Transgenic plant generation and segregation analysis. *Agrobacterium tumefaciens* strain LBA4404 with a helper plasmid pSB1 (containing extra copies of virulence genes) harbouring the binary vector of interest was used for infecting embryogenic calli derived from scutellum. Transformation of PB1 calli was performed as per previous methods (Hiei et al. 1994; Sridevi et al. 2003). About 21 days (d) old scutellar calli maintained on the callus induction media (CIM) (Murashige-Skoog (MS) medium, 2.5 mg/L 2,4-D, 300 mg/L casein hydrolysate, 500 mg/L L-proline, pH 5.8) was incubated for 4 d after cutting and infected with *Agrobacterium* culture of OD_{600nm} around 1.0. The calli were co-cultivated with *Agrobacterium* in dark and washed with liquid CIM and then incubated in selection medium (CIM with 50 mg/L hygromycin or 5 mg/L phosphinothricin and 250 mg/L cefotaxime) for two rounds of 21 days each. The selected calli were moved to regeneration medium (MS medium, 3 mg/L kinetin, 2 mg/L NAA) in light. The regenerated shoots were moved to half-strength MS medium for rooting before hardening in growth chamber and then transferred to a transgene-compliant green house.

For segregation to subsequent generations, de-husked seeds were surface sterilised using ethanol, bleach and 0.1% mercuric chloride and placed on half-strength MS medium with 50mg/L hygromycin.

Plant phenotyping. WT and *kd* plants of different generations were grown uniformly after genotyping for the transgene. Tillers were counted on a per plant basis and the grain filling rate was calculated as number of filled grains per tiller.

Pollen staining and imaging was done as described in (Pedersen et al. 2004). Equally developed, pre dehisced anthers were collected and stained as described. Three biological replicates were chosen per transgenic line (2 transgenic lines - 6 plants) and multiple images were taken post staining under light microscope. The obtained images were analysed using custom ImageJ scripts for detecting the pollen staining and size of the pollen. Staining index was calculated as percentage of negative stain normalized to the background.

Reverse Transcriptase – quantitative Polymerase Chain Reaction (RT-qPCR). Total RNA isolated from the respective plant tissues using TRIzol method was converted to cDNA using reverse transcriptase kit (Thermo Fisher Scientific RevertAid RT kit) following manufacturer's protocol after treating with DNaseA. The cDNA template was used for the qPCRs using SYBR green qPCR master mix (Solis Biodyne - 5x HOT Firepol Evagreen qPCR Master Mix). Rice *Actin* (LOC_Os03g50885.1) was used as an internal control. The primers used for quantification are listed in Supplemental Table S3.

mRNA northern hybridisation. Total mRNA northern hybridisation was performed as described previously (Shivaprasad et al. 2006). Around 20ug of total RNA was ethanol precipitated and resuspended in denaturing loading buffer. RNA was electrophoresed in 1.5% denaturing agarose gel with 1% formaldehyde in MOPS buffer. The gel was blotted onto Hybond N+ membrane (GE healthcare) by capillary elution and UV crosslinked. The crosslinked membrane was hybridised using rRNA oligo probes described previously (Hang et al. 2018) end labelled with [γ-P32]-ATP in ultrahyb-buffer (Invitrogen). The blots are hybridised at 42^{o}C, washed, exposed to phosphor screen (GE healthcare) and scanned using Typhoon scanner. After scanning stripping and re-probing was performed as done for sRNA northern blots. The membrane was stained with methylene blue solution and shown as loading control.

Extrachromosomal Circular DNA (ECC DNA) PCR. Profiling of extrachromosomal fragments were performed as described previously (Lanciano et al. 2017). The primers were designed to amplify either only the circular DNA products or linear products. UBCE2 gene was used as loading control. To further validate the results, the total DNA (100ng) was digested with PlasmidSafe DNase (Lucigen) to eliminate the linear DNA and the same PCRs were repeated again. The primers used are listed in Supplemental Table S3.

Scanning electron microscopy. SEM imaging was performed as described previously (Das et al. 2020). Equally grown WT and *kd* spikelets were collected before anthesis and fixed in 0.1 M sodium phosphate buffer (pH 7.0) with 3% paraformaldehyde and 0.25%

glutaraldehyde. After incubation at 4°C for 24 h,spikelets were rinsed with 0.1 M sodium phosphate buffer and serially dehydrated with ethanol (30% to 100%). Samples were further dehydrated using critical point drying (LEICA EM CPD300), gold coated and the images were acquired using a Carl ZEISS scanning electron microscope at an accelerating voltage of 2 kV.

