SI Appendix

Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome-expression, and early seedling-lethality

Lanjuan Hu[a,1], Ning Li[a,1], Chunming Xu[a,1], Silin Zhong[a,b,1], Xiuyun Lin[c], Jingjing Yang[a], Tianqi Zhou[a], Anzhi Yuliang[a], Ying Wu[a], Yun-Ru Chen[b], Xiaofeng Cao[d], Assaf Zemach[e], Sachin Rustgi[f], Diter von Wettstein[f,2], and Bao Liu[a,2]

[a]Key Laboratory of Molecular Epigenetics of the Ministry of Education, Northeast Normal University, Changchun 130024, China; [b]State Key Laboratory of Agrobiotechnology, the School of Life Sciences, The Chinese University of Hong Kong, Hong Kong, China; [c]Jilin Academy of Agricultural Sciences, Changchun 130033; [d]State Key Laboratory of Plant Genomics and National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China; [e]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, United States of America; [f]Department of Crop and Soil Sciences, School of Molecular Biosciences, and Center for Reproductive Biology, Washington State University, Pullman, WA 99164, United States of America

Including: 6 Supplementary tables and 13 Supplementary figures.

SI Appendix

SI Materials and Methods

Plant Materials. The three genotypes of the *OsMET1-2* gene in rice (cv. Nipponbare), wild-type (WT, +/+), heterozygote (+/-) and homozygous mutant (-/-) were germinated and grown on the hormone-free, half-strength Murashige & Skoog (MS) medium under 16/8 hrs light/dark at 26/24℃. Eleven-day-old seedlings from each genotype were harvested for genomic DNA and total RNA extraction.

Identification of Heterozygote and Homozygote *OsMET1-2* Mutant Plants by Locus-Specific Pyrosequencing. The genomic region adjacent to the *Tos17* insertion site was highly homologous with that of the other *MET1* gene in rice (*OsMET1-1*), which obstructs identification of *OsMET1-2 +/-* and *OsMET1-2 -/-* by conventional PCR-based assay. To circumvent this problem, heterozygote (+/-) and homozygote (-/-) seedlings of the *OsMET1-2* mutant were distinguished by pyrosequencing (PyroMark™, Biotage AB, Sweden) based on a single nucleotide polymorphism (T/C) between the *OsMET1-1*and *OsMET1-2*genes (Fig. S1). The pyrosequencing primers (P1L and P1R, Table S6) were designed to anneal to the regions immediately upstream of the SNP sites (Table S6; Fig. S2), and which were verified by the pyrosequencing platform (PyroMarkQ96 ID, Qiagen, USA). Pyrosequencing was essentially according to the procedures previously described (1). Briefly, PCR amplifications were carried out using HotStar *Taq* DNA Polymerase following the manufacturers' instructions (Qiagen, USA). The thermal cycling program was as follows: 95℃ for 5 min, followed by 50 cycles at 95 ℃ for 30 s, 52 ℃ for 30 s and 72 ℃ for another 30 s and a final extension for 5 min at 72 ℃. PCR products were analyzed by electrophoresis on 1.5% agars gel.

Semi-Quantitative RT-PCR and Real-Time Quantitative (q)RT-PCR Analyses. Total RNA was isolated from 11-day-old seedlings or calli of the three *OsMET1-2* genotypes (WT +/+, heterozygous +/- and homozygous -/- mutant types) using TRIzol reagent (Invitrogen™ | Life Technologies, USA) and then treated with DNase I (Ambion® | Life Technologies, USA). One μg of treated total RNA was used to synthesize the first-strand cDNA with a reverse-transcription kit (Promega Cooperation, USA). The primers P2L and P2R (Table S6) which targeted to conserved regions between the two *OsMET1*genes (*OsMET1-1* and-*2*) was considered suitable to assay the expression level of *OsMET1-2* because *OsMET1-1* was unexpressed at this stage of the tissue (2, 3). The rice *Actin* gene (amplified with primers P3L and P3R, Table S6) was used as an internal control. The qRT-PCR assay was conducted for dual purposes: to verify reliability of the RNA-seq data, and to quantify expression of a set of DNA methylation/chromatin related genes in the mutant. The SYBR Green I PCR master mix kit (TaKaRa, Japan) was used in the qRT-PCR reactions with each of the gene-specific primer pairs (Table S6).

