HHS Public Access

Author manuscript

Biochim Biophys Acta. Author manuscript; available in PMC 2018 January 01.

Published in final edited form as:
Biochim Biophys Acta. 2017 January ; 1860(1): 140–148. doi:10.1016/j.bbagrm.2016.08.004.

The RNAs of RNA-directed DNA methylation

Jered M. Wendte[1] and **Craig S. Pikaard**[1,2,3]

[1]Department of Biology and Department of Molecular and Cellular Biochemistry, Indiana University, 915 E. Third Street, Bloomington, IN 47405, USA

[2]Howard Hughes Medical Institute, Indiana University, Bloomington, IN 47405, USA

Summary

RNA-directed chromatin modification that includes cytosine methylation silences transposable elements in both plants and mammals, contributing to genome defense and stability. In *Arabidopsis thaliana*, most RNA-directed DNA methylation (RdDM) is guided by small RNAs derived from double-stranded precursors synthesized at cytosine-methylated loci by nuclear multisubunit RNA Polymerase IV (Pol IV), in close partnership with the RNA-dependent RNA polymerase, RDR2. These small RNAs help keep transposons transcriptionally inactive. However, if transposons escape silencing, and are transcribed by multisubunit RNA polymerase II (Pol II), their mRNAs can be recognized and degraded, generating small RNAs that can also guide initial DNA methylation, thereby enabling subsequent Pol IV-RDR2 recruitment. In both pathways, the small RNAs find their target sites by interacting with longer noncoding RNAs synthesized by multisubunit RNA Polymerase V (Pol V). Despite a decade of progress, numerous questions remain concerning the initiation, synthesis, processing, size and features of the RNAs that drive RdDM. Here, we review recent insights, questions and controversies concerning RNAs produced by Pols IV and V, and their functions in RdDM. We also provide new data concerning Pol V transcript 5' and 3' ends.

Keywords

RNA silencing; epigenetic regulation; cytosine methylation; chromatin modification; gene regulation; noncoding RNA

1. Overview of RNA-directed DNA methylation (RdDM) and multisubunit RNA Polymerases IV and V

RNA-directed chromatin modification is used throughout eukaryotes as a means to prevent the transcription and movement of transposable elements, thereby protecting the genome from mutation and instability. In eukaryotes that do not methylate their DNA, such as fission

[3]Corresponding author: cpikaard@indiana.edu, phone: 812-272-9499.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

yeast, fruit flies or nematodes, small noncoding RNAs (RNAs that do not encode proteins) can guide histone modifications that help establish chromatin states refractive to transcription by RNA polymerases I, II or III. These same repressive histone modifications occur in species that utilize RNA to direct DNA methylation, including plants and humans, with cytosine methylation serving as an additional, important layer of regulation (reviewed in Law and Jacobsen 2010; He et al. 2011; Volpe and Martienssen 2011).

In plants, noncoding RNAs can direct the methylation of previously unmodified cytosines in any sequence context (CG, CHG, CHH; where H=A, C or T). The DNA methyltransferase enzyme that carries out this de novo cytosine methylation is DRM2. Importantly, DRM2 is an ortholog of human DNMT3a/b, reflecting their descent from a progenitor enzyme present in a common ancestor of plants and animals (Cao and Jacobsen 2002a; Cao and Jacobsen 2002b; Cao et al. 2003). Once cytosines have been methylated *de novo* by DNMT3 or DRM2, other DNA methyltransferases can maintain specific DNA methylation patterns following each round of DNA replication, and do so in an RNA-independent manner. For instance, CG methylation patterns, which are symmetric on the two strands of DNA, are maintained in mammals and plants (and other eukaryotes) by DNMT1 and MET1, respectively, which again are orthologous enzymes (Finnegan and Dennis 1993; Ronemus et al. 1996). Plants also encode plant-specific cytosine methyltransferases, named CMT3 and CMT2, which maintain symmetric CHG methylation (CMT3) or maintain asymmetric CHH methylation in regions of dense pericentric heterochromatin (CMT2) (Cao and Jacobsen 2002a; Cao et al. 2003; Zemach et al. 2013; Stroud et al. 2014). Proteins that recognize the different DNA methylation patterns help recruit chromatin modifying enzymes that chemically modify the associated histones. Likewise, proteins recognizing specific histone modifications can recruit cytosine methyltransferases (reviewed in Law and Jacobsen 2010). In this way, RNA biogenesis, DNA methylation and histone modification machineries work together to reinforce so-called epigenetic states, which can be defined as alternative states of gene activity that are not dictated by DNA sequence alone.

A common theme in eukaryotic RNA-directed chromatin modification is that small RNAs, such as siRNAs or piRNAs, bound to Argonaute family proteins, interact with longer chromatin-associated RNAs that act as scaffolds for assembly of chromatin modifying activities (reviewed in Holoch and Moazed 2015). In *Arabidopsis thaliana*, the model plant species in which RNA-directed DNA methylation (RdDM) is best understood, two plant-specific nuclear multisubunit RNA polymerases, Pol IV and Pol V, play major roles as generators of RNAs that program RdDM (reviewed in Haag and Pikaard 2011). Pol IV and Pol V are each composed of twelve subunits, and mass spectrometry analyses revealed that these subunits are either identical to, or paralogous to, the twelve subunits of Pol II (Ream et al. 2009) (Haag et al. 2014), indicating that Pols IV and V evolved as specialized forms of Pol II (see also He et al. 2009a; Huang et al. 2009; Lahmy et al. 2009). In Arabidopsis, Pols IV and V differ in three of their subunits, suggesting that these subunits might collectively make Pol IV and Pol V activities unique (Ream et al. 2009). However, subsequent mass spectrometry analyses in maize have shown that the only fundamental difference between Pol IV and Pol V is their use of different largest subunits (Haag et al. 2014). The catalytic centers of multisubunit RNA polymerases are formed by their two largest subunits (Cramer et al. 2008), and the second-largest subunits of Pols IV and V are the same. Thus, any

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

differences in Pol IV and Pol V activity are presumably explained by amino acid differences in their distinctive largest subunits.

The RNA products of Pols IV and V have distinct roles in separate steps of the RdDM pathway. In the conventional model, Pol IV initiates RdDM by producing precursors that are processed into small RNAs (Kanno et al. 2005b; Pontier et al. 2005). Pol IV physically associates with RNA-DEPENDENT RNA POLYMERASE 2 (RDR2) (Law et al. 2011; Haag et al. 2012). Together, these enzymes make double-stranded RNAs that are diced into 24 nt siRNAs by DICER-LIKE 3 (DCL3) (Xie et al. 2004), 2'-O-methylated at their 3' ends by HEN1 (Li et al. 2005) (which stabilizes the RNAs by preventing their uridylation and turnover), and loaded into an Argonaute protein, primarily AGO4 (Zilberman et al. 2003) or AGO6 (Zheng et al. 2007). Pol V then works downstream in the pathway by producing RNA scaffolds that recruit AGO4-siRNA complexes to the chromatin target sites (Wierzbicki et al. 2008; Wierzbicki et al. 2009). AGO4 interaction with Pol V transcripts subsequently recruits the *de novo* methyltransferase, DRM2 (Gao et al. 2010; Bohmdorfer et al. 2014; Zhong et al. 2014). As summarized in the model of Figure 1, resulting DRM2-mediated *de novo* cytosine methylation of the adjacent DNA, accompanied by histone modifications that include histone deacetylation (by HDA6) (Aufsatz et al. 2002), histone H3 lysine 9 methylation (by SUVH4, SUVH5 and SUVH6) (Jackson et al. 2002; Jackson et al. 2004; Ebbs et al. 2005; Ebbs and Bender 2006) and histone H3 lysine 4 demethylation (by JMJ14) (Deleris et al. 2010; Searle et al. 2010) all ensue, resulting in a chromatin environment refractive to conventional gene transcription by RNA Polymerases I, II or III (the RdDM pathway has been reviewed in Haag and Pikaard 2011; Matzke and Mosher 2014) (see Figure 1).

Thus, a major difference between RNA-directed chromatin modification in plants compared to other eukaryotes is that both the siRNA precursors and scaffold RNAs are produced by specialized polymerases rather than the ubiquitous polymerase, Pol II. Moreover, Pol II is an essential enzyme whereas Pols IV and V are dispensable for viability (at least in Arabidopsis), making plants valuable model systems in which to study the specific roles of multisubunit RNA polymerases and their RNA products in the epigenetic regulation of gene expression. In this review, we discuss recent advances in our understanding of Pol IV and V transcriptional regulation and transcript processing and function, as well as challenges that lie ahead in our quest to understand the RNAs of RdDM.

2. Pol IV and V Transcript Function

Pol IV transcript function: precursors for siRNAs

The requirement of Pol IV for the accumulation of 24nt siRNAs *in vivo* was one of the first phenotypes identified in the initial characterization of this enzyme (Herr et al. 2005; Onodera et al. 2005). However, the Pol IV-dependent precursors that give rise to siRNAs have only been described recently, based on their accumulation when DICERLIKE 3 (DCL3), the endonuclease primarily responsible for 24nt siRNA processing, is mutated (Blevins et al. 2015; Li et al. 2015; Zhai et al. 2015; Yang et al. 2016; Ye et al. 2016). Mutations in DCL2 and DCL4 further increase the abundance of these RNAs by blocking alternative routes of dicing into 21 or 22nt siRNAs. Incubating RNAs that accumulate in *dcl3* mutants with purified DCL3 *in vitro* results in cleavage of the RNAs into 24nt RNA

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

products, consistent with the longer RNAs being direct precursors of 24nt siRNAs (Blevins et al. 2015). Moreover, the sequences of 24nt siRNAs tend to perfectly match the sequences of precursor RNAs beginning at either the 5' ends or 3' ends of the precursors, suggesting that individual precursors give rise to siRNAs by single DCL3 cleavage events measured from either end (Blevins et al. 2015; Zhai et al. 2015).

Pol IV-dependent precursor RNAs that accumulate in *dcl3* mutants are also fully dependent on RDR2, thus we refer to them as Pol IV-RDR2 transcripts (abbreviated as P4R2 RNAs) (Blevins et al. 2015; Li et al. 2015; Zhai et al. 2015). Notably, however, Pol IV affinity-purified from *rdr2* null mutant plants is transcriptionally active in vitro (Haag et al. 2012). Likewise, recombinant RDR2 is transcriptionally active *in vitro* (Mishra and Pikaard, unpublished), showing that Pol IV-RDR2 interaction is not obligatory for the fundamental catalytic activities of either enzyme. Therefore, it is currently unclear why both enzymes are required for making either strand of siRNA precursors *in vivo*.