Micro computed tomography (CT) scan. Micro CT scanning of developing florets was performed as described earlier (Kannan et al. 2021). Florets from panicles that were equally developed before anthesis was scanned on Bruker Skyscan-1272 (Kontich, Belgium) at 40 kV, 250 µA, 4 µm image pixel size, without filters. The obtained image series was merged into a montage after colour adjustments using ImageJ.

ChIP differential enrichment analyses. ChIP differential enrichment analyses for the histone marks was performed using the tool Diffreps (Shen et al. 2013). The differentially enriched loci were identified using default parameters and the p-value cut-off of 0.001. The obtained loci were converted to bedfiles and the coverage of the ChIP signals normalized to H3 ChIP signal was obtained using BEDTools multicov followed by library normalisation. The results were plotted using custom R scripts. Chromosome resolved heatmaps are generated using shinyChromosome (Yu et al. 2019).

Nuclei immunostaining. Immunofluorescence labelling and microscopy was performed as described earlier (Yelagandula et al. 2014) with the following modifications. Freshly collected tissues were fixed in 4% PFA in PBS for 20 mins and quenched with PBS containing 0.125M Glycine for 20 mins. The tissues are chopped in lysis buffer as described for 2-3min with a scalpel and pelleted. The pelleted nuclei are washed in lysis buffer and resuspended in N buffer as described. The nuclei are dried on a positively charged slide. The dried nuclei are fixed with 4% PFA for 30 min and washed with PBS. The washed nuclei are blocked with blocking solution (1xPBS with 0.1% Triton-X 100, 1%BSA, 10% normal goat serum and 100mM glycine) for 1 hour at room temperature. Primary antibody (H3K27me3 - Acive Motif 39155 [1:300], H3K9me2 – Abcam ab1220 [1:100]) incubations were done at 4 degrees for

16 hours in antibody binding solution (1xPBS with 0.1%BSA, 1% normal goat serum, 0.1% Triton-X 100 and 100mM glycine). The nuclei are washed with 1xPBS and incubated with appropriate secondary antibodies (Invitrogen; goat raised Anti-mouse 488 and Anti-Rabbit 555) in antibody binding solution at 1:1000 dilution in dark at RT. Post washing, the nuclei are DAPI stained, mounted with fluoromount anti-fade reagent, and imaged on a Confocal microscope (Olympus FV3000). The images are analysed using ImageJ for ROI analyses.

DNA methylome analyses. Total DNA was isolated using CTAB method from the tissues mentioned. The library was constructed using the NEBNext® Enzymatic Methyl-seq Kit (Catalog no-E7120) using manufacturer's instructions and protocols described previously (Feng et al. 2020). The libraries were sequenced in a paired end mode (100bp) on a Hiseq2500 platform.

The obtained reads were quality checked and trimmed using cutadapt (Martin 2011) followed by alignment to IRGSP1.0 genome using Bismark aligner tool with default parameters (Krueger and Andrews 2011). DNA methylation status was extracted and coverage reports were generated using the Bismark tools. The obtained results are analysed using methylation package ViewBS (Huang et al. 2018). The chloroplast and mitochondrial genomes' cytosine conversion rate was used as a quality control. ViewBS tools were used for estimating DNA methylation in different contexts over the defined regions.

For *Arabidopsis* datasets (Supplemental Table S2), same analyses pipeline was used except that the reads are aligned to the TAIR10 genome.

sRNA differential expression analyses. The filtered sRNAs, size classified into 21-22nt and 23-24nt, were aligned and sRNA loci identified using Shortstack (Axtell 2013) with following parameters: --nohp --mmap f --mismatches 1 --mincov 0.2 rpmm. The raw counts obtained in each cluster was used for differential expression analyses using DESeq (Love et al. 2014). Only the clusters with p-value less than 0.05 and absolute $\log_2$ abundance fold change more than 1 were called as differential sRNA clusters (Supplemental Table S5).

