The effect of RNA polymerase V on 24-nt siRNA accumulation depends on DNA methylation contexts and histone modifications in rice

Kezhi Zheng[1*], Lili Wang[1*], Longjun Zeng[2*], Dachao Xu[1], Zhongxin Guo[3], Xiquan Gao[1], Dong-Lei Yang[1#]

1. State Key Laboratory for Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, 210095, Nanjing, China
2. Academy for Advanced Interdisciplinary Studies, Nanjing Agricultural University, 210095, Nanjing, China
3. Vector-borne Virus Research Center, Fujian Agriculture and Forestry University, Fujian, China

* These authors contributed equally.

Corresponding author: D-L Yang (dlyang@njau.edu.cn).

This PDF file includes:

Materials and Methods

SI References

Figure S1 to S9

Dataset S1 to S2

SI Materials and Methods

Materials and growth conditions

OsGA2ox1 ectopic expression (GAE) plants, *OsGA2ox1*-silencing (GAS) plants, and *fem3-1* (*osnrpe1-1)* plants were in the background of Taipei 309 (TP309) (*Oryza sativa*, japonica). The *pol iv* mutants were created using CRISPR/Cas9 technology with the pCBSG03 vector (1) in the TP309 background. The *osnrpe1-2*, *osnrpe1-3*, *osnrpf1-1*, *osnrpf1-2*, *osnrpe1-2/osnrpf1-2*, and *osnrpe1-2/osnrpf1-3* mutants were generated using CRISPR/Cas9 technology with the CRISPR/Cas9-MH vector (2) in Nipponbare. The *fem1-6* (*osrdr2-6*) mutant in the Nipponbare background was previously reported (3).

Seedlings grown for 18 days in Kimura B nutrient solution in the growth chamber (16 h/8 h light/dark with a 28℃/25℃ cycle) were used for sRNA sequencing and bisulfite sequencing. For morphological and molecular analyses, rice plants were grown in an isolated paddy field in Nanjing, Jiangsu Province.

Genetic screen of *five elements mountain* mutants

GAS seeds (about 50,000) were soaked in water for 24 h at room temperature and then transferred into 1% ethyl methyl sulfonate (EMS) for another 24 h. The EMS-treated seeds were planted in an isolated paddy field. The seeds were separately harvested from each individual plant (more than 23,000 M1 plants). The M2 plants from the same M1 line were planted together. The M2 plants that produced seedlings with GA deficient symptoms were selected and planted (about 200 M1 lines). The mRNA level of *OsGA2ox1* was measured in 3-month-old plants. Rice plants with *OsGA2ox1* mRNA levels that were 20 times greater than those of GAS were named as *five elements mountain* (*fem*). The name *fem* was inspired by "The journey to the west", in which Buddha created Five Elements Mountain ('FEM') to imprison a monkey king who disturbed the hierarchy in heaven. The monkey king is like the transgenes and transposons that often disturb genomes and result in developmental defects. Five Elements Mountain created by Buddha is analogous to the gene silencing mechanism that represses transgenes and transposons.

Map-based cloning of *FEM3*

Because *fem3-1* was fully sterile, we crossed *fem3-1*/*FEM3* heterozygous plants with TN1 (Taichuang native 1), an indica variety. The dwarf F2 plants were used for mapping. A total of 228 SSR or STS markers that are polymorphic between TP309 and TN1 were used for BSA analysis. During primary mapping, *FEM3* location was narrowed to the region between S11 and M1 using 188 dwarf F2 plants. Subsequent fine-mapping with 978 dwarf plants delimited *FEM3* to M2 and M3, which span ~68 kb. The oligos used in map-based cloning are listed in Dataset S1.

Phylogenetic analysis

Homologous protein sequences of OsNRPE1 (Arabidopsis, maize, and rice) were downloaded from NCBI. The sequences were then aligned using MUSCLE, and the neighbor-joining trees were constructed with 1000 bootstrap replications using MEGA 7.0 software.

RNA isolation and analysis of transcriptional levels

Total RNA was extracted with TRIzol reagent (Invitrogen, 15596018). About 500 ng of total RNA was used for reverse transcription (Vazyme, R123-01) in a 10-μL reaction system. The synthesized complementary DNA (cDNA) was then diluted to 200 μL, 5 μL of which was used for semi-quantitative RT-PCR or qRT-PCR. qRT-PCR was performed on the Bio-Rad CFX 96 PCR system using AceQ qPCR RT SYBR Green Master Mix (Q212-01, Vazyme). The expression levels were calculated using the comparative C_T method relative to *OsUbiquitin*, LOC_Os03g13170 (4). The oligos used in this study are listed in Dataset S1.

To detect Pol V-dependent transcripts, about 4 μg of RNA was treated with DNase I (Vazyme) at 42℃ for 20 min. The DNase I was inactivated at 85℃ for 5 min. The treated RNA was reverse-transcribed using the SuperScript® III First-Strand Synthesis System for RT-PCR (Invitrogen, 18080-051) with random hexamers. The

PCR was performed using TaKaRA Taq (R001A) and the loci-specific oligos listed in Dataset S1.

DNA methylation analyses of individual loci

For Chop-PCR, genomic DNA was extracted from rice leaves using the CTAB method. About 1 μg of genomic DNA was digested in a 20-μL reaction mixture for 10 h with the following DNA methylation-sensitive restriction endonucleases: *Hpa*II (NEB, R0171) for CG methylation; *Msp*I (NEB, R0106) for CHG methylation; and *Alu*I and *Hae*III (NEB, *Alu*I, R0137; *Hae*III, R0108) for CHH methylation. The tested loci were amplified using 1 μL of enzyme-digested DNA as template. The undigested DNA was used to amplify the indicated loci to serve as the control in Fig. 1*D* and *E*, and *SI Appendix* Fig. S1*F*. In Fig. 1*F*, *SI Appendix* Fig. S1*G* and Fig. S2*E*, *OsUbiquintin* (LOC_Os03g13170) was amplified and served as the control using the same amount of undigested DNA as template. The PCR product was separated on 1% agarose gels. The oligos used in Chop-PCR are listed in Dataset S1.

Vector construction and plant transformation

For the complementation, a 16.9 kb fragment of genomic DNA of *OsNRPE1* with a 2-kb promoter and a 102-bp terminator was amplified from TP309 with DNA polymerase (Vazyme, P505-d1). The DNA fragment was cloned into the pCAMBIA3301 vector, which was digested by *EcoR*I and *Hind*III (NEB, R3101, and R3104) using the ClonExpress MultiS One Step Cloning kit (Vazyme, C113-02) according to the manufacturer's instructions. The recombinant vector was then introduced into *Agrobacterium tumefaciens* EHA105. OsNRPE1 was then transformed into the *osnrpe1-1* mutant using the agrobacterium infection method (5).

For genome editing using the CRISPR/Cas9 vector, sgRNAs were designed online (http://skl.scau.edu.cn/). The oligo pairs of sgRNAs were cloned into the pCBSG03 vector (1) or into the CRISPR/Cas9-MH vector (2), which were digested by *Bsa*I (NEB, R0535). The recombinant vectors were introduced into *A. tumefaciens* EHA105. Nipponbare seeds were used for transformation. The oligos used in vector construction are listed in Dataset S1.

