PNAS
www.pnas.org

Supplementary Information for

DNA demethylation by ROS1a in rice vegetative cells promotes methylation in sperm

M. Yvonne Kim, Akemi Ono, Stefan Scholten, Tetsu Kinoshita, Daniel Zilberman, Takashi Okamoto, and Robert L. Fischer

Robert L. Fisher, Takashi Okamoto, Daniel Zilberman, Tetsu Kinoshita, Stefan Scholten.
Email: rfischer@berkeley.edu, okamoto-takashi@tmu.ac.jp, Daniel.Zilberman@jic.ac.uk, tkinoshi@yokohama-cu.ac.jp, stefanscholten@gmx.eu

This PDF file includes:

Other supplementary materials for this manuscript include the following:

www.pnas.org/cgi/doi/10.1073/pnas.1821435116

Kim et al., Supplemental Information

Supplemental Methods

Generation of *ros1a/+*; UBQ::H2B-GFP plants. UBQ::H2B-GFP pollen was used to fertilize the *ROS1a*/*ros1a-GUS1* flowers. The F1 embryo was rescued and grown in sterile conditions until transplanting. From the F1 plant, the embryos with the *ros1a/+* mutant phenotype were rescued as previously described (1). Pollen from F2 plants was screened for GFP expression and only the plants with 100% GFP expressing pollen were used for the experiment.

Isolation of Nipponbare sperm cells and vegetative cell nuclei. Wild-type and *ros1a/+ O. sativa* cv. Nipponbare harboring a *UBQ::H2B-GFP* transgene were grown at 26 ºC in a 13/11 hr light/dark cycled chamber. Rice sperm cells and vegetative cell nuclei were isolated as previously described (2, 3). Briefly, rice pollen was harvested from unopened flowers with elongated stamen. Pollen grains were burst open via osmotic pressure in 370 mOsmol/kg H_2O mannitol solution for 5 minutes at room temperature. Sperm cells and vegetative cell nuclei were isolated using the differential interference contrast microscopy and GFP fluorescent microscopy, respectively, under a BX-71 inverted fluorescence microscope (Olympus, Tokyo, Japan). The sperm cells and vegetative cell nuclei were washed once in 370 mOsmol/kg H_2O mannitol solution and stored at -80C.

Whole-genome Bisulfite Sequencing. Bisulfite sequencing libraries were constructed as previously described (4). Illumina sequencing was performed using the HiSeq4000 100 nt single-end read platform at Vincent J. Coates Genomic Sequencing Laboratory at UC Berkeley. Sequenced reads, as well as the previous published data used in the study, were trimmed using Trim Galore and mapped to the MSU7 reference genome using BS-Seeker2 using Bowtie2 with the end-to-end mode (5). Bioinformatic analyses were carried out as previously described (6, 7).

Density plots. Density plots are showing the frequency distribution of DNA methylation differences between 50-bp window of two samples (combination of wild-type vegetative cell, wild-type sperm, *ros1a/+* vegetative cell, and *ros1a/+* sperm) with at least 20 informative sequenced cytosines in both samples and 70% CG, 30% CHG, or 10% CHH methylation in either of the samples

Box plots. The 50-bp windows with at least 20 sequenced cytosines (10 sequenced cytosines for egg) and with methylation greater than 50% or 30% in one of the samples in CG/CHG or CHH contexts were used for the analysis. For wild-type and *ros1a/+* sperm comparison, only 50-bp windows with methylation greater than 30% or 10% in either wild-type or *ros1a/+* in CG or CHG/CHH methylation, respectively, with at least 20 sequenced cytosines in each window were used.

CG DMRs. The CG DMRs for all comparisons are 50-bp windows with at least 20 sequenced cytosines and with methylation greater than 70% in one of the samples in CG context were used for the analysis. The difference between two different samples with at least 0.5 (50% methylation) were identified as the hypomethylated loci. The only exception is that for egg and central cell DMRs we required only 10 sequenced cytosines for egg and central cell (6, 8, 9).

Low-stringency DMRs. CG fractional methylation in 50-bp windows (each window averages methylation on both strands) was compared between vegetative cell and sperm. Windows with sperm fractional methylation at least 0.2 greater than that in vegetative cell and a Fisher's exact test $p < 0.05$ were merged if they occurred within 300 bp. Merged DMRs were retained if the fractional methylation in sperm across the DMR was at least 0.3 greater than that in vegetative cell and the Fisher's exact test was $p < 10^{-5}$ (10).

Chromatin feature analyses. Previously published sequencing data (11) were aligned using Bowtie2 and non-uniquely aligned reads were removed. The scores of aligned reads within 50 bp windows were summed. The five groups used for *Fig. S4* are as follows: for TE length, <0.5, 0.5-1, 1-2, 2-5, and >5 kb; for GC ratio, 0.4-0.5, 0.5-0.6, 0.6-0.7, and 0.8-0.9; for H3K9me2, 1-3, 3-5, 5-7, 7-9, and >9 reads; for H3K9me1, 1-3, 3-5, 5-7, 7-9, and >9 reads; for H3K27me3, 1-3, 3-5, 5-7, 5-9, and >9 reads; for H3K4ac, 1-3, 3-5, 5-7, 7-9, and >9 reads; for H3K9ac, 1-2, 2-3, 3-5, 5-8, and >8 reads; and for H3K27ac, 1-3, 3-5, 5-7, 7-9, and >9 reads.

