RNA-directed DNA methylation: an epigenetic pathway of increasing complexity

Marjori A. Matzke[1] *and Rebecca A. Mosher*[2]

Abstract | RNA-directed DNA methylation (RdDM) is the major small RNA-mediated epigenetic pathway in plants. RdDM requires a specialized transcriptional machinery that comprises two plant-specific RNA polymerases — Pol IV and Pol V — and a growing number of accessory proteins, the functions of which in the RdDM mechanism are only partially understood. Recent work has revealed variations in the canonical RdDM pathway and identified factors that recruit Pol IV and Pol V to specific target sequences. RdDM, which transcriptionally represses a subset of transposons and genes, is implicated in pathogen defence, stress responses and reproduction, as well as in interallelic and intercellular communication.

Dicer
(DCR). A ribonuclease III enzyme that cleaves double-stranded RNA precursors into small RNAs of 20–30 nucleotides. In plants, homologues of Dicer are referred to as DICER-LIKE (DCL). Of the four DCL enzymes in *Arabidopsis thaliana*, DCL3 produces 24-nucleotide small interfering RNAs (siRNAs) that act in the canonical RNA-directed DNA methylation pathway.

[1]*Institute of Plant and Microbial Biology, Academia Sinica, Taipei 115, Taiwan.*
[2]*School of Plant Sciences, University of Arizona, Tucson, Arizona 85721–0036, USA.*
e-mails: marjorimatzke@gate.sinica.edu.tw; rmosher@email.arizona.edu
doi:10.1038/nrg3683
Published online 8 May 2014

RNA interference (RNAi) is an umbrella term that describes gene silencing phenomena triggered by small RNAs[1]. In the cytoplasm, small RNAs induce post-transcriptional gene silencing (PTGS) by targeting complementary mRNAs for degradation or translational repression. In the nucleus, small RNAs elicit transcriptional gene silencing (TGS) by directing repressive epigenetic modifications, such as DNA cytosine methylation and histone methylation, to homologous regions of the genome. Small RNA-mediated epigenetic modifications are observed in plants, fungi and metazoans[1]. Epigenetically active small RNAs, which are typically 20–30 nucleotides in length, are generally classified into two groups: small interfering RNAs (siRNAs) and PIWI-interacting RNAs (piRNAs)[1].

siRNA pathways rely on core proteins of the RNAi machinery: Dicer (DCR) processes long double-stranded RNAs (dsRNAs) into siRNAs, and Argonaute (AGO) proteins are involved in siRNA effector functions, including small RNA-guided sequence-specific chromatin modifications. siRNA-mediated epigenetic modifications are described most thoroughly in plants and in the fission yeast *Schizosaccharomyces pombe*, and there is growing evidence for their occurrence in somatic cells of metazoans[1]. In contrast to siRNAs, piRNAs are not present in plants or fungi. In metazoans, piRNAs do not require DCR activity for their biogenesis and are incorporated into members of the germline-specific PIWI subfamily of AGO proteins[1] in *Drosophila melanogaster*, *Caenorhabditis elegans* and mice, in which they protect genome integrity by guiding repressive chromatin modifications that transcriptionally silence transposons[1,2].

In plants, the major siRNA-mediated epigenetic pathway is RNA-directed DNA methylation (RdDM), which is the subject of this Review. Initially detected in plants that were infected with RNA pathogens[3,4], RdDM was later shown to require siRNAs[5,6] and core RNAi proteins[1,7]. RdDM in plants is unique among small RNA-mediated chromatin modifications because it depends on a specialized transcriptional machinery that is centred around two plant-specific RNA polymerase II (Pol II)-related enzymes called Pol IV and Pol V[8] (BOX 1). In brief, the canonical view of RdDM involves the following steps. Transcripts from Pol IV are copied into long dsRNAs, processed by DICER-LIKE 3 (DCL3) into siRNAs and exported to the cytoplasm. Following loading of one strand of these siRNAs onto AGO4, they are re-imported to the nucleus, where the siRNA guides the targeting of nascent scaffold transcripts from Pol V by sequence complementarity. Ultimately, this targeting recruits DNA methyltransferase activity to mediate *de novo* methylation of cytosines in all classes of sequence contexts (that is, CG, CHG and CHH, where H represents A, T or G). This results in transcriptional silencing at the genomic loci that are transcribed by Pol V, particularly transposons and other repetitive DNA. This mechanism contrasts with TGS in *S. pombe*[1], in which Pol II synthesizes a nascent transcript that serves as both the precursor of

Box 1 | **RNA polymerases IV and V**

Eukaryotes have three multisubunit, nuclear DNA-dependent RNA polymerases: RNA polymerase I (Pol I) transcribes large ribosomal RNAs, Pol II transcribes mRNA precursors, and Pol III transcribes tRNA and 5S rRNA. Plants have two additional RNA polymerases known as Pol IV and Pol V, both of which evolved from Pol II and are specialized for RNA-directed DNA methylation (RdDM)[129]. Pol II, Pol IV and Pol V each have 12 subunits, many of which are shared by the 3 polymerases, but each also has specialized subunits[129,130]. Subunits are named nuclear RNA polymerase B (NRPB) for Pol II subunits, NRPD for Pol IV subunits and NRPE for Pol V subunits.

The largest subunits in Pol IV and Pol V are NRPD1 and NRPE1, respectively, and they bind to a shared subunit NRPD2/NRPE2 to form the catalytic cores. NRPD1 and NRPE1 differ from NRPB1 (which is the largest subunit of Pol II) through numerous substitutions or deletions of conserved amino acids in the catalytic centre[16,130,131] and in their carboxy-terminal domains (CTDs)[8], which probably contribute to their specialized functions in RdDM. Whereas the CTD of NRPB1 comprises multiple copies of a heptapeptide repeat, these repeats are absent in the CTDs of NRPD1 and NRPE1, which contain a motif found in a group of proteins called defective chloroplasts and leaves (DeCL). Authentic DeCL proteins are involved in processing rRNA[132], but the function of the DeCL motif in NRPD1 and NRPE1 remains unknown[8]. The extended CTD of NRPE1 also contains Trp-Gly or Gly-Trp repeats, which form an Argonaute (AGO) hook region that can bind to AGO4 (REF. 133), thus contributing to the specific role of Pol V in siRNA-directed *de novo* methylation (FIG. 1).

In Pol II, the NRPB1 and NRPB2 subunits combine with NRPB5 and NRPB9A or NRPB9B to create the 'jaw' region that grips DNA during transcription[5,134]. In addition to differences between NRPE1 and NRPB1, unique contributions of NRPE5 (REFS 5,38,130,135) and NRPE9B[5,134] to Pol V function imply that it may be adapted for transcribing templates with specific structural features or chromatin modifications[131]. *In vitro*, Pol IV and Pol V can carry out RNA-primed transcription of DNA and transcribe from bipartite RNA–DNA templates; in addition, Pol IV can transcribe bipartite RNA–RNA templates[16]. Until the *in vivo* templates of Pol IV and Pol V are known[131], deviations from their traditional activities, such as acting as an unconventional endonuclease or exonuclease[131], and the possibility of extrachromosomal templates should be considered. Similarly to Pol II, Pol IV and Pol V may require factors that assist entry into the nucleus from the cytoplasm, in which the subunits are synthesized and assembled. The identification of homologues of yeast IWR1 (interacts with RNA polymerase II, which facilitates nuclear import of Pol II[136]) in screens for RdDM-defective mutants[137,138] indicates that Pol IV and Pol V might be imported into the nucleus in a similar way to that of Pol II.

siRNAs and a scaffold RNA that interacts with siRNAs, and recruits the silencing effector complex[1]. Rather than inducing DNA methylation, this complex in *S. pombe* induces TGS and heterochromatin formation through histone H3 lysine 9 methylation (H3K9me) at pericentromeric repeats[1], as well as at some developmentally regulated genes and retrotransposons[9].

In addition to providing one of the first examples of a small RNA-mediated epigenetic modification[1], the RdDM pathway represents an impressive extension to the transcriptional capabilities of a eukaryotic organism. It remains incompletely understood why plants require two additional RNA polymerases and a host of auxiliary factors to carry out RdDM, and how these transcriptional complexes are harnessed for this particular epigenetic pathway. Therefore, the mechanism, biological roles and evolutionary importance of RdDM are topics of active investigation. The pace of discovery in RdDM research has been accelerated by technological advances, particularly next-generation sequencing, which has revealed genome-wide patterns of DNA methylation at single-nucleotide resolution in both wild-type and mutant plants[10], as well as natural epigenomic variation in different plant strains[11].

In this Review, we focus on RdDM in *Arabidopsis thaliana*, which has been studied most intensively, and include contributions from other plants where appropriate. We describe components of the pathway and their contributions to the mechanism of RdDM, as well as factors that promote targeting of Pol IV and Pol V to specific genomic regions. In addition, we discuss emerging variations in the canonical RdDM pathway and the growing list of biological processes that involve RdDM.

Canonical RdDM pathway mechanisms

Components of the RdDM pathway, including particular subunits of Pol IV and Pol V (BOX 1), have been discovered using genetic and biochemical approaches (TABLE 1). Current models assign these factors to Pol IV-dependent siRNA biogenesis, Pol V-mediated *de novo* DNA methylation or chromatin alterations that involve histone modifications, nucleosome positioning and higher-order chromatin conformations (FIG. 1).

Pol IV-dependent siRNA biogenesis. Analyses of Pol IV-defective mutants have shown that this polymerase is responsible for producing the precursor of >90% of 24-nucleotide siRNAs, which guide methylation in the canonical RdDM pathway[12,13]. Recruitment of Pol IV to target sequences, which are primarily transposons and other repeats, is not fully understood. Pol IV is recruited to a subset of its genomic targets by the Pol IV-interacting protein SAWADEE HOMEODOMAIN HOMOLOGUE 1 (SHH1), which binds to H3K9me and unmethylated H3K4 through its unique tandem Tudor-like fold[14,15]. Although Pol IV transcripts have not yet been observed *in vivo*, Pol IV is assumed to transcribe single-stranded RNAs (ssRNAs) at its target loci. The ssRNA is copied by the RNA-dependent RNA polymerase RDR2, which physically associates with Pol IV[16,17], to produce dsRNAs. The chromatin remodeller CLASSY 1 (CLSY1)[17,18] participates at some point in these steps, presumably to ease the passage of Pol IV along the genomic locus. DCL3 processes dsRNAs to 24-nucleotide siRNAs, which are stabilized by methylation at their 3′-OH groups by HUA ENHANCER 1 (HEN1)[19] and loaded onto AGO4 (FIG. 1). Other members of the AGO4 clade include AGO6 (which is partially redundant with AGO4 (REFS 20,21)) and AGO9 (which is expressed specifically in reproductive tissue[22]).

Pol V-mediated* de novo *methylation. Pol V transcripts, which are either triphosphorylated or capped at the 5′ ends and lack poly(A) tails, are thought to provide scaffold RNAs that interact with siRNAs and that recruit other silencing factors[23]. Similarly to Pol IV, the chromatin features responsible for recruiting Pol V to its target sequences remain incompletely understood, and no consensus DNA sequence has emerged. However, some insights into binding site preferences for Pol V-mediated RdDM are being revealed from genome-wide studies. Experiments using chromatin immunoprecipitation followed by sequencing (ChIP–seq) localized most Pol V at transposons and repeats that are associated with 24-nucleotide siRNAs and with cytosine methylation,

Argonaute
(AGO). A family of effector proteins of RNA interference that bind to small interfering RNAs (siRNAs) through their PAZ (PIWI–AGO–ZWILLE) and MID (middle) domains and, in some cases, slice RNA through their PIWI domain. Of the ten AGOs in *Arabidopsis thaliana*, AGO4, AGO6 and AGO9 act in canonical RNA-directed DNA methylation and/or transcriptional gene silencing.

Transposons
Invasive genetic elements that move within a genome and that are sometimes associated with replicative movement which produces many copies. Transposons include retrotransposons, DNA transposons and helitrons.