Volcano plots generated using custom R scripts. The clusters identified were further subcategorized into predominant 21-22nt clusters (if the 21-22nt population in a particular cluster exceeds 65% of total small RNAs), predominant 23-24nt (if 23-24nt sRNAs exceed 65%) and mixed clusters (all the other DE clusters).

Genome window-based sRNA analyses and plotting. The genome was split into 100bp non-overlapping windows using BEDTools makewindows (Quinlan and Hall 2010) and the sRNAs (after dividing into 21-22nt and 23-24nt) were counted in each of the bins in the respective sized alignment files using BEDTools multicov . The raw counts were normalized to RPM by considering the total reads mapped in each size class. All the bins that accumulate non-zero RPM values in both genotypes and at least minimum of 2 RPM when summed were taken for differential analysis. 2-fold change in RPM and absolute difference in RPM as minimum 5 were considered as the cut-off for differential calling for a bin (up-regulated or downregulated in *kd* w.r.t. WT) (Supplemental Table S6). The bins identified from different tissues were taken for further analyses of mapping to genomic feature (overlap with the annotation), 5'-nucleotide bias of the sRNAs mapping to them, circos plots and the cumulative sum plots.

ChIPseekeR (Yu et al. 2015) was used for plotting the Vennpie diagrams of annotations of the suppressed and dependent sRNA bins after merging the different size category bins. Distal intergenic regions were the ones beyond 3 kb of any annotated PCGs and intergenic were the ones that are within this window. Circos plots displaying the heatmaps of the sRNA abundance of size-categorized bins were plotted using shinyCircos (Yu et al. 2018). The Upset plots were created in R using Intervene (Khan and Mathelier 2017).

Arabidopsis sRNA analysis was also performed the same way with the same thresholds for differential bin calling and associated analyses as handled for rice bins. The differential expressed sRNA bins were calculated with the same cut-off as used for rice when comparing WT and *nrpd1* sRNAs in different tissues (Supplemental Table S7). The chromosome-wide heatmaps of the sRNA bins were plotted using shinyChromosome (Yu et

al. 2019). All the *Arabidopsis* datasets were processed the same way as rice and the BEDTools multicov was used to estimate the sRNA abundance in each of the up-regulated and downregulated sRNA bins. The data was plotted as a boxplot using ggplot in R. Statistical analyses are performed using R.

Degradome analysis. The obtained reads were processed for adapter removal and size filtering from 18 to 21 nt using UEA small RNA workbench (Stocks et al. 2018). CleaveLand version 4 (Addo-Quaye and Axtell 2008) with default parameters was used to identify the degradome validated target genes. Three fasta files were given as input to CleaveLand, sRNA from the Pol IV-suppressed loci, genes overlapping with Pol IV-suppressed loci and degradome reads. The cut-offs of Allen score 8 and mfe ratio of 0.65 were used to obtain the valid gene targeting list. The list of targeted genes is in Supplemental Table S8. The target sites are mapped onto the target transcripts and deepTools was used to generate the metaplots as described earlier centering at the slicing site.

AGO-IP data analyses. AGO IP datasets of various AGOs from rice and *Arabidopsis* were processed in the same way as the sRNA datasets, mapping to their corresponding genomes. The IP sRNAs mapping to the up-regulated and downregulated bins/clusters were counted using BEDTools multicov and box-plots were plotted using custom R scripts.

RNA-seq and analyses. RNA-seq was performed in pre-dehisced anthers and pre-emerged panicle tissues. The total RNA was extracted using TRIzol method and was poly(A) enriched before library preparation. Library preparation was done with NEBNext® Ultra™ II Directional RNA Library Prep kit (E7765L) as per manufacturer's instructions. The obtained libraries were sequenced in paired end mode (100bp) on a Illumina Hiseq2500 platform.

The obtained reads were adapter trimmed using Trimmomatic (Bolger et al. 2014) and rRNA depletion was done using SortmeRNA (Kopylova et al. 2012). The reads are mapped to IRGSP1.0 genome using HISAT2 (Kim et al. 2015) with default parameters. Cufflinks (Trapnell et al. 2012) was used to perform differential gene expression analyses and

statistical testing. The volcano plots were generated for DEGs using custom R scripts with the p-value cut-off of less than 0.05 and absolute $\log_2$ (fold change) expression cut-off of more than 2. For quantifying the expression of genes and transposons, BEDTools multicov (Quinlan and Hall 2010) was used to obtain raw abundance and then normalized to RPKM values. These values are plotted as box-plots using custom R scripts.