Imprinted gene RNA analyses. Previously published vegetative cell and sperm transcriptome data (11) were aligned using TopHat (12) and differentially expressed genes between vegetative cell and sperm cells were identified using DESeq (13). Previously identified imprinted genes from (14) with at least 2-fold expression difference and p-value < 0.05 were used for the analysis.

Supplementary Figures

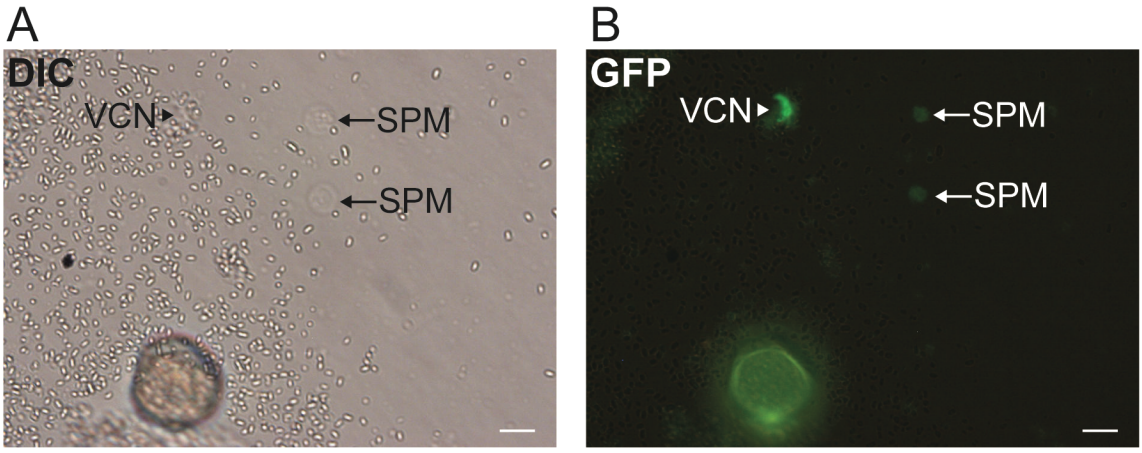

Fig. S1. GFP expression in rice sperm and vegetative cell nucleus. (*A-B*) Differential interference contrast microscopy (*A*, DIC) and GFP fluorescence microscopy (*B,* GFP) images of burst pollen from plants harboring the *UBQ::H2B-GFP* transgene. Strong GFP fluorescence expression is localized in the nucleus of rice sperm (SPM) and vegetative cell nucleus (VCN). Scale Bars, 10 um.

