2009

Intergenic transcription by RNA Polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*

Binglian Zheng
University of California - Riverside

Zhengming Wang
University of California - Riverside

Shengben Li
University of California - Riverside

Bin Yu
University of Nebraska-Lincoln, byu3@unl.edu

Jin -Yuan Liu
Tsinghua University

See next page for additional authors

Intergenic transcription by RNA Polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*

Binglian Zheng,[1] Zhengming Wang,[1,2] Shengben Li,[1] Bin Yu,[1,3] Jin-Yuan Liu,[2] and Xuemei Chen[1,4]

[1]Department of Botany and Plant Sciences and Institute of Integrative Genome Biology, University of California at Riverside, Riverside, California 92521, USA; [2]Laboratory of Molecular Biology and Protein Science Laboratory of the Ministry of Education, Department of Biological Sciences and Biotechnology, Tsinghua University, Beijing 100084, Peoples Republic of China

Intergenic transcription by RNA Polymerase II (Pol II) is widespread in plant and animal genomes, but the functions of intergenic transcription or the resulting noncoding transcripts are poorly understood. Here, we show that *Arabidopsis* Pol II is indispensable for endogenous siRNA-mediated transcriptional gene silencing (TGS) at intergenic low-copy-number loci, despite the presence of two other polymerases—Pol IV and Pol V—that specialize in TGS through siRNAs. We show that Pol II produces noncoding scaffold transcripts that originate outside of heterochromatic, siRNA-generating loci. Through these transcripts and physical interactions with the siRNA effector protein ARGONAUTE4 (AGO4), Pol II recruits AGO4/siRNAs to homologous loci to result in TGS. Meanwhile, Pol II transcription also recruits Pol IV and Pol V to different locations at heterochromatic loci to promote siRNA biogenesis and siRNA-mediated TGS, respectively. This study establishes that intergenic transcription by Pol II is required for siRNA-mediated TGS, and reveals an intricate collaboration and division of labor among the three polymerases in gene silencing.

[*Keywords*: RNA polymerase; siRNA; DNA methylation; transcriptional gene silencing; noncoding transcripts]

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

Received September 26, 2009; revised version accepted October 26, 2009.

One prominent role of transcriptional gene silencing (TGS) is defense against the proliferation of transposable elements and repeats (Moazed 2009). In *Arabidopsis*, silencing at endogenous repeat loci involves histone modifications, such as dimethylation at histone H3 Lys 9 (H3K9me2), and RNA-directed DNA methylation (RdDM) guided by homologous siRNAs (Hamilton et al. 2002; Zilberman et al. 2003; Xie et al. 2004). Players in RdDM in *Arabidopsis* include RNA-DEPENDENT RNA POLYMERASE2 (RDR2), ARGONAUTE4 (AGO4), AGO6, and DICER-LIKE3 (DCL3) (Zilberman et al. 2003; Xie et al. 2004; Zheng et al. 2007). RDR2 is thought to copy ssRNA from a heterochromatic locus into dsRNA. DCL3 cleaves the dsRNA into 24-nucleotide (nt) siRNA duplexes, one strand of which associates with AGO4 (or AGO6) to form an RNA-induced silencing complex (RISC). AGO4 RISC recruits, directly or indirectly, the de novo DNA methyltransferase DRM2 to genomic loci homologous to the siRNAs to trigger DNA methylation (Cao et al. 2003). An AGO4 RISC complex may also guide H3K9 methylation by recruiting the SUVH family of histone methyltransferases (Malagnac et al. 2002; Ebbs et al. 2005; Naumann et al. 2005; Ebbs and Bender 2006).

In fission yeast, siRNA-mediated formation of heterochromatin at pericentromeric repeats depends on RNA Polymerase II (Pol II) transcription of the repeats (Volpe et al. 2002; Djupedal et al. 2005; Kato et al. 2005). The Pol II-generated noncoding RNAs have a dual function in heterochromatin assembly, serving as both precursors to siRNAs and scaffolds that interact with siRNAs to recruit chromatin-modifying factors (Djupedal et al. 2005; Kato et al. 2005). Nonlethal mutations that disrupt siRNA-mediated gene silencing and/or siRNA accumulation in *Schizosaccharomyces pombe* have been mapped to RPB2 and RPB7, two subunits of Pol II (Djupedal et al. 2005; Kato et al. 2005).

Plants have evolved from Pol II two additional RNA polymerases—Pol IV and Pol V—to specialize in siRNA production and siRNA-mediated gene silencing, respectively (Herr et al. 2005; Kanno et al. 2005; Onodera et al. 2005; Pontier et al. 2005). Many subunits of Pol IV or Pol V have identical or paralogous counterparts in Pol II,

[3]Present address: Center for Plant Science innovation, University of Nebraska, Lincoln Lincoln, NE 68588, USA.
[4]Corresponding author.
E-MAIL xuemei.chen@ucr.edu; FAX (951) 827-4437.
Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.1868009.

indicating that they are derived from Pol II (Huang et al. 2009; Lahmy et al. 2009; Ream et al. 2009). The largest subunits of Pol II, Pol IV, and Pol V are distinct from one another (encoded by *NRPB1*, *NRPD1*, and *NRPE1*, respectively). Pol IV and Pol V share the same second largest subunit (encoded by *NRPD2*/*NRPE2*), which is distinct from that of Pol II (encoded by *NRPB2*). While null mutations in Pol II genes are lethal (Onodera et al. 2008), those in Pol IV- or Pol V-specific genes do not lead to strong visible defects, but result in transcriptional derepression of transposons and repeat loci throughout the genome. Pol IV acts at an early step in the RdDM pathway to generate 24-nt siRNAs, probably by transcribing heterochromatic loci to produce precursor RNAs (Herr et al. 2005; Kanno et al. 2005; Onodera et al. 2005). Pol V is required for siRNA-mediated silencing of target sequences, probably by recruiting AGO4 RISCs through RNA transcripts that it generates at the target loci and by physically interacting with AGO4 through its GW/WG motifs (Li et al. 2006; El-Shami et al. 2007; Wierzbicki et al. 2008, 2009). In *nrpe1* mutants, siRNA accumulation is reduced in a subset of the Pol IV-dependent loci but is unaffected in other loci (Pontier et al. 2005; Huettel et al. 2006). It is thought that the role of Pol V in siRNA accumulation at some loci is an indirect consequence of its function in heterochromatin formation, which in turn promotes siRNA production.

The presence of two polymerases specializing in TGS in plants raises the question of whether or not the plant Pol II has any role in TGS. The isolation of a weak allele in the gene encoding the second largest subunit of Pol II, *NRPB2*, allowed us to assess Pol II's role in TGS. Despite the weak nature of the mutation, it results in the derepression of a set of intergenic low-copy-number repeat sequences. Pol II acts in siRNA-mediated DNA and histone methylation by recruiting AGO4 and Pol V to the silenced loci through both physical interactions with AGO4 and scaffold transcripts that it generates adjacent to the silenced loci. In addition, Pol II, but not Pol V, acts in the feed-forward loop to promote siRNA accumulation from these loci by recruiting Pol IV to chromatin. Our studies show that Pol II, Pol IV, and Pol V have nonredundant roles, thus revealing a more complex TGS mechanism with noncoding transcription by Pol II playing a central role in coordinating the other two polymerases.

Results

Isolation of a weak mutant allele of NRPB2

In an ethyl methanesulfonate mutagenesis screen aimed at the isolation of mutants compromised in microRNA (miRNA) biogenesis, we isolated a mutant with pleiotropic developmental phenotypes that are also exhibited by known mutants in miRNA biogenesis genes. The defects included distorted leaf shape, delayed leaf emergence, reduced number of floral organs, abnormal phyllotaxy, and shorter siliques (Fig. 1A; data not shown). Genetic studies showed that these morphological defects were controlled by a single, recessive mutation. The mutation

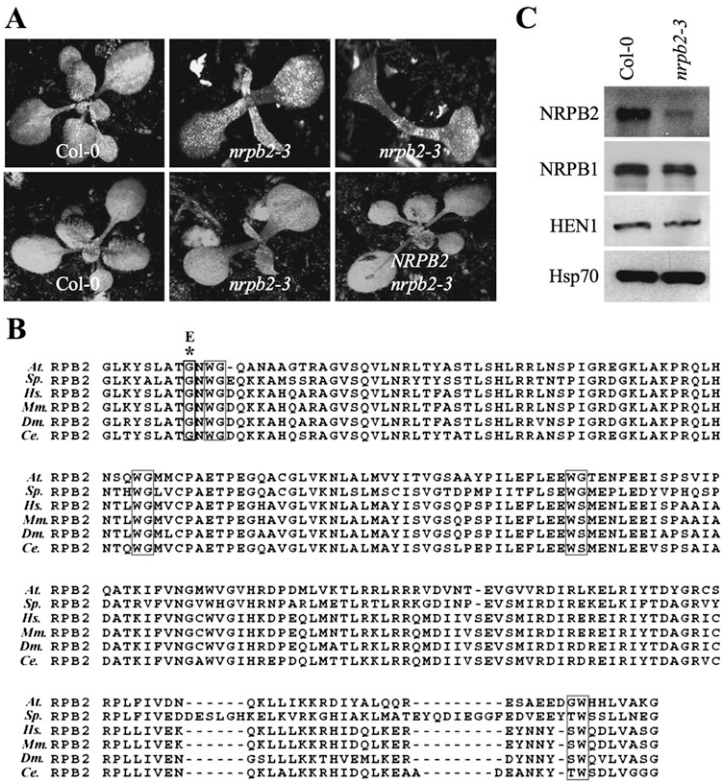


Figure 1. Isolation of an *nrpb2* mutant. (*A*) Three-week-old seedlings of various genotypes as indicated. (*NRPB2 nrpb2-3*) An *nrpb2-3* mutant carrying an *NRPB2* transgene. Note that the photos of the mutant plants were taken at a higher magnification than those of the wild-type and the rescued line. (*B*) Partial amino acid sequence of RPB2 in various species. The conserved glycine, which was mutated to glutamic acid in *nrpb2-3*, is marked by a rectangle and an asterisk. The other rectangles show GW/WG motifs in RPB2 proteins. (*At*) *Arabidopsis thaliana*; (*Sp*) *Schizosaccharomyces pombe*; (*Hs*) *Homo sapiens*; (*Mm*) *Mus musculus*; (*Dm*) *Drosophila melanogaster*; (*Ce*) *Caenorhabditis elegans*. (*C*) Western blotting to determine the levels of NRPB2 and NRPB1 in wild type (Col-0) and *nrpb2-3*. HEN1 and Hsp70 served as loading controls.

responsible for the morphological defects was mapped to a 206-kb region on chromosome IV. Sequencing candidate genes in this region identified a G-to-A mutation in the coding region of *NRPB2*, which encodes the second-largest subunit of Pol II. This mutation caused a conversion of a highly conserved glycine in RPB2 proteins from *S. pombe*, animals, and plants to glutamic acid (Fig. 1B). An *NRPB2* genomic fragment, when introduced into this mutant, completely rescued the phenotypes in 51 out of 58 T1 transgenic lines. (Fig. 1A; data not shown). Therefore, the mutant is an allele of *NRPB2*. Two previously isolated loss-of-function alleles of *NRPB2—nrpb2-1* and *nrpb2-2*—are embryo-lethal, consistent with the essential role of Pol II in nuclear gene expression (Onodera et al. 2005, 2008). We named this new mutant *nrpb2-3*. Consistent with its being a weak allele, NRPB2 protein still accumulated, but at a reduced level in the mutant (Fig. 1C). Levels of NRPB1, the largest subunit of Pol II, remained largely unchanged (Fig. 1C). To minimize differences in tissue composition between wild-type and *nrpb2-3* samples, we used inflorescences, which appeared to be the least affected in *nrpb2-3*, for all studies described below.

The expression of a small set of genes is affected in nrpb2-3

We next examined the effects of the *nrpb2-3* mutation on gene expression at the transcriptome level using ATH1 Affymetrix microarrays. RNAs from inflorescence tissues from three biological replicates of wild type and *nrpb2-3* were compared. A total of 448 genes were reduced in expression by twofold in *nrpb2-3* (Supplemental Table S1), and 95 genes were increased in expression by twofold in *nrpb2-3* (Supplemental Table S2). Most affected genes encode metabolic enzymes, and no genes known to play a role in siRNA biogenesis or siRNA-mediated TGS were affected. Real-time RT–PCR confirmed that genes with known roles in siRNA biogenesis, DNA methylation or demethylation, and histone H3K9me2 were not affected by the *nrpb2-3* mutation (Supplemental Fig. S1). Therefore, the effects of *nrpb2-3* on small RNA biogenesis or TGS (see below) are unlikely to be attributable to indirect effects of the mutation on the expression of these genes. We suspect that some of the developmental defects of the mutant were caused by reduced miRNA levels (data not shown).

Role of Pol II in the accumulation of heterochromatic siRNAs

We classified heterochromatic loci into two types according to their dependence on Pol IV and Pol V for siRNA accumulation (Zhang et al. 2007; Mosher et al. 2008). Type I loci are high-copy-number repeats or transposons, such as AtSN1, siR1003, AtREP2, SimpleHAT2, and AtCopia2. siRNA levels from these loci are Pol IV- and Pol V-dependent (Herr et al. 2005; Onodera et al. 2005; Pontier et al. 2005). Type II loci are represented by low-copy-number repeats and intergenic sequences, such as siR02, Cluster4, TR2558, Cluster2, and soloLTR. siRNAs from these loci are Pol IV-dependent but Pol V-independent (Pontier et al. 2005; Huettel et al. 2006).

Five type I siRNAs accumulated to comparable levels in *nrpb2-3* and wild type, but were absent in *sde4-3* (a Pol IV mutant) and greatly reduced in abundance in *nrpe1-1* (a Pol V mutant) (Fig. 2A). For type II siRNAs, the levels of all tested siRNAs were slightly reduced in *nrpb2-3* compared with wild type (Fig. 2B,C). Although the differences were small, the reduction was highly reproducible in three to five biological replicates for each siRNA loci tested. siR02 was the most severely affected locus, with a 67% reduction in siRNA levels in *nrpb2-3* (Fig. 2B). Note that the reduction in type II siRNA levels in the *nrpb2-3* mutant was rescued by *NRPB2* genomic DNA (Fig. 2B,C), indicating that the siRNA defects were caused by the *nrpb2-3* mutation. In addition, we did not observe any reduction in type II siRNA levels in *nrpe1-1* (Fig. 2B,C), consistent with findings from previous studies (Pontier et al. 2005; Huettel et al. 2006; Mosher et al.

