

# Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*

Dominique Pontier,<sup>1,5</sup> Galina Yahubyan,<sup>1,5,6</sup> Danielle Vega,<sup>1</sup> Agnès Bulski,<sup>2</sup> Julio Saez-Vasquez,<sup>1</sup> Mohamed-Ali Hakimi,<sup>3</sup> Silva Lerbs-Mache,<sup>4</sup> Vincent Colot,<sup>2</sup> and Thierry Lagrange<sup>1,7</sup>

<sup>1</sup>LGDP, UMR 5096, Université de Perpignan, 66860 Perpignan Cedex, France; <sup>2</sup>Unité de Recherche en Génomique Végétale (URGV), INRA/CNRS/UEVE, 91057 Evry Cedex, France; <sup>3</sup>Institut Jean Roget, 38000 Grenoble, France; <sup>4</sup>PDC, UMR 5575, Université Joseph Fourier, 38000 Grenoble, France

Recent genetic and biochemical studies have revealed the existence in plants of a fourth RNA polymerase, RNAPIV, which mediates siRNA accumulation and DNA methylation-dependent silencing of endogenous repeated sequences. Here, we show that *Arabidopsis* expresses, in fact, two evolutionarily related forms of RNAPIV, hereafter referred to as RNAPIVa and RNAPIVb. These two forms contain the same second-largest subunit (NRPD2), but differ at least by their largest subunit, termed NRPD1a and NRPD1b. Unlike NRPD1a, NRPD1b possesses a reiterated CTD, a feature that also characterizes the largest subunit of RNAPII. Our data indicate that RNAPIVb is the most abundant form of RNAPIV in *Arabidopsis*. Selective disruption of either form of RNAPIV indicates that RNAPIVa-dependent siRNA accumulation is not sufficient per se to drive robust silencing at endogenous loci and that high levels of DNA methylation and silencing depend on siRNA that are accumulated through a pathway involving the concerted action of both RNAPIV forms. Taken together, our results imply the existence of a novel two-step mechanism in siRNA synthesis at highly methylated loci, with RNAPIVb being an essential component of a self-reinforcing loop coupling *de novo* DNA methylation to siRNA production.

[*Keywords*: RNAPIV; reiterated CTD; siRNA; DNA methylation; silencing]

Supplemental material is available at <http://www.genesdev.org>.

Received April 25, 2005; revised version accepted June 29, 2005.

A major evolutionary distinction separating prokaryotes from eukaryotes is the passage from a unique multisubunit DNA-dependent RNA polymerase enzyme (RNAP) to three complexes (Roeder and Rutter 1969), each responsible for the transcription of a subclass of nuclear DNA sequences (Sentenac 1985). RNA polymerase I (RNAPI) transcribes the repeated genes encoding the large ribosomal RNAs, which represent up to four-fifths of total RNA. RNA polymerase II (RNAPII) transcribes all of the cell protein-coding messenger RNAs (mRNAs) as well as some small nuclear RNAs (snRNAs). RNA polymerase III (RNAPIII) is dedicated to the transcription of a collection of genes whose main common feature is that they encode structural or catalytic RNAs (tRNAs,

5S RNA, snRNA) that are components of protein synthesis, splicing, and tRNA processing apparatuses. It is believed that this triplication event provided the eukaryotic cell with a greater flexibility toward energy-consuming cellular functions such as ribosome synthesis, as well as with more sophisticated means for the regulation of gene expression.

Prokaryotic and eukaryotic RNA polymerases are multisubunit enzymes that are evolutionarily related to each other through their largest and second-largest subunits (Ebright 2000; Cramer 2002). The largest subunit ( $\approx 160$  kDa:  $\beta'$ ; A; RPA1; RPB1; RPC1) contains eight regions conserved in order and sequence (A to H), while the second-largest subunit ( $\approx 150$  kDa:  $\beta$ ; B; RPA2; RPB2; RPC2), contains nine such regions (A to I) (Allison et al. 1985; Sweetser et al. 1987). Although the role of these conserved domains is not yet fully understood, the structure determination of RNAPII suggests that they cooperate in the formation of a single fold cleft containing the active site of the enzyme (Cramer et al. 2001).

<sup>5</sup>These authors contributed equally to this work.

<sup>6</sup>Present address: Department of Plant Physiology and Molecular Biology, University of Plovdiv, Plovdiv-4000, Bulgaria.

<sup>7</sup>Corresponding author.

E-MAIL [lagrange@univ-perp.fr](mailto:lagrange@univ-perp.fr); FAX 0033-0033468668499.

Article and publication are at <http://www.genesdev.org/cgi/doi/10.1101/gad.348405>.

A specific feature of RNAPII is the presence at the C-terminal end of its largest subunit (RPB1) of variable numbers of a conserved heptapeptide repeat motif (consensus YSPTSPS), which form the so-called C-terminal domain (CTD) (Corden 1990; Young 1991). In vivo and in vitro studies indicated that this CTD is essential for cell viability and transcriptional activation and that it undergoes a cycle of extensive phosphorylation and dephosphorylation through each round of transcription (Dahmus 1996; Carlson 1997; Greenblatt 1997). Indeed, depending on its phosphorylation state, the CTD recruits essential proteins, which help regulate the many steps of transcription from initiation to termination (Greenblatt 1997; Otero et al. 1999; Kim et al. 2004), as well as chromatin modification (Hampsey and Reinberg 2003), and pre-mRNA processing (Bentley 1999; Hirose and Manley 2000). The acquisition of the CTD by RNAPII can therefore be considered as an important step in the evolution of the multiple levels of transcriptional and post-transcriptional regulation that are the hallmark of eukaryotic organisms.

An unexpected outcome of the *Arabidopsis* genome sequence analysis was the identification of genetic information encoding novel large subunits of a putative fourth type of eukaryotic RNAP (*Arabidopsis* Genome Initiative 2000). Specifically, this information consisted in two pairs of related genes potentially coding for the largest and second-largest subunits. Recent genetic data have implicated the product of two of these genes (*NRPD1a* and *NRPD2a*) in the silencing of transposable elements and other repeats via short-interfering RNAs (siRNAs) (Herr et al. 2005; Onodera et al. 2005). In the present work, we have extended this analysis by characterizing the full set of potential RNAPIV largest and second-largest subunit genes. We show that *Arabidopsis* expresses, in fact, two forms of RNAPIV, RNAPIVa and RNAPIVb, which contain the same second-largest subunit (NRPD2) but differ at least by the nature of their largest subunits. Furthermore, we show that NRPD1b possesses a reiterated CTD, unlike NRPD1a, and that RNAPIVb is the most abundant form of RNAPIV in *Arabidopsis*. Selective disruption of RNAPIVa and RNAPIVb indicates that efficient silencing at transposons and highly repeated sequences requires the concerted action of both RNAPIV forms, while a basal level of silencing at low repetitive DNA is only dependent on RNAPIVa. Taken together, our results indicate that a self-reinforcing silencing pathway based on the nonredundant action of two different RNAPIV forms, probably via a novel two-step mechanism in siRNA synthesis, is required to maintain a high level of methylation at transposons and highly repeated sequences.