Table 1 | **Components of the canonical RdDM pathway**

Proteins	AGI number of gene	Description	Refs
NUCLEAR RNA POLYMERASE D1 (NRPD1)	AT1G63020	The unique, largest subunit of Pol IV	130,150–152
NRPE1	AT2G40030	The unique, largest subunit of Pol V	130,152,153
NRPD2/NRPE2*	AT3G23780	The shared second largest subunit of Pol IV and Pol V	130,151–153
NRPD4/NRPE4*	AT4G15950	The shared fourth largest subunit of Pol IV and Pol V	130,154
NRPE5	AT3G57080	A special fifth largest subunit of Pol V	5,38,130,135
NRPE9B	AT4G16265	A special ninth largest subunit required for Pol V activity	5,134
NRPB1	AT4G35800	The largest subunit of Pol II	71
RNA-DEPENDENT RNA POLYMERASE 2 (RDR2)	AT4G11130	An RNA-dependent RNA polymerase	7,16
DICER-LIKE 3 (DCL3)	AT3G43920	A Dicer endonuclease that produces 24-nucleotide siRNAs	7,155,156
HUA ENHANCER 1 (HEN1)	AT4G20910	A small RNA methyltransferase	19
ARGONAUTE 4 (AGO4)	AT2G27040	An Argonaute protein	21
AGO6	AT2G32940	An Argonaute protein	5,21
AGO9	AT5G21150	An Argonaute protein	21,22
CLASSY 1 (CLSY1)	AT3G42670	A putative SWI/SNF chromatin remodeller; involved in the Pol IV pathway	17,18
DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1)	AT2G16390	A putative SWI/SNF chromatin remodeller; part of the DDR complex; involved in the Pol V pathway	156,157
DEFECTIVE IN MERISTEM SILENCING 3 (DMS3)	AT3G49250	A SMC solo hinge protein; part of the DDR complex; involved in the Pol V pathway	158,159
RNA-DIRECTED DNA METHYLATION 1 (RDM1)	AT3G22680	An AGO4- and Pol II-interacting protein; part of the DDR complex; involved in the Pol V pathway	35,37
KOW DOMAIN-CONTAINING TRANSCRIPTION FACTOR 1 (KTF1)	AT5G04290	Contains the AGO hook motif; involved in Pol V transcription	38–40
INVOLVED IN DE NOVO 2 (IDN2)	AT3G48670	A dsRNA-binding protein in the Pol V pathway	158
IDN2 PARALOGUE 1 (IDP1)	AT1G15910	Forms a complex with IDN2	43–45
IDP2	AT4G00380	Forms a complex with IDN2	43–45
DMS4	AT2G30280	Associated with Pol V and Pol II	137,138
DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2)	AT5G14620	A *de novo* DNA methyltransferase	7,156,160
SUVH2	AT2G33290	An SRA domain protein that binds to methylated DNA and recruits Pol V	31–34
SUVH9	AT4G13460	An SRA domain protein that binds to methylated DNA and recruits Pol V	31–34
SUVR2	AT5G43990	A putative histone methyltransferase	10
MICRORCHIDIA 1 (MORC1)	AT4G36290	A MORC-type ATPase	42,55,56
MORC6	AT1G19100	A MORC-type ATPase	56
SAWADEE HOMEODOMAIN HOMOLOGUE 1 (SHH1)	AT1G15215	Binds to methylated H3K9 and recruits Pol IV	14,15
HISTONE DEACETYLASE 6 (HDA6)	AT5G63110	A histone deacetylase	49,50
JUMONJI 14 (JMJ14)	AT4G20400	A histone demethylase	51,52
LYSINE-SPECIFIC HISTONE DEMETHYLASE 1 (LDL1)	AT1G62830	A histone demethylase	53
LDL2	AT3G13682	A histone demethylase	53
UBIQUITIN-SPECIFIC PROTEASE 26 (UBP26)	AT3G49600	A histone H2B deubiquitinase	54

de novo methylation
Methylation of a previously unmodified DNA sequence. Small interfering RNAs (siRNAs) in the RNA-directed DNA methylation pathway are well-known triggers of sequence-specific *de novo* methylation of cytosines in all sequence contexts.

Silencing effector complex
A multiprotein complex that elicits RNA interference and related small RNA-mediated gene silencing pathways. It is composed of an Argonaute protein (which binds to the small RNA guide) and, in the case of RNA-directed DNA methylation, cofactors that aid in directing DNA methylation to the small RNA-targeted region of the genome.

Pericentromeric
Pertaining to the region surrounding the centromere, which is the chromosomal region where two sister chromatids are joined.

RNA-dependent RNA polymerase
(RDR). A cellular enzyme that copies single-stranded RNAs to produce double-stranded RNA precursors, which are processed by Dicer-like proteins to generate small interfering RNAs (siRNAs). Of the six RDRs in *Arabidopsis thaliana*, RDR2 is associated with the canonical RdDM pathway.

Table 1 (cont.) | **Components of the canonical RdDM pathway**

Proteins	AGI number of gene	Description	Refs
Additional factors			
NEEDED FOR RDR2-INDEPENDENT DNA METHYLATION (NERD)	AT2G16485	Involved in non-canonical RdDM	70
CHROMOMETHYLASE 2 (CMT2)	AT4G19020	A DNA methyltransferase specific to CHH‡	27,30
CMT3	AT1G69770	A DNA methyltransferase specific to CHG‡	48
METHYLTRANSFERASE 1 (MET1)	AT5G49160	A DNA methyltransferase specific to CG	49,50
SUVH4	AT5G13960	A H3K9 methyltransferase	48
DECREASED DNA METHYLATION 1 (DDM1)	AT5G66750	A chromatin remodeller	27

AGI; *Arabidopsis* Genome Initiative; DDR complex, DRD1–DMS3–RDM1 complex; dsRNA, double-stranded RNA; H3K9, histone H3 lysine 9; Pol, RNA polymerase; RdDM, RNA-directed DNA methylation; siRNA, small interfering RNA; SMC, structural maintenance of chromosomes; SRA, SET and RING-associated. *For shared subunits, NRPD refers to the protein when it is a component of Pol IV, whereas NRPE refers to the same protein when it is a component of Pol V. ‡H represents A, T or G.

thus indicating that Pol V is mediating RdDM at these sites[24,25]. However, ~25% of genomic sites occupied by Pol V lack these features, which suggests that Pol V occupancy alone is not sufficient for RdDM; these unmethylated sites are biased towards genes, some of which contain repetitive sequences in their coding region[25].

Overall, Pol V-mediated RdDM has been found to act at a wide range of locations throughout the genome but with preferences for euchromatic regions, particularly at small, 'young' (that is, recently acquired) intergenic transposons and at genes that contain transposons or other repeats in their promoters, introns or coding regions[24–28]. The location of many RdDM targets in euchromatin is consistent with the proposed evolution of both Pol IV and Pol V from Pol II (BOX 1), which transcribes genes primarily in euchromatic contexts. Furthermore, Pol V enrichment at gene promoters may reflect a retention of Pol II binding preferences for regions upstream of genes[24].

RdDM seems to be excluded to some extent from pericentromeric heterochromatin[27,29], which is enriched in larger transposons. Instead, the modifications at pericentromeric heterochromatin (that is, H3K9me and DNA methylation) mostly occur in an siRNA-independent manner and rely on the chromatin remodeller DECREASED DNA METHYLATION 1 (DDM1); the DNA (cytosine-5)-methyltransferase 1 (DNMT1) class enzyme METHYLTRANSFERASE 1 (MET1; also known as DMT1); and the plant-specific DNA methyltransferases CHROMOMETHYLASE 2 (CMT2) and CMT3 (REFS 27,30).

Recruitment of Pol V to some target sequences is aided by SUVH2, SUVH9 and SUVR2, which are members of the SU(VAR)3-9 histone methyltransferase family. SUVH2 and SUVH9 are unable to catalyse histone methylation, but they can bind to methylated DNA through their SRA (SET and RING-associated) domain and thus contribute to a platform for interaction of Pol V with chromatin that contains some pre-existing methylation[31–34]. Pol V transcription and association with chromatin are facilitated by the DDR complex[24,35]. This complex comprises the CLSY1-related putative chromatin remodeller DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1)[5,23]; DEFECTIVE IN MERISTEM SILENCING 3 (DMS3), which is a structural maintenance of chromosomes solo hinge protein[5,36]; and RNA-DIRECTED DNA METHYLATION 1 (RDM1)[37] (FIG. 1), which is a small plant-specific protein that may have multiple roles in the RdDM pathway (see below).

Pol V recruits AGO4 through the AGO hook region in the carboxy-terminal domain of its unique largest subunit NUCLEAR RNA POLYMERASE E1 (NRPE1) (BOX 1), which interacts with KOW DOMAIN-CONTAINING TRANSCRIPTION FACTOR 1 (KTF1; also known as SPT5L)[38] — a putative transcription elongation factor that also contains an AGO hook motif[39,40]. During Pol V-mediated transcription, the AGO4-bound siRNA is believed to base-pair with the nascent Pol V transcript and recruit DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) — a member of the DNMT3 family of DNA methyltransferases — to catalyse *de novo* methylation at the homologous genomic sites. A key part in the recruitment of DRM2 may be played by RDM1, which is the only protein that has been reported to interact with both AGO4 and DRM2, and to bind to methylated single-stranded DNA (ssDNA)[37]. The role of RDM1 in the RdDM mechanism might thus extend beyond its participation in the DDR complex (FIG. 1).

INVOLVED IN DE NOVO 2 (IDN2) is a dsRNA-binding protein that is related to SUPPRESSOR OF GENE SILENCING 3 (SGS3) and is involved in PTGS[41–43]. IDN2 forms a complex with two partially redundant paralogues, IDP1 and IDP2 (REFS 41,43,44). The IDN2–IDP complex may stabilize base-pairing between siRNAs and Pol V scaffold RNAs[41,44,45], and facilitate RdDM by altering nucleosome positioning through interactions with the SWI/SNF chromatin remodelling complex[46] (FIG. 1).

Interplay with chromatin features

Histone modifications. Around 70% of RdDM targets are modified by H3K9me[47], which can act in a feedback loop with DNA methylation to reinforce TGS. For

Structural maintenance of chromosomes
(SMC). A large family of ATPases that can manipulate chromosome-sized molecules and that contribute to higher-order chromatin structure and dynamics.

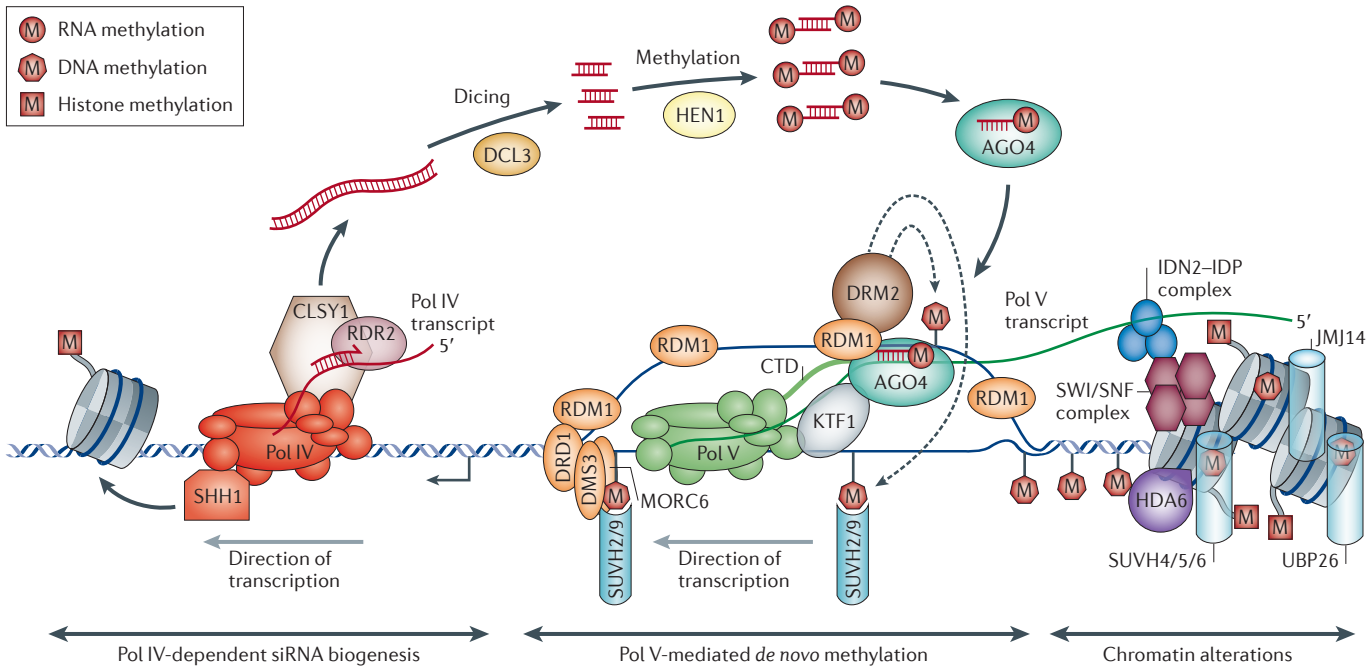


Figure 1 | **Canonical RdDM pathway.** A transcription fork model for RNA-directed DNA methylation (RdDM) is shown[148]. In RNA polymerase IV (Pol IV)-dependent small interfering RNA (siRNA) biogenesis (left panel), Pol IV transcribes a single-stranded RNA (ssRNA) that is copied into a double-stranded RNA (dsRNA) by RNA-DEPENDENT RNA POLYMERASE 2 (RDR2) with the assistance of the chromatin remodeller CLASSY 1 (CLSY1). The dsRNA is processed by DICER-LIKE 3 (DCL3) into 24-nucleotide siRNAs that are methylated at their 3′ ends by HUA ENHANCER 1 (HEN1) and incorporated into ARGONAUTE 4 (AGO4). SAWADEE HOMEODOMAIN HOMOLOGUE 1 (SHH1), which binds to histone H3 methylated at lysine 9 (H3K9me), interacts with Pol IV and recruits it to some target loci. In Pol V-mediated *de novo* methylation (middle panel), Pol V transcribes a scaffold RNA that base-pairs with AGO4-bound siRNAs. AGO4 is recruited through interactions with the AGO hook regions in the carboxy-terminal domain of the largest subunit of Pol V and with KOW DOMAIN-CONTAINING TRANSCRIPTION FACTOR 1 (KTF1). RNA-DIRECTED DNA METHYLATION 1 (RDM1) links AGO4 and DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), which catalyses *de novo* methylation of DNA. Pol V transcription may be enabled by the duplex unwinding activity of the chromatin remodeller DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1), whereas the single-stranded DNA-binding activity of RDM1 and the putative cohesin-like roles of DEFECTIVE IN MERISTEM SILENCING 3 (DMS3) and MICRORCHIDIA 6 (MORC6) may help to generate and stabilize the unwound state. Pol V recruitment is potentially aided by SUVH2 or SUVH9, both of which bind to methylated DNA. Nucleosome positioning (right panel) is adjusted by the SWI/SNF complex, which interacts with the IDN2 (INVOLVED IN DE NOVO 2)–IDP (IDN2 PARALOGUE) complex that binds to Pol V scaffold RNAs. Deposition of repressive histone modifications — such as H3K9me by SUVH4, SUVH5 and SUVH6 — is facilitated following removal of active marks by HISTONE DEACETYLASE 6 (HDA6), JUMONJI 14 (JMJ14) and UBIQUITIN-SPECIFIC PROTEASE 26 (UBP26). Higher-order chromatin conformations that reinforce the silent state are established through the ATPase activities of MORC1 and MORC6 (not shown). Adapted with permission from REF. 148, Cold Spring Harbor Laboratory Press.

example, the SUVH histone methyltransferases SUVH4 (also known as KRYPTONITE), SUVH5 and SUVH6 catalyse H3K9me, which is closely associated with non-CG methylation catalysed by CMT3 (REF. 48). The SRA domain of the SUVH enzymes binds to methylated DNA, whereas the chromodomain of CMT3 binds to H3K9me, thus creating a self-reinforcing loop that perpetuates both epigenetic modifications[31,48].