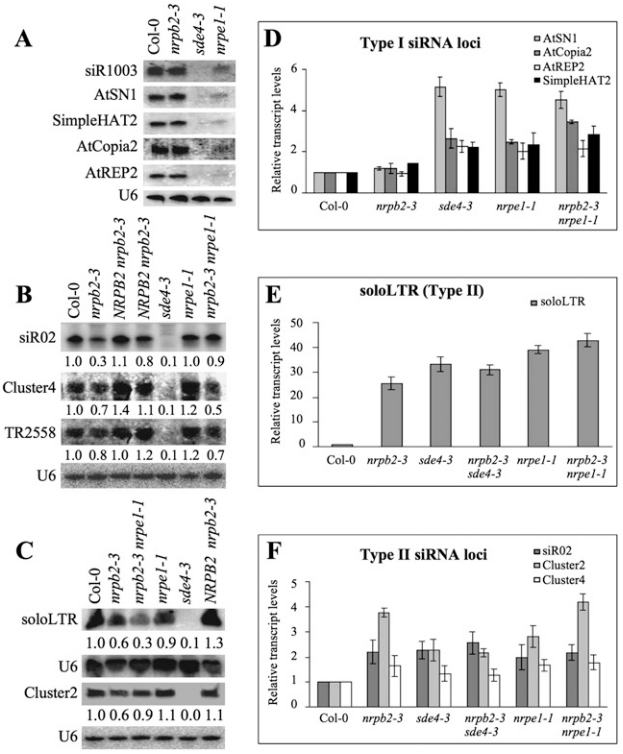


Figure 2. siRNA accumulation and TGS at heterochromatic loci in various genotypes. (*A–C*) siRNA accumulation in various genotypes. (Col-0) Wild type; (*nrpb2-3*) a Pol II mutant; (*sde4-3*) a Pol IV mutant; (*nrpe1-1*) a Pol V mutant; (*NRPB2 nrpb2-3*) the *nrpb2-3* mutant complemented with *NRPB2* genomic DNA; two independent transgenic lines were included in *B*. The numbers *below* the gel images indicate the relative levels of an siRNA among the different genotypes. The U6 blots served as loading controls for the small RNA blots *above*. (*D–F*) Real-time RT–PCR analysis of the expression of siRNA loci in various genotypes. Transcripts from the siRNA-generating region (region A in Fig. 3A) were detected. Transcript levels were normalized to those of *UBQ5* and the mutants were compared with wild type. Standard deviations were calculated from three technical replicates. The results shown were reproduced with three biological replicates.

2008). These results indicate that Pol II, but not Pol V, promotes the biogenesis of type II siRNAs.

Role of Pol II in TGS of type II siRNA loci

We next examined whether Pol II plays a role in siRNA-mediated TGS. We divided the siRNA loci into regions A and B, which represent the immediate siRNA-generating regions and the adjacent regions, respectively (Fig. 3A). We determined transcript levels from the siRNA-generating region A by real-time RT–PCR as a measure of the silencing status of the loci, as done in previous studies (Xie et al. 2004; Herr et al. 2005; Kanno et al. 2005; Huettel et al. 2006). Transcripts from type I loci accumulated at two to six times the wild-type levels in *sde4-3* and *nrpe1-1*, indicating derepression of these loci in Pol IV and Pol V mutants (Fig. 2D; Xie et al. 2004; Herr et al. 2005; Kanno et al. 2005). In contrast, transcript levels from type I loci in *nrpb2-3* were similar to those in wild type. However, since *nrpb2-3* is a weak allele, the potential role of Pol II in TGS at type I loci cannot be excluded. In fact, we observed a small increase in expression at SimpleHAT2 and AtCopia2 in the *nrpb2-3 nrpe1-1* double mutant as compared with the *nrpe1-1* single mutant (Fig. 2D).

For type II siRNA loci such as soloLTR, siR02, Cluster4, and Cluster2, we found similar amplitudes of derepression in *nrpb2-3* as compared with *sde4-3* or *nrpe1-1* (Fig. 2E,F). We next asked whether the silencing effect of *NRPB2* at type II loci was mediated by siRNAs or by an siRNA-independent mechanism. We examined the amplitudes of derepression in the *nrpb2-3 sde4-3* double mutant in comparison with the *sde4-3* single mutant, which completely lacks siRNAs from these loci. If the derepression effects observed in *nrpb2-3* are siRNA-dependent, there should not be any additive effects in the double mutant. There was no difference in transcript levels at the type II loci between *sde4-3* and *nrpb2-3 sde4-3* (Fig. 2E,F), suggesting that *NRPB2* acts through an siRNA-dependent mechanism. The similar amplitudes of derepression between Pol II and Pol IV mutants also suggested that the lack of TGS in the Pol II mutant was unlikely attributable to the small reduction in siRNA accumulation. Rather, Pol II likely acted downstream from siRNAs in TGS. Despite normal siRNA accumulation, type II siRNA loci were derepressed in *nrpe1-1* (Fig. 2E,F), consistent with previous reports (Huettel et al. 2006). Therefore, Pol V also acted downstream from the siRNAs in the silencing of type II loci. The amplitudes of type II loci derepression were comparable between the *nrpb2-3 nrpe1-1* double mutant and either single mutant (Fig. 2E,F), suggesting that Pol II and Pol V had non-redundant functions in the silencing of type II loci. Indeed, we observed a severe reduction in DNA methylation at siR02 (Supplemental Fig. S2A) and a reduction in H3K9me2 at most type II loci examined (Supplemental Fig. S2B) in both *nrpb2-3* and *nrpe1-1*.

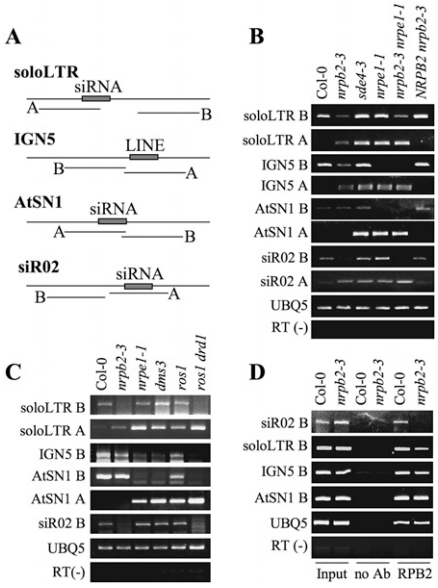


Figure 3. Detection of noncoding transcripts at siRNA loci. (*A*) Diagrams of soloLTR, IGN5, AtSN1, and siR02 genomic regions. The soloLTR, IGN5, and AtSN1 regions are based on analysis of transcription units by Wierzbicki et al. (2008). (*B*,*C*) RT–PCR analysis of noncoding transcripts at siRNA loci in various genotypes. (*D*) RIP with anti-RPB2 antibodies to test the association of Pol II with noncoding transcripts at siRNA loci. In *B–D*, *UBQ5* served as a loading control. The RT (−) control PCR was performed with *UBQ5* primers. The results shown were reproduced at least three times.

Pol II generates transcripts from adjacent regions of type II siRNA loci

Pol V generates specific RNA transcripts originating outside of some siRNA-producing loci, and these transcripts likely serve as scaffolds to recruit AGO4/siRNA RISCs to homologous genomic loci to facilitate RdDM (Wierzbicki et al. 2008, 2009). Intriguingly, RNA transcripts originating outside of the soloLTR locus (i.e., the scaffold transcripts) were present in wild-type plants and were just as abundant in the *nrpe1-1* mutant, and chromatin immunoprecipitation (ChIP) detected the presence of not only Pol V but also Pol II at the soloLTR chromatin (Wierzbicki et al. 2008). Given our finding that soloLTR was greatly derepressed in *nrpb2-3*, we hypothesized that Pol II was responsible for transcribing soloLTR-specific scaffold transcripts. To test this hypothesis, we detected transcripts from different positions near the soloLTR locus by RT–PCR (Fig. 3A,B). For the soloLTR B fragment (representing the scaffold transcript-generating region), we observed greatly reduced transcript levels in *nrpb2-3* (Fig. 3B). Consistent with previous findings (Wierzbicki et al. 2008), the transcript levels were not affected in either the *nrpe1-1* or the *sde4-3* mutant (Fig. 3B). Transcript levels from region A were increased in *nrpb2-3*, *sde4-3*, and *nrpe1-1* mutants relative to wild type (Fig. 3B), consistent with the derepression of this locus in mutants of all three polymerases.

We tested the accumulation of other Pol V-dependent scaffold transcripts detected by Wierzbicki et al. (2008) in *nrpb2-3*. At AtSN1 (a type I locus), transcripts from region A were greatly elevated in *sde4-3*, *nrpe1-1*, and *nrpb2-3*

nrpe1-1, but not in *nrpb2-3* (Fig. 3A,B), consistent with derepression of AtSN1 in the first three genotypes but not in *nrpb2-3* (Fig. 2D). The scaffold transcripts from region B were diminished in *nrpe1-1* but not affected in *nrpb2-3* (Fig. 3B). Surprisingly, at IGN5, an intergenic noncoding region for which scaffold transcripts were Pol V-dependent (Wierzbicki et al. 2008), we also detected reduced transcript levels in the *nrpb2-3* mutant (IGN5 B) (Fig. 3A,B). Consistently, derepression of IGN5 was observed in *nrpb2-3* as well as *nrpe1-1* (IGN5 A) (Fig. 3A,B).

Next, we tested whether Pol II generated transcripts from siR02, another type II locus like soloLTR and IGN5. For the immediate siRNA-producing region (region A), transcripts were found at elevated levels in *nrpb2-3*, *sde4-3*, and *nrpe1-1* mutants (Fig. 3B), consistent with derepression of siR02 in the three genotypes (Fig. 2F). Potential scaffold transcripts from region B (Fig. 3A) were also detected, and they were greatly reduced in *nrpb2-3* but unaffected in either Pol IV or Pol V mutants (Fig. 3B). Note that at all type II loci tested, the lack of, or reduced levels of, transcripts from region B was always accompanied by derepression at region A in the *nrpb2-3* mutant, implying that the Pol II transcripts were important for silencing. The fact that both defects were rescued simultaneously by the *NRPB2* genomic DNA (Fig. 3B) confirmed that the molecular defects were due to the *nrpb2-3* mutation.

The accumulation of Pol V-dependent scaffold transcripts requires DRD1, an SNF2-like chromatin remodeling protein, and DMS3, a structural maintenance of chromosome hinge domain-containing protein, both of which are required for RdDM (Kanno et al. 2004, 2008). We tested whether the two genes are required for the accumulation of Pol II-dependent transcripts at type II loci by RT–PCR. Consistent with previous observations, Pol V-dependent transcripts from IGN5 B and AtSN1 B were diminished in the *dms3* mutant (Fig. 3C; Wierzbicki et al. 2009). However, the Pol II-dependent transcripts at soloLTR and siR02 were not affected (Fig. 3C). The effects of *drd1* were assayed in the *ros1* background; compared with the *ros1* single mutant, the *ros1 drd1* double mutant lacked both Pol II-dependent transcripts at type II loci as well as Pol V-dependent transcripts at type I loci (Fig. 3C). DRD1 was found previously to be required for the accumulation of Pol V-dependent scaffold transcripts (Wierzbicki et al. 2008). Taken together, this study and previous studies show that *DMS3* is specifically required for the accumulation of Pol V transcripts, and *DRD1* is required for the accumulation of both Pol II- and Pol V-dependent scaffold transcripts.

Next, we tested whether Pol II was associated with the Pol II-dependent transcripts at type II loci by RNA immunoprecipitation (RIP). Anti-RPB2 polyclonal antibodies were used to immunoprecipitate NRPB2, and the associated RNAs were detected by RT–PCR. Two biological replicates gave identical results. While no transcripts were detected in the "no antibody" immunoprecipitate, Pol II-dependent transcripts at siR02 and soloLTR were coprecipitated with Pol II (Fig. 3D). In the *nrpb2-3* mutant, the abundance of the Pol II-associated RNAs was reduced, as expected (Fig. 3D). Pol II was also associated with Pol V-dependent scaffold transcripts at AtSN1 (Fig. 3D).

Association of Pol II with the chromatin at siRNA loci

To determine whether Pol II has a direct role in the production of the Pol II-dependent transcripts and in TGS at type II loci, we first examined Pol II occupancy at type II chromatin by ChIP using an anti-RPB1 antibody. eIF4A1 served as a positive control, whereas the same ChIP procedure performed without an antibody served as a negative control. Real-time PCR of coprecipitated DNA showed that all tested loci were enriched in the Pol II immunoprecipitate relative to the "no antibody" immunoprecipitate, suggesting that Pol II was present at all silenced loci in both regions A and B (Fig. 4). The presence of Pol II at soloLTR was also detected in a previous study (Wierzbicki et al. 2008). Interestingly, Pol II was also detected at regions A and B of AtSN1, a type I locus in two biological replicates (Fig. 4A, B; data not shown), whereas a previous study did not observe Pol II occupancy at AtSN1 above the levels for their "no antibody" control (Wierzbicki et al. 2009). The differences in the two studies could be due to different antibodies used against Pol II. At all tested loci, Pol II occupancy at region A was independent of Pol IV or Pol V (Fig. 4A) and, in the case of soloLTR, was slightly increased in the Pol II, Pol IV, and Pol V mutants, suggesting that Pol II transcribes the soloLTR region A when it is derepressed. Pol II occupancy at region B at all tested loci was also independent of Pol IV or Pol V (Fig. 4B).

Recruitment of AGO4 to type II siRNA loci by Pol II

Since 24-nt siRNAs are bound by the effector protein AGO4 to result in TGS, and Pol II is required for TGS at type II loci, we asked whether Pol II recruits AGO4/siRNAs

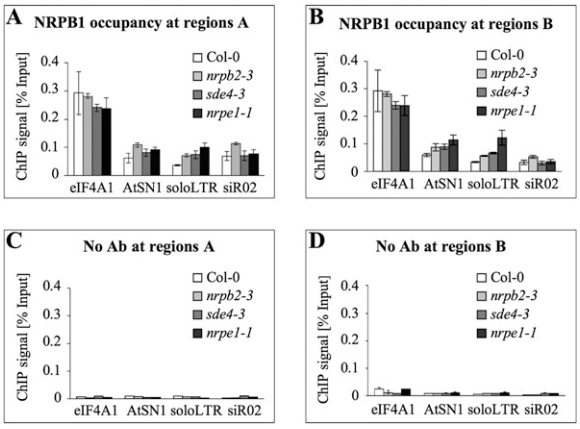


Figure 4. Pol II is present at siRNA loci. (*A*,*B*) Real-time PCR of DNA that copurified with Pol II in ChIP performed in various genotypes with an anti-RPB1 antibody. The eIF4A1 gene served as a positive control. (*C*,*D*) The "No Ab" (no antibody) immunoprecipitate served as negative controls. The ChIP signal was quantified as the percentage of total input DNA. Two biological replicates were performed and identical results were obtained. Standard deviations were calculated from three technical repeats.

to type II loci. We crossed a functional *myc-AGO4* transgene (Li et al. 2006) into the *nrpb2-3* background and performed ChIP using an anti-myc antibody to examine the occupancy of myc-AGO4 at type I and type II loci. Region A at type I (AtSN1) and type II (soloLTR and siR02) loci, but not the gene eIF4A1, was enriched in the myc-AGO4 ChIP (Fig. 5A). AGO4 association to region A of AtSN1 and soloLTR was also detected, and was found to be Pol V-dependent in a previous study (Wierzbicki et al. 2009). Interestingly, myc-AGO4 was not associated with region B of AtSN1, soloLTR, or siR02 (Fig. 5B). myc-AGO4 occupancy at region A of soloLTR and siR02 was strongly reduced in the *nrpb2-3* mutant, but that at region A of AtSN1 was not (Fig. 5A). Therefore, the dependence on Pol II for AGO4 recruitment at siRNA loci correlates with the production of scaffold transcripts by Pol II. Western blotting showed that myc-AGO4 protein levels were not detectably different between wild type and *nrpb2-3* (Supplemental Fig. S3A).