## Results

### *Arabidopsis* expresses two class IV largest subunit genes

We have initiated the molecular characterization and the functional analysis of the class IV largest and second-

largest subunit genes present in the *Arabidopsis* genome, namely, At1g63020, At2g40030, and At3g23780, At3g18090. Using a combination of PCRs and cDNA library screens, full-length cDNA sequences were obtained for these genes, indicating that they are all transcribed in *Arabidopsis*. The corresponding cDNA sequences have been deposited in the GenBank database and recently renamed according to the revised nomenclature (Herr et al. 2005; Onodera et al. 2005) as *AtNRPD1a* (accession number AY826515), *AtNRPD1b* (AY826516), and *AtNRPD2a* (AY935711), *AtNRPD2b* (AY935712) (Fig. 1A; Supplementary Figs. S1A, S2A). While producing a full-length RNA with the same exon boundaries as *NRPD2a*, *NRPD2b* is unlikely to encode a functional protein due to a premature stop codon in the first coding exon of the full-length cDNA, and it is therefore likely to be an expressed pseudogene (Supplementary Fig. S2B). Accordingly, we refer to the 1172-amino-acid product of the *NRPD2a* gene as NRPD2. NRPD1a and NRPD2 correspond to the largest and second-largest subunits of the recently identified RNAPIV enzyme (Herr et al. 2005; Onodera et al. 2005). Comparison of the full-length cDNA sequence of *AtNRPD1b* with that of the genome sequence revealed that it spans, in fact, the two misannotated genes At2g40030 and At2g40040 (Fig. 1A). Thus, *AtNRPD1b* contains 17 exons and encodes a 1976-amino-acid polypeptide with a molecular mass of 235 kDa (AtNRPD1b) (Fig. 1A). RT-PCR analysis indicates that *AtNRPD1b* as well as *AtNRPD1a* are expressed in all tissues tested (Supplementary Fig. S1B).

As outlined in Figure 1A, AtNRPD1b presents sequence similarities to AtRPB1 and AtNRPD1a in the evolutionary conserved regions A through H that are conserved between all known largest RNAP subunits. Except for the B and C regions, AtNRPD1a and AtNRPD1b proteins are more similar to each other than to AtRPB1, suggesting that they originate from a common ancestor (Fig. 1A; Supplementary Fig. S3A). Of note, region G is located near the conserved region F in both AtNRPD1a and b, which suggests that these two proteins are missing the structural element known as the foot domain in the RNAPII structure (Cramer et al. 2001). Library screening and database searches indicate that *NRPD1b* genes are also present in other plants, including rice (*OsNRPD1b*; OsAP004365.3) and spinach (*SoNRPD1b*; accession number AY826517). As expected for bona fide orthologs, these proteins share a significant level of sequence identity throughout their RPB1-like region, with values ranging from 55% to 84% (Fig. 1A). A striking difference between NRPD1s and the other large subunits of RNAP is the amino acid substitutions observed at the first position (for the NRPD1b proteins) and at the first two positions (for the NRPD1a proteins) of the invariant NADFDGD motif found in the conserved region D (Supplementary Fig. S3B). Although these two positions are not directly engaged in the coordination of the Mg<sup>2+</sup> ions that participate in catalysis and can tolerate conservative substitutions (Dieci et al. 1995), the structure of the RNAPII elongation complex has recently suggested a possible role for the first residue N in the

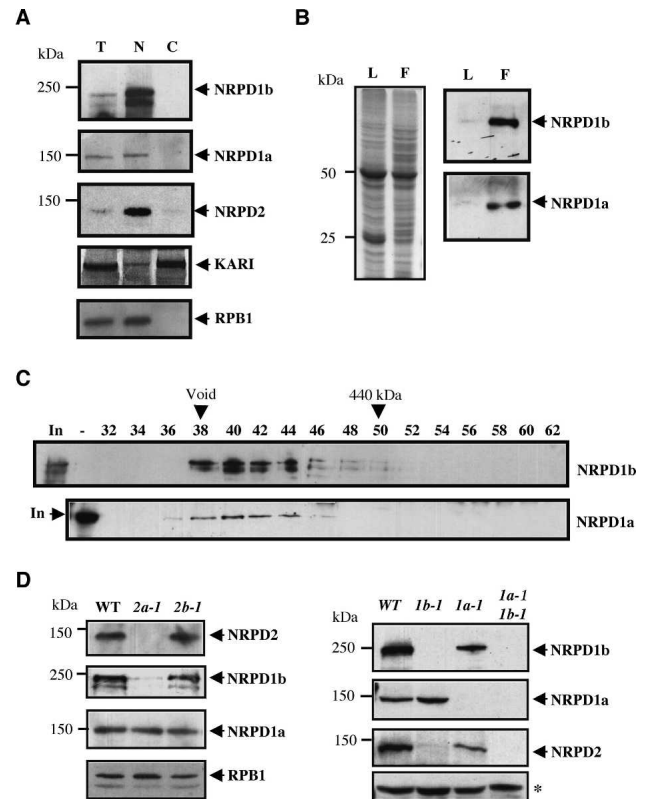


terminated. Beside the overall sequence conservation, AtNRPD1b presents several features that make it a likely component of a functional multimeric RNAP, as was previously shown for AtNRPD1a (Herr et al. 2005): All the invariant aspartate residues known to be directly involved in the catalytic activity are conserved (Fig. 1A), and the zinc-binding motifs (cc) that are critical for the assembly of the largest and second-largest subunits are maintained (Fig. 1A). When compared with yeast RNAPII, AtNRPD1b and AtNRPD2 (which compose RNAPIVb [see below]) present blocks of sequence homology that cluster around the active center of the enzyme (Fig. 1B, green panel) in a way that is similar to bacterial RNA polymerase (Fig. 1B, orange panel). This suggests the conservation of an RNA polymerase core structure in RNAPIVb.

Compared with NRPD1a, NRPD1b shows an additional feature: namely the presence of a long CTD that extends beyond the DCL-like motif (defective chloroplast and leaves) and comprises a highly hydrophilic domain including 10 complete repeats of a 16-amino-acid consensus sequence with multiple potential phosphorylation sites (Fig. 1C,D). This latter arrangement is reminiscent of the reiterations that characterize the CTD of RPB1 (Corden 1990; Young 1991) and was also found in rice and spinach NRPD1b, although the number of reiterations and the primary sequence of the repeat motifs appear specific to each species (data not shown).

#### *Arabidopsis harbors two forms of RNAPIV, RNAPIVa and RNAPIVb*

The identification of two functional *NRPD1* genes and one functional *NRPD2* gene raised the question of the exact nature of RNAPIV(s) in *Arabidopsis*. To address this question, we raised specific antibodies against NRPD1a, NRPD1b, and NRPD2 proteins, and we characterized several *Arabidopsis* lines (Alonso et al. 2003) containing the T-DNA-disrupted mutant alleles *nrpd1a-1* (*1a-1*), *nrpd1a-2* (*1a-2*), *nrpd1b-1* (*1b-1*), *nrpd1b-2* (*1b-2*), *nrpd2a-1* (*2a-1*), *nrpd2a-2* (*2a-2*), and *nrpd2b-1* (*2b-1*) (Fig. 1A; Supplementary Figs. S1A, S2A). Plants homozygous for these mutant alleles were identified by PCR analysis of segregating families and were used as controls in Western blot experiments. In wild-type extracts, proteins with an apparent mass identical to the predicted molecular sizes of NRPD1b (~250 kDa), NRPD1a (~150 kDa), and NRPD2 (~135 kDa) were specifically detected as judged by the absence of signal in corresponding KO insertion lines (Supplementary Figs. S1C, S2C). This analysis also shows that the *AtNRPD2b* gene does not contribute to NRPD2 protein accumulation, in agreement with the cDNA sequence data (Supplementary Figs. S2B, S2C). Finally, Western analysis of fractions corresponding to total leaf extracts or highly purified nuclei and chloroplasts indicated that, as expected for eukaryotic-type RNAP subunits, NRPD1a and NRPD1b proteins are specifically detected in nuclei (Fig. 2A). NRPD2 was also found in the nucleus (Fig. 2A),