One of the most intriguing recent findings came from whole-genome bisulfite sequencing experiments that indicate that P4R2 RNAs can guide RNA-directed DNA methylation without being diced into siRNAs (Yang et al. 2016; Ye et al. 2016). This deduction stems from the detection of Pol IV and RDR2-dependent cytosine methylation that persists in quadruple dicer mutants that are null for *dcl2, dcl3 and dcl4*, and hypomorphic for *dcl1* (a hypomorphic mutation was needed due to the lethality of *dcl1* null mutants). This raises the question of whether siRNAs are truly needed for RdDM, a provocative idea.

Pol V transcript function: scaffolds for recruitment of chromatin modifying complexes

Pol V-dependent transcripts were first identified through a simple, trial and error approach that consisted of identifying RT-PCR products that were detected in wild type plants but absent in Pol V mutants. These non-coding transcripts were found to be derived from regions of the genome associated with siRNAs and DNA methylation, and were shown to extend from intergenic regions into the promoter regions of adjacent transposons, correlating with the silencing of these promoters. Importantly, the intergenic (IGN) transcripts could be cross-linked to Pol V, were generated at loci occupied by Pol V, and were no longer generated if the Pol V active site was mutated (Wierzbicki et al. 2008). Collectively, these lines of evidence indicated that IGN RNAs are the direct transcription products of Pol V.

The major functions of Pol V transcripts are to serve as scaffolds for recruitment of proteins to chromatin. As noted above (Figure 1), siRNA-AGO complexes find their target sites as a consequence of Pol V transcription (Wierzbicki et al. 2009; Zheng et al. 2013). This process is best understood in the context of AGO4-medaited RdDM. AGO4 association with chromatin is lost if the catalytic site of the Pol V largest subunit (NRPE1) is mutated, preventing Pol V transcription but not Pol V subunit assembly (Wierzbicki et al. 2009). These observations suggest that siRNA-Pol V transcript basepairing is the primary mode of AGO4 recruitment to target sites. However, protein-protein interactions between AGO4 and the C-terminal domain (CTD) of the Pol V largest subunit (El-Shami et al. 2007; He et al. 2009b) and between AGO4 and SPT5L/KTF1 (a paralog of the Pol II elongation factor SPT5 that binds Pol V and Pol V transcripts, as a heterodimer with SPT4) likely generate a

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

metastable multimeric complex (Bies-Etheve et al. 2009; He et al. 2009b; Rowley et al. 2011; Kollen et al. 2015).

An additional protein that binds Pol V scaffold transcripts and helps mediate RdDM is the RNA binding protein, IDN2 (Bohmdorfer et al. 2014). IDN2 is related to SGS3, an RNA binding protein involved in dsRNA synthesis by RDR6, which generates precursors for DCL4-mediated 21 nt siRNA production (Mourrain et al. 2000). However, IDN2 is not involved in 24 nt siRNA production, yet is required for DNA methylation at a subset of RdDM target loci (Ausin et al. 2009; Zheng et al. 2010). IDN2 preferentially binds dsRNA molecules with 5' overhangs *in vitro* (Ausin et al. 2009), but the significance of this activity *in vivo* is unclear. Importantly, IDN2 forms a complex with the paralogous proteins, IDP1 (also known as IDNL1 or FDM1) and IDP2 (also known as IDNL2 or FDM2,) mutants of which have RdDM phenotypes partially redundant with those of IDN2 (Ausin et al. 2012; Xie et al. 2012a; Xie et al. 2012b; Zhang et al. 2012).

IDN2's binding to Pol V transcripts *in vivo* occurs in an AGO4-dependent manner (Zhu et al. 2013; Bohmdorfer et al. 2014). This binding event is subsequently required for the direct or indirect association of DRM2 with the Pol V scaffold transcripts, providing a mechanistic link between IDN2-RNA interactions and DNA methylation at loci that are dependent on IDN2 (Ausin et al. 2012; Bohmdorfer et al. 2014). Exactly how IDN2 is recruited downstream of AGO4 and how DRM2 is recruited downstream of IDN2 is unclear, given that there appears to be no direct physical interaction between IDN2 and AGO4 or DRM2 (Ausin et al. 2012; Xie et al. 2012b; Zhang et al. 2012). But at least one of the IDN2 interacting proteins, IDP1 (IDNL1/FDM1), has been shown to be capable of interacting with DNA as well as RNA (Xie et al. 2012b), suggesting that the IDN2 complex might bind to both Pol V scaffold transcripts and their DNA templates, possibly stabilizing AGO4-DRM2 effector complexes and promoting DNA methylation (Ausin et al. 2012; Bohmdorfer et al. 2014). *In vitro* reconstitution or structural analyses of complexes involving RNA, IDN2, AGO4, and DRM2 would no doubt be enlightening.

In addition to being required for recruitment of DRM2, IDN2 interactions with Pol V transcripts are also required for recruitment of the SWI-SNF chromatin remodeling complexes to regions via direct physical interaction with the SWI/SNF subunit, SWI3b, which stabilizes nucleosomes at RdDM target loci (Zhu et al. 2013). The function of nucleosome stabilization is not entirely clear but has been proposed to provide DRM2 with stably positioned target cytosines around the nucleosome (Zhu et al. 2013). Alternatively, nucleosome stabilization may be related to IDN2's interaction with the MORC6/DMS11 ATPase protein (which also interacts with SWI3b) that functions in transcriptional silencing downstream of DNA methylation (Liu et al. 2016).

Other Argonaute proteins in the cell, including AGO2, AGO6, and AGO9, have also been shown to bind siRNAs and participate in RdDM (Havecker et al. 2010; Olmedo-Monfil et al. 2010; Pontier et al. 2012). These AGO-siRNA complexes are presumably recruited to chromatin through interactions with Pol V transcripts in a manner similar to AGO4. Both AGO6 and AGO9 transgenes, when placed under the native AGO4 promoter, can partially rescue *ago4* mutants (Havecker et al. 2010). Moreover, AGO2 and AGO6 have been shown

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

to be required for DNA methylation at loci that also require Pol V (Zheng et al. 2007; Havecker et al. 2010; Eun et al. 2011; Pontier et al. 2012; Nuthikattu et al. 2013; Duan et al. 2015; McCue et al. 2015) and are capable of interacting with the AGO-hook motifs of the Pol V largest subunit's C-terminal domain (Pontier et al. 2012). Additionally, AGO6 recruitment to chromatin has been shown to be dependent on Pol V both in the context of Pol IV-derived, 24nt siRNA-mediated RdDM and RDR6-derived, 21–22nt siRNA-mediated RdDM (McCue et al. 2015). These data indicate that Pol V transcription can be used to recruit diverse, specialized AGO-siRNA complexes, and not necessarily for RdDM. For instance, AGO2 recruitment by Pol V is implicated in DNA double-strand break (DSB) repair because DSBs are associated with production of 21 and 24nt siRNAs bound by AGO2, and both AGO2 and Pol V are required for efficient DSB repair (Wei et al. 2012). However, it is not yet clear if Pol V transcribes DSB loci to produce scaffold RNAs to which AGO2 complexes bind. Nonetheless, the possibility that small RNA- Pol V scaffold RNA interactions guide a variety of processes in the nucleus is intriguing.

3. How do Pols IV and V know where to transcribe?

Given the importance of Pols IV and V in RNA-directed DNA methylation and gene silencing, a critical question is: how do these polymerases target specific loci? DNA sequences associated with Pol IV and Pol V have been identified genome-wide by chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) (Wierzbicki et al. 2012; Zhong et al. 2012; Law et al. 2013). No consensus sequences that are highly correlated with Pol IV or Pol V-associated regions have been identified. This suggests that Pol IV and Pol V recruitment may not involve conventional promoter sequences that are recognized by sequence specific DNA-binding proteins, although use of diverse, or sequence-tolerant, promoter elements could make such sequences difficult to identify. Instead, evidence has accumulated that chromatin marks play an important role in specifying sites of Pol IV and Pol V transcription. Pol IV and Pol V are both recruited to sites modified by HDA6-dependent histone deacetylation, maintenance methylation (by MET1 and/or CMT3), and histone H3K9 dimethylation (H3K9me2), a histone modification known to occur in crosstalk with maintenance cytosine methylation (Law et al. 2011; Law et al. 2013; Zhang et al. 2013; Blevins et al. 2014; Johnson et al. 2014). This realization has come from experimental evidence generated by multiple laboratories. For instance, in studies investigating the role of HDA6 in RdDM, Blevins et al. found a surprising requirement for HDA6 in Pol IV recruitment and 24nucleotide siRNA biogenesis at many loci (Blevins et al. 2014). Interestingly, siRNA biogenesis lost in *hda6* mutants is not regained at these loci simply by restoring HDA6 activity, indicating that a heritable epigenetic memory required for Pol IV recruitment is erased in *hda6* mutants, and is not easily regained once lost. Knowing that HDA6 is required for CG and CHG maintenance methylation at rRNA genes (Earley et al. 2010), as well as other loci, *met1* null mutants were tested and found to have the same loss of siRNA biogenesis and loss of epigenetic memory observed in *hda6* mutants. Thus, MET1-dependent maintenance methylation can account for the heritable epigenetic memory that marks loci for Pol IV recruitment and silencing by RdDM, establishing a chromatin state responsible for a region's "silent locus identity" (Blevins et al. 2014).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In independent studies, the Jacobsen and Zhu labs found that a protein that physically interacts with Pol IV, SHH1/DTF1, binds histones displaying the repressive H3K9me2 mark, and Pol IV association with many of its sites of action genome-wide are lost in *shh1* mutants (Law et al. 2011; Law et al. 2013; Zhang et al. 2013). It has long been known that a subset of heterochromatin regions displaying MET1-dependent cytosine methylation are enriched for H3K9me2, a histone mark that can be lost when mutation of *met1* leads to transcriptional derepression (Johnson et al. 2002; Soppe et al. 2002). Mechanistically, the relationship between H3K9 methylation and cytosine methylation is best understood in relation to CHG methylation, which is the sequence context that is specifically bound by the SUVH4 histone methyltransferase (Johnson et al. 2007). Thus it is possible that *met1* effects on H3K9 methylation are related to decreases in non-CG rather than CG methylation—a hypothesis that has been supported by genome-wide cytosine methylation analyses (Stroud et al. 2014). H3K9 methylation is also highly correlated with maintenance methylation established by CMT3, which methylates DNA in the CHG context and specifically binds to methylated H3K9, thus forming a positive-feedback loop with SUVH4 (Du et al. 2012; Stroud et al. 2014). Mutations in *cmt3* also cause a loss of H3K9 methylation and Pol IV-SHH1 dependent siRNAs (Stroud et al. 2014), but whether Pol IV occupancy fails to be regained upon re-introduction of a wild type CMT3 gene, similar to MET1 or HDA6, remains to be tested. Collectively, the observations of the different laboratories fit the hypothesis that maintenance DNA methylation allows for the inheritance of an epigenetic signal that is translated into H3K9 dimethylation, providing a docking site for SHH1-Pol IV complexes (see Figure 1). Consistent with this hypothesis, RdDM loci dependent on HDA6, MET1 and SHH1 significantly overlap (Blevins et al. 2014).