While the Pol II-dependent scaffold transcripts may serve to recruit siRNA/AGO4 RISCs to type II chromatin, physical interaction between Pol II and AGO4 may also contribute to the recruitment of AGO4 to chromatin. The extended C-terminal domain (CTD) of NRPE1 harbors 16 GW/WG motifs that provide an AGO4-binding platform essential for RdDM (Li et al. 2006; El-Shami et al. 2007). We found four GW/WG motifs in NRPB2, three of which are conserved with those in fission yeast RPB2 (Fig. 1B). Since one GW/WG motif is sufficient for AGO4 binding in vitro (El-Shami et al. 2007), it is possible that the GW/WG motifs in NRPB2 mediate the interaction between Pol II and AGO4. We first performed coimmunoprecipitation (co-IP) to determine whether Pol II and AGO4 were associated in vivo. Anti-RPB1 antibodies were able to immunoprecipitate NRPB2 as well as NRPB1, as expected (Fig. 5C). Anti-RPB1 antibodies also coprecipitated myc-AGO4, but not HEN1, a negative control (Fig. 5C). Conversely, an anti-myc antibody specifically immunoprecipitated myc-AGO4 and NRPB1, but not HEN1 (a negative control) (Fig. 5D).

We next performed GST pull-down experiments with recombinant NRPB2 expressed in *Escherichia coli*. A 900-amino-acid region containing the GW/WG motifs (GWR) was fused to GST and purified with glutathione beads. GST-NRPE1 CTD (Li et al. 2006; El-Shami et al. 2007) and GST were also included as a positive and a negative control, respectively. In multiple biological replicates, we observed myc-AGO4, but not HEN1, binding to the NRPB2 GWR and to NRPE1 CTD (Supplemental Fig. S4). The GWR containing the *nrpb2-3* mutation (NRPB2-3 GWR) also pulled down myc-AGO4 as efficiently as wild-type GWR (Supplemental Fig. S4). Recombinant GWR with four GW/WG motifs mutated (NRPB2 GWRm) showed reduced binding to AGO4 (Supplemental Fig. S4). These data, together with the co-IP results, strongly suggest that Pol II and AGO4 interact with each other.

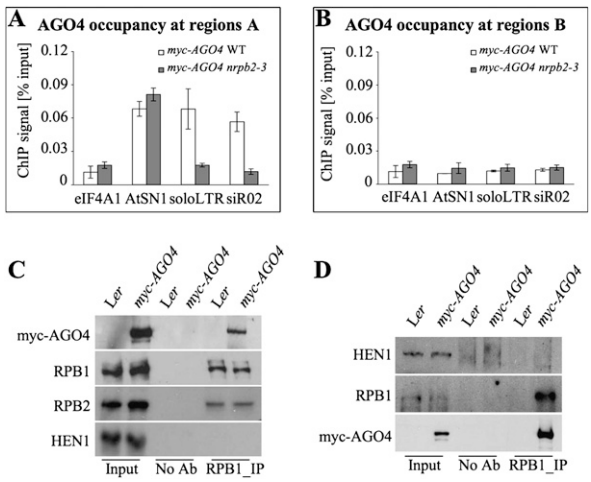


Figure 5. Recruitment of AGO4 by Pol II. (*A*,*B*) AGO4 occupancy at regions A and B of siRNA loci. ChIP was performed using an anti-myc antibody and DNA that copurified with myc-AGO4 was quantified by real-time PCR. Region A, but not B, was enriched in the myc-AGO4 immunoprecipitate, as compared with the eIF4A1 negative control. (WT) Wild type. (*C*,*D*) Co-IP to test the association between AGO4 and Pol II in vivo with anti-RPB1 and anti-myc antibodies, respectively. (L*er*) Wild type; a negative control for *myc-AGO4*. Proteins from one-sixth of "No Ab" (no antibody) immunoprecipitate and one-sixth of immunoprecipitate were analyzed by Western blotting. Proteins from 1/100 of input were included for anti-RPB1, anti-RPB2, and anti-HEN1 blots, while proteins from 1/1000 of input were used for the anti-myc blot.

Pol II recruits Pol V to type II loci

At soloLTR, Pol V binds to scaffold transcripts (Wierzbicki et al. 2008), which we showed to be produced by Pol II. Therefore, it is likely that Pol V is recruited to soloLTR chromatin by Pol II-dependent transcripts or the act of Pol II transcription itself. To determine whether the association of Pol V with soloLTR and other type II loci depends on Pol II, we determined Pol V occupancy at regions A and B at AtSN1 and two type II loci in wild type and *nrpb2-3*. A functional *NRPE1-Flag* transgene (Pontes et al. 2006) was crossed into *nrpb2-3* and ChIP was conducted with anti-Flag antibodies. Pol V occupancy at AtSN1 was not affected by the *nrpb2-3* mutation in either region A or B (Fig. 6A,B). However, Pol V occupancy at soloLTR and siR02 was strongly reduced at region B (Fig. 6B) but largely unaffected at region A (Fig. 6A). This suggests that the Pol II-dependent transcripts or Pol II transcription recruits Pol V to region B of type II loci. Western blots showed that the levels of NRPE1-Flag protein were not different between wild type and *nrpb2-3* (Supplemental Fig. S3B).

Pol II recruits Pol IV to type II loci

We asked how Pol II promotes siRNA biogenesis at type II loci through Pol IV. The molecular function of Pol IV in siRNA biogenesis is still unknown. In one model, Pol IV transcribes heterochromatic chromatin into single-stranded transcripts that serve as siRNA precursors. In another model, Pol IV transcribes double-stranded transcripts generated by RDR2. In nuclei, numerous Pol IV

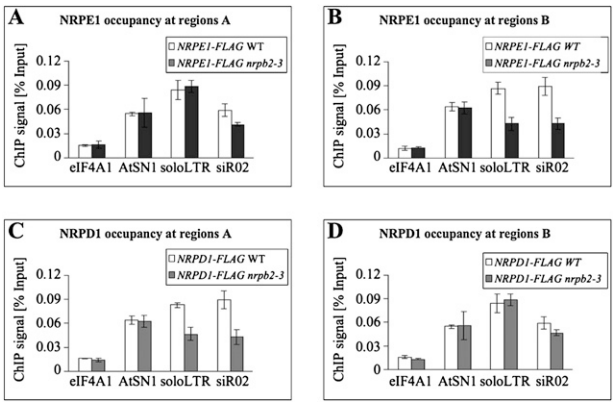


Figure 6. Pol V and Pol IV occupancy at regions A and B of siRNA loci. ChIP was performed against NRPE1-Flag (*A*,*B*) or NRPD1-Flag (*C*,*D*), and the coprecipitated DNA was quantified by real-time PCR and expressed as the percentage of input DNA. eIF4A1 served as a negative control. Standard deviations were calculated from three technical repeats. The results shown were reproduced in two biological replicates. (WT) Wild type.

foci are sensitive to RNase treatment, suggesting that Pol IV is anchored by RNA (Pontes et al. 2006).

We examined whether Pol IV is present at heterochromatic chromatin, which has not been tested before, and whether its occupancy depends on Pol II. ChIP using anti-Flag antibodies against a wild-type line with a functional *NRPD1-Flag* transgene (Pontes et al. 2006) showed that both regions A and B at type I and II loci, but not the negative control eIF4A, were enriched in the Pol IV ChIP (Fig. 6C,D). Therefore, Pol IV is present at type I and type II loci, consistent with the known role of Pol IV in siRNA biogenesis at heterochromatic loci but not at genes. We next evaluated whether Pol IV occupancy at these loci depended on Pol II by ChIP from *NRPD1-Flag nrpb2-3*. Pol IV occupancy at either region A (Fig. 6C) or B (Fig. 6D) of AtSN1 was not affected in the *nrpb2-3* mutant, consistent with the lack of an observable effect of the *nrpb2-3* mutation on siRNA biogenesis at this locus. At the two type II loci, Pol IV occupancy at region A was strongly reduced in *nrpb2-3* (Fig. 6C), whereas that at region B (Fig. 6D) was not. The levels of NRPD1-Flag were largely the same between wild type and *nrpb2-3* (Supplemental Fig. S3C). Therefore, Pol II recruits Pol IV to region A of type II loci for siRNA production.

Discussion

Pol II has a direct role in TGS

The isolation of the reduction-of-function *nrpb2-3* allele allowed us to address the role of Pol II in heterochromatic siRNA-mediated TGS in *Arabidopsis*. Pol IV and Pol V share the same second-largest subunit, NRPD2/NRPE2, which is distinct from NRPB2 (Kanno et al. 2005; Onodera et al. 2005; Pontier et al. 2005). Immunoprecipitation using anti-Flag antibodies from an *NRPB2-Flag* line recovered all 12 Pol II subunits and no subunits specific to Pol I, Pol III, Pol IV, or Pol V (Ream et al. 2009), suggesting that NRPB2 does not reside in polymerases other than Pol II. Targeted co-IP experiments between NRPB2 and Pol V-specific subunits also demonstrated that NRPB2 is not found in Pol V (Ream et al. 2009). In addition, our findings that some scaffold transcripts are reduced in *nrpb2-3* but not in *nrpe1-1* or *sde4-3* also imply that NRPB2 does not function together with the largest subunits of Pol IV or Pol V. Therefore, the functions of *NRPB2* discovered in our study are likely those of Pol II.

Our results strongly indicate a direct role of Pol II in TGS at low-copy-number, intergenic loci. First, whole-genome microarray as well as targeted real-time RT–PCR analysis did not reveal reduced levels of expression of known genes in siRNA biogenesis or siRNA-mediated TGS. Second, Pol II is present at the chromatin of type II loci, where Pol II-dependent transcripts are made, and is associated with these transcripts. This suggests that Pol II directly generates these noncoding transcripts.

Relationship between Pol II and Pol V

Our studies show that both Pol II and Pol V are required for siRNA-mediated TGS at type II loci. What is the relationship between Pol II and Pol V? One possibility is that the two have partially redundant functions and both contribute to TGS; however, we rarely observed an additive effect of *nrpb2-3* and *nrpe1-1* mutations (in terms of the amplitudes of derepression), which suggests that the two have separate, nonredundant functions. The separate roles of the two polymerases at these loci are currently unknown, but we show that the presence of Pol V at region B of type II loci requires Pol II, whereas the presence of Pol II at type II loci is independent of Pol V. Moreover, both Pol II and Pol V are required for the recruitment of AGO4 to these loci (Wierzbicki et al. 2009; this study). We propose two models that are consistent with these observations to explain the nonredundant functions of Pol II and Pol V in TGS at type II loci. The first model is based on an additional observation that Pol V and AGO4 are associated in the Cajal body during siRNA biogenesis but are not colocalized in the nucleoplasm (Pontes et al. 2006), where TGS occurs. We hypothesize that Pol V delivers AGO4/siRNA RISCs to type II loci to allow AGO4/siRNA RISCs to interact with Pol II or Pol II transcripts, which originate in region B and perhaps extend into region A (Fig. 7A). In this model, Pol V is recruited to type II loci at region B by either Pol II transcripts or the act of Pol II transcription in region B, and its role in TGS lies mainly in the delivery of AGO4/siRNAs to Pol II. AGO4 is only present in region A due to the pairing between siRNAs, which are only found in region A, to Pol II transcripts. In the second model, Pol II and Pol V act sequentially to recruit AGO4/siRNA RISCs to region A (Fig. 7B). First, Pol II transcription in region B or the Pol II transcripts recruit Pol V to region B. Pol V then transcribes through region A, and the Pol V transcripts recruit AGO4/siRNA RISCs to region A. This model is consistent with the observation that mutations in the putative active site residues in Pol V abolish TGS,

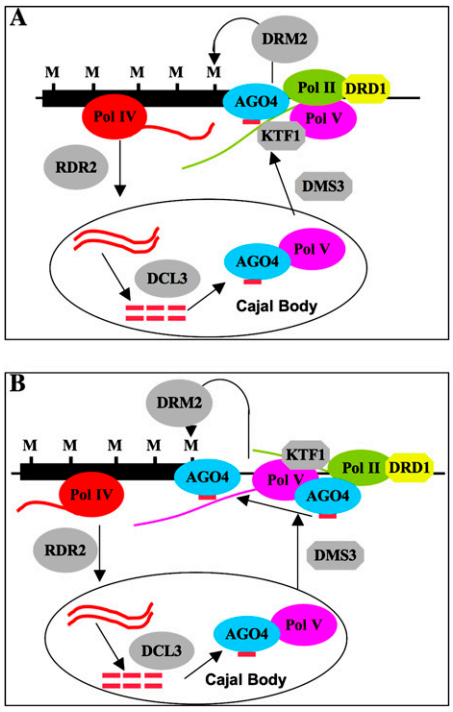


Figure 7. Models for the roles of the three polymerases in TGS at type II loci in *Arabidopsis*. In both models, no modifications to the prevailing views on the roles of Pol IV, RDR2, DCL3, and AGO4 are made except that we showed that Pol IV is associated with the chromatin at siRNA loci. The two models differ in the roles of Pol II and Pol V in TGS. (*A*) Pol II generates scaffold transcripts that originate in region B and that traverse region A. The Pol II transcripts or Pol II transcription itself recruits Pol V, which is in complex with AGO4 during siRNA biogenesis in the Cajal body, to region B. Pol V delivers AGO4 to Pol II and AGO4/siRNA RISCs localize to region A due to the sequence complementarity between siRNAs, which are only derived from region A, to Pol II transcripts. (*B*) Pol II generates scaffold transcripts that correspond only to region B. These transcripts recruit Pol V to region B and Pol V makes new scaffold transcripts that traverse region A. These transcripts help anchor AGO4/siRNAs to region A. In both models, AGO4/siRNA RISCs then recruit, directly or indirectly, the DNA methyltransferase DRM2 or histone methyltransferases to result in epigenetic modifications and silencing. KTF1 may bind to Pol II transcripts to facilitate the recruitment of Pol V by Pol II transcripts. DRD1 is required for the production of Pol II transcripts. DMS3 acts downstream from Pol II transcripts to facilitate the recruitment of AGO4. (M) DNA methylation.

which suggests that Pol V transcription is necessary for TGS (Wierzbicki et al. 2008).

Roles of several known genes in TGS

Our findings also prompt the re-evaluation of current models of how several proteins—such as KTF, DRD1, and DMS3—act in TGS. In either of our models (Fig. 7), the recently identified factor KTF1, which associates with Pol V and AGO4 and binds ssRNAs (He et al. 2009; Huang et al. 2009), may be the link between Pol II and Pol V, such that KTF1 binds Pol II transcripts to recruit Pol V to region B. The putative chromatin remodeling factor DRD1 was thought previously to program Pol V transcription at heterochromatic loci, since Pol V occupancy at these loci (including soloLTR) requires DRD1 (Wierzbicki et al. 2008). We show that DRD1 is not specific to Pol V—it is also required for the accumulation of Pol II transcripts. Given that Pol II is required for the recruitment of Pol V to soloLTR, and that Pol II transcripts at soloLTR are diminished in the *drd1* background, the previously observed dependence on DRD1 for Pol V occupancy at soloLTR is likely an indirect effect of the lack of Pol II-dependent transcripts in the mutant. DMS3, however, is not required for the production of Pol II transcripts at type II loci (this study), but is necessary for AGO4 recruitment to siRNA loci including soloLTR (Wierzbicki et al. 2009). DMS3 must act at a step downstream from Pol II transcription in the recruitment of AGO4.

