

Regulation and function of DNA methylation in plants and animals

Xin-Jian He¹, Taiping Chen², Jian-Kang Zhu^{3,4}

¹National Institute of Biological Sciences, Beijing 102206, China; ²Developmental and Molecular Pathways, Novartis Institutes for BioMedical Research, 100 Technology Square, Cambridge, MA 02139, USA; ³Plant Stress Genomics Research Center, 4700 King Abdullah University of Science and Technology, Thuwal 23955-6900, Kingdom of Saudi Arabia; ⁴Horticulture and Landscape Architecture Department, Purdue University, West Lafayette, IN 47907, USA

DNA methylation is an important epigenetic mark involved in diverse biological processes. In plants, DNA methylation can be established through the RNA-directed DNA methylation pathway, an RNA interference pathway for transcriptional gene silencing (TGS), which requires 24-nt small interfering RNAs. In mammals, *de novo* DNA methylation occurs primarily at two developmental stages: during early embryogenesis and during gametogenesis. While it is not clear whether establishment of DNA methylation patterns in mammals involves RNA interference in general, *de novo* DNA methylation and suppression of transposons in germ cells require 24-32-nt piwi-interacting small RNAs. DNA methylation status is dynamically regulated by DNA methylation and demethylation reactions. In plants, active DNA demethylation relies on the repressor of silencing 1 family of bifunctional DNA glycosylases, which remove the 5-methylcytosine base and then cleave the DNA backbone at the abasic site, initiating a base excision repair (BER) pathway. In animals, multiple mechanisms of active DNA demethylation have been proposed, including a deaminase- and DNA glycosylase-initiated BER pathway. New information concerning the effects of various histone modifications on the establishment and maintenance of DNA methylation has broadened our understanding of the regulation of DNA methylation. The function of DNA methylation in plants and animals is also discussed in this review.

Keywords: DNA methylation; DNA demethylation; RNA; histone modifications

Cell Research (2011) 21:442-465. doi:10.1038/cr.2011.23; published online 15 February 2011

Introduction

DNA methylation refers to the addition of a methyl group to the cytosine bases of DNA to form 5-methylcytosine. DNA methylation occurs in both prokaryotes and eukaryotes. In bacteria, DNA methylation differentiates the genome DNA from invading phage DNA, so that phage DNA is cleaved by the host restriction enzymes [1]. DNA methylation is conserved in most major eukaryotic groups, including plants, and many fungi and animals, although it has been lost in some organisms such as the budding yeast *Saccharomyces cerevisiae* and the nematode worm *Caenorhabditis elegans* [2]. In plants, DNA methylation

occurs in the contexts of CG, CHG, and CHH (H = A, C, or T). In mammals, DNA methylation is restricted to the symmetric CG context, although non-CG methylation is prevalent in embryonic stem (ES) cells [3, 4]. In both mammals and plants, centromeric and pericentromeric regions, as well as other repetitive elements are heavily methylated. Many genic regions also show high degrees of methylation. In contrast, promoter regions mostly lack DNA methylation [5, 6].

Mammals have three active DNA methyltransferases: DNMT1, DNMT3A, and DNMT3B. DNMT1 maintains DNA methylation at hemi-methylated DNA after DNA replication during cell divisions, whereas DNMT3A and DNMT3B are responsible for establishing *de novo* DNA methylation [7]. A third member of the DNMT3 family, DNMT3-like (DNMT3L), which has no catalytic activity, functions as a regulator of DNMT3A and DNMT3B [8, 9]. Mammalian DNMT2 is a tRNA methyltransferase rather than a DNA methyltransferase and has been re-

Correspondence: Xin-Jian He^a, Taiping Chen^b

^aTel: 86-10-80707712; Fax: 86-10-80707715

E-mail: hexinjian@nibs.ac.cn

^bTel: 617-871-7414; Fax: 617-871-7057

E-mail: taiping.chen@novartis.com

named tRNA aspartic acid methyltransferase 1 [10]. In plants, three DNA methyltransferases have been characterized: MET1, CMT3, and DRM2 [11-13]. MET1, a homolog of mammalian DNMT1, is responsible for maintenance of symmetric CG methylation [14]. CMT3 is a plant-specific DNA methyltransferase that is required for the maintenance of DNA methylation at CHG sites [12]. DRM2 is responsible for *de novo* DNA methylation at all sequence contexts; its role in CHH methylation is most prominent, since CHH methylation cannot be maintained and thus must rely entirely on *de novo* methylation. The DRM2 activity is highly regulated by the RNA-directed DNA methylation (RdDM) pathway [15, 16].

In plants, DNA demethylation depends on four bifunctional 5-methylcytosine glycosylases: Repressor of silencing 1 (ROS1), Demeter (DME), DME-like 2 (DML2), and DML3, which remove methylated bases and cleave the DNA backbone at abasic sites. The resulting gap is filled by a DNA polymerase and a DNA ligase [17-19]. ROS1 is ubiquitously expressed and functions to prevent the TGS of transgenes and some endogenous genes, and to maintain some levels of expression of transposons [17, 20, 21]. DME is required for gene imprinting and genome-wide DNA demethylation in central cells and endosperms [19, 22-24]. In mammals, active demethylation contributes to the genome-wide erasure of DNA methylation marks observed in preimplantation embryos and primordial germ cells (PGCs). Active demethylation also takes place at specific loci in somatic cells under certain circumstances [25]. However, the mechanisms underlying active demethylation in mammals have been highly controversial. It is possible that various mechanisms function in a context-dependent manner [26, 27].

Recent studies have deepened our understanding of the establishment and maintenance of DNA methylation as well as of active DNA demethylation pathways in plants and animals. The high-throughput sequencing and ChIP-on-chip technologies have been applied to study the profiles of DNA methylation at the whole genome level [28]. In this review, we will discuss recent advances in DNA methylation and demethylation, the affects of the RNA interference pathway and histone modifications on DNA methylation, and the diverse functions of DNA methylation.

De novo DNA methylation

RNA-directed DNA methylation in plants

In the *Arabidopsis* genome, DNA methylation is highly concentrated in centromeric regions and repetitive sequences throughout the genome [5, 29]. The methylated DNA loci are frequently accompanied by siRNAs. Ap-

proximately one-third of methylated DNA loci are rich in siRNAs, supporting an important role of siRNAs in DNA methylation [5, 30]. Whole-genome tiling array studies in *Arabidopsis* suggested that most non-coding genomic sequences are capable of generating RNA transcripts. Recent studies suggested that both siRNAs and long non-coding RNAs are involved in *de novo* DNA methylation [31].

RdDM is a conserved *de novo* DNA methylation mechanism in plants. The phenomenon was first found in transgenic potato containing viroid genes, in which the recombinant viroid sequence in the plant genome was methylated when the RNA viroid was replicated [32], indicating an RdDM mechanism. RdDM was subsequently recognized as a general transcriptional silencing mechanism in plants. It is involved in various epigenetic phenomena, which include transgene silencing, transposon suppression, gene imprinting, and paramutation [33-36].

Many components of the RdDM pathway have been recently identified (Figure 1). The pathway contains two plant-specific, atypical DNA-dependent RNA polymerases whose largest subunits, which were independently identified by different groups, are NRPD1 and NRPE1 [31, 34, 37-39]. NRPD1 and NRPE1 are similar to the largest subunits of canonical DNA-dependent RNA polymerases but specifically function in *de novo* DNA methylation and transcriptional gene silencing. Other characteristics of the two plant-specific, atypical, DNA-dependent RNA polymerases have been explored by forward genetic screens and affinity purification [40-43]. Similar to DNA-dependent RNA polymerase II (Pol II), the two atypical plant-specific RNA polymerases are also multi-subunit protein complexes and have been named (in the manner of DNA-dependent RNA polymerases) Pol IV and Pol V [42, 43]. Pol IV and Pol V have their own specific subunits but share some subunits with each other and share other subunits with Pol II [42].

Pol IV was proposed to synthesize RNA transcripts by using the RdDM target loci as templates (Figure 1). The single-stranded RNA is converted into double-stranded RNA by the RNA-dependent-RNA polymerase RDR2 [33]. The double-stranded RNA is cleaved into 24-nt primary siRNAs by a Dicer-like protein DCL3 [33]. The maturation of siRNAs requires HUA ENHANCER 1 (HEN1), a conserved *S*-adenosyl-L-methionine-dependent RNA methyltransferase that methylates the 2'-OH group on the 3'-terminal nucleotides of siRNAs [44, 45]. In addition to siRNAs, RdDM also requires non-coding RNA transcripts produced by Pol V (Figure 1) [31]. When siRNAs are loaded onto the ARGONAUTE proteins AGO4 and/or AGO6 [16, 46, 47], KTF1 binds the nascent non-coding RNA transcripts produced by Pol

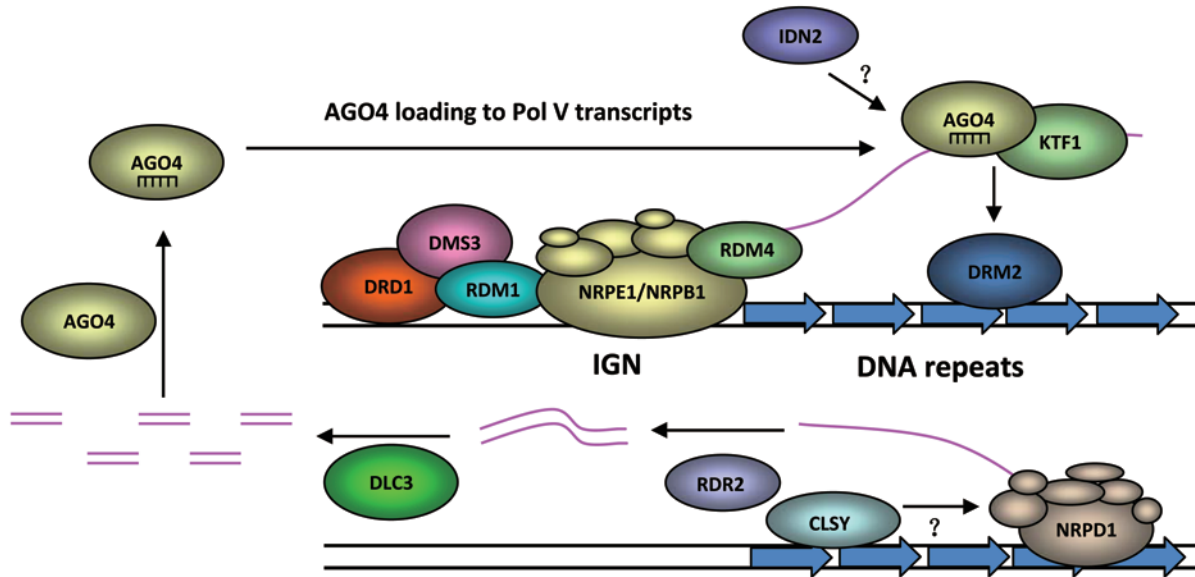


Figure 1 The RNA-directed DNA methylation pathway in plants. In transposons and other DNA repeat regions, aberrant single-stranded RNAs are proposed to be produced by DNA-dependent RNA polymerase IV (Pol IV). The chromatin remodeling protein CLSY may facilitate Pol IV transcription. RNA-dependent RNA polymerase RDR2 converts the aberrant single-stranded RNAs to double-stranded RNAs, which are then cleaved into 24-nt siRNAs by the Dicer-like protein DCL3. The 24-nt siRNAs are bound by an ARGONAUTE protein AGO4, AGO6, or AGO9. In intergenic non-coding (IGN) regions, DNA-dependent RNA polymerase V (Pol V) generates single-stranded scaffold RNA transcripts. Generation of Pol V RNA transcripts requires RDM4/DMS4, DRD1, DMS3, and RDM1. RDM1 may bind single-stranded methylated DNA and help recruit Pol V and Pol II to appropriate chromatin regions. DRD1, DMS3, and RDM1 form a stable protein complex, named DDR. KTF1 is an RNA-binding protein, which tethers AGO4 to nascent Pol V or Pol II RNA transcripts to form the RNA-directed DNA methylation effector complex. IDN2 may stabilize the base-pairing between the nascent scaffold transcripts and 24-nt siRNAs. The effector complex directs the *de novo* DNA methyltransferase DRM2 to specific chromatin regions to catalyze new DNA methylation.

V, and is tethered to AGO4 by the base pairing between the non-coding RNA transcripts and the siRNAs [48, 49]. The interaction between AGO4 and the WG/GW motif of KTF1 reinforces the association between siRNAs and Pol V transcripts, forming a functional RdDM effector complex that directs the *de novo* DNA methyltransferase DRM2 to genomic loci being transcribed by Pol V [31, 47, 48, 50]. Moreover, AGO4 has a slicer activity by which the scaffold non-coding RNA transcripts are cleaved. The cleaved RNA fragments may be copied by RDR2 and then processed into 24-nt secondary siRNAs by DCL3. IDN2, a recently identified RdDM component, is an SGS3-like protein containing XS and XH domains separated by a coiled-coil region [51, 52]. The XS domain in IDN2 binds double-stranded RNA *in vitro* [51], suggesting a role in siRNA production or in the RdDM effector complex. RDM1 was recently identified from forward genetics screens for *Arabidopsis* mutants with reduced TGS [53]. RDM1 functions together with AGO4 and DRM2 as a part of the RdDM effector complex. RDM1 appears to bind single-stranded methylated DNA, and may contribute to the recruitment of the

RdDM effector complex to methylated DNA [53]. Such recruitment helps to explain how siRNA production and amplification occur preferentially at methylated loci, and support the notion of a self-reinforcing feedback loop between siRNA generation and DNA methylation.

In eukaryotes, DNA-dependent RNA polymerases are generally composed of 12 subunits, some of which are shared by Pol I, II, and III [54]. Similar to RNA polymerases I, II, and III, the plant-specific Pol IV and Pol V are multi-subunit complexes. The subunits of Pol IV and Pol V are paralogous or identical to the 12 subunits of Pol II [34, 37-40]. The functions of subunits NRPE5 as well as the largest subunits NRPD1/E1 have been demonstrated by both genetic and biochemical analyses [34, 37-42]. NRPD/E2 and NRPD/E4 are common to Pol IV and Pol V [37, 40]. NRPE5 functions exclusively in Pol V [41, 42]. The absence of genetic evidence for the function of the other subunits indicates that there are redundancy or they are non-essential for RdDM. It appears that Pol IV and Pol V have evolved from the general RNA polymerases and function specifically in the production of non-coding RNAs for

RdDM and TGS.