Some RdDM targets require histone-modifying enzymes that remove active marks — such as acetylation, H3K4me and H2B ubiquitylation — to maintain DNA methylation and promote methylation of H3K9 (FIG. 1). HISTONE DEACETYLASE 6 (HDA6) acts in conjunction with MET1 to maintain CG methylation and to promote H3K9me by deacetylating histones[49,50]. The Jumonji C-type histone demethylase JUMONJI 14 (JMJ14)[51,52] and two related histone demethylases, LYSINE-SPECIFIC HISTONE DEMETHYLASE 1 (LDL1) and LDL2 (REF. 53), cooperate to maintain DRM2-mediated DNA methylation, presumably by removing H3K4me at target loci. Deubiquitylation of H2B by UBIQUITIN-SPECIFIC PROTEASE 26 (UBP26) is needed for non-CG methylation and H3K9me at certain target sites[54] (FIG. 1).

Higher-order chromatin structure. Information is beginning to emerge on the influence of RdDM on higher-order chromatin structure. Two microrchidia (MORC)-type ATPases are important for gene silencing in the RdDM pathway at a subset of genes and transposons[42,55,56]. In *A. thaliana*, MORC6 has been proposed to provide an ATPase function for DMS3 to form a functional analogue

of SMCHD1, which is a cohesin-like protein required for X chromosome inactivation in mice[42] (FIG. 1). Mutant plants deficient in MORC1 and MORC6 show only modest changes in DNA methylation and repressive histone modifications despite loss of TGS at target loci[42,55,56], which suggests downstream roles in higher-order chromatin condensation. Consistent with this proposal, one study demonstrated that these mutants show decondensation of pericentromeric heterochromatin[56].

DNA methylation maintenance and dynamics

Maintenance during replication. Symmetrical methylation at CG and CHG can be maintained independently of siRNAs during subsequent rounds of DNA replication by MET1 and CMT3, respectively, both of which act on hemimethylated DNA to copy the modifications to the other strand. For the most part symmetrical methylation is maintained through meiosis[57,58], even at loci that lack siRNAs[59,60]. DRM2-catalysed asymmetrical CHH methylation — which, unlike symmetrical methylation, is not found in both daughter DNA strands — cannot be maintained in the absence of the siRNA trigger and requires re-establishment following each DNA replication cycle[61]. Recent findings revealed additional complexity of maintenance methylation, in which multiple methyltransferases contribute to preservation of cytosine methylation in all three sequence contexts[10,27]. Moreover, although CHH methylation was initially thought to result primarily from DRM2 activity, CMT2 maintains a substantial amount of CHH methylation in heterochromatic regions in a manner that relies on H3K9me and that is independent of siRNAs[27,30].

Dynamic demethylation. Although RdDM contributes to transcriptional repression of transposons[24,26,27], it has a smaller role overall than the DDM1 pathway[9,15]. Instead of stably silencing transposons, RdDM may establish a more dynamic modification that is removable from euchromatic targets through passive or active demethylation[24,28,62,63]. Supporting this suggestion, REPRESSOR OF SILENCING 1 (ROS1) — a DNA glycosylase involved in active demethylation through a base excision repair pathway — preferentially counteracts RdDM-induced methylation[63]. An antagonistic relationship between active demethylation and RdDM is reinforced by the finding of many known RdDM components in genetic screens for *ros1* suppressors[63] and by the reduced expression of *ROS1* in RdDM mutants[62–64]. Additional work is needed to clarify the link between RdDM and ROS1 activity, and to assess the possible involvement of short or long non-coding RNAs in targeting active demethylation to specific loci[63].

Emerging non-canonical RdDM mechanisms

Several RdDM mechanisms have recently been reported to deviate from the canonical pathway that involves 24-nucleotide siRNAs, Pol IV, RDR2, DCL3, Pol V and AGO4. These mechanisms partly incorporate components that are typically associated with PTGS (such as Pol II), characteristic types and sizes of small RNAs, and specific silencing factors.

miRNA and tasiRNA-induced DNA methylation. Plant microRNAs (miRNAs) and *trans*-acting siRNAs (tasiRNAs, which are are plant specific) function in PTGS by guiding cleavage or translational repression of complementary mRNAs. In some cases, these small RNAs can also have roles in the nucleus by acting in RdDM.

Normally, 21-nucleotide miRNAs are produced by DCL1 cleavage of Pol II-transcribed hairpin RNA precursors and loaded onto AGO1 to interact with target mRNAs to elicit PTGS. However, in rice, the hairpin RNA can also be processed by DCL3 to produce a longer class of 24-nucleotide miRNA (called long miRNA (lmiRNA)) that is loaded onto AGO4. Some of the AGO4-bound lmiRNAs can guide methylation of target genes, which shows a role for lmiRNAs in chromatin-based silencing[65]. A type of miRNA-mediated transcriptional regulation and DNA methylation that relies on the ratio between miRNA and its mRNA target has been reported for the moss *Physcomitrella patens*[66] but not yet for other plants.

Biogenesis of tasiRNAs in land plants is a distinct process that is initiated by miRNA-guided cleavage of various larger Pol II-generated *TAS* RNA precursors, which are copied into dsRNAs by RDR6 and stabilized by SGS3. These dsRNAs are normally processed by DCL4 to produce 21-nucleotide siRNAs that are loaded onto AGO1 to interact with target mRNAs and induce PTGS. However, the dsRNAs can also be processed by DCL1 to produce 21-nucleotide tasiRNAs that are loaded onto AGO4 or AGO6, which act in Pol V-mediated RdDM of the corresponding *TAS* loci[67]. These results demonstrate a role for DCL1, which usually processes miRNA precursors, in producing 21-nucleotide tasiRNAs that can induce RdDM.

RDR6-dependent RdDM. A recently identified RDR6-dependent pathway of RdDM provides a link between PTGS of transposon transcripts and *de novo* methylation of transposon DNA. Young transposons are initially transcribed by Pol II to produce mRNAs for transposon-encoded proteins. Some of these Pol II transcripts can be copied by RDR6 to produce dsRNAs, which are processed by DCL2 and DCL4 into 21–22-nucleotide siRNAs, resulting in AGO1-mediated PTGS of transposon mRNAs (FIG. 2). However, these dsRNAs can also initiate low levels of *de novo* DNA methylation in a manner that is dependent on AGO2, Pol V scaffold transcripts and DRM2. This initial methylation activates the canonical RdDM pathway, which leads to the biogenesis of 24-nucleotide siRNAs (mediated by Pol IV, RDR2 and DCL3) that enhance methylation and reinforce TGS[68] (FIG. 2). A variation of this pathway was observed in a system that reconstructed *de novo* silencing of a newly integrated retrotransposon. A transition from PTGS to TGS occurred when high levels of Pol II- and RDR6-dependent dsRNAs saturated DCL2 and DCL4 enzymes and became available for processing by DCL3, which produces 24-nucleotide siRNAs that triggered canonical RdDM and TGS of the transposon[69].

Symmetrical methylation
Cytosine methylation at CG:GC and CHG:GHC nucleotide groups in both DNA strands. As a result of complementary base pairing, CG and CHG are base-paired to GC and GHC, respectively, on the opposite DNA strand and hence considered symmetrical.

Maintenance methylation
The preservation of pre-existing methylation at symmetrical CG and CHG sites after DNA replication by the DNA methyltransferases MET1 and CMT3, which recognize hemimethylated substrates (that is, those methylated on one strand but not the other).

MicroRNAs
(miRNAs). Small non-coding RNAs (~21–23 nucleotides) that silence gene expression by mRNA degradation or translational repression through complementarity with the target transcripts.

Trans-acting siRNAs
(tasiRNAs). A class of small interfering RNAs (siRNAs) that silences gene expression in land plants by targeting complementary mRNAs for cleavage. Their biogenesis depends on microRNA (miRNA)-mediated cleavage of longer *TAS* RNA precursors that are further acted on by RNA-DEPENDENT RNA POLYMERASE 6 (RDR6) and DICER-LIKE 4 (DCL4). The miRNA-triggered initiation followed by DCL4 cleavage results in a phased pattern of accumulation, in which small RNAs are in an exact head-to-tail arrangement. tasiRNAs are one category of 'phased' siRNA (phasiRNA).

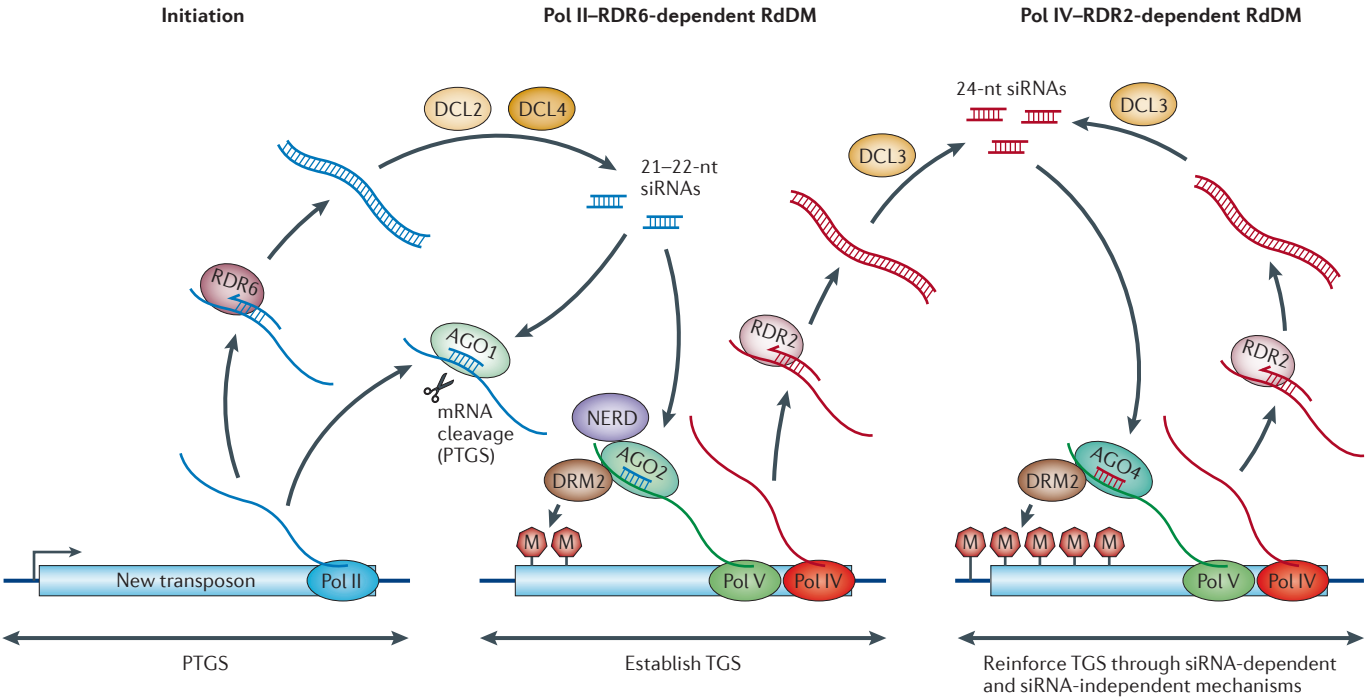


Figure 2 | **Non-canonical Pol II–RDR6-dependent RdDM pathway.** This pathway provides a means to establish RNA-directed DNA methylation (RdDM) and eventually ensure stable transcriptional gene silencing (TGS) of a newly acquired transposon that is originally a target of post-transcriptional gene silencing (PTGS). In PTGS (left panel), a newly inserted transposon is initially active and transcribed by RNA polymerase II (Pol II). Some of the transcripts are copied by RNA-DEPENDENT RNA POLYMERASE 6 (RDR6) to produce double-stranded RNAs (dsRNAs), which are processed by DICER-LIKE 2 (DCL2) and DCL4 into 21–22-nucleotide (nt) small interfering RNAs (siRNAs). These siRNAs are loaded onto ARGONAUTE 1 (AGO1) and guide cleavage of transposon transcripts in a typical PTGS pathway. In a deviation from the canonical RdDM pathway (middle panel), some of the 21–22-nt siRNAs can also trigger low levels of DNA methylation in a manner that is dependent on DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), Pol V and AGO2, which interacts with NEEDED FOR RDR2-INDEPENDENT DNA METHYLATION (NERD) through its AGO hook motif. The sparsely methylated DNA recruits Pol IV, which initiates the canonical RdDM pathway by transcribing a single-stranded RNA (ssRNA). The ssRNA is copied by RDR2 into a dsRNA that is processed by DCL3 into 24-nt siRNAs. Following incorporation into AGO4 (right panel), the 24-nt siRNAs base-pair with Pol V scaffold transcripts, which results in DRM2 recruitment and dense methylation. siRNAs are continuously produced from the methylated template by Pol IV pathway components, which reinforces TGS that can be maintained in an siRNA-independent manner by METHYLTRANSFERASE 1 (MET1), CHROMOMETHYLASE 3 (CMT3) and DECREASED DNA METHYLATION 1 (DDM1)[149] (not shown).