References

Axtell MJ. 2013. ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA* **19**: 740–751.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bolte & F SP. A guided tour into subcellular colocalization analysis in light microscopy. *Journal of Microscopy* **224**: 213–232.

Das S, Swetha C, Pachamuthu K, Nair A, Shivaprasad PV. 2020. Loss of function of Oryza sativa Argonaute 18 induces male sterility and reduction in phased small RNAs. *Plant Reprod* **33**: 59–73.

Feng S, Zhong Z, Wang M, Jacobsen SE. 2020. Efficient and accurate determination of genome-wide DNA methylation patterns in *Arabidopsis* thaliana with enzymatic methyl sequencing. *Epigenetics Chromatin* **13**: 42.

Hang R, Wang Z, Deng X, Liu C, Yan B, Yang C, Song X, Mo B, Cao X. 2018. Ribosomal RNA biogenesis and its response to chilling stress in Oryza sativa. *Plant Physiol* **177**: 381–397.

Hiei Y, Ohta S, Komari T, Kumashiro T. 1994. Efficient transformation of rice (Oryza sativa L.) mediated by Agrobacterium and sequence analysis of the boundaries of the T-DNA. *Plant J* **6**: 271–282.

Huang X, Zhang S, Li K, Thimmapuram J, Xie S. 2018. ViewBS: a powerful toolkit for visualization of high-throughput bisulfite sequencing data. *Bioinformatics* **34**: 708–709.

Kannan P, Chongloi GL, Majhi BB, Basu D, Veluthambi K, Vijayraghavan U. 2021. Characterization of a new rice OsMADS1 null mutant generated by homologous recombination-mediated gene targeting. *Planta* **253**: 39.

Khan A, Mathelier A. 2017. Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. *BMC Bioinformatics* **18**. http://dx.doi.org/10.1186/s12859-017-1708-7.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**: 357–360.

Kopylova E, Noé L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: 3211–3217.

Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**: 1571–1572.

Lanciano S, Carpentier M-C, Llauro C, Jobet E, Robakowska-Hyzorek D, Lasserre E, Ghesquière A, Panaud O, Mirouze M. 2017. Sequencing the extrachromosomal circular mobilome reveals retrotransposon activity in plants. *PLoS Genet* **13**: e1006630.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**: 10.

Pedersen JF, Bean SR, Funnell DL, Graybosch RA. 2004. Rapid iodine staining techniques for identifying the waxy phenotype in sorghum grain and waxy genotype in sorghum pollen. *Crop Sci* **44**: 764–767.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.

Shen L, Shao N-Y, Liu X, Maze I, Feng J, Nestler EJ. 2013. diffReps: detecting differential chromatin modification sites from ChIP-seq data with biological replicates. *PLoS One* **8**: e65598.

Shivaprasad PV, Thillaichidambaram P, Balaji V, Veluthambi K. 2006. Expression of full-length and truncated Rep genes from Mungbean yellow mosaic virus-Vigna inhibits viral replication in transgenic tobacco. *Virus Genes* **33**: 365–374.

Sridevi G, Sabapathi N, Meena P, Nandakumar R, Samiyappan R, Muthukrishnan S, Veluthambi K. 2003. Transgenic indica rice variety Pusa basmati 1 constitutively expressing a rice chitinase gene exhibits enhanced resistance to Rhizoctonia solani. *J Plant Biochem Biotechnol* **12**: 93–101.

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.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**: 562–578.

Yelagandula R, Stroud H, Holec S, Zhou K, Feng S, Zhong X, Muthurajan UM, Nie X, Kawashima T, Groth M, et al. 2014. The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in *Arabidopsis*. *Cell* **158**: 98–109.

Yu G, Wang L-G, He Q-Y. 2015. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**: 2382–2383.

Yu Y, Ouyang Y, Yao W. 2018. shinyCircos: an R/Shiny application for interactive creation of Circos plot. *Bioinformatics* **34**: 1229–1231.

Yu Y, Yao W, Wang Y, Huang F. 2019. ShinyChromosome: An R/Shiny application for interactive creation of non-circular plots of whole genomes. *Genomics Proteomics Bioinformatics* **17**: 535–539.