

# Plant polymerase IV sensitizes chromatin through histone modifications to preclude spread of silencing into protein-coding domains

Vivek Hari Sundar G,<sup>1</sup> Chenna Swetha,<sup>1</sup> Debjani Basu,<sup>1</sup> Kannan Pachamuthu,<sup>1</sup> Steffi Raju,<sup>1</sup> Tania Chakraborty,<sup>2</sup> Rebecca A. Mosher,<sup>2</sup> and P.V. Shivaprasad<sup>1</sup>

<sup>1</sup>National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bangalore 560065, India;

<sup>2</sup>School of Plant Sciences, The University of Arizona, Tucson, Arizona 85721, USA

Across eukaryotes, gene regulation is manifested via chromatin states roughly distinguished as heterochromatin and euchromatin. The establishment, maintenance, and modulation of the chromatin states is mediated using several factors including chromatin modifiers. However, factors that avoid the intrusion of silencing signals into protein-coding genes are poorly understood. Here we show that a plant specific paralog of RNA polymerase (Pol) II, named Pol IV, is involved in avoidance of facultative heterochromatic marks in protein-coding genes, in addition to its well-established functions in silencing repeats and transposons. In its absence, H3K27 trimethylation (me3) mark intruded the protein-coding genes, more profoundly in genes embedded with repeats. In a subset of genes, spurious transcriptional activity resulted in small(s) RNA production, leading to post-transcriptional gene silencing. We show that such effects are significantly pronounced in rice, a plant with a larger genome with distributed heterochromatin compared with *Arabidopsis*. Our results indicate the division of labor among plant-specific polymerases, not just in establishing effective silencing via sRNAs and DNA methylation but also in influencing chromatin boundaries.

[Supplemental material is available for this article.]

The chromatin is decorated with modifications to DNA and histones constituting epigenetic modifications (Law and Jacobsen 2010; Feng and Michaels 2015). The proportion of genomic regions coding for proteins called euchromatin reduces with an increase in genome size across organisms, and such alterations and genome expansion are mainly owing to the proliferation of repeats and transposons (Pellicer et al. 2018). Plant genomes are especially enriched with latent transposons poised for transcription located proximal to the protein-coding genes (PCGs) in the euchromatic domains (Hirsch and Springer 2017).

The activity of RNA polymerase II (Pol II) on the euchromatic domains is regulated locally by transcription factors, chaperones, and chromatin remodeling enzymes that transduces the local epigenetic status (Hahn 2004; Gibney and Nolan 2010; Schier and Taatjes 2020). The evolution of plants with increased instances of gene-proximal repeats in the genome necessitates articulation of silencing states to the Pol II with precision. This locally silenced state is called facultative heterochromatin and is enriched with Polycomb Repressive Complex (PRC)-dependent H3K27me3 marks in contrast to the constitutive heterochromatin enriched with H3K9me2 modifications (Zhang et al. 2007). Studies in early land plant genomes like that of *Marchantia polymorpha* suggest that H3K27me3 evolved as a predominant silencing mark in plants (Montgomery et al. 2020), and H3K9me2 mark took over the constitutive silencing, leading to the observed dichotomy between facultative and constitutive heterochromatin (Délérís et al. 2021). Mutants of *MET1*, a major player in constitutive hetero-

chromatin establishment, showed compensation by H3K27me3 marks at repeats and transposons (Soppe et al. 2002; Mathieu et al. 2005; Deleris et al. 2012; Rougée et al. 2021; Zhao et al. 2022), indicating that unknown players monitor heterochromatin states in specific domains to initiate compensatory marks in constitutive heterochromatic regions. Demarcation of facultative and constitutive heterochromatic boundaries is paramount in avoiding the intrusion of silencing states into neighboring PCGs, warranting the evolution of novel machineries curtailing silencing overshoot and simultaneously defending against genotoxic repeats. Although these mechanisms have been envisaged across eukaryotes, major upstream players involved in these beyond the obvious epigenetic readers, writers, and erasers are unknown.

Plants have evolved RNA silencing as an efficient mode of robust and targeted silencing of repeats and genes, at both transcriptional and post-transcriptional levels (Baulcombe 2004). Small RNAs (sRNAs), predominant effectors of plant RNA silencing, confer both specificity and amplification modality in silencing. The production of sRNAs associated with transcriptional silencing is primarily initiated by plant-specific RNA Pol II paralog RNA polymerase IV (Pol IV) and, in peculiar cases, by Pol II itself (Nuthikattu et al. 2013; Cuerda-Gil and Slotkin 2016). The Pol IV transcripts, majorly originating from the repeats and transposons, are acted upon by RNA-dependent RNA polymerase 2 (RDR2), whereas the Pol II aberrant transcripts are acted upon by RDR6, thereby converting them to double-stranded duplexes that become substrates for several Dicer-like proteins (DCLs) (Nuthikattu et al. 2013). The resultant short duplex sRNAs are picked up by specific Argonaute (AGO) effectors, specified by their length and by one of the strands

**Corresponding author:** shivaprasad@ncbs.res.in

Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.277353.122>. Freely available online through the *Genome Research* Open Access option.

© 2023 Hari Sundar G et al. This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

possessing the preferred 5'-nucleotide (Mi et al. 2008). For post-transcriptional gene silencing (PTGS), majorly undertaken in plants by cleavage of target mRNA, 21- to 22-nt size class sRNAs with 5'-uracil (U) are loaded into the AGO1, leading to slicing of the target mRNA between positions 10 and 11 in the sRNA: mRNA duplex regions (Baumberger and Baulcombe 2005). In contrast, for transcriptional silencing, Pol IV–derived repeat-associated 24-nt sRNAs are loaded into AGO4 (Zilberman et al. 2003) with a 5'-adenosine (A) preference, and this ribonucleoprotein complex binds to the complementary region of the long scaffold transcript produced by RNA polymerase V (Pol V) (Wierzbicki et al. 2008). This triggers recruitment of DNA methyl transferases like Domains Rearranged Methyltransferase 2 (DRM2) and majorly results in asymmetric DNA methylation at the target locus (Matzke and Mosher 2014). Numerous noncanonical modalities in the concerted activity of these RNA polymerase variants have also been reported (Cuerda-Gil and Slotkin 2016).

Even though RNA silencing is an effective process to accurately delimit silencing, the loci over which different polymerases operate are dependent on the chromatin states. For instance, Pol IV is recruited by the readers of H3K9me2 marks named Sawadee Homeodomain Homolog 1 and 2 (SHH1 and -2) in coordination with CLASSY-type chromatin remodelers (Law et al. 2013; Zhang et al. 2013; Zhou et al. 2018). Pol V recruitment involves DNA methylation readers like SUVH2 and SUVH9 (Liu et al. 2014), a chromatin remodeling complex called the DDR complex (Law et al. 2010), and several other proteins that are modulated by chromatin states. The activity of these polymerases is tightly linked to the epigenetic status of the locus of interest and, hence, the RNA silencing. A few reports have probed the effects of the loss of DNA methylation machineries on the chromatin states in *Arabidopsis* and observed excessive reorganization of the chromatin domains. For example, in *met1* mutants, chromatin is decondensed with a loss of clear heterochromatin and euchromatin boundaries, exemplifying the impact of the cross talk between these epigenetic states (Soppe et al. 2002; Mathieu et al. 2005; Zhong et al. 2021). Similarly, *pol iv* shows permanent loss of silencing at selected loci that are recalcitrant to rescue by complementation (CO) of Pol IV, and this phenomenon is owing to an associated loss of H3K9me2 marks (Li et al. 2020). Unrestricted to silencing, it is also evidenced that loss of Pol IV leads to transcriptional misregulation as observed by the accumulation of atypical nascent transcripts in maize (Erhard et al. 2015) and by increased Pol II activity at the 3'-ends of Pol II transcriptional units (McKinlay et al. 2018). The indirect facilitation of spurious transcription by virtue of loss of epigenetic players triggers cryptic transcription by Pol II in both genes and transposons (Le et al. 2020). The prolific presence of repeat fragments in the gene coding units like introns in rice has been shown to trigger transcriptional silencing-like features (Espinas et al. 2020). Taken together, the epigenetic landscape not only modulates the silencing of repeats directly but also encumbers the cryptic transcriptional activity.

Such counter-balancing reinforcement loops between epigenetic states must contribute to molecular and morphological phenotypes upon perturbation, especially in monocots with higher proportion of transposons. It has been conclusively documented that reproductive structures in rice, including gametes, undergo massive reprogramming in terms of sRNA production and DNA methylation in a locus-specific manner (Chenxin et al. 2020). Indeed, unlike *Arabidopsis*, loss of silencing players display exacerbated phenotypes in the monocot model rice in the cases of *drm2*, *met1b*, *pol iv*, *pol v*, and *dcl3* (Moritoh et al. 2012; Wei et al. 2014;

Yamauchi et al. 2014; Xu et al. 2020; Zheng et al. 2021). In agreement with this, grass family (*Poaceae*) members have evolved a specific neofunctionalized RNA polymerase paralog called RNA polymerase VI (Trujillo et al. 2018). These phenomena substantiate that a complex gene arrangement interleaved by repeats mandates a very robust mechanism of epigenetic silencing.

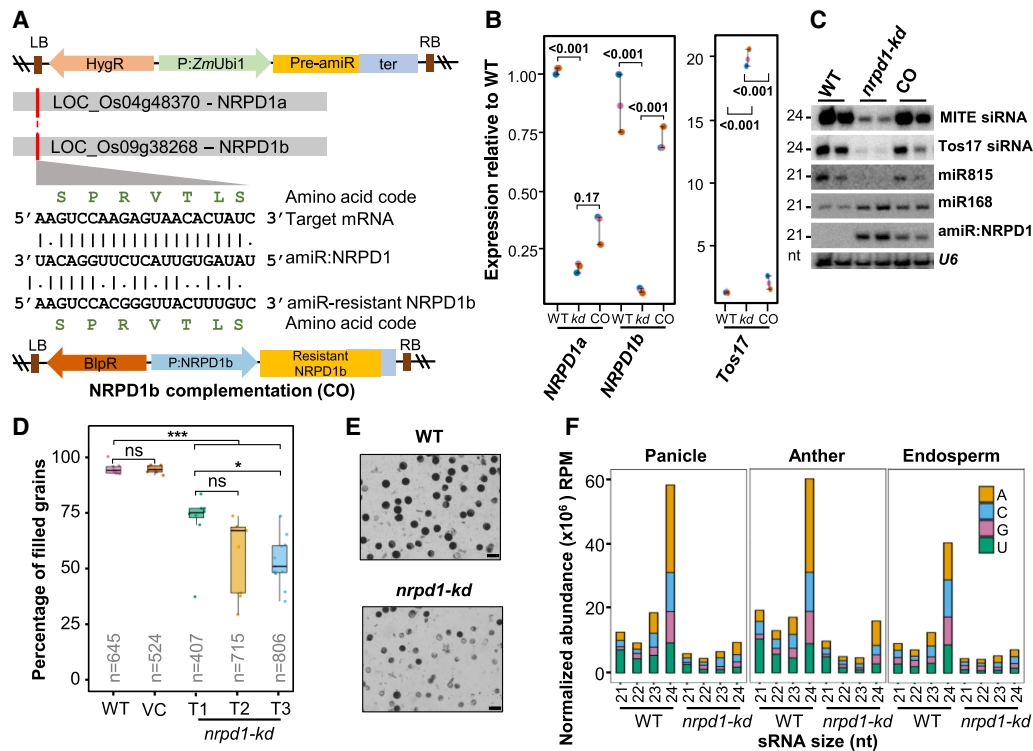
Pol IV is well known to be involved in sRNA biogenesis and silencing via DNA methylation. However, the roles of Pol IV in regulating other epigenetic layers such as chromatin marks, specifically in plants with complex genomes, are unknown. This study aims to explore the loss-of-function lines of Pol IV in rice and *Arabidopsis*, mainly focusing on the less well known roles of Pol IV in demarcating genome-wide silencing, chromatin boundaries, and aberrant transcription.