Whole-genome bisulfite sequencing and data analysis

Genomic DNA was extracted from 18-day-old seedlings using the standard CTAB method. For construction of bisulfite-seq libraries, 1 μg of DNA was randomly broken into 200 to 300 bp fragments with a Covaris ultrasonicator; the fragmented DNAs were purified by agarose gel electrophoresis with the MiniElute PCR Purification kit (QIAGEN, 28004). The fragmented DNA was then combined, and A-Tailing was added to the 3' end, finally it was ligated with adapters. The libraries were purified with the MiniElute PCR Purification kit. The purified products were subjected to bisulfite treatment and purification with the Methylation-Gold kit (ZYMO, D5005). After agarose gel electrophoresis, fragments ranging in size from 320 to 420 bp were selected. The DNA fragments were then amplified by 15 rounds of PCR, and the fragments ranging from 320 to 420 bp were purified with the QIAquick Gel Extraction kit (QIAGEN, 28704). The qualities of the libraries were tested with the ABI StepOnePlus RealTime PCR System and the Agilent Technologies 2100 bioanalyzer. High throughput sequencing was performed on the Illumina X-Ten platform at the Beijing Genomics Institute (Shenzhen, China).

For the data analysis, the adaptor and low-quality sequences ($q < 20$) were removed, and the clean reads were mapped to the rice genome (RGAP7.0) using Bismark (6) (version v0.18.1) with default parameters. The PCR duplicates were removed by "deduplicate_bismark". The uniquely mapped reads of the chloroplast genome were used to calculate the conversion rate. The average methylation levels of CG, CHG, and CHH on genes and TEs, and in their 3 kb upstream and downstream flanking regions were analyzed in 100 bp windows.

For identification of differentially methylated regions (DMRs), fractional DNA methylation levels of 100 bp sliding windows with 50 bp steps in the genome were first calculated. Windows with at least four effective cytosines were retained. Pairwise comparisons of DNA methylation levels in the WT and mutants were made using Fisher's exact test, and the p-values were adjusted for multiple comparisons using the Benjamini-Hochberg

method (FDR < 0.01). Regions with an absolute methylation difference > 0.7, 0.5, and 0.1 for CG, CHG, and CHH, respectively, were scored as DMRs. Adjacent DMRs were merged if the gap was < 100 bp. All analyses were conducted in R (version 3.1.0). The overlap between DMRs was calculated with bedtools (7) (version v2.25.0).

Small RNA sequencing and analysis

Total RNA was extracted from 18-day-old seedlings using TRIzol reagent (Invitrogen, 15596018). For construction of sRNA-seq libraries, 18 to 45 nt fractions were selected and purified from total RNA using PAGE. After the purified products were ligated with 3' and 5' adaptors, the sRNAs with adapters were reversed transcribed by SuperScript™ III First-Strand Synthesis SuperMix (Invitrogen, 18080400) to generate the cDNA. The cDNA fragments were amplified by 15 rounds of PCR and were then purified by PAGE. The double-stranded PCR products were converted into single-stranded circular DNA (ssCir DNA) by heat denaturation and were circularized with the splint oligo sequence. Finally, the DNA (ssCir DNA) was formatted as the sRNA library. The qualities of sRNA-seq libraries were measured with an Agilent Technologies 2100 bioanalyzer. High throughput sequencing was performed on the BGISEQ-500 platform at the Beijing Genomics Institute (Shenzhen, China).

For the data analysis, the adaptors of the reads were trimmed, and the low-quality reads of the raw data were removed. The remaining clean sRNA reads were mapped to the rice genome (RGAP7.0) using BOWTIE software (8) (version 1.1.2) allowing no mismatches. The structural RNA (tRNAs, rRNAs, snRNAs, and snoRNAs) reads were excluded, and the abundance of siRNA was normalized using the reads per ten million mapped reads (RPTM) based on the uniquely mapped 24-nt reads. For identification of the 24-nt siRNA clusters, the distance of uniquely mapped 24-nt reads ≤100 bp were merged as the candidate clusters by BEDTools (7) (version v2.25.0); the candidate clusters with RPTM value > 12 were defined as the 24-nt siRNA clusters.

Analysis of ChIP-seq

The following data were downloaded from NCBI: H3K4me2 (GSM658110)(9), H3K9me1 (GSE79033), H3K4ac (GSE79033), H3K27me3 (GSE79033), H3K27ac (GSE79033), H3K9ac (GSE79033) (10), H4K16ac (GSE69426) (11), H3K9me2 (GSM2152477) (12), and MNase-seq (SRP045236)(13). For the ChIP-seq and MNase-seq analysis, the reads were mapped to the rice genome (RGAP7.0) using BOWTIE2 software (8) (version 2.3.2) with default parameters. The bam files were generated using samtools, and the expression levels were normalized using RPTM based on the bam file reads using bedtools (7) (version v2.25.0).

Gene expression analysis

For RNA-seq data, low quality reads of the raw data were removed with Trimmomatic (version 0.33), and clean reads were mapped to the rice genome (RGAP7.0) using TopHat-2.1.1 (14). PCR duplicates were removed by samtools rmdup. The expression levels of each gene were normalized into fragments per kilobase of exon per million mapped reads (FPKM).

SI References

1. H. Zhang*, et al.,* The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol. J.* **12**,797-807 (2014).
2. X. Ma*, et al.,* A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Mol. Plant* **8**,1274-1284 (2015).
3. L. Wang*, et al.,* Global reinforcement of DNA methylation through enhancement of RNA-directed DNA methylation ensures sexual reproduction in rice. *bioRxiv*[Preprint](2020). www.biorxiv.org/content/10.1101 (accessed 03 July 2020).
4. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**,402-408 (2001).
5. Y. Hiei, T. Komari, Agrobacterium-mediated transformation of rice using immature embryos or calli

induced from mature seed. *Nat. Protoc.* **3**,824-834 (2008).

6. F. Krueger, S. R. Andrews, Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**,1571-1572 (2011).
7. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**,841-842 (2010).
8. B. Langmead, Aligning short sequencing reads with Bowtie. *Curr. Protoc. Bioinformatics* **Chapter 11**,Unit 11.17 (2010).
9. W. Zhang*, et al.,* High-resolution mapping of open chromatin in the rice genome. *Genome Res.* **22**,151-162 (2012).
10. Y. Fang*, et al.,* Histone modifications facilitate the coexpression of bidirectional promoters in rice. *BMC Genomics* **17**,768 (2016).
11. L. Lu, X. Chen, D. Sanders, S. Qian, X. Zhong, High-resolution mapping of H4K16 and H3K23 acetylation reveals conserved and unique distribution patterns in Arabidopsis and rice. *Epigenetics* **10**,1044-1053 (2015).
12. F. Tan*, et al.,* Analysis of Chromatin Regulators Reveals Specific Features of Rice DNA Methylation Pathways. *Plant Physiol.* **171**,2041-2054 (2016).
13. T. Zhang, W. Zhang, J. Jiang, Genome-Wide Nucleosome Occupancy and Positioning and Their Impact on Gene Expression and Evolution in Plants. *Plant Physiol.* **168**,1406-1416 (2015).
14. C. Trapnell*, et al.,* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**,562-578 (2012).