**Figure 2.** *Arabidopsis* contains two RNAPIV forms. (A) Western blot analysis of *Arabidopsis* proteins from total extracts (T), percoll-purified nuclei (N), and chloroplasts (C) with NRPD1b, NRPD1a, NRPD2, KARI, and RPB1 antibodies. KARI was used as a plastid control protein, whereas RPB1 was used as a nuclear control. (B) Western blot analysis and Coomassie blue staining of leaf (L) and flower (F) extracts (~20  $\mu$ g) with NRPD1a and NRPD1b antibodies. (C) Gel filtration chromatographic analysis of the NRPD1-containing complexes. The eluted fractions were analyzed by Western blots using the NRPD1b and NRPD1a antibodies. Numbered lanes correspond to size-fractionated protein fractions ranging from 250 kDa (lane 58) to 2 MDa (lane 38). The peak position of the ferritin protein marker (440 kDa) is indicated. (In) Input material. (D) Mutual stability of the class IV largest and second-largest subunits in various null mutants. The steady state of NRPD1b and NRPD1a was analyzed in various *nrpd2* mutants (*left* panel). RPB1 protein was used as a loading control. The steady state of NRPD2 was analyzed in various *nrpd1* mutants (*right* panel). A nonspecific cross-reacting band indicated by an asterisk was used as a loading control.

consistent with immunolocalization experiments (Onodera et al. 2005). Analysis of the accumulation of these proteins in various plant organs also reveals a very abundant accumulation of the NRPD1a and NRPD1b proteins in flowers compared with leaves (Fig. 2B).

Multisubunit RNAP holoenzymes are high-molecular-weight protein complexes (Cramer et al. 2001). To test whether NRPD1a and NRPD1b are also part of large complexes in vivo, flower extracts were size-fractionated by sephacryl S-300HR gel filtration chromatography, and

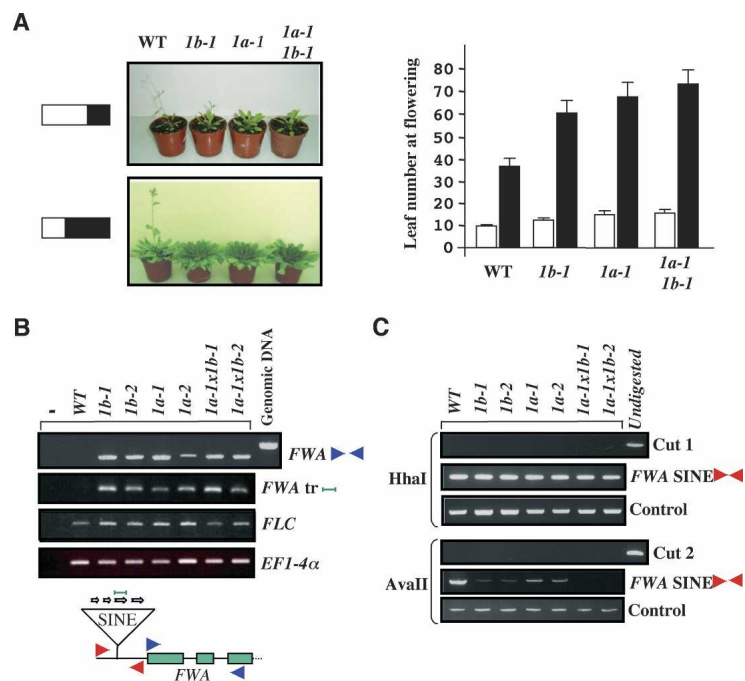
the resulting fractions were analyzed for the presence of NRPD1a and NRPD1b by Western blot. As shown in Figure 2C, both proteins were eluted close to the void volume, therefore indicating that they are present exclusively as complexed forms in vivo. The average molecular masses of the NRPD1a/b-containing complexes are close to 1 MDa, a typical size for nuclear RNAP holoenzymes (Greenblatt 1997).

To determine if NRPD1a, NRPD1b and NRPD2 are part of one or several multisubunit RNAP complexes, we analyzed the costability of these proteins in vivo using different null mutant backgrounds. The level of NRPD1b was strongly reduced in *nRPD2a-1* mutant extract, indicating that NRPD2 is required to stabilize NRPD1b (Fig. 2D, left panel). In contrast, NRPD1a accumulation was unaffected by the absence of NRPD2, which suggests that NRPD1a is either stable as a sub-complex in the mutant extract or that it does not form a tight complex with NRPD2 in vivo. In order to extend this analysis, NRPD2 protein stability was assessed in *nRPD1a-1*, *nRPD1b-1*, and *nRPD1a-1/nRPD1b-1* mutant plants. These experiments confirmed that NRPD1b stabilizes NRPD2, as most of the NRPD2 signal disappeared in *nRPD1b-1* mutant plants (Fig. 2D, right panel). While a weak NRPD2 signal was reproducibly observed in this background, no NRPD2 could be detected in *nRPD1a-1/nRPD1b-1* double-mutant plants (Fig. 2D, right panel), thus suggesting that a minor fraction of NRPD2 is complexed with NRPD1a. Taken together, our results are consistent with most of NRPD2 being engaged in a complex with NRPD1b, therefore defining a bona fide novel RNAPIV form in *Arabidopsis thaliana*. We refer to this novel RNA polymerase complex as RNAPIVb and renamed the RNAPIV form harboring NRPD1a as RNAPIVa.

### RNAPIVa and RNAPIVb act in the same pathway and direct silencing and DNA methylation at endogenous repeated loci

To investigate the functional role of the RNAPIVa and RNAPIVb complexes in vivo, two independent *nRPD1a* and *nRPD1b* homozygous mutant lines were further characterized. To rule out any possible functional redundancy of *NRPD1a* and *NRPD1b*, we also generated double homozygous lines for *nRPD1a-1/nRPD1b-1* and *nRPD1a-1/nRPD1b-2* alleles. None of the mutant alleles showed readily discernable morphological changes other than a nonadditive delay in flowering that was more pronounced when plants were grown under short days compared with long days (Fig. 3A).