Feed-forward loops in siRNA biogenesis and TGS mediated by three polymerases

Our findings provide new insights into the molecular mechanisms underlying the feed-forward loop involving siRNA biogenesis and siRNA-mediated TGS. At both type I and type II loci, Pol IV is essential for siRNA biogenesis, probably by transcribing these loci to generate siRNA precursors (Herr et al. 2005; Onodera et al. 2005; this study). The greatly reduced accumulation of type I siRNAs in *nrpe1* mutants has been hypothesized to be an indirect effect of compromised DNA methylation (Pontier et al. 2005). It is thought that DNA methylation recruits Pol IV to generate more siRNAs in a feed-forward loop. siRNA accumulation at type II loci is not affected in Pol V mutants (Fig. 2; Herr et al. 2005; Pontier et al. 2005; Huettel et al. 2006), but is reduced in *nrpb2-3* (Fig. 2). Therefore, Pol II, instead of Pol V, plays a role in the feed-forward loop to promote siRNA production at type II loci. Since DNA methylation and/or H3K9me2 are affected at these loci in both Pol II and Pol V mutants, and yet Pol V mutants are not affected in siRNA accumulation, it is unlikely that it is the chromatin marks that recruit Pol IV in the feed-forward loop for siRNA biogenesis. Instead, we propose that it is Pol II (or the RNAs it generates) that recruits Pol IV. In fact, we show for the first time that Pol IV is associated with siRNA-generating loci, and that Pol IV occupancy at the immediate siRNA-generating region (region A) requires Pol II. By analogy, we propose that at type I loci, it is not DNA methylation that recruits Pol IV in the feed-forward loop. Instead, it is Pol V (or the Pol V-dependent transcripts) that recruits Pol IV. Alternatively, the Pol II or Pol V scaffold transcripts serve as templates for Pol IV at type II and type I loci, respectively.

Evolutionary implications

In fission yeast, Pol II is required for the biogenesis of heterochromatic siRNAs and for siRNA-directed heterochromatin assembly at pericentromeric repeats (Djupedal et al. 2005; Kato et al. 2005). Release of gene silencing, accompanied by loss of siRNA accumulation and reduced H3K9 methylation, was observed in an *rpb2* mutant.

Pol II-dependent RNA transcripts at pericentromeric repeats probably serve a dual function of being precursors to siRNAs and scaffold transcripts to recruit siRNAs. In plants, Pol IV is thought to transcribe heterochromatic loci to produce transcripts that are channeled into siRNA biogenesis (Herr et al. 2005; Onodera et al. 2005; Pontier et al. 2005). Pol V transcribes the surrounding regions of siRNA-producing loci to generate transcripts that recruit AGO4/siRNAs (Wierzbicki et al. 2008, 2009). Therefore, Pol IV and Pol V in plants appear to specialize in one of the two functions of fission yeast Pol II in TGS. Our studies showed that the plant Pol II has a central role in siRNA biogenesis and siRNA-mediated TGS at low-copy-number intergenic loci. Pol II recruits Pol IV and Pol V to regions A and B at these loci to promote siRNA biogenesis and TGS, respectively. We propose that TGS is one of the ancestral functions of the plant Pol II. If Pol II does not act at type I loci, then the silencing function of Pol II has been devolved to Pol IV and Pol V at these high-copy-number loci. This would imply that Pol IV and Pol V have evolved to recognize certain features (perhaps copy number) of these loci. Alternatively, Pol II also has a crucial role in TGS at these loci, which is not evident from the weak *nrpb2-3* mutant.

A silencing role of intergenic noncoding Pol II transcripts

Transcriptome studies have revealed widespread intergenic transcription, presumably by Pol II, in animal and plant genomes to result in numerous stable and unstable noncoding transcripts (Kapranov et al. 2002, 2007; Davis and Ares 2006; Chekanova et al. 2007; Neil et al. 2009; Xu et al. 2009). Noncoding transcripts from Pol II regulate the epigenome by targeting chromatin modification machineries to specific loci. For example, many noncoding RNAs—including, but not limited to, HOTAIR and Xist—copurify with the Polycomb group PRC2 H3K27 trimethyltransferase, and, in the case of HOTAIR and Xist, the noncoding RNAs target the PRC2 to specific loci (Rinn et al. 2007; Zhao et al. 2008; Khalil et al. 2009). Our work adds to the utility of noncoding, intergenic transcription by Pol II by showing that it is a means to shape the epigenome through endogenous siRNAs.

Materials and methods

Plant mutants and antibodies

sde4-3 (SALK_128428) (Herr et al. 2005) and *nrpe1-1* (previously known as *nrpd1b-1*, SALK_029919) (Pontier et al. 2005) seeds are kind gifts from Drs. David Baulcombe and Thierry Lagrange, respectively. The *myc-AGO4* transgenic line (Li et al. 2006) was provided by Dr. Steve Jacobsen. The plasmid GST-NRPE1 CTD (Li et al. 2006; El-Shami et al. 2007) was provided by Dr. Thierry Lagrange. *dms3* (SALK_068723c), *ros1*, and *drd1 ros1* seeds were kindly provided by Dr. Jian-kang Zhu. RPB1 (catalog no. ab5408, Abcam), RPB2 (catalog no. ab10338, Abcam), H3K9me2 (catalog no. ab1220, Abcam), Flag (F7425, Sigma), myc (sc-70463, Santa Cruz Biotechnologies), and Hsp70 (SPA-818, Stressgen) antibodies were purchased from various vendors. Anti-HEN1 antibodies were generated in our laboratory. The anti-RPB1 monoclonal antibody was raised against the peptide YSPTSPS, which is repeated many times in the CTD of RPB1 from many species, including *Arabidopsis*. It only detected one band of the correct size for NRPB1 from total *Arabidopsis* protein extracts. Anti-RPB2 polyclonal antibodies were raised against a peptide in the human RPB2 with 95% sequence identity to a homologous sequence in *Arabidopsis* NRPB2. Anti-RPB2 antibodies only recognized one band of the expected size for NRPB2 from total *Arabidopsis* protein extracts.

Plasmid construction

NRPB2 genomic DNA was amplified from Col-0 genomic DNA with primers NRPB2F1 and NRPB2R1, and was cloned into the pPZP211 binary vector at the BamHI and SalI restriction sites to generate *pNRPB2::NRPB2*. The plasmid was introduced into *nrpb2-3* plants by Agro infiltration. The cDNA of the GW/WG-rich region of *NRPB2* was amplified from Col-0 and the *nrpb2-3* mutant by RT–PCR with primers NRPB2F4 and NRPB2R4, and was cloned into the pGEX-2TK vector. Each of the four GW/WG motifs was changed to AA or VA by site-directed mutagenesis with primers NRPB2F6/R6, NRPB2F7/R7, NRPB2F8/R8, or NRPB2F9/R9. Expression and purification of GST fusion proteins in *E. coli* were done as described (Yang et al. 2007). The primer sequences are shown in Supplemental Table S3.

ATH1 Affymetrix microarray analysis

GeneChip arrays were hybridized according to the manufacturer's instructions (Affymetrix). Data analysis was done according to Horan et al. (2008). Normalization of raw intensities across all probe sets was performed in R using RMA algorithms. To calculate *P*-values for increase or decrease in expression, the Wilcoxon signed-rank test was applied to each pair of chips after normalization using R&BioConductor (developed by Thomas Girke). A *P*-value of ≤ 0.05 in combination with a twofold difference was used to define changes in gene expression.

Real-time RT–PCR analysis of endogenous siRNA loci

cDNA was synthesized from 3–5 μg of total RNA using RT (Invitrogen) and oligo-dT primers. Quantitative PCR was performed in triplicate on a Bio-Rad IQ cycler apparatus with the Quantitech SYBR green kit (Bio-Rad). The primers used are listed in Supplemental Table S3.

siRNA Northern blotting

RNA isolation and hybridization for endogenous siRNAs were performed as described (Park et al. 2002). 5′-End-labeled (^{32}P) antisense LNA oligonucleotides were used to detect AtSN1, TR2558, Cluster4, AtREP2, AtCopia2, SimpleHAT2, IR-71, siRNA02, and siRNA1003. To detect Cluster2 siRNAs, a DNA fragment was amplified from genomic DNA with primers Cluster2F2/Cluster2R2, gel-purified, and labeled by random priming. To detect soloLTR siRNAs, a DNA fragment was amplified from genomic DNA with soloLTRF2/soloLTRR2 primers, gel-purified, and ligated into the pGEM-T Easy vector. The plasmid was digested by SpeI and used as a template for in vitro transcription with T7 RNA polymerase in the presence of [α^{32}P]UTP. Radioactive signals were detected and quantified with a PhosphorImager.

Co-IP

For co-IP experiments, 0.5 g of L*er* and *myc-AGO4* inflorescences were ground in liquid nitrogen and homogenized in

1 mL of protein lysis buffer (50 mM Tris-HCl at pH 7.5, 150 mM NaCl, 0.2% NP-40, 2 mM DTT, 10% glycerol, complete protease inhibitor cocktail [Roche]). The extracts were incubated for 50 min at 4°C on a rotating wheel and centrifuged at 16,000*g* to pellet any debris. The centrifugation was repeated three more times to remove the debris completely. The lysate was then precleared for 20 min with protein-A agarose. Half of the lysate was mixed with anti-RPB1 or anti-Myc antibodies precoupled to Protein A or Protein G agarose beads, respectively. The other half was mixed with beads alone. After a 2-h incubation at 4°C, the immune complexes were washed five times with 1 mL of lysis buffer. The volume was reduced to 30 μL in the end. Protein samples were analyzed on 8% SDS-PAGE. Anti-myc, anti-RPB1, anti-RPB2, and anti-HEN1 antibodies were used for Western blotting.

GST-NRPB2 pull-down

GST-NRPB2 pull-down experiments were performed according to Li et al. (2006) with slight modifications. GST, GST-NRPE1 CTD, GST-NRPB2 GWR, GST-NRPB2-3 GWR, and GST-NRPB2 GWRm were expressed in *E. coli* BL21. Expression was induced for 3 h with 0.5 mM IPTG at 30°C. Cells were disrupted by sonication, and the fusion proteins were purified by glutathione Sepharose 4B affinity chromatography. Six-hundred microliters of inflorescence extract were applied to the beads–protein mixture (30 μg of protein total) and incubated for 2 h at 4°C on a rotating wheel. The beads were washed five times with immunoprecipitation lysis buffer. Proteins bound by the beads and those in the supernatant were resolved in 8% SDS-PAGE gels and subjected to Western blotting.

ChIP

ChIP was performed according to Xie et al. (2004) and Huettel et al. (2006). The results shown were reproduced in three biological replicates. The primer sets used for the PCR are listed in Supplemental Table S3.

Detection of Pol II-dependent transcripts by RT–PCR

RT–PCR experiments were done according to Wierzbicki et al. (2008), with some modifications. cDNA was synthesized from 5 μg of total RNA using RT (Invitrogen) and gene-specific primers. PCR was performed using hot-start platinum Taq DNA polymerase (Invitrogen). For siR02, the temperature-cycling scheme was as follows: 30 sec at 94°C, 30 sec at 54°C, 30 sec at 72°C, 38 cycles, and a final extension for 10 min at 72°C. The PCR conditions for AtSN1 B, soloLTR B, and IGN5 B were according to Wierzbicki et al. (2008). The results shown were reproduced at least three times. The primers used are listed in Supplemental Table S3.

RIP

RIP was performed according to Wierzbicki et al. (2008). The primers used are listed in Supplemental Table S3.

Acknowledgments

We are grateful to David Baulcombe, Steve Jacobsen, Thierry Lagrange, Craig Pikaard, and Jian-Kang Zhu for sharing plant material and plasmids. We thank Jian-Kang Zhu, So Youn Won, Vanitharani Ramachandran, and Theresa Dinh for comments on the manuscript. The work was supported by grants from the National Institutes of Health (GM61146) and National Science Foundation (MCB-0718029) to X.C.

References

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

Chekanova JA, Gregory BD, Reverdatto SV, Chen H, Kumar R, Hooker T, Yazaki J, Li P, Skiba N, Peng Q, et al. 2007. Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the *Arabidopsis* transcriptome. *Cell* **131:** 1340–1353.

Davis CA, Ares M Jr. 2006. Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci* **103:** 3262–3267.

Djupedal I, Portoso M, Spahr H, Bonilla C, Gustafsson CM, Allshire RC, Ekwall K. 2005. RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. *Genes & Dev* **19:** 2301–2306.

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

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

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

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

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

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

Horan K, Jang C, Bailey-Serres J, Mittler R, Shelton C, Harper JF, Zhu JK, Cushman JC, Gollery M, Girke T. 2008. Annotating genes of known and unknown function by large-scale coexpression analysis. *Plant Physiol* **147:** 41–57.

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

Huettel B, Kanno T, Daxinger L, Aufsatz W, Matzke AJ, Matzke M. 2006. Endogenous targets of RNA-directed DNA methylation and Pol IV in *Arabidopsis*. *EMBO J* **25:** 2828–2836.

Kanno T, Mette MF, Kreil DP, Aufsatz W, Matzke M, Matzke AJ. 2004. Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr Biol* **14:** 801–805.

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

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

Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, Fodor SP, Gingeras TR. 2002. Large-scale transcriptional activity in chromosomes 21 and 22. *Science* **296:** 916–919.

Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermuller J, Hofacker IL, et al. 2007. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316:** 1484–1488.

Kato H, Goto DB, Martienssen RA, Urano T, Furukawa K, Murakami Y. 2005. RNA polymerase II is required for RNAi-dependent heterochromatin assembly. *Science* **309:** 467–469.

Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, et al. 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci* **106:** 11667–11672.

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

Li CF, Pontes O, El-Shami M, Henderson IR, Bernatavichute YV, Chan SW, Lagrange T, Pikaard CS, Jacobsen SE. 2006. An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. *Cell* **126:** 93–106.

Malagnac F, Bartee L, Bender J. 2002. An *Arabidopsis* SET domain protein required for maintenance but not establishment of DNA methylation. *EMBO J* **21:** 6842–6852.

Moazed D. 2009. Small RNAs in transcriptional gene silencing and genome defence. *Nature* **457:** 413–420.

Mosher RA, Schwach F, Studholme D, Baulcombe DC. 2008. PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proc Natl Acad Sci* **105:** 3145–3150.

Naumann K, Fischer A, Hofmann I, Krauss V, Phalke S, Irmler K, Hause G, Aurich AC, Dorn R, Jenuwein T, et al. 2005. Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. *EMBO J* **24:** 1418–1429.

Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A. 2009. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* **457:** 1038–1042.

Onodera Y, Haag JR, Ream T, Nunes PC, Pontes O, Pikaard CS. 2005. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120:** 613–622.

Onodera Y, Nakagawa K, Haag JR, Pikaard D, Mikami T, Ream T, Ito Y, Pikaard CS. 2008. Sex-biased lethality or transmission of defective transcription machinery in *Arabidopsis*. *Genetics* **180:** 207–218.

Park W, Li J, Song R, Messing J, Chen X. 2002. CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr Biol* **12:** 1484–1495.

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

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

Ream TS, Haag JR, Wierzbicki AT, Nicora CD, Norbeck AD, Zhu JK, Hagen G, Guilfoyle TJ, Pasa-Tolic L, Pikaard CS. 2009. Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol Cell* **33:** 192–203.

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, et al. 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129:** 1311–1323.

Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297:** 1833–1837.

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

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

Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC. 2004. Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol* **2:** E104. doi: 10.1371.journal.pbio.0020104.

Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Munster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM. 2009. Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457:** 1033–1037.

Yang Z, Vilkaitis G, Yu B, Klimasauskas S, Chen X. 2007. Approaches for studying microRNA and small interfering RNA methylation in vitro and in vivo. *Methods Enzymol* **427:** 139–154.