The transcription activity of Pol II is tightly regulated with the help of a variety of general transcription factors during initiation and elongation [55, 56]. While the transcription regulation mechanism for eukaryotic Pol II has been extensively studied, little is known about the transcription regulation of Pol IV and Pol V. KTF1, an SPT5-like elongation factor, was characterized as a scaffold protein in the RdDM effector [48], and was also identified as a Pol V-interacting component [43]. RDM4/DMS4, a homolog of the yeast transcription factor IWR1, was identified by two different genetic screens [57, 58]. RDM4/DMS4 interacts with Pol V, and is required for generation of Pol V-dependent transcripts. These results suggest that the regulation of the polymerase activity of Pol IV and Pol V requires transcriptional factors such as KTF1 and RDM4/DMS4. However, KTF1 is dispensable for generation of Pol V-dependent non-coding RNA transcripts. Pol V-dependent transcripts apparently over-accumulate in *ktf1* and *ago4* mutants [48, 49], indicating that KTF1 and AGO4 act downstream of Pol V transcription. The C-terminal region of KTF1 contains more than 40 WG/GW repeats that bind to AGO4 [48, 59]. Moreover, KTF1 also binds RNA transcripts, which may be produced by Pol V and/or Pol II. With the dual functions, KTF1 helps recruit AGO4 to the scaffold RNA transcripts to form a functional effector complex that directs the DNA methylase DRM2 to RdDM target loci [48, 49].

Non-coding RNAs produced by RNA Pol II are widespread in plants and animals. Zheng *et al.* [60] found that mutation of the second largest subunit of Pol II leads to loss of DNA methylation and TGS at specific genomic loci. The Pol II-dependent non-coding RNAs physically associate with AGO4, resulting in DNA methylation and TGS. Moreover, Pol II-dependent transcripts may also help recruit Pol IV and Pol V to promote siRNA biogenesis and function [60]. Interestingly, RDM1 interacts and co-localizes with Pol II, but it is also required for the generation of Pol V-dependent transcripts [53]. RDM1 may bind single-stranded DNA at sites of Pol II transcription, and be involved in recruiting Pol IV, Pol V, and other RdDM components. Further studies are required to clarify the functions of Pol II, Pol IV, and Pol V in the RdDM pathway.

Two other RdDM components: defective in RNA-directed DNA methylation 1 (DRD1) and defective in meristem silencing 3 (DMS3), were identified from forward genetics screens [61, 62]. DRD1 is an SNF2-like chromatin remodeler, and DMS3 contains a domain that is similar to the hinge region of structural maintenance of chromosome proteins. Both proteins are required for

Pol V-dependent transcript production and are required for the association of Pol V with chromatin [16, 31, 63]. Recent affinity purification identified a protein complex containing DRD1, DMS3, and RDM1 (DDR; Figure 1) [63]. Pol V components are enriched in the DDR complex, indicating that the complex may help recruit Pol V to specific chromatin regions [63]. Another SNF2-like protein, CLSY1, functions together with Pol IV and RDR2, and is required for the production 24-nt siRNAs and for the spreading of TGS [64].

De novo DNA methylation in animals

In mammals, DNA methylation patterns are set up early in development through a highly orchestrated process that involves genome-wide demethylation and *de novo* methylation. During preimplantation development, DNA methylation marks inherited from gametes are largely erased. On implantation, the embryo undergoes a wave of *de novo* methylation that establishes a new methylation pattern, which is copied during division of somatic cells [65-67]. Genetic and biochemical evidence indicates that DNMT3A and DNMT3B are mainly responsible for *de novo* methylation, and DNMT1 is the major enzyme for maintaining DNA methylation patterns [68-74].

Another developmental stage that exhibits substantial *de novo* DNA methylation in mammals is gametogenesis. DNA methylation in both male and female germ cells plays a critical role in the establishment of genomic imprinting. Genomic imprinting is an epigenetic process that marks alleles according to their parental origin and results in monoallelic expression of a small subset of genes [75]. Genetic studies demonstrate that DNMT3A and DNMT3L are essential for setting up DNA methylation imprints in germ cells [8, 9, 76]. Although DNMT3L has no enzymatic activity, it has been shown to interact with DNMT3A and stimulate its activity [9, 77-80]. A more recent study showed that DNMT3L binds the N-terminal tail of histone H3, suggesting a role for DNMT3L in recruiting DNMT3A to specific genomic regions [81]. Interestingly, the interaction between DNMT3L and histone H3 is inhibited by methylation at H3K4, indicating that H3K4 methylation may regulate germline imprinting [81]. Consistent with this notion, Ciccone *et al.* [82] demonstrated that KDM1B (also known as AOF1 and LSD2), a histone H3K4 demethylase highly expressed in growing oocytes, is required for *de novo* DNA methylation of maternally imprinted genes (Figure 2). Active transcription across imprinting control regions also appears to be required for the establishment of DNA methylation imprints in female germ cells [83]. It is possible that removal of H3K4 methylation and active transcription at imprinted loci create or maintain a chromatin state

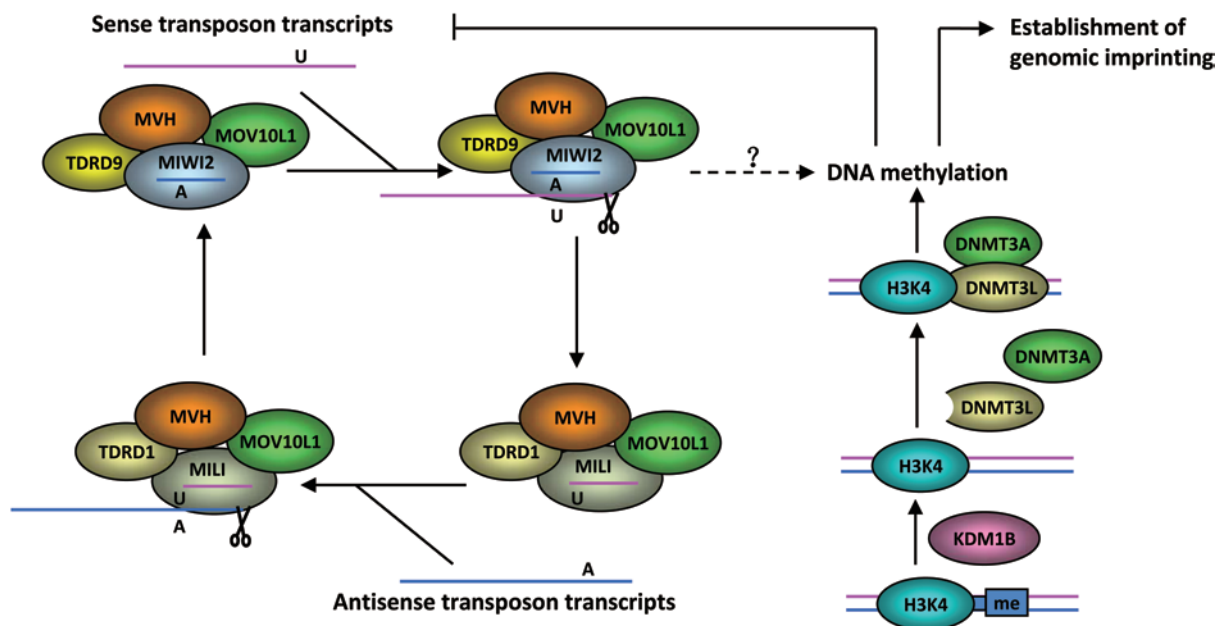


Figure 2 *De novo* DNA methylation in the mammalian germline. Two Piwi proteins, MILI and MIWI2, are required for piRNA (Piwi-interacting RNA) generation. The piRNAs are generated in fetal male gonads, and play important roles in silencing transposons by causing DNA methylation. The primary piRNAs are bound with cytoplasmic MILI, which cleaves antisense transposon transcripts. The secondary piRNAs are bound with MIWI2, which cleaves sense transposon transcripts. The sense transposon transcripts produce primary piRNAs with 5' uridine (U), whereas antisense transposon transcripts produce secondary piRNAs with an adenine (A) at position 10. Generation of piRNAs also require Tudor domain-containing (TDRD) proteins, mouse VASA homolog (MVH), and putative DExD-box helicase MOV10L1. The interaction between Piwi proteins and TDRD proteins is essential for generation of piRNAs. TDRD9-MIWI2 and TDRD1-MILI are two conserved complexes that generate primary piRNAs and secondary piRNAs, respectively. MVH and MOV10L1 are required for the activity of both MILI and MIWI2 in piRNA generation and *de novo* DNA methylation. DNMT3L interacts with unmethylated H3K4, and recruits DNMT3A to specific genomic regions for DNA methylation. The histone H3K4 demethylase KDM1B catalyzes demethylation of H3K4, by which it promotes *de novo* DNA methylation.

that facilitates the access of the DNMT3A-DNMT3L complex to these loci.

DNA methylation in mammalian germ cells is also essential for suppression of transposons. *DNMT3L*-deficient male mice, in addition to exhibiting imprinting defects, fail to establish *de novo* methylation of transposons and, as a result, show uncontrolled transposon expression and spermatogenesis failure [84, 85]. This phenotype is remarkably similar to that of mice deficient for the Piwi family members MILI or MIWI2, suggesting that the Piwi-piRNA complexes and DNA methylation machinery act together to suppress transposons in germ cells [86-88].

The Piwi proteins belong to the ARGONAUTE superfamily. In contrast to the Ago subfamilies of ARGONAUTE proteins, which complex with microRNAs or siRNAs, the Piwi subfamily members bind to piRNAs (Figure 2). Tens of thousands of piRNA species, typically 24-32 nucleotide in length, have been found in mammals, zebrafish, and *Drosophila*, but not in plants. A consider-

able portion of piRNAs maps to transposon-encoding regions. The expression of Piwi proteins and piRNAs is mainly restricted to the germline, and the Piwi-piRNA pathway has been implicated in a variety of germline functions [89-91]. In mice, two Piwi proteins, MILI and MIWI2, are involved in piRNA generation in fetal male gonads, and these piRNAs play important roles in silencing retrotransposons via DNA methylation (Figure 2) [86, 88, 89]. Genetic studies indicate that, in male germ cells, *MILI*- or *MIWI2*-deficiency impairs *de novo* methylation of retrotransposons, whereas *DNMT3L*-deficiency leaves the piRNA pathway largely intact [92, 93]. These results suggest that the Piwi-piRNA pathway acts upstream of DNA methylation. piRNA-containing complexes are proposed to guide the *de novo* DNA methylation machinery to transposon sequences. This role is analogous to that of the AGO4/6-siRNA complexes in RdDM and transposon suppression. No direct interaction between Piwi proteins and DNMT3 proteins was detected [93], suggesting that

the recruitment of DNMT3 proteins by Piwi proteins could be indirect.

Tudor-domain-containing proteins (TDRDs) have been shown to participate in the Piwi pathway to suppress retrotransposons via DNA methylation [94–98]. TDRD1 associates with MILI, whereas TDRD9 associates with MIWI2 (Figure 2) [95]. Dimethylation of arginine in the N-terminal region of MILI is essential for the interaction with TDRD1. A global mass spectrometry study suggested that all mouse Piwi protein complexes comprise specific TDRD proteins [95]. The TDRD-Piwi associations are essential for retrotransposon repression. MVH is a homolog of *Drosophila* VASA, a germ cell-specific DEAD-box RNA helicase [99]. Similar to Piwi proteins and TDRD proteins, MVH is expressed in male germ cells from early embryogenesis to around the time of spermatid development [100]. The defective spermatogenesis in *MVH*-deficient mice is similar to that in *MILI*-deficient mice [86]. *MVH* interacts with both MILI and MIWI, and plays essential roles in piRNA production and subsequent DNA methylation (Figure 2) [101]. Another putative DExD-box RNA helicase, MOV10L1, is also essential for the piRNA-dependent DNA methylation pathway [102, 103]. Absence of MOV10L1 in mice results in activation of LTR and long interspersed nuclear element-1 (LINE-1) retrotransposons, followed by cell death, causing male infertility and complete blockage of spermatogenesis [103]. MOV10L1 is required for biogenesis of piRNAs, and for the loading of piRNA to MILI and MIWI2 in mammalian male germ cells (Figure 2) [102]. These results indicate that requirement of small RNAs and ARGONAUTE proteins is a conserved mechanism for *de novo* DNA methylation in mammalian male germ cells and plant RdDM.

Active DNA demethylation in plants and animals

While DNA methylation can be established and maintained, DNA demethylation also occurs in plants and animals. When DNA methylation pathways are inactivated, DNA methylation is diluted after DNA replication, leading to passive DNA demethylation. In other cases, however, DNA methylation is removed through active DNA demethylation pathways. Passive and active DNA demethylation may simultaneously reduce DNA methylation during specific developmental stages. We now review active DNA demethylation mechanisms.

In plants, active DNA demethylation was first found in a transgene system. Our previous research showed that ROS1 counteracts the DNA methylation pathway to prevent gene silencing in plants [20]. In the wild-type (WT) genetic background, the construct harboring the *RD29A-*

LUC transgene is normally transcribed in response to appropriate environmental cues [104]. Mutations of *ROS1* cause hypermethylation of the promoter of *RD29A-LUC* transgene, leading to silencing of the transgene and its homologous endogenous gene (Figure 3A). *ROS1* is also required to suppress DNA methylation in a number of other endogenous genomic loci including many transposons [21, 105]. The DNA demethylation function of ROS1 was further confirmed by screening for second-site suppressors of *ros1*. The study recovered most of known RdDM components, including NRPD1, NRPE1, NRPD2, AGO4, DRD1, HEN1, and DMS3, but also identified several previously uncharacterized RdDM components, including RDM1, RDM2/NRPD4, RDM3/KTF1, and RDM4 [40, 48, 53, 57]. In the *ros1* background, mutation of each of these RdDM components reduces the DNA methylation level at the transgene *RD29A* promoter and its corresponding homologous endogenous gene (Figure 3A). The results strongly suggest that ROS1 counteracts *de novo* DNA methylation in transgenes by actively promoting DNA demethylation.

