliferating cell nuclear antigen (PCNA) component of the replication machinery[101]. However, disruption of this interaction only resulted in a minor reduction in DNA methylation[102–104]. Recently, it was shown that DNMT1 also interacts with another chromatin-associated protein, ubiquitin-like plant homeodomain and RING finger domain 1 (UHRF1) (FIG. 1) and that UHRF1 is required for the association of DNMT1 with chromatin[105,106]. Studies showing that mutations in UHRF1 cause severe decreases in DNA methylation[105,106] and that the SET- or RING-associated (SRA) domain of UHRF1 specifically binds to hemimethylated CG dinucleotides[105,107–110] have led to a model in which UHRF1 recruits DNMT1 to hemimethylated DNA[105,106]. In addition, UHRF1 interacts with DNMT3A and DNMT3B[111], which may suggest a role for UHRF1 in *de novo* methylation. Maintenance

© 2010 Macmillan Publishers Limited. All rights reserved

of DNA methylation also requires the chromatin-remodelling factor lymphoid-specific helicase 1 (LSH1, also known as HELLS)[112,113], although the mechanism through which LSH1 functions in DNA methylation remains unknown.

In plants, genetic analyses have shown that homologues of the above-mentioned mammalian proteins (FIG. 1) — MET1 (REF. 79), the VARIANT IN METHYLATION (VIM, also known as ORTHRUS) family of SRA domain proteins[114,115] and DDM1 (REFS 79,116) — are required for the maintenance of CG methylation, which suggests that plants and mammals maintain CG methylation in a similar manner. However, further work is needed to determine mechanistically whether these proteins are indeed functioning in a similar way to their counterparts in mammals. One known difference between plants and mammals is that mutations in DDM1, but not LSH1, cause a decrease in H3K9 methylation[117,118], a modification that is highly correlated with DNA methylation and silencing in plants[119] and mammals[5].

In *A. thaliana*, approximately one-third of genes have CG methylation in their coding region, which is maintained by MET1 (REFS 8,9,120,121). Unlike methylation at transposons, CG methylation in gene bodies does not seem to cause silencing, as these genes tend to be moderately expressed in many tissues[9,121]. Nonetheless, the expression of some body-methylated genes is upregulated in *met1* mutants[121], and genes with high or low expression tend to lack body methylation, which suggests an interplay between transcription and body methylation. The presence of body CG methylation at some genes has also been reported in other invertebrate organisms, suggesting it may be a common feature of eukaryotic genomes[6]. Initial studies in *A. thaliana* postulated that body methylation might suppress the production of antisense transcripts from cryptic promoters[121,122]. However, increases in antisense transcripts in *met1* mutants were found to be rare and uncorrelated with body methylated genes[9]. Therefore, the function of body methylation remains poorly understood.

Maintenance of non-CG methylation in plants. CHG methylation is thought to be maintained by a reinforcing loop involving histone and DNA methylation[123] (FIG. 6b). Genome-wide profiling of H3K9 and DNA methylation showed that these marks are highly correlated[119]. Furthermore, loss of CMT3, the DNA methyltransferase that is largely responsible for maintaining CHG methylation[124,125], or SUPPRESSOR OF VARIEGATION 3-9 HOMOLOGUE 4 (SUVH4, also known as KYP), the histone methyltransferase that is largely responsible for H3K9 dimethylation[126–128], results in a dramatic decrease in DNA methylation[126,127]. Two other H3K9 histone methyltransferases, SUVH5 and SUVH6, also contribute to global levels of CHG methylation[129,130]. The observed interdependence of DNA and histone modifications could arise from the multidomain structure of CMT3 and SUVH4 (FIG. 1). In addition to its histone methyltransferase domain, SUVH4 possesses an SRA domain that specifically binds CHG methylation[123], which suggests that CHG methylation recruits SUVH4. In turn, CMT3 possesses a chromodomain that binds methylated histone H3 tails[71], suggesting that histone methylation by SUVH4 may recruit CMT3. Such crosstalk between DNA and histone methylation is also observed in mammals and, in many cases, the connection between these modifications seems to involve protein–protein interactions between the histone and DNA methyltransferases[5]. Whether direct protein interactions between CMT3 and SUVH4 occur and aid in maintaining CHG methylation in plants is unknown.

Asymmetric DNA methylation is maintained by constant *de novo* methylation by DRM2 and RdDM. However, at some loci CHH methylation is controlled by CMT3 and DRM2 (REF. 131). Like the maintenance of CG and CHG methylation, RdDM requires proteins with SRA domains. SUVH9 and SUVH2 possess SRA domains that preferentially bind CHH and CG methylation, respectively, and these proteins are thought to act late in the RdDM pathway (FIG. 2), possibly functioning to recruit or retain DRM2 at loci targeted for methylation[132].

DNA demethylation

Although in most cases DNA methylation is a stable epigenetic mark, reduced levels of methylation are observed during development in plants and mammals. This net loss of methylation can either occur passively, by replication in the absence of functional maintenance methylation pathways, or actively, by removing methylated cytosines.

Active demethylation in plants. Active demethylation is achieved in plants by DNA glycosylase activity, probably in combination with the base excision repair (BER) pathway[133,134] (FIG. 7). DEMETER (DME)[135] and REPRESSOR OF SILENCING 1 (ROS1)[136] are the founding members of a family of DNA glycosylases in *A. thaliana* that also includes DEMETER-LIKE 2 (DML2) and DML3 (REFS 137,138). The *A. thaliana* glycosylases recognize and remove methylated cytosines from dsDNA oligonucleotides, irrespective of sequence context *in vitro*[137–141], and *in vivo*, mutations in these genes cause increased DNA methylation in all sequence contexts at specific genomic loci[120,136–139,142,143]. In general, DNA glycosylases involved in BER recognize and remove mutagenic substrates, including oxidized and alkylated bases, as well as T/G mismatches, which are often generated by deamination of methylated cytosines[144]. The DME/ROS1 family of glycosylases have homology to the helix–hairpin–helix-Gly-Pro-Asp (HhH-GPD) class of DNA glycosylases. They are bifunctional enzymes that can break both the *N*-glycosidic bond, removing the base, and the DNA backbone[144,145] (FIG. 7). In mammals, the resulting single-nucleotide gap is then acted on by DNA polymerase-β and DNA ligase IIIα to repair the DNA through the short-patch BER pathway. Homologues of these enzymes have not been identified in plants, raising the possibility that plants use enzymes that are involved in the long-patch BER pathway[144].

Despite similar substrate specificity, the DME/ROS1 glycosylases have distinct biological roles, with

Base excision repair
A cellular mechanism that repairs damaged DNA and is initiated by the activity of DNA glycosylases.

