Cell, Vol. 120, 613–622, March 11, 2005, Copyright ©2005 by Elsevier Inc. DOI 10.1016/j.cell.2005.02.007

Plant Nuclear RNA Polymerase IV Mediates siRNA and DNA Methylation-Dependent Heterochromatin Formation

Yasuyuki Onodera,[1,2,4] Jeremy R. Haag,[1,4] Thomas Ream,[1,4] Pedro Costa Nunes,[3] Olga Pontes,[1] and Craig S. Pikaard[1,*]
[1]Biology Department
Washington University
1 Brookings Drive
St. Louis, Missouri 63130
[2]Graduate School of Agriculture
Faculty of Agriculture
Hokkaido University
Kita 9, Nishi 9, Kita-ku
Sapporo 060-8589
Japan
[3]Secção de Genética
Centro de Botanica e Engenharia Biologica
Instituto Superior de Agronomia
Tapada da Ajuda
1349-017 Lisboa
Portugal

Summary

All eukaryotes have three nuclear DNA-dependent RNA polymerases, namely, Pol I, II, and III. Interestingly, plants have catalytic subunits for a fourth nuclear polymerase, Pol IV. Genetic and biochemical evidence indicates that Pol IV does not functionally overlap with Pol I, II, or III and is nonessential for viability. However, disruption of the Pol IV catalytic subunit genes *NRPD1* or *NRPD2* inhibits heterochromatin association into chromocenters, coincident with losses in cytosine methylation at pericentromeric 5S gene clusters and *AtSN1* retroelements. Loss of CG, CNG, and CNN methylation in Pol IV mutants implicates a partnership between Pol IV and the methyltransferase responsible for RNA-directed de novo methylation. Consistent with this hypothesis, 5S gene and *AtSN1* siRNAs are essentially eliminated in Pol IV mutants. The data suggest that Pol IV helps produce siRNAs that target de novo cytosine methylation events required for facultative heterochromatin formation and higher-order heterochromatin associations.

Introduction

In eukaryotes, three nuclear DNA-dependent RNA polymerases (RNAPs) transcribe genomic DNA into RNA. RNA polymerase I (Pol I) transcribes the ribosomal RNA (rRNA) genes clustered at nucleolus organizer regions (Grummt, 2003); RNA polymerase II (Pol II) transcribes the vast majority of genes, including protein-coding genes (Woychik and Hampsey, 2002), and RNA polymerase III (Pol III) transcribes genes encoding short (<400 nt) structural RNAs that include tRNAs and 5S rRNA (Schramm and Hernandez, 2002).

RNA polymerases I, II, and III are composed of 12–17 proteins, including subunits sharing sequence and structural homology with the eubacterial RNA polymerase subunits β′, β, α^{I}, α^{II}, and ω (Archambault and Friesen, 1993; Cramer et al., 2001; Zhang et al., 1999). RNA Pol I, II, and III (designated RPA, RPB, and RPC in yeast and N [nuclear] RPA, NRPB, and NRPC in *Arabidopsis*) largest subunits are homologous to eubacterial β′ and are encoded by different genes, (*N*)*RPA1*, (*N*)*RPB1*, and (*N*)*RPC1*. Likewise, the second-largest subunits of Pol I, II, and III are β homologs encoded by (*N*)*RPA2*, (*N*)*RPB2*, and (*N*)*RPC2*. Together, the largest and second-largest subunits form the catalytic center in which RNA synthesis occurs (Cramer et al., 2000; Zhang et al., 1999), with α^{I}, α^{II}, and ω serving regulatory or assembly functions.

Surprisingly, analysis of the *Arabidopsis thaliana* genome sequence revealed evidence for a fourth class of RNA polymerase in addition to Pol I, II, and III (CSP and Jonathan Eisen, discussed in *Arabidopsis* Genome Initiative [2000]). Specifically, two class IV largest and second-largest subunit genes were predicted, implying the existence of a nuclear RNA polymerase IV (Pol IV) distinct from eubacterial-type RNAPs of chloroplasts, from mitochondrial polymerase, or from RNA-dependent RNA polymerases (RdRP).

Here, we present evidence that RNA Pol IV is located within the nucleus and plays a role in heterochromatin formation. Dispersal of chromocenters in Pol IV mutants is correlated with the loss of cytosine methylation from pericentromeric 5S gene clusters and *AtSN1* retroelements. By contrast, methylation of constitutively heterochromatic 180 bp centromere core repeats is not appreciably affected in Pol IV mutants. We propose that Pol IV is required for the production of siRNAs that direct de novo methylation of repetitive elements that are subject to facultative heterochromatin formation, thereby facilitating higher-order heterochromatin associations.

Results

Genes for RNA Pol IV

An unrooted phylogenetic tree of DNA-dependent RNA polymerase (RNAP) largest subunits (Figure 1A) reveals distinct clades for eubacteria, cyanobacteria and chloroplasts, archaea, DNA viruses, and eukaryotic RNA polymerases I (RPA1), II (RPB1), and III (RPC1). *Arabidopsis thaliana* (At) Pol I, II, and III largest subunits group with their orthologs from rice (Os), yeast (Sp and Sc), *C. elegans* (Ce), *Drosophila* (Dm), and human (Hs). Unlike other eukaryotes, *Arabidopsis* and rice have additional genes (*NRPD1a* and *b*) that form a clade for a putative Pol IV.

An unrooted tree of RNAP second-largest subunits resembles the tree for the largest subunits (Figure 1B). Again, in addition to clades for *RPA2* (Pol I), *RPB2* (Pol II), and *RPC2* (Pol III), a plant-specific *NRPD2* (Pol IV)

*Correspondence: pikaard@biology.wustl.edu
[4]These authors contributed equally to this work.