## Results

### Knockdown of RNA Pol IV induced pleiotropic phenotypic defects

The loss of function of the catalytically active component of the Pol IV complex, NRPD1, causes a spectrum of effects in different plant species, from delayed growth transition in *Physcomitrella patens*, delayed flowering in *Arabidopsis*, and reproductive defects in *Brassica rapa*, *Capsella rubella*, and *Zea mays* to increased tillering in *Japonica* rice (Erhard et al. 2009; Coruh et al. 2015; Grover et al. 2018; Xu et al. 2020; Wang et al. 2020b). To understand the multifaceted roles of Pol IV beyond RdDM in rice, we generated artificial microRNA (amiR)–mediated knockdown (*kd*) transgenic *indica* rice lines targeting the largest subunit of Pol IV (NRPD1). Since the rice genome encodes two isoforms of Pol IV, *NRPD1a* and *NRPD1b*, sharing 88.8% sequence identity, amiR was designed to target a conserved sequence in the N-terminal region of both the transcripts, incorporating optimal design parameters as identified earlier (Fig. 1A; Narjala et al. 2020). Previous attempts targeted both the rice NRPD1 isoforms using RNA interference with suboptimal targeting or the CRISPR-Cas9–mediated knockout (KO) that was embryonic lethal or was showing drastic reduction in fertility (Debladis et al. 2020; Xu et al. 2020; Zhang et al. 2020; Zheng et al. 2021). To obtain precise targeting and to avoid the lethality and sterility effects of KO, we resorted to the amiR technology. To prove that amiR is targeting only NRPD1, leading to the effects, we supertransformed the *kd* with a construct expressing a amiR-targeting-resistant version of NRPD1b CDS driven by its cognate promoter (Fig. 1A). The expression of amiR in *kd* lines was confirmed with northern hybridization, and a reduction in transcript abundance of both the isoforms was confirmed with RT-qPCR and RNA-seq. Restoration of these transcripts was observed in the CO lines (Fig. 1B,C; Supplemental Fig. S1A,B). Rice transformation was performed twice independently to score for consistent phenotypes and to eliminate T-DNA insertion effects. Independent transgenic events were identified by junction fragment Southern analyses (Supplemental Fig. S2A,B).

As observed previously (Xu et al. 2020), *kd* plants showed increased tiller number (Supplemental Fig. S1C). The percentage of viable filled grains was significantly low at 75% in T<sub>1</sub> *kd* plants and dropped progressively over generations to reach 50% in T<sub>3</sub> (Fig. 1D; Supplemental Fig. S1D). The unfilled florets showed structural deformities, and they were consistent among the distinct transgenic events, proving that the defects observed were not owing to transgenesis (Supplemental Fig. S2C–E). Specific lineages of the *kd* lines had extreme reproductive defects without viable grains



**Figure 1.** Knockdown of RNA polymerase IV (Pol IV) in rice results in pleiotropic phenotypes. (A) T-DNA map of amiR coding binary construct with the supertransformed NRPD1b complementation (CO) construct. The alignment depicts the amiR binding region and the modifications in the amiR-resistant CDS of NRPD1b that were driven by its cognate promoter (P:NRPD1b). Amino acids encoded by the original and amiR-resistant modifications were unchanged. (BlpR) Bialaphos resistance. The precursor-amiR (Pre-amiR) is driven by maize ubiquitin promoter (P:ZmUbi1). (ter) 35S-poly(A) signal, (HygR) hygromycin selection marker, (RB and LB) right and left border. (B) Plots representing relative abundance of *NRPD1a*, *NRPD1b*, and *Tos17* transcripts in *kd* and CO with respect to WT, measured by RT-qPCRs. Student's *t*-test; *P*-values mentioned across comparisons. (C) Small RNA (sRNA) northern blots showing the accumulation of 24-nt siRNAs and miRNAs in WT, *kd*, and CO. U6 was used as loading control. (D) Boxplots showing the percentage of filled grains (n = number of grains). Dots represent the average of panicles in each plant. Tukey's test; (\*\*\*) *P*-value of 0.001, (\*) *P*-value of 0.01, (ns) nonsignificant. (VC) Vector control plants that lack amiR encoding region. (E) Representative images of pollen grains of dehiscent anther stained with iodine. Scale: 50  $\mu$ m. (F) Stacked bar plots showing the normalized abundance of sRNAs of different sizes with their 5' nucleotides. The replicates were merged.

(Supplemental Fig. S3A). These defects were not owing to zygosity of the T-DNA or the dosage of amiR, indicating a possible effect of induced epimutations (Supplemental Fig. S3B,C; Johannes and Schmitz 2019). Because NRPD1 is reported to influence pollen development (Wang et al. 2020b), we further investigated the pollen quality by measuring size, iodine staining potential, and ultrastructural morphology that revealed pollen defects in *kd* (Fig. 1E; Supplemental Fig. S1E–G). Taken together, amiR-mediated *kd* of the Pol IV complex in rice led to severe reproductive defects consistent across generations, unlike the mild abnormalities observed in *Arabidopsis*.

### Rice Pol IV is required for biogenesis of repeat-associated sRNAs

To explore the molecular effects of the loss of Pol IV in rice and to discern the reasons for extreme reproductive defects upon NRPD1 *kd*, we performed sRNA profiling in three reproductive tissues that showed drastic defects: pre-emerged panicle, anther, and endosperm. As shown earlier in other species, *kd* plants showed evident reduction of 24-nt sRNAs in all the tissues (Fig. 1C,F; Supplemental Fig. S4A; Herr et al. 2005; Coruh et al. 2015; Grover et al. 2018; Wang et al. 2020b). As expected (Mi et al. 2008; Zhai et al. 2015), the maximum reduction in 24-nt sRNAs corresponded to 5'A-containing sRNAs in all the three tissues (Fig. 1F). On the contrary, bulk of the Pol II transcribed miRNAs and the sRNAs

mapping to their precursors largely remained unchanged (Supplemental Fig. S4B,C). Further, as the Pol IV–derived sRNAs are known to be associated with transposons and repeats, we profiled the sRNAs from major annotated repeats in the rice genome. We observed a substantial loss of repeat-associated sRNAs, and this was also validated by northern blotting in different tissues (Fig. 1C; Supplemental Fig. S4D–F). The repeat-associated sRNAs were also restored to WT levels in the CO lines as proven by northern blots (Fig. 1C). In conclusion, RNA Pol IV is essential for production of repeat-associated sRNAs in different tissues of rice, and CO of one of the two isoforms of NRPD1 rescues the sRNA levels.

Given the severity of the defects found in the rice *kd* plants, we hypothesized that the loss of sRNA-mediated regulation of transposons and the concomitant misregulation of genes might be attributable to the defects. Because the 24-nt sRNAs are involved in establishing DNA methylation of the repeats, we profiled the DNA methylome at the annotated repeats in rice, which revealed a substantial decrease of methylation in *kd* compared with WT, especially in CHG and CHH contexts (Supplemental Fig. S5A). In addition, several of the transposon classes showed an increase in expression (Supplemental Fig. S6A–C). The up-regulation of repeats and loss of DNA methylation over transposons prompted us to question if they have gained proliferative potential upon *Pol IV kd*. This possibility was previously predicted to influence phenotypes and genomic integrity in a number of reports (Cui

et al. 2013; Wei et al. 2014; Debladis et al. 2020). To verify such a possibility, we performed a PCR-based assay to detect the extra-chromosomal circular DNA (ECC DNA) intermediates that are produced as by-products of transposition (Lanciano et al. 2017). We observed additional bands portraying increased proliferative potential of *PopRice* and *Tos17* transposons (Supplemental Fig. S6D). These bands were observed even after the removal of linear DNA using a specific exonuclease (Supplemental Fig. S6E; Supplemental Methods). We examined if specific transposons proliferated by performing transposon display Southern hybridizations. Among the *kd* lines tested, a specific line showed a discernible copy number increase of the LINE-1 element across generations (Supplemental Fig. S6F). This observation is in line with a previous report in which silencing of *NRPD1* caused an increase in transposon copy number (Debladis et al. 2020). The proliferation of transposons might be the reason for exacerbated reproductive defects in selected *kd* lines (Supplemental Fig. S3A). Ribosomal RNA (rRNA) precursor expression and the expression of several genes were also misregulated in the *kd* lines compared with WT (Supplemental Fig. S5B,C). The genes misregulated in anthers and panicles of the *kd* lines also showed a significant overlap with the DEGs seen in *nrdp1-RNAi* shoot bases (Xu et al. 2020), suggesting a conserved effect of loss of Pol IV in different tissues of rice.

In summary, the activity of RNA Pol IV and ensuing sRNAs are essential for regulating the expression of genes and repeats via DNA methylation, and loss of these processes results in the potentially genotoxic proliferation of transposons in the rice genome.

### Activity of RNA Pol IV potentiates silencing by H3K27me3 and H3K9me2 marks and modulates their relative occupancy over genes and repeats

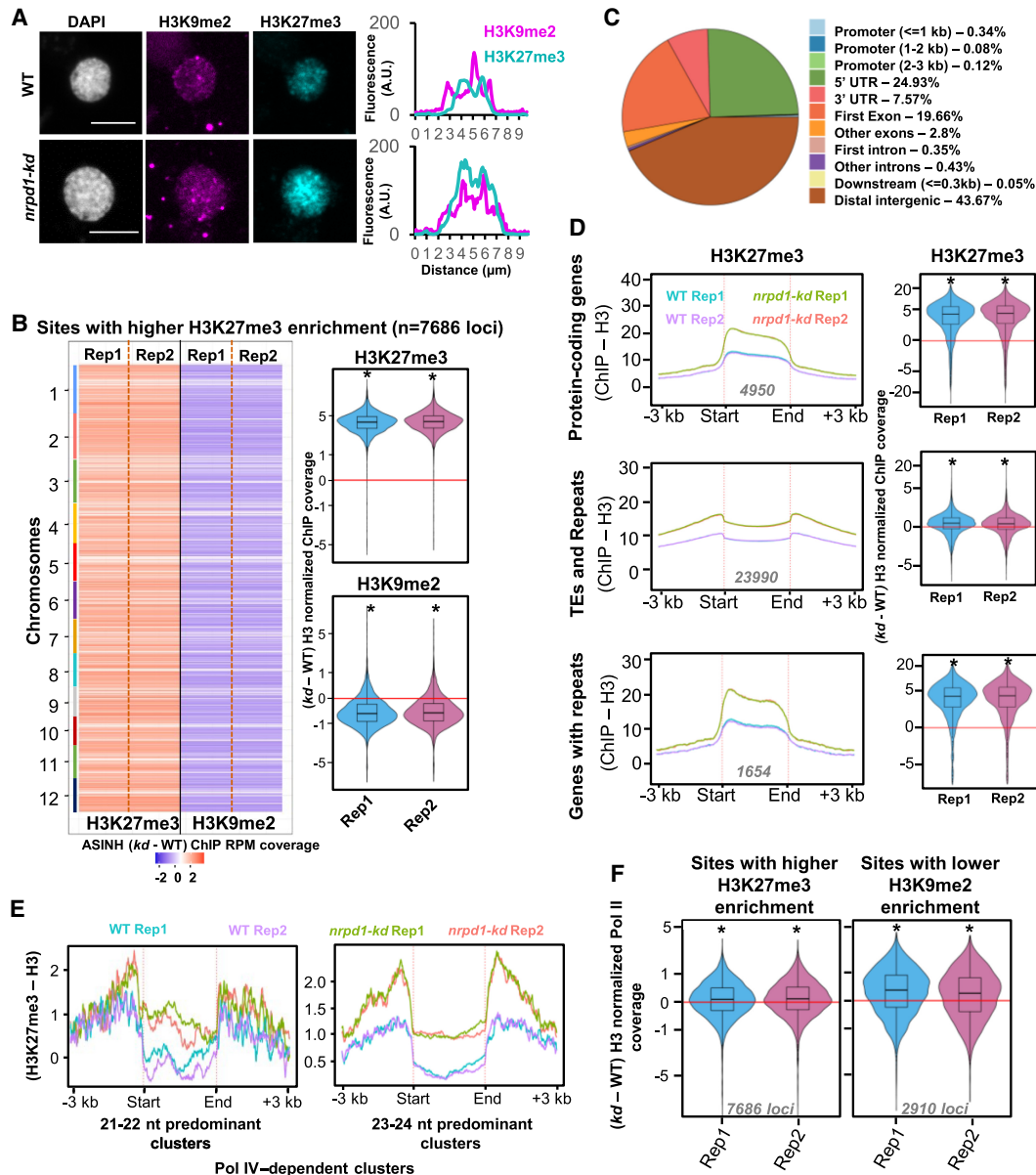
Several reports have suggested that the effective silencing is brought about by the cross talk between DNA methylation and histone modifications in a locus-specific manner in plants (Mathieu et al. 2005; Deleris et al. 2012; Gent et al. 2015; Zhou et al. 2016). These reports have found that loss of DNA methylation led to perturbation of post-translational modifications (PTMs) of histones at the respective loci, together influencing the status of transcriptional silencing (Jamge et al. 2022). Recent studies in *Arabidopsis* have suggested that the loss of Pol IV and associated sRNAs can influence the histone H3 PTMs levels at specific loci (Li et al. 2020; Parent et al. 2021). In particular, Rougée et al. (2021) have described H3K27me3 redistribution at a subset of *ddm1* hypomethylated loci in *Arabidopsis* without being able to repress the hypomethylated loci. These observations in *Arabidopsis*, along with lack of in-depth epigenetic analysis in *Arabidopsis nrdp1* mutants, prompted us to explore the multiple layers of epigenetic changes upon *kd* of Pol IV in rice.