A computational screen for AGO hook-containing proteins identified NEEDED FOR RDR2-INDEPENDENT DNA METHYLATION (NERD), which is a large plant-specific protein with plant homeodomain (PHD) and zinc-finger domains[70]. NERD-dependent RdDM requires both Pol IV and Pol V but also PTGS components, including RDR6, SGS3, SILENCING DEFECTIVE 3 (SDE3) and SDE5. NERD interacts with histone H3 and with AGO2, which incorporates 21-nucleotide siRNAs that can induce RdDM[70]. These features are similar to those reported for the RDR6-dependent mechanisms of transposon silencing described above (FIG. 2). Given that the NERD-dependent mechanism has also been suggested as a means to initiate *de novo* methylation of newly acquired transposons that initially undergo PTGS[70], the RDR6- and NERD-dependent mechanisms are likely to overlap considerably.

Pol II as a source of scaffold transcripts. At some low copy-number intergenic loci that do not produce siRNAs, Pol II synthesizes scaffold transcripts that recruit AGO4-bound siRNAs to elicit RdDM and TGS[71]. Pol II transcription or transcripts can also recruit Pol IV and Pol V to other low copy-number loci to carry out their specific functions in RdDM[71,72]. The characteristics that attract Pol II to some intergenic loci and the requirements for Pol II interaction with Pol IV and Pol V are unknown[71,72]. These results suggest an intricate collaboration between the three polymerases, which remains to be investigated in more detail. Pol II can also mediate DNA methylation independently of Pol IV and Pol V[10].

Biological roles for RdDM

PTGS by miRNAs and tasiRNAs is crucial for plant development and physiology; however, despite the loss of siRNA production from thousands of genomic sites,

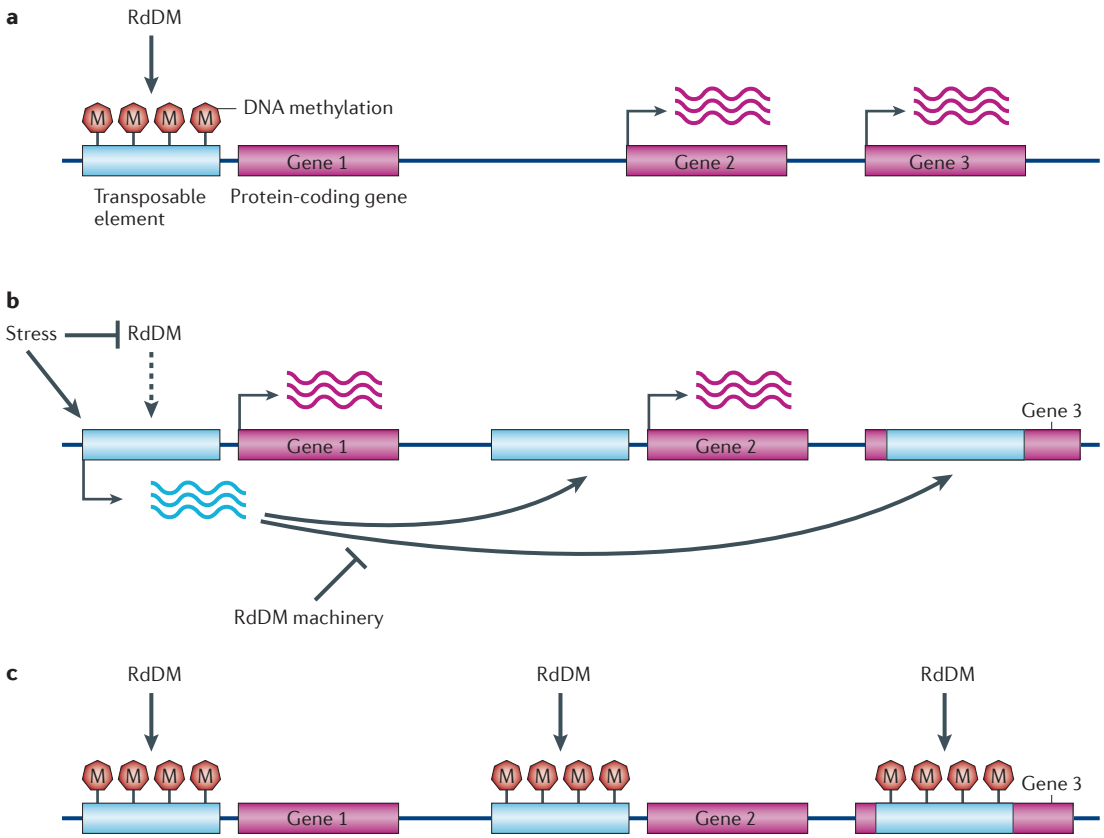


Figure 3 | **Release of RdDM during stress.** **a** | A schematic locus is shown, in which a silenced transposon leads to the silencing of a nearby protein-coding gene (Gene 1). By contrast, the activity of other protein-coding genes (Gene 2 and Gene 3) are not under control of the transposon. **b** | Transposons are activated during biotic and abiotic stress responses through a combination of loss of RNA-directed DNA methylation (RdDM) and transcriptional responses to stress[74,79,81,83,85]. The nearby Gene 1 is also activated owing to loss of methylation in promoter regions[81]. Reactivated transposons can integrate into new genomic loci, although the RdDM machinery inhibits the reinsertion of some elements through an unknown mechanism[74]. **c** | New transposon insertions can establish stress-responsive transcription at additional protein-coding genes (Gene 2) or might permanently disrupt gene function (Gene 3).

RdDM-defective *A. thaliana* have few obvious phenotypes. Recently, our understanding of the biological function of RdDM has increased through careful observation of reproduction, and of the effects of biotic and abiotic stresses on the genome. New functions in cellular communication have also been uncovered.

Transposon control. The most abundant targets of Pol IV and RdDM are repetitive elements, including all classes of transposons[12,13], particularly smaller and younger ones[24,26,27]. Proliferation of an introduced retrotransposon in *A. thaliana* induces RdDM after the element has inserted into at least four genomic sites[73], and mobilization of an epigenetically reactivated endogenous element similarly attracts the RdDM machinery when copy numbers increase[69].

As mentioned above, RdDM has a smaller role overall in DNA methylation of transposons than the DDM1 pathway. In one study, >2,000 transposons were reactivated in a *ddm1* mutant, whereas ~40 were reactivated in an RdDM-defective mutant[27]. These results are consistent with previous data that indicate infrequent mobilization of transposons in RdDM mutants compared with *ddm1* and *met1* mutants[74]. Instead of being solely involved in creating constitutive heterochromatin that stably silences transposons, RdDM may establish a more dynamic modification that is removable from euchromatic targets through passive or active demethylation[28,62,63] (see above). The activation of only a few endogenous transposons when the RdDM machinery is compromised[27] is probably due to maintenance of CG and CHG methylation independently of RdDM. In some cases, transposons might also require an environmental stimulus for activation (FIG. 3). The retrotransposon *ONSEN* is transcriptionally activated during heat stress, and this activation is enhanced in mutants that lack RdDM[74]. Surprisingly, although *ONSEN* is transcriptionally active and the presence of extrachromosomal copies suggests reverse transcription, new genomic insertions by retrotransposition occurs only in mutants that lack either Pol IV or RDR2 (REF. 74) (FIG. 3). This observation implies that the small RNAs associated with RdDM have additional roles in genome defence beyond TGS.

Retrotransposition
The process of mobilizing a retrotransposon. It involves transcription, processing of the RNA, translation, reverse transcription of the transposon RNA and integration of the reverse-transcribed DNA into a new genomic location.

Pathogen defence. RDR and DCL proteins target RNA-encoded viruses to generate siRNAs, which then catalyse additional cleavage of viral RNAs to limit viral replication and spread[75]. DNA-encoded viruses also generate small RNAs and are subject to RdDM[75]. Unsurprisingly, both RNA- and DNA-based viruses encode suppressors of silencing that limit the silencing capability of the host plant[75–78]. These silencing suppressors also reduce RdDM activity at transposons and repetitive sequences in the host genome, which potentially affects genome stability and increases the adaptive potential of the host[76,78,79].

RdDM is also involved in the host response to bacterial pathogens. *A. thaliana* mutants that lack RdDM components have altered resistance to a range of bacterial pathogens[80–82]; however, Pol IV does not seem to be involved, which suggests that a non-canonical RdDM pathway is responsible[80]. Infection with the bacterium *Pseudomonas syringae* DC3000 or treatment with the bacterial elicitor flagellin triggers active hypomethylation and hypermethylation at specific genomic sites in the host genome through simultaneous downregulation of RdDM and active demethylation[81,83]. These methylation changes increase the levels of transposon transcripts and synthesis of 21-nucleotide siRNAs, and alter gene expression of some transposon-associated genes[81,83]. Pathogen-triggered changes in host methylation indicate that hosts and pathogens might alter RdDM efficiency to affect pathogen infection, although it is unclear whether such changes are a mechanism of pathogenesis or a part of the host defence response.

The pathogenic bacterium *Agrobacterium tumefaciens*, which inserts tumour-causing genes into plant host genomes, presents a unique link between pathogens and genome defence. During tumour formation the host genome undergoes global hypermethylation, which is potentially a defence mechanism driven by host RdDM, as *ago4* mutants and *cmt3–drm1–drm2* triple mutants of *A. thaliana* have enhanced tumour development[84].

Stress responses. There are numerous accounts of environmental stresses — such as extreme temperature, drought or ultraviolet radiation — triggering epigenetic changes, which in turn can affect transcription[81,85,86]. Genetic analyses directly implicate RdDM in some cases[84,86–88], whereas alteration of methylation patterns at transposons is circumstantial evidence for changes in RdDM in other cases[81,85]. Loss of transposon silencing in response to stress increases phenotypic diversity through transcriptional upregulation of neighbouring genes and though novel transposon insertions, which might either disrupt gene function or alter the transcription pattern of nearby genes[74] (FIG. 3). This increase in phenotypic diversity might increase adaptability to a changing environment.

Epialleles. Environmental cues can trigger epigenetic changes at particular loci to generate epialleles, which can be stably transmitted to the progeny for many generations[85,89–91]. Therefore, in cases in which epialleles have phenotypic consequences for the organism, they might represent a form of adaptive epigenetic inheritance known as Lamarckian inheritance, which could offer increased fitness to subsequent generations in similar stressful environments[82,92,93]. In unstressed *A. thaliana*, the rate of spontaneous gains or losses of DNA methylation is 1,000 times higher than the genetic mutation rate[57], and hypermethylated alleles are associated with siRNA production and TGS[58]. If epialleles are inherited and selected in a similar way to genetic mutations, then random RdDM-mediated epiallele formation could have a greater role in evolution than genetic variation.

Despite the adaptive potential of epialleles, plants must balance epigenetic variation with faithful transmission of established epigenetic states. Spontaneous changes in DNA methylation are less frequent at established sites of siRNA production, which indicates that RdDM stabilizes methylation patterns[57]. siRNAs also re-establish methylation patterns that have been lost[60]. Interestingly, this remethylation occurs in a stepwise manner in each generation[60], which implies that RdDM is especially active during reproduction. Indeed, *de novo* methylation of transposable elements occurs in embryos following their demethylation in male and female gametophytes[94,95].