ROS1 encodes a 5-methylcytosine DNA glycosylase/demethylase, which is a nuclear protein containing a C-terminal DNA glycosylase domain and an N-terminal histone H1-like basic region [17, 20]. ROS1 is a member of a small protein family comprising four DNA glycosylases that also includes DME, DML2, and DML3 [25, 106]. DNA glycosylases function in DNA repair processes by excising damaged or mismatched bases. The bifunctional DNA glycosylase ROS1 not only removes 5-methylcytosine from deoxyribose but also cleaves the DNA backbone at the abasic site [17, 18, 20]. Overexpression of *ROS1* in transgenic plants leads to a reduced level of cytosine methylation and an increased expression of target genes [17]. Another *Arabidopsis* DNA glycosylase, DME, is preferentially expressed in central cells of the female gametophyte and endosperms of developing seeds and is required for gene imprinting in endosperms [19, 24]. DME also shows 5-methylcytosine glycosylase activity, which is required for the reduced methylation and expression of maternal alleles at imprinted loci [18, 19]. Recent studies showed that DME is responsible for genome-wide DNA demethylation in endosperms (Figure 3B) [23, 24]. These results demonstrated that ROS1 and DME are 5-methylcytosine DNA glycosylases and function as DNA demethylases in *Arabidopsis*. DML2 and DML3 are in the same family as ROS1 and DME. Whole-genome DNA methylation assays in *ros1dml2dml3*-triple mutants showed that hundreds of endogenous loci are hypermethylated, suggesting that these DNA demethylases function throughout the genome [30, 106]. Most hypermethylated loci are located in promoters and

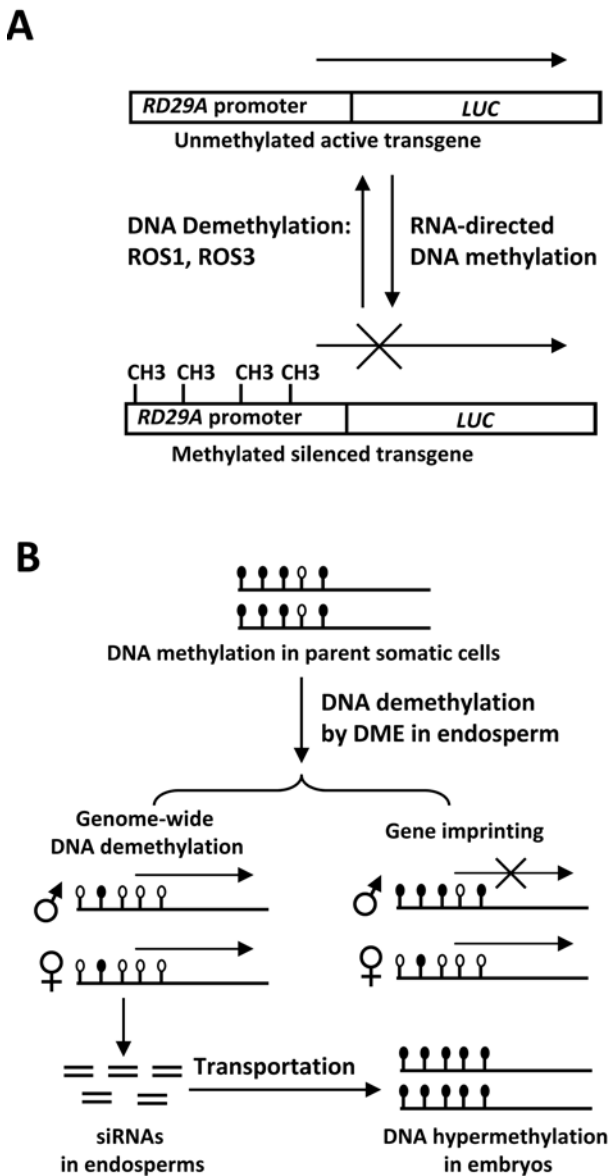


Figure 3 Active DNA demethylation and its function in plants. The plant 5-methylcytosine DNA glycosylases ROS1, DME, DML2, and DML3 function as active DNA demethylases. **(A)** ROS1 was discovered by screening for repressor of silencing in *Arabidopsis* plants expressing the *RD29A* promoter-driven luciferase reporter gene. ROS1 prevents transgene silencing that is caused by RNA-directed DNA methylation. ROS1 also functions to prevent over-methylation and alleviate the silencing of some endogenous genes and transposons. ROS3 is an RNA-binding protein that may direct ROS1 to specific genome targets. **(B)** DME is preferentially expressed in endosperms, and is responsible for genome-wide DNA demethylation and gene imprinting. Genome-wide DNA demethylation activates transposons and other repetitive DNA sequences, leading to the enhanced production of siRNAs in endosperms. These siRNAs might be transported into embryos, and contribute to DNA hypermethylation, to ensure genome stability in embryos. Black and white circles represent methylated and unmethylated cytosines, respectively.

3'-UTRs. These results indicate that ROS1, DML2, and DML3 might specifically demethylate 5' and 3' ends of genes but not gene-body methylation [106]. Besides *ros1*, our forward genetic screen identified another ROS mutant *ros3* [107]. ROS3 also counteracts the DNA methylation function of RdDM components. ROS3 is an essential regulator of DNA demethylation and acts in the same genetic pathway as ROS1 to prevent DNA hypermethylation (Figure 3A). ROS3 and ROS1 co-localize in discrete foci dispersed throughout the nucleus. ROS3 is capable of binding single-stranded RNAs, which might guide ROS1-mediated DNA demethylation [107].

ROS1-mediated active demethylation and RdDM pathways have opposing functions but have an interesting inter-dependent relationship. Mutations in RdDM components significantly reduce the *ROS1* transcript level [105, 108, 109], suggesting a feedback mechanism for the regulation of DNA methylation. The hypomethylation status of some genomic loci in RdDM mutants somehow reduces the expression of *ROS1* to protect the genome from further demethylation. The central cell-specific expression of DME causes genome-wide DNA demethylation in endosperms relative to embryos, but compared with WT endosperms, non-CG DNA methylation is lower in *dme* mutant endosperms [23, 24]. This interesting result suggests that *DME* promotes DNA methylation indirectly at non-CG sites, possibly by up-regulating RdDM components. Moreover, increased DNA methylation occurs at specific genome loci in the RdDM mutants *nrdp1* and *nrdp1*, while reduced DNA methylation occurs at many repetitive DNA sequences [110]. The increased DNA methylation at certain loci in the RdDM mutants may be due to downregulation of *ROS1* in these mutants. Together, these results indicate that methylation of genomic DNA in *Arabidopsis* is dynamically regulated by both active DNA demethylation and DNA methylation mechanisms.

Compared with the compelling evidence for the active DNA demethylation mechanism in plants, the evidence for active DNA demethylation mechanisms in animals is controversial. No orthologs of ROS1 and DME were found in mammals [25]. Global DNA demethylation occurs at two stages of embryogenesis in mammals. One is the paternal demethylation in zygotes, and the other is demethylation in PGCs from embryonic day E10.5 to E12.5 (Figure 4) [27]. After fertilization, the paternal genome in zygotes shows active DNA demethylation and remains demethylated until the implantation of the blastocyst, while the maternal genome undergoes passive demethylation due to exclusion of DNMT1 from the nucleus [111]. Imprinted genes are resistant to this wave of DNA demethylation. Several proteins, including

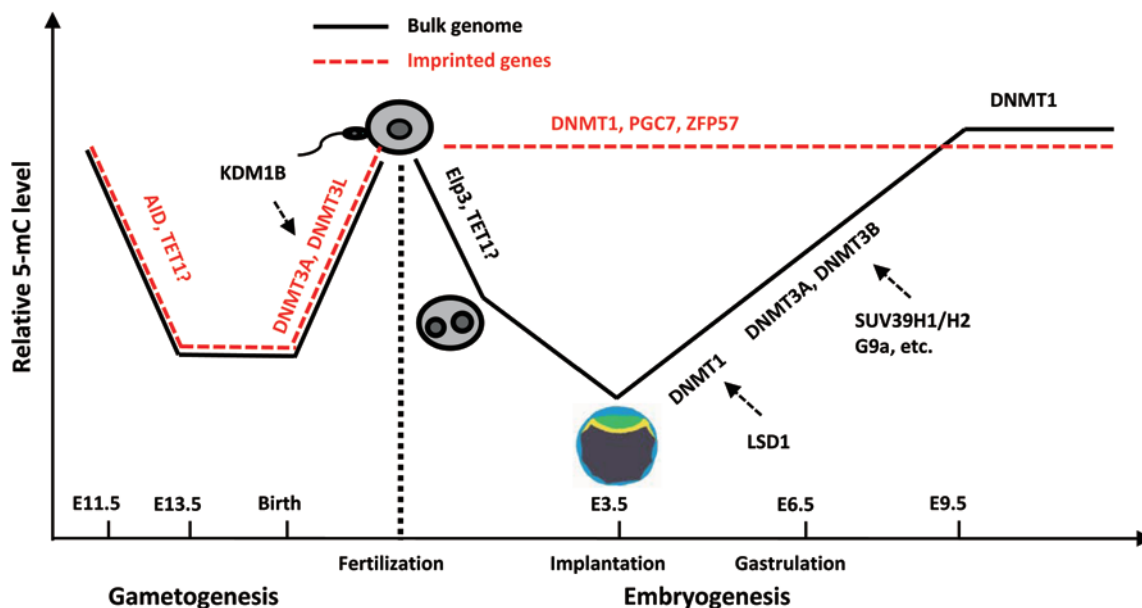


Figure 4 Dynamic changes in DNA methylation during mouse development. Shortly after fertilization, the paternal genome undergoes active demethylation, and the maternal genome is passively demethylated during subsequent cleavage divisions. After implantation, the embryo undergoes *de novo* methylation that establishes a new methylation pattern. Imprinted genes escape the waves of demethylation and *de novo* methylation during embryogenesis. Genome-wide demethylation and *de novo* methylation also occur in the male and female germ cells during gametogenesis, which are critical for the establishment of genomic imprinting. DNMTs and other regulatory factors involved in these processes are indicated.

DNMT1, ZFP57, and PGC7, are required for maintaining the DNA methylation imprints or protecting them from demethylation (Figure 4) [112-116]. PGCs appear at E7.5, begin to migrate at E8.5, and arrive at the genital ridge at E11.5. During the migration, global active DNA demethylation occurs in the presence of DNMT1 [117]. Besides the global DNA demethylation described above, sequence-specific DNA demethylation also occurs in somatic cells in response to various signals [118-121].

Glycosylase-dependent DNA demethylation was first proposed in animals [122, 123]. The 5-methylcytosine glycosylase activity was initially detected in chicken embryo extracts, which contain thymine-DNA glycosylase (TDG). However, the glycosylase activity of TDG is much lower against 5-methylcytosine than against mismatched thymine [124]. Thus, some other components in the chicken embryo extracts could be involved in active DNA demethylation. In addition to TDG, a methylcytosine-binding protein (MBD4) also has DNA glycosylase activity. Similar to the activity of TDG, the 5-methylcytosine glycosylase activity of MBD4 is also much lower than the TDG activity [125]. Recently, MBD4 was reported to carry out active DNA demethylation in the promoter of CYP27B1 [126]. Thus, TDG and MBD4 might require some other proteins to activate DNA demethyla-

tion pathways *in vivo*.

Activation-induced deaminase AID and apolipoprotein B mRNA editing enzyme APOBEC can deaminate 5-methylcytosine to thymine *in vitro* [127]. A study of zebrafish embryos suggested that AID, MBD4, and the DNA-repair protein GADD45A can cooperate to promote DNA demethylation, which is required for the development of zebrafish embryos [128]. The results suggested that AID might deaminate 5-methylcytosine to thymine, which is followed by excision of the T/G mismatch by MBD4. As a DNA-repair-related protein, GADD45A might help to coordinate the function of AID and MBD4. A recent study suggests that AID deficiency interferes with genome-wide erasure of DNA methylation marks in mouse PGCs [129]. However, no clear developmental phenotypes were found in AID- or MBD4-deficient mice [130-132], possibly due to functional redundancy.

DNMT3A and DNMT3B have been shown to deaminate 5-methylcytosine *in vitro*, although they are commonly known as DNA methyltransferases [121]. The 5-methylcytosine deamination function of DNMTs may occur when *S*-adenosylmethionine (SAM) concentrations are very low. DNMT3A/DNMT3B associate with TDG/MBD4, and this interaction stimulates the glycosylase activity, allowing for the repair of the T/G mismatch

[133, 134]. The interaction between DNMTs and TDG/MBD4 facilitates active DNA demethylation by coupling the functions of 5-methylcytosine deaminase and thymine glycosylase. Although this mechanism has been shown to demethylate the *pS2/TFF1* gene promoter upon activation by estrogens [121], its significance in mediating demethylation in other contexts remains to be determined.

Consistent with the DNA repair-driven DNA demethylation hypothesis, Hajkova *et al.* [135] recently showed that components of the base excision repair (BER) pathway (i.e., *Parp1*, *Ape1*, and *Xrcc1*) are upregulated in mouse PGCs and zygotes at time points when active DNA demethylation takes place. While it remains to be determined what triggers BER, based on the lack of expression of AID, APOBEC, or DNMT3A/DNMT3B in E11.5 PGCs, the authors suggest that deamination of 5-methylcytosine may not play a major role in PGCs [135]. Another possible DNA demethylation mechanism in mammals depends on oxidation of the 5-methyl group of cytosine [136]. Trypanosome base J, a modified thymine (β -D-glucosyl hydroxymethyluracil), is produced by sequential hydroxylation and glucosylation of the methyl group of thymine [27, 137]. The process requires JBP1 and JBP2, enzymes of the 2OG- and Fe(II)-dependent oxygenases [138, 139]. A computational search for homologs of JBP1 and JBP2 identified TET1, a mammalian enzyme that has now been shown to catalyze the conversion of 5-methylcytosine to 5-hydroxymethylcytosine [136]. A recent study showed that TET1 is involved in maintaining the expression of *Nanog* in ES cells [140]. Downregulation of *Nanog* via TET1 knockdown correlates with hypermethylation of the *Nanog* promoter, indicating that TET1 has an important role in DNA demethylation. It is possible that 5-hydroxymethylcytosine may act as an intermediate that is finally replaced by an unmethylated cytosine through DNA-repair pathways [27, 140]. TET1 is highly expressed in preimplantation embryos and PGCs, and TET1 depletion in preimplantation embryos favors cell specification toward the trophoderm lineage [135, 140]. It would be interesting to determine whether TET1 plays a major role in genome-wide erasure of DNA methylation marks in PGCs and zygotes.

Recently, Okada *et al.* [141] provided evidence for the involvement of the ELP3-containing elongator complex in demethylating the paternal genome in mouse zygotes. The human ELP3 contains a histone acetyltransferase (HAT) domain and a SAM domain. Knockdown of ELP3, as well as of ELP1 and ELP4, impairs paternal DNA demethylation in zygotes. Importantly, the ELP3 radical SAM domain, but not the HAT domain, is re-

quired for paternal DNA demethylation, indicating that the SAM radical domain is essential for DNA demethylation. Demethylation may thus be mediated through a reaction that requires an intact radical SAM domain [141]. Alternatively, the role of the elongator complex in active DNA demethylation may be indirect.