Zhang X, Henderson IR, Lu C, Green PJ, Jacobsen SE. 2007. Role of RNA polymerase IV in plant small RNA metabolism. *Proc Natl Acad Sci* **104:** 4536–4541.

Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. 2008. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322:** 750–756.

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

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

Supplemental Material (Zheng et al.)

FIGURE LEGENDS

Figure S1. Real-time RT-PCR analysis of transcript levels of genes involved in endogenous siRNA biogenesis (A), DNA methylation or demethylation (B), or histone methylation (C). RNAs from inflorescences from wild type (Col-0) and *nrpb2-3* were used.

Figure S2. Analysis of DNA methylation and H3K9 dimethylation in various genotypes. (A) CG, CHG, and CHH methylation (H being A, C, or T) at siR02 was analyzed by bisulfite sequencing. Methylation levels are shown as the percentage of methylated cytosine among total cytosine in all sequenced clones. Methylation levels were reduced in both *nrpb2-3* and *nrpe1-1* mutants. (B) Analysis of H3K9 dimethylation at type II loci by ChIP. After IP using antibodies against dimethyl histone H3K9 (H3K9me2), associated genomic DNA was analyzed by real-time PCR. The ChIP signal was quantified as the percentage of total input DNA. Three biological replicates were performed and gave identical results. Standard deviations were calculated from three technical repeats.

Figure S3. Western blotting to determine the levels of myc-AGO4 (A), NRPE1-FLAG (B), and NRPD1-FLAG (C) levels in wild type and *nrpb2-3* inflorescences. Anti-myc (A) and anti-FLAG (B and C) antibodies were used. Hsp70 served as a loading control. L*er* and Col-0 were the wild-type backgrounds into which the transgenes had been introduced.

Figure S4. AGO4 associates with the NRPB2 GW/WG-rich domain *in vitro*. (A) GST pull-down experiments with various purified GST fusion proteins and myc-AGO4. Inflorescence extracts from *myc-AGO4* plants were applied to GST, GST-NRPE1 CTD, GST-NRPB2 GWR, or GST-NRPB2 GWRm beads. After incubation and extensive washes, the bound and unbound (supernatant) proteins were detected by western blotting with anti-myc and anti-HEN1 antibodies. 1/6 of pellet fractions, and 1% and 0.1% of the supernatant fractions (for HEN1 and myc-AGO4 blots, respectively) were used. (B) A stained protein gel showing the proteins used for the GST pull-down assay. 1/30 of the amount of each protein used in (A) was resolved in an

SDS-PAGE gel. GST-NRPE1 CTD, the CTD domain of NRPE1 fused to GST; it served as a positive control. GST-NRPB2 GWR, the GW/WG-rich region of NRPB2 fused to GST; GST-NRPB2-3 GWR, the GW/WG-rich region of NRPB2-3 fused to GST; GST-NRPB2 GWRm, same as GST-NRPB2 except that the four GW/WG motifs were mutated. The full-length proteins were marked by arrowheads.

SUPPLEMENTAL TABLES

Table S1. Genes down-regulated in the *nrpb2-3* mutant by Affymetrix microarray analysis (see supplemental spreadsheets in a separate file)

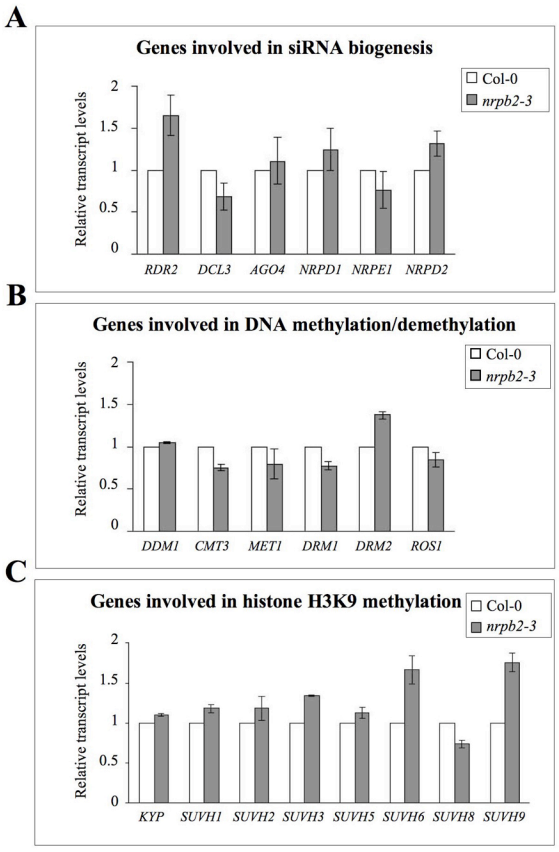
Table S2. Genes up-regulated in the *nrpb2-3* mutant by Affymetrix microarray analysis (see supplemental spreadsheets in a separate file)

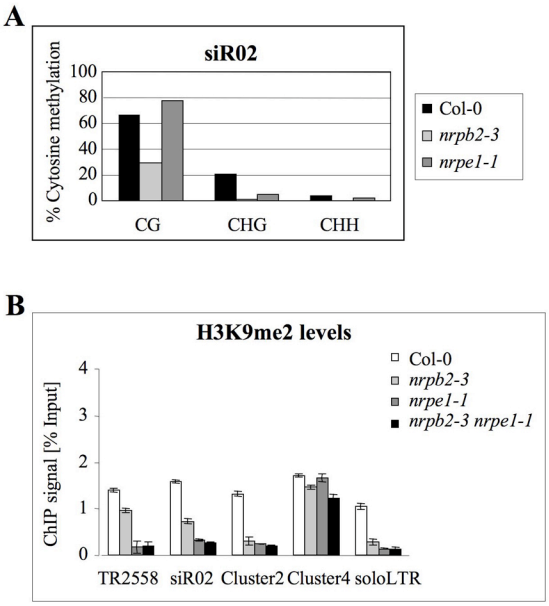
Table S3. Oligonucleotides used in this study

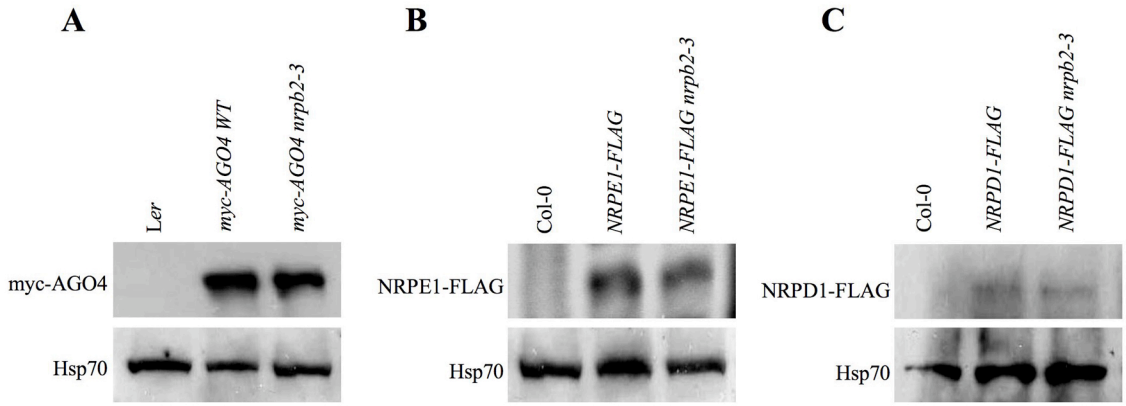
Name	Sequence (5'-3')	Application
N_UBQ5	GGTGCTAAGAAGAGGAAGAAT	
C_UBQ5	CTCCTTCTTTCTGGTAAACGT	Real-time PCR/RT-PCR
RDR2F1	GTTTGAGTTCTTGGCTTTTT	
RDR2R1	TAATCAGCACCATCAGTTGT	Real-time PCR
DCL3F1	CGAGTTTCTCTCTTTTGCTT	
DCL3R1	CTCTTTGTTCCCTGTGACTT	Real-time PCR
AGO4F1	ATTTCTCTGCTCGCTGTAAT	
AGO4R1	TCACTGTTCTTTTTGTCAGG	Real-time PCR
NRPD1F1	GAGAGACAAAAATCCGCTAC	
NRPD1R1	GGACAAAGAAACTGCTGAAT	Real-time PCR
NRPE1F1	CGATGGAACTGTAAGAAACA	
NRPE1R1	TATTTGGAGAGGAATCAAGC	Real-time PCR
NRPD2F1	ACTTCTCTTTGACTGCTTCG	
NRPD2R1	AAAAACCCTTGATTGCTTA	Real-time PCR
KYPF1	TCGCTTGCTACTGTGGAG	
KYPR1	CCGATAGATAAGATGCCAAT	Real-time PCR
SUVH1F1	ATAAGGAAAAATGGAGGTGA	
SUVH1R1	ATAAAAGAACCAGCACGAAT	Real-time PCR
SUVH2F1	AATGGGGATGTTATGGTCTA	
SUVH2R1	TGGTTCTTTGCTGTGACTAA	Real-time PCR
SUVH3F1	GGGTAGAGAAAGGGAAATCT	
SUVH3R1	TCCTCATCAACATCATTCAC	Real-time PCR
SUVH5F1	CACAGAAGGGAAAGAGACAT	
SUVH5R1	GGCTCCATCGTAATAAGGTA	Real-time PCR
SUVH6F1	TCTTCTTCTCCACCATTTTT	
SUVH6R1	AACTGACACTCCTCGTTTCT	Real-time PCR
SUVH8F1	TATCTGGTCTCTGATTGCTG	
SUVH8R1	GTAAACCCTCTTCTCCAAAA	Real-time PCR
SUVH9F1	GAACCGAATCCTTACTCAAC	

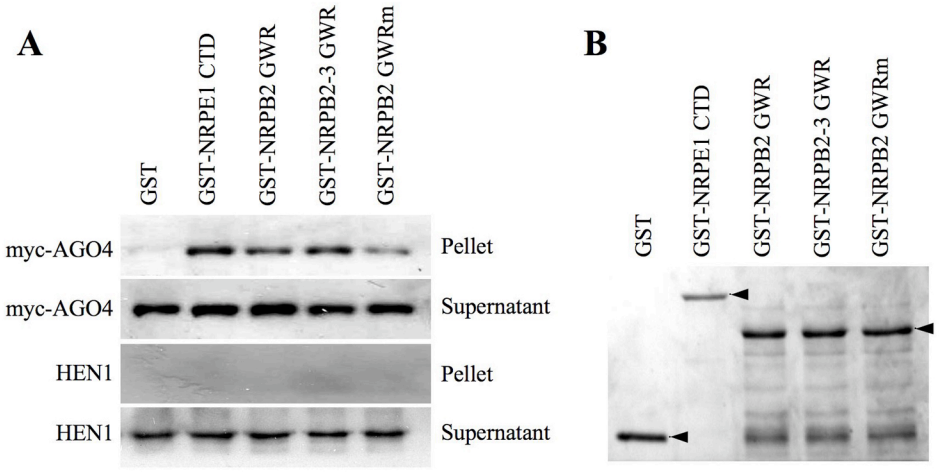
SUVH9R1	TTCTTCCATCATCAAAAACA	Real-time PCR
CMT3F	CACTTCTTACTCTGCCTGAA	
CMT3R	CCACAAATCCATAACATCTG	Real-time PCR
MET1F	AGGGAACATAAATCACAAGC	
MET1R	CAAAAGTATCAGCACCAATG	Real-time PCR
DDM1F	AGAGGAAGAGGACATACTGG	
DDM1R	TGTTCAGGGAAGACAGCA	Real-time PCR
DRM2F1	CAAATACTCAAGCGACTGTG	
DRM2R1	TGCCATCTTCTTCCTCTACT	
DRM1F1	CAGACGAATCAAAAGGGTAT	
DRM1R1	GAAGAAGAGTGACGAGTGCT	
ROS1F1	GTTGAAGAAAGACGAGAAGG	
ROS1R1	TATTATTTTGCTGTGGCTTG	
TR2558F	AAACCATCTTCCTCCTCATA	
TR2558R	AGCAAAAAGTTCACACGAA	ChIP/Realtime-PCR
cluster4F	CGTCCTCAAAGTTCCAGA	
cluster4R	GGTATTCTCCATCCCAAAG	ChIP/Realtime-PCR
cluster2F2	TTGCTGATTTGTATTTTATGCAT	
cluster2R2	CTTTTTCAAACCATAAACCAGAAA	ChIP/Realtime-PCR/Probe
soloLTR F1	AACTAACGTCATTACATACACATCTTG	
soloLTR R1	AATTAGGATCTTGTTTGCCAGCTA	Real-time PCR
soloLTR F2	AATGCATTACAAAAACCTTCTGA	
soloLTR R2	GGATTCACGATTAGAGAACGTAGA	ChIP-PCR/probe
siR02F2	CAATATGTTCTTCACCATCG	
siR02R2	ATTTGCGAAACTAATGGAAG	ChIP/Realtime-PCR
AtSN1F1	ACCAACGTGCTGTTGGCCCAGTGGTAAATC	
AtSN1R1	AAAATAAGTGGTGGTTGTACAAGC	ChIP/Realtime-PCR
AtCopia2F1	TTGCCCCAACAACAAAAA	
AtCopia2R1	CAGAGAAAGAGATAGAAGAAATGA	ChIP/Realtime-PCR
AtREP2 F1	TTTTTATGGGACGGGTTG	
AtREP2 R1	ATTCTATTTTGGCGTGCTC	ChIP/Realtime-PCR
SimpleHAT2 F1	ATTTTCCTCCATTCACACTT	
SimpleHAT2 R1	TAACAACTCAACCCATCAAA	ChIP/Realtime-PCR
eIF4A1_F	TCTTGGTGAAGCGTGATGAG	
eIF4A1_R	GCTGAGTTGGGAGATCGAAG	ChIP/Realtime-PCR
LNA-siR02	G+TTG+ACC+AGT+CCG+CCA+GCC+GAT	Probe
LNA-siR1003	A+TGC+CAA+GTT+TGG+CCT+CAC+CGT+C	Probe
LNA-AtSN1	ACCA+ACG+TGTTG+TTGG+CCCAGTG+GTAAA+TCTCTCA+GAT	Probe
LNA-Cluster4	AA+GATC+AAAC+ATCA+GCA+GCGTC+AG+AGG+CTT	Probe
LNA-TR2558	AA+GCTA+TCGG+TCATG+CTGATG+AATATG+AGG+AGG+AA	Probe
LNA-Cluster4	AA+GATC+AAAC+ATCA+GCA+GCGTC+AG+AGG+CTT	Probe
LNA-TR2558	AA+GCTA+TCGG+TCATG+CTGATG+AATATG+AGG+AGG+AA	Probe
LNA-AtCopia2	TC+ATCTA+GAT+GTT+TCTA+GGGT+TTT+ACCC+TTTT+CCTTGT+ACTCC	Probe
LNA-AtREP2	G+CGGG+ACGG+GTTT+GGCA+GGAC+GTTA+CTT+AAT	Probe
LNA-SimpleHAT2	T+GGG+TTA+CCC+ATTT+TGAC+ACCC+CTA	Probe
siR02F4	TGAAAACCACTCTACGTA	
siR02R4	CGATGGTGAAGAACATAT	RT-PCR for transcript B
IGN5_B F	TCCCGAGAAGAGTAGAACAAATGCTAAAA	
IGN5_B R	CTGAGGTATTCCATAGCCCCTGATCC	RT-PCR for transcript B
IGN5_A F	CGCAGCGGAATTGACATCCTATC	
IGN5_A R	TCGGAAAGAGACTCTCCGCTAGAAA	RT-PCR for transcript A
AtSN1_B F	TGAGAGATTTACCACTGGGCCAACA	
AtSN1_B R	TGAGGAGCTCAACACATAAATGGCAATA	RT-PCR for transcript B
soloLTR A A221	ATCAATTATTATGTCATGTTAAAACCGATTG	
soloLTR A A222	ATCAATTATTATGTCATGTTAAAACCGATTG	RT-PCR for transcript A

soloLTR B A217	CATATAACCGAAGCCGAAGGATGTGAAA	
soloLTR B A218	CAGAAACCTAAGGAACCATTACACGCTAAACC	RT-PCR for transcript B
siRNA02-BisF	ATYATATTTAGATAGATAGGAGAAAG	
siRNA02-BisR	AARAATATRCTAATTTTTCTATCATA	Bilsulfite Sequencing
NRPB2F1	CGCGGATCCCGAGTACACTACTATTACACCAATTCATTAACTTGG	
NRPB2R1	ACGCGTCGACCGGTTCCGGCGAAGAAGATAGTCAAAGATTGAG	Complemention test
NRPB2F4	CGCGGATCCGCTGGTCCTTTACTTGGA	
NRPB2R4	CGCGGATCCGCGGTTTTCAGTATAAGCCTC	Protein expression
NRPB2F6	CTTGCTACTGGGAACGCTGCGCAGGCAAACGCTGC	
NRPB2R6	GCAGCGTTTGCCTGCGCAGCGTTCCCAGTAGCAAG	Site mutageneisis
NRPB2F7	CTGCACAACTCACAGGTCGCGATGATGTGCCCTGCT	
NRPB2R7	AGCAGGGCACATCATCGCGACCTGTGAGTTGTGCAG	Site mutageneisis
NRPB2F8	GAATTTTTGGAAGAAGTGGCCACTGAGAATTTTGAG	
NRPB2R8	CTCAAAATTCTCAGTGGCCACTTCTTCCAAAAATTC	Site mutageneisis
NRPB2F9	GCAGAAGAAGATGCTGCGCATCATCTGGTTGC	
NRPB2R9	GCAACCAGATGATGCGCAGCATCTTCTTCTGC	Site mutageneisis