Pol V recruitment also requires maintenance methylation by MET1 (Johnson et al. 2014). However, in this case, the methylated DNA is recognized directly by the methylcytosine binding proteins, SUVH2 and SUVH9 (Johnson et al. 2008; Johnson et al. 2014; Liu et al. 2014; Jing et al. 2016; Liu et al. 2016). These proteins are members of the histone methyltransferase family that have SRA domains for binding methylated DNA as well as domains for methylating histone H3 within associated nucleosomes. Interestingly, amino acid changes have rendered SUVH2 and SUVH9 non-functional as histone methyltransferases, yet they retain their ability to bind methylated cytosines (Johnson et al. 2014). Pol V recruitment to its sites of action is impaired in *suvh2/9* mutants—an interaction that is presumably mediated by bridging proteins. In the studies that first identified Pol V transcripts (Wierzbicki et al. 2008; Wierzbicki et al. 2009), production of the RNAs was shown to require DRD1, a putative ATP-dependent DNA translocase (Kanno et al. 2005a), and DMS3, a protein related to the hinge domains of cohesins and condensins (Kanno et al. 2008). DRD1 and DMS3 were subsequently shown to associate with one another and with a single-stranded DNA binding protein, RDM1 to form the so-called DDR complex (Law et al. 2010; Lorkovic et al. 2012). A protein that interacts with the DDR complex is the MORC ATPase, DMS11 (Lorkovic et al. 2012), also known as AtMORC6. For DRD1, DMS3, and DMS11/MORC6, there is evidence for direct physical interaction with SUVH2 and/or SUVH9 *in vivo* (Johnson et al. 2014; Liu et al. 2014). Collectively, these observations suggest that methylated DNA is recognized by SUVH2 or SUVH9, recruiting the DDR-DMS11/MORC6 complex that facilitates Pol V transcription, via mechanisms that remain

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

unclear. We have speculated that the DDR complex, and its associated proteins, may play a role in unwinding DNA to facilitate Pol V transcription, analogous to the helix unwinding that occurs at DNA replication forks (Pikaard et al. 2013). However, direct biochemical tests of this hypothesis are still needed.

Although the majority of RdDM is mediated by the 24 nt siRNAs that account for approximately 90% of all siRNAs present in Arabidopsis cells, an important role for 21–22 nt siRNAs in the early establishment of RdDM has come to light in recent years (Sigman and Slotkin 2016). Transposon-derived transcripts synthesized by Pol II are recognized by the cell, in part by cellular microRNAs (miRNAs) that basepair with the transposon RNAs in association with AGO1, which then slices the mRNAs (Creasey et al. 2014). Via mechanisms that remain biochemically undefined, sliced mRNAs become templates for dsRNA synthesis by the RNA-dependent RNA polymerase, RDR6, and these dsRNAs are then cut, primarily by DCL4, into 21 nt secondary siRNAs which can further direct mRNA slicing in association with AGO1. In this way, messenger RNAs encoding transposon proteins are degraded, thus curtailing transposon activity. However, genetic and genomic evidence also indicate that RDR6-dependent siRNAs can associate with AGO6 and direct *de novo* cytosine methylation of targeted transposons' DNA (Nuthikattu et al. 2013; Panda and Slotkin 2013; McCue et al. 2015). In this way, methylation marks that can potentiate subsequent Pol IV and Pol IV recruitment can be established. The ensuing Pol IV-RDR2 and Pol V transcription, and 24 nt siRNA-mediated RdDM, can then maintain silencing.

4. P4R2 and Pol V transcript characteristics

Because Pols IV and V are evolutionarily derivatives of Pol II, might their transcripts undergo processing like Pol II derived mRNAs? Major processing steps for Pol II transcripts include the addition of a 7- methylguanosine cap on the 5' end, a 3' poly-A tail, and splicing of intronic sequences. All of these modifications are mediated by protein-protein interactions between Pol II and processing enzymes. Of particular importance is the C-terminal domain (CTD) of the Pol II largest subunit, which interacts with components of the splicing, capping, and cleavage/polyadenylation complexes, allowing co-transcriptional processing (reviewed in Hocine et al. 2010; Hsin and Manley 2012; Bentley 2014). Other Pol II domains can also interact with RNA processing factors, including the foot domain of the largest subunit, which interacts with the capping complex (Suh et al. 2010), and the flap loop domain of the second-largest subunit, which interacts with both the cleavage/ polyadenylation and capping machineries (Pearson and Moore 2014; Martinez-Rucobo et al. 2015). Intriguingly, the Pol IV and Pol V largest subunits, NRPD1 and NRPE1, have CTDs whose sequences have completely diverged from the CTD of the Pol II largest subunit, NRPB1 (Pontier et al. 2005; Haag and Pikaard 2011; Huang et al. 2015; Trujillo et al. 2016). Moreover, NRPD1 and NRPE1 are missing the foot domain (Pontier et al. 2005; Luo and Hall 2007). Likewise, the second largest subunit, NRPD/E2, which is shared by Pols IV and V, is missing the flap loop domain. The flap loop domain is also absent in Pols I and III, which has been suggested to at least partially explain why Pol I and III transcripts do not undergo capping (Figure 2A) (Martinez-Rucobo et al. 2015). Collectively, the amino acid sequence divergence of Pol IV and V subunits, compared to Pol II subunits, suggests that

Author Manuscript

these polymerases may have lost the ability to interact with the RNA processing factors that modify Pol II transcripts.

P4R2 RNAs are relatively short, double-stranded RNAs typically ranging in size from ~26–45 bp, with the peak of their size distribution occurring at approximately 30 bp (Blevins et al. 2015; Zhai et al. 2015; Yang et al. 2016; Ye et al. 2016). This short size likely precludes splicing. Furthermore, consideration of P4R2 RNA clusters at the loci from which they are derived revealed no evidence for discontinuities consistent with splicing of longer precursors (Li et al. 2015).

Currently, our knowledge of Pol V transcription units is limited to regions that have been amplified by reverse transcription and PCR (RT-PCR), generally limited to intervals of ~200bp or less (Wierzbicki et al. 2008; Wierzbicki et al. 2009; Rowley et al. 2011; Wierzbicki et al. 2012; Zhong et al. 2012; Zheng et al. 2013; Zhu et al. 2013; Bohmdorfer et al. 2014; Zhang et al. 2014). Thus far, RT-PCR product sizes have matched the sizes predicted from the DNA template, with no unexpected smaller products attributable to splicing. Clearly, this does not rule out the possibility of splicing, but at present there is no evidence for it.

The 3’ ends of P4R2 RNAs are not modified with a poly-A tail, but often contain one or two untemplated nucleotides (Blevins et al. 2015; Li et al. 2015; Zhai et al. 2015; Yang et al. 2016). Different opinions exist as to how the 3’ ends of P4R2 RNAs are generated, and why the RNAs are so short (~30 bp). One hypothesis is that cytosine methylation in the DNA template can cause misincorporation of a nucleotide other than G in the RNA, inducing Pol IV termination and thus limiting the size of Pol IV transcripts (Zhai et al. 2015). However, Pol IV transcripts generated *in vitro* using unmethylated single-stranded bacteriophage M13 DNA as the template, are similar in size to P4R2 RNAs made *in vivo*, albeit slightly longer (Blevins et al. 2015). This suggests that the short size of P4R2 RNAs reflects an intrinsically poor processivity of Pol IV compared to other known RNA polymerases, even on non-methylated DNA. Moreover, RDR2 has terminal transferase activity that will add one or two non-templated nucleotides to the ends of an RNA molecule, suggesting an alternative explanation for non-templated nucleotides at the ends of P4R2 RNAs, as opposed to Pol IV mis-incorporation (Blevins et al. 2015). Another recent paper has suggested that the 3’ ends of P4R2 RNAs might be produced by sequential exonuclease action that removes one nucleotide at a time from longer transcripts, resulting in P4R2 RNAs that can have the same 5’ end, but variable 3’ ends (Ye et al. 2016). This hypothesis stems from the observation that defects in RdDM are observed in mutants for RRP6-LIKE 1 (RRP6L1), an ortholog of a yeast enzyme known to be a 3’ exoribonuclease. However, as the authors themselves point out, there is no direct evidence that RRP6L1 trims P4R2 RNAs, such that the effects of RRP6L1 on RdDM may not be related to small RNA biosynthesis. Because Pol IV transcripts generated *in vitro* have variable 3’ end positions (Blevins et al. 2015), random termination rather than sequential exonuclease action may provide an alternative explanation for 3’ ends that can differ by increments of single nucleotides.

Pol V transcript 3’ ends have been characterized, to a limited extent, relative to Pol II-derived mRNAs (Wierzbicki et al. 2008). Upon fractionation of RNA into poly-A enriched

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and poly-A depleted fractions, Pol V IGN5 transcripts were detected only in the poly-A depleted fraction, suggesting that Pol V transcripts lack a 3' poly-A tail (Wierzbicki et al. 2008). An expanded examination of Pol V transcripts from other loci indicates that the lack of a poly-A tail is a general characteristic of these RNAs (Figure 2B).

P4R2 RNAs tend have a purine (A or G) at their 5' ends (the +1 position), with a pyrimidine (C or T) present in the DNA at −1 (Blevins et al. 2015; Zhai et al. 2015). This is a conserved signature of transcripts generated by all known multisubunit RNA polymerases, from bacteria to archaea to eukaryotes (Basu et al. 2014), thereby implicating Pol IV rather than RDR2. However, as first noted by Li et al., the P4R2 RNAs lack triphosphate groups at their 5' ends, as is expected for an initiating nucleotide; instead the RNAs have 5' monophosphates (Li et al. 2015). This raises the possibility that 5' ends of P4R2 RNAs might be generated by processing, perhaps involving an endonuclease that cleaves RNA between adjacent pyrimidine-purine motifs. However, transcripts generated *in vitro* using affinity purified Pol IV (isolated in an *rdr2* mutant background) from long single-stranded DNA templates also obey the −1 pyrimidine/ +1 purine rule at their 5' ends (Blevins et al. 2015) arguing against P4R2 end consensus sequences arising through processing. Instead, an unknown activity that removes a terminal pyrophosphate group of P4R2 RNAs is implicated.