Reproduction. Although there is no obvious fertility defect associated with loss of RdDM in *A. thaliana*, there is evidence that Pol IV-dependent siRNAs are important for germ cell specification. *A. thaliana* strains that carry mutations in AGO9 — a reproductive-specific AGO4 family member — inappropriately express a marker of gametic identity in somatic cells that surround the functional spore (which is the precursor of the female gametophyte)[22], whereas maize *ago9* mutants fail to complete meiosis and generate functional diploid gametes[96]. It is unclear why similar mutations would both promote (in *A. thaliana*) and repress (in maize) the germ cell programme, but it is possible that these mutations cause derepression of transposons that are adjacent to different developmental regulators in each species, which results in alterations to distinct developmental genes and thus different developmental outcomes.

Alterations in germ cell development are required for diplosporous apomixis — an asexual reproductive strategy that generates maternal clones through seeds. In both maize and *A. thaliana*, apomixis is correlated with downregulation or heterochronic expression of non-CG methyltransferases that are associated with RdDM during ovule development[97], which suggests that RdDM promotes sexual reproduction.

Genomic imprinting. Genomic imprinting occurs in mammals and flowering plants[98], and there are numerous indications that genomic imprinting is associated with RdDM, although the nature of this interaction is unknown. All known imprinted genes in *A. thaliana* are either proximal to or overlapping siRNA-encoding loci[99]. Pol IV-dependent siRNA expression is uniparental at many loci in developing *A. thaliana* and rice seeds[100,101], and the imprinting control gene *DEMETER* establishes differential methylation patterns by demethylating maternal transposons, which are normally targets of RdDM[102,103]. These observations led to the

Epialleles
Alleles that differ in transcriptional level from other genetically identical alleles, frequently owing to DNA methylation. Some epialleles are faithfully transmitted to the progeny.

Lamarckian inheritance
The hypothesis that an organism can pass on traits acquired during its lifetime to its progeny.

Gametophytes
The multicellular structure formed through mitosis from a single haploid spore. Male and female gametophytes contain sperm and egg cells, respectively.

Diplosporous apomixis
A process of reproduction whereby failure of meiosis produces an unreduced female gametophyte. An embryo then develops from the diploid egg cell and forms a clone of the maternal plant.

Heterochronic
Pertaining to an evolutionary change in the timing of a developmental process so that a character or process occurs earlier or later in ontogeny, or grows at a different rate.

Genomic imprinting
A phenomenon whereby differential epigenetic marks on maternally and paternally derived alleles result in uniparental gene expression.

hypotheses that imprinted genes and siRNA-producing loci are coordinately regulated[99,102,104], or that uniparentally expressed siRNAs function in *cis* to initiate or maintain differential methylation. However, *DEMETER* does not control uniparental siRNA production, and only a subset of imprinted genes become biallelically expressed in RdDM mutants[105,106], which makes the link between RdDM and genomic imprinting unclear. One possibility is that RdDM generates the parental epigenetic asymmetry that is necessary for the evolution of genomic imprinting[107].

Interallelic communication. In contrast to imprinted loci, which maintain different epigenetic states at each allele, during paramutation silent alleles transmit TGS to active sister alleles[108]. In maize, strains that carry mutations in one of several components of the Pol IV- and Pol V-mediated RdDM pathway are defective for paramutation; these components include orthologues of RDR2, NRPD1 and NRPD2, as well as a SWI/SNF nucleosome remodelling protein that is similar to DRD1 and CLSY1 (REF. 108). The production of siRNAs and the presence of DNA methylation at sites of paramutation further support a role for RdDM in paramutation and suggest a model whereby siRNAs produced at a silent allele trigger both methylation and additional siRNA production at the initially active sister allele. Indeed, production of siRNAs from transgenic hairpin RNAs facilitates paramutation at loci that match the siRNA sequence[109]. However, there is little difference in methylation and siRNA production between active and silent alleles[109], and not all mutants of RdDM components show the same defects in paramutation[108]. Although paramutation is only characterized at a few loci, similar interallelic communication might be widespread during intraspecific hybridization (see below).

Communication of epigenetic states also occurs at non-allelic sites as demonstrated in crosses between *A. thaliana* strains with active transposon copies and strains with silent copies. When the copy number of the long terminal repeat (LTR) retrotransposon *EVD* increases above a threshold, RdDM is triggered to silence Pol II transcription of the element and to initiate Pol IV-mediated production of siRNAs[69]. These silent copies can trigger RdDM at Pol II-active *EVD* copies that are introduced through crossing, even when total *EVD* copy number in the hybrid is below the threshold needed to initiate RdDM[69]. However, experiments using a different LTR retrotransposon indicate that silent elements are unable to maintain silencing when transposon copy number drops[73], which indicates that there might be additional requirements for *trans*-homologue communication.

Intercellular communication. siRNAs also facilitate communication between nuclei. Release of transposon silencing in the nucleus of the pollen vegetative cell induces production of siRNAs that move to sperm cells, in which they might affect silencing in the following generation[110]. Similarly, in the endosperm, demethylation of transposons and coincident siRNA production imply that parallel movement might occur between the endosperm and the embryo[104], but direct evidence for this movement is currently lacking. In *A. thaliana* AGO9 accumulates at high levels in somatic cells that surround the female gametophyte, but transposon silencing is disrupted in the gametophyte of *ago9* mutants[22], which suggests movement of the AGO9–siRNA complex between the soma and the germ line.

Intercellular movement of siRNAs is better understood in vegetative tissues, in which cytoplasmic tunnels connect neighbouring cells and eventually lead to long-distance movement through the vascular system[111,112]. Intercellularly mobile siRNAs can elicit RdDM at genomic loci in recipient tissues[111,112], including the growing root meristems[113]. This systemic spread of RdDM may provide a mechanism for the transmission of stress responses from the affected cells to the germ line, in which they can affect stress responses in subsequent generations.

Genome interactions. The abundance of Pol IV-dependent siRNAs during seed development suggests that RdDM mediate epigenetic interactions between maternal and paternal genomes during hybridization. Generally, hybrids are expected to show additive gene expression, which results in expression level of each gene that is the average of the expression in the two parental lines (known as the mid-parental value (MPV)). Interactions between maternal and paternal genomes cause deviations from the MPV that can be similar to the high-expressing parent or the low-expressing parent, or that can be transgressive (that is, beyond the levels detected in the parents). In interspecific hybrids, non-additive expression can result in either hybrid vigour (that is, increased fitness relative to that expected from the parental lines) or the emergence of novel phenotypes. There is active research concerning how Pol IV-dependent siRNAs and RdDM are influenced by hybridization, and how they in turn influence hybrid gene expression.

Intraspecific hybrids in several species have non-additive levels of 24-nucleotide siRNAs and DNA methylation relative to their parental species[114–119]. This effect is strongest at loci where the parental siRNA expression and DNA methylation levels differ greatly[114,118–120]. The link between non-additive methylation and siRNA expression is probably due to interallelic RdDM. If one allele produces a high level of siRNAs, then these siRNAs can induce methylation of the sister allele in *trans* and trigger additional siRNA production. Conversely, if the expressing allele produces a low level of siRNAs, then these might be insufficient to maintain methylation and siRNA production on either allele when paired with a non-expressing allele[118,120].

Levels of 24-nucleotide siRNAs are also altered, and transposons are upregulated in F_1 interspecific hybrids between *A. thaliana* and *Arabidopsis arenosa* or in nascent wheat hybrids[121–124]. In tomato introgression lines that are generated by hybridizing *Solanum lycopersicum* (cultivated tomato) with *Solanum pennellii* (wild tomato) followed by repeated backcrossing to *S. lycopersicum*, small

Paramutation
A process whereby a transcriptionally silent allele confers meiotically heritable silencing on an active sister allele.

Hybridization
Crossing of two different plant varieties to combine valuable traits from each variety.

Vegetative cell
A haploid cell in the male gametophyte (that is, the pollen grain) that assists fertilization but that does not directly contribute to the zygote.

Endosperm
A tissue in the seed that supports the growth of the embryo. Endosperm is produced after fertilization of the diploid (2N) central cell by a haploid (1N) sperm cell, which creates a maternal:paternal genome ratio of 2:1.

Meristems
Regions of undifferentiated cells at the shoot or root apex that is responsible for cell division and organogenesis. All aerial tissues, including the germ line, arise from the shoot meristem, and all root tissues arise from the root meristem.

Additive gene expression
Gene expression in a hybrid that is the average of the expression levels in the two parental lines.

Interspecific hybrids
Crosses between two closely related but distinct species.

Hybrid vigour
Increase in fitness associated with crosses between distinct inbred strains.

Introgression lines
Lines into which defined DNA segments have been introduced from a different line through backcrossing.

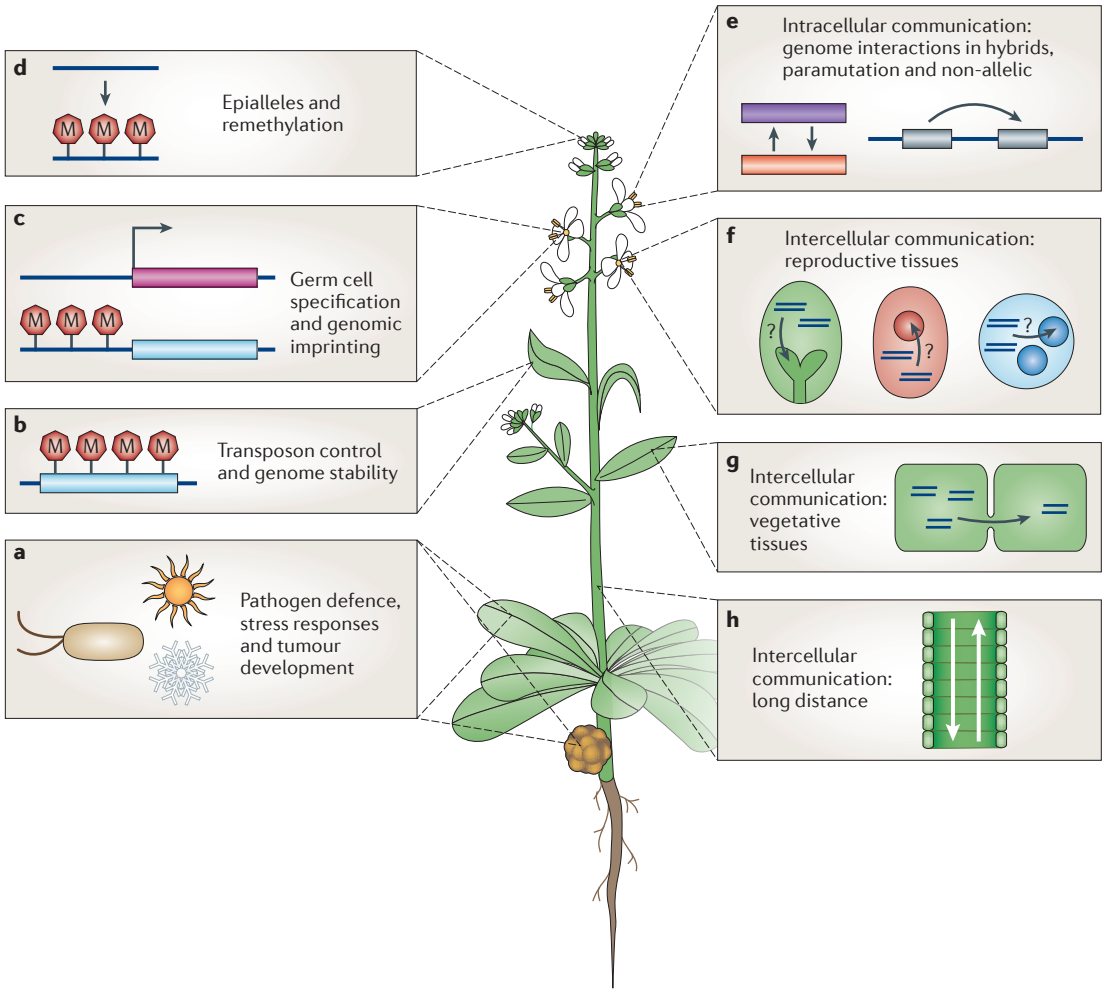


Figure 4 | **Biological processes involving siRNAs and RdDM components.** **a** | RNA-directed DNA methylation (RdDM) is implicated in defence against some viral and bacterial pathogens, and affects the development of pathogenic tumours. Biotic and abiotic stress responses include RdDM-mediated changes in DNA methylation. **b** | High copy-number transposable elements (shown in blue) are transcriptionally silenced by RdDM, which promotes genome stability. **c** | Small interfering RNAs (siRNAs) and the RdDM machinery are also involved in specification of female germ cells and might be involved in establishment or maintenance of parent-of-origin genomic imprints. **d** | By initiating methylation that can be passed to progeny and stably maintained, RdDM might establish epialleles. Conversely, production of siRNAs can remethylate DNA after loss of epigenetic marks. **e** | siRNAs can also mediate interactions between maternal and paternal genomes upon hybridization, particularly at sites of paramutation. RdDM can mediate both allelic and non-allelic communication within a cell. **f** | Alternatively, RdDM can function in a non-cell autonomous way to communicate information between cells. In reproductive tissues, siRNA movement might occur between the pollen vegetative nucleus and the sperm cells (blue), between somatic tissues and the megaspore mother cell (pink), or between the endosperm and the embryo after fertilization (green). **g** | Intercellular transport of siRNAs in vegetative tissues is likely to occur through cytoplasmic tunnels that connect neighbouring cells. **h** | Intercellular transport of siRNAs can eventually occur over long distances and bidirectionally through the vascular system to induce methylation at remote sites.