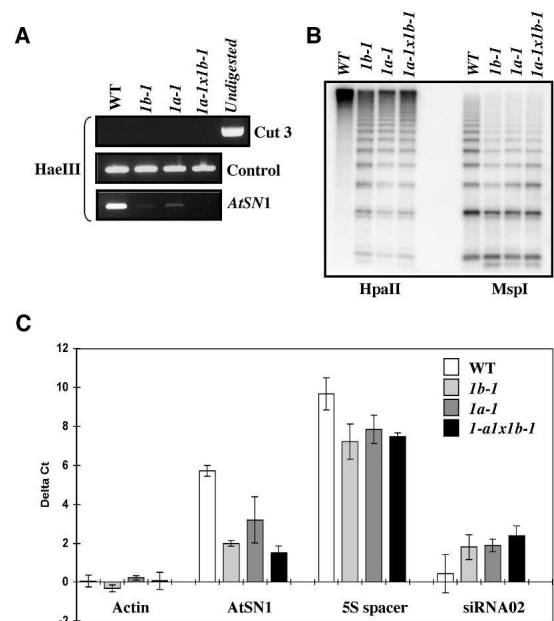
In order to understand the molecular basis of this phenotype, we analyzed several candidate genes whose expression is known to delay flowering time under both light conditions. While the regulation of most genes analyzed was not affected, as exemplified by the floral gene regulator *FLC*, all independent mutants showed significant expression of the imprinted *FWA* gene, whose ectopic expression has been shown to induce a late flowering phenotype in *Arabidopsis* (Fig. 3B, see panels *FLC* and *FWA*; Soppe et al. 2000). Consistent with the ectopic activation of *FWA*, we found that the long tandem repeat that encompasses the *FWA* promoter and the first two untranslated exons is transcriptionally active in all mutants analyzed (Fig. 3B, panel *FWA tr*; Lippman et al. 2004). The promoter region of *FWA* in wild-type plants has been shown to be methylated in all sequence contexts (CG, CNG, and CNN) within the two direct repeats region, causing *FWA* to be silenced (Soppe et al. 2000). In this regard, *FWA* activation in our mutants suggested a possible loss of cytosine methylation and con-



**Figure 3.** Phenotypes of *nRPD1* mutant lines. (A) Phenotype of *nRPD1* plants, grown under long days (top) and short days (bottom, left). The picture represents one out of nine plants per line. (Right) Comparison of the leaf number at flowering of wild-type (WT) and *nRPD1* single and double mutants. Similar results were obtained for *1a-2* and *1b-2* mutants. (B) RT-PCR analysis of flowering-related gene expression in homozygous single and double mutants. The *FWA* panel corresponds to the amplification of the coding region spanning exons 3–5 of the gene. The *FWAtr* corresponds to the tandem repeats located upstream of *FWA*. The cartoon shows the location of *FWA*-derived primer pairs used in panels B and C. (C) Analysis of DNA methylation of the *FWA* tandem repeats. DNA was digested with the methylation-sensitive enzymes HhaI (top) and AvaII (bottom).

sequent heterochromatin disruption. To determine if *npr1a* and *npr1b* mutant alleles affect *FWA* cytosine methylation, DNA methylation assays using methyl-sensitive restriction enzymes and PCR amplification were performed on the *FWA* tandem repeats region in genomic DNA extracted from wild type and *npr1* mutants. If the enzyme site is methylated, the DNA is not cut and can therefore be amplified by PCR, whereas demethylation of the site will allow DNA digestion and preclude PCR amplification due to restriction of the DNA. While no difference in CG methylation level (as judged by the use of the *HhaI* enzyme, which does not recognize GCGC motifs specifically methylated on the inner C) was observed in *npr1* mutants, a strong reduction of CNN methylation (as judged by the use of *AvaII* enzyme, which is blocked by methylation of the outer C of the recognition motif GGTC) was observed in all mutants (Fig. 3C). Taken together, our results indicate that while they retain pre-existing CG methylation at the endogenous *FWA* locus, *npr1* mutants are blocked in the perpetuation of CNN methylation, an effect that would probably account for the transcriptional activation of this gene. Based on the effects of mutations on flowering delay, *FWA* transcript accumulation, and DNA cytosine methylation, our results are consistent with a nonadditive effect of both mutations, suggesting that both proteins act in the same regulatory pathway leading to de novo methylation at the *FWA* locus.

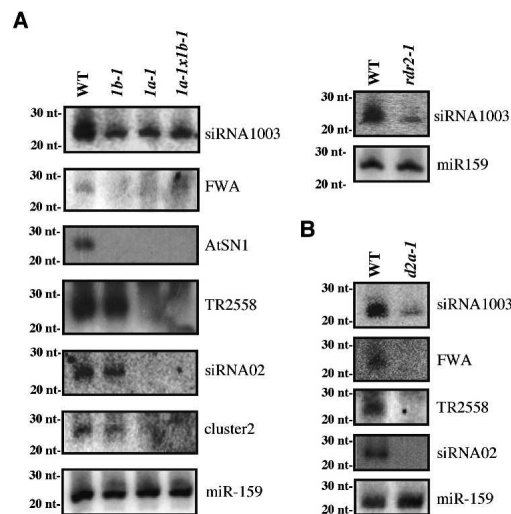
To determine if *npr1* mutants have a more general effect on the methylation of repeated endogenous DNA, we tested the methylation status at retrotransposon *AtSN1* and the 5S rDNA cluster, two repeated loci that are extensively methylated in all sequence contexts (CG, CNG, and CNN) (Hamilton et al. 2002; Zilberman et al. 2003; Xie et al. 2004). In wild-type plants, the *AtSN1* retroelement is resistant to cleavage by the methylation-sensitive restriction enzyme *HaeIII*, whereas a strong reduction of CNN methylation is observed in the mutants (Fig. 4A). This result was confirmed and extended by quantitative PCR analysis of genomic DNA digested with the methylation-dependent enzyme *McrBC* (Lippman et al. 2003), which confirms a nonadditive effect of the mutations (Fig. 4C; see Materials and Methods). As for the 5S rDNA intergenic spacers, *HpaII* and *MspI* digestions indicated again a loss of methylation in both CG and CNG sequence contexts in mutant plants (Fig. 4B). Comparison of the single to the double mutant showed a similar sensitivity to *HpaII* and *MspI*. This result was confirmed by quantitative PCR analysis of genomic DNA digested with the methylation-dependent enzyme *McrBC* with no apparent additive effect in this case (Fig. 4C). That both RNAPIV forms act in a nonredundant manner to direct methylation at 5S rDNA was further supported by the observation that the extent of methylation loss in the *npr2a-1* mutant (which likely harbors no functional RNAPIV forms because of *NRPD2* depletion) is similar to that seen in *npr1a* and *npr1b* single mutants (Supplementary Fig. S2D).



**Figure 4.** RNAPIVa and RNAPIVb act nonredundantly in the same pathway controlling methylation at *AtSN1* and 5s rDNA. (A) Analysis of DNA methylation of *AtSN1*. *HaeIII*-digested DNA was used as a template for PCR reactions using *AtSN1* and control primers. The *cut3* primers would not amplify DNA if the digestion were complete. Undigested corresponds to undigested DNA. (B) Blot analysis of 5S rDNA digested with methylation-sensitive restriction enzymes *HpaII* (left) and *MspI* (right) in *npr1* mutants and hybridized to a 5S probe. (C) DNA methylation at various repeated loci was assessed in different mutant backgrounds by quantitative *McrBC*-PCR. Delta Ct corresponds to the difference in Ct between digested and undigested samples. Heavily methylated sequences give large Delta Ct, whereas unmethylated sequences give a Delta Ct value of 0.