A
Genes involved in siRNA biogenesis
Col-0
nrpb2-3
Relative transcript levels
2
1.5
1
0.5
0
RDR2
DCL3
AGO4
NRPD1
NRPE1
NRPD2
B
Genes involved in DNA methylation/demethylation
Col-0
nrpb2-3
Relative transcript levels
2
1.5
1
0.5
0
DDM1
CMT3
MET1
DRM1
DRM2
ROS1
C
Genes involved in histone H3K9 methylation
Col-0
nrpb2-3
Relative transcript levels
2
1.5
1
0.5
0
KYP
SUVH1
SUVH2
SUVH3
SUVH5
SUVH6
SUVH8
SUVH9


A
siR02
% Cytosine methylation
100
80
60
40
20
0
CG
CHG
CHH
Col-0
nrpb2-3
nrpe1-1
B
H3K9me2 levels
ChIP signal [% Input]
4
3
2
1
0
TR2558
siR02
Cluster2
Cluster4
soloLTR
Col-0
nrpb2-3
nrpe1-1
nrpb2-3 nrpe1-1


A
Ler
myc-AGO4 WT
myc-AGO4 nrpb2-3
myc-AGO4
Hsp70
B
Col-0
NRPE1-FLAG
NRPE1-FLAG nrpb2-3
NRPE1-FLAG
Hsp70
C
Col-0
NRPD1-FLAG
NRPD1-FLAG nrpb2-3
NRPD1-FLAG
Hsp70


A
GST
GST-NRPE1 CTD
GST-NRPB2 GWR
GST-NRPB2-3 GWR
GST-NRPB2 GWRm
myc-AGO4
Pellet
myc-AGO4
Supernatant
HEN1
Pellet
HEN1
Supernatant
B
GST
GST-NRPE1 CTD
GST-NRPB2 GWR
GST-NRPB2-3 GWR
GST-NRPB2 GWRm

Supplemental Table S1. Genes down-regulated in *nrpb2-3*

Affimetrix code	relative levels (wt/*nrpb2-3*)	AGI code	description
244943_at	2.506031482	ATMG00070	NADH dehydrogenase subunit 9
245036_at	6.440138681	AT2G26410	IQD4,calmodulin binding
245053_at	4.969283934	AT2G26450	pectinesterase family protein
245057_at	4.134615536	AT2G26490	transducin family protein / WD-40 repeat family protein
245151_at	2.694601064	AT2G47550	pectinesterase family protein
245159_at	3.850129991	AT2G33100	cellulose synthase
245232_at	4.769024425	AT4G25590	ADF7 (ACTIN DEPOLYMERIZING FACTOR 7); actin binding
245246_at	3.687802749	AT1G44224	Encodes a ECA1 gametogenesis related family protein
245247_at	2.633255806	AT4G17230	SCL13 (SCARECROW-LIKE 13); transcription factor
245322_at	4.072624958	AT4G14815	lipid transfer protein (LTP) family protein
245335_at	2.286720596	AT4G16160	protein transmembrane transporter
245376_at	2.484100833	AT4G17690	peroxidase
245423_at	2.535276737	AT4G17483	palmitoyl protein thioesterase family protein
245468_at	3.095029263	AT4G15980	pectinesterase family protein
245477_at	4.269311281	AT4G16110	ARR2 (ARABIDOPSIS RESPONSE REGULATOR 2); transcription factor
245577_at	2.28553368	AT4G14780	protein kinase
245622_at	3.601645951	AT4G14080	hydrolase
245636_at	3.750345719	AT1G25240	epsin N-terminal homology (ENTH) domain-containing protein
245658_at	5.577215112	AT1G28270	RALFL4 (RALF-LIKE 4)
245674_at	3.110233545	AT1G56680	glycoside hydrolase family 19 protein
245700_at	4.759829529	AT5G04180	carbonic anhydrase family protein
245737_at	3.562272282	AT1G44160	DNAJ chaperone C-terminal domain-containing protein
245901_at	2.731872753	AT5G11060	KNAT4 (KNOTTED1-LIKE HOMEOBOX GENE 4); transcription factor
245946_at	4.113024551	AT5G19580	glyoxal oxidase
245970_at	2.011637777	AT5G20710	BGAL7,beta-galactosidase
246065_at	3.859779967	AT5G19360	CPK34,calmodulin-dependent protein kinase
246075_at	2.577431311	AT5G20410	MGD2,monogalactosyldiacylglycerol synthase 2
246106_at	4.864259242	AT5G28680	protein kinase family protein
246208_at	4.464667055	AT4G36490	SEC14 cytosolic factor
246392_at	4.88398451	AT1G58120	similar to hypothetical protein
246392_at	4.88398451	AT1G58122	CPuORF45,Conserved peptide upstream open reading frame 45
246416_at	5.001546638	AT5G16920	similar to unknown protein
246418_at	3.371452369	AT5G16960	NADP-dependent oxidoreductase
246431_at	4.090174863	AT5G17480	calcium-binding pollen allergen
246491_at	2.485054432	AT5G16100	similar to hypothetical protein
246540_at	2.541510447	AT5G15600	SP1L4 (SPIRAL1-LIKE4)
246545_at	5.848569572	AT5G15110	pectate lyase family protein

246582_at	2.737748941	AT1G31750	proline-rich family protein
246592_at	2.893627279	AT5G14890	NHL repeat-containing protein
246616_at	3.429583178	AT5G36260	aspartyl protease family protein
246646_at	2.843998128	AT5G35090	unknown protein
246751_at	5.726695011	AT5G27870	pectinesterase family protein
246761_at	5.223402504	AT5G27980	seed maturation family protein
246841_at	5.680868461	AT5G26700	germin-like protein
246878_at	6.055691228	AT5G26060	S1 self-incompatibility protein
247055_at	3.379438485	AT5G66740	similar to unknown protein
247136_at	2.483648253	AT5G66170	similar to unknown protein
247178_at	2.041163357	AT5G65205	short-chain dehydrogenase
247253_at	3.602661012	AT5G64790	glycosyl hydrolase family 17 protein
247293_at	3.101783441	AT5G64510	similar to unnamed protein product
247402_at	5.854774082	AT5G62750	unknown protein
247424_at	4.892125974	AT5G62850	ATVEX1 (VEGETATIVE CELL EXPRESSED1)
247462_at	5.650194881	AT5G62080	lipid transfer protein (LTP) family protein
247512_at	4.128388356	AT5G61720	similar to unknown protein
247519_at	2.132191926	AT5G61430	Arabidopsis NAC domain containing protein 100,transcription factor
247539_at	3.663707568	AT5G61710	similar to unknown protein
247546_at	3.184102813	AT5G61605	similar to unknown protein
247639_at	4.428269197	AT5G60500	undecaprenyl pyrophosphate synthetase family protein
247639_at	4.428269197	AT5G60510	undecaprenyl pyrophosphate synthetase family protein
247657_at	2.049808395	AT5G59845	gibberellin-regulated family protein
247736_at	3.990977592	AT5G59370	ACT4 (ACTIN 4)
247804_at	5.577768608	AT5G58170	glycerophosphoryl diester phosphodiesterase family protein
247843_at	3.273373896	AT5G58050	glycerophosphoryl diester phosphodiesterase family protein
247965_at	2.448642935	AT5G56540	AGP14 (ARABINOGALACTAN PROTEIN 14)
247981_at	3.933306787	AT5G56640	MIOX5 (MYO-INOSITOL OXYGENASE 5)
248194_at	5.706235421	AT5G54095	similar to unknown protein
248227_at	4.208202368	AT5G53820	similar to unknown protein
248245_at	2.223555699	AT5G53190	nodulin MtN3 family protein
248367_at	4.798958925	AT5G52360	actin-depolymerizing factor
248470_at	6.09734021	AT5G50830	unknown protein
248484_at	3.176050378	AT5G51030	short-chain dehydrogenase
248534_at	4.255051227	AT5G50030	methylesterase inhibitor family protein
248638_at	2.695170886	AT5G49070	CoA synthase family protein
248714_at	6.97485482	AT5G48140	polygalacturonase
248716_at	4.508842601	AT5G48210	similar to unknown protein
248761_at	2.120256289	AT5G47635	similar to unknown protein
248804_at	3.101649183	AT5G47470	nodulin MtN21 family protein
248822_at	4.881097252	AT5G47000	peroxidase
248840_at	5.47602943	AT5G46770	unknown protein

248916_at	3.149196207	AT5G45840	leucine-rich repeat transmembrane protein kinase
248926_at	3.557810572	AT5G45880	pollen allergen and extensin family protein
249046_at	2.756190066	AT5G44400	FAD-binding domain-containing protein
249048_at	4.090863197	AT5G44300	dormancy/auxin associated family protein
249194_at	4.411554537	AT5G42490	kinesin motor family protein
249226_at	2.81769506	AT5G42170	carboxylesterase
249401_at	2.884268856	AT5G40260	nodulin MtN3 family protein
249402_at	4.317762024	AT5G40155	Encodes a defensin-like (DEFL) family protein.
249429_at	5.055128907	AT5G39880	similar to unknown protein
249447_at	5.217912553	AT5G39400	pollen specific phosphatase
249503_at	2.842258409	AT5G39310	ATEXPA24 (ARABIDOPSIS THALIANA EXPANSIN A24)
249536_at	4.575376856	AT5G38760	similar to unknown protein
249584_at	2.535090188	AT5G37810	water channel
249584_at	2.535090188	AT5G37820	water channel
249852_at	5.764152659	AT5G23270	STP11 (SUGAR TRANSPORTER 11)
249950_at	5.384108054	AT5G18910	protein kinase family protein
250091_at	3.249193039	AT5G17340	similar to unknown protein
250121_at	7.335629088	AT5G16500	protein kinase family protein
250154_at	4.993920537	AT5G15140	aldose 1-epimerase family protein
250174_at	4.997059819	AT5G14380	AGP6 (ARABINOGALACTAN PROTEINS 6)
250204_at	3.060191976	AT5G13990	ATEXO70C2 (exocyst subunit EXO70 family protein C2)
250294_at	3.757329706	AT5G13380	auxin-responsive GH3 family protein
250308_at	6.283955508	AT5G12180	CPK17 (calcium-dependent protein kinase 17)
250349_at	4.400121471	AT5G12000	kinase
250514_at	5.163749115	AT5G09550	RAB GDP-dissociation inhibitor
250606_at	4.787188597	AT1G69940	pectinesterase
250606_at	4.787188597	AT5G07410	pectinesterase family protein
250608_at	4.173690574	AT5G07420	pectinesterase family protein
250610_at	3.007693241	AT5G07550	GRP19 (Glycine rich protein 19)
250619_at	4.65812873	AT5G07230	lipid transfer protein (LTP) family protein
250631_at	4.123848976	AT5G07430	pectinesterase family protein
250635_at	4.460315414	AT5G07510	GRP14 (Glycine rich protein 14)
250636_at	3.970741014	AT5G07520	GRP18 (Glycine rich protein 18)
250637_at	4.387049084	AT5G07530	GRP17 (Glycine rich protein 17)
250638_at	4.584627383	AT5G07540	GRP16 (Glycine rich protein 16)
250639_at	5.483529282	AT5G07560	GRP20 (Glycine rich protein 20)
250917_at	2.438755208	AT5G03690	fructose-bisphosphate aldolase
251180_at	4.497248	AT3G62640	similar to unknown protein
251228_at	5.183672365	AT3G62710	glycosyl hydrolase family 3 protein
251250_at	4.960585768	AT3G62180	methylesterase inhibitor family protein
251252_at	5.93451167	AT3G62230	F-box family protein
251258_at	4.209537152	AT3G62170	VGDH2 (VANGUARD 1 HOMOLOG 2); pectinesterase