To examine the polycomb-mediated H3K27me3 and constitutive heterochromatic mark H3K9me2 profiles, we performed nuclear immunostaining that revealed broader H3K27me3 occupancy in *kd* over the H3K9me2 regions, unlike the WT, corroborating observations in *Arabidopsis* DNA methylation mutants (Fig. 2A; Supplemental Figs. S8, S9). Unlike *Arabidopsis* nuclei with distinct chromocenters, rice nuclei showed a wide distribution of H3K9me2 signal, suggesting a broad distribution of chromatin states in monocots (Liu et al. 2017). Further, to examine the H3K27me3 marks and other epigenetic marks in a locus-specific manner, we performed ChIP-seq with H3K27me3-, H3K4me3-, and H3K9me2-specific antibodies in replicates that showed consistent enrichment profiles (Supplemental Fig. S7). We analyzed the

profiles of these marks genome-wide, showing clear enrichment of H3K4me3 over PCGs, the absence of H3K9me2 at the genic loci, and the presence of H3K27me3 at both genic and nongenic regions (Methods) (Supplemental Fig. S10A). Differential enrichment analyses revealed 7686 loci possessing significantly higher H3K27me3 occupancy in *kd* compared with WT, validating the immunostaining observations (Supplemental Fig. S10B). We found that the loci that showed enhanced occupancy of H3K27me3 marks also showed loss of H3K9me2 marks (Fig. 2B). This observation points out that Pol IV activity is quintessential not only for maintaining the DNA methylation and sRNA profiles but also for buffering the relative locations of gene-specific H3K27me3 marks and repeat-specific H3K9me2 marks. In the absence of Pol IV, H3K27me3 marks intruded into new domains. Several sites showed reduced enrichment of H3K27me3 and misregulation of other marks (Supplemental Fig. S10B). On the other hand, H3K9me2 marks were reduced over the loci intruded by H3K27me3 marks, likely redistributing the silencing signatures (Fig. 2B).

Furthermore, annotation of the H3K27me3 gained loci showed that almost 55% corresponded to coding genes, with ~45% of the loci concentrated at the 5' UTR (24.93%) and first exons (19.66%) of the PCGs (Fig. 2C). All the non-PCGs accounted for 44% of the H3K27me3 overenriched loci (Fig. 2C). Analysis of H3K27me3 peaks further revealed that it was globally overenriched at PCGs, TEs, and repeat loci, thus implicating the spread of this mark into both PCGs and non-PCGs in *kd* (Fig. 2D). Moreover, the activation mark H3K4me3 also showed differential enrichment and overabundance over the PCGs and non-PCGs in *kd* (Supplemental Fig. 10B,C). Because monocot genomes are relatively enriched with repeats where they can pronouncedly influence the proximal genes to a greater extent than dicots (Hirsch and Springer 2017), we selected peaks containing PCGs that have at least 10% of their length overlapping with repeats (“genes with repeats”) and explored the redistribution of histone marks. As expected, this subset of PCGs recapitulated the hyper-occupancy of the H3K27me3 marks over the gene bodies to a greater degree than PCGs (Fig. 2D). Moreover, the hallmark RdDM sites enriched in Pol IV-dependent sRNAs gained H3K27me3 marks (Fig. 2E). This is in line with the observation that, even though the silencing by H3K9me2 and H3K27me3 is compartmentalized, it can, to a certain degree, substitute and overlap in cases of perturbation of other silencing signatures (Deleris et al. 2012; Délérís et al. 2021).

In spite of the spreading of H3K27me3 silencing in *kd*, we observed enhanced transcription of transposons and repeats in *kd* lines (Supplemental Fig. S6), indicating that H3K27me3 overshoot and intrusion did not completely compensate for the loss of repressive signatures of H3K9me2, DNA methylation, and sRNAs, supporting earlier investigations in *Arabidopsis* (Rougée et al. 2021). To test if occupancy of H3K27me3 marks can effectively prevent transcription at the loci that lost H3K9me2 marks, we performed ChIP-seq with RNA Pol II antibody. This analysis revealed the overoccupancy of Pol II ChIP signals over the sites with increased H3K27me3 or reduced H3K9me2 occupancy (Fig. 2F), suggesting inefficient suppression of transcription by H3K27me3 as established in *ddm1* mutants in *Arabidopsis* (Rougée et al. 2021). Taken together, we identify a novel effect of loss of Pol IV in buffering the occupancy of H3K9me2 and H3K27me3 marks over the PCGs and repeats. Failure to establish this balance led to a more permissive chromatin state, promoting misregulation of PCGs and repeats.



**Figure 2.** Pol IV maintains the genome-wide distribution of H3K27me3-mediated silencing states. (A) Immunostaining images of H3K27me3 and H3K9me2 marks in the nuclei extracted from WT and *kd* seedlings. DAPI-stained DNA. Scale bar: 5 μm. At least 30 nuclei from each genotype were examined. Fluorescence signals over the regions of interest are plotted. Regions of interest are also shown in Supplemental Figures S8 and S9. (B) Chromosome-resolved heatmaps showing the levels of difference in H3K9me2 enrichment at sites with higher H3K27me3 enrichment in *kd* compared with WT. Box-violin plots show the distribution of the differences in H3K9me2 and H3K27me3 levels. The y-axis is scaled to the inverse sine hyperbolic function of differences. (C) Pie chart showing the feature annotation of the sites shown in B. All the sites that do not fall in the ±3-kb windows of genes were categorized as distal intergenic regions. (D) Metaplots showing the levels of H3K27me3 enrichment (normalized to H3) over the annotated features overlapping with the identified H3K27me3 peaks. Numbers in gray depict the number of loci. Box-violin plots show the difference in enrichment in *kd* compared with the WT. The y-axis is scaled to the inverse sine hyperbolic function of H3 ChIP-normalized values. (E) Metaplots showing the abundance normalized enrichment of H3K27me3 marks w.r.t. H3 over the Pol IV-dependent sRNA clusters size categorized as 21- to 22-nt and 23- to 24-nt predominant clusters. (F) Box-violin plots showing the difference in Pol II coverage over the H3K27me3 higher-enriched sites and H3K9me2 lower-enriched sites in *kd* compared with the WT. Numbers in gray describe the number of loci taken for analyses. The y-axis is scaled to inverse sine hyperbolic function of enrichment values. (B,D,F) A Mann–Whitney *U* test was used to test statistical significance against WT Rep1. (\*) *P*-value < 0.0001.

**Aberrant transcription in *kd* lines gives rise to atypical sRNAs that are suppressed by Pol IV**

Relaxed chromatin state and the reduction of silencing signatures in the *kd* lines might favor the enhanced transcription. Such observations were routinely reported in model dicot *Arabidopsis* and

other organisms when mutants with altered chromatin state were analyzed (Zheng et al. 2009; McKinlay et al. 2018; Ishihara et al. 2021; Jamge et al. 2022). To address this possibility, we performed a MNase accessibility assay from the chromatin of WT and *kd* plants. This analysis indicated relaxed chromatin status in *kd* compared with WT as *kd* lines accumulated shorter digested

fragments (Supplemental Fig. S11A). Also, we profiled the Pol II status over PCGs and genes with repeats using ChIP-seq. Even though we did not observe a significant difference of Pol II occupancy over all PCGs, we observed a significant hyperoccupancy over the genes with repeats (Supplemental Fig. S11B,C). Because the antibody captures all the states of Pol II, we postulated whether aberrant assembly, stalling, or elongation of the Pol II in *kd* might trigger the accumulation of misformed, potentially aberrant transcripts. We explored if these transcripts over the subset of PCGs might trigger production of sRNAs from these regions. As anticipated, sRNA analysis showed that the numerous ShortStack (Axtell 2013)-identified sRNA clusters (“clusters”) were up-regulated in *kd* in all the three tissues tested in both of the size classes (Supplemental Fig. S12A–C,E; Supplemental Methods). The sRNAs that are up-regulated in *kd* were denoted as Pol IV-suppressed sRNAs (“suppressed sRNAs”), and those that are reduced in *kd* were named as Pol IV-dependent sRNAs (“dependent sRNAs”; Supplemental Methods). Suppressed sRNAs did not show any first nucleotide bias (Supplemental Fig. S12D). Further analysis of the 21- to 24-nt sRNAs from the sRNA clusters revealed that these suppressed sRNA clusters had 21- to 22-nt predominant, 23- to 24-nt predominant, and mixed clusters of both of these size classes in almost equal proportions (Supplemental Fig. S13A; Supplemental Methods). The suppressed and dependent clusters had uniform length distribution (Supplemental Fig. S13B).

To have an unbiased estimation of the relative distribution of these sRNAs, we performed a genome window (100 bp)-based analysis (Supplemental Methods). Although dependent bins were higher in number in all the three Pol IV *kd* tissues, suppressed sRNA loci were fewer (Supplemental Fig. S14A–C). For exploring the size distribution of all inserts in the panicle sRNA data sets, we plotted size-density distribution of 16- to 45-nt sRNAs from suppressed and dependent bins (Fig. 3A). As expected, Pol IV-dependent sRNA bins accumulated sRNAs of the predominantly 24-nt size class, whereas suppressed sRNA bins accumulated wide fragment sizes ranging from 21–35 nt. We verified that suppressed sRNAs were not owing to oversampling of the residual sRNAs in the *kd* tissues and associated normalization artifacts by comparing the raw abundance of sRNAs in each library to the sum of miRNAs (Supplemental Fig. S15). The suppressed bins were concentrated at fewer selective loci (Supplemental Figs. S16A, S17A, S18A). Further, cumulative sum plots describing the relation between the abundance of sRNAs and the number of uniformly sized bins also indicated that suppressed sRNAs were abundantly concentrated in fewer bins (Supplemental Figs. S16B, S17B, S18B).

Further, to validate if suppressed sRNAs in *kd* are associated with other Pol IV machineries, we analyzed the published sRNA data sets derived from rice *nrdp1*, *rdr2*, and *nrdp1* seedlings (Zheng et al. 2021; Wang et al. 2022). We indeed observed the accumulation of suppressed sRNAs in *nrdp1* and *rdr2* but to a lesser degree in *nrdp1* seedlings (Fig. 3B). Further, we observed that the *nrdp1(d/e)2* sRNA data sets (Chakraborty et al. 2022) indeed retained the Pol IV-suppressed loci, suggesting only a minor role of Pol V in altering the chromatin state and the accumulation of these sRNAs (Fig. 3B). We also detected the Pol IV-suppressed sRNAs by sRNA northern hybridizations at two distinct loci (Fig. 3C; Supplemental Fig. S12E), proving the existence of the suppressed sRNAs in *kd* lines, and they were not observed upon NRDP1 CO (Fig. 3D). Furthermore, repeat features that are hallmarks of dependent sRNAs were completely devoid of suppressed sRNAs (Fig. 3E). Annotation of the differentially expressed sRNA bins from different tissues identified that nearly half of the Pol IV-suppressed sRNA bins