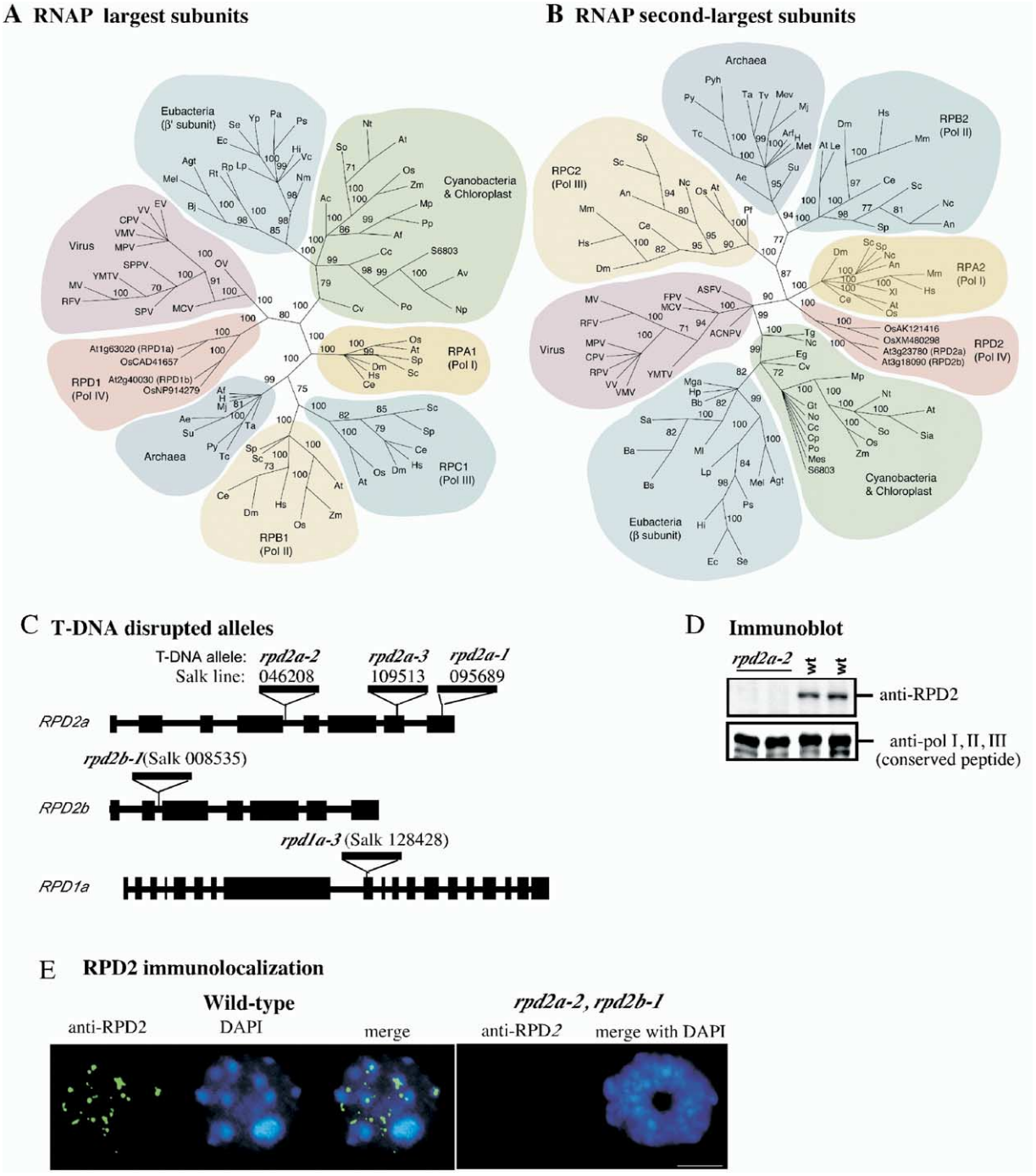

Figure 1. Evidence for RNA Pol IV in Plants

(A and B) Unrooted neighbor-joining phylogenies based on conserved domains A, C, D, and F of DNA-dependent RNA polymerase largest subunits and conserved domains A, C, D, F, G, H, and I of DNA-dependent RNA polymerase second-largest subunits. Bootstrap values are given for branch nodes. Species designations and GenBank accession numbers for the sequences analyzed are provided in Tables S1 and S2.

(C) Diagrams of T-DNA-disrupted *nrpd2* and *nrpd1* alleles. Exons are denoted by black rectangles.

(D) Immunoblot showing no detectable NRPD2 protein in two *nrpd2a-2* mutant individuals, unlike their wild-type siblings. A control immunoblot utilized an antibody raised against a peptide conserved in Pol I, II, and III second-largest subunits.

(E) NRPD2 localizes to the nucleus. On the left is a wild-type interphase nucleus showing immunolocalization of NRPD2 relative to ten DAPI-positive chromocenters. On the right is a homozygous *nrpd2a-1 nrpd2b-1* nucleus. The dark, DAPI-negative region is the nucleolus. The wild-type and mutant plants were progeny of homozygous siblings. The size bar corresponds to 5 μm.

Arabidopsis pol IV subunit names are abbreviated from NRPD to RPD in this and all subsequent figures.

clade exists. In both *Arabidopsis* and rice, there are two *NRPD2* genes (*NRPD2a* and *NRPD2b*) that were apparently duplicated after monocots and dicots diverged.

Multiple alignments revealed that NRPD2 proteins closely resemble their Pol I–III homologs, whereas NRPD1 sequences frequently lack amino acids that are invariant in Pol I–III largest subunits, including amino acids near the active site (see Figures S1–S4 in the Supplemental Data available with this article online). Therefore, we focused our studies on NRPD2 but also subjected *nrpd1a* mutants to a subset of the same assays. *NRPD1b* was ignored because existing annotation suggested that this gene lacks essential C-terminal domains.

Only *NRPD2a* appears to be expressed in *Arabidopsis*, based on existing EST (cDNA) sequences and by our inability to amplify *NRPD2b* RNA using RT-PCR or 5′ RACE. By contrast, *NRPD2a* sequences were readily amplified by PCR and by primer extension (Figure S5) to yield a full-length mRNA sequence (GenBank accession number AY862891).

Salk lines 046208, 109513, and 095689 contain the T-DNA-disrupted mutant alleles *nrpd2a-2*, *nrpd2a-3*, and *nrpd2a-1*, respectively. Salk lines 008535 and 128428 contain the *nrpd2b-1* and *nrpd1a-3* alleles (Figure 1C). Plants homozygous for these alleles were identified by PCR or Southern blot analysis of segregating families. The *nrpd2a* and *nrpd1a* alleles are all recessive and cause equivalent molecular phenotypes (data below and data not shown).

NRPD2 Expression and Nuclear Localization

RNA and protein blot analyses showed that NRPD2a is expressed throughout the plant but is most highly expressed in flowers and roots (data not shown). In homozygous *nrpd2a-2* mutants, no NRPD2 protein is detectable (Figure 1D), indicating that *nrpd2a-2* is a null allele. Immunolocalization of NRPD2 showed it to be a nuclear protein that is concentrated in numerous distinct foci (Figure 1E). Examination of 56 interphase nuclei revealed 10–15 NRPD2 signals in 71% of the nuclei and fewer than ten signals in 29% of the nuclei. In the nucleus shown, there are ten prominent DAPI-positive heterochromatic chromocenters, which are made up of centromeric repeats for the ten chromosomes, dispersed pericentromeric repeats, and four NORs (nucleolus organizer regions) (Fransz et al., 2002). Approximately 15 NRPD2 signals of varying size are apparent in Figure 1E, five of which are located at chromocenters and five of which are at the edges of chromocenters. Similar association of NRPD2 with chromocenters was observed in all nuclei.

Genetic Analysis of NRPD Mutants

To rule out any possible functional redundancy of *NRPD2a* and *NRPD2b*, we generated lines homozygous for both the *nrpd2a-2* and *nrpd2b-1* alleles, which was laborious, because the genes are linked (~10 cM genetic distance). We first crossed *nrpd2a-2* and *nrpd2b-1* homozygotes to generate F1 individuals that were hemizygous for each allele. The F1 was then outcrossed with a wild-type plant such that all resulting progeny had a wild-type chromosome 3 and either an *nrpd2a-2* or an *nrpd2b-1* allele but not both, unless a meiotic recombination event occurred between the two genes. We then identified the latter rare recombinants that had one wild-type chromosome 3 and one chromosome 3 bearing both the *nrpd2a-2* and *nrpd2b-1* alleles, allowed these to self-fertilize, and genotyped their progeny. Plants homozygous for both *nrpd2a -2* and *nrpd2b-1* (referred to as *nrpd2* double mutants or simply *nrpd2* in the remainder of the paper) were recovered, demonstrating that *NRPD2* is nonessential for viability. Siblings that were homozygous for the wild-type *NRPD2* gene were also identified and used as controls in subsequent assays. This genetic strategy is likely to have segregated away any potential T-DNAs unlinked to *NRPD2*, but, if such T-DNAs persist, they are as likely in the wild-type control plants as in their double mutant siblings.