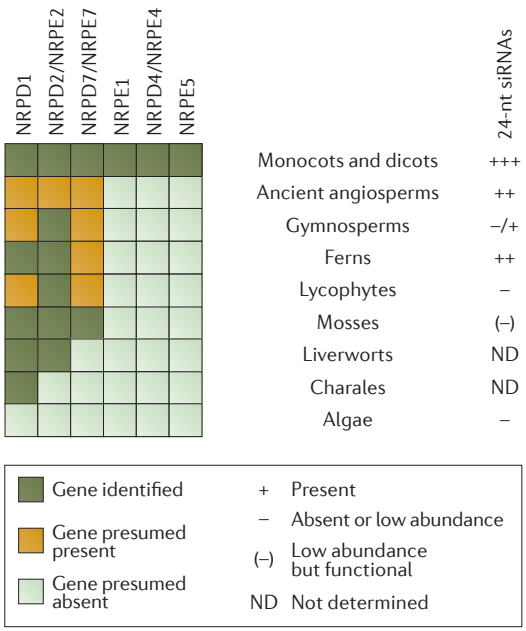
RNA-producing loci are either upregulated or downregulated beyond parental levels and can be correlated with increased DNA methylation and decreased expression of overlapping genes[125]. These observations imply that interspecific crosses might misregulate RdDM and contribute to the novel phenotypes observed in hybrids.

Conclusions

From fairly humble beginnings, research on RdDM over the past two decades has revealed a complex and multifaceted epigenetic pathway that reaches into many aspects of plant biology (FIG. 4). The canonical RdDM pathway — in which Pol IV-, RDR2- and DCL3-dependent 24-nucleotide siRNAs trigger AGO4- and Pol V-mediated DNA methylation — has emerged as only one variant of small RNA-guided DNA methylation in plants. The variation in RdDM pathways reflects considerable overlap with components and small RNA size classes that are traditionally associated with PTGS pathways. The areas of convergence probably account for the reported redundancy among DCL enzymes in RdDM[126] and early reports of associations between PTGS and RdDM[5].

Box 2 | **Evolution of RdDM**

RNA polymerase IV (Pol IV) and Pol V are plant-specific DNA-dependent RNA polymerases that are specialized for RNA-directed DNA methylation (RdDM) and transcriptional gene silencing. Similar to all eukaryotic polymerases, Pol IV and Pol V are holoenzymes composed of at least 12 individual subunits (known as nuclear RNA polymerase D (NRPD) subunits for Pol IV and NRPE subunits for Pol V) (BOX 1). Many of these subunits are shared with Pol II (in which the same proteins are known as NRPB subunits), whereas others arose through duplication of Pol II components[129,130,139]. The first evidence for emergence of a specialized subunit is in Charales — an order of algae that is thought to be the closest ancestor of land plants — which contains a gene encoding NRPD1 (REF. 139) (see the figure). Another component of the catalytic core, the shared NRPD2/NRPE2 subunit of Pol IV and Pol V, is absent from Charales but present in all groups of non-flowering land plants, including liverworts, mosses, lycophytes, ferns and gymnosperms[139]. Similarly, NRPD7 and NPRE7 are found in mosses and presumed to be present in all land plants[129]. Other specialized subunits, including NRPE1, NRPD4/NRPE4 and NRPE5, have been identified only in monocots and dicots[129,139]. This stepwise evolution suggests that the unique properties of Pol IV and Pol V arose gradually throughout plant evolution rather than as a result of a rapid evolutionary event.



Small interfering RNAs (siRNAs) of 24 nucleotides (nt) are abundant in monocots and dicots, and are present at lower levels in ancient angiosperms, gymnosperms and ferns[140] (see the figure). It has been suggested that among gymnosperms, conifers have lost RdDM owing to an apparent absence of 24-nt siRNAs[140,141]. However, 24-nt siRNAs were recently identified in *Picea abies* (Norway spruce) flowers[142] and *Larix leptolepis* (Japanese larch) somatic embryos[143], which indicates that RdDM might be restricted to reproductive tissues in conifers. siRNAs of 24 nt were also identified in *Cunninghamia lanceolata* (China fir) libraries that were produced from a mixture of seed, seedling, leaf, stem and callus RNA[144]. Conifer 24-nt siRNAs are highly complex[144] and are associated with genome repeats[142], which suggests that they are Pol IV-dependent siRNAs involved in RdDM and transcriptional gene silencing.

siRNAs of 24 nt are absent from *Chlamydomonas reinhardtii* (green alga)[145] and are not readily detected in *Selaginella moellendorffii* (lycophyte) or *Physcomitrella patens* (moss)[146]. However, an analysis of siRNAs in *P. patens* indicates that low levels of siRNAs (22–24 nt) are produced from genomic regions that are enriched for transposable elements and DNA methylation[147]. In *P. patens*, DICER-LIKE 3 (DCL3) is required for accumulation of these siRNAs, and silencing of two long terminal repeat (LTR) retrotransposons is relieved in *dcl3* mutants[147]. These observations suggest that RdDM and transcriptional silencing through Pol IV comprise an ancient process that is conserved among all land plants, although Pol IV activity varies in intensity and developmental extent. The continued evolution of specialized components, including NRPE1, suggests mechanistic differences in RdDM across plant genera. These differences and the evolutionary pressure behind remain to be defined.

Further investigation is required to answer important questions concerning Pol IV and Pol V recruitment, and the mechanism of *de novo* methylation. The preferential targeting of RdDM to transposons and repetitive genomic regions relies on chromatin signatures or other factors that have not yet been fully defined. The Pol IV recruitment factor SHH1 recognizes pre-existing H3K9me marks, but the trigger for this histone methylation remains unclear. As SHH1 is involved in recruiting Pol IV to only a subset of loci, additional factors must recruit Pol IV to loci that are not associated with this protein. In some cases, Pol V recruitment relies on a certain level of pre-existing DNA methylation that is recognized by SUVH2 or SUVH9. An emerging theme illustrated by non-canonical RDR6-dependent RdDM is that Pol II-dependent pathways for DNA methylation can establish an epigenetic landscape that is favourable for Pol IV and Pol V recruitment.

Although current models depict siRNAs base-pairing with Pol V-generated scaffold RNAs, it is conceivable that uniquely in plants, siRNAs interact directly with the target DNA during Pol V transcription. Such a specialized mechanism would explain the restriction of methylation to the DNA region with sequence homology to siRNAs and account for the elaborate and mainly plant-specific transcriptional machinery required for RdDM. Despite the attractiveness of current models, further work is needed to ascertain the precise roles of the plant-specific factors in the mechanism of RdDM. Unusual features of Pol IV and Pol V, which are likely to affect their template choices and enzymatic activities (BOX 1), merit further investigation.

After initial characterization of siRNAs and methylation, research has begun to investigate RdDM in response to a range of biotic and abiotic stimuli, and how RdDM might mediate physiological responses to

stress. Particularly interesting is the potential to prime these responses in subsequent generations, which gives rise to progeny that are better adapted to their environment. As research in this area moves from observation to mechanism, we might eventually understand how stress triggers genome remodelling.

Future research is also likely to unravel the mechanism of intercellular siRNA movement. Cell-to-cell movement in vegetative tissues can occur passively through cytoplasmic tunnels; however, the vascular system is cytoplasmically isolated from surrounding cells, which suggests that active transport is required to move siRNAs systemically. Whether cytoplasmic tunnels exist between the embryo and the endosperm is unclear[127], and direct evidence for movement of endogenous 24-nucleotide siRNAs between the vegetative nucleus and the sperm cells is also lacking. Constraints on siRNA movement and function in recipient tissues will also be interesting areas for future research.

Another promising area of future RdDM research is potential variation between species: developmental phenotypes of mutants of single RdDM components are weaker in *A. thaliana* than in maize. RdDM might be more important in maize owing to the increased repeat content of the maize genome or to its tendency to outcross. Intraspecific differences in RdDM might underlie hybrid vigour, and it will be interesting to discover how genomic differences in siRNA accumulation might affect species barriers. Further details on the evolution of RdDM components in the plant kingdom (BOX 2) and comparative analyses of RdDM mutants from multiple species[128] will be crucial to fully understand the range of RdDM pathways in plants.

Our understanding of RdDM has progressed greatly over the past decades, including detailed descriptions of the molecular mechanism and a growing knowledge of its biological importance. However, many questions still remain, which will ensure that the next decades of RdDM research will be fruitful and fascinating.

1. Castel, S. E. & Martienssen, R. A. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nature Rev. Genet.* **14**, 100–112 (2013).
2. Luteijn, M. J. & Ketting, R. F. PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nature Rev. Genet.* **14**, 523–534 (2013).
3. Jones, A. L., Thomas, C. L. & Maule, A. J. *De novo* methylation and co-suppression induced by a cytoplasmically replicating plant RNA virus. *EMBO J.* **17**, 6385–6393 (1998).
4. Wassenegger, M., Heimes, S., Riedel, L. & Sanger, H. L. RNA-directed *de novo* methylation of genomic sequences in plants. *Cell* **76**, 567–576 (1994).
5. Eun, C. *et al.* Use of forward genetic screens to identify genes required for RNA-directed DNA methylation in *Arabidopsis thaliana*. *Cold Spring Harb. Symp. Quant. Biol.* **77**, 195–204 (2012).
6. Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A. & Matzke, A. J. M. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J.* **19**, 5194–5201 (2000).
7. Chan, S. W. *et al.* RNA silencing genes control *de novo* DNA methylation. *Science* **303**, 1336 (2004).
8. Haag, J. R. & Pikaard, C. S. Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. *Nature Rev. Mol. Cell Biol.* **12**, 483–492 (2011).
9. Yamanaka, S. *et al.* RNAi triggered by specialized machinery silences developmental genes and retrotransposons. *Nature* **493**, 557–560 (2013).
10. Stroud, H., Greenberg, M. V., Feng, S., Bernatavichute, Y. V. & Jacobsen, S. E. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell* **152**, 352–364 (2013).
This methylome analysis of 86 *A. thaliana* gene silencing mutants provides a resource for epigenetics researchers.
11. Schmitz, R. J. *et al.* Patterns of population epigenomic diversity. *Nature* **495**, 193–198 (2013).
12. Mosher, R. A., Schwach, F., Studholme, D. & Baulcombe, D. C. PolIVb influences RNA-directed DNA-methylation independently of its role in siRNA biogenesis. *Proc. Natl Acad. Sci. USA* **105**, 3145–3150 (2008).
13. Zhang, X., Henderson, I. R., Lu, C., Green, P. J. & Jacobsen, S. E. Role of RNA polymerase IV in plant small RNA metabolism. *Proc. Natl Acad. Sci. USA* **104**, 4536–4541 (2007).
14. Law, J. A. *et al.* Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* **498**, 385–389 (2013).
Using ChIP–seq, this study identifies 928 high-confidence Pol IV peaks. The Pol IV-interacting protein SHH1 is shown to interact with H3K9me and recruit Pol IV to chromatin at 44% of Pol IV target sites.
15. Zhang, H. *et al.* DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of Pol IV. *Proc. Natl Acad. Sci. USA* **110**, 8290–8295 (2013).
16. Haag, J. R. *et al. In vitro* transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. *Mol. Cell* **48**, 811–818 (2012).
17. Law, J. A., Vashisht, A. A., Wohlschlegel, J. A. & Jacobsen, S. E. SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. *PLoS Genet.* **7**, e1002195 (2011).
18. Smith, L. M. *et al.* An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in *Arabidopsis*. *Plant Cell* **19**, 1507–1521 (2007).
19. Ji, L. J. & Chen, X. M. Regulation of small RNA stability: methylation and beyond. *Cell Res.* **22**, 624–636 (2012).
20. Eun, C. *et al.* AGO6 functions in RNA-mediated transcriptional gene silencing in shoot and root meristems in *Arabidopsis thaliana*. *PLoS ONE* **6**, e25730 (2011).
21. Havecker, E. R. *et al.* The *Arabidopsis* RNA-directed DNA methylation Argonautes functionally diverge based on their expression and interaction with target loci. *Plant Cell* **22**, 321–334 (2010).
22. Olmedo-Monfil, V. *et al.* Control of female gamete formation by a small RNA pathway in *Arabidopsis*. *Nature* **464**, 628–632 (2010).
23. Wierzbicki, A. T., Haag, J. R. & Pikaard, C. S. Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* **135**, 635–648 (2008).
24. Zhong, X. *et al.* DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. *Nature Struct. Mol. Biol.* **19**, 870–875 (2012).
25. Wierzbicki, A. T. *et al.* Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the *Arabidopsis* epigenome. *Genes Dev.* **26**, 1825–1836 (2012).
Using ChIP–seq, this study identifies 1,157 high-confidence Pol V peaks that are distributed throughout the genome. Most of these peaks are associated with 24-nucleotide siRNAs and CHH methylation, but a consensus sequence for Pol V binding could not be discerned.
26. Lee, T. F. *et al.* RNA polymerase V-dependent small RNAs in *Arabidopsis* originate from small, intergenic loci including most SINE repeats. *Epigenetics* **7**, 781–795 (2012).
27. Zemach, A. *et al.* The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**, 193–205 (2013).
This paper shows that most transposons in *A. thaliana* are silenced by DDM1 and RdDM pathways that act preferentially in highly heterochromatic and more euchromatic regions, respectively. CMT2 is identified as a CHH methyltransferase that functions independently of siRNAs in heterochromatin.
28. Zheng, Q. *et al.* RNA polymerase V targets transcriptional silencing components to promoters of protein-coding genes. *Plant J.* **73**, 179–189 (2013).
29. Schoft, V. K. *et al.* Induction of RNA-directed DNA methylation upon decondensation of constitutive heterochromatin. *EMBO Rep.* **10**, 1015–1021 (2009).
30. Stroud, H. *et al.* Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. *Nature Struct. Mol. Biol.* **21**, 64–72 (2014).
31. Johnson, L. M., Law, J. A., Khattar, A., Henderson, I. R. & Jacobsen, S. E. SRA-domain proteins required for DRM2-mediated *de novo* DNA methylation. *PLoS Genet.* **4**, e1000280 (2008).
32. Kuhlmann, M. & Mette, M. F. Developmentally non-redundant SET domain proteins SUVH2 and SUVH9 are required for transcriptional gene silencing in *Arabidopsis thaliana*. *Plant Mol. Biol.* **79**, 623–633 (2012).
33. Johnson, L. M. *et al.* SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. *Nature* **507**, 124–128 (2014).
34. Liu, Z. W. *et al.* The SET domain proteins SUVH2 and SUVH9 are required for Pol V occupancy at RNA-directed DNA methylation loci. *PLoS Genet.* **10**, e1003948 (2014).
References 33 and 34 show that Pol V is recruited to methylated DNA at some loci by methyl DNA-binding proteins SUVH2 and SUVH9.
35. Law, J. A. *et al.* A protein complex required for polymerase V transcripts and RNA-directed DNA methylation in *Arabidopsis*. *Curr. Biol.* **20**, 951–956 (2010).
36. Wierzbicki, A. T., Ream, T. S., Haag, J. R. & Pikaard, C. S. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nature Genet.* **41**, 630–634 (2009).
37. Gao, Z. H. *et al.* An RNA polymerase II- and AGO4-associated protein acts in RNA-directed DNA methylation. *Nature* **465**, 106–109 (2010).
38. Huang, L. F. *et al.* An atypical RNA polymerase involved in RNA silencing shares small subunits with RNA polymerase II. *Nature Struct. Mol. Biol.* **16**, 91–93 (2009).
39. Bies-Etheve, N. *et al.* RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. *EMBO Rep.* **10**, 649–654 (2009).