© 2010 Macmillan Publishers Limited. All rights reserved

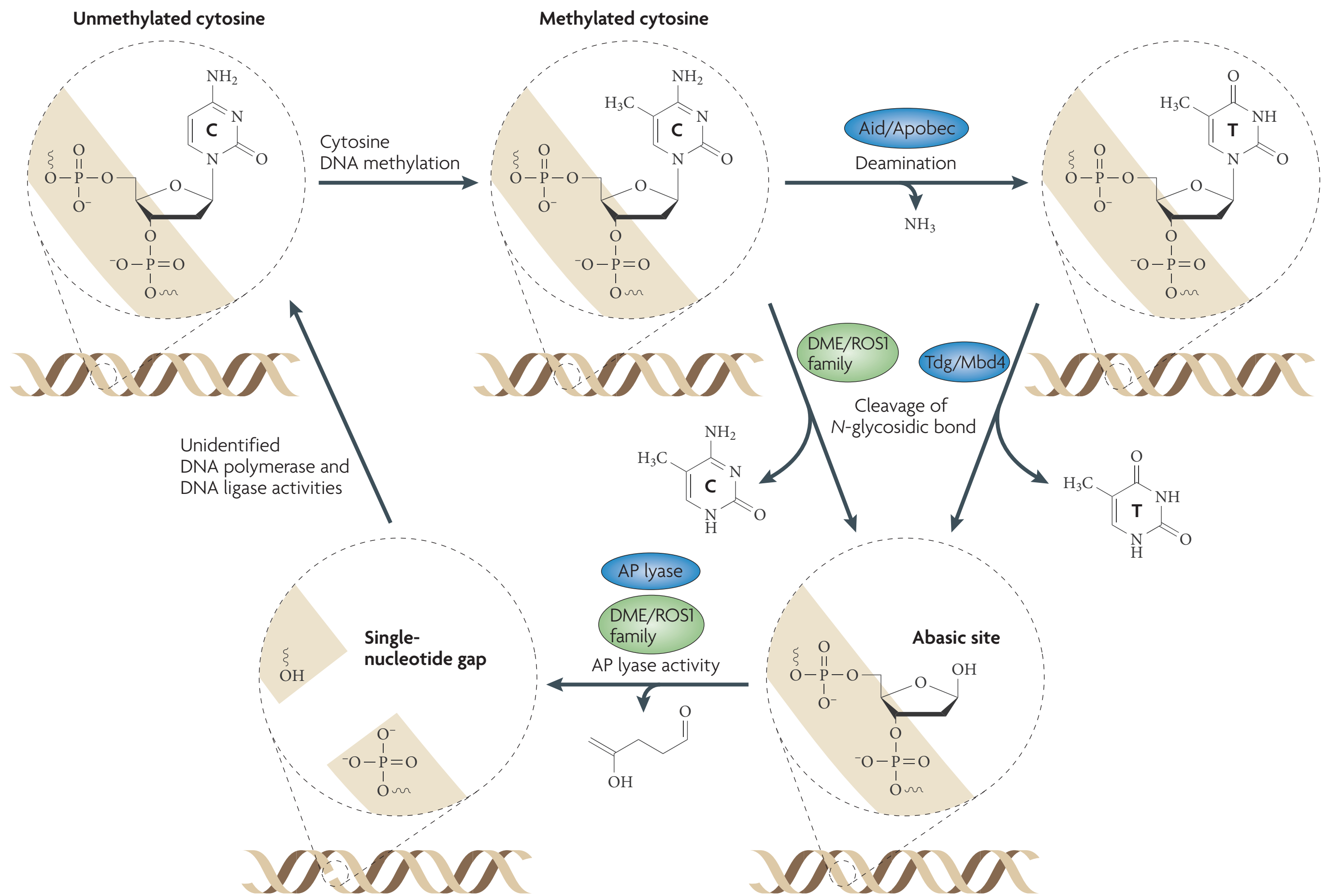

Figure 7 | **Active DNA demethylation through DNA glycosylase activity and base excision repair.** In *Arabidopsis thaliana* (green proteins), methylated (CH_3) cytosine (bold C) bases are removed by the DEMETER (DME)/REPRESSOR OF SILENCING 1 (ROS1) family of bifunctional 5-methylcytosine glycosylases. First, the methylated cytosine base is released by cleavage of the *N*-glycosidic bond, generating an abasic site. Next, the phosphodiester linkage is broken both 3′ and 5′ of the abasic site through apyrimidic (AP) lysase activity, generating a single-nucleotide gap in the DNA. The DNA is then proposed to be repaired by unknown DNA polymerase and DNA ligase activities, resulting in a net loss of cytosine methylation. In zebrafish and mammals, no efficient 5-methylcytosine glycoslases have been identified. However, in zebrafish and mammals (blue proteins), it has been proposed that the activation-induced cytosine deaminase (Aid)/apolipoprotein B mRNA-editing enzyme (Apobec) family of deaminases first convert methylated cytosines into thymines (bold T), generating thymine/guanine (T/G) mismatches. These mismatches could be recognized by the methyl-CpG-binding domain 4 (Mbd4) glycosylase or thymine DNA glycosylase (Tdg), resulting in removal of the thymine base and generation of an abasic site. Unlike the bifunctional DME/ROS1 glycosylases, Mbd4 and Tdg are monofunctional DNA glycosylases, so other unidentified proteins are likely to be required to provide the AP lyase activity to remove the sugar ring to generate a single-nucleotide gap. As in *A. thaliana*, this substrate is proposed to be repaired by unidentified DNA polymerase and DNA ligase activities.

DME functioning during gametogenesis to establish imprinting[77] and the other family members functioning in vegetative tissues, possibly to counteract robust DNA methylation by the RdDM pathway[137,142,143,146]. Unlike in mammals, in which imprinting is established by the addition of methylation in an allele-specific manner and is observed in both the placenta and the developing embryo, in plants imprinting is restricted to the endosperm, the plant equivalent of the placenta[77], and is established by allele-specific removal of DNA methylation by DME in the central cell before fertilization, such that only the maternal allele is expressed in the resulting endosperm[77] (FIG. 4C).

Until recently, DME was only known to activate the maternal alleles of three genes: *MEDEA* (*MEA*), *FLOWERING WAGENINGEN* (*FWA*) and *FERTILIZATION INDEPENDENT SEED 2* (*FIS2*), whereas in mammals ~80 imprinted genes have been identified (see 'Genomic Imprinting' at the Medical Research Council Harwell website). However, the genome-wide decrease in CG methylation observed in the endosperm was found to be largely dependent on DME, which suggests that this glycosylase acts as a global regulator of DNA methylation[75,76]. By comparing DNA methylation levels in embryo and endosperm tissues, Gehring *et al.*[76] were able to identify DMRs and confirm

© 2010 Macmillan Publishers Limited. All rights reserved

parent-of-origin expression of five genes, doubling the number of known imprinted genes in *A. thaliana*. Approximately 40 other candidate imprinted genes were identified[76], and although imprinting at these loci remains to be experimentally verified, these findings suggest that the number of imprinted loci may be more similar in plants and mammals than previously thought.

Recent characterization of Pol IV-dependent siRNA populations, which are generated from dispersed loci corresponding to >1% of the *A. thaliana* genome[18,25], suggests that they may also be maternally imprinted[81]. After reciprocal crosses between two *A. thaliana* ecotypes, siRNAs from the resultant silique tissue (which contains the developing embryos) were sequenced, and nearly all of the Pol IV-dependent siRNAs that could be uniquely distinguished between the two ecotypes were maternal in origin[81]. What causes these loci to be maternally imprinted, whether this imprinting requires DME and what function this massive extent of imprinting serves remain unknown. One hypothesis is that such maternal imprinting would allow recognition of self from non-self and have a suppressive effect on hybrids[81]. For example, maternal siRNAs could fail to target and silence a TE present in another *A. thaliana* ecotype or could target and silence a functional gene. Indeed, in *D. melanogaster*, piRNAs corresponding to TEs and other repeat sequences are maternally inherited — much like the Pol IV-dependent class of siRNAs in *A. thaliana* — and if female flies lacking piRNAs to a particular TE are crossed with male flies harbouring that element, the offspring are largely inviable[147].