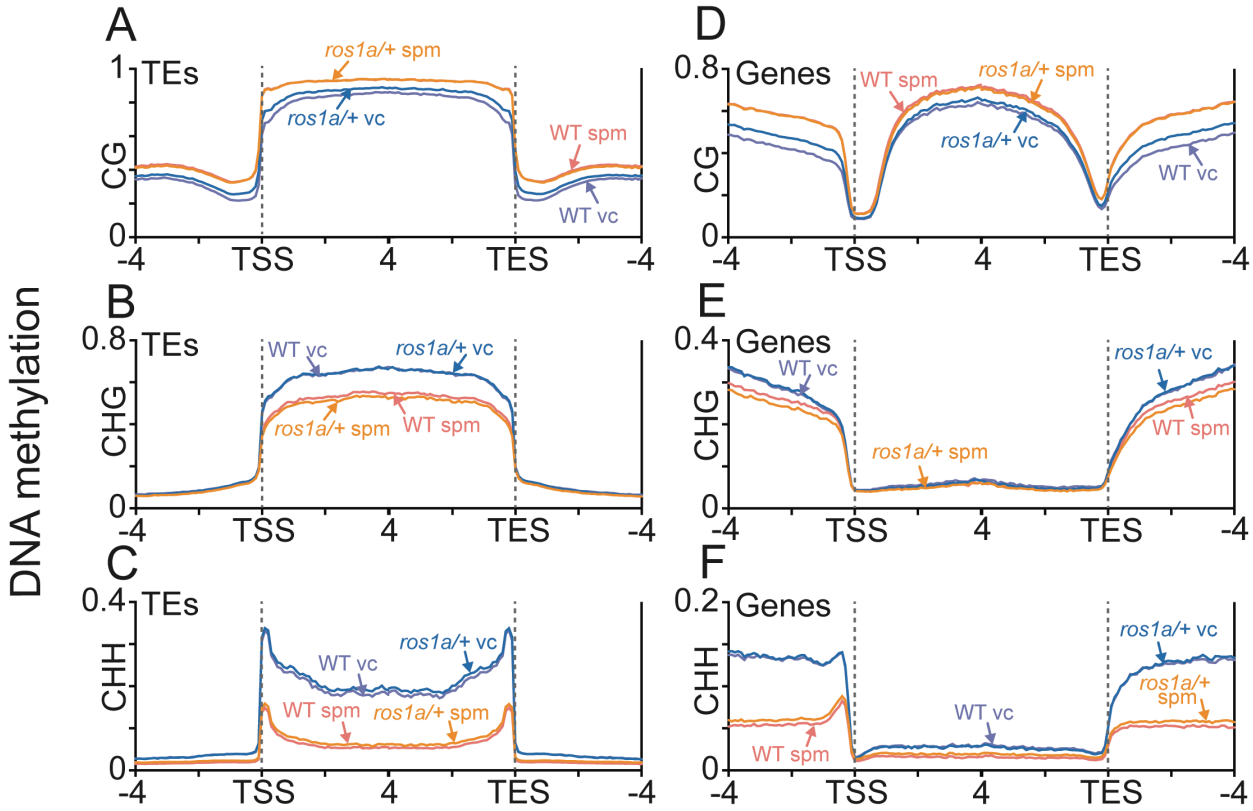

Fig. S2. Average TE and gene methylation in wild-type and *ros1a/+* sperm and vegetative cell. (*A-C*) TEs were aligned at the 5' (TE start site, TSS) and 3' (TE end site, TES) ends. Methylation within each 100-bp interval was averaged and plotted from 4 kb away from the annotated TE (negative numbers) to 4 kb into the annotated TE (positive numbers). The dashed lines represent the points of alignment. (*D-F*) Genes were aligned at the 5' (transcription start site, TSS) and 3' (transcription termination site, TES) ends. Methylation within each 100-bp interval was averaged and plotted from 4 kb away from the annotated gene (negative numbers) to 4 kb into the annotated gene (positive numbers). The dashed lines represent the points of alignment. spm, sperm; vc, vegetative cell.

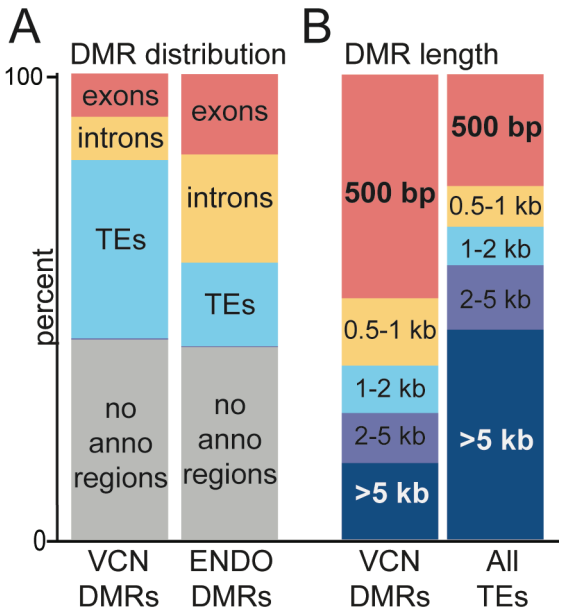

Fig. S3. Local hypomethylation frequently occurs in short transposons in vegetative cells. *(A-B)* Genomic distribution of differentially methylated regions (DMRs) between sperm and vegetative cell. *(A)* 50-bp windows were assigned to exons, introns, TEs, and non-annotated regions. The number of windows that overlapped defined vegetative cell DMRs (VCN DMRs, compared to sperm) or endosperm DMRs (ENDO DMRs, compared to embryo) were identified. *(B)* 50-bp windows were assigned to the length of their corresponding TE and the number of windows that overlapped vegetative cell DMRs (VCN DMRs) and all TEs (All TEs) were counted.