We tested whether NRPD2 might be functionally redundant with the NRPA2, NRPB2, or NRPC2 subunits of Pol I–III by asking if any of these subunits were nonessential. We identified hemizygous individuals bearing T-DNA insertions in *NRPA2*, *NRPB2*, or *NRPC2* and genotyped 60–80 of their progeny. Only homozygous wild-type and hemizygous progeny were obtained; no homozygous mutants were recovered (data not shown). These results indicate that *NRPA2*, *NRPB2*, and *NRPC2* are essential genes, unlike *NRPD2a* and *NRPD2b*, and that *NRPD2* genes do not complement *nrpa2*, *nrpb2*, or *nrpc2* mutations. The *nrpd2* double mutation also failed to induce haploinsufficiency in plants hemizygous for *nrpa2*, *nrpb2*, or *nrpc2* mutations, consistent with the interpretation that NRPD2 does not overlap functionally with Pol I, II, or III.

NRPD2 Does Not Copurify with DNA-Dependent RNA Polymerases I–III

Among *Arabidopsis* RNAP second-largest subunits, NRPD2 is most similar to NRPB2 (Figure 2A). Therefore, we asked if NRPD2 copurified with RNA Pol II activity, as might be expected if NRPD2 is an alternative Pol II subunit. Nuclear extract was fractionated by anion exchange chromatography, and fractions were tested for DNA-dependent RNA polymerase activity (Figure 2B) and for the presence of NRPD2, NRPB2, or a 24 kDa polymerase subunit (RPB5) that is shared by Pol I, II, and III (Larkin et al., 1999; Saez-Vasquez and Pikaard, 2000).

The DNA-dependent RNA polymerase assay measures the incorporation of radioactive nucleotide triphosphates into RNA using sheared template DNA, which allows polymerase initiation from broken DNA ends in a promoter-independent fashion (Schwartz and Roeder, 1974). Duplicate reactions were performed with and without α-amanitin, a potent inhibitor of RNA Pol II, and mean values were plotted (Figure 2B). Comparison of the RNA polymerase activity profiles reveals a peak of activity that is inhibited by α-amanitin (fractions 29–37), indicative of Pol II (Figure 2B). As expected, NRPB2 eluted in these fractions (Figure 2C). By contrast, NRPD2 eluted in fractions 15–18, suggesting that NRPD2 is not an alternative Pol II subunit. Immunoblotting of column fractions using an antibody against the 24 kDa subunit that is shared by Pol I, II, and III revealed a good correspondence between the presence of the

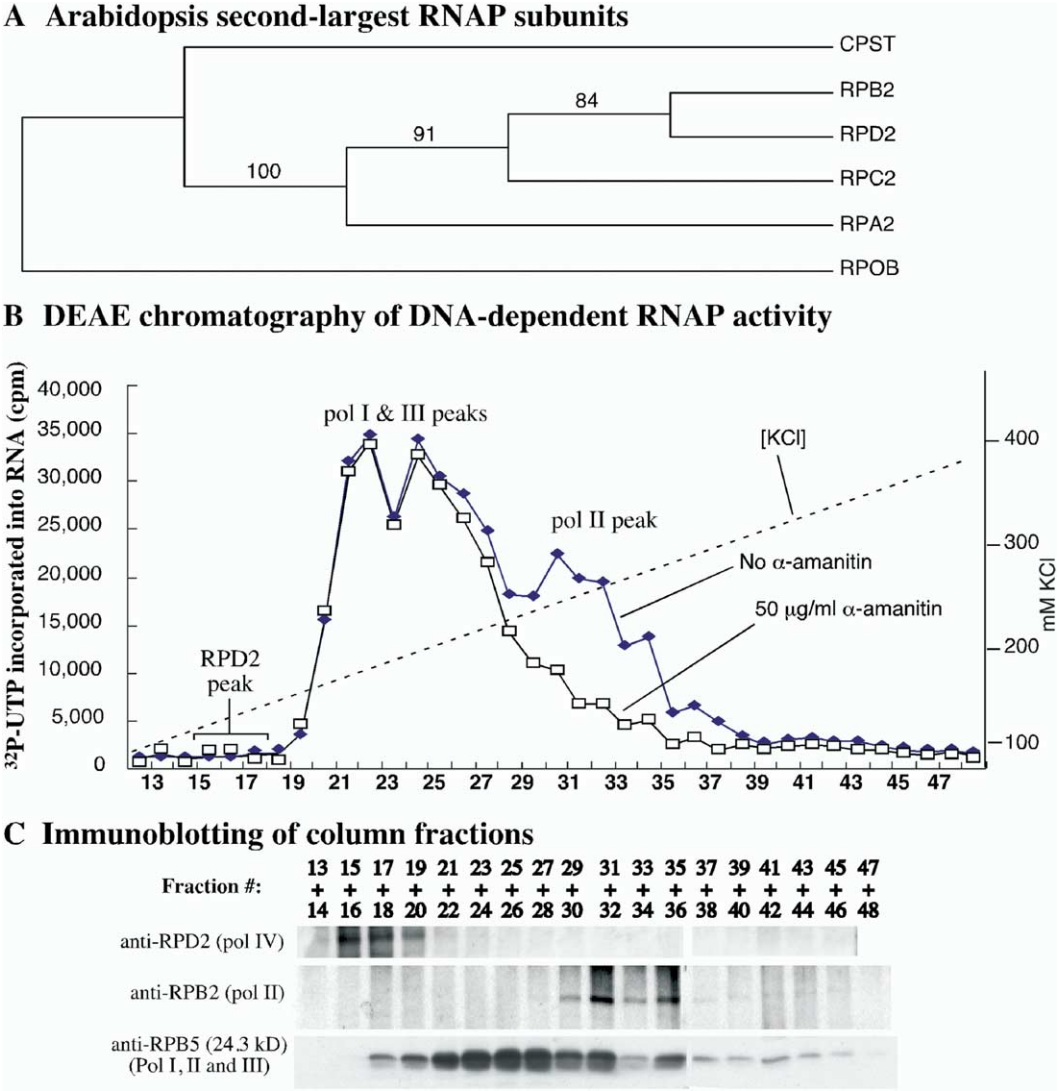

Figure 2. NRPD2 Does Not Cofractionate with Pol II or with DNA-Dependent RNA Polymerase Activity

(A) Neighbor-joining tree (with bootstrap values based on 1000 replications) for second-largest subunits of *Arabidopsis* chloroplast RNAP and RNA polymerases I, II, and III. The *E. coli* RpoB subunit serves as the outgroup.

(B) Fractionation of DNA-dependent RNA polymerase activity by DEAE-Sepharose chromatography. Fractions eluted with a linear KCl gradient were tested for RNA polymerase activity both with and without α-amanitin.

(C) Immunoblot detection of NRPD2, NRPB2, and NRPB5 in fractions eluted from the DEAE column.

Arabidopsis pol IV subunit names are abbreviated from NRPD to RPD in this and all subsequent figures.

24 kDa subunit and RNAP activity. Surprisingly, the peak fractions for NRPD2a displayed no detectable RNAP activity. We conclude that NRPD2 is not an alternative subunit of a conventional DNA-dependent RNA polymerase.