Interplay between histone modifications and DNA methylation

Both DNA methylation and histone modifications are important epigenetic marks for gene regulation. Histone H3 lysine 9 dimethylation (H3K9me2) is an important histone modification for TGS. In *Arabidopsis*, H3K9me2 is catalyzed by the histone methyltransferase SUVH4/KYP. SUVH4/KYP is also required for maintenance of non-CG methylation [142]. SUVH4 is a member of a protein family, in which the proteins contain an SET- and RING-associated (SRA) domain and a SET domain. The SRA domain of SUVH4 can directly bind to methylated DNA. Therefore, DNA methylation is required for recruitment of SUVH4. On the other side, the non-CG DNA methyltransferase CMT3 can directly interact with the N-terminal tail of histone H3, but only when it is simultaneously methylated at both the H3K9 and H3K27 positions [143]. The results suggest that the histone methylation at H3K9 and H3K27, catalyzed by SUVH4 and an unknown enzyme, provides a histone code for the recruitment of CMT3 to methylate the DNA loci. The results also suggest a self-reinforcing loop between histone H3K9 methylation and DNA methylation [144].

VIM1 is a methylcytosine-binding protein that interacts with methylated cytosine at CG and CHG sites via its SRA domain [145]. The *Arabidopsis vim1* mutation causes centromeric DNA hypomethylation and centromeric heterochromatin decondensation in interphase. Moreover, the localization of centromere-specific histone H3 variant, HTR12, was altered in the *vim1* mutant. VIM1, which is a member of a small family of proteins containing PHD, RING, and SRA domains [145, 146], is similar to mammalian ubiquitin-like, containing PHD and RING finger domains 1 (UHRF1). Recent studies demonstrated that UHRF1 is required for maintenance of CG DNA methylation [147, 148]. Similar to UHRF1 in mammals, VIM1 may also be required for recruitment of the maintenance DNA methyltransferase DNMT1/MET1 to hemimethylated DNA loci after DNA replication. VIM1 associates with histone proteins [145], suggesting it is an interface between histone proteins and DNA methylation machinery. In *Arabidopsis*, the *vim1vim2vim3*-triple mutant is late flowering, which is associated with hypomethylation at CG sites in the 5' region of the

flower repressor *FWA* and release of *FWA* gene silencing [146]. Importantly, while the *vim1vim3*-double mutant showed decreased DNA CG methylation in the 5S rRNA genes, it gained ectopic CHH methylation [146]. A similar phenomenon was also found in the *met1* and *ddm1* mutants, in which some specific genome loci are hypermethylated, especially at CHH sites, through the RdDM pathway [149, 150]. Thus, the symmetric CG hypomethylation caused by mutation of *MET1*, *DDM1*, or *VIMs* leads to increased CHH methylation. The work of Mathieu *et al.* [109] suggested that this increase in CHH methylation is caused by RdDM combined with a lack of active DNA demethylation due to downregulation of *ROS1* expression.

De novo DNA methylation in plants is established by *DRM2*, a protein homologous to *DNMT3* in animals. In plants, targeting of *DRM2* relies on the RdDM pathway. However, two SRA-SET proteins, *SUVH2* and *SUVH9*, are also essential for *DRM2*-mediated *de novo* DNA methylation in *Arabidopsis* [151]. In the *svuh2svuh9*-double mutant, DNA methylation of the RdDM targets, *FWA*, and suppressor of *drm1drm2cmt3* (*SDC*), was significantly reduced, especially at asymmetric CHH sites, leading to the release of silencing. Unlike histone methylation in the *svuh4* mutant, histone methylation in the *svuh2svuh9* mutant was apparently not affected [151]. These results indicate that the SRA-domain proteins in *Arabidopsis* are not only involved in symmetric CG and CHG methylation, but also in asymmetric CHH methylation.

Involvement of histone modifications in DNA methylation is also indicated by the study of increase in bonsai methylation 1 (*IBM1*). *IBM1* is a JmjC domain-containing protein in the *JHDM2* family, which consists of histone H3K9 demethylases [152, 153]. *IBM1* was identified by screening for mutants that show ectopic DNA methylation in the genic region of *BONSAI* (*BNS*) [154]. Mutation of *IBM1* induces CHG hypermethylation at *BNS*, which is suppressed by mutations of CHG methylase *CMT3* and histone H3K9 methylase *KYP* [154]. The results are consistent with the positive correlation between CHG methylation and H3K9 methylation [142–144]. A genome-wide analysis revealed that the *ibm1* mutation induces CHG hypermethylation in thousands of genes, while transposons and other repetitive DNA elements are unaffected. *IBM1* might specifically counteract the function of *CMT3* and *KYP*, preventing CHG methylation but not CG methylation in open reading frames on a genome-wide scale [155]. Further study is required to investigate how *IBM1* targets open reading frames and specifically functions at CHG sites. Another *Arabidopsis* JmjC protein, *JMJ14* (a homolog of human *KDM5/JAR-*

ID1 protein), contributes to *de novo* DNA methylation through the RdDM pathway [156]. Mutation of *JMJ14* causes DNA hypomethylation and releases gene silencing at RdDM targets. *JMJ14* seems to promote DNA methylation through demethylation of H3K4me3. *JMJ14* acts downstream from RdDM components *RDR2* and *AGO4* to demethylate histone H3K4me3 residues and to thereby promote DNA methylation [156].

HDA6 is a histone deacetylase that catalyzes the deacetylation of histone H4 at some specific genome loci. Mutation of *HDA6* results in loss of TGS at some RdDM targets, which suggests a function for *HDA6* in RdDM [40, 157]. In *hda6* mutants, symmetric cytosine methylation at CG and CHG sites is reduced, which supports the function of *HDA6* in maintaining DNA methylation [157]. Aberrant RNA Pol II transcripts occur throughout the 45S intergenic spacers [158]. In *hda6* mutants, the transcripts are upregulated, which results in an overproduction of siRNAs. The overproduced siRNAs direct *de novo* DNA methylation at CHH sites, but the increased *de novo* methylation fails to suppress Pol I or Pol II transcription in *hda6* mutants [158]. The results suggest that the histone deacetylation catalyzed by *HDA6* induces symmetric DNA methylation at specific chromatin regions while it suppresses asymmetric DNA methylation. Our previous forward genetic screens identified *SUP32/UBP26*, which encodes an *Arabidopsis* deubiquitination enzyme. Mutation of *SUP32* decreases RNA-directed DNA methylation and histone H3K9 dimethylation, and releases TGS [159]. The study showed that *SUP32* deubiquitinates histone H2B at lysine143. The histone H2B deubiquitination by *SUP32* represents an important histone modification that coordinates heterochromatic histone modifications with DNA methylation [159].

In the *Arabidopsis* genome, DNA methylation is highly correlated with the histone H2A variant, H2A.Z. H2A.Z is absent at hypermethylated DNA regions in the bodies of actively transcribed genes and in methylated transposons [160]. Mutation of the *MET1* DNA methyltransferase causes both loss and gain of DNA methylation and leads to opposite changes (gains and losses) in H2A.Z deposition, whereas mutation of the *PIE1* subunit of the *SWR1* complex that deposits H2A.Z leads to genome-wide hypermethylation [160]. The results indicate that DNA methylation can influence chromatin structure and affect gene silencing by excluding H2A.Z, and that H2A.Z may protect genes from DNA methylation. It is unclear whether this protection from DNA methylation involves active DNA demethylation. The negative correlation of H2A.Z and methylated DNA is conserved between plants and animals [161].

In mammals, the maintenance DNA methyltransferase

DNMT1 can recognize hemimethylated DNA as templates and catalyze DNA methylation on newly synthesized DNA strands. The proliferating cell nuclear antigen of the replication machinery has been known as a component that recruits DNMT1 to hemimethylated DNA replication foci [162]. However, disruption of the interaction causes only minor reduction in DNA methylation [163]. Recent studies provided evidence for the involvement of the multidomain protein UHRF1, also known as NP95 and ICP90, in recruiting DNMT1 to hemimethylated DNA during S phase. UHRF1 directly interacts with DNMT1 and, during S phase, co-localizes with DNMT1. Disruption of *Uhrfl* in mouse ES cells results in severe loss of global DNA methylation [147, 148]. Structural studies reveal that the SRA domain of UHRF1 recognizes hemimethylated DNA and flips 5-methylcytosine out of the DNA helix [164-166].

Polycomb group proteins are required for epigenetic transcriptional repression. The Polycomb repressive complex 2 component enhancer of zeste homolog 2 (EZH2), which catalyzes histone methylation on H3K27, has been shown to associate with DNA methyltransferases. EZH2 seems to be important for DNA methylation of EZH2-target promoters, and may serve as a recruitment platform for DNA methyltransferases [167]. In ES cells, the majority of Polycomb-target genes are marked by both the repressive H3K27me3 and the activating H3K4me3 [168-171]. This 'bivalent' modification pattern is believed to confer a gene the potential to be either activated or repressed. Interestingly, genes marked by H3K27me3 in ES cells often undergo *de novo* DNA methylation in cancer cells [172-174]. Other studies, however, indicated that EZH2 is not required for DNA methylation, suggesting that some additional events are required for *de novo* DNA methylation [175, 176].

Numerous studies have shown a close relationship between histone H3K9 methylation and DNA methylation in animals. The H3K9 methyltransferase G9a is involved in the recruitment of *de novo* DNA methyltransferases DNMT3A and DNMT3B, as well as the maintenance DNA methyltransferase DNMT1 [177, 178]. G9a protein has two distinct functional domains, SET domain and ankyrin domain. The SET domain is responsible for histone H3K9 methyltransferase activity, whereas the ankyrin domain recruits *de novo* DNA methyltransferases DNMT3A and DNMT3B independently of the histone methyltransferase activity. The two activities of G9a are required for silencing of the pluripotency-determining gene *Oct3/4* in early embryos [179]. The mammalian HP1 family members mediate communication between G9a and DNMT1. The methylated H3K9 tails create a binding platform for HP1, which is sufficient for the re-

cruitment of DNMT1 and the silencing of euchromatic genes [180]. In adult neural stem cells, knockdown of G9a leads to demethylation of the *Oct4* promoter and partial reactivation of *Oct4* expression, supporting the correlation between histone methylation and DNA methylation [181]. In zebrafish, DNMT3 is required for proper neurogenesis. The neurogenesis regulator, *LEF1*, is a DNMT3-specific target gene that is demethylated and upregulated in the *dnmt3* mutant. DNMT3 cooperates with G9a in regulating DNA methylation of *LEF1*. DNA methylation and H3K9 trimethylation on the *LEF1* promoter were reduced in both *dnmt3* and *G9a* mutants. The results suggested that the cooperation between the DNA methyltransferase DNMT3 and the histone methyltransferase G9a is required for the silencing of critical regulators during neurogenesis [182].

SUV39H1 and SUV39H2, which mediate H3K9 trimethylation at pericentric heterochromatin [183], have also been shown to regulate DNA methylation (Figure 4). In *svu39h1/h2*-double null mouse ES cells, DNA methylation is impaired at pericentric satellite repeats, but not at other sequences examined. Co-immunoprecipitation experiments showed that DNMT3B interacts with HP1, suggesting that HP1 may act as a bridge between H3K9me3 and DNMT3B. The data indicate that histone methyltransferases SUV39H1/H2 direct H3K9 trimethylation and DNMT3B-dependent DNA methylation to pericentric repeats to reinforce the stability of the heterochromatin regions [184].

Similar to that in plants, in animals, while histone methylation seems to act upstream of DNA methylation in some cases, DNA methylation also affects histone methylation. For example, the methylcytosine-binding protein, MBD1, has been shown to recruit the H3K9 methyltransferase SETDB1 to the large subunit of chromatin assembly factor CAF-1 to form a CAF-1/MBD1/SETDB1 complex that facilitates H3K9 methylation during DNA replication [185]. Moreover, SETDB1 interacts with *de novo* DNA methyltransferases DNMT3A and DNMT3B. SETDB1 and DNMT3A simultaneously occupy the promoter of *RASSF1A*, which is essential for the silencing of this gene in human cancer cells. The results support the functional connection between the histone H3K9 methyltransferase SETDB1 and DNA methyltransferase DNMT3A in epigenetic transcriptional repression [186].

In contrast to H3K9 methylation, which usually facilitates DNA methylation, H3K4 methylation appears to protect DNA from *de novo* methylation. The finding that the ADD domain of DNMT3L binds histone H3 tail only if H3K4 is unmethylated suggests a possible molecular mechanism for the inhibitory effect of H3K4 methylation

on DNA methylation [81, 187]. The recent discovery that H3K4 demethylation by KDM1B is a prerequisite for the establishment of genomic imprinting in female germ cells provides genetic evidence supporting such a mechanism (Figure 2B) [82]. It is worth noting that *DNMT3L*-deficient mice show severe phenotypes in the germline, but not in somatic tissues. There is evidence that DNMT3A and DNMT3B, via their ADD domain, can also bind H3K4-unmodified H3 tail, suggesting that a similar mechanism may be operative in somatic cells [188, 189].

Recent studies have revealed more complex relationships between the histone and DNA methylation systems. Wang *et al.* [190] recently identified DNMT1 as a substrate for the lysine demethylase LSD1 (also known as KDM1A). *Lsd1*-deficient mouse ES cells show progressive loss of global DNA methylation. This loss correlates with a decrease in DNMT1 protein, due to reduced DNMT1 stability. DNMT1 protein is methylated *in vivo*, and LSD1 deficiency enhances DNMT1 methylation (Figure 4) [190]. Furthermore, DNMT1 can be methylated by Set7/9 (an H3K4 methyltransferase) and demethylated by LSD1 *in vitro* [190, 191]. These results suggest that DNMT1 stability is regulated by lysine methylation. LSD1 and Set7/9 (and perhaps other histone methyltransferases and demethylases as well), by acting directly on both histones and DNMT1, may play a role in coordinating histone methylation and DNA methylation.

Function of DNA methylation

Function of DNA methylation in plants

The genomes of higher plants comprise many transposable elements that potentially disrupt genome stability. In plants, high DNA methylation generally occurs in these transposable elements [5]. Methylation of symmetric cytosines mainly relies on the methylase MET1 and chromatin remodeler DDM1, whereas methylation at CHG and CHH sites mainly depends on the methylases CMT3 and DRM2, respectively [36]. Loss of DNA methylation in mutants of these genes leads to the release of transcriptional silencing of transposable elements and other repetitive DNA sequences. Mutation of *DDM1* causes a variety of developmental abnormalities. One of the *ddm1*-induced abnormalities in the *clam* mutant is caused by transposition of an endogenous CACTA family transposon *CAC1* from its original site to the *DWF4* locus [192]. The *bns* mutant is due to silencing of *APC13* in the *ddm1* background, in which a LINE retrotransposon is inserted [193]. Transposition of several LTR retrotransposons was also found in the *ddm1*-induced defective mutants through genomic tiling arrays [194]. In the epigenetic recombinant inbred lines derived from

a cross between WT and *met1*, movement of the LTR transposon *EVD* was detected. The 5'-LTR of *EVD* has a high methylation level in WT plants, but the methylation is erased in *met1* and in those lines with transposition. The *EVD* RNA transcript is present in *met1* but not in WT. The transcript is translated and subjected to reverse transcription, producing extrachromosomal DNA [195]. The extrachromosomal DNA and *EVD* movement are absent in the first generation of the *met1* mutant, but they occur in the second and subsequent generations, which is similar to what happens in the *ddm1* mutant. In the *met1nrpd2a*-double mutant, the *EVD* RNA transcript is synergistically increased, and *EVD* extrachromosomal DNA accumulates. Moreover, a high rate of *EVD* transposition is found in the double mutant [196]. The results indicated that NRPD2, as a component of RdDM, cooperates with MET1 to inhibit *EVD* transposition. In *nrpd2* and other RdDM-single mutant plants, transposons show increased transcript levels but transposition generally does not happen.