Pol V transcripts at the IGN5 locus were found to have purines at their 5' ends (Wierzbicki et al. 2008). Furthermore, full-length Pol V IGN5 transcripts were deduced to have either 5' caps or triphosphate moities (Wierzbicki et al. 2008), but tests to distinguish between these possibilities were not conducted. Thus, we have revisited this question by using three enzyme treatment schemes that can distinguish between a 5' cap, a 5' triphosphate, or a 5' monophosphate (Figure 2C). Treatment A consisted of an initial treatment with Tobacco Acid Pyrophosphatase (TAP), which will convert 5' caps or 5' triphosphate groups to 5' monophosphates, followed by treatment with Terminator exonuclease, which specifically degrades RNAs that have 5' monophosphates. In Treatment B, an initial treatment with 5' polyphosphatase, which converts 5' triphosphates into monophosphates, but has no effect on 5' caps, was followed by treatment with Terminator exonuclease. Treatment C consisted of treatment with Terminator exonuclease only (Figure 2C). As a control, an equal amount of total RNA was also subjected to each treatment step, but with no enzyme added. After the treatments, transcript levels were determined using quantitative RT-PCR and ratios of transcript levels in treated versus untreated samples were compared for treatments A, B, or C.

We predicted that a transcript with a 5' cap should be most severely depleted by Treatment A but would be relatively resistant to Treatments B and C. By contrast, transcripts with 5' triphosphate groups should be equally depleted by Treatments A and B, but resistant to C. Transcripts with a 5' monophosphate group should be equally depleted in all treatments. These predictions hold true for actin and tubulin (ACT2 and TUB8) mRNA controls, transcribed by Pol II, which were significantly more depleted by Treatment A than by Treatments B or C, consistent with these mRNAs bearing 7-methylG caps (Figure 2D). By contrast, a Pol I-transcribed control transcript, 45S pre-rRNA, which is expected to have a 5' triphosphate (Batts-Young and Lodish 1978), was equally depleted by Treatments A and B, as expected for RNAs with triphosphate groups at their 5' ends, and was less depleted by

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Treatment C (Figure 2D). Pol V transcribed RNAs respond similar to the Pol I 45S pre-rRNA control, being equally depleted by Treatments A and B, and less depleted by Treatment C (Figure 2D). This indicates that the majority of Pol V transcripts have a 5' triphosphate, as expected of a primary transcript, and are not modified by addition of a 7-methyguanosine cap. A portion of the Pol V transcripts (averaging ~30%) tested have 5' monophosphate groups (Figure 2D), presumably generated by cleavage or partial degradation.

Collectively, the evidence suggests that P4R2 and Pol V transcripts do not undergo RNA processing steps characteristic of Pol II derived mRNAs, namely capping, splicing or polyadenylation. This is consistent with the amino acid changes, relative to Pol II, in domains that interact with processing enzymes, as discussed previously.

6. Outstanding questions for Pol IV and Pol V dependent RNAs

As noted above, Pol IV and RDR2 are both capable of transcription independently of one another *in vitro*, yet both are equally required for the accumulation of P4R2 RNAs *in vivo*. Thus, a major question remaining for P4R2 RNAs is the extent to which Pol IV and RDR2 each contribute to the population of precursor RNAs present *in vivo*. The presence of a 5' monophosphate on P4R2 RNAs strongly suggests there are additional steps in P4R2 RNA processing that are yet to be discovered. Indeed, a recent study described a role for RNaseIII-like enzyme, RTL2, in cleaving a subset of P4R2 RNAs prior to processing by Dicers (Elvira-Matelot et al. 2016). Identification of additional P4R2 RNA processing factors may allow for an experimental uncoupling of Pol IV and RDR2 *in vivo* and provide insight to their individual roles.

Since their initial discovery and characterization (Wierzbicki et al. 2008), the number of known Pol V-transcribed loci has increased substantially, due largely to genome-wide Pol V localization by ChIP-seq (Wierzbicki et al. 2012; Zhong et al. 2012). However, only tens of Pol V-produced RNAs have been confirmed by RT-PCR and no full-length primary transcripts of Pol V have yet been described. An understanding of Pol V transcription start and termination sites, thus defining the lengths of Pol V transcripts and the sequences that flank them, is sorely needed. Such information would also assist efforts to determine the fates of Pol V transcripts following transcription, including their potential cleavage and degradation.

There are currently only two proteins with potential ribonuclease activity that have been identified as interactors with Pol V transcripts, AGO4 and RRP6L1 (Wierzbicki et al. 2009; Zhang et al. 2014). AGO4 is capable of slicing RNA *in vitro*, and this activity is at least partially required for RdDM *in vivo* (Qi et al. 2006). However, it is not clear that AGO4 slices Pol V transcripts, nor is it clear what the consequences of Pol V transcript slicing might be. An initial hypothesis was that Pol V transcript slicing may contribute to Pol V dependent siRNA production (Qi et al. 2006). However, recent studies in yeast have shown that, at least in *S. pombe*, ago slicer activity is dispensable for secondary siRNA amplification (Jain et al. 2016). RRP6L1 is a putative 3' to 5' exoribonuclease related to the yeast nuclear exosome component, Rrp6 (Lange et al. 2008), which one might think would

Author Manuscript

make it an excellent candidate for an enzyme involved in Pol V transcript degradation and turnover. However, contrary to this hypothesis, a recent study has demonstrated that Pol V transcripts, are less stable in *rrp6L1* mutants, not more stable (Zhang et al. 2014).

Mass spectrometry analyses in maize indicate that the distinct largest subunits of Pols IV and V are the only fundamental difference between the two enzymes (Haag et al. 2014). The major difference between the largest subunits of Pols IV and V is the presence of a long C-terminal domain (CTD) in the Pol V largest subunit and a short CTD in the Pol IV largest subunit. The Pol II CTD plays a critical role in all aspects of Pol II transcription (Hocine et al. 2010; Hsin and Manley 2012; Bentley 2014). Although the long Pol V CTD has no amino acid sequence similarity with the Pol II CTD , structural features reminiscent of the Pol II CTD are present, including a (predicted) unstructured series of peptide repeats and multiple amino acid residues that could potentially be subjected to post-translational modification, such as phosphorylation (Pontier et al. 2005; Haag and Pikaard 2011; Huang et al. 2015; Trujillo et al. 2016). Targeted investigations of Pol IV and Pol V CTD domains to explore their roles in RdDM and Pol V transcription, their post-translational modification, or their interactions with other proteins seems likely to provide new insights into the unique functions of these fascinating and enigmatic enzymes.

Methods

RNA extraction and RT-PCR

Total RNA was extracted from 2–2.5 week old above-ground tissue of Col-0 WT or *nrpe1–11* (Pontes et al. 2006) using Trizol and RT-PCR was conducted as described in (Wierzbicki et al. 2008). Primers for IGN22 (RT: 5’ CGGGTCCTTGGACTCCTGAT 3’; PCR: 5’ TCGTGACCGGAATAATTAAATGG 3’), IGN23 (RT: 5’ GCCATTAGTTTTAGATGGACTGCAA 3’; PCR: 5’ GGGCGAACCTGGAGAAAGTT 3’), IGN25 (RT: 5’ CTTCTTATCGTGTTACATTGAGAACTCTTTCC 3’; PCR: 5’ ATTCGTGTGGGCTTGGCCTCTT 3’), and IGN26 (RT: 5’ CGTGACATTAGAAGCTCTACGAGAA 3’; PCR: 5’ TTCCTGGCCGTTGATTGGT 3’) are from (Rowley et al. 2011). Primers for IGN24 (RT: 5’ CGCATACGATGGTCGGAGAGTT 3’; PCR: 5’ GCTTATCATTATCCAAACTTGATCCTATCCTAAA 3’) are from (Wierzbicki et al. 2012). Primers for IGN29 (RT: 5’ CATGTGTTGTTGTGTGTTTCACTAT 3’; PCR: 5’ TAAAACTTTTCCCGCCAACCA 3’) are from this study and (Zhu et al. 2013). Primers for IGN34 (RT: 5’ CCCTTTCATCGACACTGACA 3’; PCR: 5’ ATGAATAACAAATTTGGAGTCGTC 3’), IGN35 (RT: 5’ GACGGACCAAACGATTTCAT 3’; PCR: 5’ TTCCTCTTTGAGCTTGACCA 3’), IGN36 (RT: 5’ CAGTTTTGGGTGCGGTTTAT 3’; PCR: 5’ GACAAAAATTGCTTTAGACCATGA 3’) are from (Bohmdorfer et al. 2014). Primers for 45S rRNA (RT: 5’ CAAGCAAGCCCATTCTCCTC 3’; PCR: 5’ GAGTCTGGGCAGTCCGTGG 3’) are from (Chandrasekhara et al. 2016). Primers for ACT2 (RT: 5’ CTGTACTTCCTTTCAGGTGGT 3’; PCR: 5’ GCTGACCGTATGAGCAAAGA 3’) and TUB8 (RT: 5’ TCGAATCGATCCCGTGCT 3’; PCR: 5’ TCACATACAAGGTGGCCAATG 3’) are from this study.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Poly-A fractionation and 5' ends analysis of RNA

RNA was fractionated into Poly-A enriched and Poly-A depleted fractions using the Sigma FastTrack MAG mRNA Isolation Kit according to the manufacturer's instructions. RNA was precipitated from the Poly-A depleted fraction using 2 volumes 100% ethanol and 1/10th volume sodium acetate (pH 5.2). RT-PCR products from each fraction were quantified by running 10ul on a 2% agarose gel stained with ethidium bromide, followed by measuring band intensities with Quantity One software. Ratios of band intensities in the Poly-A enriched and Poly-A depleted fractions were calculated for each target transcript across at least 3 biologic replicates.