40. He, X. J. *et al.* An effector of RNA-directed DNA methylation in *Arabidopsis* is an ARGONAUTE 4- and RNA-binding protein. *Cell* **137**, 498–508 (2009).
41. Ausin, I. *et al.* INVOLVED IN DE NOVO 2-containing complex involved in RNA-directed DNA methylation in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **109**, 8374–8381 (2012).
42. Lorkovic, Z. J., Naumann, U., Matzke, A. J. & Matzke, M. Involvement of a GHKL ATPase in RNA-directed DNA methylation in *Arabidopsis thaliana*. *Curr. Biol.* **22**, 933–938 (2012).
43. Zhang, C. J. *et al.* IDN2 and its paralogs form a complex required for RNA-directed DNA methylation. *PLoS Genet.* **8**, e1002693 (2012).
44. Xie, M., Ren, G., Costa-Nunes, P., Pontes, O. & Yu, B. A subgroup of SGS3-like proteins act redundantly in RNA-directed DNA methylation. *Nucleic Acids Res.* **40**, 4422–4431 (2012).
45. Finke, A., Kuhlmann, M. & Mette, M. F. IDN2 has a role downstream of siRNA formation in RNA-directed DNA methylation. *Epigenetics* **7**, 950–960 (2012).
46. Zhu, Y., Rowley, M. J., Bohmdorfer, G. & Wierzbicki, A. T. A. SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing. *Mol. Cell* **49**, 298–309 (2013).
47. Bernatavichute, Y. V., Zhang, X., Cokus, S., Pellegrini, M. & Jacobsen, S. E. Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in *Arabidopsis thaliana*. *PLoS ONE* **3**, e3156 (2008).
48. Enke, R. A., Dong, Z. & Bender, J. Small RNAs prevent transcription-coupled loss of histone H3 lysine 9 methylation in *Arabidopsis thaliana*. *PLoS Genet.* **7**, e1002350 (2011).
49. Liu, X. *et al.* HDA6 directly interacts with DNA methyltransferase MET1 and maintains transposable element silencing in *Arabidopsis*. *Plant Physiol.* **158**, 119–129 (2012).
50. To, T. K. *et al.* *Arabidopsis* HDA6 regulates locus-directed heterochromatin silencing in cooperation with MET1. *PLoS Genet.* **7**, e1002055 (2011).
51. Deleris, A. *et al.* Involvement of a Jumonji-C domain-containing histone demethylase in DRM2-mediated maintenance of DNA methylation. *EMBO Rep.* **11**, 950–955 (2010).
52. Searle, I. R., Pontes, O., Melnyk, C. W., Smith, L. M. & Baulcombe, D. C. JMJ14, a JmjC domain protein, is required for RNA silencing and cell-to-cell movement of an RNA silencing signal in *Arabidopsis*. *Genes Dev.* **24**, 986–991 (2010).
53. Greenberg, M. V. *et al.* Interplay between active chromatin marks and RNA-directed DNA methylation in *Arabidopsis thaliana*. *PLoS Genet.* **9**, e1003946 (2013).
54. Sridhar, V. V. *et al.* Control of DNA methylation and heterochromatic silencing by histone H2B deubiquitination. *Nature* **447**, 735–738 (2007).
55. Brabbs, T. R. *et al.* The stochastic silencing phenotype of *Arabidopsis morc6* mutants reveals a role in efficient RNA-directed DNA methylation. *Plant J.* **75**, 836–846 (2013).
56. Moissiard, G. *et al.* MORC family ATPases required for heterochromatin condensation and gene silencing. *Science* **336**, 1448–1451 (2012).
57. Becker, C. *et al.* Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* **480**, 245–249 (2011).
58. Schmitz, R. J. *et al.* Transgenerational epigenetic instability is a source of novel methylation variants. *Science* **334**, 369–373 (2011).
59. Jones, L., Ratcliff, F. & Baulcombe, D. C. RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. *Curr. Biol.* **11**, 747–757 (2001).
60. Teixeira, F. K. *et al.* A role for RNAi in the selective correction of DNA methylation defects. *Science* **323**, 1600–1604 (2009).
61. Law, J. A. & Jacobsen, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Rev. Genet.* **11**, 204–220 (2010).
62. Huettel, B. *et al.* Endogenous targets of RNA-directed DNA methylation and Pol IV in *Arabidopsis*. *EMBO J.* **25**, 2828–2836 (2006).
63. Zhang, H. & Zhu, J. K. Active DNA demethylation in plants and animals. *Cold Spring Harb. Symp. Quant. Biol.* **77**, 161–173 (2012).
64. Li, X. *et al.* Antisilencing role of the RNA-directed DNA methylation pathway and a histone acetyltransferase in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **109**, 11425–11430 (2012).
65. Wu, L. *et al.* DNA methylation mediated by a microRNA pathway. *Mol. Cell* **38**, 465–475 (2010).
66. Khraiwesh, B. *et al.* Transcriptional control of gene expression by microRNAs. *Cell* **140**, 111–122 (2010).
67. Wu, L., Mao, L. & Qi, Y. Roles of Dicer-like and Argonaute proteins in *TAS*-derived small interfering RNA-triggered DNA methylation. *Plant Physiol.* **160**, 990–999 (2012).
68. Nuthikattu, S. *et al.* The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21–22 nucleotide small interfering RNAs. *Plant Physiol.* **162**, 116–131 (2013).
69. Mari-Ordonez, A. *et al.* Reconstructing *de novo* silencing of an active plant retrotransposon. *Nature Genet.* **45**, 1029–1039 (2013).
This study follows a newly activated retrotransposon through many generations. Initially silenced through PTGS, saturation of DCL2 and DCL4 activity leads to 24-nucleotide siRNA production by DCL3 and initiates canonical RdDM.
70. Pontier, D. *et al.* NERD, a plant-specific GW protein, defines an additional RNAi-dependent chromatin-based pathway in *Arabidopsis*. *Mol. Cell* **48**, 121–132 (2012).
This paper identifies NERD as an AGO hook protein and, together with reference 68, defines more fully the non-canonical RDR6-dependent RdDM pathway that is independent of Pol IV and that relies instead on Pol II.
71. Zheng, B. *et al.* Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. *Genes Dev.* **23**, 2850–2860 (2009).
72. You, W., Lorkovic, Z. J., Matzke, A. J. & Matzke, M. Interplay among RNA polymerases II, IV and V in RNA-directed DNA methylation at a low copy transgene locus in *Arabidopsis thaliana*. *Plant Mol. Biol.* **82**, 85–96 (2013).
73. Perez-Hormaeche, J. *et al.* Invasion of the *Arabidopsis* genome by the tobacco retrotransposon *Tnt1* is controlled by reversible transcriptional gene silencing. *Plant Physiol.* **147**, 1264–1278 (2008).
74. Ito, H. *et al.* An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* **472**, 115–119 (2011).
This study shows that heat stress triggers transcription of the *ONSEN* retrotransposon, but such transposition is only possible in some RdDM mutant backgrounds. Insertion of *ONSEN* near genes can confer heat-responsive transcription.
75. Ruiz-Ferrer, V. & Voinnet, O. Roles of plant small RNAs in biotic stress responses. *Annu. Rev. Plant Biol.* **60**, 485–510 (2009).
76. Buchmann, R. C., Asad, S., Wolf, J. N., Mohannath, G. & Bisaro, D. M. Geminivirus AL2 and L2 proteins suppress transcriptional gene silencing and cause genome-wide reductions in cytosine methylation. *J. Virol.* **83**, 5005–5013 (2009).
77. Yang, L. P. *et al.* C2-mediated decrease in DNA methylation, accumulation of siRNAs, and increase in expression for genes involved in defense pathways in plants infected with beet severe curly top virus. *Plant J.* **73**, 910–917 (2013).
78. Zhang, Z. *et al.* BSCTV C2 attenuates the degradation of SAMDC1 to suppress DNA methylation-mediated gene silencing in *Arabidopsis*. *Plant Cell* **23**, 273–288 (2011).
79. Romanel, E. *et al.* Global alteration of microRNAs and transposon-derived small RNAs in cotton (*Gossypium hirsutum*) during Cotton leafroll dwarf polerovirus (CLRDV) infection. *Plant Mol. Biol.* **80**, 443–460 (2012).
80. Lopez, A., Ramirez, V., Garcia-Andrade, J., Flors, V. & Vera, P. The RNA silencing enzyme RNA polymerase V is required for plant immunity. *PLoS Genet.* **7**, e1002434 (2011).
81. Dowen, R. H. *et al.* Widespread dynamic DNA methylation in response to biotic stress. *Proc. Natl Acad. Sci. USA* **109**, E2183–E2191 (2012).
This genome-wide analysis of DNA methylation during pathogen infection reveals transposon-associated demethylation that affects expression of neighbouring genes.
82. Luna, E., Bruce, T. J., Roberts, M. R., Flors, V. & Ton, J. Next-generation systemic acquired resistance. *Plant Physiol.* **158**, 844–853 (2012).
83. Yu, A. *et al.* Dynamics and biological relevance of DNA demethylation in *Arabidopsis* antibacterial defense. *Proc. Natl Acad. Sci. USA* **110**, 2389–2394 (2013).
84. Gohlke, J. *et al.* DNA methylation mediated control of gene expression is critical for development of crown gall tumors. *PLoS Genet.* **9**, e1003267 (2013).
85. Ou, X. *et al.* Transgenerational inheritance of modified DNA methylation patterns and enhanced tolerance induced by heavy metal stress in rice (*Oryza sativa* L.). *PLoS ONE* **7**, e41143 (2012).
86. Tricker, P. J., Gibbings, J. G., Rodriguez Lopez, C. M., Hadley, P. & Wilkinson, M. J. Low relative humidity triggers RNA-directed *de novo* DNA methylation and suppression of genes controlling stomatal development. *J. Exp. Bot.* **63**, 3799–3813 (2012).
87. Popova, O. V., Dinh, H. Q., Aufsatz, W. & Jonak, C. The RdDM pathway is required for basal heat tolerance in *Arabidopsis*. *Mol. Plant* **6**, 396–410 (2013).
88. Yao, Y., Bilichak, A., Golubov, A., Blevins, T. & Kovalchuk, I. Differential sensitivity of *Arabidopsis* siRNA biogenesis mutants to genotoxic stress. *Plant Cell Rep.* **29**, 1401–1410 (2010).
89. Verhoeven, K. J., Jansen, J. J., van Dijk, P. J. & Biere, A. Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytol.* **185**, 1108–1118 (2010).
90. Cubas, P., Vincent, C. & Coen, E. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* **401**, 157–161 (1999).
91. Manning, K. *et al.* A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nature Genet.* **38**, 948–952 (2006).
92. Suter, L. & Widmer, A. Environmental heat and salt stress induce transgenerational phenotypic changes in *Arabidopsis thaliana*. *PLoS ONE* **8**, e60364 (2013).
93. Luna, E. & Ton, J. The epigenetic machinery controlling transgenerational systemic acquired resistance. *Plant Signal. Behav.* **7**, 615–618 (2012).
94. Calarco, J. P. *et al.* Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell* **151**, 194–205 (2012).
95. Jullien, P. E., Susaki, D., Yelagandula, R., Higashiyama, T. & Berger, F. DNA methylation dynamics during sexual reproduction in *Arabidopsis thaliana*. *Curr. Biol.* **22**, 1825–1830 (2012).
96. Singh, M. *et al.* Production of viable gametes without meiosis in maize deficient for an ARGONAUTE protein. *Plant Cell* **23**, 443–458 (2011).
97. Garcia-Aguilar, M., Michaud, C., Leblanc, O. & Grimanelli, D. Inactivation of a DNA methylation pathway in maize reproductive organs results in apomixis-like phenotypes. *Plant Cell* **22**, 3249–3267 (2010).
98. Garnier, O., Laoueille-Duprat, S. & Spillane, C. Genomic imprinting in plants. *Epigenetics* **3**, 14–20 (2008).
99. Gehring, M., Bubb, K. L. & Henikoff, S. Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science* **324**, 1447–1451 (2009).
100. Mosher, R. A. *et al.* Uniparental expression of PolIV-dependent siRNAs in developing endosperm of *Arabidopsis*. *Nature* **460**, 283–286 (2009).
101. Rodrigues, J. A. *et al.* Imprinted expression of genes and small RNA is associated with localized hypomethylation of the maternal genome in rice endosperm. *Proc. Natl Acad. Sci. USA* **110**, 7934–7939 (2013).
102. Hsieh, T. F. *et al.* Genome-wide demethylation of *Arabidopsis* endosperm. *Science* **324**, 1451–1454 (2009).
103. Ibarra, C. A. *et al.* Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* **337**, 1360–1364 (2012).
104. Mosher, R. A. & Melnyk, C. W. siRNAs and DNA methylation: seedy epigenetics. *Trends Plant Sci.* **15**, 204–210 (2010).
105. Mosher, R. A. *et al.* An atypical epigenetic mechanism affects uniparental expression of Pol IV-dependent siRNAs. *PLoS ONE* **6**, e25756 (2011).
106. Vu, T. M. *et al.* RNA-directed DNA methylation regulates parental genomic imprinting at several loci in *Arabidopsis*. *Development* **140**, 2953–2960 (2013).
107. Mosher, R. A. Maternal control of Pol IV-dependent siRNAs in *Arabidopsis* endosperm. *New Phytol.* **186**, 358–364 (2010).
108. Erhard, K. F. Jr & Hollick, J. B. Paramutation: a process for acquiring *trans*-generational regulatory states. *Curr. Opin. Plant Biol.* **14**, 210–216 (2011).
109. Arteaga-Vazquez, M. *et al.* RNA-mediated *trans*-communication can establish paramutation at the *b1* locus in maize. *Proc. Natl Acad. Sci. USA* **107**, 12986–12991 (2010).
110. Slotkin, R. K. *et al.* Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* **136**, 461–472 (2009).