Heterochromatin Association Is Impaired in nrpd2 Mutants

In *nrpd2* mutants, we noted an increased number and decreased size of DAPI-positive heterochromatic foci in interphase nuclei relative to wild-type siblings (Figure 1E), prompting further investigation. Histone H3 dimethylated on lysine 9 ($H3^{dimethyl}K9$) is a marker of heterochromatin (Richards and Elgin, 2002) that colocalizes with chromocenters in wild-type nuclei (Figure 3A). However, in *nrpd2* mutant siblings, the $H3^{dimethyl}K9$ signals are dispersed and colocalize with the numerous, small DAPI-positive foci (Figure 3A; Table S3).

Chromocenters involving NORs are relatively resistant to dispersal (Figure 3B). It is noteworthy that there are four NORs in a diploid nucleus, located at the tips of chromosomes 2 and 4. However, 36% of wild-type and 19% of *nrpd2* interphase nuclei show only two NOR fluorescence in situ hybridization (FISH) signals (as in Figure 3B) due to association of pairs of NORs and their linked centromeres. Nuclei with either three or four NOR FISH signals are also observed in wild-type and *nrpd2* mutants, but only *nrpd2* mutants frequently

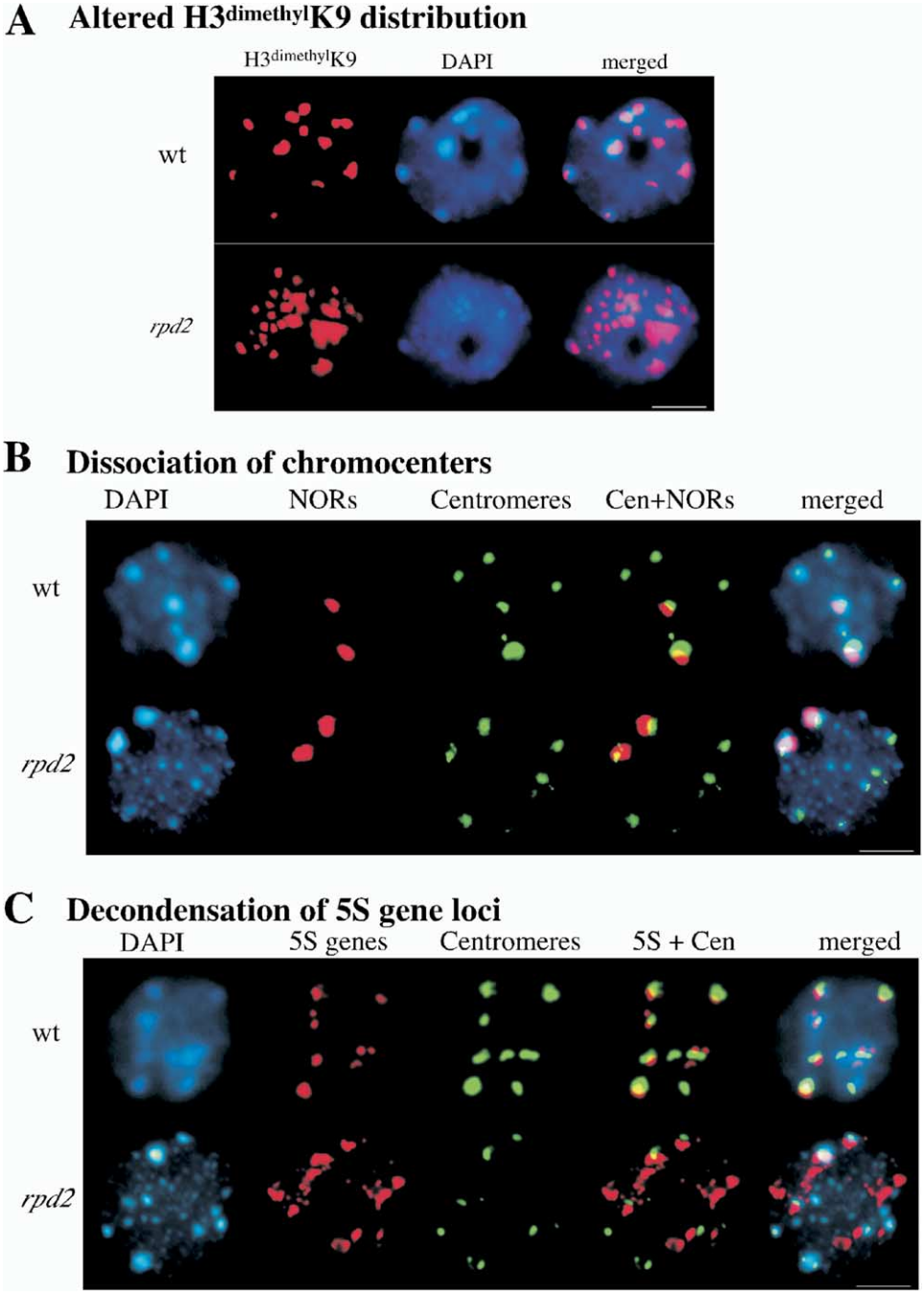

Figure 3. Heterochromatin Is Disrupted in *nrpd2* Mutants

(A) Immunolocalization of histone H3 dimethylated on lysine 9 in interphase cells of wild-type and the *nrpd2a-2 nrpd2b-1* mutant. Chromatin was counterstained with DAPI.

(B) Chromocenters containing NORs are relatively resistant to dispersal in *nrpd2a-2 nrpd2b-1* mutants. Centromeres and NORs (45S rRNA gene loci) were detected by FISH. Chromatin was counterstained with DAPI.

(C) 5S gene loci become decondensed and dissociated from centromeres in *nrpd2a-2 nrpd2b-1*double mutants. 5S genes and centromeres were detected by FISH. Wild-type and mutant plants were progeny of homozygous siblings. Size bars in all panels correspond to 5 μm. Arabidopsis pol IV subunit names are abbreviated from NRPD to RPD in this and all subsequent figures.

(23%) show >4 NOR signals (Table S3), presumably due to dissociation of facultative heterochromatin subdomains of the ~4 Mbp NORs.

5S rRNA gene repeats are tandemly arranged in pericentromeric regions of chromosomes 3, 4, and 5 in *Arabidopsis* ecotype Col-0 such that dual FISH typically reveals substantial overlap of 5S and 180 bp centromere repeat signals in wild-type cells (Figure 3C).

However, in *nrpd2* double mutant siblings, the 5S genes are typically decondensed and show significantly less ($p = 0.0012$) colocalization with centromeres, consistent with the interpretation that pericentromeric facultative heterochromatin is dispersed away from the constitutively heterochromatic centromeres (see Table S3 for quantitation).

Pol IV Participates in the siRNA-Chromatin Modification Pathway

Heterochromatin disruption and 5S gene dispersal in Pol IV mutants suggested a possible loss of cytosine methylation (Soppe et al., 2002). To determine if *nrpd2* or *nrpd1a* mutants affect 5S gene cytosine methylation, we performed Southern blotting using methylation-sensitive restriction endonucleases. HpaII and MspI cut CCGG motifs, but HpaII will not cut if the inner C is methylated, and MspI will not cut if the outer C is methylated (McClelland et al., 1994). HaeIII recognizes GGCC but won't cut if the inner C is methylated. Digestion of 5S genes with these three enzymes reports on methylation at CG (HpaII), CNG (MspI), and CNN (in the ecotype Col-0, the 5S HaeIII site is a CNN site). The Southern blots reveal ladders of bands at ~500 bp intervals (Figure 4A), the size of a 5S gene repeat (Campell et al., 1992). High levels of methylation cause most of the hybridization signal to be near the top of the ladder, whereas loss of methylation results in more signal near the bottom.