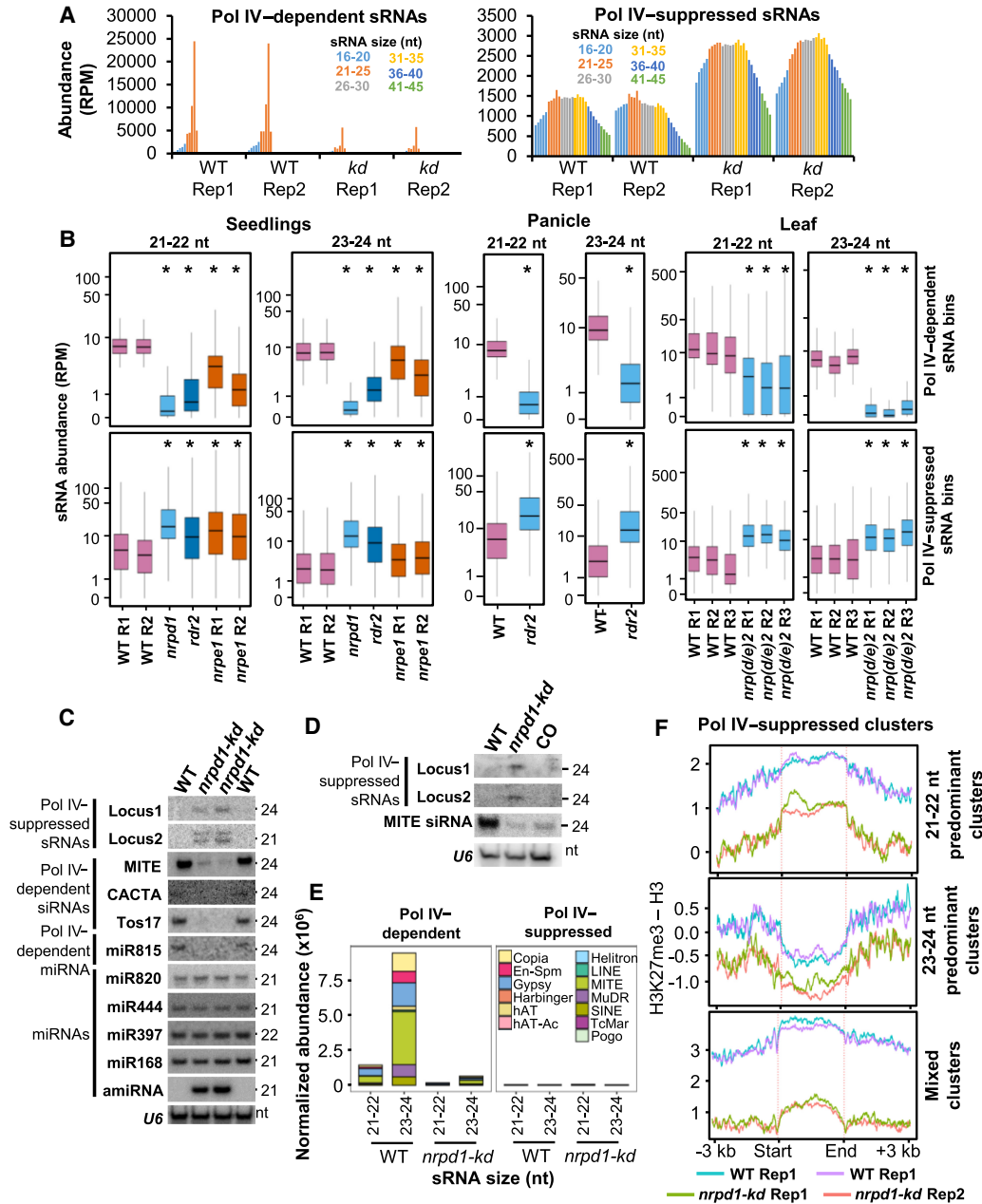
mapped to coding regions, whereas dependent counterparts mapped to the noncoding regions (Supplemental Fig. S19A). Pol IV-suppressed sRNAs from the coding regions can be either generated from hairpin structures of aberrant transcripts or generated from duplex RNAs of aberrant transcripts by the activity of RDRs. Specific AGOs will stabilize the preferred sRNAs. To test if dsRNAs are generated by the activity of RDRs (which will not show strand bias with respect to the coding gene), we counted suppressed sRNAs from bins mapping to coding genes. Suppressed sRNAs uniformly mapped to both the same mRNA encoding strand and the opposite strand, suggesting activity of one of the RDRs (Supplemental Fig. S19B). Unlike dependent sRNAs, suppressed sRNAs were within the gene body PCGs (Supplemental Fig. S19C).

Furthermore, occupancy of H3K27me3 marks at the suppressed loci of all size classes were lower than the WT in *kd* lines, suggesting a chromatin misregulation associated with these loci (Fig. 3F). Moreover, genes overlapping with the suppressed sRNA loci showed reduced H3K27me3 and H3K9me2 levels and increased H3K4me3 levels (Supplemental Fig. S20). In summary, we revealed a population of novel sRNAs that were suppressed by Pol IV showing atypical genomic origins and molecular signatures.

#### Pol IV-suppressed sRNAs are conserved in *Arabidopsis* and are dependent on RDR6 and DDM1

We speculated that altered chromatin state owing to loss of Pol IV might be conserved across plants and might be leading to the accumulation of suppressed sRNAs. In agreement with this, we observed that loss of *Arabidopsis* Pol IV resulted in suppressed sRNAs, very similar to rice, but from fewer bins in inflorescence tissues (Supplemental Fig. S21A; Zhou et al. 2018). Unlike Pol IV-dependent bins, suppressed sRNAs in *Arabidopsis* showed nonuniform distribution like rice tissues (Supplemental Fig. S21B). The availability of various genetic mutants in *Arabidopsis* prompted us to check the mechanistic origins of these sRNAs and possible regulators of chromatin state when Pol IV is not functional. In this direction, we examined several siRNA biogenesis mutants, especially connected to RdDM, and found that the suppressed sRNAs were found mainly upon loss of Pol IV itself or associated proteins using northern blots with various mutants (Fig. 4A; Supplemental Fig. S22). The suppressed sRNAs were unique to *nrdp1* and were not observed in sRNA-processing mutants such as *dcl3*, *dcl234*, or *ago4* or the mutant of chromatin remodeler *ddm1* (Fig. 4A). We also used sRNA data sets from the similar stage inflorescence tissues from additional set of mutants involved in RNA silencing (Zhai et al. 2015). This independent set of *nrdp1* data sets also confirmed the increase of suppressed sRNAs compared with the WT mainly from nonrepeat regions (Supplemental Fig. S22A,B). The suppressed sRNAs were also seen in other *nrdp1*-associated mutants like *rdr2*, *nrdp1(d/e)2*, *nrdp1dcl3*, and *rdr2dcl3* (Supplemental Fig. S22C). This indicates that the suppressed sRNAs are directly coupled to the absence of Pol IV complex (NRDP1 or RDR2). In addition, most of the suppressed sRNAs were brought back to the WT levels upon NRDP1 CO (Supplemental Fig. S22D).

To probe which part of the RdDM machinery is mainly involved in generation of Pol IV-suppressed sRNAs, we analyzed published sRNA data sets from inflorescence tissues of 28 different genotypes of *Arabidopsis* mutants, pertaining to three categories: Pol IV-associated mutants, Pol V and histone modification-associated mutants, and DNA methylation mutants (Supplemental Table S2). For the set of loci identified as suppressed and dependent loci, we evaluated the abundance of sRNAs across the

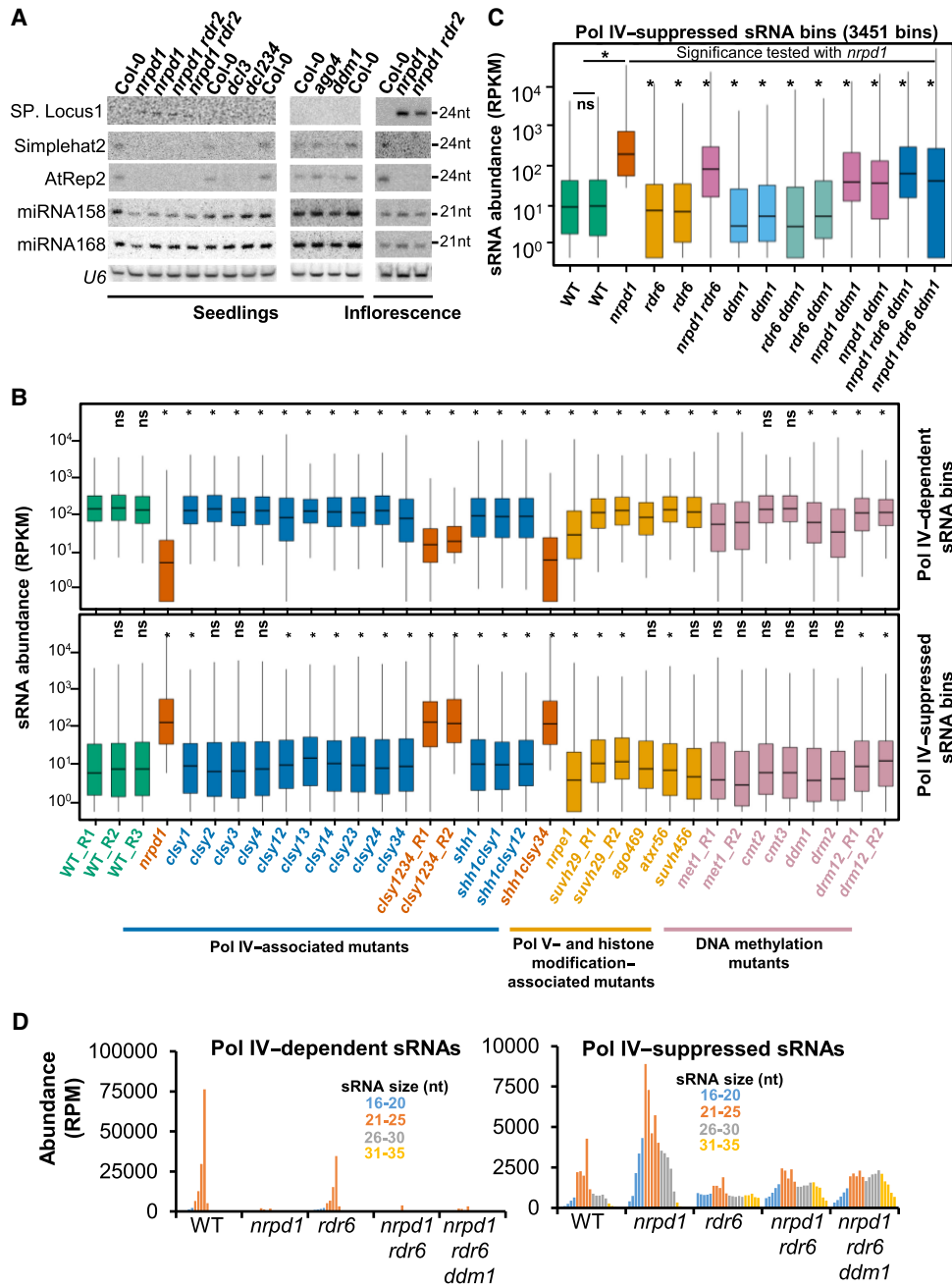


**Figure 3.** Pol IV complex suppresses sRNA production from several loci. (A) Bar plots showing the abundance of Pol IV–dependent and –suppressed sRNAs of 16- to 45-nt size range. (B) Boxplots showing the abundance of sRNAs from different mutants in rice. sRNAs were size-categorized into 21- to 22-nt and 23- to 24-nt and counted in 100-bp nonoverlapping windows. Plots depict the abundance of sRNAs in each size class over *nprp(d/e)2* (leaf), *nprp1* (seedlings), and *rdr2* (panicle) dependent and suppressed bins. The data sets were obtained from the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) under accession numbers GSE158709 (seedlings: *nprp1*, *nprp1*) and GSE130166 (seedlings and panicle: *rdr2*). The y-axis is scaled to inverse sine hyperbolic function of RPM values. Mann–Whitney *U* test was used to test statistical significance against WT R1. (\*) *P*-value < 0.0001. (C) sRNA northern blots validating the presence of Pol IV–suppressed sRNAs, –dependent sRNAs, and –independent miRNAs. *U6* was used as loading control. (D) sRNA northern blots showing the abundance of sRNAs in WT, *kd*, and *NRPD1* CO lines. *U6* was used as loading control. (E) Stacked bar plots showing abundance of Pol IV–dependent and –suppressed sRNAs from transposon categories. (F) Metaplots describing histone H3K27me3 occupancy normalized to total H3 signal over the Pol IV–suppressed sRNA clusters, categorized as 21- to 22-nt predominant clusters, 23- to 24-nt predominant clusters, and the mixed sized clusters.

mutants. We observed evident and drastic accumulation of suppressed sRNAs comparable to *nprp1* only in *clsy1234* quadruple mutants and in the *shh1clsy34* triple mutant (Fig. 4B). This observation points out that the suppressed sRNAs are the effect of ablation of Pol IV enzyme complex along with the machinery responsible for its assembly. On the contrary, the Pol IV–sup-

pressed sRNAs are not initiated upon removal of either the Pol V arm of RdDM or the DNA methyltransferases (Fig. 4B).