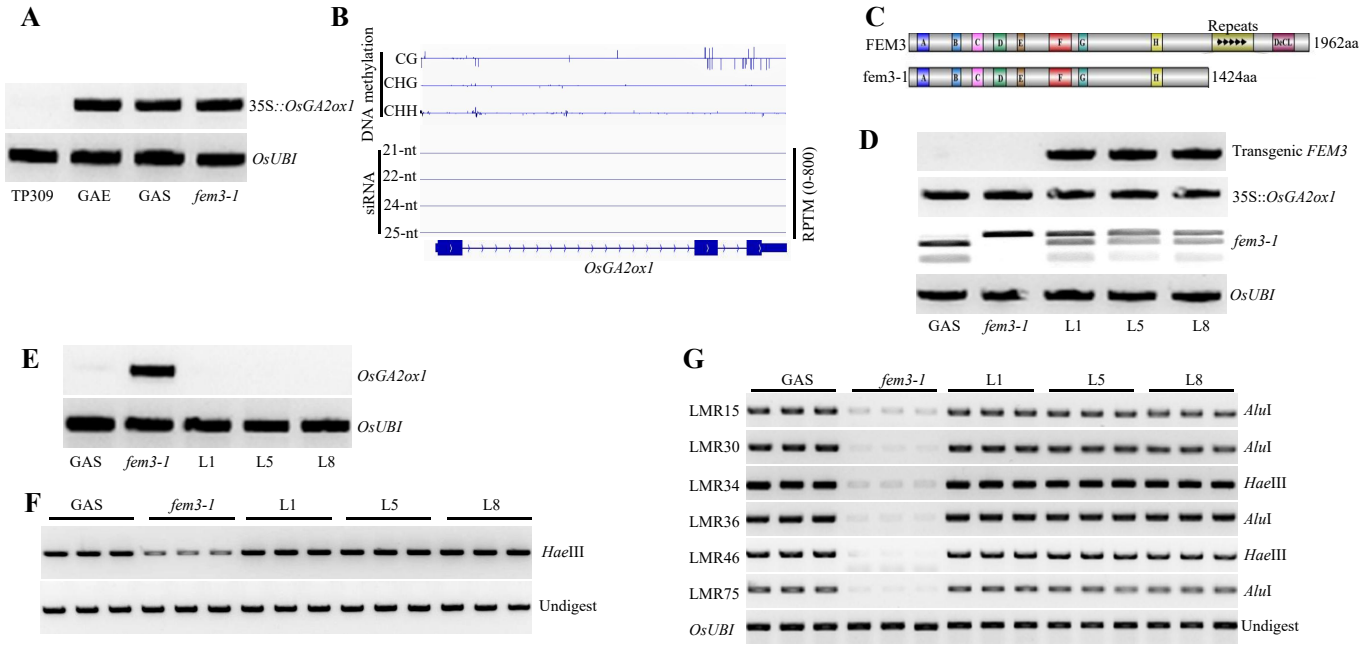

Fig. S1. Genetic complementation of *fem3*.

(*A*) Genotyping of 35S::*OsGA2ox1* in the various genotypes. *OsUbiquitin* (LOC_Os03g13170) served as the control. (*B*) Integrative genome browser image showing the methylation levels and siRNA abundance on *OsGA2ox1* in TP309 WT plants. (*C*) Comparison of domains between FEM3 and fem3-1. (*D*) Genotyping of complemented *FEM3* in GAS, *fem3-1*, and three independent transgenic lines (L1, L5, and L8). 35S::*OsGA2ox1*, *fem3-1*, and *OsUbiquitin* (LOC_Os03g13170) served as controls. (*E*) Transcriptional levels of *OsGA2ox1* in the indicated genotypes as determined by semi-quantitative RT-PCR. *OsUbiquitin* (LOC_Os03g13170) served as the control. (*F*) CHH methylation levels on the 35S promoter in GAS, *fem3-1*, and three independent transgenic lines as determined by Chop-PCR assay. (*G*) DNA methylation levels on six RdDM loci in the indicated genotypes as determined by Chop-PCR assay.

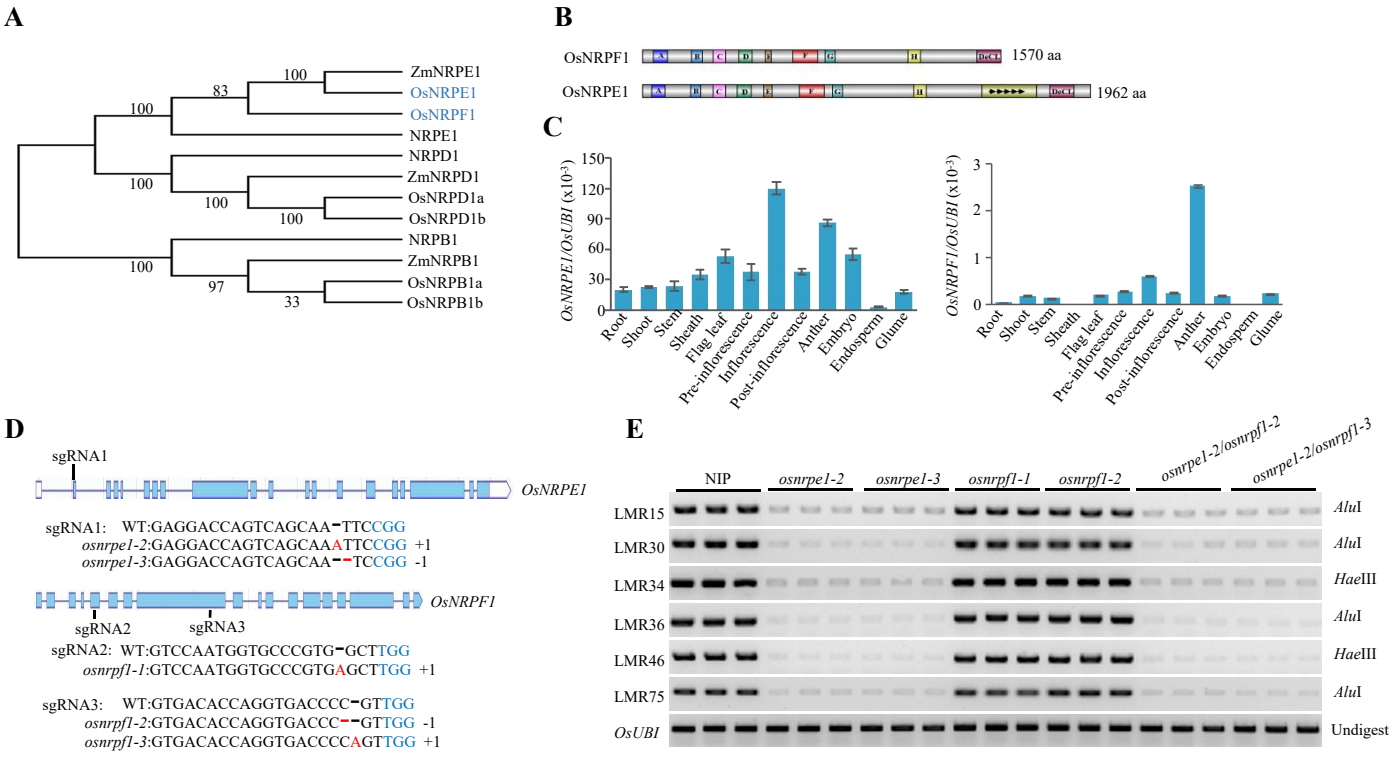

Fig. S2. Functional characterization of OsNRPE1 and OsNRPF1.