5S gene methylation at HpaII, MspI, and HaeIII sites is decreased in *nrpd1a-3* and *nrpd2* mutants (Figure 4A, lanes 3, 5, 18, 20, 22, and 24) relative to their wild-type siblings (lanes 2, 4, 19, 21, 23, and 25), with HaeIII digestion showing the largest effect. Comparison of *nrpd1* and *nrpd2* to the DNA methylation mutants *ddm1*, *met1*, *cmt3*, and *drm1drm2* showed that HpaII digestion of 5S genes in *nrpd1* and *nrpd2* mutants occurred to the same extent as in a *drm1drm2* double mutant (compare lanes 3, 5, and 6) but to a lesser extent than in a *ddm1* (lane 10) or *met1* (lane 11) mutant. DRM2 is responsible for de novo methylation in all sequence contexts (CG, CNG, and CNN); DDM1 is involved in maintenance of methylation in all sequence contexts, and MET1 is primarily responsible for maintenance of CG methylation (reviewed in Bender [2004]). DRM1 has no known function. CMT3 is primarily responsible for maintenance of CNG methylation, so a *CMT3* mutant has little effect on HpaII digestion (lane 7) but has a profound effect on MspI digestion (lane 16). Collectively, the results indicate that Pol IV affects 5S gene methylation in all sequence contexts (CG, CNG, and CNN). Interestingly, the highly methylated 180 bp centromere repeats are unaffected by *nrpd1* and *nrpd2* mutations (Figure 4B), suggesting that Pol IV does not affect global cytosine methylation levels but acts on only a subset of methylated genomic sequences.

Methylation of *AtSN1*, a well-characterized retroelement family (Hamilton et al., 2002; Xie et al., 2004), was assayed using HaeIII digestion followed by PCR (Figure 4C) (Hamilton et al., 2002). If HaeIII sites are methylated, the DNA is not cut and can be amplified. However, if CNN methylation is lost at any of three HaeIII sites (see diagram), HaeIII digestion precludes PCR amplification. In wild-type Col-0, Ler, or Ws (the genetic backgrounds for the mutants tested), *AtSN1* elements are heavily methylated and resistant to HaeIII cleavage. Methylation is unaffected by *met1* or *cmt3* mutants but is substantially reduced in a *drm1 drm2* double mutant, as expected for CNN methylation. HaeIII methylation is also disrupted in mutants of the heterochromatic siRNA pathway, including *rdr2* (RNA-dependent RNA polymerase 2), *hen1* (Hua enhancer 1), or *dcl3* (Dicer-like 3), consistent with published results (Xie et al., 2004). By contrast, *AtSN1* methylation is not diminished in a mutant of *DCL1*, the dicer responsible for miRNA production. Importantly, *AtSN1* methylation is also reduced in both *nrpd1* and *nrpd2* mutants. The loss of *AtSN1* methylation in both siRNA pathway mutants and *nrpd* mutants suggests that Pol IV might also affect siRNAs. Consistent with this hypothesis, 5S gene and *AtSN1* siRNAs are significantly reduced or eliminated in *nrpd2* and *nrpd1* mutants (Figures 4D and 4E) as in *hen1*, *rdr2*, *drm*, or *ago4* mutants, confirming prior studies (Herr et al., 2005; Xie et al., 2004; Zilberman et al., 2004). By contrast, mutations of the RNA-dependent RNA polymerases *rdr1* or *rdr6* (*sgs2*, also known as *sde1*) had no effect, though *rdr6* is known to function in RNA silencing of transgenes (Baulcombe, 2004). Interestingly, 5S siRNA levels were actually increased in *ddm1* and *met1* mutants (Figure 4D), indicating that disrupted maintenance of cytosine methylation is not the explanation for loss of 5S siRNAs in *nrpd1* and *nrpd2* mutants.

Importantly, miRNA levels are unaffected in *nrpd* mutants, as shown by comparison of miR163, 159, 164, 171, and 172 levels in mutant and wild-type siblings (Figure 4F), indicating that Pol IV acts only in the siRNA pathway and not in the miRNA pathway.

Discussion

Loss of NRPD1 or NRPD2 function causes the loss of cytosine methylation at pericentromeric 5S genes and *AtSN1* retroelements yet has no discernible effect on centromere repeat methylation. These observations suggest that Pol IV primarily affects facultative heterochromatin rather than constitutive heterochromatin, consistent with the localization of NRPD2 at foci that overlap or are adjacent to chromocenters but are not fully coincident with chromocenters. We propose that Pol IV acts on genes that cycle between decondensed, euchromatic states and condensed, chromocenter-associated heterochromatic states, playing a key role in the amplification of siRNAs that direct cytosine methylation to these genes when they become activated (Aufsatz et al., 2002; Wassenegger, 2000).

Interestingly, the total amount of $H3^{dimethyl}K9$, a reliable marker of heterochromatin, does not appear to be reduced in Pol IV mutant nuclei. Instead, the $H3^{dimethyl}K9$ is simply dispersed into a larger number of heterochromatic foci. Collectively, these data, combined with data showing disruption of chromocenters in *ddm1* and *met1* mutants (Soppe et al., 2002), suggest that loss of cytosine methylation from either pericentromeric repeats or centromeric repeats is sufficient to disrupt