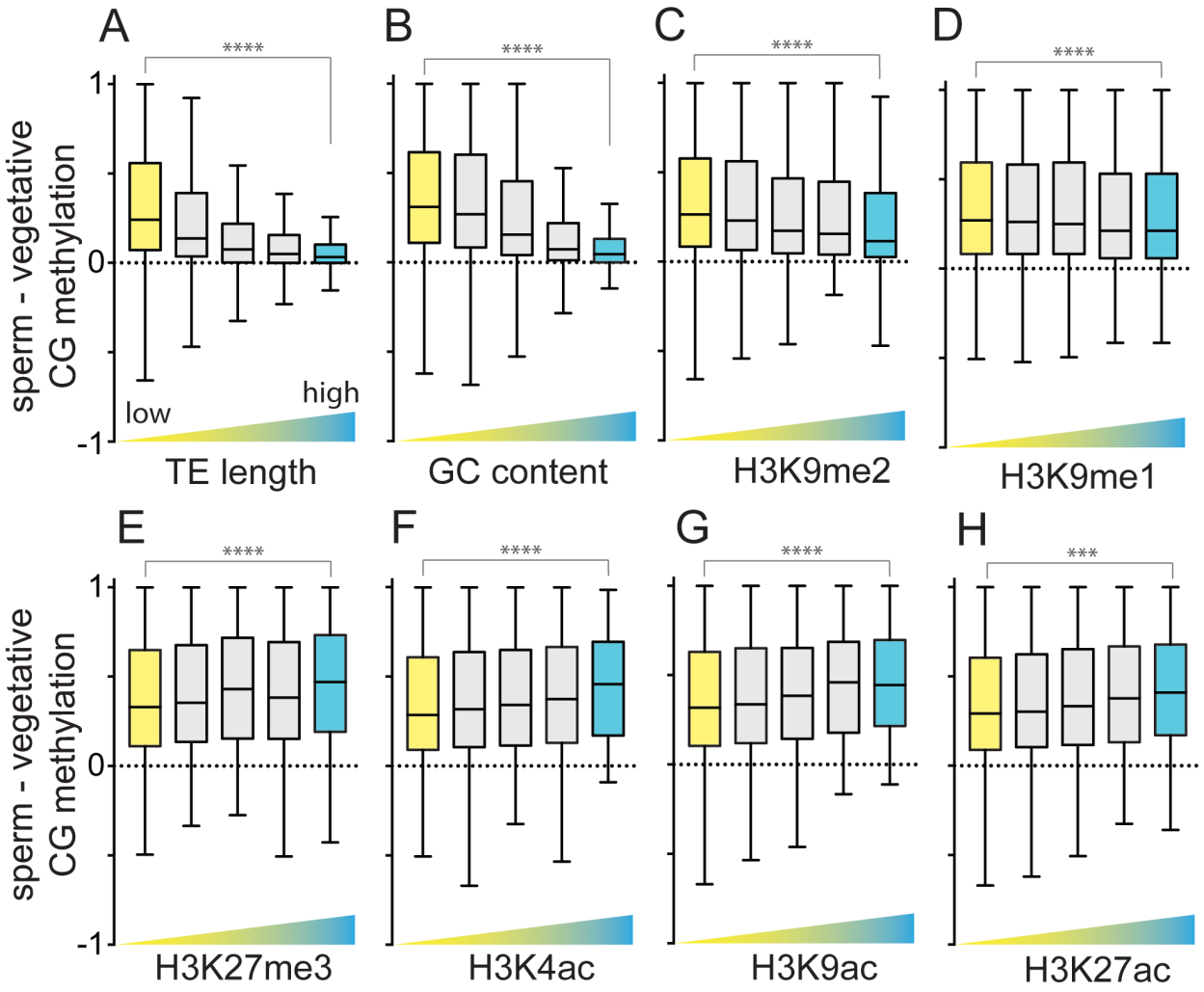

Fig. S4. Local hypomethylation occurs in euchromatin. (*A-H*) The box plots show fractional CG methylation differences between sperm and vegetative cells in all TEs (*A* and *B*) or in TEs smaller than 500 bp (*C-H*). The windows are grouped in ascending order according to TE length (*A*), CG content (*B*), enrichment of heterochromatin marks (H3K9me2 and H3K9me1) (15, 16) (*C-D*), and enrichment of euchromatin marks (H3K27me3, H3K4ac, H3K9ac, and H3K27ac) (16) (*E-H*). Each box encloses the middle 50% of the distribution, with the horizontal line marking the median and vertical lines marking the minimum and maximum values that fall within 1.5 times the height of the box. The difference between yellow and blue shaded box plots is significant (****P < 0.0001 & ***P=0.0002, Mann-Whitney test).

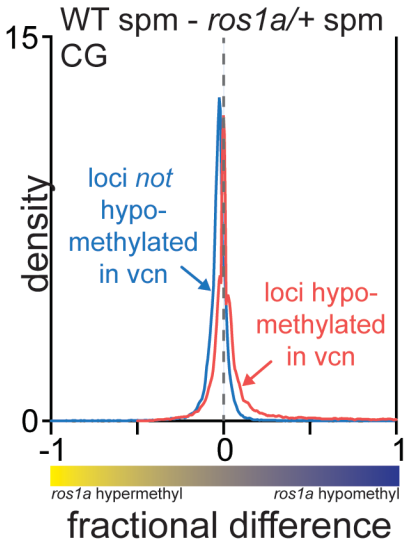

Fig. S5. ROS1a indirectly has minimal effect on CG methylation in sperm. Density plots of CG methylation differences between wild-type sperm and *ros1a*/+ sperm. The blue traces display control loci that are not hypomethylated in the vegetative cell, with the sperm and vegetative cell methylation difference less than zero in CG, CHG, and CHH (see Fig. 1A-C, gray trace). The red traces in Fig. 3A and B contain the CG DMRs in *Fig. 1A* (shaded region), that also have sperm and vegetative CHG and CHH methylation differences bigger than zero, as shown in Fig. 1B (red trace) and Fig. 1C (red trace), respectively.