(*A*) Phylogenetic analysis of NRPB1, NRPD1, and NRPE1 from Arabidopsis, rice, and maize. The neighbor-joining tree was constructed using MEGA 7.0 software. (*B*) Domain features of OsNRPE1 and OsNRPF1. (*C*) Relative expression levels of *OsNRPE1* and *OsNRPF1* in different tissues as determined by qRT-PCR. Values are means ± SD of three biological replicates. (*D*) Diagram indicating the locations of sgRNAs on *OsNRPE1* and *OsNRPF1*. Mutations caused by CRISPR/Cas9 are indicated in red. (*E*) DNA methylation levels on six RdDM loci in various genotypes as determined by Chop-PCR assay.

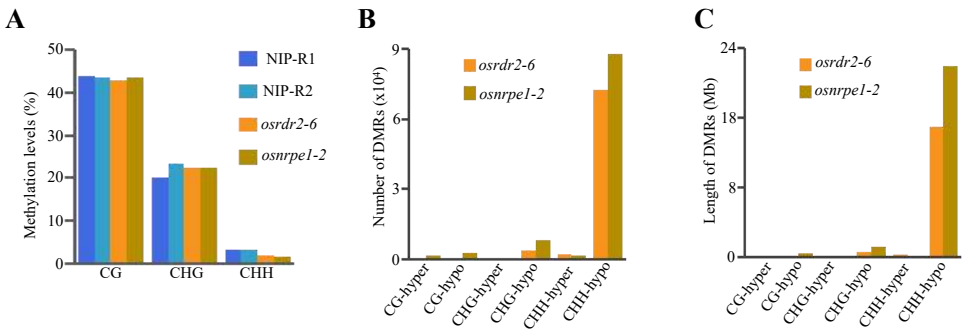

Fig. S3. OsNRPE1 controls genome-wide CHH methylation in Nipponbare.

(*A*) Whole-genome methylation levels of CG, CHG, and CHH in various genotypes. (*B*) DMR number of CG, CHG, and CHH in *osrdr2-6* and *osnrpe1-2*. (*C*) Total length of different kinds of DMRs in *osrdr2-6* and *osnrpe1-2*.

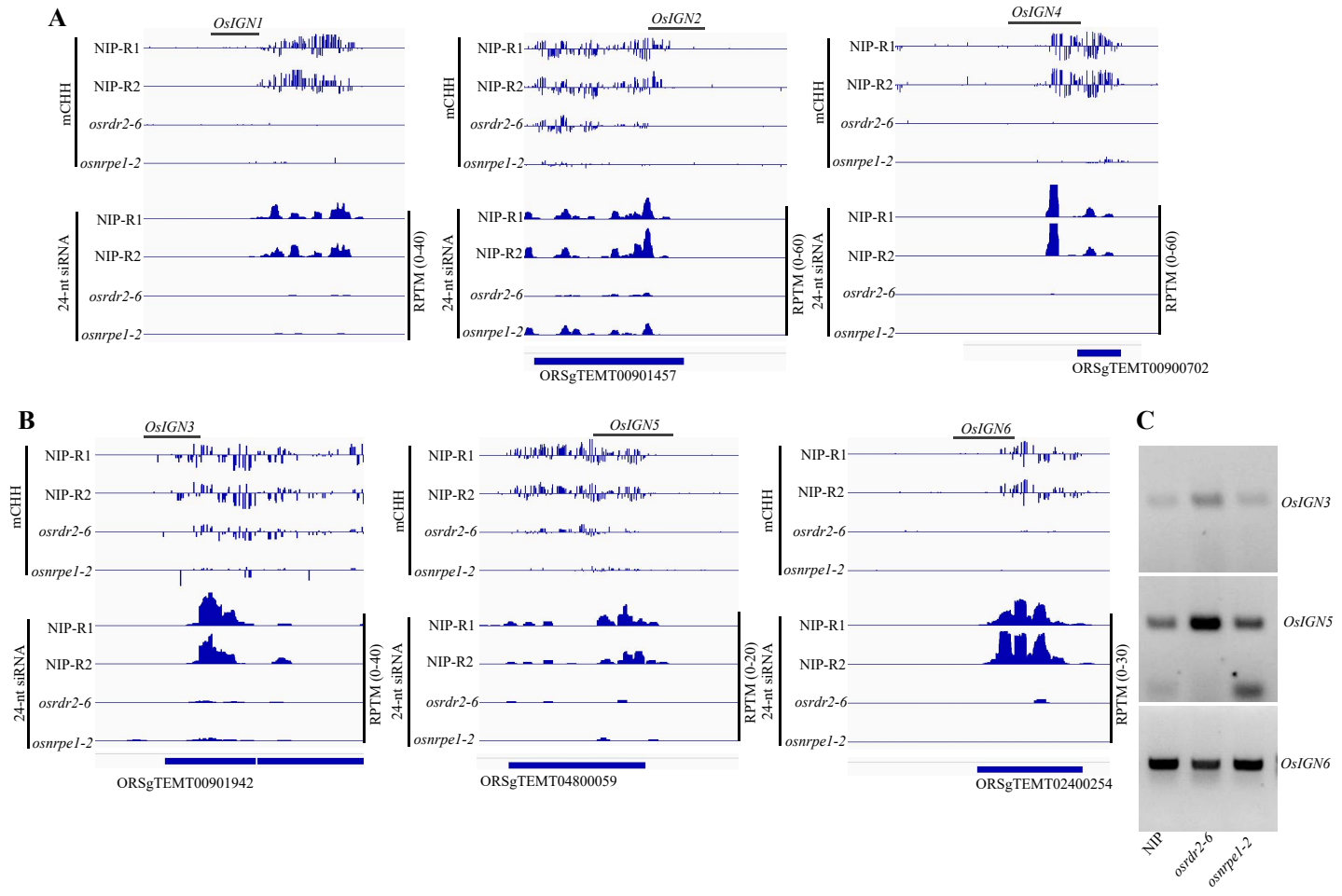

Fig. S4. Detection of intergenic Pol V-dependent transcripts in rice.

(*A*) Integrative genome browser image showing the CHH methylation levels and siRNA abundance on *OsIGN1*, *OsIGN2*, and *OsIGN4.* (*B*) Integrative genome browser image showing the CHH methylation levels and siRNA abundance on *OsIGN3*, *OsIGN5*, and *OsIGN6.* (*C*) RT-PCR analysis of transcripts on *OsIGN3*, *OsIGN5*, and *OsIGN6* in Nipponbare and in *osrdr2-6* and *nrpe1-2* mutants.