Gene imprinting is widespread in mammals as well as in flowering plants, although the underlying mechanisms are different. In *Arabidopsis*, parent-of-origin-specific gene expression is primarily found in the endosperm during seed development. The previously characterized imprinted genes, including *FWA*, *MEA*, *FIS2*, and *PHERES1*, are specifically expressed from the maternal genome of the endosperm while the alleles from the paternal genome are silenced [19, 197-199]. *FWA*, a suppressor of flowering, was initially identified from late-lowering epigenetic mutants that show ectopic *FWA* expression. The hypomethylation of the DNA repeat region in the *FWA* promoter is required for the normal expression of *FWA* [200]. *FWA* is expressed in the endosperm but not in any other tissues. The *FWA* imprint depends on *DME* [197]. Unlike the constitutive expression of *ROS1*, the expression of *DME* is concentrated in the central cells and endosperm during seed development [22]. *DME* specifically demethylates the maternal genome in central cells (progenitors of endosperm) and leads to the expression of imprinted genes (Figure 3B) [19, 23, 24].

Although maternal-origin-specific gene expression was found in the endosperm of *Arabidopsis*, only a small number of imprinted genes (such as *FWA*, *MEA*, *FIS2*, and *PHERES1*) had previously been identified. New imprinted genes were subsequently found when imprinted sequences in the endosperm and embryo were explored genome-wide [23, 24]. The new imprinted genes also show hypomethylation and high expression in endosperms relative to other parts of the plant. Moreover, at the whole-genome level, transposable elements are extensively demethylated in endosperms (Figure

3B). These data suggest that imprinting in plants might have evolved from targeted methylation of transposable elements inserted near genic regulatory elements [23, 24]. The genome-wide demethylation in endosperms is accompanied by extensive non-CG hypermethylation of siRNA-targeted transposon sequences.

Mosher *et al.* [201] found that the 24-nt siRNAs, which are capable of inducing DNA methylation, are highly expressed during early embryogenesis. The predominant phase of Pol IV-dependent 24-nt siRNA accumulation is initiated in the maternal gametophyte and continues during seed development. This discovery of maternally expressed Pol IV-dependent siRNAs in the endosperms of developing seeds greatly expands the catalog of imprinted loci, which now include siRNA-producing sequences throughout the *Arabidopsis* genome [110]. Because the genome-wide demethylation and transcription occur in the genome of central cells and persist in endosperms after fertilization, the widespread demethylation in the maternal-origin genome may contribute to the overexpression of 24-nt siRNAs in endosperms during embryogenesis (Figure 3B). Several recent studies indicated that the genome-wide hypomethylation caused by mutation of *DDMI* or *MET1* induces the expression of 24-nt siRNAs, and thereby activates *de novo* methylation pathways at the hypomethylated sites [149, 150, 202]. The highly expressed 24-nt siRNAs in endosperms may be translocated to embryos and help reinforce silencing of transposable elements in the embryonic genome (Figure 3B).

Transposable elements in the *Arabidopsis* genome are normally methylated and silenced. Pollen grains contain one vegetative cell and two accompanying sperm cells. Transposable elements can be reactivated and transpose in the pollen vegetative cell but not in the sperm cells, which provide DNA to the fertilized zygote [203]. DNA in the pollen vegetative nucleus has reduced methylation, which correlates with the reactivation and transposition of transposable elements. The expression of transposable elements in the vegetative cell may be important to ensure the silencing of these elements in sperm cells [203]. The silencing status of transposable elements in the male gametes is essential for genome stability and integrity. Many genes involved in the RdDM pathway have reduced expression in mature pollen [203, 204]. This reduced expression coincides with the downregulation of 24-nt siRNAs in the vegetative cell nucleus. However, the epigenetic reactivation of transposable elements in the pollen vegetative nucleus leads to the accumulation of 21-nt siRNAs from transposable elements. These 21-nt siRNAs could move to the sperm cell to posttranscriptionally silence transposons that may have escaped TGS

[203]. Although passive DNA demethylation contributes to the reactivation of transposable elements in the pollen vegetative nucleus [203], it is unclear whether active DNA demethylation is also involved.

Mutation of *DDMI* leads to severe loss of DNA methylation in the *Arabidopsis* genome. This loss persists in F1 plants obtained in crosses with the WT, although the *ddm1* mutation is recessive [202]. This finding suggests that DDM1 is involved in the maintenance of DNA methylation but is not sufficient for establishment of DNA methylation. In subsequent generations, however, the hypomethylated DNA loci can be partially remethylated by siRNA-dependent DNA methylation [202]. These remethylatable sequences are characterized by an abundance of siRNAs. Methylation of the remethylatable sequences requires the RNA-directed DNA methylation pathway [202]. The results suggest an important role of RdDM in protecting genomes against long-term epigenetic defects [202, 205].

In *Arabidopsis*, siRNA-dependent *de novo* DNA methylation mainly targets transposons and other repetitive DNA sequences in pericentromeric regions and the end of chromosomes where few genes are distributed. RdDM, however, also functions in gene regulation, especially for those genes flanked by transposons and other repetitive DNA sequences. Some RdDM mutants display late-flowering phenotypes under short-day conditions [39]. *FWA* is a flowering suppressor in *Arabidopsis*. The promoter region of *FWA* is hypermethylated and silenced in WT plants [206]. In *fwa* mutant plants, however, *FWA* is hypomethylated and expressed. *De novo* DNA methylation of *FWA* depends on the RdDM components such as DRM2, NRPD1, RDR2, and DCL3 [207]. The results partially explain the late-flowering phenotype of the RdDM mutants. In the *drm1drm2cmt3* (*ddc*)-triple mutant, DNA methylation in most of the non-symmetrical (CHG and CHH) sequence contexts is lost, and the mutant has defective developmental phenotypes. The gene responsible for the developmental phenotypes has been identified as *SDC*, which encodes an F-box protein and has seven promoter tandem repeats upstream of *SDC*. In WT, the tandem repeats of *SDC* is methylated. The methylation requires DRM2 and CMT3, which are recruited to the repeats in an siRNA-dependent manner. In the *ddc* mutant, DNA methylation of the repeats is reduced and thereby the expression of *SDC* is activated, which causes the developmental phenotypes [208]. For most RdDM components, however, no severe developmental phenotype was found in their defective alleles. An exception is RDM4/DMS4. The *rdm4/dms4* mutants show pleiotropic developmental phenotypes [57, 58]. As a putative transcriptional regulator, RDM4/DMS4 is not only required

for the RdDM pathway by affecting Pol V transcription, but is also critical for Pol II transcription of some coding sequences including developmentally important genes.

In contrast to the lack of strong developmental phenotypes in most RdDM mutants in *Arabidopsis*, mutation of the RDR2 and NRPD1 orthologs in maize causes significant developmental defects [209, 210]. The different effects of RdDM on development between *Arabidopsis* and maize may be caused by the different abundance and distribution of RdDM targets at the genome level. Compared with the small-sized *Arabidopsis* genome, the maize genome has more transposons and other non-coding DNA repeats, and thus more developmentally important genes are flanked and influenced by repetitive sequences that are targeted by RdDM. Moreover, a genome-wide DNA methylation analysis in rice indicated that DNA methylation in rice is enriched in the promoter regions of some endogenous genes. The promoter DNA methylation is associated with transcriptional repression [211]. As a monocot, maize may have a similar DNA methylation pattern with rice. The defective development phenotype in the maize *rdr2* and *nRPD1* mutants is likely due to reduced DNA methylation at the promoter regions of development-related genes. Investigating DNA methylation in plants with larger genomes will enhance our understanding on the regulation and function of DNA methylation.

Paramutation is a well-studied epigenetic phenomenon first described in maize. It is defined as an interaction between two alleles of a single locus, resulting in a heritable change of one allele that is induced by the other allele. The booster 1 (*b1*) is a well-studied locus that shows paramutation. The *b1* locus encodes a bHLH transcription factor, which activates genes involved in anthocyanin biosynthesis in maize [212]. The two alleles of the *b1* locus are *B-I* (paramutable allele, active) and *B'* (paramutagenic allele, inactive), and these alleles possess the same DNA sequence but differ in DNA methylation. When *B-I* and *B'* are crossed, the inactive *B'* allele silences the *B-I* allele and leads to the lack of expression of the bHLH transcription factor gene in the hybrid plants. The tandem repeats ~100 kb upstream of the *b1* locus are required for paramutation [213]. Recent studies with maize determined that MOP1, ZmRPD1, and RMR7 (the orthologs of *Arabidopsis* RDR2, NRPD1, and NRPD2, respectively) are required for paramutation, suggesting that the RdDM pathway is required for establishment and maintenance of silencing in paramutation [210, 214, 215].

In rice, the spontaneous mutant *Epi-d1* shows a metastable dwarf phenotype [216]. The phenotype is mitotically and meiotically heritable and corresponds to the

metastable epigenetic silencing of the *DWARF1* (*D1*) gene. The silenced state is correlated with DNA hypermethylation in the *D1* promoter region. The epigenetic state of *D1* is bidirectionally mutable, from active to repressed and from repressed to active. The bidirectional epigenetic state might indicate that the epigenetic regulation of *D1* is due to *de novo* DNA methylation caused by RNA. Epigenetic alleles such as *Epi-d1* could provide for a mechanism of rapid adaptation to changing environmental conditions [216].

Function of DNA methylation in animals

In mammals, DNA methylation and demethylation are involved in diverse processes including early embryogenesis [217, 218], stem cell differentiation [219, 220], genomic imprinting [75, 221], X chromosome inactivation [222, 223], and silencing of repetitive elements [224]. DNA methylation is also involved in regulating neuronal development and development of cancers [225, 226].

In mammals, complex changes in DNA methylation levels occur during embryonic development. Immediately after fertilization, the paternal and maternal genomes in zygotes differ in DNA methylation levels, i.e., the DNA methylation level is initially higher in the paternal than in the maternal genome [217]. Within 3 to 6 h after fertilization, however, the maternal genome is rapidly methylated through *de novo* DNA methylation, while the paternal genome is actively demethylated in an elongator-dependent manner (Figure 4) [141]. Demethylation of the paternal genome persists in preimplantation embryos. After rapid *de novo* DNA methylation in the maternal genome of zygotes, passive DNA demethylation occurs gradually during DNA replication cycles [217, 227]. The replication-dependent demethylation is due to exclusion of maintenance methyltransferase DNMT1 from the nucleus of embryos [111, 228]. At the morula stage, the parental and maternal genomes have equally low DNA methylation levels (Figure 4). Embryonic DNA methylation patterns are established after implantation through lineage-specific *de novo* methylation. DNA methylation levels increase rapidly in the primitive ectoderm, which gives rise to the entire embryo, whereas methylation is either inhibited or not maintained in the trophoblast and the primitive endoderm lineage, which give rise to the placenta and yolk sac membrane, respectively [217, 218, 229, 230]. Demethylation and *de novo* methylation also occur during gametogenesis, which, as discussed above, are critical for the establishment of genomic imprinting and for the suppression of transposons in germ cells [65].

Our knowledge about the roles of DNA methylation in mammalian development comes mainly from genetic manipulations of DNMTs in mice. Studies of the zygotic

functions of DNMTs have shown that the establishment of embryonic methylation patterns requires both *de novo* and maintenance methyltransferase activities, and that the maintenance of DNA methylation above a threshold level is essential for embryonic development (Figure 4) [71, 72, 74]. Complete elimination of *DNMT1* function results in embryonic lethality around E9.5, with extensive loss of global DNA methylation [72]. *DNMT3B* is also essential for embryogenesis. *DNMT3B*-deficient embryos show growth impairment and multiple developmental defects after E9.5 and die after E12.5. *DNMT3A*-mutant mice die around at 4 weeks of age. *DNMT3A/DNMT3B*-double knockout embryos die around E9.5, similar to *DNMT1*-null mutants [74].

The underlying mechanisms for the developmental defects observed in the *DNMT* mutants are not fully understood. Loss of DNA methylation does not affect ES proliferation and viability, and the effect of demethylation only becomes apparent during or after gastrulation when the pluripotent embryonic cells begin to differentiate [71, 72, 74]. Consistent with these data, *DNMT1*-deficient or *DNMT3A/DNMT3B*-double mutant ES cells show severe differentiation defects [231]. Conditional disruption of *DNMT1* in mouse embryonic fibroblasts (MEFs) results in severe demethylation and cell death, and *DNMT3B*-deficient MEFs show moderate demethylation, chromosomal instability, and abnormal proliferation [232, 233]. These findings suggest that DNA methylation is essential for cellular differentiation and the normal functioning of differentiated cells.

The pluripotency genes are generally hypomethylated in stem cells and gain methylation during cell differentiation [234]. *Oct4* and *Nanog* are two such pluripotency genes, and they are essential for the pluripotency of ES cells. Recently, Kim *et al.* [235] demonstrated that the transcription factor *Oct4* is sufficient to reprogram human neural stem cells to pluripotency. *De novo* DNA methylation and silencing of the pluripotency genes contributes to loss of pluripotency in differentiated cells. Genome-wide DNA methylation analysis indicated widespread differences in DNA methylation between the genomes of ES cells and differentiated cells [4]. Non-CG methylation is widespread in ES cells and induced pluripotent stem cells but not in differentiated cells. Methylation in non-CG contexts shows enrichment in gene bodies and depletion in protein binding sites and enhancers [4]. The results suggest that DNA methylation is highly associated with mammalian development and cell pluripotency. AID is involved in active DNA demethylation through its deaminase activity. It is required for promoter demethylation and induction of *Oct4* and *Nanog* gene expression during reprogramming toward pluripotency

in single heterokaryons [236]. AID protein binds silent methylated *Oct4* and *Nanog* promoters in fibroblasts but not active and demethylated promoters in ES cells. AID-mediated active DNA demethylation is required for nuclear reprogramming toward pluripotency in human somatic cells [236].