To explore the molecular machinery involved in the biogenesis of suppressed sRNAs upon loss of *nprp1*, we explored the abundance of sRNAs in combination mutants of *nprp1*. The sRNA abundance in the Pol IV–suppressed bins (*n* = 3451) identified in



**Figure 4.** Pol IV-suppressed sRNAs are conserved in *Arabidopsis*. (A) sRNA northern blots showing the levels of Pol IV-suppressed sRNAs (SP. Locus1) and other sRNAs from *Arabidopsis* seedlings and inflorescence tissues. (B) Boxplots showing the sRNA abundance from different genotypes of *Arabidopsis* with counts from Pol IV-dependent and -suppressed 23- to 24-nt size class bins. The sRNA data sets were obtained from GEO GSE165574, GSE45368, and GSE99694. Analyses depicts sRNAs from inflorescence tissues. (C) Boxplots showing the sRNA abundance from different genotypes of *Arabidopsis* with counts from Pol IV-suppressed 23- to 24-nt size class bins depicted in B. The sRNA data sets were obtained from GEO GSE79780 (inflorescence tissues). The y-axis is scaled to the inverse sine hyperbolic function of RPKM values. Mann-Whitney *U* test was used to test statistical significance against WT R1, unless stated otherwise. (\*) *P*-value < 0.0001. (D) Bar plots showing the abundance of Pol IV-dependent and -suppressed sRNAs of 16- to 35-nt size range.

previous analyses was counted from *ddm1*, *rdr6*, *nrpd1 rdr6*, and *nrpd1 rdr6 ddm1* data sets (Panda et al. 2016). This revealed that the production of suppressed sRNAs in the *nrpd1* mutant was partially dependent on RDR6 and DDM1, as the double or triple mutants of these genes in a *nrpd1* background showed a reduction in suppressed sRNAs compared with the *nrpd1* mutant (Fig. 4C; Supplemental Fig. S22E). Dependence on RDR6 and DDM1 was

also observed across all the sRNA size classes that mapped to suppressed sRNA bins but not for the dependent sRNA bins (Fig. 4D). Importantly, the dependence of suppressed sRNAs on silencing machinery genes establishes an active mode of suppressed sRNA biogenesis in the absence of the Pol IV complex.

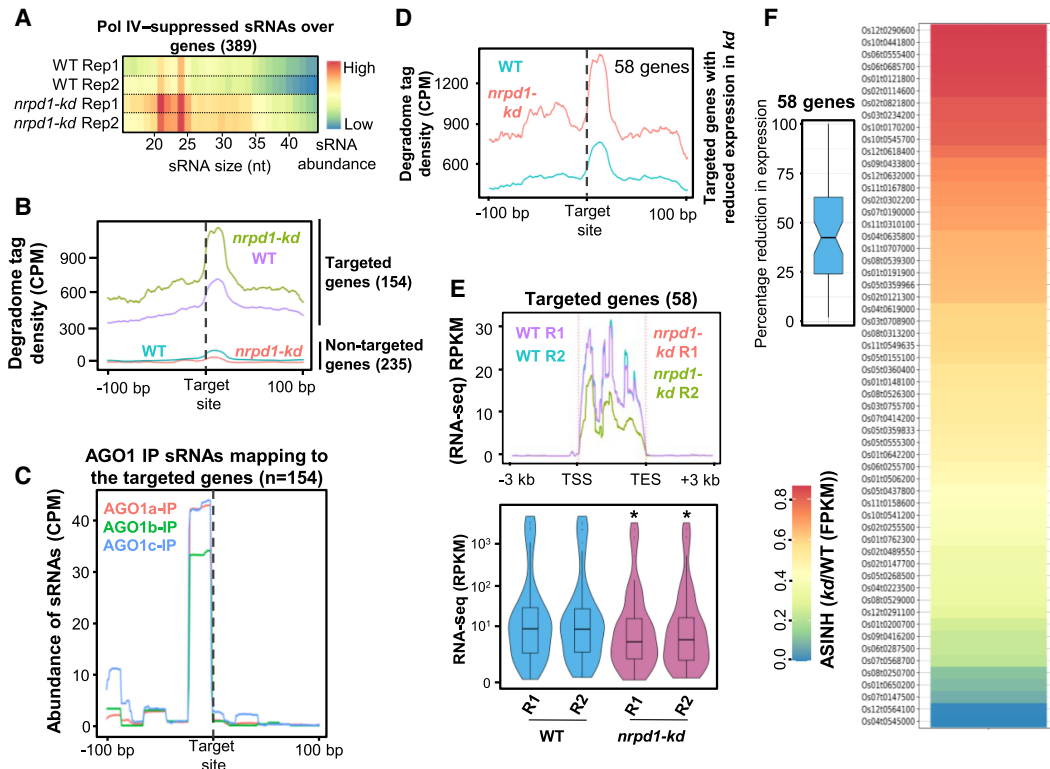
Probing the contributing molecular signatures associated with the suppressed sRNAs, we also tested the status of the DNA

methylation in reproductive tissues over the suppressed sRNA clusters and found that methylation over the CG, CHG, or CHH contexts in both rice and *Arabidopsis* remains largely unchanged (Supplemental Figs. S23, S24). Pol II occupancy over the suppressed and dependent loci also remained unchanged, likely owing to the transient interaction by Pol II in generating these transcripts (Supplemental Fig. S25). On the other hand, reduction of H3K27me3 over specific genomic regions upon loss of Pol IV might have triggered production of suppressed sRNAs (Supplemental Fig. S26), which are conserved across monocots and dicots.

### Pol IV-suppressed sRNAs are loaded into AGO1 to induce PTGS at protein-coding loci

Given that these abundant atypical sRNAs are from gene coding regions and they did not change the DNA methylation signature of the loci, we explored if they can target genes post-transcriptionally. Pol IV-suppressed sRNAs from PCGs are predominantly of the 21- or 24-nt size class unlike the broad size range of the bulk of suppressed sRNAs (Fig. 5A), suggesting processing by sRNA machinery. To test if AGO1-loaded suppressed sRNAs targeted genes post-transcriptionally, we performed degradome sequencing in WT and *kd* panicles. Among the genes that overlapped with the suppressed sRNAs (389 genes), target prediction tools identified 154 genes as potential targets of suppressed sRNAs (Supplemental Table S8; Supplemental Fig. S27B). The degradome tag density at

the predicted target loci was substantially increased in *kd* lines, and nontargets did not accumulate degradome tags despite being sources of suppressed sRNAs (Fig. 5B; Supplemental Fig. S27C). In further agreement with the targeting process, the WT AGO1 IP was enriched with suppressed sRNAs that mediated slicing as seen in degradome analysis (Fig. 5C; Supplemental Fig. S27A; Wu et al. 2009). The transcript-to-degradome read depth ratio for targets and nontargets showed expected profiles (Supplemental Fig. S27D,E). It would be interesting to explore the loading distribution of suppressed sRNAs in *kd* lines into specific AGO1 isoforms (out of AGO1a, -b, -c, and -d). Many of the target RNAs had significantly reduced expression (~40% mean), suggesting precise PTGS at these sites mediated by suppressed sRNAs (Fig. 5D-F; Supplemental Fig. S28). We also observed that many of the genes that underwent PTGS by suppressed sRNA-mediated targeting in rice were previously implicated in reproductive growth and development (Supplemental Table S8). For instance, *OsMADS18* (*APETALA1* homolog in rice), a member involved in floral architecture establishment (Wang et al. 2020a), a close homolog of fertility restorer (*RF*; Os10t0497366), a glycine-rich interaction partner of RF5 (Os12t0632000) (Hu et al. 2012), and a pollen-specific desiccation-associated protein (Os11t0167800) were targeted for degradation in *kd* lines (Supplemental Table S8). Targeting of aberrant transcripts originating from atypical loci specifically in *kd*, but not in WT, by suppressed sRNAs is an area for further investigation.



**Figure 5.** Pol IV-suppressed sRNAs get loaded into AGO1 to mediate PTGS. (A) Heatmaps showing size distribution of sRNAs from Pol IV-suppressed sRNA bins overlapping with genes (389 genes). (B) Metaplots showing the degradome tag density over the genes identified as targets in *nrpd1-kd* panicle degradome (green and purple; 154 genes) and the same for the genes that are not effectively targeted (category > 5) but produce Pol IV-suppressed sRNAs (cyan and orange; 235 genes). Targeted locations are centered, and 100 bp on either side is displayed. (C) Metaplots showing the abundance of AGO1 IP-enriched sRNAs over the identified targeted genes centered at the targeting site. (D) Metaplots showing the degradome tag density of the 58 genes that showed targeting and significant reduction in expression. (E) Metaplots and box-violin plots showing the expression status of 58 targeted genes described in D. (F) Heatmap showing the fold changes (scaled to inverse sine hyperbolic function) in expression (FPKM) of the 58 genes that showed reduction in expression compared with the WT. The *inset* boxplot shows the distribution of reduction in expression (FPKM) observed.

On the other hand, the suppressed sRNAs in *Arabidopsis* did not accumulate in the AGO1-enriched fractions in both Col-0 and *mpd1* plants more than the mock IP controls (Supplemental Fig. S29). On the contrary, Pol IV-dependent sRNAs were abundantly found in AGO4 IP data sets (Supplemental Fig. S29A; Zhai et al. 2015; Panda et al. 2020). AGO1 targeting is unlikely or is rare to be detected in bulk analyses in *Arabidopsis* in which suppressed sRNAs are lowly abundant nor loaded into AGO1 (Supplemental Fig. S29C). It is worth exploring the AGO1-mediated targets of suppressed sRNAs in other plant species in which perturbation of NRPD1 showed strong phenotypes (Grover et al. 2018; Wang et al. 2020b). Such a targeting mediated by suppressed sRNAs might act as strong deterrent for the loss of Pol IV activity in plants with complex genomes.

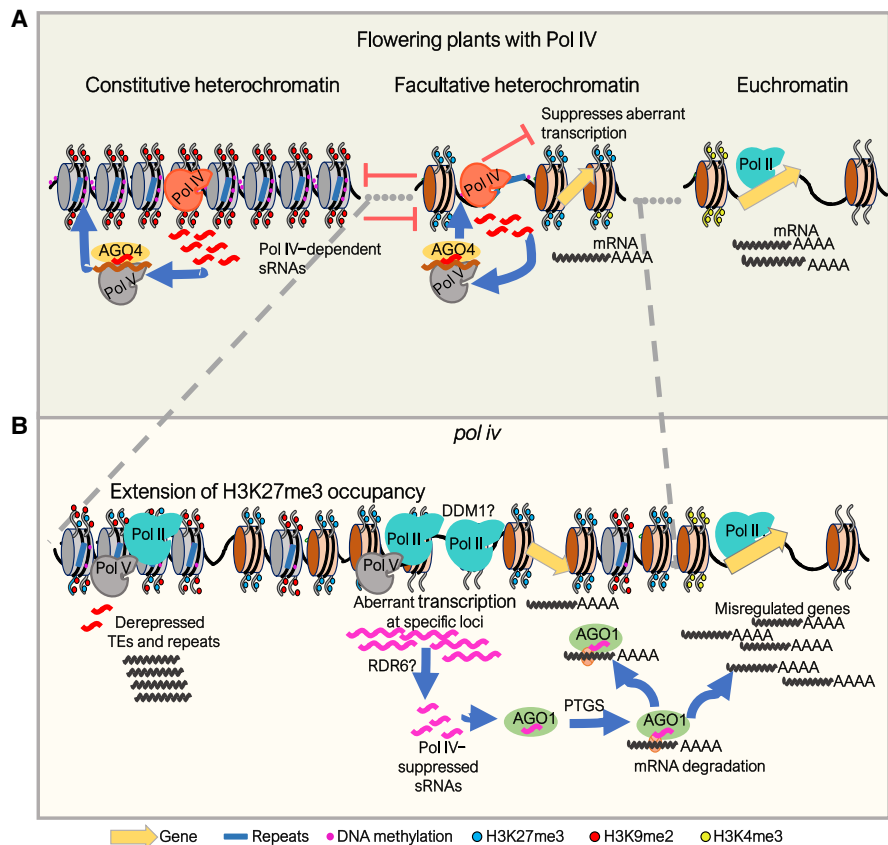
**Discussion**

The precise and dynamic regulation of epigenetic modifications in a coordinated manner hallmarks the center stage of regulation. Our investigation uncovers the role of plant-specific RNA polymerase, Pol IV, in control over locus-specific epigenetic marks and prevention of illegit transcription.

Evolutionary analyses suggest that the plant-specific RNA polymerases IV, V, and, in specific cases, VI are novel machineries evolved in conjunction with the complexity of plant genomes, and the degree of importance of these polymerases varies across plants. For instance, in early land plants like *P. patens*, the pathway is completely redundant with other sRNA-independent DNA methyltransferases introducing de novo DNA methylation (Yaari et al. 2019) as opposed to the widespread defects observed in rice (Fig. 1; Supplemental Figs. S1–S3; Xu et al. 2020; Zheng et al. 2021). The functions of these polymerases are pronounced in reproductive tissues as they are involved in faithful transmission of epigenetic information, antagonizing genome dosage aberrations, and hybridization (Zhang et al. 2016; Erdmann et al. 2017; Martinez et al. 2018; Satyaki and Gehring 2019). It is apparent that the green-lineage-specific RNA polymerases have neofunctionalized to perform additional roles with increasing genome complexity. The comparisons we present with respect to the redistribution of epigenetic marks between rice and *Arabidopsis* serve as strong evidence to these predictions.