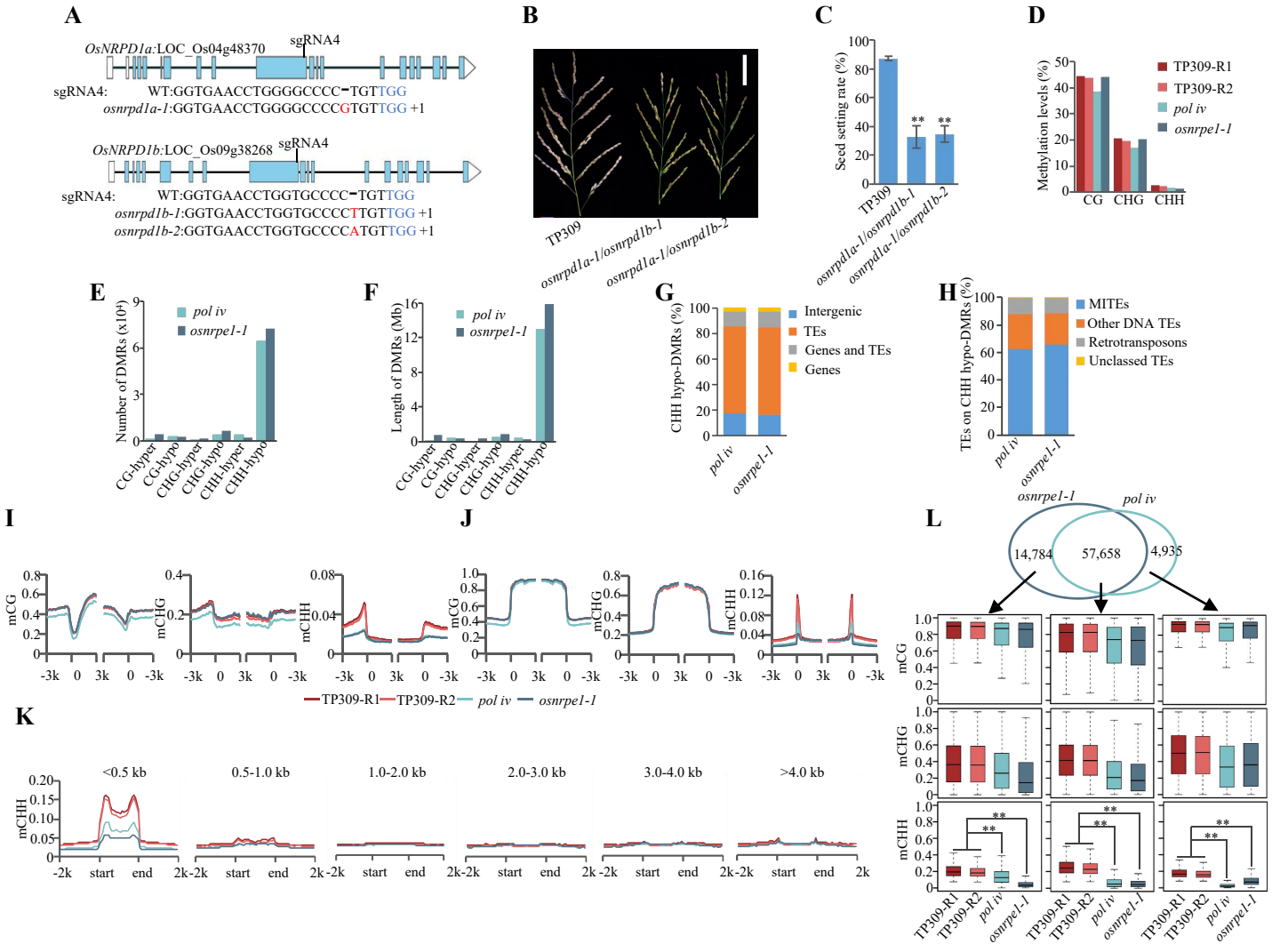

Fig. S5. OsNRPE1 controls genome-wide CHH methylation in TP309.

(*A*) Location of sgRNA4 on *OsNRPD1a* and *OsNRPD1b*. The mutation caused by CRISPR/Cas9 is indicated in red. (*B*) Comparison of the seed setting phenotype in *pol iv* and the WT. Scale bar=5 cm. (*C*) Seed setting rate in *pol iv* and the WT (n=10). Values are means ± SD; ** indicates that $P < 0.01$ according to Student's *t*-test. (*D*) Whole-genome methylation levels of CG, CHG, and CHH in various genotypes. (*E*) DMR numbers of CG, CHG, and CHH in *pol iv* and *osnrpe1-1*. (*F*) Total length of different kinds of DMRs in *pol iv* and *osnrpe1-1*. (*G*) Genomic location of CHH hypo-DMRs in *pol iv* and *osnrpe1-1*. (*H*) TE categories associated with CHH hypo-DMRs in *pol iv* and *osnrpe1-1*. (*I* and *J*) DNA methylation levels of CG, CHG, and CHH on genes (*I*) and on TEs (*J*) in TP309, *pol iv*, and *osnrpe1-1*. The average methylation levels within each 100 bp interval are plotted. (*K*) CHH methylation levels on TEs with different lengths in TP309, *pol iv*, and *osnrpe1-1*. (*L*) Top: Venn diagram showing the overlap of CHH hypo-DMRs in *osnrpe1-1* and *pol iv*. Bottom: DNA methylation levels of CG, CHG, and CHH on *osnrpe1-1*-specific, *pol iv*-specific, and overlapped CHH hypo-DMRs as indicated by box plots. ** indicates that compared plots are significantly different at $P < 0.01$ (Fisher's LSD).

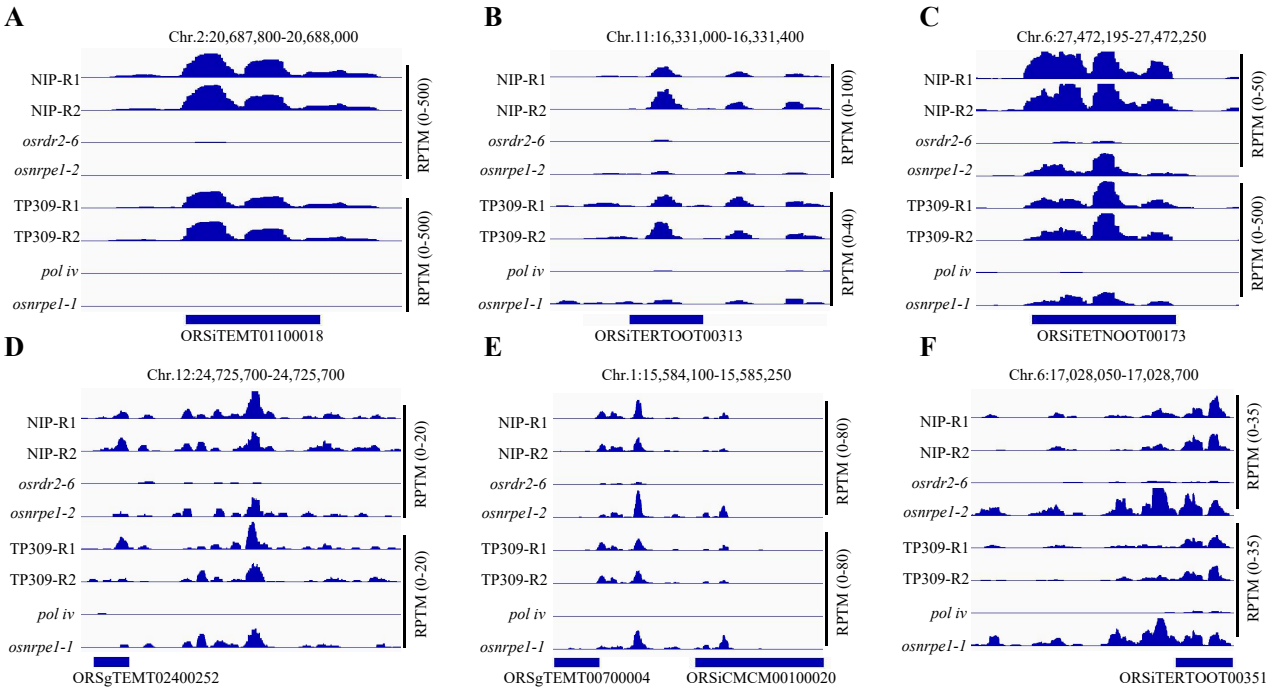

Fig. S6. OsNRPE1 plays dual roles in producing 24-nt siRNAs in rice.