251358_at	4.489204494	AT3G61160	shaggy-related protein kinase beta / ASK-beta (ASK2)
251361_at	3.861758066	AT3G61230	LIM domain-containing protein
251397_at	3.172656529	AT3G60570	ATEXPB5 (ARABIDOPSIS THALIANA EXPANSIN B5)
251590_at	3.819861889	AT3G57690	AGP23/ATAGP23 (ARABINOGALACTAN-PROTEIN 23)
251595_at	2.59457935	AT3G57620	glyoxal oxidase
251697_at	4.635746177	AT3G56600	inositol or phosphatidylinositol kinase
251854_at	6.175746024	AT3G54800	pleckstrin homology (PH) domain-containing protein
251874_at	3.2944226	AT3G54240	hydrolase
252005_at	3.786862738	AT3G52810	ATPAP21/PAP21 (purple acid phosphatase 21)
252035_at	3.69181021	AT3G52160	CoA synthase family protein
252052_at	4.570680752	AT3G52600	hydrolase
252061_at	3.646415551	AT3G52620	unknown protein
252063_at	3.319789995	AT3G51590	LTP12 (LIPID TRANSFER PROTEIN 12)
252085_at	5.318549332	AT3G52000	SCPL36 (serine carboxypeptidase-like 36)
252085_at	5.318549332	AT3G52010	SCPL37 (serine carboxypeptidase-like 37)
252090_at	3.329387168	AT3G52130	lipid transfer protein (LTP) family protein
252095_at	2.326904308	AT3G51000	epoxide hydrolase
252140_at	2.597949353	AT3G51070	dehydration-responsive protein
252212_at	2.908602455	AT3G50310	MAPKKK20 (Mitogen-activated protein kinase kinase kinase 20)
252364_at	2.784934038	AT3G48450	nitrate-responsive NOI protein
252440_at	3.086141112	AT3G47440	water channel
252531_at	3.882155109	AT3G46520	ACT12 (ACTIN-12)
252571_at	3.610634625	AT3G45280	SYP72 (SYNTAXIN OF PLANTS 72)
252710_at	5.716027658	AT3G43860	hydrolase
252733_at	5.829213049	AT3G43120	auxin-responsive protein
252780_at	3.372069496	AT3G42960	oxidoreductase
252820_at	3.842296453	AT3G42640	ATPase
252881_at	4.661126999	AT4G39610	similar to unknown protein
252908_at	4.533504045	AT4G39670	glycolipid transporter
252933_at	5.114923614	AT4G39110	protein kinase family protein
253012_at	4.060952497	AT4G37900	glycine-rich protein
253034_at	3.066750976	AT4G38230	CPK26 (calcium-dependent protein kinase 26)
253087_at	2.013611894	AT4G36350	ATPAP25/PAP25 (purple acid phosphatase 25)
253181_at	2.765680026	AT4G35180	amino acid transmembrane transporter
253195_at	2.369234464	AT4G35420	dihydroflavonol 4-reductase family
253222_at	2.744119004	AT4G34850	chalcone and stilbene synthase family protein
253226_at	5.858091419	AT4G35010	BGAL11 (beta-galactosidase 11)
253240_at	2.053456819	AT4G34510	KCS2 (CoA synthase 2); acyltransferase
253641_at	2.461033419	AT4G29980	similar to unknown protein
253721_at	2.746127298	AT4G29250	transferase family protein
253725_at	4.692147238	AT4G29340	PRF4 (PROFILIN 4); actin binding
253772_at	3.609753943	AT4G28395	ATA7 (Arabidopsis thaliana anther 7)

253807_at	3.626788844	AT4G28280	similar to unknown protein
253831_at	5.93374424	AT4G27580	unknown protein
253833_at	3.435604953	AT4G27790	calcium-binding EF hand family protein
253846_at	2.741801335	AT4G28000	AAA-type ATPase family protein
253863_at	2.132986677	AT4G27420	ABC transporter family protein
254024_at	4.237682014	AT4G25780	pathogenesis-related protein
254096_at	2.183767021	AT4G25150	acid phosphatase
254104_at	2.034092558	AT4G25040	integral membrane family protein
254123_at	5.804598297	AT4G24640	APPB1; pectinesterase inhibitor
254465_at	3.022282543	AT4G20420	tapetum-specific protein
254468_at	2.326002453	AT4G20460	NAD-dependent epimerase/dehydratase family protein
254490_at	2.079646209	AT4G20320	CTP synthase
254494_at	3.598067816	AT4G20050	QRT3 (QUARTET 3)
254627_at	5.626744563	AT4G18395	unknown protein
254716_at	4.237195726	AT4G13560	UNE15 (unfertilized embryo sac 15)
254762_at	3.707305823	AT4G13230	late embryogenesis abundant domain-containing protein
254886_at	4.412321217	AT4G11760	LCR17 (Low-molecular-weight cysteine-rich 17)
254956_at	2.061975216	AT4G10850	nodulin MtN3 family protein
255101_at	4.905217234	AT4G08670	lipid transfer protein (LTP) family protein
255175_at	3.249563145	AT4G07960	ATCSLC12 (Cellulose synthase-like C12)
255251_at	2.754329578	AT4G04980	similar to proline-rich family protein
255276_at	5.765468953	AT4G04930	DES-1-LIKE (fatty acid desaturase 1-like)
255345_at	2.31669115	AT4G04460	aspartyl protease family protein
255423_at	6.614501709	AT4G03290	calcium-binding protein
255515_at	4.472121345	AT4G02250	methylesterase inhibitor family protein
255530_at	2.638692943	AT4G02140	similar to unknown protein
255580_at	2.575047171	AT4G01470	water channel
255717_at	4.659572451	AT4G00350	MATE efflux family protein
255806_at	2.25385925	AT4G10260	pfkB-type carbohydrate kinase family protein
255815_at	2.108780714	AT1G19890	ATMGH3/MGH3 (MALE-GAMETE-SPECIFIC HISTONE H3)
255835_at	4.396195022	AT2G33420	similar to unknown protein
255837_at	3.981421995	AT2G33460	RIC1 (ROP-INTERACTIVE CRIB MOTIF-CONTAINING PROTEIN 1)
255847_at	3.99679566	AT2G33270	thioredoxin family protein
255956_at	2.330741036	AT1G22015	DD46; transferase
256059_at	4.142903344	AT1G06990	GDSL-motif lipase/hydrolase family protein
256297_at	2.257537403	AT1G69500	oxygen binding
256307_at	5.053502321	AT1G30350	pectate lyase family protein
256381_at	3.869345514	AT1G66850	lipid transfer protein (LTP) family protein
256392_at	5.150185401	AT3G06260	GATL4 (Galacturonosyltransferase-like 4)
256481_at	3.038658882	AT1G33430	galactosyltransferase family protein
256580_at	6.167110878	AT3G28810	similar to unknown protein
256580_at	6.167110878	AT3G28820	similar to unknown protein

256581_at	5.011662767	AT3G28830	similar to unknown protein
256582_at	4.101515755	AT3G28840	similar to unknown protein
256584_at	4.996431766	AT3G28750	similar to unknown protein
256587_at	5.009652955	AT3G28780	similar to unknown protein
256588_at	5.078780174	AT3G28790	similar to unknown protein
256589_at	3.994262716	AT3G28740	cytochrome P450 family protein
256662_at	3.579040243	AT3G11980	MS2 (MALE STERILITY 2)
256860_at	2.182453413	AT3G23840	transferase family protein
256945_at	3.786794838	AT3G19020	leucine-rich repeat family protein
256955_at	4.588649057	AT3G13390	copper ion binding / oxidoreductase
256966_at	5.222324072	AT3G13400	copper ion binding / oxidoreductase
257065_at	4.27241732	AT3G18220	phosphatidic acid phosphatase family protein
257082_at	4.721808607	AT3G20580	COBL10 (COBRA-LIKE PROTEIN 10 PRECURSOR)
257090_at	4.80251291	AT3G20530	protein kinase family protein
257102_at	4.793213327	AT3G25050	hydrolase
257119_at	6.47276475	AT3G20190	leucine-rich repeat transmembrane protein kinase
257121_at	6.413702458	AT3G20220	auxin-responsive protein
257170_at	2.918570472	AT3G23770	glycosyl hydrolase family 17 protein
257256_at	3.031199302	AT3G21970	receptor-like protein kinase
257261_at	2.395954933	AT3G21920	[AT3G21920, pollen coat receptor kinase
257261_at	2.395954933	AT3G21930	[AT3G21930, receptor-like protein kinase
257392_at	5.360969506	AT2G24450	FLA3 (FASCICLIN-LIKE ARABINOGALACTAN PROTEIN 3 PRECURSOR)
257399_at	2.217151733	AT1G23690	similar to unknown protein
257402_at	4.868318866	AT1G23570	similar to unknown protein
257405_at	4.388173154	AT1G24620	calcium-binding pollen allergen
257469_at	6.907306855	AT1G49290	similar to unknown protein
257491_at	2.418788821	AT1G06170	basic helix-loop-helix (bHLH) family protein
257532_at	4.61127127	AT3G04700	similar to unknown protein
257558_at	2.663709824	AT3G22000	receptor-like protein kinase
257558_at	2.663709824	AT4G20680	receptor-like protein kinase
257633_at	3.317009872	AT3G26125	cytochrome P450
257804_at	3.104488555	AT3G18810	protein kinase family protein
257819_at	4.614174027	AT3G25165	RALFL25 (RALF-LIKE 25)
257821_at	4.879606346	AT3G25170	RALFL26 (RALF-LIKE 26)
257850_at	2.847439262	AT3G13065	ATP binding /protein serine/threonine kinase
257862_at	2.463969673	AT3G17760	GAD5 (GLUTAMATE DECARBOXYLASE 5)
257862_at	2.463969673	AT3G17720	pyridoxal-dependent decarboxylase family protein
257886_at	5.643873707	AT3G17060	pectinesterase family protein
257925_at	2.538090527	AT3G23170	similar to unknown protein
257951_at	3.21282451	AT3G21700	GTP binding
257986_at	4.775645122	AT3G20865	AGP40 (ARABINOGALACTAN-PROTEIN 40)
258005_at	3.048059222	AT3G19390	cysteine proteinase

258012_at	5.091065461	AT3G19310	phospholipase C
258035_at	3.068251863	AT3G21180	calcium-transporting ATPase
258058_at	5.394226549	AT3G28980	similar to unknown protein
258077_at	5.286308635	AT3G26110	similar to BCP1
258168_at	4.495642015	AT3G21570	unknown protein
258216_at	4.997041454	AT3G17980	C2 domain-containing protein
258278_at	7.695561318	AT3G26860	similar to unknown protein
258337_at	5.797605447	AT3G16040	similar to unknown protein
258392_at	3.596988556	AT3G15400	ATA20 (Arabidopsis thaliana anther 20)
258408_at	2.27443543	AT3G17630	monovalent cation:proton antiporter
258465_at	2.651583276	AT3G17220	methylesterase inhibitor family protein
258498_at	3.587491473	AT3G02480	ABA-responsive protein
258523_at	3.779043967	AT3G06830	pectinesterase family protein
258600_at	4.521719092	AT3G02810	protein kinase family protein
258605_at	6.284283724	AT3G02970	phosphate-responsive 1 family protein
258639_at	5.577250389	AT3G07820	polygalacturonase 3 (PGA3)
258645_at	4.02814964	AT3G07850	exopolygalacturonase
258645_at	4.02814964	AT3G14040	exopolygalacturonase
258671_at	2.495453386	AT3G08560	hydrogen-exporting ATPase
258748_at	5.255561314	AT3G05930	manganese ion binding
258753_at	4.681283451	AT3G09530	ATEXO70H3 (exocyst subunit EXO70 family protein H3)
258889_at	5.744850313	AT3G05610	pectinesterase family protein
258968_at	3.647072169	AT3G10460	self-incompatibility protein
259044_at	3.637848861	AT3G03430	calcium-binding pollen allergen
259063_at	3.714943559	AT3G07450	lipid transfer protein (LTP) family protein
259189_at	5.545135578	AT3G01700	AGP11 (ARABINOGALACTAN PROTEIN 11)
259265_at	4.561417647	AT3G01250	unknown protein
259266_at	5.077127453	AT3G01240	similar to unknown protein
259269_at	5.166114128	AT3G01270	pectate lyase family protein
259338_at	4.160381228	AT3G03800	SYP131 (syntaxin 131)
259346_at	2.274380643	AT3G03910	glutamate dehydrogenase
259425_at	3.549902223	AT1G01460	1-phosphatidylinositol-4-phosphate 5-kinase
259451_at	3.411881029	AT1G13890	SNAP30 (synaptosomal-associated protein 30)
259464_at	5.74292194	AT1G18990	similar to unknown protein
259615_at	4.594057898	AT1G47980	similar to unknown protein
259693_at	2.741461506	AT1G63060	similar to unknown protein
259719_at	3.489892601	AT1G61070	LCR66/PDF2.4 (Low-molecular-weight cysteine-rich 66)
259744_at	3.800494931	AT1G71160	beta-ketoacyl-CoA synthase family protein
259820_at	4.752709307	AT1G66210	subtilase family protein
259836_at	4.354897823	AT1G52240	ATROPGEF11/ROPGEF11 (KINASE PARTNER PROTEIN-LIKE)
259886_at	5.088365216	AT1G76370	protein kinase
259975_at	3.370034946	AT1G76470	cinnamoyl-CoA reductase

259987_at	2.482880254	AT1G75030	ATLP-3 (Arabidopsis thaumatin-like protein 3)
260001_at	5.221714569	AT1G67990	caffeoyl-CoA 3-O-methyltransferase
260011_at	4.014377331	AT1G68110	clathrin assembly protein
260022_at	5.085212986	AT1G30020	similar to unknown protein
260038_at	3.73759654	AT1G68875	unknown protein
260161_at	3.42154256	AT1G79860	Rho guanyl-nucleotide exchange factor
260228_at	2.999423607	AT1G74540	cytochrome P450
260233_at	3.523537255	AT1G74550	cytochrome P450
260306_at	4.107671072	AT1G70540	EDA24 (embryo sac development arrest 24)
260386_at	2.875459212	AT1G74010	strictosidine synthase family protein
260527_at	2.854227819	AT2G47270	transcription factor
260577_at	3.068230856	AT2G47340	methylesterase inhibitor family protein
260661_at	5.46811745	AT1G19500	similar to unknown protein
260737_at	3.747617778	AT1G17540	kinase
260788_at	4.23300655	AT1G06260	cysteine proteinase
260791_at	3.047169829	AT1G06250	lipase class 3 family protein
260888_at	4.273888042	AT1G29140	pollen allergen and extensin family protein
261015_at	4.334104209	AT1G26480	protein phosphorylated amino acid binding
261045_at	5.742444691	AT1G01310	allergen related family protein
261051_at	3.04259041	AT1G01280	oxidoreductase
261096_at	2.709155798	AT1G62940	CoA ligase family protein
261222_at	2.610420882	AT1G20120	family II extracellular lipase
261244_at	2.813209263	AT1G20150	subtilase family protein
261245_at	2.176366809	AT1G20135	hydrolase
261264_at	3.332477249	AT1G26710	similar to unknown protein
261392_at	2.037257026	AT1G79780	similar to integral membrane family protein
261442_at	4.658724212	AT1G28375	unknown protein
261499_at	4.011249561	AT1G28430	cytochrome P450
261528_at	5.964914852	AT1G14420	lyase
261532_at	2.545786843	AT1G71680	amino acid transmembrane transporter
261559_at	3.49965321	AT1G01780	LIM domain-containing protein
261623_at	4.397078557	AT1G01980	electron carrier
261673_at	3.430624699	AT1G18280	lipid transfer protein (LTP) family protein
261687_at	5.134592263	AT1G47280	unknown protein
261712_at	2.305224167	AT1G32780	alcohol dehydrogenase
261875_at	3.740106487	AT1G50610	leucine-rich repeat transmembrane protein kinase
261943_at	6.210559554	AT1G80660	hydrogen-exporting ATPase
262022_at	4.637855361	AT1G35490	bZIP family transcription factor
262115_at	4.613686872	AT1G02813	similar to unknown protein
262122_at	4.866212235	AT1G02790	polygalacturonase
262146_at	3.581089283	AT1G52580	rhomboid family protein
262147_at	2.002376927	AT1G52570	phospholipase D