#### Differential role of the two RNAPIV forms on siRNA accumulation

Silencing and DNA methylation at *FWA*, *AtSN1*, and 5S gene clusters are associated with siRNAs (Hamilton et al. 2002; Zilberman et al. 2003; Lippman et al. 2004; Xie et al. 2004), which prompted us to examine the involvement of RNAPIVa and RNAPIVb in their accumulation. We found that the 24-nt siRNAs corresponding to *FWA* and *AtSN1* retrotransposons were absent in *npr1* single- and double-mutant alleles (Fig. 5A), which is consistent with a nonredundant role of RNAPIVa and RNAPIVb in the production of endogenous siRNAs. Likewise, siRNAs corresponding to 5S rDNA (siRNA1003) were also greatly reduced down to a level that reveals a 21-nt nonspecific band also observed in the *rdr2-1* mutant control (Fig. 5A, left and right panels). This was in sharp contrast to miR-159 accumulation, which remained unchanged in the various mutants (Fig. 5A). Thus, RNAPIVa and RNAPIVb appear to be nonredundant components of a pathway that generates endogenous siRNAs that guide DNA methylation at transposons and highly repeated endogenous sequences.



**Figure 5.** Differential role of the two RNAPIVs on siRNA accumulation. (A) Small RNA blot assays for miR-159 and various endogenous siRNAs in various *nrpd1* mutants. Blots in A were stripped and reprobed multiple times as indicated on the right. (Right panel) Small RNA blot assays for miR-159 and siRNA 1003 in the *rdr2-1* mutant is shown. (B) Small RNA blot assays for miR-159 and various endogenous siRNAs in *nrpd2a-1* mutant.

To test if the requirement of both RNAPIV forms for endogenous siRNA formation was general, we tested three additional endogenous loci present in few copies in the genome and that are associated with detectable levels of siRNAs: siRNA02, cluster 2, and TR2558 loci (Lippman et al. 2004; Xie et al. 2004). Whereas siRNAs corresponding to these loci were eliminated in *nrpd1a* and *nrpd1a/nrpd1b* mutants, they remained mostly unaffected in *nrpd1b* (Fig. 5A). We conclude from this result that RNAPIVb engagement in the siRNA synthesis pathway is target-dependent. DNA methylation assays indicate that RNAPIVb-independent loci (TR2558, siRNA02, and cluster 2) are less methylated than 5S rDNA, *FWA*, and *AtSN1* sequences (Fig. 4C; Lippman et al. 2004; Zilberman et al. 2004).

The identification of genomic sequences that require only RNAPIVa for the accumulation of corresponding siRNAs gave us a functional way to confirm the association between NRPD1a and NRPD2. The accumulation of siRNAs was decreased at all loci tested in the *nrpd2a* mutant (Fig. 5B; Herr et al. 2005; Onodera et al. 2005), which is consistent with the presence of NRPD2 in both RNAPIV forms.

## Discussion

The present work demonstrates that, besides the recently characterized RNAPIVa (Herr et al. 2005; Onodera et al. 2005), plants contain a novel form of RNAPIV, RNAPIVb. Although the two forms contain the same second-largest subunit (NRPD2), they differ by their largest subunit, termed NRPD1a and NRPD1b. Unlike

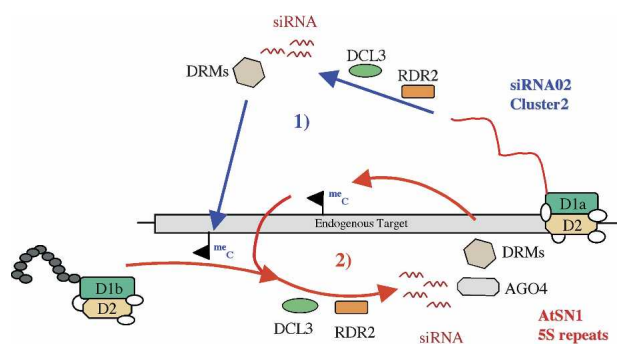
NRPD1a, NRPD1b is characterized by the presence of a long, repeat-containing CTD, and thus possesses a structure that is remarkably similar to RPB1, the largest subunit of RNAPII. The conservation of *NRPD1a* and *NRPD1b* genes in both monocot and dicot plants suggests that the emergence of these two forms predated the divergence between these two phyla.

Recent studies have implicated RNAPIVa in the silencing of endogenous DNA in a siRNA-dependent regulatory pathway (Herr et al. 2005; Onodera et al. 2005). Our work supports this conclusion but also reveals an unexpected level of complexity by demonstrating an additional requirement of RNAPIVb for siRNA accumulation and silencing at specific endogenous repeated loci. Indeed, our data indicate that RNAPIVa is necessary but not sufficient for the accumulation of endogenous siRNAs that guide de novo methylation at 5S rDNA and retroelements. At loci that are characterized by a high level of methylation, the synthesis of siRNAs and the perpetuation of de novo methylation also necessitates RNAPIVb. This is not the case at loci such as siRNA02, cluster 2, and TR2558, which are not densely methylated (Fig. 4C; Lippman et al. 2004; Zilberman et al. 2004). These observations indicate a correlation between RNAPIVb engagement and the capacity of the siRNA-dependent regulatory pathway to drive robust DNA methylation. This suggests, in turn, that RNAPIVb is an essential component of a reinforcing silencing mechanism acting downstream of RNAPIVa. Considering that each of the two RNAPIV forms is likely to produce long transcripts that can feed into the RNAi pathway involving RDR2 and DCL3 (Chan et al. 2004; Herr et al. 2005; Onodera et al. 2005), their requirement for siRNA synthesis at highly methylated loci implies the existence of two rounds of siRNA production. Involvement of primary and secondary siRNAs has already been proposed to explain heterochromatin formation at centromeric sequences in fission yeast (Sugiyama et al. 2005) and at highly repeated loci in plants (Onodera et al. 2005; Vaughn and Martienssen 2005). Our study suggests, however, that in plants the primary siRNAs that initiate silencing via de novo methylation derive from primary transcripts produced by RNAPIVa and not by euchromatic RNAP as previously proposed (Onodera et al. 2005; Vaughn and Martienssen 2005). The secondary siRNAs and the subsequent amplification of methylation would then result from the recognition, either direct or indirect, of the methylated target by RNAPIVb. This hypothesis is further supported by the observation that siRNA accumulation at endogenous loci is differentially affected in DNA methyltransferase mutants (Lippman et al. 2003; Zilberman et al. 2004; Onodera et al. 2005).

The involvement of two forms of RNAPIV in siRNA-dependent silencing pathway at specific targets implies that their ability to direct silencing must be somehow different. One major difference between them is the presence, in the CTD of RNAPIVb, of an additional region mostly composed of tandemly repeated motifs. Taking into account that the acquisition of a reiterated CTD has

dramatically enhanced the capacity of RNAPII to couple transcription to post-transcriptional events by recruiting essential transcription-related proteins (Hirose and Manley 2000; Hampsey and Reinberg 2003), it is tempting to propose, by analogy, that the RNAPIVb reiterated CTD would facilitate the coupling between the transcription and silencing pathways by providing a platform for RNAi- and/or silencing-related proteins. The observation that *ago4* and *nrdp1b* mutants show similar molecular phenotypes with respect to siRNA accumulation at all loci tested (Zilberman et al. 2004) makes the AGO4 protein a good candidate for being a specific component of the RNAPIVb-dependent pathway (Fig. 6). Based on the above considerations, we propose a model for siRNA-dependent silencing at repeated loci with two key steps (Fig. 6). The first one (blue arrows) implicates RNAPIVa, RDR2, and DCL3 in the synthesis of endogenous primary siRNA at all repeated loci. Their incorporation into putative effector complexes and subsequent recruitment of the DRM methyltransferase would trigger low level of de novo methylation of the homologous target loci. This would be the only operating step at loci such as TR2558, siRNA02, and cluster 2, whose siRNA accumulation is independent of RNAPIVb. The second step would subsequently concern targets that are characterized by high levels of methylation (red arrows). On such sequences (5S rDNA, retroelements), DNA methylation directed by primary siRNAs would create binding sites for the recruitment of RNAPIVb that, together with RDR2 and DCL3, will trigger production of secondary siRNA. Subsequent methylation via a pathway involving both AGO4 and DRMs would further recruit RNAPIVb: This would constitute the first round of a self-reinforcing RNAi loop linking siRNA synthesis to heterochromatin formation.