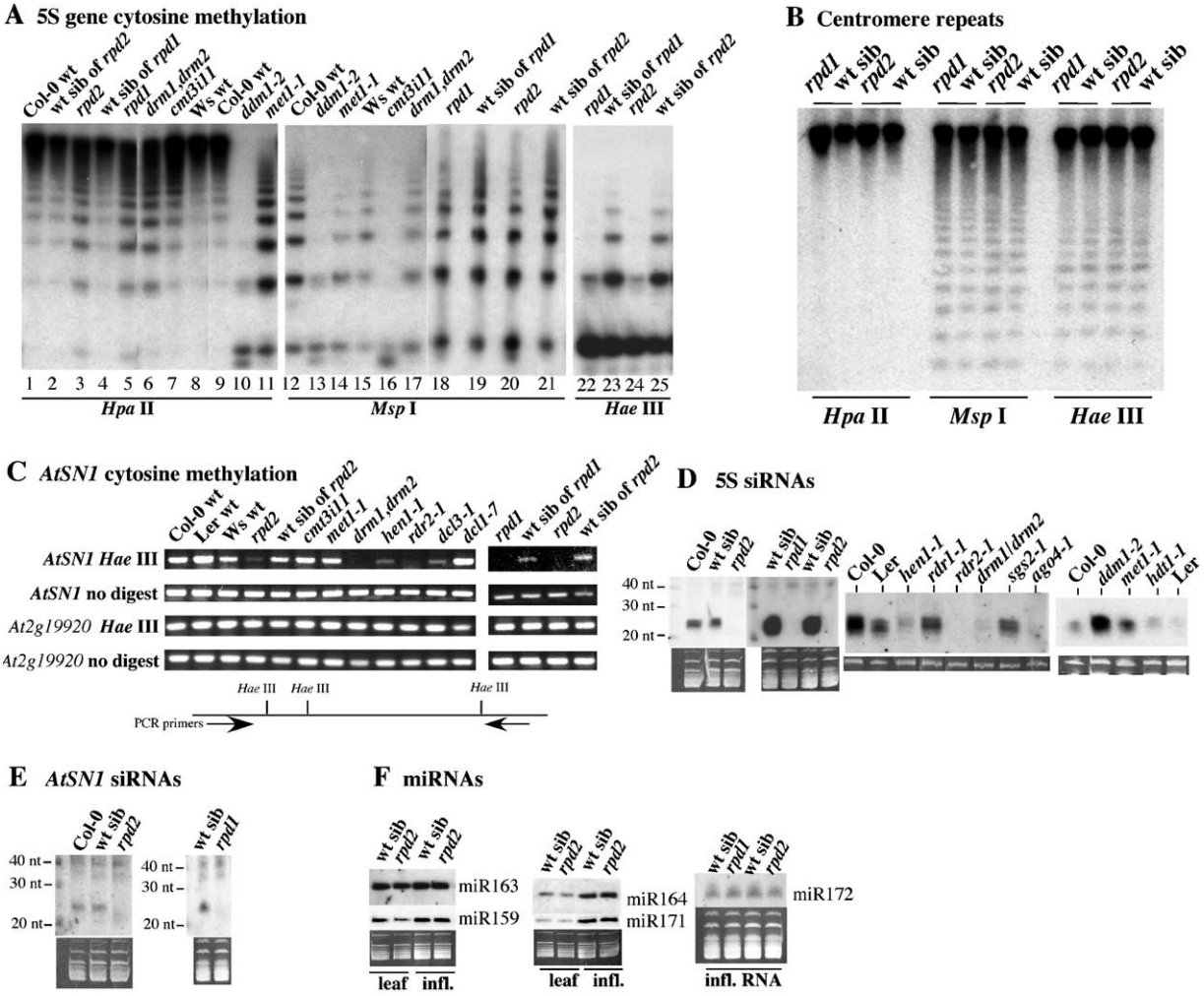

Figure 4. NRPD1 and NRPD2 Are Required for 5S Gene and *AtSN1* Cytosine Methylation and siRNA Accumulation

(A) Analysis of 5S gene repeats in *nrpd1a-3* and *nrpd2a-2 nrpd2b-1* double mutants relative to wild-type siblings and methylation mutants. Genomic DNA digested with HpaII, MspI, or HaeIII was hybridized to a 5S gene probe. *nrpd1*, *nrpd2*, *ddm1*, and *met1* mutants are in the Col-0 genetic background; *drm1drm2* and *cmt3* are in the WS background.

(B) Methylation of 180 bp centromere repeats is apparently unaffected in *nrpd1* and *nrpd2* mutants relative to wild-type siblings.

(C) *nrpd1* and *nrpd2* mutations cause decreased *AtSN1* cytosine methylation. PCR was used to amplify a portion of an *AtSN1* retroelement that includes three HaeIII sites. Undigested DNA and a gene lacking HaeIII sites served as PCR controls.

(D) 5S siRNAs in *nrpd1*, *nrpd2*, and mutants affecting siRNA production. Small RNA blots were probed for 5S siRNA sequences. Ethidium-stained gel bands serve as loading controls. The *hdt1* mutant is an ecotype Col-0 line with a T-DNA insertion in a nucleolar histone deacetylase; it serves as a T-DNA control in the blot at far right.

(E) *AtSN1* siRNAs are reduced or eliminated in *nrpd1* and *nrpd2* mutants.

(F) miRNAs 159, 163, 164, and 171 are unaffected in *nrpd1* and *nrpd2* mutants.

Arabidopsis pol IV subunit names are abbreviated from NRPD to RPD in this and all subsequent figures.

higher-order heterochromatin association into chromocenters. One possibility is that methylcytosine binding domain proteins and/or their associated proteins might act as linkers or bridges that help bring together dispersed heterochromatin domains.

At 5S genes, Pol IV affects cytosine methylation in all sequence contexts (CG, CNG, and CNN). Importantly, CG, CNG, and CNN de novo methylation is accomplished by DRM methyltransferase activity (Cao et al., 2003; Cao and Jacobsen, 2002). DRM is also responsible for siRNA-directed DNA methylation (in all sequence contexts) in *Arabidopsis* (Cao et al., 2003). We have shown that Pol IV and DRM activities are both needed for CNN methylation at *AtSN1* retroelements, as are genes of the siRNA pathway. These facts, combined with our demonstration that 5S and *AtSN1* siRNAs are essentially eliminated in Pol IV mutants, are most parsimonious with the hypothesis that Pol IV is involved in production of siRNAs that guide DRM-mediated cytosine methylation to repeated sequences complementary to the siRNAs (Chan et al., 2004). This would explain why loss of cytosine methylation in Pol IV mutants is most apparent at CNN (HaeIII in our experiments) sites, which would be dependent on continuous de novo

methylation due to the lack of a dedicated CNN maintenance methyltransferase (reviewed in Bender [2004]). By contrast, preexisting methylation at CG and CNG sites would be perpetuated by the MET1 and CMT3 maintenance methyltransferases, explaining the lesser effect of Pol IV or *drm* mutations on HpaII and MspI-sensitive 5S gene methylation (Figure 4A).

One could argue that DNA methylation is upstream of siRNA production, as suggested by the decrease in *AtSN1* siRNAs in *ddm1* and *met1* mutants (Lippman et al., 2003). However, this hypothesis does not fit with the fact that *ddm1* and *met1* cause dramatic decreases in 5S gene methylation yet actually increase 5S siRNA levels, possibly due to derepression of silenced 5S genes, thereby increasing the number of transcripts from which to generate dsRNAs and siRNAs. By contrast, Pol IV and *drm* mutations cause only modest decreases in total methylation yet essentially eliminate 5S siRNAs.

So how can loss of de novo methylation in a *drm* mutant eliminate siRNAs (Figure 4D) if siRNAs are upstream of de novo methylation? This apparent paradox might be explained if initial, primary siRNAs direct de novo methylation events that then trigger a massive amplification of siRNAs, and more extensive methylation, by a mechanism requiring Pol IV. Presumably, it is this second wave that yields the high levels of siRNAs and methylation that we detect. One possibility is that methylated DNA serves as the template for Pol IV-mediated transcription of aberrant RNAs. Another possibility is that methylation stalls elongating polymerases, as suggested by studies in *Neurospora* (Rountree and Selker, 1997), providing RDR2 with an opportunity to make dsRNAs from incomplete transcripts and leading to local production of aberrant RNAs or siRNAs that prime Pol IV transcription. Testing such hypotheses will be priorities for future studies.

Experimental Procedures

Plant Strains

Arabidopsis mutants *hen1-1*, *rdr2-1*, *dcl3-1*, and *dcl1-7* were provided by Jim Carrington. *met1-1* was provided by Eric Richards. *cmt3i11* was provided by Judith Bender. *sgs2-1* (alias *sde1*; *rdr6*) was provided by Herve Vaucheret. Salk T-DNA insertion lines and other mutants were obtained from the *Arabidopsis* Biological Resource Center (ABRC).