ROS1, DML2 and DML3, unlike DME, are expressed in vegetative tissues[136–138]. Comparative analysis of methylation patterns in *ros1*, *dml2* and *dml3* single mutants showed that these glycosylases function redundantly, although some locus specificity was observed[137]. In a *ros1 dml2 dml3* triple mutant, 179 loci with increased methylation relative to wild-type controls were identified, despite the fact that no global increase in methylation was observed[137]. These loci are enriched for transposons, repetitive DNA elements and siRNA-generating loci. In addition, ~80% are near or overlap annotated genes, and the increase in DNA methylation at genes is primarily located at their 5′ and 3′ ends[137,143]. Together, these studies suggest that ROS1, DML2 and DML3 act both at normally silenced loci (that is, transposons) and at the boundaries between euchromatin and heterochromatin (that is, genes that reside in or near heterochromatic environments). At such boundaries, these glycosylases may function to protect genes that are targeted for methylation through RdDM from silencing by removing DNA methylation. At normally silenced loci, they may be required to maintain a silenced but readily adaptable state[142,143], and perhaps this is important to allow efficient reactivation of transposons during gametogenesis.

The mechanism (or mechanisms) through which the DME/ROS1 glycosylases are targeted to specific loci to carry out DNA demethylation is (are) unknown. These *A. thaliana* glycosylases are quite different from most other glycosylases: they are much larger and contain two conserved domains of unknown function[144] (FIG. 1). Whether these domains are required for targeting demethylation remains unknown. For ROS1, it has been proposed that ROS3 — a protein that binds small ssRNAs (21–26 nt) in a sequence-specific manner and acts in the same demethylation pathway as ROS1 — may be involved in targeting ROS1 to certain loci[134,148]. The recent finding that DME participates in genome-wide demethylation suggests that its targeting may be less specific.

Active demethylation in zebrafish and mammals. In mammals, genome-wide decreases in DNA methylation are observed in PGCs and on the paternal genome of the zygote[47,48]. Although mechanisms for passive demethylation seem to have a role in achieving the observed hypomethylated states, the timing of methylation loss suggests that active mechanisms may also be required[149–153]. Notably, DNA methylation imprints in the zygote and pre-implantation embryo — but not in the PGCs — are resistant to demethylation, and several proteins, including Stella[154], zinc finger protein 57 (ZFP57)[155], and methyl-CpG-binding domain 3 (MBD3)[156] are proposed to protect specific imprinted loci from demethylation[157,158].

Proteins that are orthologous to the DME/ROS1 family of glycosylases have not been identified in mammals, and the existence of other enzymes that can directly remove methylated cytosines is controversial[145,159]. However, early work in mammals showed that activation-induced cytosine deaminase (AID) and apolipoprotein B mRNA-editing enzyme 1 (APOBEC1) are expressed in cells that are thought to undergo active DNA demethylation and catalyse 5-methylcytosine deamination, which results in T/G mismatches. This finding led to a model for demethylation that involves the coupling of 5-methylcytosine deaminase and thymine DNA glycosylase activities[160]. Such a model is supported by recent findings in zebrafish (*Danio rerio*)[161]. Rai *et al.*[161] showed that Mbd4, an HhH-GPD thymine glycosylase that is related to the *A. thaliana* DME/ROS1 family of glycosylases and that has active mammalian homologues[162,163], is involved in demethylation during zebrafish development. In addition, they showed that three proteins belonging to the AID/APOBEC family (FIG. 1), Aid, Apobec2a and Apobec2b, are involved in DNA demethylation[161].

By overexpressing the Aid, Apobec2a and Apobec2b deaminases in the absence or presence of overexpressed human MBD4 in zebrafish embryos, Rai *et al.*[161] found that loss of DNA methylation, as well as deamination of methylated cytosines, seems to be limited by the abundance of the MBD4 glycosylase, which suggests that mechanisms are in place to ensure that deamination does not occur unless the resultant T/G mismatch can be efficiently removed. This is an important finding as previous models for DNA demethylation involving 5-methylcytosine deamination have been discounted owing to the large mutagenic potential of an uncoupled deamination step. Growth arrest and DNA-damage-inducible protein 45α (Gadd45α) may aid in coupling these processes. Gadd45α interacts with MBD4 and with Aid and Apobec family members

Ecotype
A genetically distinct population within a widely spread species.

Silique
An elongated seed capsule that is formed after fertilization.

Hybrids
Offspring that are produced by crossing two different populations within a single species.

Zygote
A single diploid cell formed by the union of two haploid germ cells.

© 2010 Macmillan Publishers Limited. All rights reserved

in vitro and stimulates demethylation of plasmid DNA transfected into zebrafish embryos, as well as the association of MBD4 and Aid with methylated DNA[161]. In addition, MBD4 possesses a methyl-binding domain (FIG. 1) that may aid in recruitment of the demethylation machinery to methylated DNA. Together, these findings suggest a model (FIG. 7) in which tight coupling of 5-methylcytosine deamination by Aid and Apobec to T/G mismatch repair by MBD4 results in DNA demethylation[161]. Importantly, recent genetic evidence in mice has shown that AID is needed for DNA demethylation in PGCs[164]. Therefore, in the case of zebrafish and mammals, there seems to be an additional deamination step in the demethylation pathway compared with the pathway in plants. However, the downstream events that lead to a net loss of cytosine methylation may be similar (FIG. 7).

In mammals, recent data presented by Kim *et al.*[165] suggest that MBD4 may be able to directly remove methylated cytosines at the cytochrome p450 27B1 (CYP27B1) promoter upon hormone-induced MBD4 phosphorylation. Whereas previous *in vitro* analysis of MBD4 glycosylase activity revealed a strong preference for T/G mismatches over methylated cytosines[166], Kim *et al.*[165] found that upon phosphorylation, the activity of MBD4 on methylated cytosines is stimulated. *In vivo,* the observed decrease in methylation at the CYP27B1 promoter can occur in the absence of DNA replication — which suggests an active mechanism — and is dependent on the presence of a catalytically active MBD4 protein that contains the serine residues that are targeted for phosphorylation[165]. These recent findings need to be confirmed, and whether such a mechanism for the direct removal of methylated cytosines could account for DNA demethylation on a larger scale remains unknown.

A role for the 5-hydroxymethylcytosine modification in mammalian DNA demethylation has also been proposed. 5-hydroxymethylcytosine is present in mouse Purkinje neurons, brain tissue and ES cells[167,168] and can be generated from methylated cytosines through hydroxylation of the methyl group. Findings that ten-eleven translocation 1 (TET1) is able to catalyse the conversion of methylated cytosines into 5-hydroxymethylcytosines *in vitro* and that targeted depletion of TET1 by RNAi in mouse ES cells results in decreased levels of 5-hydroxymethylcytosine[167] have led to the hypothesis that TET1 and possibly other TET family members generate 5-hydroxymethylcytosines. As proteins that are known to interact with methylated cytosines — namely methyl-CpG-binding protein 2 (MECP2) and DNMT1 (REFS 169,170) — have reduced affinity for 5-hydroxymethylcytosine *in vitro*, potential roles for this modification in the regulation of chromatin structure and in passive DNA demethylation have been proposed[167,168]. It has been suggested that 5-hydroxymethylcytosine could be an intermediate in an active DNA demethylation pathway involving DNA repair[167], as 5-hydroxymethylcytosine-specific DNA glycosylase activity has been reported in mammalian extracts[171].