To characterize 5' ends, RNA was subjected to three enzyme treatment schemes. To target RNAs with a 5' cap, 5' triphosphate, or 5' monophosphate group RNA was first treated with Tobacco Acid Pyrophosphatase, followed by a treatment with Terminator exonuclease. To target RNAs with a 5' triphosphate or 5' monophosphate, RNA was treated with 5' polyphosphatase, followed by a treatment with Terminator exonuclease. To target only 5' monophosphate RNAs, RNA was treated with Terminator exonuclease only. All enzymes were obtained from Epicentre and treatments were conducted according to the manufacturer's instructions. After each treatment step, RNA was cleaned using the Zymogen RNA Clean and Concentrate Kit. As a control RNA was subjected to each treatment step with no enzyme added. For each treatment scheme, RT-PCR products were quantified by running 10ul on a 2% agarose gel stained with ethidium bromide, followed by measuring band intensities with Quantity One software. Ratios of band intensities in the treatment (+) enzyme and treatment (-) enzyme fractions were calculated for each target transcript across at least 3 biologic replicates for each treatment scheme.

Amino Acid sequence alignments

Amino acid sequences for the second largest subunits of Pols I, II, III, and IV/V for each of the species listed in Figure 2 were obtained from public databases and (Huang et al. 2015). Sequences were aligned using T-Coffee (Notredame et al. 2000).

Acknowledgments

Research in the Pikaard lab is supported by National Institutes of Health grant GM077590 (to C.S.P.) and support to C.S.P. as an Investigator of the Howard Hughes Medical Institute and the Gordon and Betty Moore Foundation. JMW received support from the NIH departmental training Grant, T32GM007757, and the National Institute of General Medical Sciences of the NIH under Award Number F31GM116346. The content of this work is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

List of abbreviations

AGO1	ARGONAUTE 1
AGO2	ARGONAUTE 2
AGO4	ARGONAUTE 4
AGO6	ARGONAUTE 6

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

CLSY1	CLASSY 1
CMT2	CHROMOMETHYLASE 2
CMT3	CHROMOMETHYLASE 3
CTD	carboxy-terminal domain
DCL1	DICER-LIKE 1
DCL2	DICER-LIKE 2
DCL3	DICER-LIKE 3
DCL4	DICER-LIKE 4
DDR	DRD1-DMS3-RDM1 complex
DMS11/ATMORC6	DEFECTIVE IN MERISTEM SILENCING 11/ MICRORCHIDIA 6
DMS3	DEFECTIVE IN MERISTEM SILENCING 3
DNMT1	DNA METHYLTRANSFERASE 1
DNMT3	DNA METHYLTRANSFERASE 3
DRD1	DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1
DRM2	DOMAINS REARRANGED METHYLTRANSFERASE 2
dsRNA	double stranded RNA
HDA6	HISTONE DEACETYLASE 6
HEN1	HUA ENHANCER 1
IDN2	INVOLVED IN DE NOVO 2
IDP1/IDNL1/FDM1	IDN2 PARALOG 1/INVOLVED IN DE NOVO 2-LIKE 1/ FACTOR OF DNA METHYLATION 1
IDP2/IDNL2/FDM2	IDN2 PARALOG 2/INVOLVED IN DE NOVO 2-LIKE 1/ FACTOR OF DNA METHYLATION 1
IGN	Intergenic non-coding RNA
JMJ14	JUMONJI 14
MET1	DNA METHYLTRANSFERASE 1
miRNA	micro RNA
NRPD1	NUCLEAR RNA POLYMERASE IV, subunit 1

NRPE1	NUCLEAR RNA POLYMERASE V, subunit 1
P4R2 RNA	Pol IV-RDR2 dependent RNA
piRNA	piwi-associated RNA
Pol IV	Plant NUCLEAR MULTISUBUNIT RNA POLYMERASE IV
Pol V	Plant NUCLEAR MULTISUBUNIT RNA POLYMERASE V
RdDM	RNA-directed DNA Methylation
RDM1	RNA DIRECTED DNA METHYLATION 1
RDR2	RNA-DEPENDENT RNA POLYMERASE 2
RDR6	RNA DEPENDENT RNA POLYMERASE 6
RRP6L1	RRP6-LIKE 1
siRNA	short interfering RNA
SPT4	SUPRESSOR OF TY4
SPT5L/KTF1	SUPRESSOR OF TY INSERTION 5-LIKE / KOW DOMAIN-CONTAINING TRANSCRIPTION FACTOR 1
SUVH2	SU(VAR)3–9 HOMOLOG 2
SUVH4	SU(VAR)3–9 HOMOLOG 4
SUVH5	SU(VAR)3–9 HOMOLOG 5
SUVH6	SU(VAR)3–9 HOMOLOG 6
SUVH9	SU(VAR)3–9 HOMOLOG 9
SWI3B	SWITCH SUBUNIT 3B

References

Aufsatz W, Mette MF, van der Winden J, Matzke M, Matzke AJ. HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA. EMBO J. 2002; 21:6832–6841. [PubMed: 12486004]

Ausin I, Greenberg MV, Simanshu DK, Hale CJ, Vashisht AA, Simon SA, Lee TF, Feng S, Espanola SD, Meyers BC, et al. INVOLVED IN DE NOVO 2-containing complex involved in RNA-directed DNA methylation in Arabidopsis. Proc Natl Acad Sci U S A. 2012; 109:8374–8381. [PubMed: 22592791]

Ausin I, Mockler TC, Chory J, Jacobsen SE. IDN1 and IDN2 are required for de novo DNA methylation in Arabidopsis thaliana. Nat Struct Mol Biol. 2009; 16:1325–1327. [PubMed: 19915591]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Basu RS, Warner BA, Molodtsov V, Pupov D, Esyunina D, Fernandez-Tornero C, Kulbachinskiy A, Murakami KS. Structural basis of transcription initiation by bacterial RNA polymerase holoenzyme. J Biol Chem. 2014; 289:24549–24559. [PubMed: 24973216]

Batts-Young B, Lodish HF. Triphosphate residues at the 5' ends of rRNA precursor and 5S RNA from Dictyostelium discoideum. Proc Natl Acad Sci U S A. 1978; 75:740–744. [PubMed: 204930]

Bentley DL. Coupling mRNA processing with transcription in time and space. Nat Rev Genet. 2014; 15:163–175. [PubMed: 24514444]

Bies-Etheve N, Pontier D, Lahmy S, Picart C, Vega D, Cooke R, Lagrange T. RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. EMBO Rep. 2009; 10:649–654. [PubMed: 19343051]

Blevins T, Podicheti R, Mishra V, Marasco M, Tang H, Pikaard CS. Identification of Pol IV and RDR2-dependent precursors of 24 nt siRNAs guiding de novo DNA methylation in Arabidopsis. Elife. 2015; 4:e09591. [PubMed: 26430765]

Blevins T, Pontvianne F, Cocklin R, Podicheti R, Chandrasekhara C, Yerneni S, Braun C, Lee B, Rusch D, Mockaitis K, et al. A two-step process for epigenetic inheritance in Arabidopsis. Mol Cell. 2014; 54:30–42. [PubMed: 24657166]

Bohmdorfer G, Rowley MJ, Kucinski J, Zhu Y, Amies I, Wierzbicki AT. RNA-directed DNA methylation requires stepwise binding of silencing factors to long non-coding RNA. Plant J. 2014; 79:181–191. [PubMed: 24862207]

Cao X, Aufsatz W, Zilberman D, Mette MF, Huang MS, Matzke M, Jacobsen SE. Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. Curr Biol. 2003; 13:2212–2217. [PubMed: 14680640]

Cao X, Jacobsen S. Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proc Natl Acad Sci U S A. 2002a; (99 Suppl 4):16491–16498. [PubMed: 12151602]

Cao X, Jacobsen SE. Role of the Arabidopsis DRM Methyltransferases in De Novo DNA Methylation and Gene Silencing. Curr Biol. 2002b; 12:1138–1144. [PubMed: 12121623]

Chandrasekhara C, Mohannath G, Blevins T, Pontvianne F, Pikaard CS. Chromosome-specific NOR inactivation explains selective rRNA gene silencing and dosage control in Arabidopsis. Genes Dev. 2016; 30:177–190. [PubMed: 26744421]

Cramer P, Armache KJ, Baumli S, Benkert S, Brueckner F, Buchen C, Damsma GE, Dengl S, Geiger SR, Jasiak AJ, et al. Structure of eukaryotic RNA polymerases. Annu Rev Biophys. 2008; 37:337–352. [PubMed: 18573085]

Creasey KM, Zhai J, Borges F, Van Ex F, Regulski M, Meyers BC, Martienssen RA. miRNAs trigger widespread epigenetically activated siRNAs from transposons in Arabidopsis. Nature. 2014; 508:411–415. [PubMed: 24670663]

Deleris A, Greenberg MV, Ausin I, Law RW, Moissiard G, Schubert D, Jacobsen SE. Involvement of a Jumonji-C domain-containing histone demethylase in DRM2-mediated maintenance of DNA methylation. EMBO Rep. 2010; 11:950–955. [PubMed: 21052090]

Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, Vashisht AA, Terragni J, Chin HG, Tu A, et al. Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell. 2012; 151:167–180. [PubMed: 23021223]

Duan CG, Zhang H, Tang K, Zhu X, Qian W, Hou YJ, Wang B, Lang Z, Zhao Y, Wang X, et al. Specific but interdependent functions for Arabidopsis AGO4 and AGO6 in RNA-directed DNA methylation. EMBO J. 2015; 34:581–592. [PubMed: 25527293]

Earley KW, Pontvianne F, Wierzbicki AT, Blevins T, Tucker S, Costa-Nunes P, Pontes O, Pikaard CS. Mechanisms of HDA6-mediated rRNA gene silencing: suppression of intergenic Pol II transcription and differential effects on maintenance versus siRNA-directed cytosine methylation. Genes Dev. 2010; 24:1119–1132. [PubMed: 20516197]

Ebbs ML, Bartee L, Bender J. H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. Mol Cell Biol. 2005; 25:10507–10515. [PubMed: 16287862]

Ebbs ML, Bender J. Locus-specific control of DNA methylation by the Arabidopsis SUVH5 histone methyltransferase. Plant Cell. 2006; 18:1166–1176. [PubMed: 16582009]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

El-Shami M, Pontier D, Lahmy S, Braun L, Picart C, Vega D, Hakimi MA, Jacobsen SE, Cooke R, Lagrange T. Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. Genes Dev. 2007; 21:2539–2544. [PubMed: 17938239]