Integrative genome browser image showing one representative example for six subsets of 24-nt siRNA clusters in Nipponbare, *osrdr2-6,* and *osnrpe1-2*, TP309, *pol iv*, and *osnrpe1-1* (*A* for the 0 subset; *B* for the 0-25% subset; *C* for the 25-50% subset; *D* for the 50-75% subset; *E* for the 75-100% subset; and *F* for the >100% subset). Subset percentages indicate the abundance on siRNA clusters in the *nrpe1* mutant relative to the abundance in the WT.

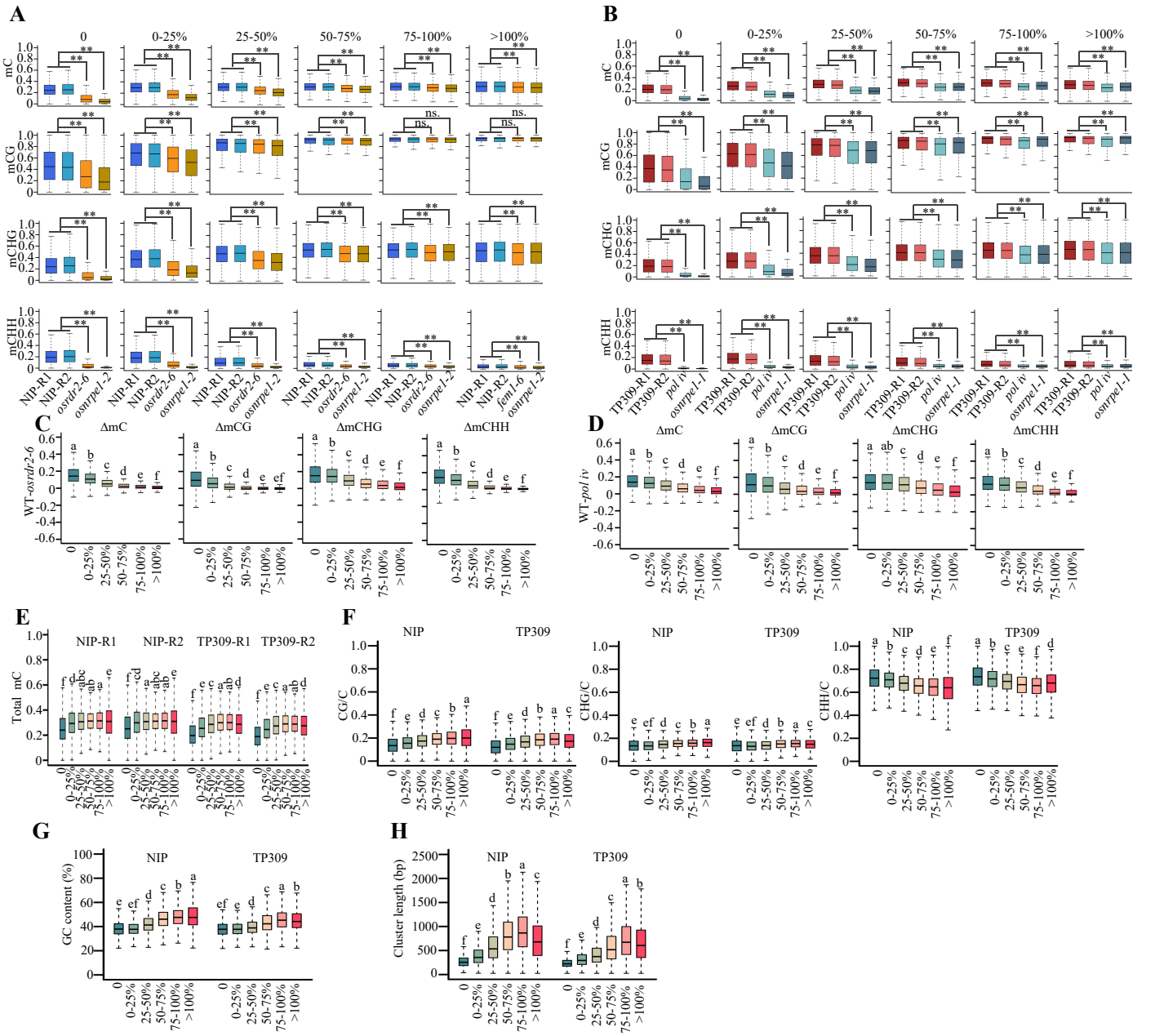

Fig. S7. DNA methylation ratio in different types of 24-nt siRNA.

(*A*) DNA methylation levels of total C, CG, CHG, and CHH in Nipponbare, *osrdr2-6*, and *osnrpe1-2* in six subsets of 24-nt siRNA clusters. ** indicates that compared plots are significantly different at $P < 0.01$; ns. indicates not significantly different (Fisher's LSD). (*B*) DNA methylation levels of total C, CG, CHG, and CHH in TP309, *pol iv*, and *osnrpe1-1* in six subsets of 24-nt siRNA clusters. ** indicates that compared plots are significantly different at $P < 0.01$ (Fisher's LSD). (*C* and *D*) Reduction of methylation levels in total C, CG, CHG, and CHH in six subsets of siRNA clusters in *osrdr2-6* (*C*) and *pol iv* (*D*). (*E*) Methylation levels of total C in six subsets of 24-nt siRNA clusters in Nipponbare and TP309. (*F*) Ratios of CG/C, CHG/C, and CHH/C in six subsets of 24-nt siRNA clusters in backgrounds of Nipponbare and TP309. (*G*) GC content in six subsets of 24-nt siRNA clusters in backgrounds of Nipponbare and TP309. (*H*) Lengths of six subsets of 24-nt siRNA clusters in Nipponbare and TP309. Within each panel of *C*–*H*, different letters indicate a significant difference at $P < 0.05$ according to Fisher's LSD.