Several investigations in the model plant *Arabidopsis* have established that Pol IV initiates biogenesis of sRNAs from repeats and transposons establishing de novo DNA methylation aiding against their genotoxic proliferation. Our studies in rice corroborate this function of Pol IV in multiple reproductive tissues (Figs. 1, 6A; Supplemental Fig. S4). Several reports have mentioned the synergistic effects and cross dependence of histone

modifications and global CG DNA methylation in establishing epigenetic modalities in *Arabidopsis* (Soppe et al. 2002; Mathieu et al. 2005; Deleris et al. 2012; Li et al. 2018; Zhong et al. 2021; Zhao et al. 2022). These studies had used *met1*, *ddm1*, and other methyltransferase mutants, resulting in a loss of compaction at the pericentromeres and constitutive heterochromatin domains. On the other hand, the RdDM is not restricted to heterochromatin, and its purview extends into the gene-rich regions as well. Whether modulation of the sRNAs can directly perturb the chromatin states attributable to the gene expression was unclear. Especially, in monocots like rice, the interspersion of repeat fragments within genes mandates gene regulation not at extended length scales but in localized compartments even within the euchromatin (Espinas et al. 2020). Probing the effect of the loss of RdDM on the chromatin states over the genes in our study uncovered a distinct redistribution of H3K9me2 and H3K27me3 marks (Figs. 2, 6B). Predominantly, H3K9me2 marks, attributed with well-suppressed constitutively silenced loci and H3K27me3 associated with silenced gene coding units by action of PRC2 showed overoccupancy in *kd* over the transposons and genes (Fig. 6B). The authentic Pol IV transcribing loci that result in sRNAs showed significant H3K27me3 occupancy, likely an ectopic compensatory mode of silencing (Fig. 2E). This feature is evolutionarily conserved between plants and animals, and loss of transposon methylation can contribute to the redistribution of



**Figure 6.** Pol IV complex precludes the spread of aberrant silencing. (A) Transcription by Pol IV produces sRNAs that guide silencing via DNA methylation at euchromatic TEs. Its activity also suppresses unengaged transcription by other polymerases and optimal distribution of H3K27me3 over coding genes. (B) Upon loss of Pol IV, distributions of H3K27me3, H3K9me2 marks are perturbed, and several transposons and genes are marked by H3K27me3. In addition, upon loss of Pol IV, unengaged aberrant transcription leads to production of Pol IV-suppressed sRNAs, partly dependent on RDR6 (post-transcriptionally) and DDM1 (transcriptionally). In rice, these transcripts are loaded into AGO1 and can target genes.

these marks into transposon territories (Deleris et al. 2012; Délérís et al. 2021). The proximity of the PCGs and transposons might be the trigger causing the intrusion of facultative silencing marks on the genes in species with larger genomes such as rice. This is well supported by the fact that upon *kd* of the Pol IV, genes that are dispersed with the repeat fragments showed an increased degree of H3K27me3 intrusion (Fig. 2E). PRC2-mediated H3K27me3 on genes impacts reproductive success in plants by controlling several imprinted genes that scale the genome dosage (Köhler et al. 2003; Roszak and Köhler 2011; Kradolfer et al. 2013; Jiang et al. 2017). Such indirect effects on the PCGs by unwarranted H3K27me3 marks, to a certain extent, might explain the defects in *kd*.

Loss of silencing signals and redistribution of H3K27me3 marks over the PCGs upon loss of RdDM in rice should result in differential accessibility of these loci to RNA polymerases. Ungauged transcription in the absence of Pol IV might feed to the sRNA pools via activity of DCLs that are devoid of Pol IV precursor load. Exploration of sRNA pools in the *nripd1* mutants of rice and *Arabidopsis* indeed showed resultant aberrant sRNAs, likely triggered by spurious transcripts. This pool of Pol IV-suppressed sRNAs is partly dependent on RDR6 and DDM1, as the double mutants of these genes in the *nripd1* background showed a reduction in their abundance. Such spuriously transcribed loci showed a reduction in H3K27me3 occupancy (Fig. 3F). This commonality of spurious sRNA transcripts in RdDM mutants was observed earlier (Zheng et al. 2021), where a similar notion of chromatin relaxation triggering Pol IV and Pol II transcription is promulgated. Similarly, studies in maize and *Arabidopsis* have suggested atypical transcripts in the *nripd1* or *rdr2* mutation (Lu et al. 2006; Kasschau et al. 2007; Erhard et al. 2015; McKinlay et al. 2018). Our studies mechanistically delineate the causative chromatin features in the *nripd1* mutant. Not limiting to that, we find that the resultant Pol IV-suppressed sRNAs are capable of targeting genes post-transcriptionally by loading into specific AGOs (Fig. 6B). The ability of suppressed sRNAs to induce phased sRNAs and their activity *in trans* is worthy of investigation.

These analyses reveal that the atypical Pol IV-suppressed sRNAs in diverse plants such as rice and *Arabidopsis* were a result of misregulated chromatin states that are maintained by Pol IV. In rice, suppressed sRNAs were capable of targeting PCGs post-transcriptionally. We catalog reciprocity of the silencing between transcriptional and post-transcriptional modes, repurposing an existing RNA targeting machinery. We believe this acts as a second line of defense when the conventional transcriptional silencing via RdDM is impaired. This is in turn regulated by aberrant transcription and production of suppressed sRNAs capable of PTGS. *Arabidopsis nripd1*-related mutants also showed production of these sRNAs; nevertheless, they did not get loaded into AGO1 for active targeting. The absence of efficient PTGS in *Arabidopsis nripd1* mutants might have alleviated the reproductive defects observed in rice. This difference between rice and *Arabidopsis* might be indicating stronger selection pressure for Pol IV function in complex genomes. Our results suggest a strong and multifaceted impact of Pol IV on the expression of PCGs, while enabling evolution of additional genomic complexities, architecture, and heterogeneity.

## Methods

### Plant material

*Indica* variety rice (*Oryza sativa indica* sp.) Pusa Basmati 1 (PB1) plants were grown in a growth chamber at 24°C/70% RH with a

16-h–8-h light–dark cycle for hardening before transferring to a greenhouse maintained at 28°C with a natural day–night cycle. *Arabidopsis* plants were grown in a growth chamber maintained at 22°C/70% RH with a 16-h–8-h light–dark cycle. The different mutants used are reported in Supplemental Table S4.

### Binary vector construction and *Agrobacterium* mobilization

For the amiR construct, the artificial miRNA was designed using the WMD3 tool (Ossowski et al. 2008). The amiR was chosen so that it targets both of the NRPD1 isoforms (NCBI IDs: NRPD1a, XM\_015781553.2; NRPD1b, XM\_015756207.1). The mature amiR sequence (5'-UAUAGUGUACUCUUGGACAU-3') was embedded in the *OsamiR528* precursor in the pNW55 plasmid (Supplemental Table S3; Ossowski et al. 2008; Warthmann et al. 2008). The precursor region was subcloned into a binary vector (pCAMBIA1300 backbone) between *ZmUbiquitin1* promoter and 35S poly(A) signal and mobilized into *Agrobacterium* using the electroporation method.

For the CO construct, the NRPD1b promoter and 5' UTR were amplified from the genomic DNA (Chr 9: 22,015,503–22,019,281) and fused to the amiR-resistant CDS of OsNRPD1b, and this was cloned into pCAMBIA3300 (with bialaphos selection, BIPR). The construct was supertransformed to the HygR *kd* calli.

### Southern hybridization

Southern hybridization was performed as previously described (Ramanathan and Veluthambi 1995; G and Shivaprasad 2022). Total DNA (5 µg) was isolated and digested using 30 units of appropriate restriction enzyme. The digested DNA was electrophoresed on a 0.8% agarose gel in 1× TBE buffer and capillary-transferred to a Zeta probe nylon membrane (Bio-Rad). The probes (Supplemental Table S3) were labeled using [ $\alpha$ -P<sup>32</sup>] dCTP (BRIT India) using a Rediprime labeling kit (GE Healthcare). The blots were exposed to phosphor screen and scanned using a Typhoon scanner (GE Healthcare).

### sRNA northern hybridization

sRNA northern blots were performed as described earlier (Shivaprasad et al. 2012; Tirumalai et al. 2020). Around 15 µg of TRIzol extracted total RNA from different tissues was electrophoresed on a denaturing 15% acrylamide gel. The gel was electroblotted onto Hybond N+ membrane (GE Healthcare) and UV cross-linked. The membrane was hybridized with the T4 PNK end-labeled oligonucleotides (with [ $\gamma$ -P<sup>32</sup>]-ATP) in Ultrahyb buffer (Invitrogen). The blots were washed, exposed to phosphor screen, and scanned using a Typhoon scanner (GE Healthcare). Post scanning, blots were stripped at 80°C in stripping buffer and proceeded with repeat hybridizations with subsequent probes.

### Chromatin immunoprecipitation-sequencing (ChIP-seq) and analyses

ChIP was performed as described earlier (Saleh et al. 2008; Song et al. 2016). Around 1.2 g of pre-emerged panicle tissues was taken and cross-linked with 1% formaldehyde. The tissues were pulverized in liquid nitrogen, and nuclei were isolated. Equal numbers of nuclei were lysed and sheared using ultrasonication (Covaris) until fragments reached 150–350 bp in size. The sheared chromatin was incubated overnight with 50 µL of antibody (H3K4me3–Merck 07-473; H3K9me2–Abcam ab1220; H3K27me3–Active Motif 39155; H3–Merck 07-10254; Pol II–Abcam ab817) bound protein–G Dynabeads (Thermo Fisher Scientific) at 4°C. The purified IP products were taken for library preparation using NEBNext

Ultra II DNA library prep kit with sample purification beads (NEB E7103L) as per the manufacturer's protocol. The libraries (with replicates) were sequenced on an Illumina HiSeq 2500 platform (Supplemental Table S1).

The obtained data sets were adapter-trimmed using cutadapt (Martin 2011) and aligned to the IRGSP 1.0 genome using the Bowtie 2 tool (Langmead and Salzberg 2012) with the following parameters: `-v 1 -k 1 -y -a --best --strata`. PCR duplicates were removed. The alignment files were converted to coverage files and compared (difference) to total H3 signal (for histone H3 PTMs) using bamcompare utility of deepTools (Ramírez et al. 2014). The average signals over the desired regions/ annotations were estimated using computematrix (deepTools), and the coverage signal meta-plots were plotted using plotprofile that were modified using a custom R script (R Core Team 2021). The peak calling was performed using MACS2 (Zhang et al. 2008), with broad peak calling for all ChIP data sets except for H3K4me3. The peaks with enrichment above threefold compared with H3 ChIP were taken as valid peaks, and the peaksets were merged using BEDTools across genotypes and replicates. The composite peak sets were intersected with the annotated PCGs and repeats to get the peak overlapping features used for signal counting (deepTools). The replicate concordance was measured across the samples using deepTools plot correlation feature.

### sRNA sequencing and analyses

sRNA sequencing was performed from pre-emerged panicle, anther, and endosperm of WT and *kd* plants. The size fractionation and library preparation were performed as previously described (Tirumalai et al. 2019). Reads were quality checked, adapter-removed, and size-selected using UEA sRNA workbench (Stocks et al. 2018). The reads are aligned after categorization into 21- to 22-nt or 23- to 24-nt sizes to IRGSP1.0 genome using Bowtie 2 (Langmead and Salzberg 2012) with the following parameters: `-v 1 -m 100 -y -a --best --strata`. Step-wise analyses for genomic annotations and 5'-nucleotide abundance were performed as previously described (Swetha et al. 2018). Only the mapped reads were used for further analyses, including differential expression analyses (Supplemental Table S5–S7). miRNA abundance was calculated using the miRProf tool (Stocks et al. 2018). *Arabidopsis* data sets were aligned to the TAIR10 genome.

### Data access

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE180457. All the necessary bioinformatics analyses scripts have been provided as Supplemental Code.

### Competing interest statement

The authors declare no competing interests.

### Acknowledgments

We thank Prof. K. Veluthambi for *Agrobacterium* strains, PB1 seeds, and binary plasmids. We thank Prof. David Baulcombe for the *Arabidopsis* mutants. We thank ClIFF, genomics, electron microscopy, IT, radiation, greenhouse, and laboratory-kitchen facilities at the NCBS. We acknowledge help from Anushree Narjala in bioinformatic analysis. We thank Dr. Aswin Sai Narain Seshasayee for his support. We thank Dr. Dimple Notani for sharing reagents. We thank Rahul Singh, M. Rajagopalan, and S. Goyal for

subcloning. We acknowledge the R-based metaplotting script from the GitHub page of Jeffrey Grover (<https://github.com/groverj3>). We thank Nitish Dua and Mohammad Shariq for the help with immunostaining. We thank all the laboratory members for discussions and comments. This work was supported by NCBS-TIFR core funding, Department of Atomic Energy, Government of India, project identification no. RTI 4006 (1303/3/2019/R&D-II/DAE/4749 dated 16.7.2020) and grants (BT/IN/Swiss/47/JGK/2018-19; BT/PR25767/GET/ 119/151/2017) from Department of Biotechnology (DBT), Government of India. C.S. and K.P. acknowledge fellowship from DBT, India. These funding agencies did not participate in the designing of experiments, analysis, or interpretation of data or in writing the manuscript.