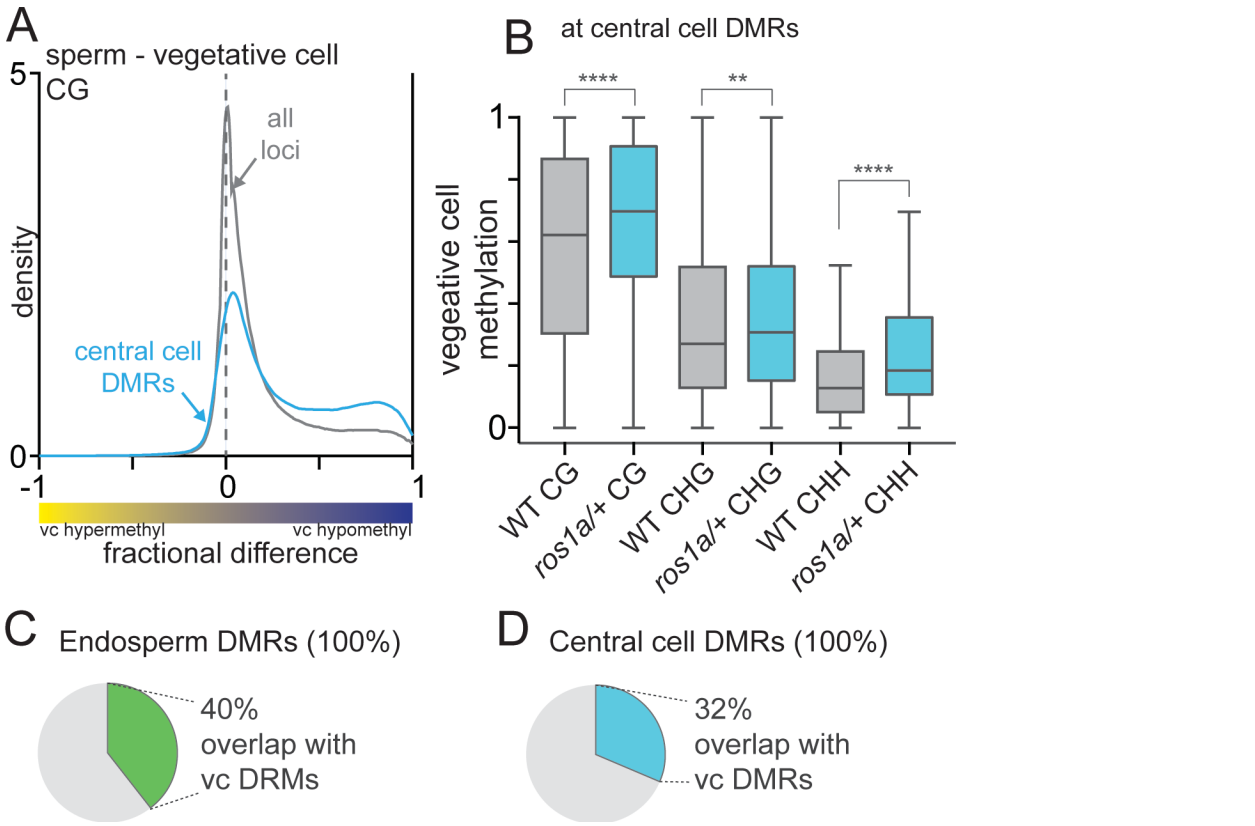

Fig. S6. ROS1-targeted DMRs in the vegetative cell overlap with central cell DMRs. (*A*) Density plots showing the fractional distribution of CG methylation differences between sperm and vegetative cell with at least 20 sequenced cytosines with 70% methylation in either of the samples. The gray trace represents all loci and the blue trace contains the sequences that overlap with hypomethylated sites in central cells compared to eggs (central cell hypo-loci). *(B)* Box plots displaying absolute DNA methylation levels of wild-type and *ros1a/+* vegetative cell 50-bp windows with at least 20 sequenced cytosines and 30% CG or 10% non-CG methylation within central cell hypo-loci (****P<0.0001 and **P=0.0069, Wilcoxon's matched-pairs signed rank test). *(C-D)* Pie charts showing the percent overlap of endosperm (*C*) or central cell (*D*) DMRs with the vegetative cell DMRs. Previously published central cell and egg cell data are from Park et al (2006) (6).