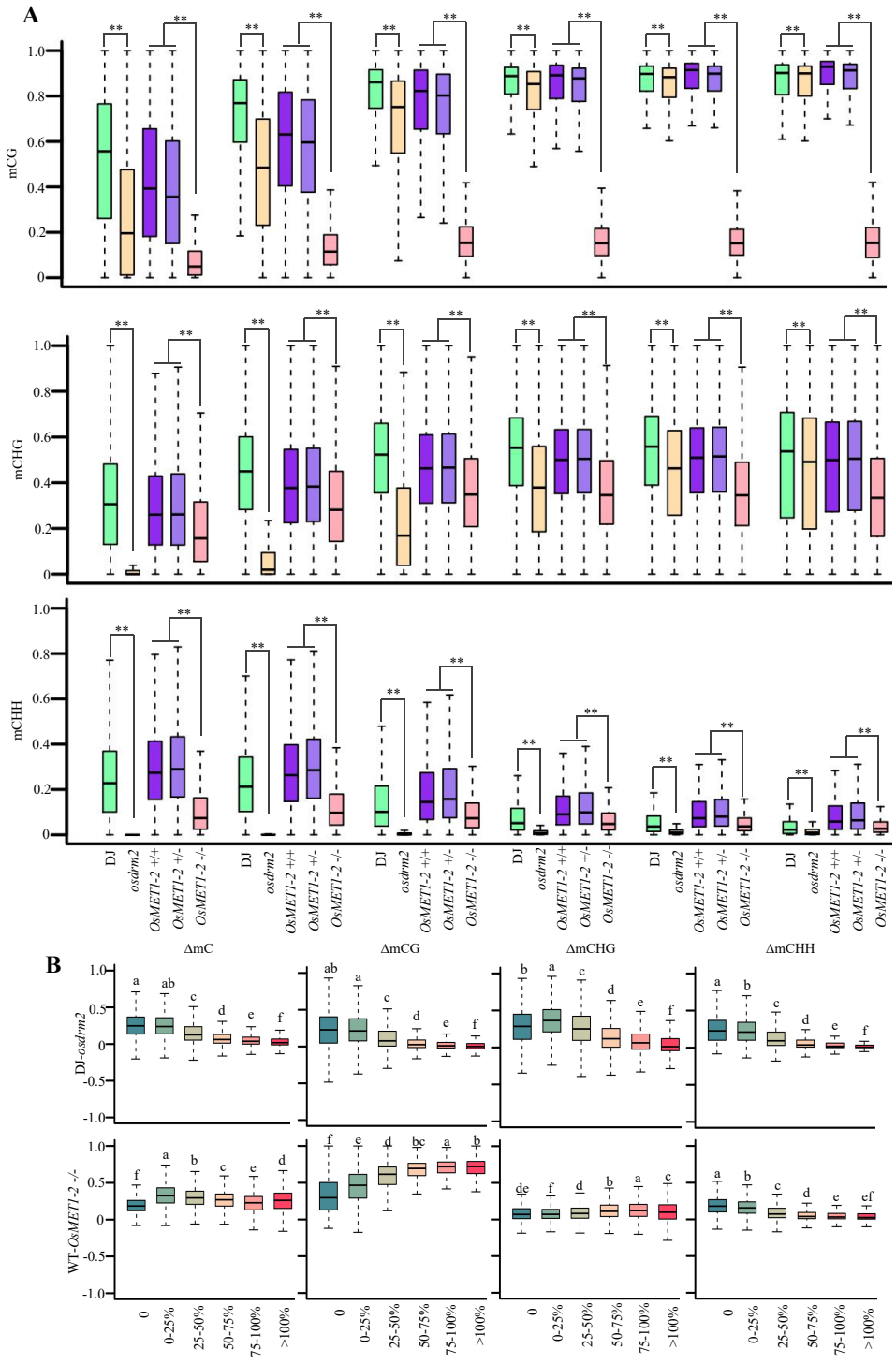

Fig. S8. *OsDRM2* and *OsMET1* control DNA methylation on six subsets of 24-nt siRNA clusters.

(*A*) The CG, CHG, and CHH methylation in the indicated genotypes in six subsets of 24-nt siRNA clusters of *osnrpe1-2* as indicated by box plots. ** indicates that compared plots are significantly different at $P < 0.01$ (Fisher's LSD). (*B*) Reduction of methylation levels in total C, CG, CHG, and CHH in six subsets of 24-nt siRNA clusters in *osdrm2* and *osmet1-2* as indicated by box plots. Different letters indicate a significant difference at $P < 0.05$ according to Fisher's LSD.

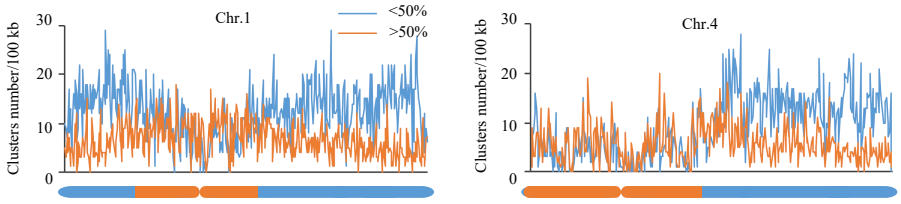

Fig. S9. Effect of Pol V on 24-nt siRNA accumulation.

The density of Pol V-dependent (<50%) and Pol V-independent (>50%) 24-nt siRNA clusters on chromosome 1 and chromosome 4; the <50% group included siRNA clusters 0, 0-25%, and 25-50%; the >50% groups included siRNA clusters 50-75%, 75-100%, and >100%. The 24-nt siRNA clusters were counted within a sliding 100 kb window. Euchromatin and heterochromatin are indicated by blue and orange, respectively.

Fig. S10. The Yin-Yang model of DNA sequence and epigenetic factors correlated with the effects of Pol V on the accumulation of 24-nt siRNAs.

Yin-Yang model showing the base bias, DNA methylation contexts, histone modifications, and gene expression on Pol V-dependent and -independent siRNA loci.

Table S1. Oligonucleotides used in this study.