What determines the engagement of RNAPIVb in the siRNA-synthesizing pathway leading to silencing of a subset of targets remains unclear. One can consider that 5S rDNA and retroelements are highly repeated sequences and consequently that the dependence on RNAPIVb would rely on copy number. However, the



**Figure 6.** Hypothetical model for siRNA-dependent DNA methylation at endogenous repeated loci in plants. The blue arrows correspond to the RNAPIVa-dependent pathway that concerns siRNA02 and cluster 2. On targets such as *AtSN1* and 5S repeats it would be followed by a self-reinforcing loop implicating RNAPIVb (red arrows).

concerted action of both RNAPIVs is required for siRNA synthesis at *FWA* tandem repeat, which is not appreciably repeated elsewhere in the genome. This suggests the existence of another criterion discriminating RNAPIVb-dependent elements that could be the transcriptional potential of the endogenous locus. In this regard, it is interesting to note that in *ddm1* mutant plants, which lose most DNA methylation, *AtSN1* and *FWA* tandem repeat are active loci while the TR2558 tandem repeat remains transcriptionally silent (Lippman et al. 2004). One can therefore imagine that, in the process of heterochromatin formation, DNA methylation, together with another chromatin mark left by the transcribing euchromatic RNAP, could serve as specific binding sites for the recruitment of RNAPIVb.

An alternative model for RNAPIV activity has been recently proposed by Vaughn and Martienssen (2005), who raises the possibility that RNAPIVa may, in fact, synthesize primary transcripts from double-stranded RNA templates. Similar template specificity could be inferred to RNAPIVb since both polymerases present a similar extent of divergence compared with RNAPII. This would imply in our two-step model, that RNAPIVa would synthesize primary transcripts in *trans* from mitotically transmissible double-stranded RNA templates while RNAPIVb would probably be acting in *cis*. One appealing aspect of this model is that it readily accounts for the engagement of RNAPIVb at only a subset of loci that lead to siRNA accumulation. In this model, diffusible primary siRNA and associated effector complexes would guide, via DRM-directed methylation, RNAPIVb and the associated RNAi machinery to the DNA target sequence. Once located at the transcriptionally active target sites, RDR2 will convert nascent euchromatic transcripts into dsRNA that then will be transcribed in *cis* by RNAPIVb, therefore maintaining the sustained production of secondary transcripts and subsequent secondary siRNAs. Future biochemical and functional studies will be necessary to understand precisely the interplay between both RNAPIV forms and the mechanism by which RNAPIV triggers a high level of DNA methylation and heterochromatin assembly.

## Materials and methods

### Arabidopsis growth and genetic analysis

For plant growth, seeds were stratified at 4°C for 2 d before growth at 23°C with 16 h light/8 h dark cycles on soil or under continuous light on plates with MS medium. All mutants originated from the Salk Institute collection of T-DNA insertions (<http://signal.salk.edu/cgi-bin/tdnaexpress>) obtained via the Nottingham *Arabidopsis* Stock Center (<http://nasc.life.nott.ac.uk>). The *nrdp1a-1* (*1a-1*) and *nrdp1a-2* (*1a-2*) alleles correspond, respectively, to the lines SALK\_583051 and SALK\_628428. The *nrdp1b-1* (*1b-1*) and *nrdp1b-2* (*1b-2*) alleles correspond, respectively, to the lines SALK\_029919 and SALK\_033852. Genomic DNA was amplified with the LBA1 and specific primers to detect the T-DNA (5'-CGCTGTCTC GTTGCTGTAAACATGG-3' for *1b-1* and *1b-2* lines, 5'-AACCTCGAAGCAACAAGAATCTCCG-3' for the *1a-1* line, and 5'-



CGGATCCGCAATTATCAAAGCTTCTCC-3' for the 1a-2 line). To follow the presence of the endogenous copy of the genes and identify homozygous lines, genomic DNA was PCR-amplified with specific couples of primers (1b-1, 5'-CAAAGTG GTGATGCATGGAGG-3' and 5'-ATGTAAATTTTGGAAAGT CGGC-3'; 1b-2, 5'-GCGCTCTAGACTTCAAATGGTCATAT CCACG-3' and 5'-CGCTGTCTCGTTGCTGTAACATGG -3'; 1a-1, 5'-AACCTCGAAGCAACAAGAATCTCCG-3' and 5'-ATCACGGTGAAGTGTAGTTGAAGCC-3'; 1a-2, 5'-GCGG ATCCGCAATTATCAAAGCTTCTCC-3' and 5'-CTCGTTA AAGATGAAACAACGG-3'). Homozygous lines were propagated by repeated self-pollination.

For crosses, homozygous plants were grown on soil to maturity and five to eight flowers were hand-pollinated after removal of their immature stamens. The F2 generation was analyzed by PCR to characterize double-homozygous plants from which seeds were collected. The F3 generation was used for further experiments.

#### Bioinformatics and molecular visualization

Gene searching was done on the *Arabidopsis* genome sequence database or the nonredundant databases by using the TBLASTN program (Altschul et al. 1997) with the amino acid sequence of the human RPB1 protein. The spinach *NRPD1b* ortholog gene was isolated by screening of a  $\lambda$ Zap library (Clontech), using a labeled DNA fragment corresponding to a part of the large exon 7 of *AtNRPD1b* as a probe. Phylogenetic studies were performed with the conserved region comparison as described previously (Lagrange et al. 2003). Molecular visualization of the RNAP conserved regions was performed using the WebLab ViewerLite software from Accelrys.

#### Protein isolation, cell fractionation, antibodies, and Western blotting

Total plant protein extracts were obtained following the Tanaka method (Hurkman and Tanaka 1986). Coomassie staining was used to calibrate loadings. Intact *Arabidopsis* chloroplasts were purified as described previously (Kunst 1998). Intact nuclei from *Arabidopsis* leaves were purified as described previously onto percoll gradients (Watson and Thompson 1986).

Proteins were separated onto SDS/PAGE gels and blotted onto PVDF membrane (Immobilon-P, Millipore). Rabbit antisera were raised against peptides designed in NRPD1a, NRPD1b, and NRPD2 and affinity-purified. Dilutions of 1/1000 were used to perform Western blot analysis of the indicated protein extracts. Dilutions of 1:1000 and 1:500 of KARI (Dumas et al. 2001) and RPB1 monoclonal antibody 8WG16, respectively, were used for plastid and nuclei controls.