111. Dunoyer, P. *et al.* An endogenous, systemic RNAi pathway in plants. *EMBO J.* **29**, 1699–1712 (2010).
112. Molnar, A. *et al.* Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells. *Science* **328**, 872–875 (2010).
113. Melnyk, C. W., Molnar, A., Bassett, A. & Baulcombe, D. C. Mobile 24 nt small RNAs direct transcriptional gene silencing in the root meristems of *Arabidopsis thaliana*. *Curr. Biol.* **21**, 1678–1683 (2011).
114. Groszmann, M. *et al.* Changes in 24-nt siRNA levels in *Arabidopsis* hybrids suggest an epigenetic contribution to hybrid vigor. *Proc. Natl Acad. Sci. USA* **108**, 2617–2622 (2011).
115. Li, Y., Varala, K., Moose, S. P. & Hudson, M. E. The inheritance pattern of 24 nt siRNA clusters in *Arabidopsis* hybrids is influenced by proximity to transposable elements. *PLoS ONE* **7**, e47043 (2012).
116. He, G. *et al.* Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell* **22**, 17–33 (2010).
117. Barber, W. T. *et al.* Repeat associated small RNAs vary among parents and following hybridization in maize. *Proc. Natl Acad. Sci. USA* **109**, 10444–10449 (2012).
118. Greaves, I. K. *et al. Trans* chromosomal methylation in *Arabidopsis* hybrids. *Proc. Natl Acad. Sci. USA* **109**, 3570–3575 (2012).
This paper analyses DNA methylation in *Arabidopsis* intraspecific hybrids and reveals interactions between chromosomes that result in non-additive methylation. This methylation can alter gene expression and potentially contributes to heterosis.
119. Shen, H. *et al.* Genome-wide analysis of DNA methylation and gene expression changes in two *Arabidopsis* ecotypes and their reciprocal hybrids. *Plant Cell* **24**, 875–892 (2012).
120. Chodavarapu, R. K. *et al.* Transcriptome and methylome interactions in rice hybrids. *Proc. Natl Acad. Sci. USA* **109**, 12040–12045 (2012).
121. Ha, M. *et al.* Small RNAs serve as a genetic buffer against genomic shock in *Arabidopsis* interspecific hybrids and allopolyploids. *Proc. Natl Acad. Sci. USA* **106**, 17835–17840 (2009).
122. Kenan-Eichler, M. *et al.* Wheat hybridization and polyploidization results in deregulation of small RNAs. *Genetics* **188**, 263–272 (2011).
123. Josefsson, C., Dilkes, B. & Comai, L. Parent-dependent loss of gene silencing during interspecies hybridization. *Curr. Biol.* **16**, 1322–1328 (2006).
124. Kashkush, K., Feldman, M. & Levy, A. A. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nature Genet.* **33**, 102–106 (2003).
125. Shivaprasad, P. V., Dunn, R. M., Santos, B. A., Bassett, A. & Baulcombe, D. C. Extraordinary transgressive phenotypes of hybrid tomato are influenced by epigenetics and small silencing RNAs. *EMBO J.* **31**, 257–266 (2012).
126. Henderson, I. R. *et al.* Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nature Genet.* **38**, 721–725 (2006).
127. Kozieradzka-Kiszkurno, M. & Plachno, B. J. Are there symplastic connections between the endosperm and embryo in some angiosperms?—a lesson from the Crassulaceae family. *Protoplasma* **249**, 1081–1089 (2012).
128. Huang, Y., Kendall, T. & Mosher, R. Pol IV-Dependent siRNA production is reduced in *Brassica rapa*. *Biology* **2**, 1210–1223 (2013).
129. Tucker, S. L., Reece, J., Ream, T. S. & Pikaard, C. S. Evolutionary history of plant multisubunit RNA polymerases IV and V: subunit origins via genome-wide and segmental gene duplications, retrotransposition, and lineage-specific subfunctionalization. *Cold Spring Harb. Symp. Quant. Biol.* **75**, 285–297 (2010).
This bioinformatic analysis of transcriptomes across the plant kingdom reveals stepwise evolution of Pol IV- and Pol V-specific subunits.
130. Ream, T. S. *et al.* Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol. Cell* **33**, 192–203 (2009).
131. Landick, R. Functional divergence in the growing family of RNA polymerases. *Structure* **17**, 323–325 (2009).
132. Bellaoui, M. & Gruissem, W. Altered expression of the *Arabidopsis* ortholog of DCL affects normal plant development. *Planta* **219**, 819–826 (2004).
133. Li, C. F. *et al.* An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. *Cell* **126**, 93–106 (2006).
134. Tan, E. H., Blevins, T., Ream, T. S. & Pikaard, C. S. Functional consequences of subunit diversity in RNA polymerases II and V. *Cell Rep.* **1**, 208–214 (2012).
135. Lahmy, S. *et al.* PolV(PolIVb) function in RNA-directed DNA methylation requires the conserved active site and an additional plant-specific subunit. *Proc. Natl Acad. Sci. USA* **106**, 941–946 (2009).
136. Czeko, E., Seizl, M., Augsberger, C., Mielke, T. & Cramer, P. Iwr1 directs RNA polymerase II nuclear import. *Mol. Cell* **42**, 261–266 (2011).
137. He, X. J. *et al.* A conserved transcriptional regulator is required for RNA-directed DNA methylation and plant development. *Genes Dev.* **23**, 2717–2722 (2009).
138. Kanno, T. *et al.* RNA-directed DNA methylation and plant development require an IWR1-type transcription factor. *EMBO Rep.* **11**, 65–71 (2010).
139. Luo, J. & Hall, B. D. A multistep process gave rise to RNA polymerase IV of land plants. *J. Mol. Evol.* **64**, 101–112 (2007).
140. Dolgosheina, E. V. *et al.* Conifers have a unique small RNA silencing signature. *RNA* **14**, 1508–1515 (2008).
141. Morin, R. D. *et al.* Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res.* **18**, 571–584 (2008).
142. Nystedt, B. *et al.* The Norway spruce genome sequence and conifer genome evolution. *Nature* **497**, 579–584 (2013).
143. Zhang, J. *et al.* Dynamic expression of small RNA populations in larch (*Larix leptolepis*). *Planta* **237**, 89–101 (2013).
144. Wan, L. C. *et al.* Identification and characterization of small non-coding RNAs from Chinese fir by high throughput sequencing. *BMC Plant Biol.* **12**, 146 (2012).
145. Molnar, A., Schwach, F., Studholme, D. J., Thuenemann, E. C. & Baulcombe, D. C. miRNAs control gene expression in the single-cell alga *Chlamydomonas reinhardtii*. *Nature* **447**, 1126–U1115 (2007).
146. Axtell, M. J., Snyder, J. A. & Bartel, D. P. Common functions for diverse small RNAs of land plants. *Plant Cell* **19**, 1750–1769 (2007).
147. Cho, S. H. *et al. Physcomitrella patens* DCL3 is required for 22–24 nt siRNA accumulation, suppression of retrotransposon-derived transcripts, and normal development. *PLoS Genet.* **4**, e1000314 (2008).
148. Pikaard, C. S., Haag, J. R., Pontes, O. M., Blevins, T. & Cocklin, R. A transcription fork model for Pol IV and Pol V-dependent RNA-directed DNA methylation. *Cold Spring Harb. Symp. Quant. Biol.* **77**, 205–212 (2012).
149. Panda, K. & Slotkin, R. K. Proposed mechanism for the initiation of transposable element silencing by the RDR6-directed DNA methylation pathway. *Plant Signal. Behav.* **8**, e25206 (2013).
150. Herr, A. J., Jensen, M. B., Dalmay, T. & Baulcombe, D. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**, 118–120 (2005).
151. Onodera, Y. *et al.* Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**, 613–622 (2005).
152. Pontier, D. *et al.* Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. *Genes Dev.* **19**, 2030–2040 (2005).
153. Kanno, T. *et al.* Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nature Genet.* **37**, 761–765 (2005).
154. He, X. J. *et al.* NRPD4, a protein related to the RPB4 subunit of RNA polymerase II, is a component of RNA polymerases IV and V and is required for RNA-directed DNA methylation. *Genes Dev.* **23**, 318–330 (2009).
155. Daxinger, L. *et al.* A stepwise pathway for biogenesis of 24-nt secondary siRNAs and spreading of DNA methylation. *EMBO J.* **28**, 48–57 (2009).
156. Greenberg, M. V. *et al.* Identification of genes required for *de novo* DNA methylation in *Arabidopsis*. *Epigenetics* **6**, 344–354 (2011).
157. Kanno, T. *et al.* Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr. Biol.* **14**, 801–805 (2004).
158. Ausin, I., Mockler, T. C., Chory, J. & Jacobsen, S. E. IDN1 and IDN2 are required for *de novo* DNA methylation in *Arabidopsis thaliana*. *Nature Struct. Mol. Biol.* **16**, 1325–1327 (2009).
159. Kanno, T. *et al.* A structural-maintenance-of-chromosomes hinge domain-containing protein is required for RNA-directed DNA methylation. *Nature Genet.* **40**, 670–675 (2008).
160. Naumann, U. *et al.* Genetic evidence that DNA methyltransferase DRM2 has a direct catalytic role in RNA-directed DNA methylation in *Arabidopsis thaliana*. *Genetics* **187**, 977–979 (2011).

Acknowledgements
M.A.M. thanks Academia Sinica for financial support and C. Ying for editorial assistance. R.A.M. is supported by the US National Science Foundation under grant MCB-1243608. The authors apologize to colleagues whose publications are not cited owing to space limitations.

Competing interests statement
The authors declare no competing interests.

FURTHER INFORMATION
***Arabidopsis* epigenome maps:** http://neomorph.salk.edu/epigenome/epigenome.html
Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome: http://genomes.mcdb.ucla.edu/AthBSseq/
Epigenesys: http://www.epigenesys.eu/
Epigenomics of Plants International Consortium: https://www.plant-epigenome.org/

ALL LINKS ARE ACTIVE IN THE ONLINE PDF