**Author contributions:** P.V.S. and V.H.S. designed all experiments, discussed results, and wrote the manuscript. V.H.S. performed most of the experiments and bioinformatics analyses. C.S. performed bioinformatics analysis. D.B. generated transgenic lines. K.P. performed electron microscopy and micro-CT scanning. S.R. performed confocal imaging. T.C. and R.A.M. generated and analyzed the *nrp(d/e)2* sRNA data sets. All authors have read and approved the manuscript.

### References

- Axtell MJ. 2013. ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA* **19**: 740–751. doi:10.1261/rna.035279.112
- Baulcombe D. 2004. RNA silencing in plants. *Nature* **431**: 356–363. doi:10.1038/nature02874
- Baumberger N, Baulcombe DC. 2005. *Arabidopsis* ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proc Natl Acad Sci* **102**: 11928–11933. doi:10.1073/pnas.0505461102
- Chakraborty T, Trujillo JT, Kendall T, Mosher RA. 2022. A null allele of the pol IV second subunit impacts stature and reproductive development in *Oryza sativa*. *Plant J* **111**: 748–755. doi:10.1111/tpj.15848
- Chenxin L, Hengping X, Fang-Fang F, Scott DR, Venkatesan S, Jonathan IG. 2020. Genome-wide redistribution of 24-nt siRNAs in rice gametes. *Genome Res* **30**: 173–184. doi:10.1101/gr.253674.119
- Coruh C, Cho SH, Shahid S, Liu Q, Wierzbicki A, Axtell MJ. 2015. Comprehensive annotation of *Physcomitrella patens* small RNA loci reveals that the heterochromatic short interfering RNA pathway is largely conserved in land plants. *Plant Cell* **27**: 2148–2162. doi:10.1105/tpc.15.00228
- Cuerda-Gil D, Slotkin RK. 2016. Non-canonical RNA-directed DNA methylation. *Nat Plants* **2**: 16163. doi:10.1038/nplants.2016.163
- Cui X, Jin P, Cui X, Gu L, Lu Z, Xue Y, Wei L, Qi J, Song X, Luo M, et al. 2013. Control of transposon activity by a histone H3K4 demethylase in rice. *Proc Natl Acad Sci* **110**: 1953–1958. doi:10.1073/pnas.1217020110
- Debladis E, Lee T-F, Huang Y-J, Lu J-H, Mathioni SM, Carpentier M-C, Llauro C, Pierron D, Mieulet D, Guiderdoni E, et al. 2020. Construction and characterization of a knock-down RNA interference line of *OsNRPDI* in rice (*Oryza sativa* ssp *japonica* cv Nipponbare). *Philos Trans R Soc Lond B Biol Sci* **375**: 20190338. doi:10.1098/rstb.2019.0338
- Deleris A, Stroud H, Bernatavichute Y, Johnson E, Klein G, Schubert D, Jacobsen SE. 2012. Loss of the DNA methyltransferase MET1 induces H3K9 hypermethylation at PcG target genes and redistribution of H3K27 trimethylation to transposons in *Arabidopsis thaliana*. *PLoS Genet* **8**: e1003062. doi:10.1371/journal.pgen.1003062
- Délérís A, Berger F, Duhaucourt S. 2021. Role of Polycomb in the control of transposable elements. *Trends Genet* **37**: 882–889. doi:10.1016/j.tig.2021.06.003
- Erdmann RM, Satyaki PRV, Klosinska M, Gehring M. 2017. A small RNA pathway mediates allelic dosage in endosperm. *Cell Rep* **21**: 3364–3372. doi:10.1016/j.celrep.2017.11.078
- Erhard KF, Stonaker JL, Parkinson SE, Lim JP, Hale CJ, Hollick JB. 2009. RNA polymerase IV functions in paramutation in *Zea mays*. *Science* **323**: 1201–1205. doi:10.1126/science.1164508
- Erhard KF Jr, Talbot J-ERB, Deans NC, McClish AE, Hollick JB. 2015. Nascent transcription affected by RNA polymerase IV in *Zea mays*. *Genetics* **199**: 1107–1125. doi:10.1534/genetics.115.174714
- Espinosa NA, Tu LN, Furci L, Shimajiri Y, Harukawa Y, Miura S, Takuno S, Saze H. 2020. Transcriptional regulation of genes bearing intronic heterochromatin in the rice genome. *PLoS Genet* **16**: e1008637. doi:10.1371/journal.pgen.1008637
- Feng W, Michaels SD. 2015. Accessing the inaccessible: the organization, transcription, replication, and repair of heterochromatin in plants.

- Annu Rev Genet* **49**: 439–459. doi:10.1146/annurev-genet-112414-055048
- G VHS, Shivaprasad PV. 2022. Investigation of transposon DNA methylation and copy number variation in plants using Southern hybridisation. *Bio Protoc* **12**: e4432. doi:10.21769/BioProtoc.4432
- Gent JI, Madzima TF, Bader R, Kent MR, Zhang X, Stam M, McGinnis KM, Dawe RK. 2015. Accessible DNA and relative depletion of H3K9me2 at maize loci undergoing RNA-directed DNA methylation. *Plant Cell* **26**: 4903–4917. doi:10.1105/tpc.114.130427
- Gibney ER, Nolan CM. 2010. Epigenetics and gene expression. *Heredity (Edinb)* **105**: 4–13. doi:10.1038/hdy.2010.54
- Grover JW, Kendall T, Baten A, Burgess D, Freeling M, King GJ, Mosher RA. 2018. Maternal components of RNA-directed DNA methylation are required for seed development in *Brassica rapa*. *Plant J* **94**: 575–582. doi:10.1111/tpj.13910
- Hahn S. 2004. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* **11**: 394–403. doi:10.1038/nsmb763
- Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. 2005. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**: 118–120. doi:10.1126/science.1106910
- Hirsch CD, Springer NM. 2017. Transposable element influences on gene expression in plants. *Biochim Biophys Acta Gene Regul Mech* **1860**: 157–165. doi:10.1016/j.bbagr.2016.05.010
- Hu J, Wang K, Huang W, Liu G, Gao Y, Wang J, Huang Q, Ji Y, Qin X, Wan L, et al. 2012. The rice pentatricopeptide repeat protein RF5 restores fertility in Hong-Lian cytoplasmic male-sterile lines via a complex with the glycine-rich protein GRP162. *Plant Cell* **24**: 109–122. doi:10.1105/tpc.111.093211
- Ishihara S, Sasagawa Y, Kameda T, Yamashita H, Umeda M, Kotomura N, Abe M, Shimono Y, Nikaido I. 2021. Local states of chromatin compaction at transcription start sites control transcription levels. *Nucleic Acids Res* **49**: 8007–8023. doi:10.1093/nar/gkab587
- Jamge B, Lorković Z, Axelsson E, Yelagandula R, Akimcheva S, Berger F. 2022. Transcriptional activity is shaped by the chromatin landscapes in *Arabidopsis*. bioRxiv doi:10.1101/2022.06.02.494419
- Jiang H, Moreno-Romero J, Santos-González J, De Jaeger G, Gevaert K, Van De Slijke E, Köhler C. 2017. Ectopic application of the repressive histone modification H3K9me2 establishes post-zygotic reproductive isolation in *Arabidopsis thaliana*. *Genes Dev* **31**: 1272–1287. doi:10.1101/gad.299347.117
- Johannes F, Schmitz RJ. 2019. Spontaneous epimutations in plants. *New Phytol* **221**: 1253–1259. doi:10.1111/nph.15434
- Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Carrington JC. 2007. Genome-wide profiling and analysis of *Arabidopsis* siRNAs. *PLoS Biol* **5**: e57. doi:10.1371/journal.pbio.0050057
- Köhler C, Hennig L, Spillane C, Pien S, Gruissem W, Grossniklaus U. 2003. The Polycomb group protein MEDEA regulates seed development by controlling expression of the MADS-box gene *PHERES1*. *Genes Dev* **17**: 1540–1553. doi:10.1101/gad.257403
- Kradolfer D, Hennig L, Köhler C. 2013. Increased maternal genome dosage bypasses the requirement of the FIS Polycomb repressive complex 2 in *Arabidopsis* seed development. *PLoS Genet* **9**: e1003163. doi:10.1371/journal.pgen.1003163
- Lanciano S, Carpentier M-C, Llauro C, Jobet E, Robakowska-Hyzorek D, Lasserre E, Ghesquière A, Panaud O, Mirozou M. 2017. Sequencing the extrachromosomal circular mobilome reveals retrotransposon activity in plants. *PLoS Genet* **13**: e1006630. doi:10.1371/journal.pgen.1006630
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359. doi:10.1038/nmeth.1923
- Law JA, Jacobsen SE. 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* **11**: 204–220. doi:10.1038/nrg2719
- Law JA, Ausin I, Johnson LM, Vashisht AA, Zhu J-K, Wohlschlegel JA, Jacobsen SE. 2010. A protein complex required for polymerase V transcripts and RNA-directed DNA methylation in *Arabidopsis*. *Curr Biol* **20**: 951–956. doi:10.1016/j.cub.2010.03.062
- Law JA, Du J, Hale CJ, Feng S, Krajewski K, Palanca AMS, Strahl BD, Patel DJ, Jacobsen SE. 2013. Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* **498**: 385–389. doi:10.1038/nature12178
- Le NT, Harukawa Y, Miura S, Boer D, Kawabe A, Saze H. 2020. Epigenetic regulation of spurious transcription initiation in *Arabidopsis*. *Nat Commun* **11**: 3224. doi:10.1038/s41467-020-16951-w
- Li X, Harris CJ, Zhong Z, Chen W, Liu R, Jia B, Wang Z, Li S, Jacobsen SE, Du J. 2018. Mechanistic insights into plant SUVH family H3K9 methyltransferases and their binding to context-biased non-CG DNA methylation. *Proc Natl Acad Sci* **115**: E8793–E8802. doi:10.1073/pnas.1809841115
- Li J, Yang D-L, Huang H, Zhang G, He L, Pang J, Lozano-Durán R, Lang Z, Zhu J-K. 2020. Epigenetic memory marks determine epiallele stability at loci targeted by de novo DNA methylation. *Nat Plants* **6**: 661–674. doi:10.1038/s41477-020-0671-x
- Liu Z-W, Shao C-R, Zhang C-J, Zhou J-X, Zhang S-W, Li L, Chen S, Huang H-W, Cai T, He X-J. 2014. The SET domain proteins SUVH2 and SUVH9 are required for Pol V occupancy at RNA-directed DNA methylation loci. *PLoS Genet* **10**: e1003948. doi:10.1371/journal.pgen.1003948
- Liu K, Yu Y, Dong A, Shen W-H. 2017. *SET DOMAIN GROUP701* encodes a H3K4-methyltransferase and regulates multiple key processes of rice plant development. *New Phytol* **215**: 609–623. doi:10.1111/nph.14596
- Lu C, Kulkarni K, Souret FF, MuthuValliappan R, Tej SS, Poethig RS, Henderson IR, Jacobsen SE, Wang W, Green PJ, et al. 2006. MicroRNAs and other small RNAs enriched in the *Arabidopsis* RNA-dependent RNA polymerase-2 mutant. *Genome Res* **16**: 1276–1288. doi:10.1101/gr.5530106
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**: 10. doi:10.14806/ej.17.1.200
- Martinez G, Wolff P, Wang Z, Moreno-Romero J, Santos-González J, Conze LL, DeFraia C, Slotkin RK, Köhler C. 2018. Paternal easiRNAs regulate parental genome dosage in *Arabidopsis*. *Nat Genet* **50**: 193–198. doi:10.1038/s41588-017-0033-4
- Mathieu O, Probst AV, Paszkowski J. 2005. Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in *Arabidopsis*. *EMBO J* **24**: 2783–2791. doi:10.1038/sj.emboj.7600743
- Matzke MA, Mosher RA. 2014. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet* **15**: 394–408. doi:10.1038/nrg3683
- McKinlay A, Podicheti R, Wendte JM, Cocklin R, Rusch DB. 2018. RNA polymerases IV and V influence the 3' boundaries of polymerase II transcription units in *Arabidopsis*. *RNA Biol* **15**: 269–279. doi:10.1080/15476286.2017.1409930
- Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, et al. 2008. Sorting of small RNAs into *Arabidopsis* Argonaute complexes is directed by the 5' terminal nucleotide. *Cell* **133**: 116–127. doi:10.1016/j.cell.2008.02.034
- Montgomery SA, Tanizawa Y, Galik B, Wang N, Ito T, Mochizuki T, Akimcheva S, Bowman JL, Cognat V, Maréchal-Drouard L, et al. 2020. Chromatin organization in early land plants reveals an ancestral association between H3K27me3, transposons, and constitutive heterochromatin. *Curr Biol* **30**: 573–588.e7. doi:10.1016/j.cub.2019.12.015
- Moritoh S, Eun C-H, Ono A, Asao H, Okano Y, Yamaguchi K, Shimatani Z, Koizumi A, Terada R. 2012. Targeted disruption of an orthologue of *DOMAINS REARRANGED METHYLASE 2*, *OsDRM2*, impairs the growth of rice plants by abnormal DNA methylation. *Plant J* **71**: 85–98. doi:10.1111/j.1365-313X.2012.04974.x
- Narjala A, Nair A, Tirumalai V, Hari Sundar GV, Shivaprasad PV. 2020. A conserved sequence signature is essential for robust plant miRNA biogenesis. *Nucleic Acids Res* **48**: 3103–3118. doi:10.1093/nar/gkaa077
- Nuthikattu S, McCue AD, Panda K, Fultz D, DeFraia C, Thomas EN, Slotkin RK. 2013. The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21–22 nucleotide small interfering RNAs. *Plant Physiol* **162**: 116–131. doi:10.1104/pp.113.216481
- Ossowski S, Schwab R, Weigel D. 2008. Gene silencing in plants using artificial microRNAs and other small RNAs. *Plant J* **53**: 674–690. doi:10.1111/j.1365-313X.2007.03328.x
- Panda K, Ji L, Neumann DA, Daron J, Schmitz RJ, Slotkin RK. 2016. Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. *Genome Biol* **17**: 170. doi:10.1186/s13059-016-1032-y
- Panda K, McCue AD, Slotkin RK. 2020. *Arabidopsis* RNA polymerase IV generates 21–22 nucleotide small RNAs that can participate in RNA-directed DNA methylation and may regulate genes. *Philos Trans R Soc Lond B Biol Sci* **375**: 20190417. doi:10.1098/rstb.2019.0417
- Parent J-S, Cahn J, Herridge RP, Grimanelli D, Martienssen RA. 2021. Small RNAs guide histone methylation in *Arabidopsis* embryos. *Genes Dev* **35**: 841–846. doi:10.1101/gad.343871.120
- Pellicer J, Hidalgo O, Dodsworth S, Leitch IJ. 2018. Genome size diversity and its impact on the evolution of land plants. *Genes (Basel)* **9**: 88. doi:10.3390/genes9020088
- Ramanathan V, Veluthambi K. 1995. Transfer of non-T-DNA portions of the *Agrobacterium tumefaciens* Ti plasmid pTiA6 from the left terminus of TL-DNA. *Plant Mol Biol* **28**: 1149–1154. doi:10.1007/BF00032676
- Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**: W187–W191. doi:10.1093/nar/gku365
- R Core Team. 2021. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>.
- Rozsak P, Köhler C. 2011. Polycomb group proteins are required to couple seed coat initiation to fertilization. *Proc Natl Acad Sci* **108**: 20826–20831. doi:10.1073/pnas.1117111108