#### Gel filtration chromatography

Total homogenates were prepared in BC500 buffer (Maldonado et al. 1996) and centrifuged for 30 min at 20,000 rpm (rotor JA 25.50), and the supernatant was filtered through a 0.45- $\mu$ m filter (Gelman Sciences). About 400  $\mu$ g of total soluble proteins were fractionated through a HiLoad 16/60 Sephacryl S-300 HR FPLC column (Pharmacia), with BC500 buffer at a flow rate of 1 mL/min. All fractionations were carried out in the 4°C cold room. Fractions of 1 mL each were collected and concentrated by acetone precipitation. An equal volume of individual fractions was used for immunoblot analysis. The protein standards for size estimation of the NRPD1-containing complex(es) were as follows: blue dextran, 2 MDa; ferritin, 440 kDa.

#### RNA extraction and analysis

Total RNA was isolated from different organs from plants grown in soil using the Invisorb Spin Plant RNA Mini Kit following the manufacturer's instructions (Invitex) followed by a DNase treatment using the RQ1 RNase-free DNase (Promega). A PCR reaction was performed on RNA with primers for the *RPL21* gene [5'-TCCACTGCGTTCGACGTCTCG-3' and 5'-CT CCAACTGCAACATTGGGCG-3') to check for the absence of genomic DNA. Reverse transcription was performed on 3–5  $\mu$ g of total RNA using the ProSTAR First-Strand RT-PCR kit following the manufacturer's instructions (Stratagene). Semiquantitative RT-PCR was calibrated by nonsaturating PCR reaction (22 cycles) using primers designed to amplify the *EF1-4 $\alpha$*  cDNAs (5'-CTGCTAACTTCACCTCCCAG-3' and 5'-TGGTGGGTA CTCAGAGAAGG-3'). A parallel set of reactions without addition of reverse transcriptase was run as a quality control. To analyze *NRPD1b* and *NRPD1a* expression in wild-type plants, semiquantitative RT-PCR was performed with the following primers: 5'-CAAAGTGGTGATGCATGGAGG-3' and 5'-CG GCTCATTATTTTCTGGGAC-3' for *AtNRPD1b* and 5'-ACT TGCGCTGCTTGAATGAC-3' and 5'-CAAAGACATTGTC TGTTGTGC-3' for *AtNRPD1a*. Primers used to amplify *FLC* and *FWA* cDNAs were, respectively, 5'-AAGATCCTTGATC GATATGGG-3' and 5'-AAGTAGTGGGAGAGTCACCGG-3' (27 cycles), and 5'-GCGGTTGAAACATTCCAAAACC-3' and 5'-AGTGTCTCCACAATTAGTTGCC-3' (40 cycles). The primer pair 5'-TCCATTCAACATTCATACGAGCGCCGC -3' and 5'-TCTGATATTTGCTGGAAAAACAACAATAAT C-3' was used to detect cDNAs corresponding to the long tandem repeat that encompasses the *FWA* promoter and the first two untranslated exons (40 cycles).

For siRNA analysis, total RNA samples were extracted from inflorescences using TRIzol (Invitrogen). Thirty micrograms of total RNA were separated on a 15% polyacrylamide gel containing 7 M urea and electroblotted onto Hybond N<sup>+</sup> membranes (Amersham Pharmacia Biotech Inc.). Hybridization was performed in PerfectHyb solution (Sigma) following the supplier's instruction. siRNA1003, miR159, and siRNA02 were detected using end-labeled DNA oligonucleotides AS1003, AS159, and AS-02, respectively (Xie et al. 2004). The probes for cluster2, TR2558, and *FWA* siRNA detection were riboprobes synthesized using the Riboprobe system T7 Kit (Promega) following the supplier's instruction. Primers used to amplify cluster2- and TR2558-specific sequences were, respectively, 5'-CTTTTT CAAACCATAAACCAGAAA-3' and 5'-TTGCTGATTTGTAT TTTATGCAT-3', and 5'-CAATATCTAAAATTATTTCTATG AGTCGG-3' and 5'-GCATAACATAACTATGCAAAACATG AGC-3'. Primers used to generate the *FWA* riboprobe template were the ones used for amplifying the tandem repeat in RT-PCR experiments. *AtSN1* siRNA were detected using end-labeled DNA oligonucleotides JP2107 (Zilberman et al. 2004).

#### Genomic DNA extraction and methylation detection assays

Genomic DNA was extracted from seedlings using the Wizard Genomic DNA extraction kit (Promega). DNA was digested with different restriction enzymes (HhaI, AvaII for *FWA* analysis and HaeIII for *AtSN1* analysis). HhaI recognizes the sequence GCGC and is inhibited by methylation of either cytosine. AvaII recognizes the sequence GGA/TCC and is inhibited by methylation at the outer cytosine, allowing detection of asymmetric methylation. PCR amplification was subsequently done on 150 ng of digested DNA using the following primers. The 5'-GAAG GTTCTCATCATATACCG-3' and 5'-GGTTTTGGAATGTTT CCAACCGC-3' primer pair (*FWA SINE*) was used to identify the methylation status of the tandem repeats that encompass

the FWA promoter and the first two untranslated of FWA. Primers designed for RT-PCR amplification of the FWA coding region flanked an unmethylated HhaI site and were used to check the completion of genomic DNA digestion (cut1). Since the amplified DNA sequence is devoid of AvaII, this pair was used to control equal template concentration in the AvaII experiment (control panel). Internal PCR control primers (5'-GGAGAGAG GCTTGTTGGATACTGC-3' and 5'-GAACACGCATGACA GTGGGTGGAG-3') span a region lacking HhaI as well as HaeIII sites and were used to control equal template concentration in both HhaI and HaeIII experiments (control panels). In contrast, this control region contains an unmethylated AvaII site, and, therefore, those primers were used to check the completion of genomic DNA digestion by this enzyme (cut2). For *AtSN1* methylation status, the following primers were used: 5'-ACTTAATTAGCACTCAAATTAACAAAATAAG T-3' and 5'-TTTAAACATAARAARAARTTCCTTTTTTCATC TAC-3' (Hamilton et al. 2002). The primers FWA SINE described above flank a DNA sequence containing unmethylated HaeIII sites and are used to check the completion of HaeIII digestion (cut3 panel).

The Southern experiment was performed on 5 µg of genomic DNA digested by the methylation-sensitive enzymes HpaII or MspI and separated on 0.9% agarose gel. After blotting on Hybond N<sup>+</sup> membranes (Amersham Pharmacia Biotech Inc.), hybridization was performed in PerfectHyb solution (Sigma) following the supplier's instructions, with the transcribed region of the 5S rRNA gene.

McrBC-based methylation analysis was performed as follows, based on a previously published protocol (Lippman et al. 2003). Genomic DNA (500 ng) was digested for 2 h at 37°C with 30 U of McrBC enzyme (NEB). A mock digestion was done in parallel, with no enzyme. Quantitative PCR (ABI Prism 79000 HT) was directly performed on 1/50th of the digested and mock samples, using the ABI SybrGreen PCR master mix. Reactions were performed on two independent digests, and were usually duplicated. Differences in Ct values between digested and undigested samples were averaged over three or four sets of PCR reactions. Primers used for quantitative PCR were as follows: 5'-AACGT GCTGTTGGCCAGT-3' and 5'-CTGGAAGTCAAGCCC AAAG-3' for *AtSN1*; 5'-CTTTTCGGGCNTTTTNGTG-3' and 5'-CGAAAAGGTATCACATGCC-3' for the 5S spacer; 5'-TTC CGAAACAGTAAACCATCG-3' and 5'-TCAAAGTAAAAGT GGTCTTGG-3' for siRNA02up; 5'-GCCATCCCAAGCTG TTCTCTC-3' and 5'-CCCTCGTAGATTGGCACAGT-3' for Actin 2.