Although *DNMT1* is dispensable for ES cell maintenance, it is required for maintaining the somatic progenitor state through cell divisions. Depletion of the maintenance DNA methyltransferase *DNMT1* in epidermal progenitors leads to premature differentiation. Genome-wide DNA methylation analysis showed that some epidermal differentiation gene promoters were methylated in self-renewing progenitor cells but were subsequently demethylated during differentiation [237]. *DNMT1* and *UHRF1*, which target *DNMT1* to hemi-methylated DNA, are involved in suppressing epidermal differentiation gene induction. In contrast, *Gadd45a* and *Gadd45b*, which promote DNA demethylation, are required for full epidermal differentiation gene induction [237]. The results suggest that the dynamic regulation of DNA methylation is important for maintenance and differentiation of progenitor cells.

Somatic tissues have specific gene expression patterns. Because DNA methylation in promoters of genes suppresses gene expression, the correlation between gene expression and DNA methylation was investigated in diverse tissues. *SERPINB5* is a potential tumor suppressor gene expressed in specific human tissues. Its expression level in different tissues highly correlates with the DNA methylation level of its promoter region [238]. Global DNA methylation studies suggest that DNA methylation is critical for regulating the expression of some tissue-specific genes [239-241]. Furthermore, comparison of DNA methylation levels in embryonic tissues derived from different germ layers revealed that differentially methylated regions located about 2 kb apart from CG islands may be involved in tissue-specific gene expression [242]. The results suggested that different tissue types have unique DNA methylation patterns that can contribute to their lineage specificity. As noted earlier, DNA methylation helps regulate neuronal development and development of cancers [225, 226]. In the case of neuronal development, this methylation involves *DNMT3A*, which is expressed in postnatal neural stem cells and is required for neurogenesis [243].

Defects in DNA methylation can cause several developmental disorders and diseases in humans. Deficiency of the *de novo* DNA methyltransferase *DNMT3B* leads to a rare autosomal recessive disorder, ICF syndrome [74, 244]. ICF syndrome is characterized by mental retardation, reduced growth, distinct facial abnormalities, and

immunodeficiency in children [245, 246]. Cells from ICF patients show hypomethylation of classical satellite DNA [247]. MECP2 is a methyl-CG-binding protein, which was identified as the target of mutations that cause Rett syndrome [248]. Rett syndrome is characterized by loss of acquired skills in affected girls. The X-linked *MECP2* contributes to repression of many genome targets. One of the targets is brain-derived neurotrophic factor, an important protein in neuronal plasticity [119]. Gadd45 proteins were identified as key factors that promote DNA demethylation through DNA repair [249]. Mice with Gadd45b deletion exhibit specific deficits in neural activity-induced proliferation of neural progenitors [250]. Gadd45b is required for activity-induced DNA demethylation of specific promoters and expression of corresponding genes critical for adult neurogenesis [250]. Thus, Gadd45b may link neuronal circuit activity to DNA demethylation and gene expression in the mammalian brain.

A hallmark of many cancers is global hypomethylation and regional hypermethylation of CG islands. In mouse models, hypomethylation has been shown to induce genomic instability and tumorigenesis [251, 252]. Although the causes of hypomethylation in cancer cells are poorly understood, a recent study suggests that upregulation of a DNA demethylase system may play a role in some cases [182]. Hypermethylation of CG islands can lead to silencing of tumor suppressor genes and, thus, may also contribute to tumorigenesis. Cancer cells often show overexpression of DNMTs, especially DNMT1, which may cause abnormal hypermethylation. DNMT1 has been shown to be essential for the survival and proliferation of human cancer cells [253].

Conclusions

DNA methylation is a relatively stable epigenetic mark. Nevertheless, it is still dynamically regulated. Plants and animals share some common mechanisms for the regulation of DNA methylation and demethylation. In both plants and animals, maintenance of DNA methylation depends on the maintenance DNA methyltransferases during DNA replication. RNA-directed DNA methylation is a key mechanism that silences repetitive DNA elements in plants, and this mechanism relies on the plant-specific 24-nt siRNAs for *de novo* DNA methylation. In mammals, *de novo* DNA methylation in germ cells also relies on small RNAs, although they are the animal-specific piRNAs. The requirement of small RNAs for at least some *de novo* DNA methylation is evidently shared by plants and mammals. Moreover, both types of small RNAs are loaded onto ARGONAUTE proteins,

and both mainly target transposable elements and other repetitive DNA sequences. Comparison of *de novo* DNA methylation in plants and mammals might help us understand the general aspects of RNA-guided DNA methylation mechanisms.

In mammals, genome-wide DNA methylation and demethylation occur during gametogenesis and early embryogenesis. These processes may involve multiple DNA demethylation mechanisms. ELP3 is involved in active DNA demethylation of the paternal genome in zygotes [141]. TET1 is required for ES cell maintenance and inner cell mass cell specification [140]. Although AID has no genome-wide effect in mammals, it is required for the demethylation and induction of two key pluripotency genes, *Oct4* and *Nanog*, during reprogramming toward pluripotency [236]. In plants, genome-wide DNA methylation changes are found in central cells and endosperms, which are nutrition tissues and which do not pass their DNA to the next generation. It is still unclear whether there is global resetting of DNA methylation during gametogenesis and embryogenesis in plants.

Although active DNA demethylation has been studied, how the process is initiated and regulated is still unknown. ROS3 is an RNA-binding protein that seems to cooperate with ROS1 in the plant DNA demethylation pathway [107]. However, we need to know whether and how non-coding RNAs are involved in DNA demethylation. It would also be useful to know how the DNA methylation and demethylation pathways are coordinated *in vivo*. In mammals, the DNA methyltransferases DNMT3A and DNMT3B are also reported to be involved in DNA demethylation [121]. Learning how DNA methylation and demethylation are reversed will help us understand the dynamic regulation of DNA methylation.

Gene-body methylation is widespread in plants and animals, but the function of such methylation requires further study. Whole-genome DNA methylation and transcription patterns suggest that gene-body methylation is conserved between plants and animals [161, 254]. Gene-body methylation is correlated with the transcription pattern of genes, suggesting an important role of gene-body methylation in gene regulation. A recent study indicated that the DNA methylation in non-promoter intergenic regions and gene bodies promotes gene expression by functionally antagonizing Polycomb repression. Non-promoter DNA methylation seems to be important for the maintenance of active chromatin states of genes [243]. Further studies are required to clarify how gene-body methylation is established and maintained, and how gene transcription may be regulated by gene-body methylation in animals as well as in plants.

Acknowledgements

This work was supported by the National Institutes of Health grants R01GM070795 and R01GM059138 to J-KZ.

References

- 1 Chinnusamy V, Zhu JK. RNA-directed DNA methylation and demethylation in plants. *Sci China C Life Sci* 2009; **52**:331-343.
- 2 Colot V, Rossignol JL. Eukaryotic DNA methylation as an evolutionary device. *Bioessays* 1999; **21**:402-411.
- 3 Ramsahoye B, Biniszkievicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci USA* 2000; **97**:5237-5242
- 4 Lister R, Pelizzola M, Dowen RH, *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 2009; **462**:315-322.
- 5 Zhang X, Yazaki J, Sundaresan A, *et al.* Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* 2006; **126**:1189-1201.
- 6 Weber M, Schubeler D. Genomic patterns of DNA methylation: targets and function of an epigenetic mark. *Curr Opin Cell Biol* 2007; **19**:273-280.
- 7 Chen T, Li E. Structure and function of eukaryotic DNA methyltransferases. *Curr Topics Dev Biol* 2004; **60**:55-89.
- 8 Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. Dnmt3L and the establishment of maternal genomic imprints. *Science* 2001; **294**:2536-2539.
- 9 Hata K, Okano M, Lei H, Li E. Dnmt3L cooperates with the Dnmt3 family of *de novo* DNA methyltransferases to establish maternal imprints in mice. *Development* 2002; **129**:1983-1993.
- 10 Goll MG, Kirpekar F, Maggert KA, *et al.* Methylation of tRNA^{Asp} by the DNA methyltransferase homolog Dnmt2. *Science* 2006; **311**:395-398.
- 11 Ronemus MJ, Galbiati M, Ticknor C, Chen J, Dellaporta SL. Demethylation-induced developmental pleiotropy in *Arabidopsis*. *Science* 1996; **273**:654-657.
- 12 Lindroth, AM, Cao X, Jackson JP, *et al.* Requirement of chromomethylase3 for maintenance of CpXpG methylation. *Science* 2001; **292**:2077-2080.
- 13 Cao X, Jacobsen SE. Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc Natl Acad Sci USA* 2002; **99**:16491-16498.
- 14 Jones L, Ratcliff F, Baulcombe DC. RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. *Curr Biol* 2001; **11**:747-757.
- 15 Cao X, Aufsatz W, Zilberman D, *et al.* Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. *Curr Biol* 2003; **13**:2212-2217.
- 16 Pontes O, Li C F, Nunes PC, *et al.* The *Arabidopsis* chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. *Cell* 2006; **126**:79-92.
- 17 Agius F, Kapoor A, Zhu JK. Role of the *Arabidopsis* DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proc Natl Acad Sci USA* 2006; **103**:11796-11801.
- 18 Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marín MI, Martínez-Macías MI, Ariza RR, Roldán-Arjona T. Demeter and repressor of silencing 1 encode 5-methylcytosine DNA glycosylases. *Proc Natl Acad Sci USA* 2006; **103**:6853-6858.
- 19 Gehring M, Huh JH, Hsieh TF, *et al.* Demeter DNA glycosylase establishes meade polycomb gene self-imprinting by allele-specific demethylation. *Cell* 2006; **124**:495-506.
- 20 Gong Z, Morales-Ruiz T, Ariza RR, *et al.* ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. *Cell* 2002; **111**:803-814.
- 21 Zhu J, Kapoor A, Sridhar VV, *et al.* The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation patterns in *Arabidopsis*. *Curr Biol* 2007; **17**:54-59.
- 22 Choi Y, Gehring M, Johnson L, *et al.* DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. *Cell* 2002; **110**:33-42.
- 23 Gehring M, Bubb KL, Henikoff S. Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science* 2009; **324**:1447-1451.
- 24 Hsieh TF, Ibarra CA, Silva P, *et al.* Genome-wide demethylation of *Arabidopsis* endosperm. *Science* 2009; **324**:1451-1454.
- 25 Zhu JK. Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet* 2009; **43**:143-166.
- 26 Ooi SK, Bestor TH. The colorful history of active DNA demethylation. *Cell* 2008; **133**:1145-1148.
- 27 Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol* 2010; **11**:607-620.
- 28 Zhu JK. Epigenome sequencing comes of age. *Cell* 2008; **133**:395-397.
- 29 Zilberman D, Gehring M, Tran RK, *et al.* Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* 2007; **39**:61-69.
- 30 Lister R, O'Malley RC, Tonti-Filippini J, *et al.* Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* 2008; **133**:523-536.
- 31 Wierzbicki AT, Haag JR, Pikaard CS. Non-coding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* 2008; **135**:635-648.
- 32 Wassenegger M, Heimes S, Riedel L, *et al.* RNA-directed *de novo* methylation of genomic sequences in plants. *Cell* 1994; **76**:567-576.
- 33 Xie Z, Johansen L K, Gustafson A M, *et al.* Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol* 2004; **2**:E104.
- 34 Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. RNA polymerase IV directs silencing of endogenous DNA. *Science* 2005; **308**:118-120.
- 35 Zaratiegui M, Irvine DV, Martienssen RA. Noncoding RNAs and gene silencing. *Cell* 2007; **128**:763-776.
- 36 Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Rev Genet* 2010; **11**:204-220.
- 37 Kanno T, Huettel B, Mette M F, *et al.* Atypical RNA poly-

- merase subunits required for RNA-directed DNA methylation. *Nat Genet* 2005; **37**: 761-765.
- 38 Onodera Y, Onodera Y, Haag JR, *et al.* Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* 2005; **120**:613-622.
- 39 Pontier D, Yahubyan G, Vega 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* 2005; **19**:2030-2040.
- 40 He XJ, Hsu YF, Pontes O, *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* 2009; **23**:318-330.
- 41 Lahmy S, Pontier D, Cavel E, *et al.* PolIV (PolIVb) function in RNA-directed DNA methylation requires the conserved active site and an additional plant-specific subunit. *Proc Natl Acad Sci USA* 2009; **106**:941-946.
- 42 Ream TS, Haag JR, Wierzbicki AT, *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* 2008; **33**:192-203.
- 43 Huang L, Jones AM, Searle I, *et al.* An atypical RNA polymerase involved in RNA silencing shares small subunits with RNA polymerase II. *Nat Struct Mol Biol* 2009; **16**:91-93.
- 44 Yu B, Yang Z, Li J, *et al.* Methylation as a crucial step in plant microRNA biogenesis. *Science* 2005; **307**:932-935.
- 45 Yang Z, Ebright YW, Yu B, Chen X. HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Res* 2006; **34**:667-675.
- 46 Li CF, Pontes O, El-Shami M, *et al.* An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. *Cell* 2006; **126**:93-106.
- 47 Zheng X, Zhu J, Kapoor A, Zhu JK. Role of *Arabidopsis* AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. *EMBO J* 2007; **26**:1691-1701.
- 48 He XJ, Hsu YF, Zhu S, *et al.* An effector of RNA-directed DNA methylation in *Arabidopsis* is an ARGONAUTE 4- and RNA-binding protein. *Cell* 2009; **137**:498-508.
- 49 Wierzbicki AT, Ream TS, Haag JR, Pikaard CS. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nature Genet* 2009; **41**:630-634.
- 50 Zilberman D, Cao X, Jacobsen SE. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* 2003; **299**:716-719.
- 51 Ausin I, Mockler TC, Chory J, Jacobsen SE. IDN1 and IDN2 are required for *de novo* DNA methylation in *Arabidopsis thaliana*. *Nature Struct Mol Biol* 2009; **16**:1325-1327.
- 52 Zheng Z, Xing Y, He XJ, *et al.* An SGS3-like protein functions in RNA-directed DNA methylation and transcriptional gene silencing in *Arabidopsis*. *Plant J* 2010; **62**:92-99.
- 53 Gao Z, Liu HL, Daxinger L, *et al.* An RNA polymerase II- and AGO4-associated protein acts in RNA-directed DNA methylation. *Nature* 2010; **465**:106-109.
- 54 Cramer P. Multisubunit RNA polymerases. *Curr Opin Struct Biol* 2002; **12**:89-97.
- 55 Kornberg RD. The molecular basis of eukaryotic transcription. *Proc Natl Acad Sci USA* 2007; **104**:12955-12961.
- 56 Wade JT, Struhl K. The transition from transcriptional initiation to elongation. *Curr Opin Genet Dev* 2008; **18**:130-136.
- 57 He XJ, Hsu YF, Zhu S, *et al.* A conserved transcriptional regulator is required for RNA-directed DNA methylation and plant development. *Genes Dev* 2009; **23**:2717-2722.
- 58 Kanno T, Bucher E, Daxinger L, *et al.* RNA-directed DNA methylation and plant development require an IWR1-type transcription factor. *EMBO Rep* 2010; **11**:65-71.
- 59 Bies-Etheve N, Pontier D, Lahmy S, *et al.* RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. *EMBO Rep* 2009; **10**:649-654.
- 60 Zheng B, Wang Z, Li S, Yu B, Liu JY, Chen X. Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. *Genes Dev* 2009; **23**:2850-2860.
- 61 Kanno T, Mette MF, Kreil DP, *et al.* Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr Biol* 2004; **14**:801-805.
- 62 Kanno T, Bucher E, Daxinger L, *et al.* A structural-maintenance-of-chromosomes-hinge-domain-containing protein is required for RNA-directed DNA methylation. *Nat Genet* 2008; **40**:670-675.
- 63 Law JA, Ausin I, Johnson LM, *et al.* A protein complex required for polymerase V transcripts and RNA-directed DNA methylation in *Arabidopsis*. *Curr Biol* 2010; **20**:951-956.
- 64 Smith L M, Pontes O, Searle I, *et al.* An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in *Arabidopsis*. *Plant Cell* 2007; **19**:1507-1521.
- 65 Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. *Science* 2001; **293**:1089-1093.
- 66 Li E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet* 2002; **3**:662-673.
- 67 Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 2003; **33** Suppl:245-254.
- 68 Leonhardt H, Page AW, Weier HU, Bestor TH. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 1992; **71**:865-873.
- 69 Yoder JA, Soman NS, Verdine GL, Bestor TH. DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. *J Mol Biol* 1997; **270**:385-395.
- 70 Pradhan S, Bacolla A, Wells RD, Roberts RJ. Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of *de novo* and maintenance methylation. *J Biol Chem* 1999; **274**:33002-33010.
- 71 Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 1992; **69**:915-926.
- 72 Lei H, Oh SP, Okano M, *et al.* *De novo* DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* 1996; **122**:3195-3205.
- 73 Okano M, Xie S, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet* 1998; **19**:219-220.
- 74 Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation

- and mammalian development. *Cell* 1999; **99**:247-257.
- 75 Reik W, Walter J. Genomic imprinting: parental influence on the genome. *Nat Rev Genet* 2001; **2**:21-32.
- 76 Kaneda M, Okano M, Hata K, *et al.* Essential role for *de novo* DNA methyltransferases Dnmt3a in paternal and maternal imprinting. *Nature* 2004; **429**:900-903.
- 77 Chedin F, Lieber MR, Hsieh CL. The DNA methyltransferase-like protein DNMT3L stimulates *de novo* methylation by Dnmt3a. *Proc Natl Acad Sci USA* 2002; **99**:16916-16921.
- 78 Margot JB, Ehrenhofer-Murray AE, Leonhardt H. Interactions within the mammalian DNA methyltransferase family. *BMC Mol Biol* 2003; **4**:7.
- 79 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* 2004; **279**:27816-27823.
- 80 Gowher H, Liebert K, Hermann A, Xu G, Jeltsch A. Mechanism of Stimulation of Catalytic Activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *J Biol Chem* 2005; **280**:13341-13348.
- 81 Ooi SK, Qiu C, Bernstein E, *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature* 2007; **448**:714-717.
- 82 Ciccone DN, Su H, Hevi S, *et al.* KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* 2009; **461**:415-418.
- 83 Chotalia M, Smallwood SA, Ruf N, *et al.* Transcription is required for establishment of germline methylation marks at imprinted genes. *Genes Dev* 2009; **23**:105-117.
- 84 Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* 2004; **431**:96-99.
- 85 Kato Y, Kaneda M, Hata K, *et al.* Role of the Dnmt3 family in *de novo* methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum Mol Genet* 2007; **16**:2272-2280.
- 86 Kuramochi-Miyagawa S, Kimura T, Ijiri TW, *et al.* Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. *Development* 2004; **131**:839-849.
- 87 Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* 2007; **316**:744-747.
- 88 Carmell MA, Girard A, van de Kant HJ, *et al.* MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev Cell* 2007; **12**:503-514.
- 89 Aravin AA, Hannon GJ, Brennecke J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 2007; **318**:761-764.
- 90 Hartig JV, Tomari Y, Forstemann K. piRNAs--the ancient hunters of genome invaders. *Genes Dev* 2007; **21**:1707-1713.
- 91 Thomson T, Lin H. The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. *Annu Rev Cell Dev Biol* 2009; **25**:355-376.
- 92 Kuramochi-Miyagawa S, Watanabe T, Gotoh K, *et al.* DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev* 2008; **22**:908-917.
- 93 Aravin AA, Sachidanandam R, Bourc'his D, *et al.* A piRNA pathway primed by individual transposons is linked to *de novo* DNA methylation in mice. *Mol Cell* 2008; **31**:785-799.
- 94 Reuter M, Chuma S, Tanaka T, Franz T, Stark A, Pillai RS. Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. *Nat Struct Mol Biol* 2009; **16**:639-646.
- 95 Vagin VV, Wohlschlegel J, Qu J, *et al.* Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. *Genes Dev* 2009; **23**:1749-1762.
- 96 Chen C, Jin J, James DA, *et al.* Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. *Proc Natl Acad Sci USA* 2009; **106**:20336-20341.
- 97 Kojima K, Kuramochi-Miyagawa S, Chuma S, *et al.* Associations between PIWI proteins and TDRD1/MTR-1 are critical for integrated subcellular localization in murine male germ cells. *Genes Cells* 2009; **14**:1155-1165.
- 98 Wang J, Saxe JP, Tanaka T, *et al.* Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis. *Curr Biol* 2009; **19**:640-644.
- 99 Noce T, Okamoto-Ito S, Tsunekawa N. Vasa homolog genes in mammalian germ cell development. *Cell Struct Funct* 2001; **26**:131-136.
- 100 Toyooka Y, Tsunekawa N, Takahashi Y, Matsui Y, Satoh M, Noce T. Expression and intracellular localization of mouse Vasa-homologue protein during germ cell development. *Mech Dev* 2000; **93**:139-149.
- 101 Kuramochi-Miyagawa S, Watanabe T, Gotoh K, *et al.* MVH in piRNA processing and gene silencing of retrotransposons. *Genes Dev* 2010; **24**:887-892.
- 102 Zheng K, Xiol J, Reuter M, *et al.* Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway. *Proc Natl Acad Sci USA* 2010; **107**:11841-11846.
- 103 Frost RJ, Hamra FK, Richardson JA, Qi X, Bassel-Duby R, Olson EN. MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. *Proc Natl Acad Sci USA* 2010; **107**:11847-11852.
- 104 Ishitani M, Xiong L, Stevenson B, Zhu JK. Genetic analysis of osmotic and cold stress signal transduction in *Arabidopsis*: interactions and convergence of abscisic acid-dependent and abscisic acid-independent pathways. *Plant Cell* 1997; **9**:1935-1949.
- 105 Penterman J, Uzawa R, Fischer RL. Genetic interactions between DNA demethylation and methylation in *Arabidopsis*. *Plant physiol* 2007; **145**:1549-1557.
- 106 Penterman J, Zilberman D, Huh JH, *et al.* DNA demethylation in the *Arabidopsis* genome. *Proc Natl Acad Sci USA* 2007; **104**:6752-6757.
- 107 Zheng X, Pontes O, Zhu J, *et al.* ROS3 is an RNA-binding protein required for DNA demethylation in *Arabidopsis*. *Nature* 2008; **455**:1259-1262.
- 108 Huettel B, Kanno T, Daxinger L, Aufsatz W, Matzke AJ, Matzke M. Endogenous targets of RNA-directed DNA methylation and Pol IV in *Arabidopsis*. *EMBO J* 2006; **25**:2828-2836.
- 109 Mathieu O, Reinders J, Caikovski M, Smathajitt C, Paszkowski J. Transgenerational stability of the *Arabidopsis* epigenome is coordinated by CG methylation. *Cell* 2007;

- 130**:851-862.
- 110 Mosher RA, Schwach F, Studholme D, Baulcombe DC. 2008. PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proc Natl Acad Sci USA* 2008; **105**:3145-3150.
- 111 Carlson LL, Page AW, Bestor TH. Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes Dev* 1992; **6**:2536-2541.
- 112 Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. *Nature* 1993; **366**:362-365.
- 113 Howell CY, Bestor TH, Ding F, *et al.* Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell* 2001; **104**:829-838.
- 114 Hirasawa R, Chiba H, Kaneda M, *et al.* Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNAmethylation imprints during preimplantation development. *Genes Dev* 2008; **22**:1607-1616.
- 115 Li X, Ito M, Zhou F, *et al.* A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev Cell* 2008; **15**:547-557.
- 116 Nakamura T, Arai Y, Umehara H, *et al.* PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat Cell Biol* 2007; **9**:64-71.
- 117 Hajkova P, Erhardt S, Lane N, *et al.* Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev* 2002; **117**:15-23.
- 118 Bruniquel D, Schwartz RH. Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. *Nature Immunol* 2003; **4**:235-240.
- 119 Martinowich K, Hattori D, Wu H, *et al.* DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* 2003; **302**:890-893.
- 120 Kangaspeska S, Stride B, Métiévier R, *et al.* Transient cyclical methylation of promoter DNA. *Nature* 2008; **452**:112-115.
- 121 Métiévier R, Gallais R, Tiffoche C, *et al.* Cyclical DNA methylation of a transcriptionally active promoter. *Nature* 2008; **452**, 45-50.
- 122 Jost JP. Nuclear extracts of chicken embryos promote an active demethylation of DNA by excision repair of 5-methyldeoxycytidine. *Proc Natl Acad Sci USA* 1993; **90**:4684-4688.
- 123 Jost JP, Siegmann M, Sun L, Leung R. Mechanisms of DNA demethylation in chicken embryos. Purification and properties of a 5-methylcytosine-DNA glycosylase. *J Biol Chem* 1995; **270**:9734-9739.
- 124 Zhu B, Zheng Y, Hess D, *et al.* 5-methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. *Proc Natl Acad Sci USA* 2000; **97**:5135-5139.
- 125 Zhu B, Zheng Y, Angliker H, *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* 2000; **28**:4157-4165.
- 126 Kim MS, Kondo T, Takada I, *et al.* DNA demethylation in hormone induced transcriptional derepression. *Nature* 2009; **461**:1007-1012.
- 127 Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* 2004; **279**:52353-52360.
- 128 Rai K, Huggins IJ, James SR, Karpf AR, Jones DA, Cairns BR. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and GADD45. *Cell* 2008; **135**:1201-1212.
- 129 Popp C, Dean W, Feng S, *et al.* Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 2010; **463**:1101-1105.
- 130 Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 2000; **102**:553-563.
- 131 Revy P, Muto T, Levy Y, *et al.* Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the hyper-IgM syndrome (HIGM2). *Cell* 2000; **102**:565-575.
- 132 Wong E, Yang K, Kuraguchi M, *et al.* Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation. *Proc Natl Acad Sci USA* 2002; **99**:14937-14942.
- 133 Li YQ, Zhou PZ, Zheng XD, Walsh CP, Xu GL. Association of Dnmt3a and thymine DNA glycosylase links DNA methylation with base-excision repair. *Nucleic Acids Res* 2007; **35**:390-400.
- 134 Boland MJ, Christman JK. Characterization of Dnmt3b: thymine-DNA glycosylase interaction and stimulation of thymine glycosylase-mediated repair by DNA methyltransferase(s) and RNA. *J Mol Biol* 2008; **379**:492-504.
- 135 Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, Surani MA. Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* 2010; **329**:78-82.
- 136 Tahiliani M, Koh KP, Shen Y, *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 2009; **324**:930-935.
- 137 Borst P, Sabatini R. Base J: discovery, biosynthesis, and possible functions. *Annu Rev Microbiol* 2008; **62**:235-251.
- 138 Yu Z, Genest PA, ter Riet B, *et al.* The protein that binds to DNA base J in trypanosomatids has features of a thymidine hydroxylase. *Nucleic Acids Res* 2007; **35**:2107-2115.
- 139 Cliffe LJ, Kieft R, Southern T, *et al.* JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes. *Nucleic Acids Res* 2009; **37**:1452-1462.
- 140 Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 2010; **466**:1129-1133.
- 141 Okada Y, Yamagata K, Hong K, Wakayama T, Zhang Y. A role for the elongator complex in zygotic paternal genome demethylation. *Nature* 2010; **463**:554-558.
- 142 Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Control of CpNpG DNA methylation by the kryptonite histone H3 methyltransferase. *Nature* 2002; **416**:556-560.
- 143 Lindroth AM, Shultis D, Jasencakova Z, *et al.* Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with chromomethylase3. *EMBO J* 2004; **23**:4286-