Passive demethylation in plants and mammals. In addition to active DNA demethylation by DME in the central cell of the female gametophyte, passive losses of methylation are likely to contribute to the overall decrease in methylation observed in the endosperm. Using a reporter driven by the MET1 promoter, Jullien *et al.*[172] showed that MET1 expression levels are reduced during female gametogenesis. They further showed that MULTICOPY SUPPRESSOR OF IRA1 (MSI1) and RETINOBLASTOMA-RELATED 1 (RBR1) are important for the observed repression of MET1 (REF. 172) and conclude that MET1 is transcriptionally repressed during gametogenesis by MSI1, probably through the retinoblastoma pathway and RBR1. Notably, MSI1 and RBR1 are also required for maternal expression of the imprinted *FIS2* and *FWA* genes[172], which suggests that passive DNA demethylation resulting from decreased MET1 levels and active demethylation by DME are working together to allow activation of imprinted genes. In mammals, DNMT1 expression also seems to be regulated through the retinoblastoma pathway, which involves retinoblastoma 1 (RB1) and retinoblastoma-binding protein 4 (RBBP4), which are homologues of the plant RBR1 and MSI1 proteins[173–176]. Although a direct role for these proteins in imprinting has not been established in mammals, several observations suggest that their function may be conserved[172].

The idea that passive and active demethylation pathways are working together is appealing as it fits well with several other observations. First, DME is more active on hemimethylated DNA — which would be enriched following replication in the absence of MET1 — than on fully methylated DNA *in vitro*[139,140]. Second, enrichment in hemimethylated DNA should decrease the chance of detrimental double-strand breaks (DSBs), which are predicted to arise from the removal of methylated cytosines in symmetric contexts by DME. Indeed, DME is inefficient at removing methylated cytosines across from abasic sites, which should also reduce the production of DSBs[139]. Finally, in addition to decreasing the workload for DME and the risk of DSBs, downregulation of MET1 might also function to ensure that hemimethylated CG sites generated by DME activity on one strand of the DNA are not efficiently restored to the fully methylated state by active targeting of DNA methyltransferases through interactions with SRA domain-containing proteins.

In mammals, in addition to the reported active demethylation of the paternal genome of the zygote[151–153], passive demethylation is proposed to occur during preimplantation development of the embryo[47,177]. This passive decrease in methylation is likely to be due to exclusion of the oocyte-specific form of DNMT1, DNMT1o, from nuclei until just before blastocyst formation[178,179]. This is reminiscent of the observed localization pattern of DDM1 in pollen, in which DDM1 is observed in the sperm cells but not in the vegetative nucleus[74]. Therefore, plants and mammals seem to use similar mechanisms for passive DNA demethylation, including transcriptional repression of DNA methyltransferases and exclusion of the methylation machinery from the nucleus.

Blastocyst
An embryonic stage that is characterized by the first definitive lineages.

© 2010 Macmillan Publishers Limited. All rights reserved

Conclusions

Plants and animals use similar mechanistic strategies for controlling DNA methylation. Both use small-RNA-based pathways to target DNA methylation to transposons, both require methyl-DNA-binding proteins to maintain DNA methylation patterns, and both show intimate connections between histone and DNA methylation marks. Furthermore, a growing body of evidence suggests that active demethylation may occur in animals through DNA glycosylases and the BER pathway, as has been documented in *A. thaliana*.

Several pathways that are unique to plants or mammals have also been elucidated and are likely to contribute to the observed differences in global methylation patterns between plants and mammals. For example, in mammals DNA methylation is not restricted to repeat elements and the DNA methylation machinery is recruited to specific genomic loci through interactions with chromatin marks as well as through interactions with the chromatin-modifying enzymes themselves. In addition, structural studies of the mammalian *de novo* methyltransferases suggest a mechanism in which a DNMT3A–DNMT3L tetramer may oligomerize on DNA, potentially leading to the nearly global methylation status of the mammalian genome. In plants, in which DNA methylation occurs in all sequence contexts, a plant-specific methyltransferase, CMT3, is required for the maintenance of CHG methylation, and maintenance of CHH methylation is achieved through constant *de novo* methylation by DRM2.

Despite the significant advances in our understanding of DNA methylation pathways, several key questions remain, especially surrounding the issue of targeting. How DNA methyltransferases are targeted by siRNAs and piRNAs in plants and mammals, respectively, remains elusive. In terms of DNA demethylation, whether DME is specifically targeted to many sites throughout the genome during gametogenesis or whether it non-selectively removes methylation remains unclear. Similarly, whether demethylation by the other DME/ROS1 family members is specifically directed to certain loci or whether the observed methylation pattern simply reflects a balance between the RdDM and demethylation pathways requires further investigation. Gaining further insights into mammalian DNA demethylation pathways and understanding how demethylation is targeted will be key challenges for future research.