RNA and Immunoblot Analysis of NRPD2

RNA was isolated as described previously (Chen et al., 1998). RNA blots were hybridized to a probe generated by random priming of the *NRPD2a* 5′ RACE cDNA product using standard methods (Sambrook and Russell, 2001). For immunoblotting, plant tissue was homogenized in SDS sample buffer (125 mM Tris-HCl [pH 6.8], 2% SDS, 10% glycerol, and 0.7 M β-mercaptoethanol) and 40 μg of protein, determined using a BCA (bicinchoninic acid) protein assay kit (PIERCE), subjected to SDS-PAGE on a 7.5% gel, and electroblotted to a PVDF membrane. Anti-NRPD2 and anti-NRPB2 antisera were raised in rabbits against peptides DMDIDVKDLEEFEA and MEYNEYEPEEPQYVE of NRPD2a (At3g23780) and *A. thaliana* NRPB2 (At4g21710), respectively. Anti-Pol I+II+III rabbit antiserum was raised against peptide GDKFSSRHGQKG, which is conserved in Pol I, II, and III second-largest subunits. Sera were affinity purified using peptides covalently linked to NHS-activated Sepharose resin (Pharmacia Biotech). Columns were washed with 3–5 column volumes of PBS (pH 7.0), 0.05% Tween-20; antibodies were eluted using 0.1 M glycine-HCl (pH 3.0) neutralized by addition of Tris-HCl (pH 8.0) and stored at −80°C. Antisera were diluted 1:250 for probing immunoblots. The secondary antibody, diluted 1:5000, was peroxidase-linked donkey anti-rabbit IgG (Amersham). Immunoblots were visualized by chemiluminescence (ECL Western Blotting Detection kit; Amersham).

Screening of T-DNA Knockout Lines

T-DNA insertions in *NRPD2a*, *NRPD2b*, and *NRPD1a* were verified by PCR and sequencing using a T-DNA left border primer (5′-CGTCCGCAATGTGTTATTAAG-3′) and primers specific for *NRPD2a*, *NRPD2b*, or *NRPD1a* as suggested by the suppliers of the Salk lines. Screening by Southern blot analysis was according to standard methods (Sambrook and Russell, 2001).

Anion Chromatography and DNA-Dependent RNA Polymerase Assay

Arabidopsis plants were grown for 10 days at 25°C in 3 liter flasks containing 1 liter of liquid 1× Gamborg B5 medium, 1× Gamborg vitamins (Sigma), and 2% sucrose shaken at moderate speed. Tissue (200 g) was homogenized, and crude nuclear proteins were fractionated by DEAE-Sepharose chromatography and tested for RNA polymerase activity as described previously (Saez-Vasquez and Pikaard, 1997).

Phylogenetic Analyses

RNAP subunits were identified by blastp searches using *E. coli* RPOC and RPOB, *S. cerevisiae* RPB1 and RPB2, and *A. thaliana* NRPD1a and NRPD2a protein sequences. Sequences were aligned, using Clustal X (version 1.81). Conserved sequences were highlighted using BOXSHADE. (http://bioweb.pasteur.fr/seqanal/interfaces/boxshade.html). Phylogenetic analysis was by the neighbor-joining method, with 1000 bootstrap replications, using PAUP (version 4.0b10).

Cytosine Methylation Assays

Genomic DNA (100 ng) was digested with HpaII, MspI, or HaeIII. Following agarose gel electrophoresis, DNA was blotted to uncharged nylon membranes. Probes were generated by random priming, and blots were hybridized using standard methods (Sambrook and Russell, 2001).

AtSN1 methylation assays used ~100 ng of DNA digested with HaeIII (or undigested for controls). Approximately 5% of digestion reaction DNA was then used for each PCR reaction. PCR conditions were 2 min at 94°C, followed by 35 cycles of 94°C for 30 s, 53°C for 30 s, and 72°C for 30 s. Primer sequences for *AtSN1* were the following: 5′-ACTTAATTAGCACTCAAATTAAACAAAATAAGT-3′ and 5′-TTTAAACATAAGAAGAAGTTCCTTTTTCATCTAC-3′. The *At2g19920* control was amplified using 5′-TCACCCGAACAGTTGGAAGAA GAG-3′ and 5′-GTGAGGAACCGGTCCATTATTGCT-3′. PCR products were subjected to agarose gel electrophoresis.

In Situ Hybridization and Immunolocalization

Emerging leaves of 21-day-old plants were fixed in ethanol:acetic acid (3:1, v/v). Nuclei were prepared as described (Schwarzacher and Mosgoeller, 2000). FISH using biotin-dUTP or digoxygenin-dUTP labeled 180 bp *A. thaliana* pericentromeric repeat, 5S gene or 45S rRNA gene intergenic spacer sequence probes was as described previously (Pontes et al., 2004).

For immunolocalization experiments, nuclei were fixed in 4% paraformaldehyde. $H3^{dimethyl}K9$ was localized using published methods (Houben et al., 1996) with antibody purchased from Upstate Biotechnology. For NRPD2, slides were permeabilized with 10% DMSO, 3% NP-40 in PBS, before blocking with 1% BSA in PBS. Primary antibodies were diluted 1:100 in PBS, 1% BSA, and slides were incubated overnight at 4°C. Secondary antibodies were conjugated to rhodamine or fluorescein (Sigma). Chromatin was counterstained with DAPI in antifade buffer (Vector Laboratories). Nuclei were examined using a Nikon Eclipse E600 epifluorescence microscope and images collected using a Q-Imaging Retiga EX digital camera.

siRNA and miRNA Detection

RNA was isolated using the mirVana miRNA isolation kit (Ambion). RNA (2–6 μg) was resolved by denaturing polyacrylamide gel electrophoresis on a 20% (w/v) gel. Gels were electroblotted (20 mA/cm^2 for 2 hr) to Magnacharge nylon membranes (0.22 μm; Osmonics) using a semidry transfer apparatus. An end-labeled RNA ladder was used as a molecular weight marker (Decade Marker System, Ambion). The AtSN1 riboprobe was synthesized from a NdeI-linearized plasmid DNA template (Zilberman et al., 2003). All other riboprobes were generated according to the mirVana probe construction kit (Ambion) using oligonucleotides specific for a given small RNA and labeling by T7 polymerase transcription in the presence of α-^{32}P CTP. DNA oligonucleotides for 5S and miRNA probes were the following: siR1003T7 (5S) (5′-AGACCGTGAGGCCAAACTTGG CATcctgtctc-3′; small letters are complementary to the T7 promoter oligonucleotide), miR159T7 (5′-TTTGGATTGAAGGGAGCTC TAcctgtctc-3′), miR163T7 (5′-TTGAAGAGGACTTGGAACTTCGAT cctgtctc-3′), and miR164T7 (5′-TGGAGAAGCAGGGCACGTGCA cctgtctc-3′). Unincorporated nucleotides were removed using Performa DTR Gel Filtration Cartridges (Edgebiosystems). Blot hybridization was in 50% formamide, 0.25 M Na_2HPO_4 (pH 7.2), 0.25 M NaCl, 7% SDS at 42°C (14–16 hr) followed by two 15 min washes at 37°C in 2× SSC, two 15 min washes at 37°C in 2× SSC, 0.1% SDS, and a 10 min wash in 0.5× SSC, 1% SDS.

Supplemental Data

Supplemental Data include five figures, three tables, Supplemental Experimental Procedures, and Supplemental References and can be found with this article online at http://www.cell.com/cgi/content/full/120/5/613/DC1/.