- 4296.
- 144 Johnson LM, Bostick M, Zhang X, *et al.* The SRA methylcytosine-binding domain links DNA and histone methylation. *Curr Biol* 2007; **17**:379-384.
 - 145 Woo HR, Pontes O, Pikaard CS, Richards EJ. VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. *Genes Dev* 2007; **21**:267-277.
 - 146 Woo HR, Dittmer TA, Richards EJ. Three SRA-domain methylcytosine-binding proteins cooperate to maintain global CpG methylation and epigenetic silencing in *Arabidopsis*. *PLoS Genet* 2008; **4**:e1000156.
 - 147 Bostick M, Kim JK, Esteve PO, *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 2007; **317**:1760-1764.
 - 148 Sharif J, Muto M, Takebayashi S, *et al.* The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 2007; **450**:908-912.
 - 149 Saze H, Mittelsten Scheid O, Paszkowski J. Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nat Genet* 2003; **34**:65-69.
 - 150 Blevins T, Pontes O, Pikaard CS, Meins F Jr. Heterochromatic siRNAs and DDM1 independently silence aberrant 5S rDNA transcripts in *Arabidopsis*. *PLoS One* 2009; **4**:e5932.
 - 151 Johnson LM, Law JA, Khattar A, Henderson IR, Jacobsen SE. SRA-domain proteins required for DRM2-mediated *de novo* DNA methylation. *PLoS Genet* 2008; **4**:e1000280.
 - 152 Yamane K, Toumazou C, Tsukada Y, *et al.* JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell* 2006; **125**:483-495.
 - 153 Klose RJ, Yamane K, Bae Y, *et al.* The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* 2006; **442**:312-316.
 - 154 Saze H, Shiraishi A, Miura A, Kakutani T. Control of genic DNA methylation by a jmjC domain-containing protein in *Arabidopsis thaliana*. *Science* 2008; **319**:462-465.
 - 155 Miura A, Nakamura M, Inagaki S, Kobayashi A, Saze H, Kakutani T. An *Arabidopsis* jmjC domain protein protects transcribed genes from DNA methylation at CHG sites. *EMBO J* 2009; **28**:1078-1086.
 - 156 Searle IR, Pontes O, Melnyk CW, Smith LM, Baulcombe DC. MJM14, a JmjC domain protein, is required for RNA silencing and cell-to-cell movement of an RNA silencing signal in *Arabidopsis*. *Genes Dev* 2010; **24**:986-991.
 - 157 Aufsatz W, Mette MF, van der Winden J, *et al.* HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA. *EMBO J* 2002; **21**:6832-6841.
 - 158 Earley KW, Pontvianne F, Wierzbicki AT, *et al.* Mechanisms of HDA6-mediated rRNA gene silencing: suppression of intergenic Pol II transcription and differential effects on maintenance versus siRNA-directed cytosine methylation. *Genes Dev* 2010; **24**:1119-1132.
 - 159 Sridhar VV, Kapoor A, Zhang K, *et al.* Control of DNA methylation and heterochromatic silencing by histone H2B deubiquitination. *Nature* 2007; **447**:735-738.
 - 160 Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S. Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* 2008; **456**:125-129.
 - 161 Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* 2010; **328**:916-919.
 - 162 Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF. Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* 1997; **277**:1996-2000.
 - 163 Egger G, Jeong S, Escobar SG, *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* 2006; **103**:14080-14085.
 - 164 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* 2008; **455**:818-821.
 - 165 Avvakumov GV, Walker JR, Xue S, *et al.* Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature* 2008; **455**:822-825.
 - 166 Hashimoto H, Horton JR, Zhang X, Bostick M, Jacobsen SE, Cheng X. The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* 2008; **455**:826-829.
 - 167 Vire E, Brenner C, Deplus R, *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 2006; **439**:871-874.
 - 168 Bernstein BE, Mikkelsen TS, Xie X, *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 2006; **125**:315-326.
 - 169 Mikkelsen TS, Ku M, Jaffe DB, *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 2007; **448**:553-560.
 - 170 Pan G, Tian S, Nie J, *et al.* Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* 2007; **1**:299-312.
 - 171 Zhao XD, Han X, Chew JL, *et al.* Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Cell Stem Cell* 2007; **1**:286-298.
 - 172 Schlesinger Y, Straussman R, Keshet I, *et al.* Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for *de novo* methylation in cancer. *Nat Genet* 2007; **39**:232-236.
 - 173 Ohm JE, McGarvey KM, Yu X, *et al.* A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nat Genet* 2007; **39**:237-242.
 - 174 Widschwendter M, Fiegl H, Egle D, *et al.* Epigenetic stem cell signature in cancer. *Nat Genet* 2007; **39**:157-158.
 - 175 Kondo Y, Shen L, Cheng AS, *et al.* Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. *Nat Genet* 2008; **40**:741-750.
 - 176 Rush M, Appanah R, Lee S, Lam LL, Goyal P, Lorincz MC. Targeting of EZH2 to a defined genomic site is sufficient for recruitment of Dnmt3a but not *de novo* DNA methylation. *Epigenetics* 2009; **4**:404-414.
 - 177 Feldman N, Gerson A, Fang J, *et al.* G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* 2006; **8**:188-194.
 - 178 Esteve PO, Chin HG, Smallwood A, *et al.* Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev* 2006; **20**:3089-3103.

- 179 Epsztejn-Litman S, Feldman N, Abu-Remaileh M, *et al.* *De novo* DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol* 2008; **15**:1176-1183.
- 180 Smallwood A, Esteve PO, Pradhan S, Carey M. Functional cooperation between HP1 and DNMT1 mediates gene silencing. *Genes Dev* 2007; **21**:1169-1178.
- 181 Ma DK, Chiang CH, Ponnusamy K, Ming GL, Song H. G9a and Jhdm2a regulate embryonic stem cell fusion-induced reprogramming of adult neural stem cells. *Stem Cells* 2008; **26**:2131-2141.
- 182 Rai K, Sarkar S, Broadbent TJ, *et al.* DNA demethylase activity maintains intestinal cells in an undifferentiated state following loss of APC. *Cell* 2010; **142**:930-942.
- 183 Peters AH, O'Carroll D, Scherthan H, *et al.* Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* 2001; **107**:323-337.
- 184 Lehnertz B, Ueda Y, Derijck AA, *et al.* Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* 2003; **13**:1192-1200.
- 185 Sarraf SA, Stancheva I. Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Mol Cell* 2004; **15**:595-605.
- 186 Li H, Rauch T, Chen ZX, Szabó PE, Riggs AD, Pfeifer GP. The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. *J Biol Chem* 2006; **281**:19489-19500.
- 187 Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. Structure of Dnmt3a bound to Dnmt3L suggests a model for *de novo* DNA methylation. *Nature* 2007; **449**:248-251.
- 188 Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M, Shirakawa M. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep* 2009; **10**:1235-1241.
- 189 Zhang Y, Jurkowska R, Soeroes S, *et al.* Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res* 2010; **38**:4246-4253.
- 190 Wang J, Hevi S, Kurash JK, *et al.* The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nat Genet* 2009; **41**:125-129.
- 191 Esteve PO, Chin HG, Benner J, *et al.* Regulation of DNMT1 stability through SET7-mediated lysine methylation in mammalian cells. *Proc Natl Acad Sci USA* 2009; **106**:5076-5081.
- 192 Miura A, Yonebayashi S, Watanabe K, Toyama T, Shimada H, Kakutani T. Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. *Nature* 2001; **411**:212-214.
- 193 Saze H, Kakutani T. Heritable epigenetic mutation of a transposon-flanked *Arabidopsis* gene due to lack of the chromatin-remodeling factor DDM1. *EMBO J* 2007; **26**:3641-3652.
- 194 Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T. Bursts of retrotransposition reproduced in *Arabidopsis*. *Nature* 2009; **461**:423-426.
- 195 Beauregard A, Curcio MJ, Belfort M. The take and give between retrotransposable elements and their hosts. *Annu Rev Genet* 2008; **42**:587-617.
- 196 Mirouze M, Reinders J, Bucher E, *et al.* Selective epigenetic control of retrotransposition in *Arabidopsis*. *Nature* 2009; **461**:427-430.
- 197 Kinoshita T, Miura A, Choi Y, *et al.* One-way control of FWA imprinting in *Arabidopsis* endosperm by DNA methylation. *Science* 2004; **303**:521-523.
- 198 Jullien PE, Katz A, Oliva M, Ohad N, Berger F. Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in *Arabidopsis*. *Curr Biol* 2006; **16**:486-492.
- 199 Kohler C, Page DR, Gagliardini V, Grossniklaus U. The *Arabidopsis thaliana* MEDEA Polycomb group protein controls expression of PHERES1 by parental imprinting. *Nat Genet* 2005; **37**:28-30.
- 200 Kinoshita Y, Saze H, Kinoshita T, *et al.* Control of FWA gene silencing in *Arabidopsis thaliana* by SINE-related direct repeats. *Plant J* 2007; **49**:38-45.
- 201 Mosher RA, Melnyk CW, Kelly KA, Dunn RM, Studholme DJ, Baulcombe DC. Uniparental expression of PolIV-dependent siRNAs in developing endosperm of *Arabidopsis*. *Nature* 2009; **460**:283-286.
- 202 Teixeira FK, Heredia F, Sarazin A, *et al.* A role for RNAi in the selective correction of DNA methylation defects. *Science* 2009; **323**:1600-1604.
- 203 Slotkin RK, Vaughn M, Borges F, *et al.* Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* 2009; **136**:461-472.
- 204 Pina C, Pinto F, Feijó JA, Becker JD. Gene family analysis of the *Arabidopsis* pollen transcriptome reveals biological implications for cell growth, division control, and gene expression regulation. *Plant Physiol* 2005; **138**:744-756.
- 205 Henderson IR, Jacobsen SE. Epigenetic inheritance in plants. *Nature* 2007; **447**:418-424.
- 206 Soppe WJ, Jacobsen SE, Alonso-Blanco C, *et al.* The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol Cell* 2000; **6**:791-802.
- 207 Chan SW, Zhang X, Bernatavichute YV, Jacobsen SE. Two-step recruitment of RNA-directed DNA methylation to tandem repeats. *PLoS Biol* 2006; **4**:e363.
- 208 Henderson IR, Jacobsen SE. Tandem repeats upstream of the *Arabidopsis* endogene SDC recruit non-CG DNA methylation and initiate siRNA spreading. *Genes Dev* 2008; **22**:1597-1606.
- 209 Parkinson SE, Gross SM, Hollick JB. Maize sex determination and abaxial leaf fates are canalized by a factor that maintains repressed epigenetic states. *Dev Biol* 2007; **308**:462-473.
- 210 Erhard KF Jr, Stonaker JL, Parkinson SE, *et al.* RNA polymerase IV functions in paramutation in *Zea mays*. *Science* 2009; **323**:1201-1205.
- 211 Li X, Wang X, He K, *et al.* High-resolution mapping of epigenetic modifications of the rice genome uncovers interplay between DNA methylation, histone methylation, and gene expression. *Plant Cell* 2008; **20**:259-276.
- 212 Chandler VL, Stam M. Chromatin conversations: mechanisms and implications of paramutation. *Nat Rev Genet* 2004; **5**:532-544.

- 213 Stam M, Belele C, Dorweiler JE, Chandler VL. Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. *Genes Dev* 2002; **16**:1906-1918.
- 214 Alleman M, Sidorenko L, McGinnis K, *et al.* An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature* 2006; **442**:295-448.
- 215 Stonaker JL, Lim JP, Erhard KF Jr, Hollick JB. Diversity of Pol IV function is defined by mutations at the maize *rmr7* locus. *PLoS Genet* 2009; **5**:e1000706.
- 216 Miura K, Agetsuma M, Kitano H, *et al.* A metastable DWARF1 epigenetic mutant affecting plant stature in rice. *Proc Natl Acad Sci USA* 2009; **106**:11218-11223.
- 217 Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. Demethylation of the zygotic paternal genome. *Nature* 2000; **403**:501-502.
- 218 Haaf T. Methylation dynamics in the early mammalian embryo: implications of genome reprogramming defects for development. *Curr Top Microbiol Immunol* 2006; **310**:13-22.
- 219 Watanabe D, Suetake I, Tada T, Tajima S. Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. *Mech Dev* 2002; **118**:187-190.
- 220 Latham T, Gilbert N, Ramsahoye B. DNA methylation in mouse embryonic stem cells and development. *Cell Tissue Res* 2008; **331**:31-55.
- 221 Ideraabdullah FY, Vigneau S, Bartolomei MS. Genomic imprinting mechanisms in mammals. *Mutat Res* 2008; **647**:77-85.
- 222 Heard E, Disteché CM. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev* 2006; **20**:1848-1867.
- 223 Senner CE, Brockdorff N. Xist gene regulation at the onset of X inactivation. *Curr Opin Genet Dev* 2009; **19**:122-126.
- 224 Chen RZ, Pettersson U, Beard C, Jackson-Grusby L, Jaenisch R. DNA hypomethylation leads to elevated mutation rates. *Nature* 1998; **395**:89-93.
- 225 Dulac C. Brain function and chromatin plasticity. *Nature* 2010; **465**:728-735.
- 226 Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. *Nat Rev Genet* 2006; **7**:21-33.
- 227 Rougier N, Bourc'his D, Gomes DM, *et al.* Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev* 1998; **12**:2108-2113.
- 228 Cardoso MC, Leonhardt H. DNA methyltransferase is actively retained in the cytoplasm during early development. *J Cell Biol* 1999; **147**:25-32.
- 229 Kafri T, Ariel M, Brandeis M, *et al.* Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev* 1992; **6**:705-714.
- 230 Santos F, Hendrich B, Reik W, Dean W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* 2002; **241**:172-182.
- 231 Tucker KL, Talbot D, Lee MA, Leonhardt H, Jaenisch R. Complementation of methylation deficiency in embryonic stem cells by DNA methyltransferase minigene. *Proc Natl Acad Sci USA* 1996; **93**:12920-12925.
- 232 Jackson-Grusby L, Beard C, Possemato R, *et al.* Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat Genet* 2001; **27**:31-39.
- 233 Dodge JE, Okano M, Dick F, *et al.* Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. *J Biol Chem* 2005; **280**:17986-17991.
- 234 Farthing CR, Ficz G, Ng RK, *et al.* Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS Genet* 2008; **4**:e1000116.
- 235 Kim JB, Greber B, Araúzo-Bravo MJ, *et al.* Direct reprogramming of human neural stem cells by OCT4. *Nature* 2009; **461**:649-653.
- 236 Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* 2010; **463**:1042-1047.
- 237 Sen GL, Reuter JA, Webster DE, Zhu L, Khavari PA. DNMT1 maintains progenitor function in self-renewing somatic tissue. *Nature* 2010; **463**:563-567.
- 238 Futscher BW, Oshiro MM, Wozniak RJ, *et al.* Role for DNA methylation in the control of cell type specific maspin expression. *Nat Genet* 2002; **31**:175-179.
- 239 Shen L, Kondo Y, Guo Y, *et al.* Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. *PLoS Genet* 2007; **3**:2023-2036.
- 240 Illingworth R, Kerr A, Desousa D, *et al.* A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol* 2008; **6**:e22.
- 241 Mohn F, Weber M, Rebhan M, *et al.* Lineage-specific polycomb targets and *de novo* DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* 2008; **30**:755-766.
- 242 Irizarry RA, Ladd-Acosta C, Wen B, *et al.* The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 2009; **41**:178-186.
- 243 Wu H, Coskun V, Tao J, *et al.* Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* 2010; **329**:444-448.
- 244 Hansen RS, Wijmenga C, Luo P, *et al.* The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci USA* 1999; **96**:14412-14417.
- 245 Smeets DF, Moog U, Weemaes CM, *et al.* ICF syndrome: a new case and review of the literature. *Hum Genet* 1994; **94**:240-246.
- 246 Ehrlich M, Sanchez C, Shao C, *et al.* ICF, an immunodeficiency syndrome: DNA methyltransferase 3B involvement, chromosome anomalies, and gene dysregulation. *Autoimmunity* 2008; **41**:253-271.
- 247 Jeanpierre M, Turleau C, Aurias A, *et al.* An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Hum Mol Genet* 1993; **2**:731-735.
- 248 Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999; **23**:185-188.
- 249 Barreto G, Schafer A, Marhold A, *et al.* Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 2007; **445**:671-675.
- 250 Ma DK, Jang MH, Guo JU, *et al.* Neuronal activity-induced

- Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* 2009; **323**:1074-1077.
- 251 Gaudet F, Hodgson JG, Eden A, *et al.* Induction of tumors in mice by genomic hypomethylation. *Science* 2003; **300**:489-492.
- 252 Eden A, Gaudet F, Waghmare A, Jaenisch R. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science* 2003; **300**:455.
- 253 Chen T, Hevi S, Gay F, *et al.* Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. *Nat Genet* 2007; **39**:391-396.
- 254 Feng S, Cokus SJ, Zhang X, *et al.* Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci USA* 2010; **107**:8689-8694.