## Acknowledgments

We thank E. Lam, A.M. Henry, and our lab colleagues for fruitful discussions. We thank O. Voinnet for providing mutant seed stocks. We thank M. Block for providing KARI antibodies. G.Y. was the recipient of a post-doctoral fellowship from CNRS. This work was supported by a grant from CNRS (Programme ATIP to T.L.) as well as by a grant from the European Union (Epigenome Network of Excellence to V.C.).

## Note added in proof

A genetic screen for mutants defective in RNA-directed DNA methylation and silencing of a transgene promoter in *Arabidopsis* identified *nrd1b* and *nrd2a* as *drd3* and *drd2* (Kanno et al. 2005). These findings are in agreement with those reported here that RNA-mediated DNA methylation at endogenous repeated loci requires the action of two forms of RNAPIV in *Arabidopsis*.

## References

- Allison, L.A., Moyle, M., Shales, M., and Ingles, C.J. 1985. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* **42**: 599–610.
- Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, D.K., Barajas, P., Cheuk, R., et al. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**: 653–657.
- Altschul, S.F., Thomas, L.D., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipmann, D.J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**: 3389–3402.
- Arabidopsis* Genome Initiative 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 796–815.
- Bentley, D. 1999. Coupling RNA polymerase II transcription with pre-mRNA processing. *Curr. Opin. Cell Biol.* **11**: 347–351.
- Carlson, M. 1997. Genetics of transcriptional regulation in yeast: Connections to the RNA polymerase CTD. *Annu. Rev. Cell. Dev. Biol.* **13**: 1–23.
- Chan, S.W.-L., Zilberman, D., Xie, Z., Johansen, L.K., Carrington, J.C., and Jacobsen, S.E. 2004. RNA silencing genes control de novo DNA methylation. *Nature* **303**: 1336.
- Corden, J.L. 1990. Tails of RNA polymerase II. *Trends Biochem. Sci.* **15**: 383–387.
- Cramer, P. 2002. Multisubunit RNA polymerases. *Curr. Opin. Struct. Biol.* **12**: 89–97.
- 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.
- Dahmus, M.E. 1996. Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J. Biol. Chem.* **271**: 19009–19012.
- Dieci, G., Hermann-Le Denmat, S., Lukhtanov, E., Thuriaux, P., Werner, M., and Sentenac, A. 1995. A universally conserved region of the largest subunit participates in the active site of RNA polymerase III. *EMBO J.* **14**: 3766–3776.
- Dumas, R., Biou, V., Halgand, F., Douce, R., and Duggleby, R. 2001. Enzymology, structure, and dynamics of acetoxy acid isomeroreductase. *Accounts Chem. Res.* **34**: 399–408.
- Ebright, R.H. 2000. RNA polymerase: Structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *J. Mol. Biol.* **304**: 687–698.
- Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. 2001. Structural basis of transcription: An RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**: 1876–1882.
- Greenblatt, J. 1997. RNA polymerase II holoenzyme and transcriptional regulation. *Curr. Opin. Cell Biol.* **9**: 310–319.
- Hamilton, A.J., Voinnet, O., Chappell, L., and Baulcombe, D. 2002. Two classes of short interfering RNA in RNA silencing. *EMBO J.* **21**: 4671–4679.
- Hampsey, M. and Reinberg, D. 2003. Tails of intrigue: Phosphorylation of RNA polymerase II mediates histone methylation. *Cell* **113**: 429–432.
- Herr, A.J., Jensen, M.B., Dalmay, T., and Baulcombe, D.C. 2005. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**: 118–120.
- Hirose, Y. and Manley, J.L. 2000. RNA polymerase II and the integration of nuclear events. *Genes & Dev.* **14**: 1415–1429.
- Hurkman, W.J. and Tanaka, C.K. 1986. Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol.* **116**: 1209–1218.
- Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E.,

- Daxinger, L., Kreil, D.P., Matzke, M., and Matzke, A.J. 2005. Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat. Genet.* **37**: 761–765.
- Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedeá, E., Greenblatt, J.F., and Buratowski, S. 2004. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**: 517–522.
- Kunst, L. 1998. Preparation of physiologically active chloroplasts from *Arabidopsis*. *Methods Mol. Biol.* **82**: 43–48.
- Lagrange, T., Hakimi, M.-A., Pontier, D., Courtois, F., Alcaraz, J.P., Grunwald, D., Lam, E., and Lerbs-Mache, S. 2003. Transcription factor IIB (TFIIB)-related protein (pBrp), a plant-specific member of the TFIIB-related protein family. *Mol. Cell Biol.* **23**: 3274–3286.
- Lippman, Z., May, B., Jordan, C., Singer, T., and Martienssen, R. 2003. Distinct mechanisms determine transposon inheritance and methylation via small interfering RNAs and histone modification. *PLoS Biol.* **1**: E67.
- Lippman, Z., Gendrel, A.-V., Black, M., Vaughn, M.W., Dedhia, N., McCombie, W.R., Lavine, K., Mittal, V., May, B., Kasschau, K.D., et al. 2004. Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430**: 471–476.
- Maldonado, E., Drapkin, R., and Reinberg, D. 1996. Purification of human RNA polymerase II and general transcription factors. *Methods Enzymol.* **274**: 72–100.
- Onodera, Y., Haag, J.R., Ream, T., Nunes, P.C., Pontes, O., and Pikaard, C.S. 2005. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**: 513–522.
- Otero, G., Fellows, J., Li, Y., de Bizemont, T., Dirac, A.M., Gustafsson, C.M., Erdjument-Bromage, H., Tempst, P., and Svendsen, J.Q. 1999. Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation. *Mol. Cell* **3**: 109–118.
- Roeder, R.G. and Rutter, W.J. 1969. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**: 234–237.
- Sentenac, A. 1985. Eukaryotic RNA polymerases. *CRC Crit. Rev. Biochem.* **18**: 31–90.
- Soppe, W.J.J., Jacobsen, S.E., Alonso-Blanco, C., Jackson, J.P., Kakutani, T., Koornneef, M., and Peeters, A.J.M. 2000. The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol. Cell* **6**: 791–802.
- Sugiyama, T., Cam, H., Verdel, A., Moazed, D., and Grewal, S.I. 2005. RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc. Natl. Acad. Sci.* **102**: 152–157.
- Sweetser, D., Nonet, M., and Young, R.A. 1987. Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. *Proc. Natl. Acad. Sci.* **84**: 1192–1196.
- Vaughn, M.W. and Martienssen, R.A. 2005. Finding the right template: RNA Pol IV, a plant-specific RNA polymerase. *Mol. Cell* **17**: 754–756.
- Watson, J.C. and Thompson, W.F. 1986. Purification and restriction endonuclease analysis of plant nuclear DNA. *Methods Enzymol.* **118**: 57–75.
- Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, 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**: E104.
- Young, R.A. 1991. RNA polymerase II. *Annu. Rev. Biochem.* **60**: 689–715.
- Zilberman, D., Cao, X., and Jacobsen, S.E. 2003. Argonaute 4 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.