1. Ehrlich, M. *et al.* Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res.* **10**, 2709–2721 (1982).
2. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
3. Ramsahoye, B. H. *et al.* Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl Acad. Sci. USA* **97**, 5237–5242 (2000).
4. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
5. Cedar, H. & Bergman, Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nature Rev. Genet.* **10**, 295–304 (2009).
6. Suzuki, M. M. & Bird, A. DNA methylation landscapes: provocative insights from epigenomics. *Nature Rev. Genet.* **9**, 465–476 (2008).
7. Henderson, I. R. & Jacobsen, S. E. Epigenetic inheritance in plants. *Nature* **447**, 418–424 (2007).
8. Cokus, S. J. *et al.* Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452**, 215–219 (2008).
This study, along with reference 120, provided single-nucleotide resolution, genome-wide mapping of DNA methylation patterns in *A. thaliana*.
9. Zhang, X. *et al.* Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* **126**, 1189–1201 (2006).
10. Kim, J. K., Samaranayake, M. & Pradhan, S. Epigenetic mechanisms in mammals. *Cell. Mol. Life Sci.* **66**, 596–612 (2009).
11. Goll, M. G. & Bestor, T. H. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* **74**, 481–514 (2005).
12. Cheng, X. & Blumenthal, R. M. Mammalian DNA methyltransferases: a structural perspective. *Structure* **16**, 341–350 (2008).
13. Chan, S. W., Henderson, I. R. & Jacobsen, S. E. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nature Rev. Genet.* **6**, 351–360 (2005).
14. Wassenegger, M., Heimes, S., Riedel, L. & Sanger, H. L. RNA-directed *de novo* methylation of genomic sequences in plants. *Cell* **76**, 567–576 (1994).
15. Matzke, M., Kanno, T., Daxinger, L., Huettel, B. & Matzke, A. J. RNA-mediated chromatin-based silencing in plants. *Curr. Opin. Cell Biol.* **21**, 367–376 (2009).
16. Huettel, B. *et al.* RNA-directed DNA methylation mediated by DRD1 and Pol IVb: a versatile pathway for transcriptional gene silencing in plants. *Biochim. Biophys. Acta* **1769**, 358–374 (2007).
17. Pikaard, C. S., Haag, J. R., Ream, T. & Wierzbicki, A. T. Roles of RNA polymerase IV in gene silencing. *Trends Plant Sci.* **13**, 390–397 (2008).
18. Mosher, R. A., Schwach, F., Studholme, D. & Baulcombe, D. C. PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proc. Natl Acad. Sci. USA* **105**, 3145–3150 (2008).
19. Lahmy, S. *et al.* PolV(PolIVb) function in RNA-directed DNA methylation requires the conserved active site and an additional plant-specific subunit. *Proc. Natl Acad. Sci. USA* **106**, 941–946 (2009).
20. Ream, T. S. *et al.* Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol. Cell* **33**, 192–203 (2009).
21. Huang, L. *et al.* An atypical RNA polymerase involved in RNA silencing shares small subunits with RNA polymerase II. *Nature Struct. Mol. Biol.* **16**, 91–93 (2009).
22. He, X. J. *et al.* NRPD4, a protein related to the RPB4 subunit of RNA polymerase II, is a component of RNA polymerases IV and V and is required for RNA-directed DNA methylation. *Genes Dev.* **23**, 318–330 (2009).
23. Bies-Etheve, N. *et al.* RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. *EMBO Rep.* **10**, 649–654 (2009).
24. Herr, A. J., Jensen, M. B., Dalmay, T. & Baulcombe, D. C. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**, 118–120 (2005).
25. Zhang, X., Henderson, I. R., Lu, C., Green, P. J. & Jacobsen, S. E. Role of RNA polymerase IV in plant small RNA metabolism. *Proc. Natl Acad. Sci. USA* **104**, 4536–4541 (2007).
26. Haag, J. R., Pontes, O. & Pikaard, C. S. Metal A and metal B sites of nuclear RNA polymerases Pol IV and Pol V are required for siRNA-dependent DNA methylation and gene silencing. *PLoS ONE* **4**, e4110 (2009).
27. Onodera, Y. *et al.* Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**, 613–622 (2005).
28. Pontier, D. *et al.* Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. *Genes Dev.* **19**, 2030–2040 (2005).
29. Chan, S. W., Zhang, X., Bernatavichute, Y. V. & Jacobsen, S. E. Two-step recruitment of RNA-directed DNA methylation to tandem repeats. *PLoS Biol.* **4**, e363 (2006).
30. Huettel, B. *et al.* Endogenous targets of RNA-directed DNA methylation and Pol IV in *Arabidopsis*. *EMBO J.* **25**, 2828–2836 (2006).
31. El-Shami, M. *et al.* Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. *Genes Dev.* **21**, 2539–2544 (2007).
32. Li, C. F. *et al.* An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. *Cell* **126**, 93–106 (2006).
33. Li, C. F. *et al.* Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in *Arabidopsis thaliana*. *PLoS Genet.* **4**, e27 (2008).
34. Wierzbicki, A. T., Haag, J. R. & Pikaard, C. S. Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* **135**, 635–648 (2008).
This study identifies RNA transcripts that are dependent on Pol V and DRD1 and shows they are required for DNA methylation and silencing but not for siRNA production.
35. Grewal, S. I. & Elgin, S. C. Transcription and RNA interference in the formation of heterochromatin. *Nature* **447**, 399–406 (2007).
36. He, X. J. *et al.* An effector of RNA-directed DNA methylation in *Arabidopsis* is an ARGONAUTE 4- and RNA-binding protein. *Cell* **137**, 498–508 (2009).
37. Wierzbicki, A. T., Ream, T. S., Haag, J. R. & Pikaard, C. S. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nature Genet.* **41**, 630–634 (2009).
38. Ausin, I., Mockler, T. C., Chory, J. & Jacobsen, S. E. IDN1 and IDN2 are required for *de novo* DNA methylation in *Arabidopsis thaliana*. *Nature Struct. Mol. Biol.* **16**, 1325–1327 (2009).
This study identifies IDN2 as a novel component of the RdDM pathway. The authors show that IDN2 encodes an RNA-binding protein that can interact with dsRNA duplexes containing 5′ overhangs *in vitro*.
39. Zheng, B. *et al.* Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. *Genes Dev.* **23**, 2850–2860 (2009).
40. Kanno, T. *et al.* RNA-directed DNA methylation and plant development require an IWR1-type transcription factor. *EMBO Rep.* **11**, 65–71 (2009).