Elvira-Matelot E, Hachet M, Shamandi N, Comella P, Saez-Vasquez J, Zytnicki M, Vaucheret H. Arabidopsis RNASE THREE LIKE2 Modulates the Expression of Protein-Coding Genes via 24-Nucleotide Small Interfering RNA-Directed DNA Methylation. Plant Cell. 2016; 28:406–425. [PubMed: 26764378]

Eun C, Lorkovic ZJ, Naumann U, Long Q, Havecker ER, Simon SA, Meyers BC, Matzke AJ, Matzke M. AGO6 functions in RNA-mediated transcriptional gene silencing in shoot and root meristems in Arabidopsis thaliana. PLoS One. 2011; 6:e25730. [PubMed: 21998686]

Finnegan EJ, Dennis ES. Isolation and identification by sequence homology of a putative cytosine methyltransferase from Arabidopsis thaliana. Nucleic Acids Res. 1993; 21:2383–2388. [PubMed: 8389441]

Gao Z, Liu HL, Daxinger L, Pontes O, He X, Qian W, Lin H, Xie M, Lorkovic ZJ, Zhang S, et al. An RNA polymerase II- and AGO4-associated protein acts in RNA-directed DNA methylation. Nature. 2010; 465:106–109. [PubMed: 20410883]

Haag JR, Brower-Toland B, Krieger EK, Sidorenko L, Nicora CD, Norbeck AD, Irsigler A, LaRue H, Brzeski J, McGinnis K, et al. Functional Diversification of Maize RNA Polymerase IV and V Subtypes via Alternative Catalytic Subunits. Cell Rep. 2014; 9:378–390. [PubMed: 25284785]

Haag JR, Pikaard CS. Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nat Rev Mol Cell Biol. 2011; 12:483–492. [PubMed: 21779025]

Haag JR, Ream TS, Marasco M, Nicora CD, Norbeck AD, Pasa-Tolic L, Pikaard CS. 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. 2012; 48:811–818. [PubMed: 23142082]

Havecker ER, Wallbridge LM, Hardcastle TJ, Bush MS, Kelly KA, Dunn RM, Schwach F, Doonan JH, Baulcombe DC. The Arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. Plant Cell. 2010; 22:321–334. [PubMed: 20173091]

He XJ, Chen T, Zhu JK. Regulation and function of DNA methylation in plants and animals. Cell Res. 2011; 21:442–465. [PubMed: 21321601]

He XJ, Hsu YF, Pontes O, Zhu J, Lu J, Bressan RA, Pikaard C, Wang CS, Zhu JK. 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. 2009a; 23:318–330. [PubMed: 19204117]

He XJ, Hsu YF, Zhu S, Wierzbicki AT, Pontes O, Pikaard CS, Liu HL, Wang CS, Jin H, Zhu JK. An effector of RNA-directed DNA methylation in arabidopsis is an ARGONAUTE 4- and RNA-binding protein. Cell. 2009b; 137:498–508. [PubMed: 19410546]

Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. RNA polymerase IV directs silencing of endogenous DNA. Science. 2005; 308:118–120. [PubMed: 15692015]

Hocine S, Singer RH, Grunwald D. RNA processing and export. Cold Spring Harb Perspect Biol. 2010; 2:a000752. [PubMed: 20961978]

Holoch D, Moazed D. RNA-mediated epigenetic regulation of gene expression. Nat Rev Genet. 2015; 16:71–84. [PubMed: 25554358]

Hsin JP, Manley JL. The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. 2012; 26:2119–2137. [PubMed: 23028141]

Huang L, Jones AM, Searle I, Patel K, Vogler H, Hubner NC, Baulcombe DC. An atypical RNA polymerase involved in RNA silencing shares small subunits with RNA polymerase II. Nat Struct Mol Biol. 2009; 16:91–93. [PubMed: 19079263]

Huang Y, Kendall T, Forsythe ES, Dorantes-Acosta A, Li S, Caballero-Perez J, Chen X, Arteaga-Vazquez M, Beilstein MA, Mosher RA. Ancient Origin and Recent Innovations of RNA Polymerase IV and V. Mol Biol Evol. 2015; 32:1788–1799. [PubMed: 25767205]

Jackson JP, Johnson L, Jasencakova Z, Zhang X, PerezBurgos L, Singh PB, Cheng X, Schubert I, Jenuwein T, Jacobsen SE. Dimethylation of histone H3 lysine 9 is a critical mark for DNA

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

methylation and gene silencing in Arabidopsis thaliana. Chromosoma. 2004; 112:308–315. [PubMed: 15014946]

Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature. 2002; 416:556–560. [PubMed: 11898023]

Jain R, Iglesias N, Moazed D. Distinct Functions of Argonaute Slicer in siRNA Maturation and Heterochromatin Formation. Mol Cell. 2016

Jing Y, Sun H, Yuan W, Wang Y, Li Q, Liu Y, Li Y, Qian W. SUVH2 and SUVH9 couple two essential steps for transcriptional gene silencing in Arabidopsis. Mol Plant. 2016

Johnson L, Cao X, Jacobsen S. Interplay between two epigenetic marks. DNA methylation and histone H3 lysine 9 methylation. Curr Biol. 2002; 12:1360–1367. [PubMed: 12194816]

Johnson LM, Bostick M, Zhang X, Kraft E, Henderson I, Callis J, Jacobsen SE. The SRA methyl-cytosine-binding domain links DNA and histone methylation. Curr Biol. 2007; 17:379–384. [PubMed: 17239600]

Johnson LM, Du J, Hale CJ, Bischof S, Feng S, Chodavarapu RK, Zhong X, Marson G, Pellegrini M, Segal DJ, et al. SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. Nature. 2014; 507:124–128. [PubMed: 24463519]

Johnson LM, Law JA, Khattar A, Henderson IR, Jacobsen SE. SRA-domain proteins required for DRM2-mediated de novo DNA methylation. PLoS genetics. 2008; 4:e1000280. [PubMed: 19043555]

Kanno T, Aufsatz W, Jaligot E, Mette MF, Matzke M, Matzke AJ. A SNF2-like protein facilitates dynamic control of DNA methylation. EMBO Rep. 2005a; 6:649–655. [PubMed: 15947783]

Kanno T, Bucher E, Daxinger L, Huettel B, Bohmdorfer G, Gregor W, Kreil DP, Matzke M, Matzke AJ. A structural-maintenance-of-chromosomes hinge domain-containing protein is required for RNA-directed DNA methylation. Nat Genet. 2008; 40:670–675. [PubMed: 18425128]

Kanno T, Huettel B, Mette MF, Aufsatz W, Jaligot E, Daxinger L, Kreil DP, Matzke M, Matzke AJ. Atypical RNA polymerase subunits required for RNA-directed DNA methylation. Nat Genet. 2005b; 37:761–765. [PubMed: 15924141]

Kollen K, Dietz L, Bies-Etheve N, Lagrange T, Grasser M, Grasser KD. The zinc-finger protein SPT4 interacts with SPT5L/KTF1 and modulates transcriptional silencing in Arabidopsis. FEBS Lett. 2015; 589:3254–3257. [PubMed: 26424658]

Lahmy S, Pontier D, Cavel E, Vega D, El-Shami M, Kanno T, Lagrange T. PolV(PolIVb) function in RNA-directed DNA methylation requires the conserved active site and an additional plant-specific subunit. Proc Natl Acad Sci U S A. 2009; 106:941–946. [PubMed: 19141635]

Lange H, Holec S, Cognat V, Pieuchot L, Le Ret M, Canaday J, Gagliardi D. Degradation of a polyadenylated rRNA maturation by-product involves one of the three RRP6-like proteins in Arabidopsis thaliana. Mol Cell Biol. 2008; 28:3038–3044. [PubMed: 18285452]

Law JA, Ausin I, Johnson LM, Vashisht AA, Zhu JK, Wohlschlegel JA, Jacobsen SE. A protein complex required for polymerase V transcripts and RNA- directed DNA methylation in Arabidopsis. Curr Biol. 2010; 20:951–956. [PubMed: 20409711]

Law JA, Du J, Hale CJ, Feng S, Krajewski K, Palanca AM, Strahl BD, Patel DJ, Jacobsen SE. Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. Nature. 2013; 498:385–389. [PubMed: 23636332]

Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet. 2010; 11:204–220. [PubMed: 20142834]

Law JA, Vashisht AA, Wohlschlegel JA, Jacobsen SE. SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. PLoS Genet. 2011; 7:e1002195. [PubMed: 21811420]

Li J, Yang Z, Yu B, Liu J, Chen X. Methylation protects miRNAs and siRNAs from a 3’-end uridylation activity in Arabidopsis. Curr Biol. 2005; 15:1501–1507. [PubMed: 16111943]

Li S, Vandivier LE, Tu B, Gao L, Won SY, Li S, Zheng B, Gregory BD, Chen X. Detection of Pol IV/RDR2-dependent transcripts at the genomic scale in Arabidopsis reveals features and regulation of siRNA biogenesis. Genome Res. 2015; 25:235–245. [PubMed: 25414514]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Liu ZW, Shao CR, Zhang CJ, Zhou JX, Zhang SW, Li L, Chen S, Huang HW, Cai T, He XJ. The SET domain proteins SUVH2 and SUVH9 are required for Pol V occupancy at RNA-directed DNA methylation loci. PLoS Genet. 2014; 10:e1003948. [PubMed: 24465213]

Liu ZW, Zhou JX, Huang HW, Li YQ, Shao CR, Li L, Cai T, Chen S, He XJ. Two Components of the RNA-Directed DNA Methylation Pathway Associate with MORC6 and Silence Loci Targeted by MORC6 in Arabidopsis. PLoS Genet. 2016; 12:e1006026. [PubMed: 27171427]

Lorkovic ZJ, Naumann U, Matzke AJ, Matzke M. Involvement of a GHKL ATPase in RNA-directed DNA methylation in Arabidopsis thaliana. Curr Biol. 2012; 22:933–938. [PubMed: 22560611]

Luo J, Hall BD. A multistep process gave rise to RNA polymerase IV of land plants. J Mol Evol. 2007; 64:101–112. [PubMed: 17160640]

Martinez-Rucobo FW, Kohler R, van de Waterbeemd M, Heck AJ, Hemann M, Herzog F, Stark H, Cramer P. Molecular Basis of Transcription-Coupled Pre-mRNA Capping. Mol Cell. 2015; 58:1079–1089. [PubMed: 25959396]

Matzke MA, Mosher RA. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet. 2014; 15:394–408. [PubMed: 24805120]

McCue AD, Panda K, Nuthikattu S, Choudury SG, Thomas EN, Slotkin RK. ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. EMBO J. 2015; 34:20–35. [PubMed: 25388951]