262156_at	3.510115017	AT1G52680	late embryogenesis abundant protein-related
262282_at	3.691151744	AT1G68610	similar to unknown protein
262350_at	2.217849667	AT2G48150	glutathione peroxidase
262393_at	3.794066288	AT1G49490	leucine-rich repeat family protein
262463_at	3.270621967	AT1G50310	monosaccharide transporter (STP9)
262585_at	2.086099813	AT1G15460	anion exchanger
262664_at	3.948928273	AT1G13970	similar to unknown protein
262674_at	4.250401092	AT1G75910	EXL4 (extracellular lipase 4)
262675_at	4.502481282	AT1G75930	EXL6 (extracellular lipase 6)
262683_at	4.003067175	AT1G75920	family II extracellular lipase 5 (EXL5)
262696_at	3.327630059	AT1G75870	unknown protein
262697_at	4.710431396	AT1G75940	hydrolase
262723_at	2.42640508	AT1G43630	similar to unknown protein
262742_at	4.032869667	AT1G28550	AtRABA1i (Arabidopsis Rab GTPase homolog A1i)
262760_at	5.415843677	AT1G10770	methylesterase inhibitor family protein
262764_at	4.018722641	AT1G13140	cytochrome P450
262765_at	3.662080389	AT1G13150	cytochrome P450
262817_at	5.74658319	AT1G11770	electron carrier
262949_at	3.572731392	AT1G75790	copper ion binding
262987_at	6.053451484	AT1G23240	caleosin-related family protein
263011_at	2.211463559	AT1G23250	caleosin-related
263043_at	4.06290526	AT1G23350	methylesterase inhibitor family protein
263062_at	4.034950451	AT2G18180	SEC14 cytosolic factor
263084_at	4.597967016	AT2G27180	unknown protein
263126_at	3.394617729	AT1G78460	SOUL heme-binding family protein
263144_at	6.374590672	AT1G54070	dormancy/auxin associated protein
263166_at	7.117771499	AT1G03050	clathrin assembly protein
263215_at	5.014782854	AT1G30710	FAD-binding domain-containing protein
263363_at	3.519548084	AT2G03850	late embryogenesis abundant domain-containing protein
263450_at	5.866748093	AT2G31500	CPK24 (calcium-dependent protein kinase 24)
263453_at	3.437709863	AT2G22180	hydroxyproline-rich glycoprotein family protein
263718_at	2.351715122	AT2G13570	CCAAT-box binding transcription factor
263753_at	5.36541933	AT2G21490	LEA (DEHYDRIN LEA)
263781_at	2.982241511	AT2G46360	unknown protein
263950_at	4.433147994	AT2G36020	HVA22J (HVA22-LIKE PROTEIN J)
264039_at	3.90846337	AT2G03740	late embryogenesis abundant domain-containing protein
264112_at	2.023096082	AT2G13680	CALS5 (CALLOSE SYNTHASE 5)
264175_at	2.855852712	AT1G02050	chalcone and stilbene synthase family protein
264187_at	3.975956805	AT1G54860	similar to unknown protein
264208_at	2.591868905	AT1G22760	PAB3 (POLY(A) BINDING PROTEIN 3)
264284_at	4.719334367	AT1G61860	protein kinase
264365_at	4.656972776	AT1G03220	extracellular dermal glycoprotein

264365_at	4.656972776	AT1G03230	extracellular dermal glycoprotein
264395_at	3.471541238	AT1G12070	Rho GDP-dissociation inhibitor family protein
264529_at	3.046331769	AT1G30820	CTP synthase
264601_at	4.008430354	AT1G04540	C2 domain-containing protein
264603_at	4.481145887	AT1G04670	similar to pollen-preferential protein
264760_at	2.572275944	AT1G61290	SYP124 (syntaxin 124)
264813_at	3.409725262	AT2G17890	CPK16 (calcium-dependent protein kinase 16)
264882_at	2.809563888	AT1G61110	Arabidopsis NAC domain containing protein 25,transcription factor
264923_at	2.410235995	AT1G65970	TPX2 (THIOREDOXIN-DEPENDENT PEROXIDASE 2)
264923_at	2.410235995	AT1G60740	peroxiredoxin type 2
264993_at	4.32820709	AT1G67290	glyoxal oxidase
265002_at	2.364932516	AT1G24400	amino acid transmembrane transporter
265007_at	3.596437664	AT1G61563	RALFL8 (RALF-LIKE 8)
265007_at	3.596437664	AT1G61566	RALFL9 (RALF-LIKE 9)
265022_at	4.556562158	AT1G24520	BCP1 (Brassica campestris pollen protein 1)
265024_at	2.9004983	AT1G24600	similar to unknown protein
265080_at	4.675953461	AT1G55570	copper ion binding
265127_at	5.880563885	AT1G55560	copper ion binding
265133_at	7.02886549	AT1G51240	similar to unknown protein
265133_at	7.02886549	AT1G51250	similar to unknown protein
265163_at	2.424762127	AT1G23580	similar to unknown protein
265164_at	2.348366982	AT1G23600	similar to unknown protein
265176_at	3.660007009	AT1G23520	similar to unknown protein
265180_at	2.002371285	AT1G23590	similar to unknown protein
265185_at	3.352047168	AT1G23670	similar to unknown protein
265186_at	3.179901168	AT1G23560	similar to unknown protein
265188_at	2.223788821	AT1G23800	ALDH2B7 (Aldehyde dehydrogenase 2B7)
265280_at	5.256336859	AT2G28355	LCR5 (Low-molecular-weight cysteine-rich 5)
265306_at	2.341741089	AT2G20320	DENN (AEX-3) domain-containing protein
265331_at	3.746459829	AT2G18420	Encodes a Gibberellin-regulated family protein
265368_at	4.148336542	AT2G13350	C2 domain-containing protein
265401_at	2.779138556	AT2G10970	methylesterase inhibitor family protein
265404_at	5.709898416	AT2G16730	BGAL13 (beta-galactosidase 13)
265405_at	6.245634938	AT2G16750	protein kinase family protein
265473_at	2.810818092	AT2G15535	LCR10 (Low-molecular-weight cysteine-rich 10)
265552_at	6.628058773	AT2G07560	ATPase
265672_at	3.133822227	AT2G31980	cysteine proteinase inhibitor
265681_at	2.874509981	AT2G24370	kinase
265903_at	3.198726498	AT2G25600	potassium channel
265923_at	4.147298793	AT2G18470	protein kinase family protein
266029_at	5.450616263	AT2G05850	serine carboxypeptidase
266115_at	4.778210469	AT2G02140	protease inhibitor

266143_at	2.587453185	AT2G38905	hydrophobic protein
266160_at	4.125646027	AT2G28180	monovalent cation:proton antiporter
266296_at	2.162150219	AT2G29420	glutathione transferase
266353_at	2.404730654	AT2G01520	MLP328 (MLP-LIKE PROTEIN 328)
266364_at	2.596017421	AT2G41230	unknown protein
266453_at	3.974112452	AT2G43230	serine/threonine protein kinase
266494_at	6.223971453	AT2G07040	serine/threonine kinase
266530_at	3.327467109	AT2G16910	DNA binding / transcription factor
266558_at	6.009292097	AT2G23900	glycoside hydrolase
266592_at	3.93603181	AT2G46210	delta-8 sphingolipid desaturase
266607_at	2.430033238	AT2G46300	similar to unknown protein
266654_at	3.105090799	AT2G25890	glycine-rich protein
266665_at	5.183602226	AT2G29790	Encodes a Maternally expressed gene (MEG) family protein
266697_at	5.216768739	AT2G19770	actin monomer binding
266743_at	3.305153886	AT2G02990	endoribonuclease
266750_at	4.060525688	AT2G47040	pectinesterase
266750_at	4.060525688	AT2G47030	pectinesterase
266764_at	4.676480679	AT2G47050	methylesterase inhibitor family protein
266765_at	3.873924728	AT2G46860	inorganic diphosphatase
266850_at	5.332628589	AT2G26850	F-box family protein
266918_at	6.047031769	AT2G45800	LIM domain-containing protein
267081_at	2.578052964	AT2G41210	phosphatidylinositol-4-phosphate 5-kinase family protein
267087_at	3.101654109	AT2G32460	DNA binding / transcription factor
267295_at	5.214374565	AT2G23800	farnesyltranstransferase
267322_at	3.070611959	AT2G19330	leucine-rich repeat family protein
267398_at	4.615985778	AT2G44560	hydrolase
267440_at	3.840725643	AT2G19070	transferase family protein
267443_at	4.105719365	AT2G19000	similar to glycine-rich protein
267449_at	2.387099815	AT2G33690	late embryogenesis abundant protein
267466_at	2.416892019	AT2G19010	GDSL-motif lipase/hydrolase family protein
267476_at	5.725514603	AT2G02720	pectate lyase family protein
267531_at	4.984404765	AT2G41860	CPK14 (calcium-dependent protein kinase 14)
267643_at	2.778229698	AT2G32890	RALFL17 (RALF-LIKE 17)

Supplemental Table S2. Genes up-regulated in *nrpb2-3*

Affimetrix code	relative levels (*nrpb2-3*/wt)	AGI code	description
245275_at	2.962399133	AT4G15210	beta-amylase
245300_at	3.130166545	AT4G16350	CBL6 (CALCINEURIN B-LIKE PROTEIN 6)
246145_at	8.236474521	AT5G19880	peroxidase
246256_at	2.520244392	AT4G36770	glucosyl transferase family protein
246884_at	2.143866746	AT5G26220	ChaC-like family protein
246988_at	2.33172796	AT5G67340	armadillo repeat family protein
247252_at	2.467962442	AT5G64770	similar to unknown protein
247358_at	2.338453904	AT5G63580	flavonol synthase
247377_at	3.097408181	AT5G63180	pectate lyase family protein
248218_at	2.015873447	AT5G53710	unknown protein
248263_at	2.259532425	AT5G53370	pectinesterase
248708_at	2.196224724	AT5G48560	basic helix-loop-helix (bHLH) family protein
249227_at	9.284784163	AT5G42180	peroxidase 64
249481_at	2.822574107	AT5G38900	DSBA oxidoreductase family protein
249559_at	2.838773261	AT5G38320	similar to unknown protein
249856_at	3.330094051	AT5G22980	serine carboxypeptidase
250216_at	3.311552392	AT5G14090	similar to hypothetical protein
250364_at	2.392693676	AT5G11400	protein kinase
250447_at	2.188118221	AT5G10800	RNA recognition motif (RRM)-containing protein
250646_at	2.480211342	AT5G06720	peroxidase
250690_at	2.890918306	AT5G06530	ABC transporter family protein
250764_at	4.062243548	AT5G05960	lipid transfer protein (LTP) family protein
250770_at	2.167569299	AT5G05390	copper ion binding / oxidoreductase
251017_at	2.034898832	AT5G02760	protein phosphatase 2C family protein
251065_at	3.063095782	AT5G01870	lipid transfer protein
251181_at	2.131378821	AT3G62820	methylesterase inhibitor family protein
251509_at	3.3767514	AT3G59010	pectinesterase family protein
251576_at	2.185055329	AT3G58200	MATH domain-containing protein
251746_at	2.076293414	AT3G56060	glucose-oxidoreductase family protein
251952_at	2.628700106	AT3G53650	histone H2B
252133_at	2.137298719	AT3G50900	similar to unknown protein
252199_at	2.116380585	AT3G50270	transferase family protein
252269_at	2.636369932	AT3G49580	similar to unknown protein
252317_at	4.038492145	AT3G48720	transferase family protein
252330_at	2.114160109	AT3G48770	ATP binding / DNA binding
252827_at	2.150045413	AT4G39950	cytochrome P450
253112_at	2.213233058	AT4G35970	peroxidase
253285_at	2.077609256	AT4G34250	fatty acid elongase
254098_at	2.209894185	AT4G25100	iron superoxide dismutase

254232_at	2.447838214	AT4G23600	CORI3 (JASMONIC ACID RESPONSIVE 2)
254909_at	5.24152249	AT4G11210	disease resistance-responsive family protein
254914_at	5.007544542	AT4G11290	peroxidase
255302_at	2.656185028	AT4G04830	methionine sulfoxide reductase
255648_at	2.362720348	AT4G00910	similar to unknown protein
255690_at	2.345809738	AT4G00360	oxygen binding
255884_at	3.240569525	AT1G20310	similar to unknown protein
255893_at	2.048608858	AT1G17960	threonyl-tRNA synthetase
256096_at	2.742872192	AT1G13650	similar to 18S pre-ribosomal assembly protein
256459_at	2.333475532	AT1G36180	ATP binding/ligase
256577_at	2.136985758	AT3G28220	MATH domain-containing protein
256636_at	2.693044811	AT3G12000	S-locus related protein SLR1
256940_at	3.610792242	AT3G30720	unknown protein
257021_at	2.228989747	AT3G19710	catalytic/ methionine-oxo-acid transaminase
257130_at	2.157194184	AT3G20210	cysteine-type endopeptidase
257439_at	3.9568794	AT2G17000	MS ion channel domain-containing protein
257517_at	2.111107637	AT3G16330	similar to unknown protein
257641_at	2.748252379	AT3G25760	AOC1 (ALLENE OXIDE CYCLASE 1)
258121_at	2.052620688	AT3G14530	geranylgeranyl pyrophosphate synthase
258380_at	2.351677523	AT3G16650	PP1/PP2A phosphatases pleiotropic regulator 2
258414_at	2.723565435	AT3G17380	MATH domain-containing protein
258751_at	4.008477288	AT3G05890	RCI2B (RARE-COLD-INDUCIBLE 2B)
259197_at	4.211826697	AT3G03670	peroxidase
259383_at	2.219938034	AT3G16470	JR1 (Jacalin lectin family protein)
259786_at	3.355433719	AT1G29660	GDSL-motif lipase/hydrolase family protein
259787_at	2.007846499	AT1G29460	auxin-responsive protein
259925_at	2.944062301	AT1G75040	PR5 (PATHOGENESIS-RELATED GENE 5)
260070_at	2.941305127	AT1G73830	DNA binding / transcription factor
260220_at	2.310076724	AT1G74650	DNA binding / transcription factor
260226_at	2.870751894	AT1G74660	DNA binding / transcription factor
260553_at	2.570340114	AT2G41800	similar to unknown protein
260698_at	2.008878754	AT1G32300	FAD-binding domain-containing protein
260881_at	3.013945408	AT1G21550	calcium-binding protein
261224_at	2.822054122	AT1G20160	subtilase
261385_at	3.054092878	AT1G05450	lipid transfer protein (LTP)-related
261717_at	2.309421828	AT1G18400	transcription factor
262135_at	2.307501312	AT1G78080	DNA binding / transcription factor
262232_at	2.254107337	AT1G68600	similar to unknown protein
262304_at	2.323390989	AT1G70890	MLP43 (MLP-LIKE PROTEIN 43)
262637_at	2.023074975	AT1G06640	2-oxoglutarate-dependent dioxygenase
264014_at	3.378060601	AT2G21210	auxin-responsive protein
264160_at	2.363211801	AT1G65450	transferase family protein

264163_at	2.576909354	AT1G65445	transferase
264195_at	2.285256519	AT1G22690	gibberellin-responsive protein
264495_at	3.2225923	AT1G27380	RIC2 (ROP-INTERACTIVE CRIB MOTIF-CONTAINING PROTEIN 2)
264636_at	2.322089017	AT1G65490	similar to unknown protein
264734_at	2.294985766	AT1G62280	transporter
265008_at	4.170800382	AT1G61560	calmodulin binding
265121_at	2.056912447	AT1G62560	flavin-containing monooxygenase family protein
265926_at	2.441029696	AT2G18600	RUB1-conjugating enzyme
266108_at	2.682355276	AT2G37900	proton-dependent oligopeptide transport (POT) family protein
266823_at	2.138540944	AT2G44930	similar to unknown protein
266989_at	5.184831464	AT2G39330	jacalin lectin family protein
267425_at	2.085484924	AT2G34810	FAD-binding domain-containing protein
267459_at	3.41071579	AT2G33850	similar to unknown protein
267645_at	2.277984182	AT2G32860	glycosyl hydrolase family 1 protein