- Rougée M, Quadrana L, Zervudacki J, Hure V, Colot V, Navarro L, Deleris A. 2021. Polycomb mutant partially suppresses DNA hypomethylation-associated phenotypes in *Arabidopsis*. *Life Sci Alliance* **4**: e202000848. doi:10.26508/lsa.202000848
- Saleh A, Alvarez-Venegas R, Avramova Z. 2008. An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in *Arabidopsis* plants. *Nat Protoc* **3**: 1018–1025. doi:10.1038/nprot.2008.66
- Satyaki PRV, Gehring M. 2019. Paternally-acting canonical RNA-directed DNA methylation pathway genes sensitize *Arabidopsis* endosperm to paternal genome dosage. *Plant Cell* **31**: 1563–1578. doi:10.1105/tpc.19.00047
- Schier AC, Taatjes DJ. 2020. Structure and mechanism of the RNA polymerase II transcription machinery. *Genes Dev* **34**: 465–488. doi:10.1101/gad.335679.119
- Shivaprasad PV, Chen H-M, Patel K, Bond DM, Santos BACM, Baulcombe DC. 2012. A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. *Plant Cell* **24**: 859–874. doi:10.1105/tpc.111.095380
- Song L, Koga Y, Ecker JR. 2016. Profiling of transcription factor binding events by chromatin immunoprecipitation sequencing (ChIP-seq). *Curr Protoc Plant Biol* **1**: 293–306. doi:10.1002/cppb.20014
- Soppe WJJ, Jasencakova Z, Houben A, Kakutani T, Meister A, Huang MS, Jacobsen SE, Schubert I, Fransz PF. 2002. DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in *Arabidopsis*. *EMBO J* **21**: 6549–6559. doi:10.1093/emboj/cdf657
- Stocks MB, Mohorianu I, Beckers M, Paicu C, Moxon S, Thody J, Dalmay T, Moulton V. 2018. The UEA sRNA Workbench (version 4.4): a comprehensive suite of tools for analyzing miRNAs and sRNAs. *Bioinformatics* **34**: 3382–3384. doi:10.1093/bioinformatics/bty338
- Swetha C, Basu D, Pachamuthu K, Tirumalai V, Nair A, Prasad M, Shivaprasad PV. 2018. Major domestication-related phenotypes in *indica* rice are due to loss of miRNA-mediated laccase silencing. *Plant Cell* **30**: 2649–2662. doi:10.1105/tpc.18.00472
- Tirumalai V, Swetha C, Nair A, Pandit A, Shivaprasad PV. 2019. mir828 and miR858 regulate VvMYB114 to promote anthocyanin and flavonol accumulation in grapes. *J Exp Bot* **70**: 4775–4792. doi:10.1093/jxb/erz264
- Tirumalai V, Prasad M, Shivaprasad PV. 2020. RNA blot analysis for the detection and quantification of plant microRNAs. *J Vis Exp* e61394. doi:10.3791/61394
- Trujillo JT, Seetharam AS, Hufford MB, Beilstein MA, Mosher RA. 2018. Evidence for a unique DNA-dependent RNA polymerase in cereal crops. *Mol Biol Evol* **35**: 2454–2462. doi:10.1093/molbev/msy146
- Wang F, Yano K, Nagamatsu S, Inari-Ikeda M, Koketsu E, Hirano K, Aya K, Matsuoka M. 2020a. Genome-wide expression quantitative trait locus studies facilitate isolation of causal genes controlling panicle structure. *Plant J* **103**: 266–278. doi:10.1111/tbj.14726
- Wang Z, Butel N, Santos-González J, Borges F, Yi J, Martienssen RA, Martinez G, Köhler C. 2020b. Polymerase IV plays a crucial role in pollen development in *Capsella*. *Plant Cell* **32**: 950–966. doi:10.1105/tpc.19.00938
- Wang L, Zheng K, Zeng L, Xu D, Zhu T, Yin Y, Zhan H, Wu Y, Yang D-L. 2022. Reinforcement of CHH methylation through RNA-directed DNA methylation ensures sexual reproduction in rice. *Plant Physiol* **188**: 1189–1209. doi:10.1093/plphys/kiab531
- Warthmann N, Chen H, Ossowski S, Weigel D, Hervé P. 2008. Highly specific gene silencing by artificial miRNAs in rice. *PLoS One* **3**: e1829. doi:10.1371/journal.pone.0001829
- Wei L, Gu L, Song X, Cui X, Lu Z, Zhou M, Wang L, Hu F, Zhai J, Meyers BC, et al. 2014. Dicer-like 3 produces transposable element-associated 24-nt siRNAs that control agricultural traits in rice. *Proc Natl Acad Sci* **111**: 3877–3882. doi:10.1073/pnas.1318131111
- 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. doi:10.1016/j.cell.2008.09.035
- Wu L, Zhang Q, Zhou H, Ni F, Wu X, Qi Y. 2009. Rice microRNA effector complexes and targets. *Plant Cell* **21**: 3421–3435. doi:10.1105/tpc.109.070938
- Xu L, Yuan K, Yuan M, Meng X, Chen M, Wu J, Li J, Qi Y. 2020. Regulation of rice tillering by RNA-directed DNA methylation at miniature inverted-repeat transposable elements. *Mol Plant* **13**: 851–863. doi:10.1016/j.molp.2020.02.009
- Yaari R, Katz A, Domb K, Harris KD, Zemach A, Ohad N. 2019. RdDM-independent de novo and heterochromatin DNA methylation by plant CMT and DDMT3 orthologs. *Nat Commun* **10**: 1613. doi:10.1038/s41467-019-09496-0
- Yamauchi T, Johzuka-Hisatomi Y, Terada R, Nakamura I, Iida S. 2014. The *MET1b* gene encoding a maintenance DNA methyltransferase is indispensable for normal development in rice. *Plant Mol Biol* **85**: 219–232. doi:10.1007/s11103-014-0178-9
- Zhai J, Bischof S, Wang H, Feng S, Lee T-F, Teng C, Chen X, Park SY, Liu L, Gallego-Bartolome J, et al. 2015. A one precursor one siRNA model for Pol IV-dependent siRNA biogenesis. *Cell* **163**: 445–455. doi:10.1016/j.cell.2015.09.032
- Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE. 2007. Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol* **5**: e129. doi:10.1371/journal.pbio.0050129
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137. doi:10.1186/gb-2008-9-9-r137
- Zhang H, Ma Z-Y, Zeng L, Tanaka K, Zhang C-J, Ma J, Bai G, Wang P, Zhang S-W, Liu Z-W, et al. 2013. DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of Pol IV. *Proc Natl Acad Sci* **110**: 8290–8295. doi:10.1073/pnas.1300585110
- Zhang Q, Wang D, Lang Z, He L, Yang L, Zeng L, Li Y, Zhao C, Huang H, Zhang H, et al. 2016. Methylation interactions in *Arabidopsis* hybrids require RNA-directed DNA methylation and are influenced by genetic variation. *Proc Natl Acad Sci* **113**: E4248–E4256. doi:10.1073/pnas.1607851113
- Zhang C, Wei Y, Xu L, Wu K-C, Yang L, Shi C-N, Yang G-Y, Chen D, Yu F-F, Xie Q, et al. 2020. A bunyavirus-inducible ubiquitin ligase targets RNA polymerase IV for degradation during viral pathogenesis in rice. *Mol Plant* **13**: 836–850. doi:10.1016/j.molp.2020.02.010
- Zhao L, Zhou Q, He L, Deng L, Lozano-Duran R, Li G, Zhu J-K. 2022. DNA methylation underpins the epigenomic landscape regulating genome transcription in *Arabidopsis*. *Genome Biol* **23**: 197. doi:10.1186/s13059-022-02768-x
- Zheng B, Wang Z, Li S, Yu B, Liu J-Y, Chen X. 2009. Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. *Genes Dev* **23**: 2850–2860. doi:10.1101/gad.1868009
- Zheng K, Wang L, Zeng L, Xu D, Guo Z, Gao X, Yang D-L. 2021. The effect of RNA polymerase V on 24-nt siRNA accumulation depends on DNA methylation contexts and histone modifications in rice. *Proc Natl Acad Sci* **118**: e2100709118. doi:10.1073/pnas.2100709118
- Zhong Z, Feng S, Duttke SH, Potok ME, Zhang Y, Gallego-Bartolomé J, Liu W, Jacobsen SE. 2021. DNA methylation-linked chromatin accessibility affects genomic architecture in *Arabidopsis*. *Proc Natl Acad Sci* **118**: e2023347118. doi:10.1073/pnas.2023347118
- Zhou S, Liu X, Zhou C, Zhou Q, Zhao Y, Li G, Zhou D-X. 2016. Cooperation between the H3K27me3 chromatin mark and non-CG methylation in epigenetic regulation. *Plant Physiol* **172**: 1131–1141. doi:10.1104/pp.16.01238
- Zhou M, Palanca AMS, Law JA. 2018. Locus-specific control of the de novo DNA methylation pathway in *Arabidopsis* by the CLASSY family. *Nat Genet* **50**: 865–873. doi:10.1038/s41588-018-0115-y
- Zilberman D, Cao X, Jacobsen SE. 2003. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* **299**: 716–719. doi:10.1126/science.1079695

Received September 22, 2022; accepted in revised form April 16, 2023.