Mourrain P, Beclin C, Elmayan T, Feuerbach F, Godon C, Morel JB, Jouette D, Lacombe AM, Nikic S, Picault N, et al. Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell. 2000; 101:533–542. [PubMed: 10850495]

Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol. 2000; 302:205–217. [PubMed: 10964570]

Nuthikattu S, McCue AD, Panda K, Fultz D, DeFraia C, Thomas EN, Slotkin RK. The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21–22 nucleotide small interfering RNAs. Plant Physiol. 2013; 162:116–131. [PubMed: 23542151]

Olmedo-Monfil V, Duran-Figueroa N, Arteaga-Vazquez M, Demesa-Arevalo E, Autran D, Grimanelli D, Slotkin RK, Martienssen RA, Vielle-Calzada JP. Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature. 2010; 464:628–632. [PubMed: 20208518]

Onodera Y, Haag JR, Ream T, Costa Nunes P, Pontes O, Pikaard CS. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell. 2005; 120:613–622. [PubMed: 15766525]

Panda K, Slotkin RK. Proposed mechanism for the initiation of transposable element silencing by the RDR6-directed DNA methylation pathway. Plant Signal Behav. 2013; 8

Pearson E, Moore C. The evolutionarily conserved Pol II flap loop contributes to proper transcription termination on short yeast genes. Cell Rep. 2014; 9:821–828. [PubMed: 25437538]

Pikaard CS, Haag JR, Pontes OM, 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. 2013

Pontes O, Li CF, Costa Nunes P, Haag J, Ream T, Vitins A, Jacobsen SE, Pikaard CS. The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. Cell. 2006; 126:79–92. [PubMed: 16839878]

Pontier D, Picart C, Roudier F, Garcia D, Lahmy S, Azevedo J, Alart E, Laudie M, Karlowski WM, Cooke R, et al. NERD, a plant-specific GW protein, defines an additional RNAi-dependent chromatin-based pathway in Arabidopsis. Mol Cell. 2012; 48:121–132. [PubMed: 22940247]

Pontier D, Yahubyan G, Vega D, Bulski A, Saez-Vasquez J, Hakimi MA, Lerbs-Mache S, Colot V, Lagrange T. Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. Genes Dev. 2005; 19:2030–2040. [PubMed: 16140984]

Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ. Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. Nature. 2006; 443:1008–1012. [PubMed: 16998468]

Ream TS, Haag JR, Wierzbicki AT, Nicora CD, Norbeck AD, Zhu JK, Hagen G, Guilfoyle TJ, Pasa-Tolic L, Pikaard CS. Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

their origins as specialized forms of RNA polymerase II. Mol Cell. 2009; 33:192–203. [PubMed: 19110459]

Ronemus MJ, Galbiati M, Ticknor C, Chen J, Dellaporta SL. Demethylation-induced developmental pleiotropy in Arabidopsis. Science. 1996; 273:654–657. [PubMed: 8662558]

Rowley MJ, Avrutsky MI, Sifuentes CJ, Pereira L, Wierzbicki AT. Independent chromatin binding of ARGONAUTE4 and SPT5L/KTF1 mediates transcriptional gene silencing. PLoS Genet. 2011; 7:e1002120. [PubMed: 21738482]

Searle IR, Pontes O, Melnyk CW, Smith LM, Baulcombe DC. JMJ14, a JmjC domain protein, is required for RNA silencing and cell-to-cell movement of an RNA silencing signal in Arabidopsis. Genes Dev. 2010; 24:986–991. [PubMed: 20478993]

Sigman MJ, Slotkin RK. The First Rule of Plant Transposable Element Silencing: Location, Location, Location. Plant Cell. 2016; 28:304–313. [PubMed: 26869697]

Soppe WJ, Jasencakova Z, Houben A, Kakutani T, Meister A, Huang MS, Jacobsen SE, Schubert I, Fransz PF. DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. EMBO J. 2002; 21:6549–6559. [PubMed: 12456661]

Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, Patel DJ, Jacobsen SE. Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. Nat Struct Mol Biol. 2014; 21:64–72. [PubMed: 24336224]

Suh MH, Meyer PA, Gu M, Ye P, Zhang M, Kaplan CD, Lima CD, Fu J. A dual interface determines the recognition of RNA polymerase II by RNA capping enzyme. J Biol Chem. 2010; 285:34027–34038. [PubMed: 20720002]

Trujillo JT, Beilstein MA, Mosher RA. The Argonaute-binding platform of NRPE1 evolves through modulation of intrinsically disordered repeats. New Phytol. 2016

Volpe T, Martienssen RA. RNA interference and heterochromatin assembly. Cold Spring Harb Perspect Biol. 2011; 3:a003731. [PubMed: 21441597]

Wei W, Ba Z, Gao M, Wu Y, Ma Y, Amiard S, White CI, Rendtlew Danielsen JM, Yang YG, Qi Y. A role for small RNAs in DNA double-strand break repair. Cell. 2012; 149:101–112. [PubMed: 22445173]

Wierzbicki AT, Cocklin R, Mayampurath A, Lister R, Rowley MJ, Gregory BD, Ecker JR, Tang H, Pikaard CS. Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the Arabidopsis epigenome. Genes Dev. 2012; 26:1825–1836. [PubMed: 22855789]

Wierzbicki AT, Haag JR, Pikaard CS. Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell. 2008; 135:635–648. [PubMed: 19013275]

Wierzbicki AT, Ream TS, Haag JR, Pikaard CS. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. Nat Genet. 2009; 41:630–634. [PubMed: 19377477]

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. 2012a; 40:4422–4431. [PubMed: 22302148]

Xie M, Ren G, Zhang C, Yu B. The DNA- and RNA-binding protein FACTOR of DNA METHYLATION 1 requires XH domain-mediated complex formation for its function in RNA-directed DNA methylation. Plant J. 2012b; 72:491–500. [PubMed: 22757778]

Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC. Genetic and functional diversification of small RNA pathways in plants. PLoS Biol. 2004; 2:E104. [PubMed: 15024409]

Yang DL, Zhang G, Tang K, Li J, Yang L, Huang H, Zhang H, Zhu JK. Dicer-independent RNA-directed DNA methylation in Arabidopsis. Cell Res. 2016; 26:66–82. [PubMed: 26642813]

Ye R, Chen Z, Lian B, Rowley MJ, Xia N, Chai J, Li Y, He XJ, Wierzbicki AT, Qi Y. A Dicer-Independent Route for Biogenesis of siRNAs that Direct DNA Methylation in Arabidopsis. Mol Cell. 2016; 61:222–235. [PubMed: 26711010]

Zemach A, Kim MY, Hsieh PH, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL, Zilberman D. The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell. 2013; 153:193–205. [PubMed: 23540698]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Zhai J, Bischof S, Wang H, Feng S, Lee TF, Teng C, Chen X, Park SY, Liu L, Gallego-Bartolome J, et al. A One Precursor One siRNA Model for Pol IV-Dependent siRNA Biogenesis. Cell. 2015; 163:445–455. [PubMed: 26451488]

Zhang CJ, Ning YQ, Zhang SW, Chen Q, Shao CR, Guo YW, Zhou JX, Li L, Chen S, He XJ. IDN2 and its paralogs form a complex required for RNA-directed DNA methylation. PLoS Genet. 2012; 8:e1002693. [PubMed: 22570638]

Zhang H, Ma ZY, Zeng L, Tanaka K, Zhang CJ, Ma J, Bai G, Wang P, Zhang SW, Liu ZW, 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 U S A. 2013; 110:8290–8295. [PubMed: 23637343]

Zhang H, Tang K, Qian W, Duan CG, Wang B, Zhang H, Wang P, Zhu X, Lang Z, Yang Y, et al. An Rrp6-like protein positively regulates noncoding RNA levels and DNA methylation in Arabidopsis. Mol Cell. 2014; 54:418–430. [PubMed: 24726328]

Zheng Q, Rowley MJ, Bohmdorfer G, Sandhu D, Gregory BD, Wierzbicki AT. RNA polymerase V targets transcriptional silencing components to promoters of protein-coding genes. Plant J. 2013; 73:179–189. [PubMed: 23013441]

Zheng X, Zhu J, Kapoor A, Zhu JK. Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. EMBO J. 2007; 26:1691–1701. [PubMed: 17332757]

Zheng Z, Xing Y, He XJ, Li W, Hu Y, Yadav SK, Oh J, Zhu JK. An SGS3-like protein functions in RNA-directed DNA methylation and transcriptional gene silencing in Arabidopsis. Plant J. 2010; 62:92–99. [PubMed: 20059743]

Zhong X, Du J, Hale CJ, Gallego-Bartolome J, Feng S, Vashisht AA, Chory J, Wohlschlegel JA, Patel DJ, Jacobsen SE. Molecular mechanism of action of plant DRM de novo DNA methyltransferases. Cell. 2014; 157:1050–1060. [PubMed: 24855943]

Zhong X, Hale CJ, Law JA, Johnson LM, Feng S, Tu A, Jacobsen SE. DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. Nat Struct Mol Biol. 2012; 19:870–875. [PubMed: 22864289]

Zhu Y, Rowley MJ, Bohmdorfer G, Wierzbicki AT. A SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing. Mol Cell. 2013; 49:298–309. [PubMed: 23246435]

Zilberman D, Cao X, Jacobsen SE. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science. 2003; 299:716–719. [PubMed: 12522258]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- RNA-directed DNA methylation in plants involves siRNAs and long noncoding RNAs
- Double-stranded RNA precursors of 24 nt siRNAs have recently been identified
- Properties of long noncoding RNAs made by RNA Polymerase V remain enigmatic
- Pol V transcripts have 5' triphosphate groups and lack poly-A tails
- Facts, missing information and controversies are reviewed