oligo name	oligo sequence(5'-3')	Experiment
S3-F	GCTGGTGAACAGAGAAGGAA	Mapping
S3-R	GGCTGTAATCCAGGTTATGCT	
S11-F	ATGTGTGCTTCGGTCGTGAT	Mapping
S11-R	TTCTCACATCTGAACCTCTCC	
S18-F	AACCTACCACTGCCATTGC	Mapping
S18-R	GGCATTATCCATACCAGCAG	
R1-F	CATTTCAACTAGTAAGCGTGTC	Mapping
R1-R	TTACAGCCGCTATGATAAGG	
M1-F	TGATGCTTTGTAGTAGGTGTAGC	Mapping
M1-R	TCAGAACGATCCGGGTTGAT	
M2-F	CAGAACAAGGCAGGGGAAAT	Mapping
M2-R	ACCTGTGCTGTGGAAGAAGA	
M3-F	CGGGAGGGGTAGGAGTAGTA	Mapping
M3-R	TACGATGGAAGTGGCAAGGA	
35S-F	GTAAGGGATGACGCACAATC	genotyping for 35S::*OsGA2ox1*
OsGA2ox1 -R1	GTTCTTCTAGCTAGTTAGTG	
OsGA2ox1 RT-F	TCCGAGCAAACGATGTGGAAGG	RT-PCR
OsGA2ox1 RT-R	GCTTTTCCCTCACTGGCATT	
35S Chop-PCR-F	ATGGAGTCAAAGATTCAAATAG	Chop-PCR for 35S promoter
35S Chop-PCR-R	AAGCTTGATAAACTAGAGTCCCC	
OsGA2ox1 Chop-PCR-F	GCATGTGAGATCCTGGACCT	Chop-PCR for *OsGA2ox1*
OsGA2ox1 Chop-PCR-R	TCCTGTGTGTTTTGCAGTGAG	
OsUBI-F	AACCAGCTGAGGCCCAAGA	Control
OsUBI- R	ACGATTGATTTAACCAGTCCATGA	
fem3-1 -F	AGCCAGCTCAACATGATGC	Genotyping for *fem3-1* , *Hpy188I*
fem3-1 -R	ATTGCCCCAGGACTTCTCTG	
OsNRPE1- C-F1	cggggatcctctagagtcgacCGATGATGCATTTGAGTTTTGG	Cloning of OsNRPE1
OsNRPE1 -C-R1	gcatcatcctccgtGTAGTTAAGTTGTATGATAGACTTGCTGCA	
OsNRPE1 -C-F2	aactacACGGAGGATGATGCCTTGG	
OsNRPE1 -C-R2	acgacggccagtgccaagcttCCCAAAATGACAGTACAGCATTG	
OsNRPE1 -CI-F	ATTGAACCCACTGTGCTGAG	Genotyping for transgenic *OsNRPE1*
3301-R	GAAGCTAAACTGAAGGCGGG	
OsNRPE1 -qPCR-F	GAACACCACCCTGAGAAACA	qPCR
OsNRPE1 -qPCR-R	ACAAGCACCGACTATCCTGG	
OsNRPF1 -qPCR-F	GCAACGGCAACTACATAACAGTCG	qPCR
OsNRPF1 -qPCR-R	TCTCCAGGCACTTCTTGTAGGAG	
sgRNA1-F	GCCGGAGGACCAGTCAGCAATTC	sgRNA1 of OsNRPE1 for CRISPR vector construction
sgRNA1-R	AAACGAATTGCTGACTGGTCCTC	
sgRNA2-F	GCCGGTCCAATGGTGCCCGTGGCT	sgRNA2 of OsNRPF1 for CRISPR vector construction
sgRNA2-R	AAACAGCCACGGGCACCATTGGAC	
sgRNA3-F	GCCGTCGGCAGCTGTCAGTCGTG	sgRNA3 of OsNRPF1 for CRISPR vector construction
sgRNA3-R	AAACCACGACTGACAGCTGCCGA	
sgRNA4-F	TGTGTGGTGAACCTGGGGCCCCTGT	sgRNA4 of OsNRPD1 for CRISPR vector construction
sgRNA4-R	AAACACAGGGGCCCCAGGTTCACCA	
LMR15-F	TGGTGAATGGTGGCATGAAA	Chop-PCR
LMR15-R	CGCAGCATAACTCAGGTGTG	
LMR30-F	GGGAGTTGGGGTAGGGATTC	Chop-PCR
LMR30-R	TCCAATATCCAACCCAGGCA	
LMR34-F	TTCCTGCACTCAAAACTGGC	Chop-PCR
LMR34-R	CCCTCATCTAGCTGCGTTCT	
LMR36-F	GCTTTTCAAACGGTTAAGCGG	Chop-PCR
LMR36-R	TGCAGCAGAGAAAGAATCACG	
LMR46-F	AGCCAACAGATGACTCCGAA	Chop-PCR
LMR46-R	CAAATTGCTTGCCAGTTCTCA	

LMR75-F	CTTCGTGCCAGTGAACCAAA	Chop-PCR
LMR75-R	CCAAGGCGCAGGTACAATAG	
OsIGN1 -F	CGCATGATGAGAACAGCACCACCT	RT-PCR
OsIGN1 -R	GCGAACAGCGCGGAAGAAGGAA	
OsIGN2 -F	AGTTTACAGTTCGCTTCTCCTCTCTCTCCTCTC	RT-PCR
OsIGN2 -R	ACTATGGTCACGGGCTCAATCTGTTTCTGTTG	
OsIGN3 -F	ACAAATATACTAGGAGGTGTGAATAGGAGCTTT	RT-PCR
OsIGN3 -R	GGAGAGAGTATACTTAAATTCATCTATAGTCAATCT	
OsIGN4 -F	AGGTGATAGAATACAAAAGAGGAAAGGGTGATT	RT-PCR
OsIGN4 -R	TGAGACATAGAGAGAAGAAGAAAATTGTAGCCA	
OsIGN5 -F	GTTCTCGTGTTGAAGCTGTGTCTTGCATGA	RT-PCR
OsIGN5 -R	GTTGTACACTTGTCCTACCTCCCTTCATAAAAG	
OsIGN6 -F	CGGAGGATCCACGGCCAACAT	RT-PCR
OsIGN6 -R	GAGCCTGAGCTGAAACTGTGACAA	

Table S2. High throughput sequencing datasets

Bisulfite sequencing datasets

sample	raw reads	raw bases	clean reads	clean bases	mapped reads	mapped ratio	coverage	depth(X)	conversion rate
NIP-R1	227,476,972	34,121,545,800	218,857,910	32,828,686,500	168,637,380	77.05%	94.59%	35.33	99.53%
NIP-R2	178,621,232	26,793,184,800	171,791,384	25,768,707,600	132,069,884	76.88%	94.05%	29.42	99.51%
osrdr2-6	243,416,222	36,512,433,300	234,162,858	35,124,428,700	182,309,254	77.86%	94.70%	39.09	99.55%
osnrpe1-2	223,159,486	33,473,922,900	209,677,710	31,451,656,500	164,760,120	78.6%	94.29%	34.53	99.58%
TP309-R1	163,069,870	24,460,480,500	157,335,512	23,600,326,800	119,376,714	75.90%	93.27%	29.18	99.57%
TP309-R2	143,203,494	21,480,524,100	137,859,672	20,678,950,800	104,160,102	75.60%	93.00%	25.43	99.59%
pol iv	144,358,654	21,653,798,100	125,541,120	18,831,168,000	81,249,936	64.7%	90.74%	22.40	99.40%
osnrpe1-1	156,334,956	23,450,243,400	150,092,986	22,513,947,900	112,771,728	75.1%	93.52%	29.55	99.53%

Small RNA sequencing datasets

sample	raw reads	clean reads	mapped reads	mapped ratio
NIP-R1	24,875,329	24,875,202	21,678,829	87.15%
NIP-R2	25,131,542	25,131,495	21,913,601	87.20%
osrdr2-6	24,737,577	24,737,425	19,747,442	79.83%
osnrpe1-2	24,932,994	24,932,958	16,984,194	68.12%
TP309-R1	28,321,200	28,311,900	22,601,747	79.83%
TP309-R2	27,980,509	27,965,218	22,057,092	78.87%
pol iv	23,886,458	23,886,158	19,193,951	80.36%
osnrpe1-1	29,320,568	29,315,314	24,824,433	84.68%

Genome-wide methylation levels

	CG	CHG	CHH
NIP-R1	43.6%	23.4%	3.1%
NIP-R2	41.9%	23.7%	3.1%
osrdr2-6	43.1%	22.4%	2.0%
osnrpe1-2	43.5%	22.5%	1.6%
TP309-R1	44.4%	20.7%	2.6%
TP309-R2	43.7%	19.8%	2.4%
pol iv	38.7%	17.0%	1.7%
osnrpe1-1	44.0%	20.2%	1.6%