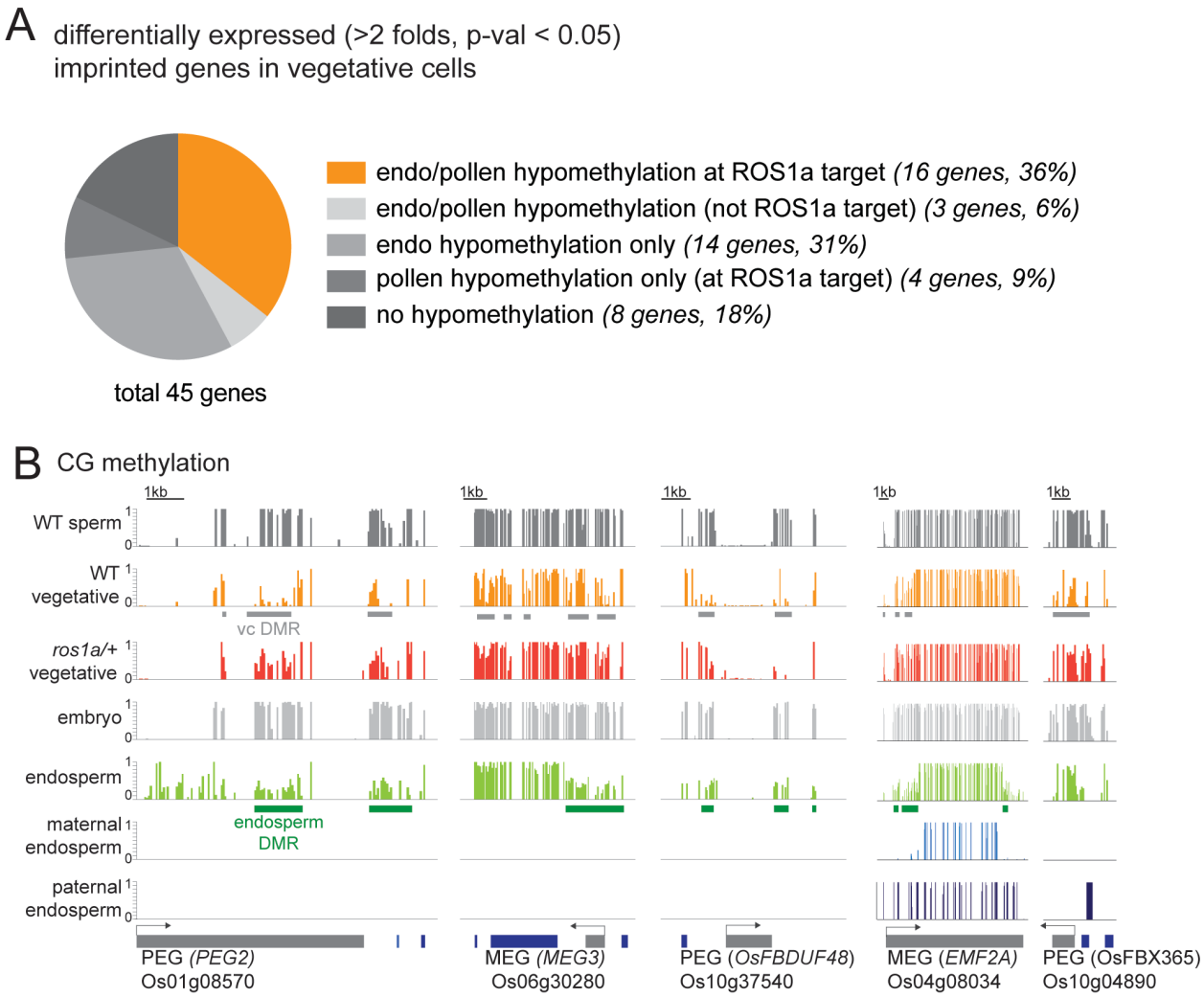

Fig. S7. ROS1a targets are found near/within the known imprinted genes. (*A*) Pie chart showing differentially expressed imprinted genes in vegetative cells compared to sperm. CG methylation within 1 kb upstream and downstream of the gene was investigated. (*F*) Snapshots of ROS1a target sites near maternally expressed genes (MEG) and paternally expressed genes (PEG).

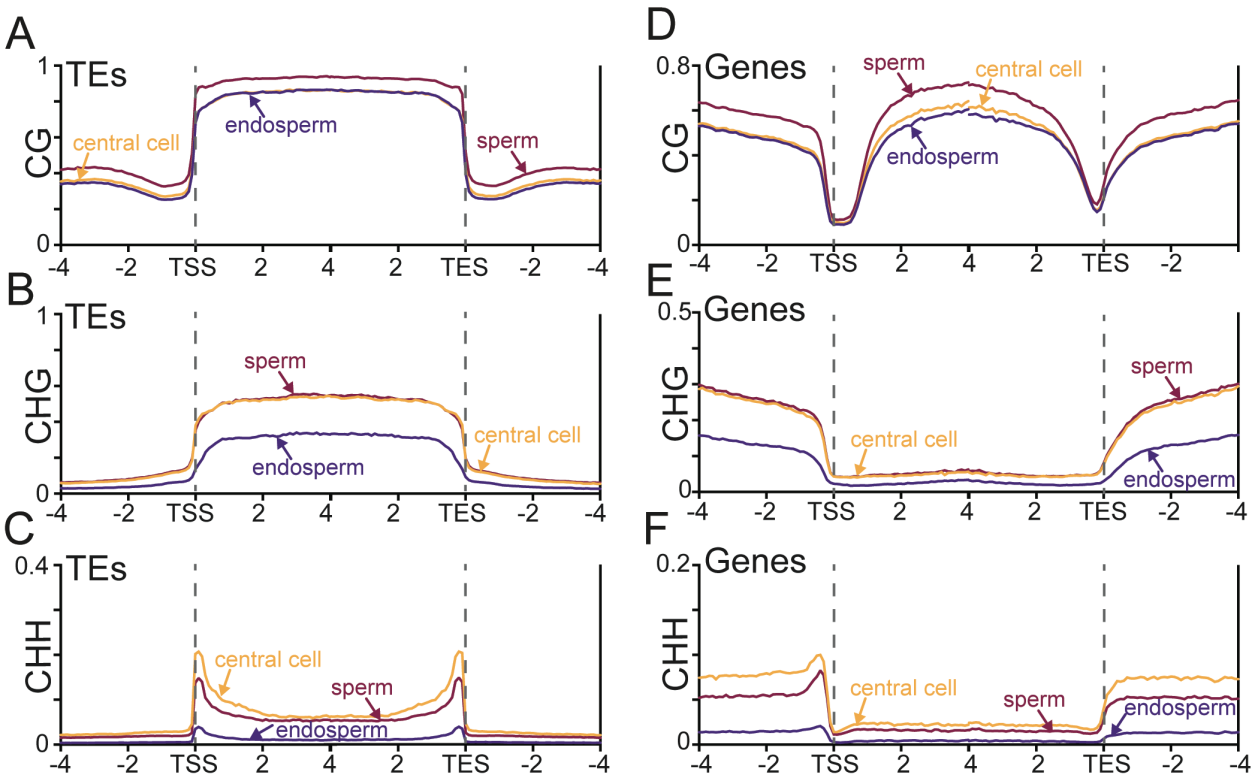

Fig. S8. Passive loss of DNA methylation occurs in endosperm. (*A-F*) TEs or genes were aligned at the 5' (TSS) and 3' (TES) ends. Methylation within each 100-bp interval was averaged and plotted from 4 kb away from the annotated TE (negative numbers) to 4 kb into the annotated TE (positive numbers). The dashed lines represent the points of alignment.