© 2010 Macmillan Publishers Limited. All rights reserved

41. He, X. J. *et al.* A conserved transcriptional regulator is required for RNA-directed DNA methylation and plant development. *Genes Dev.* **23**, 2717–2722 (2009).
42. Kanno, T. *et al.* Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr. Biol.* **14**, 801–805 (2004).
43. Kanno, T. *et al.* A structural-maintenance-of-chromosomes hinge domain-containing protein is required for RNA-directed DNA methylation. *Nature Genet.* **40**, 670–675 (2008).
44. Smith, L. M. *et al.* An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in *Arabidopsis*. *Plant Cell* **19**, 1507–1521 (2007).
45. Kafri, T. *et al.* Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev.* **6**, 705–714 (1992).
46. Monk, M., Boubelik, M. & Lehnert, S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**, 371–382 (1987).
47. Reik, W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**, 425–432 (2007).
48. Sasaki, H. & Matsui, Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nature Rev. Genet.* **9**, 129–140 (2008).
49. Ooi, S. K. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature* **448**, 714–717 (2007).
Through the biochemical purification of DNMT3L, this study revealed that this protein copurifies with the *de novo* DNA methyltransferases as well as with the four core histone proteins. The authors further characterized the interaction between DNMT3L and histones and showed it to be specific for unmethylated H3K9 tails.
50. Otani, J. *et al.* Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L domain. *EMBO Rep.* **10**, 1235–1241 (2009).
51. Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A. & Cheng, X. Structure of Dnmt3a bound to Dnmt3L suggests a model for *de novo* DNA methylation. *Nature* **449**, 248–251 (2007).
This paper provides a crystal structure showing that the C-terminal regions of DNMT3A and DNMT3L form a tetrameric complex. Superimposition of a DNA fragment from another crystal structure onto this structure suggests the DNMT3A active sites are positioned approximately ten base pairs apart.
52. Fournier, C. *et al.* Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *EMBO J.* **21**, 6560–6570 (2002).
53. Delaval, K. *et al.* Differential histone modifications mark mouse imprinting control regions during spermatogenesis. *EMBO J.* **26**, 720–729 (2007).
54. Vu, T. H., Li, T. & Hoffman, A. R. Promoter-restricted histone code, not the differentially methylated DNA regions or antisense transcripts, marks the imprinting status of IGF2R in human and mouse. *Hum. Mol. Genet.* **13**, 2233–2245 (2004).
55. Yamasaki, Y. *et al.* Neuron-specific relaxation of *Igf2r* imprinting is associated with neuron-specific histone modifications and lack of its antisense transcript *Air*. *Hum. Mol. Genet.* **14**, 2511–2520 (2005).
56. Ciccone, D. N. *et al.* KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* **461**, 415–418 (2009).
57. Okitsu, C. Y. & Hsieh, C. L. DNA methylation dictates histone H3K4 methylation. *Mol. Cell. Biol.* **27**, 2746–2757 (2007).
58. Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genet.* **39**, 457–466 (2007).
59. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).
60. Mohn, F. *et al.* Lineage-specific polycomb targets and *de novo* DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* **30**, 755–766 (2008).
61. Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77–88 (2007).
62. Zhang, X., Bernatavichute, Y. V., Cokus, S., Pellegrini, M. & Jacobsen, S. E. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol.* **10**, R62 (2009).
63. Chotalia, M. *et al.* Transcription is required for establishment of germline methylation marks at imprinted genes. *Genes Dev.* **23**, 105–117 (2009).
64. Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H. & Tajima, S. DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. *J. Biol. Chem.* **279**, 27816–27823 (2004).
65. Chedin, F., Lieber, M. R. & Hsieh, C. L. The DNA methyltransferase-like protein DNMT3L stimulates *de novo* methylation by Dnmt3a. *Proc. Natl Acad. Sci. USA* **99**, 16916–16921 (2002).
66. Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. HhaI methyltransferase flips its target base out of the DNA helix. *Cell* **76**, 357–369 (1994).
67. Jurkowska, R. Z. *et al.* Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. *Nucleic Acids Res.* **36**, 6656–6663 (2008).
This paper provides a biophysical characterization of the tetrameric complex described in reference 51. The characterization shows the formation of nucleoprotein filaments and demonstrates that this complex generates periodic DNA methylation patterns *in vitro*.
68. Glass, J. L., Fazzari, M. J., Ferguson-Smith, A. C. & Greally, J. M. CG dinucleotide periodicities recognized by the Dnmt3a–Dnmt3L complex are distinctive at retroelements and imprinted domains. *Mamm. Genome* **20**, 633–643 (2009).
69. Ferguson-Smith, A. C. & Greally, J. M. Epigenetics: perceptive enzymes. *Nature* **449**, 148–149 (2007).
70. Gowher, H. & Jeltsch, A. Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. *J. Mol. Biol.* **309**, 1201–1208 (2001).
71. Lindroth, A. M. *et al.* Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. *EMBO J.* **23**, 4286–4296 (2004).
72. Zhao, Q. *et al.* PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. *Nature Struct. Mol. Biol.* **16**, 304–311 (2009).
73. Fabbrizio, E. *et al.* Negative regulation of transcription by the type II arginine methyltransferase PRMT5. *EMBO Rep.* **3**, 641–645 (2002).
74. Slotkin, R. K. *et al.* Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* **136**, 461–472 (2009).
This study investigates the expression pattern of transposons during plant development and shows they are expressed and active in pollen. However, this reactivation probably only occurs in the vegetative nucleus and not in sperm cells.
75. Hsieh, T. F. *et al.* Genome-wide demethylation of *Arabidopsis* endosperm. *Science* **324**, 1451–1454 (2009).
This study, along with reference 76, compares the methylation status of the endosperm and embryo tissues and shows that the endosperm is globally hypomethylated.
76. Gehring, M., Bubb, K. L. & Henikoff, S. Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science* **324**, 1447–1451 (2009).
In addition to showing a global decrease in methylation in the endosperm (see reference 75), this study identifies several novel imprinted genes in the *A. thaliana* genome.
77. Huh, J. H., Bauer, M. J., Hsieh, T. F. & Fischer, R. L. Cellular programming of plant gene imprinting. *Cell* **132**, 735–744 (2008).
78. Pina, C., Pinto, F., Feijo, J. A. & Becker, J. D. Gene family analysis of the *Arabidopsis* pollen transcriptome reveals biological implications for cell growth, division control, and gene expression regulation. *Plant Physiol.* **138**, 744–756 (2005).
79. Vongs, A., Kakutani, T., Martienssen, R. A. & Richards, E. J. *Arabidopsis thaliana* DNA methylation mutants. *Science* **260**, 1926–1928 (1993).
80. Baroux, C., Pecinka, A., Fuchs, J., Schubert, I. & Grossniklaus, U. The triploid endosperm genome of *Arabidopsis* adopts a peculiar, parental-dosage-dependent chromatin organization. *Plant Cell* **19**, 1782–1794 (2007).
81. Mosher, R. A. *et al.* Uniparental expression of PolIV-dependent siRNAs in developing endosperm of *Arabidopsis*. *Nature* **460**, 283–286 (2009).
82. Teixeira, F. K. *et al.* A role for RNAi in the selective correction of DNA methylation defects. *Science* **323**, 1600–1604 (2009).
83. Aravin, A. A. *et al.* Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr. Biol.* **11**, 1017–1027 (2001).
84. Aravin, A. A., Hannon, G. J. & Brennecke, J. The Piwi–piRNA pathway provides an adaptive defense in the transposon arms race. *Science* **318**, 761–764 (2007).
85. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nature Rev. Mol. Cell Biol.* **9**, 22–32 (2008).
86. Aravin, A. *et al.* A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442**, 203–207 (2006).
87. Girard, A., Sachidanandam, R., Hannon, G. J. & Carmell, M. A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442**, 199–202 (2006).
88. Kuramochi-Miyagawa, S. *et al.* DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev.* **22**, 908–917 (2008).
This study shows that methylation defects in a Piwi clade mutant occur at the time of *de novo* methylation in male germ cells, which suggests that piRNAs are required for *de novo* DNA methylation at transposons rather than for the maintenance of pre-existing DNA methylation.
89. Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K. & Hannon, G. J. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* **316**, 744–747 (2007).
90. Aravin, A. A. *et al.* A piRNA pathway primed by individual transposons is linked to *de novo* DNA methylation in mice. *Mol. Cell* **31**, 785–799 (2008).
This study, along with reference 88, isolated piRNAs from fetal germ cells and showed they are highly enriched for transposon sequences and possess the characteristics of primary and secondary piRNAs.
91. Brennecke, J. *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**, 1089–1103 (2007).
92. Gunawardane, L. S. *et al.* A Slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. *Science* **315**, 1587–1590 (2007).
93. Carmell, M. A. *et al.* MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev. Cell* **12**, 503–514 (2007).
94. Aravin, A. A. & Bourc'his, D. Small RNA guides for *de novo* DNA methylation in mammalian germ cells. *Genes Dev.* **22**, 970–975 (2008).
95. Chen, C. *et al.* Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. *Proc. Natl Acad. Sci. USA* **106**, 20336–20341 (2009).
96. Kirino, Y. *et al.* Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. *Nature Cell Biol.* **11**, 652–658 (2009).
97. Vagin, V. V. *et al.* Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. *Genes Dev.* **23**, 1749–1762 (2009).
98. Reuter, M. *et al.* Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. *Nature Struct. Mol. Biol.* **16**, 639–646 (2009).
99. Kojima, K. *et al.* Associations between PIWI proteins and TDRD1/MTR-1 are critical for integrated subcellular localization in murine male germ cells. *Genes Cells* **14**, 1155–1165 (2009).
100. Wang, J., Saxe, J. P., Tanaka, T., Chuma, S. & Lin, H. Mili interacts with Tudor domain-containing protein 1 in regulating spermatogenesis. *Curr. Biol.* **19**, 640–644 (2009).
101. Chuang, L. S. *et al.* Human DNA-(cytosine-5) methyltransferase–PCNA complex as a target for p21WAF1. *Science* **277**, 1996–2000 (1997).
102. Schermelleh, L. *et al.* Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Res.* **35**, 4301–4312 (2007).