Acknowledgments

We are indebted to David Baulcombe for sharing his lab's unpublished evidence that *sde4* is *nrpd1a*. We thank the Carrington, Richards, Vaucheret, and Bender labs for mutant seed stocks; the Carrington lab for cloned *AtSN1*; Thomas Guilfoyle for the anti-RPB5 antibody; M.E. Gifford for phylogenetic analysis advice; Douglas Chalker for use of his fluorescence microscope; and our labmate Keith Earley for RNAP assay advice. We thank Dr. Wanda Viegas for comentoring P.C.N. This research was supported by NIH grant R01-GM60380 and USDA grant 99-35301-7865 (to C.S.P.). P.C.N. was supported by a predoctoral fellowship from the Portuguese Fundação para a Ciência e Tecnologia (SFRH/BD/6520/2001) and the Luso-American Foundation (grant 237/2004).

Received: December 23, 2004
Revised: January 24, 2005
Accepted: February 4, 2005
Published online: February 10, 2005

References

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature *408*, 796–815.

Archambault, J., and Friesen, J.D. (1993). Genetics of eukaryotic RNA polymerases I, II, and III. Microbiol. Rev. *57*, 703–724.

Aufsatz, W., Mette, M.F., van der Winden, J., Matzke, A.J., and Matzke, M. (2002). RNA-directed DNA methylation in *Arabidopsis*. Proc. Natl. Acad. Sci. USA Suppl. *99*, 16499–16506.

Baulcombe, D. (2004). RNA silencing in plants. Nature *431*, 356–363.

Bender, J. (2004). Chromatin-based silencing mechanisms. Curr. Opin. Plant Biol. *7*, 521–526.

Campell, B.R., Song, Y., Posch, T.E., Cullis, C.A., and Town, C.D. (1992). Sequence and organization of 5S ribosomal RNA-encoding genes of *Arabidopsis thaliana*. Gene *112*, 225–228.

Cao, X., and Jacobsen, S.E. (2002). Role of the *Arabidopsis* DRM methyltransferases in de novo DNA methylation and gene silencing. Curr. Biol. *12*, 1138–1144.

Cao, X., Aufsatz, W., Zilberman, D., Mette, M.F., Huang, M.S., Matzke, M., and Jacobsen, S.E. (2003). Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. Curr. Biol. *13*, 2212–2217.

Chan, S.W., Zilberman, D., Xie, Z., Johansen, L.K., Carrington, J.C., and Jacobsen, S.E. (2004). RNA silencing genes control de novo DNA methylation. Science *303*, 1336.

Chen, Z.J., Comai, L., and Pikaard, C.S. (1998). Gene dosage and stochastic effects determine the severity and direction of uniparental rRNA gene silencing (nucleolar dominance) in *Arabidopsis* allopolyploids. Proc. Natl. Acad. Sci. USA *95*, 14891–14896.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R., and Kornberg, R.D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. Science *288*, 640–649.

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 Å resolution. Science *292*, 1863–1876.

Fransz, P., De Jong, J.H., Lysak, M., Castiglione, M.R., and Schubert, I. (2002). Interphase chromosomes in *Arabidopsis* are organized as well defined chromocenters from which euchromatin loops emanate. Proc. Natl. Acad. Sci. USA *99*, 14584–14589.

Grummt, I. (2003). Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. Genes Dev. *17*, 1691–1702.

Hamilton, A., Voinnet, O., Chappell, L., and Baulcombe, D. (2002). Two classes of short interfering RNA in RNA silencing. EMBO J. *21*, 4671–4679.

Herr, A., Jensen, M.B., Dalmay, T., and Baulcombe, D.C. (2005). RNA polymerase IV directs silencing of endogenous DNA. Science, in press. Published online February 3, 2005. 10.1126/science.1106910.

Houben, A., Belyaev, N.D., Turner, B.M., and Schubert, I. (1996). Differential immunostaining of plant chromosomes by antibodies recognizing acetylated histone H4 variants. Chromosome Res. *4*, 191–194.

Larkin, R.M., Hagen, G., and Guilfoyle, T.J. (1999). *Arabidopsis thaliana* RNA polymerase II subunits related to yeast and human RPB5. Gene *231*, 41–47.

Lippman, Z., May, B., Yordan, C., Singer, T., and Martienssen, R. (2003). Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. PLoS Biol. *1(3)* e67 DOI: 10.1371/journal.pbio.0000067.

McClelland, M., Nelson, M., and Raschke, E. (1994). Effect of site-specific methylation on restriction endonucleases and DNA modification methyltransferases. Nucleic Acids Res. *22*, 3640–3659.

Pontes, O., Neves, N., Silva, M., Lewis, M.S., Madlung, A., Comai, L., Viegas, W., and Pikaard, C.S. (2004). Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis* suecica genome. Proc. Natl. Acad. Sci. USA *101*, 18240–18245.

Richards, E.J., and Elgin, S.C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell *108*, 489–500.

Rountree, M.R., and Selker, E.U. (1997). DNA methylation inhibits elongation but not initiation of transcription in *Neurospora crassa*. Genes Dev. *11*, 2383–2395.

Saez-Vasquez, J., and Pikaard, C.S. (1997). Extensive purification of a putative RNA polymerase I holoenzyme from plants that accurately initiates rRNA gene transcription in vitro. Proc. Natl. Acad. Sci. USA *94*, 11869–11874.

Saez-Vasquez, J., and Pikaard, C.S. (2000). RNA polymerase I holoenzyme-promoter interactions. J. Biol. Chem. *275*, 37173–37180.

Sambrook, J., and Russell, D.R. (2001). Molecular Cloning: A Laboratory Manual, Third Edition (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Schramm, L., and Hernandez, N. (2002). Recruitment of RNA polymerase III to its target promoters. Genes Dev. *16*, 2593–2620.

Schwartz, L.B., and Roeder, R.G. (1974). Purification and subunit structure of deoxyribonucleic acid-dependent ribonucleic acid polymerase I from mouse myeloma, MOPC 315. J. Biol. Chem. *249*, 5898–5906.

Schwarzacher, H.G., and Mosgoeller, W. (2000). Ribosome biogenesis in man: current views on nucleolar structures and function. Cytogenet. Cell Genet. *91*, 243–252.

Soppe, W.J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M.S., Jacobsen, S.E., Schubert, I., and Fransz, P.F. (2002). DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in *Arabidopsis*. EMBO J. *21*, 6549–6559.

Wassenegger, M. (2000). RNA-directed DNA methylation. Plant Mol. Biol. *43*, 203–220.

Woychik, N.A., and Hampsey, M. (2002). The RNA polymerase II machinery: structure illuminates function. Cell *108*, 453–463.

Xie, Z., Johansen, L.K., Gustafson, A.M., Kassachau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., and Carrington, J.C. (2004). Genetic and functional diversification of small RNA pathways in plants. PLoS Biol. *2(5)*:, e104 DOI: 10.1371/journal.pbio.0020104.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. Cell *98*, 811–824.

Zilberman, D., Cao, X., and Jacobsen, S.E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science *299*, 716–719.

Zilberman, D., Cao, X., Johansen, L.K., Xie, Z., Carrington, J.C., and Jacobsen, S.E. (2004). Role of *Arabidopsis* ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats. Curr. Biol. *14*, 1214–1220.

Accession Numbers

The GenBank accession number for the NRPD2a mRNA sequence determined for this paper is AY862891.

Note Added in Proof

In the early online version of the article, the genes NRPD1a, NRPD1b, NRPD2a, and NRPD2b were named RPD1a, RPD1b, RPD1a, and RPD2b, respectively. We have changed the names due to a nomenclature conflict.