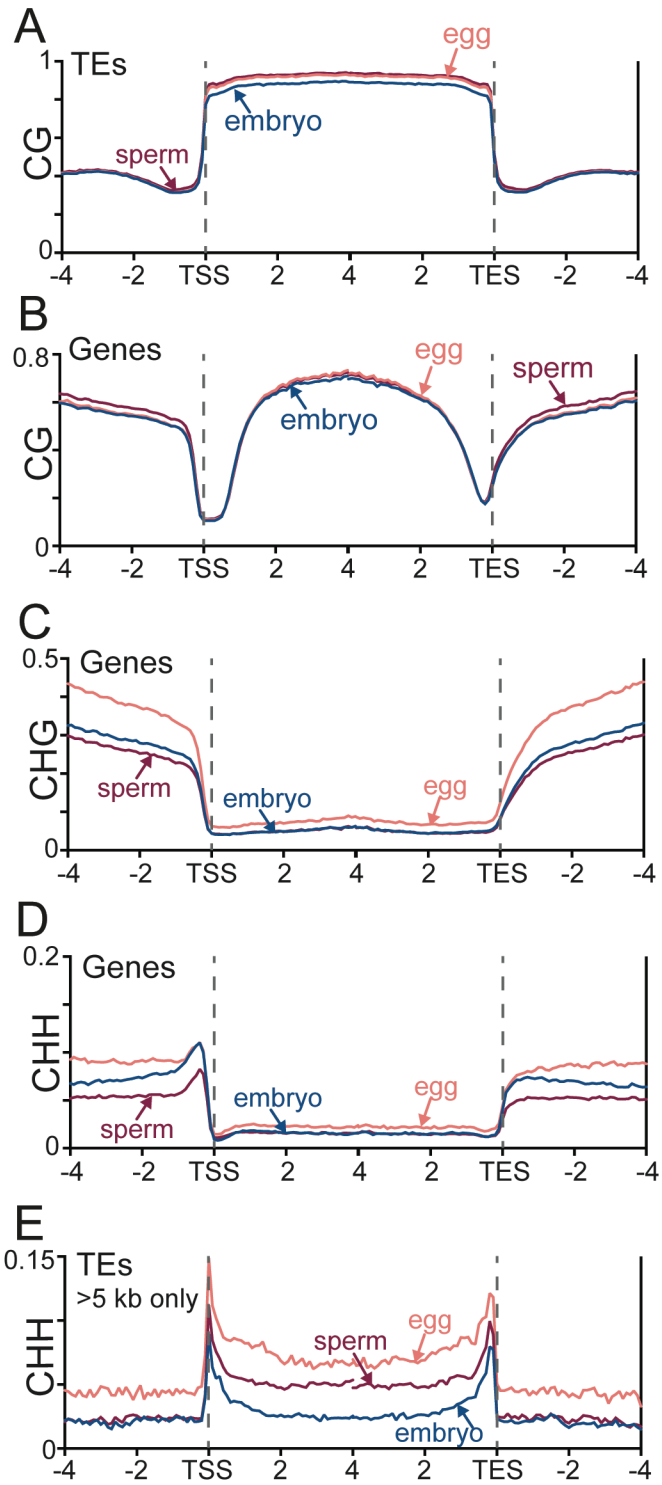

Fig. S9. DNA methylation of egg, sperm, and embryo in TEs and genes. (*A-E*) TEs or genes were aligned at the 5' (TSS) and 3' (TES) ends. Methylation within each 100-bp interval was averaged and plotted from 4 kb away from the annotated TE or gene (negative numbers) to 4 kb into the annotated TE or gene (positive numbers). The dashed lines represent the points of alignment. The analysis shown in (*A*) contains all TEs, (*B-D*) contain genes > 1.5 kb, and (*E*) contains the TEs that are bigger or equal to 5 kb.