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

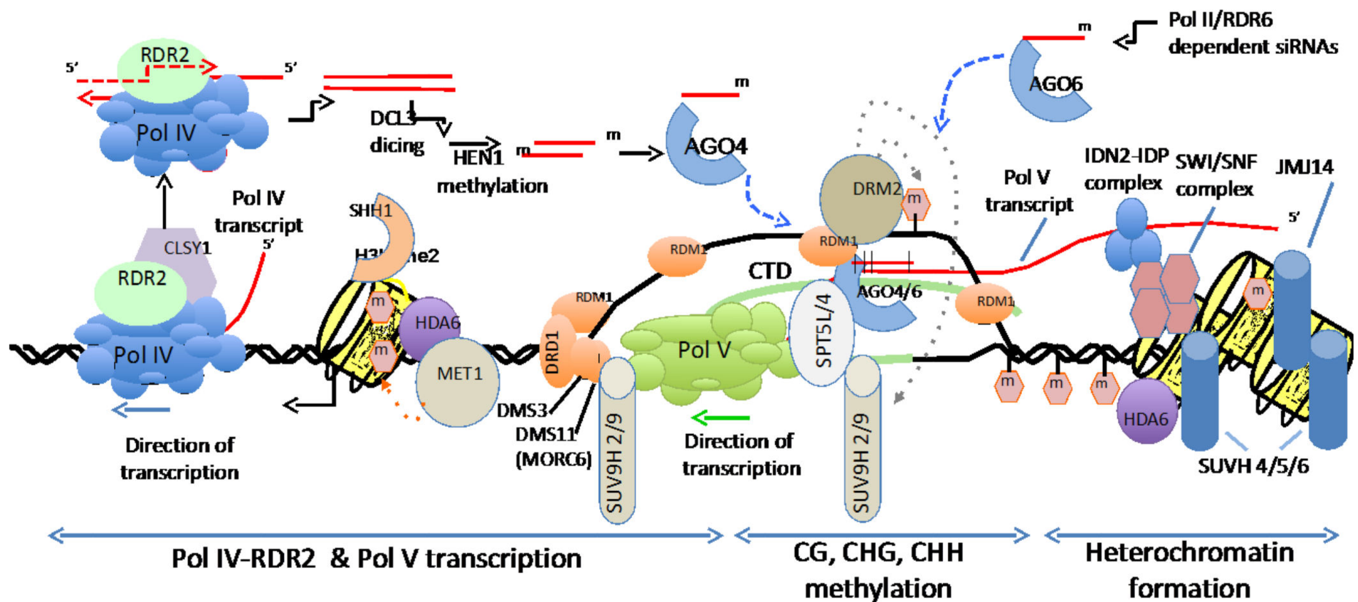

Figure 1. RNA-directed DNA methylation in Arabidopsis

It is useful to consider RdDM as having three major aspects: RNA synthesis (by Pol IV, RDR2 and Pol V), cytosine methylation, and heterochromatin formation. CG maintenance methylation, requiring the histone deacetylase, HDA6 and cytosine methyltransferase, MET1 is linked to dimethylation of histone H3 on lysine 9 (H3K9me2). The Pol IV partner protein, SHH1 binds H3K9me2, helping recruit Pol IV to the chromatin to be transcribed. Pol IV somehow gains access to the DNA and initiates transcription, generating short transcripts in partnership with RDR2 and the putative ATP-dependent DNA translocase, CLSY1 (or related CLSY proteins), whose function is unknown. MET1-dependent CG maintenance methylation also plays a role in Pol V recruitment via SUVH2 or SUVH9, which bind to methylated CG motifs and interact with MORC6/DMS11, which associates with the DDR complex (DRD1, DMS3, RDM1) known to be required for production of Pol V-transcribed RNAs. DRD1 is an ATP-dependent DNA translocase, and RDM1 has single-stranded DNA binding activity, suggesting that DRD1 may act as a helicase to provide Pol V with a single-stranded template. AGO4-siRNA complexes bind Pol V transcripts, as do SPT4/5, the IDN-IDP, and SWI/SNF complexes, helping facilitate protein-protein interactions that recruit the de novo cytosine methyltransferase, DRM2 and histone modifying enzymes, resulting in extensive cytosine methylation and heterochromatin formation.

Author Manuscript

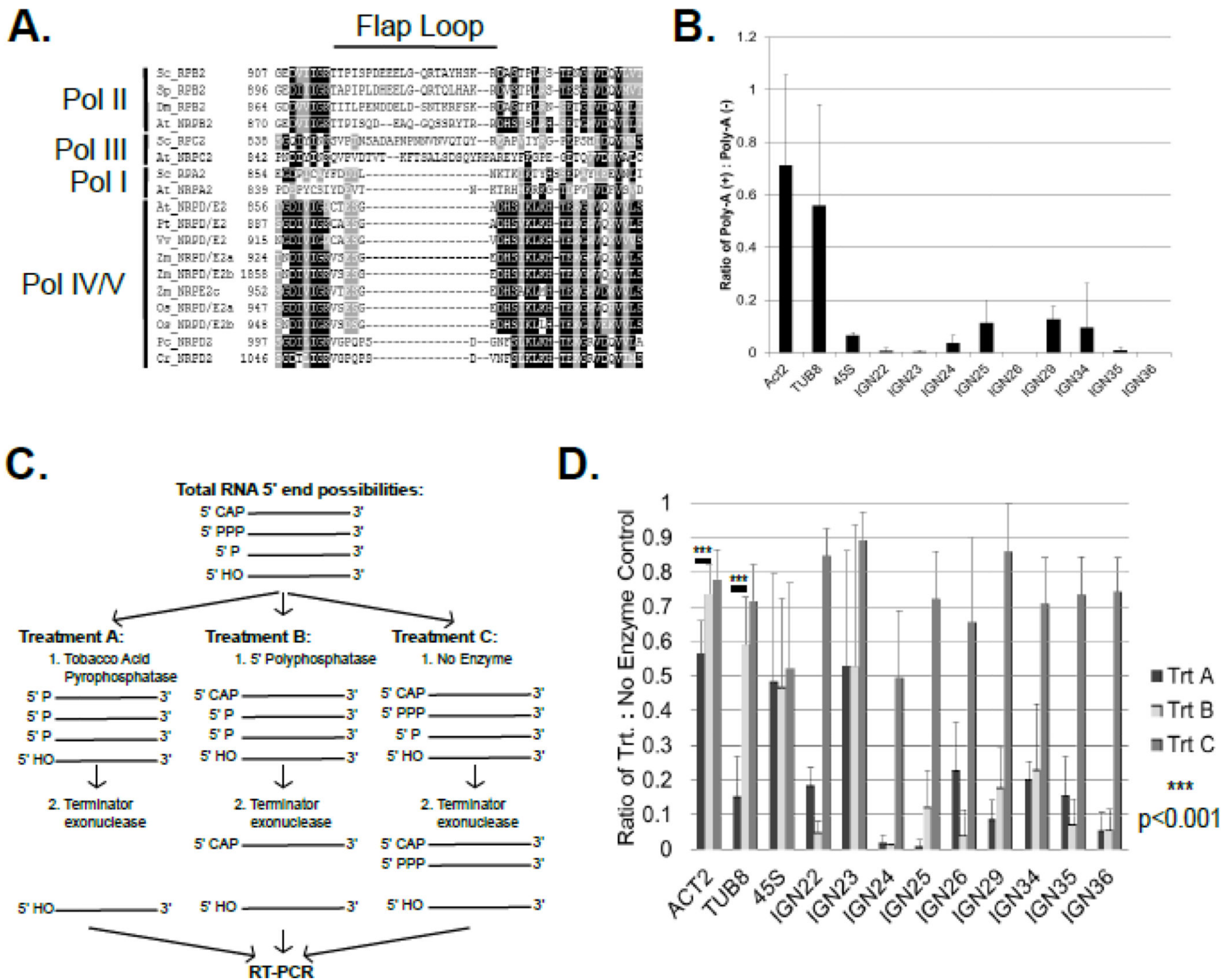

Figure 2. Pol V transcripts lack the poly A tails and 7-methylguanosine caps of Pol II transcribed mRNAs

A) Amino acid alignment of the second-largest subunits of Pols I, II, III, and IV/V in various species, highlighting the flap-loop domain. This domain of Pol II is absent in Pols I, III, and IV/V. Note that the same protein, NRP(D/E)2 serves as the second-largest subunits of Pols IV and V. For Pol II, the flap loop domain is important for interactions with both the capping and polyadenylation machinery. Sc: *Saccharomyces cerevisiae*; Sp: *Schizosaccharomyces pombe*; Dm: *Drosophila melanogaster*; At: *Arabidopsis thaliana*; Pt: *Populus trichocarpa*; Vv: *Vitis vinifera*; Zm: *Zea mays*; Os: *Oryza sativa*; Pc: *Pinus canariensis*; Cr: *Cycas revoluta*.

B) Results of oligo-dT bead fractionation of total RNA, followed by qRT-PCR. Histograms display the average ratio of transcript levels in the Poly-A enriched (bead) fraction relative to the Poly-A depleted (supernatant) fraction for three biologic replicates (error bars are one standard deviation from the mean). Note that Pol V-dependent transcripts are enriched at significantly lower levels in the Poly-A fraction compared to the Pol II-dependent transcripts ACT2 and TUB8, suggesting they Pol V transcripts are not polyadenylated.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

C) Schematic of enzyme treatments used to distinguish if Pol V transcripts possess a 7-methylguanosine cap, triphosphate, or monophosphate group on the 5' terminal nucleotide. Treatment A consisted of an initial treatment with Tobacco Acid Pyrophosphatase (TAP) to convert any 5' caps or 5' triphosphate groups to monophosphates, followed by a treatment with Terminator exonuclease, which degrades RNA with a 5' monophosphate group, but has no effect on RNA molecules with 5' caps or triphosphate groups. Treatment B included an initial treatment with 5' Polyphosphatase, which converts 5' triphosphates into monophosphates, but has no effect on 5' caps, followed by a treatment with Terminator exonuclease. Treatment C consisted of treatment with Terminator exonuclease only. As a control, an equal amount of total RNA was also subjected to each treatment step with no enzyme added.

D) After treatments, transcript levels were measured by quantitative RT-PCR and ratios of transcript levels in treatment (+) enzyme : treatment (-) enzyme were compared for each treatment group, A, B, and C (shown are the average ratios across at least three biologic replicates, with error bars showing one standard deviation from the mean). With this treatment scheme, we predicted that a transcript with a 5' cap should be most severely depleted by Treatment A and relatively resistant to Treatments B and C. A transcript with a 5' triphosphate group was predicted to be ~equally depleted by Treatments A and B and resistant to C, whereas, a transcript with a 5' monophosphate should be equally depleted in all treatments. We found these predictions to be accurate for our Pol II control transcripts, ACT2 and TUB8, which were significantly more depleted by Treatment A than Treatments B and C. Also, the Pol I control transcript, 45S, which is characterized by a 5-triphosphate in its primary state, was equally depleted by Treatments A and B. The Pol V transcripts assayed were also equally depleted by Treatments A and B, which had a greater effect than Treatment C, suggesting the majority of the Pol V transcript population is characterized by a 5' triphosphate.