© 2010 Macmillan Publishers Limited. All rights reserved

103. Egger, G. *et al.* Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. *Proc. Natl Acad. Sci. USA* **103**, 14080–14085 (2006).
104. Spada, F. *et al.* DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. *J. Cell Biol.* **176**, 565–571 (2007).
105. Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760–1764 (2007).
This study, along with reference 106, shows that UHRF1 is required for maintaining DNA methylation in mammals. Furthermore, this study demonstrates that the SRA domain of UHRF1 specifically interacts with hemimethylated CG sites and is required for the association of DNMT1 with chromatin.
106. Sharif, J. *et al.* The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908–912 (2007).
See reference 105.
107. Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y. & Shirakawa, M. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* **455**, 818–821 (2008).
108. Hashimoto, H. *et al.* The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* **455**, 826–829 (2008).
109. Qian, C. *et al.* Structure and hemimethylated CpG binding of the SRA domain from human UHRF1. *J. Biol. Chem.* **283**, 34490–34494 (2008).
110. Avvakumov, G. V. *et al.* Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature* **455**, 822–825 (2008).
111. Meilinger, D. *et al.* Np95 interacts with *de novo* DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. *EMBO Rep.* **10**, 1259–1264 (2009).
112. Dennis, K., Fan, T., Geiman, T., Yan, Q. & Muegge, K. Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes Dev.* **15**, 2940–2944 (2001).
113. Huang, J. *et al.* Lsh, an epigenetic guardian of repetitive elements. *Nucleic Acids Res.* **32**, 5019–5028 (2004).
114. Woo, H. R., Dittmer, T. A. & Richards, E. J. Three SRA-domain methylcytosine-binding proteins cooperate to maintain global CpG methylation and epigenetic silencing in *Arabidopsis*. *PLoS Genet.* **4**, e1000156 (2008).
115. Woo, H. R., Pontes, O., Pikaard, C. S. & Richards, E. J. VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. *Genes Dev.* **21**, 267–277 (2007).
116. Hirochika, H., Okamoto, H. & Kakutani, T. Silencing of retrotransposons in *Arabidopsis* and reactivation by the *ddm1* mutation. *Plant Cell* **12**, 357–369 (2000).
117. Gendrel, A. V., Lippman, Z., Yordan, C., Colot, V. & Martienssen, R. A. Dependence of heterochromatic histone H3 methylation patterns on the *Arabidopsis* gene DDM1. *Science* **297**, 1871–1873 (2002).
118. Yan, Q., Huang, J., Fan, T., Zhu, H. & Muegge, K. Lsh, a modulator of CpG methylation, is crucial for normal histone methylation. *EMBO J.* **22**, 5154–5162 (2003).
119. Bernatavichute, Y. V., Zhang, X., Cokus, S., Pellegrini, M. & Jacobsen, S. E. Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in *Arabidopsis thaliana*. *PLoS ONE* **3**, e3156 (2008).
120. Lister, R. *et al.* Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* **133**, 523–536 (2008).
See reference 8.
121. Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T. & Henikoff, S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genet.* **39**, 61–69 (2007).
122. Tran, R. K. *et al.* DNA methylation profiling identifies CG methylation clusters in *Arabidopsis* genes. *Curr. Biol.* **15**, 154–159 (2005).
123. Johnson, L. M. *et al.* The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr. Biol.* **17**, 379–384 (2007).
124. Lindroth, A. M. *et al.* Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science* **292**, 2077–2080 (2001).
125. Bartee, L., Malagnac, F. & Bender, J. *Arabidopsis cmt3* chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. *Genes Dev.* **15**, 1753–1758 (2001).
126. Jackson, J. P., Lindroth, A. M., Cao, X. & Jacobsen, S. E. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**, 556–560 (2002).
127. Malagnac, F., Bartee, L. & Bender, J. An *Arabidopsis* SET domain protein required for maintenance but not establishment of DNA methylation. *EMBO J.* **21**, 6842–6852 (2002).
128. Jackson, J. P. *et al.* Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in *Arabidopsis thaliana*. *Chromosoma* **112**, 308–315 (2004).
129. Ebbs, M. L., Bartee, L. & Bender, J. H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. *Mol. Cell Biol.* **25**, 10507–10515 (2005).
130. Ebbs, M. L. & Bender, J. Locus-specific control of DNA methylation by the *Arabidopsis* SUVH5 histone methyltransferase. *Plant Cell* **18**, 1166–1176 (2006).
131. Cao, X. *et al.* Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. *Curr. Biol.* **13**, 2212–2217 (2003).
132. Johnson, L. M., Law, J. A., Khattar, A., Henderson, I. R. & Jacobsen, S. E. SRA-domain proteins required for DRM2-mediated *de novo* DNA methylation. *PLoS Genet.* **4**, e1000280 (2008).
133. Ikeda, Y. & Kinoshita, T. DNA demethylation: a lesson from the garden. *Chromosoma* **118**, 37–41 (2009).
134. Zhu, J. K. Active DNA demethylation mediated by DNA glycosylases. *Annu. Rev. Genet.* **43**, 143–166 (2009).
135. Choi, Y. *et al.* DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. *Cell* **110**, 33–42 (2002).
136. Gong, Z. *et al. ROS1*, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. *Cell* **111**, 803–814 (2002).
137. Penterman, J. *et al.* DNA demethylation in the *Arabidopsis* genome. *Proc. Natl Acad. Sci. USA* **104**, 6752–6757 (2007).
138. Ortega-Galisteo, A. P., Morales-Ruiz, T., Ariza, R. R. & Roldan-Arjona, T. *Arabidopsis* DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. *Plant Mol. Biol.* **67**, 671–681 (2008).
139. Gehring, M. *et al.* DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* **124**, 495–506 (2006).
140. Morales-Ruiz, T. *et al.* DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc. Natl Acad. Sci. USA* **103**, 6853–6858 (2006).
141. Agius, F., Kapoor, A. & Zhu, J. K. Role of the *Arabidopsis* DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proc. Natl Acad. Sci. USA* **103**, 11796–11801 (2006).
References 139 and 140 show that DME is an active 5-methylcytosine DNA glycosylase *in vitro*, and references 140 and 141 show that ROS1 is an active 5-methylcytosine DNA glycosylase *in vitro*.
142. Zhu, J., Kapoor, A., Sridhar, V. V., Agius, F. & Zhu, J. K. The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation patterns in *Arabidopsis*. *Curr. Biol.* **17**, 54–59 (2007).
143. Penterman, J., Uzawa, R. & Fischer, R. L. Genetic interactions between DNA demethylation and methylation in *Arabidopsis*. *Plant Physiol.* **145**, 1549–1557 (2007).
144. Baute, J. & Depicker, A. Base excision repair and its role in maintaining genome stability. *Crit. Rev. Biochem. Mol. Biol.* **43**, 239–276 (2008).
145. Gehring, M., Reik, W. & Henikoff, S. DNA demethylation by DNA repair. *Trends Genet.* **25**, 82–90 (2009).
146. Kapoor, A., Agius, F. & Zhu, J. K. Preventing transcriptional gene silencing by active DNA demethylation. *FEBS Lett.* **579**, 5889–5898 (2005).
147. Malone, C. D. & Hannon, G. J. Small RNAs as guardians of the genome. *Cell* **136**, 656–668 (2009).
148 Zheng, X. *et al.* ROS3 is an RNA-binding protein required for DNA demethylation in *Arabidopsis*. *Nature* **455**, 1259–1262 (2008).
149. Hajkova, P. *et al.* Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117**, 15–23 (2002).
150. Hajkova, P. *et al.* Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* **452**, 877–881 (2008).
151. Santos, F., Hendrich, B., Reik, W. & Dean, W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev. Biol.* **241**, 172–182 (2002).
152. Oswald, J. *et al.* Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10**, 475–478 (2000).
153. Mayer, W., Niveleau, A., Walter, J., Fundele, R. & Haaf, T. Demethylation of the zygotic paternal genome. *Nature* **403**, 501–502 (2000).
154. Nakamura, T. *et al.* PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nature Cell Biol.* **9**, 64–71 (2007).
155. Li, X. *et al.* A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev. Cell* **15**, 547–557 (2008).
156. Reese, K. J., Lin, S., Verona, R. I., Schultz, R. M. & Bartolomei, M. S. Maintenance of paternal methylation and repression of the imprinted *H19* gene requires MBD3. *PLoS Genet.* **3**, e137 (2007).
157. Ciccone, D. N. & Chen, T. Histone lysine methylation in genomic imprinting. *Epigenetics* **4**, 216–220 (2009).
158. Edwards, C. A. & Ferguson-Smith, A. C. Mechanisms regulating imprinted genes in clusters. *Curr. Opin. Cell Biol.* **19**, 281–289 (2007).
159. Ooi, S. K. & Bestor, T. H. The colorful history of active DNA demethylation. *Cell* **133**, 1145–1148 (2008).
160. Morgan, H. D., Dean, W., Coker, H. A., Reik, W. & Petersen-Mahrt, S. K. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J. Biol. Chem.* **279**, 52353–52360 (2004).
161. Rai, K. *et al.* DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and Gadd45. *Cell* **135**, 1201–1212 (2008).
This study uses zebrafish embryos as a model for studying DNA demethylation and provides evidence that cytosine methylation can be removed through the coordinated activities of 5-methylcytosine deaminases and thymine mismatch DNA glycosylases.
162. Hendrich, B., Hardeland, U., Ng, H. H., Jiricny, J. & Bird, A. The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature* **401**, 301–304 (1999).
163. Petronzelli, F. *et al.* Investigation of the substrate spectrum of the human mismatch-specific DNA *N*-glycosylase MED1 (MBD4): fundamental role of the catalytic domain. *J. Cell Physiol.* **185**, 473–480 (2000).
164. Popp, C. *et al.* Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency *Nature* 22 Jan 2010 (doi:1038/nature08829).
Knocking out AID reduces the genome-wide DNA demethylation observed in PGCs.
165. Kim, M. S. *et al.* DNA demethylation in hormone-induced transcriptional derepression. *Nature* **461**, 1007–1012 (2009).
166. Zhu, B. *et al.* 5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence. *Nucleic Acids Res.* **28**, 4157–4165 (2000).
167. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
168. Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929–930 (2009).
169. Valinluck, V. *et al.* Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res.* **32**, 4100–4108 (2004).
170. Valinluck, V. & Sowers, L. C. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res.* **67**, 946–950 (2007).
171. Cannon, S. V., Cummings, A. & Teebor, G. W. 5-Hydroxymethylcytosine DNA glycosylase activity in mammalian tissue. *Biochem. Biophys. Res. Commun.* **151**, 1173–1179 (1988).
172. Jullien, P. E. *et al.* Retinoblastoma and its binding partner MSI1 control imprinting in *Arabidopsis*. *PLoS Biol.* **6**, e194 (2008).