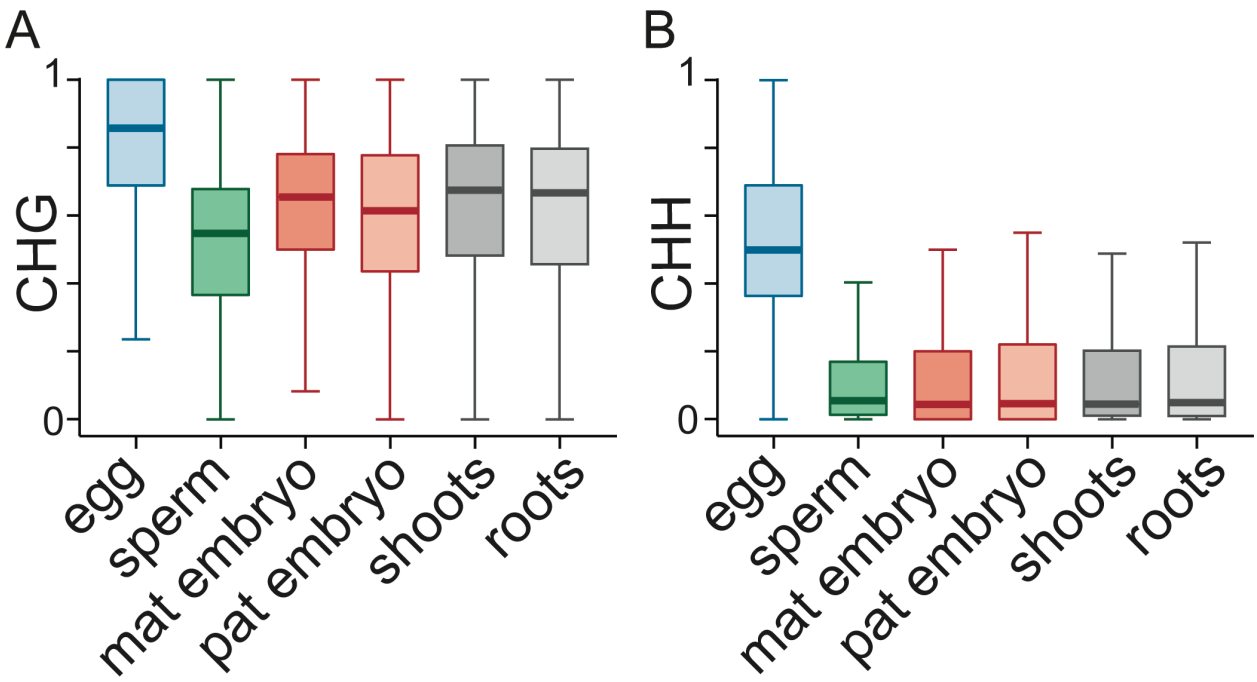

Fig. S10. DNA methylation reprogramming occurs during embryogenesis. (*A*-*B*) Box plots showing DNA methylation levels of 50-bp windows in the indicated cell or tissue type. The windows with at least 20 sequenced cytosines (10 sequenced cytosines for egg) and with methylation greater than 50% or 30% in one of the samples in CHG or CHH methylation, respectively, were kept for analyses.

Table S1. Mean genomic coverage and nuclear methylation for wild-type and *ros1a*/+ sequenced in this paper. Mean nuclear methylation is calculated by averaging methylation of individual cytosines in each context.

Sample	# of cells	Coverage	Nuclear CG	Nuclear CHG	Nuclear CHH	Chloroplast overall CHH
Wild-type sperm	341	11.2	47.9%	19.7%	3.4%	1.9%
Wild-type vegetative	432	8.7	36.3%	20.5%	7.8%	2.0%
ros1a/+ sperm	398	7.25	46.2%	18.5%	3.8%	1.1%
ros1a/+ vegetative	277	7.15	38.0%	20.5%	8.0%	1.5%

Dataset S1. CG DMRs. The location, number of methylated/unmethylated sequenced CG of sperm and vegetative cell, and methylation level of CG DMRs.

Dataset S2. High-stringency DMRs. The location, methylation level, and p-value of low-stringency DMRs.

Dataset S3. CG central cell DMRs. The location, number of methylated/unmethylated sequenced CG of egg and central cell, and methylation level of CG central cell DMRs.

Dataset S4. CG endosperm DMRs. The location, number of methylated/unmethylated sequenced CG of egg and endosperm, and methylation level of CG endosperm DMRs.

References

1. Ono A, et al. (2012) A null mutation of ROS1a for DNA demethylation in rice is not transmittable to progeny. *The Plant Journal* 71(4):564–574.
2. Toda E, Okamoto T (2016) Formation of triploid plants via possible polyspermy. *Plant Signal Behav* 11(9):e1218107.
3. Abiko M, Maeda H, Tamura K, Hara-Nishimura I, Okamoto T (2013) Gene expression profiles in rice gametes and zygotes: identification of gamete-enriched genes and up- or down-regulated genes in zygotes after fertilization. *J Exp Bot* 64(7):1927–1940.
4. Smallwood SA, et al. (2014) Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat Methods* 11(8):817–820.
5. Guo W, et al. (2013) BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data. *BMC Genomics* 14:774.
6. Park K, et al. (2016) DNA demethylation is initiated in the central cells of Arabidopsis and rice. *Proc Natl Acad Sci USA* 113(52):15138–15143.
7. Zemach A, et al. (2013) The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* 153(1):193–205.
8. Ibarra CA, et al. (2012) Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* 337(6100):1360–1364.
9. Hsieh T-F, et al. (2009) Genome-wide demethylation of Arabidopsis endosperm. *Science* 324(5933):1451–1454.
10. Rodrigues JA, et al. (2013) Imprinted expression of genes and small RNA is associated with localized hypomethylation of the maternal genome in rice endosperm. *Proc Natl Acad Sci USA* 110(19):7934–7939.
11. Anderson SN, et al. (2013) Transcriptomes of isolated Oryza sativa gametes characterized by deep sequencing: evidence for distinct sex-dependent chromatin and epigenetic states before fertilization. *The Plant Journal* 76(5):729–741.
12. Trapnell C, et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7(3):562–578.
13. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.
14. Luo M, et al. (2011) A genome-wide survey of imprinted genes in rice seeds reveals imprinting primarily occurs in the endosperm. *PLoS Genet* 7(6):e1002125.
15. Tan F, et al. (2016) Analysis of Chromatin Regulators Reveals Specific Features of Rice DNA Methylation Pathways. *Plant Physiol* 171(3):2041–2054.
16. Fang Y, et al. (2016) Functional characterization of open chromatin in bidirectional promoters of rice. *Sci Rep* 6(1):32088.