© 2010 Macmillan Publishers Limited. All rights reserved

173. McCabe, M. T., Davis, J. N. & Day, M. L. Regulation of DNA methyltransferase 1 by the pRb/E2F1 pathway. *Cancer Res.* **65**, 3624–3632 (2005).
174. Kimura, H., Nakamura, T., Ogawa, T., Tanaka, S. & Shiota, K. Transcription of mouse DNA methyltransferase 1 (Dnmt1) is regulated by both E2F-Rb-HDAC-dependent and -independent pathways. *Nucleic Acids Res.* **31**, 3101–3113 (2003).
175. McCabe, M. T., Low, J. A., Imperiale, M. J. & Day, M. L. Human polyomavirus BKV transcriptionally activates DNA methyltransferase 1 through the pRb/E2F pathway. *Oncogene* **25**, 2727–2735 (2006).
176. Nicolas, E., Ait-Si-Ali, S. & Trouche, D. The histone deacetylase HDAC3 targets RbAp48 to the retinoblastoma protein. *Nucleic Acids Res.* **29**, 3131–3136 (2001).
177. Morgan, H. D., Santos, F., Green, K., Dean, W. & Reik, W. Epigenetic reprogramming in mammals. *Hum. Mol. Genet.* **14**, R47–R58 (2005).
178. Carlson, L. L., Page, A. W. & Bestor, T. H. Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes Dev.* **6**, 2536–2541 (1992).
179. Mertineit, C. *et al.* Sex-specific exons control DNA methyltransferase in mammalian germ cells. *Development* **125**, 889–897 (1998).
180. Wilczynska, A., Minshall, N., Armisen, J., Miska, E. A. & Standart, N. Two Piwi proteins, Xiwi and Xili, are expressed in the *Xenopus* female germline. *RNA* **15**, 337–345 (2009).
181. Kawaoka, S. *et al.* The *Bombyx* ovary-derived cell line endogenously expresses PIWI/PIWI-interacting RNA complexes. *RNA* **15**, 1258–1264 (2009).
182. Lau, N. C. *et al.* Abundant primary piRNAs, endo-siRNAs, and microRNAs in a *Drosophila* ovary cell line. *Genome Res.* **19**, 1776–1785 (2009).
183. Kawaoka, S. *et al. Bombyx* small RNAs: genomic defense system against transposons in the silkworm, *Bombyx mori*. *Insect Biochem. Mol. Biol.* **38**, 1058–1065 (2008).
184. Lau, N. C., Ohsumi, T., Borowsky, M., Kingston, R. E. & Blower, M. D. Systematic and single cell analysis of *Xenopus* Piwi-interacting RNAs and Xiwi. *EMBO J.* **28**, 2945–2958 (2009).

Acknowledgements

We thank members of the Jacobsen laboratory and anonymous reviewers for useful comments and discussion. We apologize to colleagues whose research we did not have space to discuss, especially those studying DNA methylation in other systems, such as maize and *Neurospora* species. J.A.L. was supported the US National Institutes of Health National Research Service Award 5F32GM820453. This research was supported by US National Institutes of Health grant GM60398. S.E.J. is an investigator at the Howard Hughes Medical Institute.

Competing interests statement

The authors declare no competing financial interests.

DATABASES

TAIR: http://www.arabidopsis.org
DDM1
Entrez Gene: http://www.ncbi.nlm.nih.gov/gene
Piwil2 | *Piwil4*
Entrez Protein: http://www.ncbi.nlm.nih.gov/protein
AGO4 | AID
UniProtKB: http://www.uniprot.org
APOBEC1 | CMT3 | DMT1 | DNMT3A | DNMT3B | DNMT3L | DRM2 | HELLS | Mbd4 | PIWIL1 | SUVH4 | SUVH5 | SUVH6 | TDG | TDRD1 | TDRD9 | TDRKH | UHRF1

FURTHER INFORMATION

Authors' homepage: http://www.mcdb.ucla.edu/research/jacobsen/LabWebSite/P_Index.shtml
Medical Research Council Hartwell — Genomic Imprinting: http://www.har.mrc.ac.uk/research/genomic_imprinting

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

© 2010 Macmillan Publishers Limited. All rights reserved