Methylation-Sensitive Endonuclease Digestion. One μg genomic DNA each of the three genotypes (+/+, +/- and -/-) of the *MET1-2* gene was digested by 3 units of *Hpa*II (New England Biolabs, USA), which was sensitive to the CG methylation at the 5′-CCGG sites. The digestion products of each genotype at the three time points along with intact genomic DNA were fractionated by 1% agarose gel electrophoresis with ethidum bromide staining.

Locus-Specific Bisulfite Sequencing. Genomic DNA (500 ng) was treated by the EZ DNA Methylation-Gold kit (Zymo Research, http://zymoresearch.com) according to the manufacture's recommendations. For each PCR reaction, 3μl of bisulfite-treated DNA was used, and the PCR products were cloned into the pMD-18T vector and sequenced. For each

genotype, at least 12 clones were sequenced, and analyzed at the Kismeth website (http://katahdin.mssm.edu/kismeth/revpage.pl). Primers used to analyze the *Tos17* body methylation were as reported (4).

Microscopy. Eleven day-old rice seedlings were fixed overnight in FAA (3.7% formaldehyde, 5% acetic acid, 50% ethanol) at 4 ℃. Samples were dehydrated through graded ethanol series. After infiltration with xylene, samples were embedded in paraffin (Sigma-Aldrich, USA), sectioned at 10 μm, and stained with 1% safranin-fast green. Sections were observed and photographed using a bright-field microscope (OLYMPUS, Japan).

Whole Genome Bisulfite-Sequencing (BS-Seq) Library Construction and Sequencing. Genomic DNA of the three genotypes of *OsMET1-2* (+/+, +/- and -/-) was fragmented by sonication to a mean size of ~ 250 bp, followed by ligation of cytosine-methylated sequencing adaptors. Ligated DNA fragments were treated with the bisulfite reagent of ZYMO EZ DNA Methylation-Gold kit according to manufacturer's instructions. Bisulfite-treated DNAs were PCR amplified and the resultant DNAs were applied to paired-end sequencing with a read length of 90nt for each end using the HiSeq 2000 platform. Sequencing was performed at the Beijing Genomics Institute (BGI), Shenzhen, China. For each genotype, > 10 Gb clean data was generated. The average sequencing depth was 12.5 × per strand with a 99.7% conversion rate, which covered 87.4% of the rice genome on average for all samples (Table S2).

BS-Seq Data Processing. Quality Control, mapping and processing of BS-seq reads were performed as described (5). Briefly, low quality reads were removed from raw data, and then the clean data were mapped to MSU7.0 rice reference genome(ftp://ftp.plantbiology.msu.edu/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs/pseudomolecules/version_7.0/all.dir/) by SOAP2.01 according to the procedures described (6). Only uniquely mapped reads were retained for further analysis. The methylation level for each cytosine site was calculated by dividing the number of converted cytosines (T in plus strand and A in minus strand) by the sequencing depth of this site. The reads aligned to chloroplast genome were used to calculate the conversion rate for each genotype. The conversion rates were > 99.6% in all three genotypes.

Designation of DMCs, DMRs, and DMGs. The Differentially Methylated Cytosines (DMCs) between any two genotypes were defined by binomial test which was used to evaluate the methylation level changes for each cytosine in the rice reference genome against the null hypothesis that the methylation levels were equal between the two compared genotypes. Raw *p*-values were adjusted by the BH method (FDR) and only cytosine sites whose adjusted *q*-values ≤ 0.01 were designated as DMCs between any two genotypes. The Differentially Methylated Regions (DMRs) were discriminated by comparison of methylation levels of 1kbwindows/regions throughout the genome between any two genotypes. Binomial test was used to identify the methylation difference for each window between two genotypes. For the identification of Differentially Methylated Genes (DMGs), the methylation levels of the following regions which included gene body, 1 kb upstream, and 1 kb downstream flanks were compared between any two genotypes by binomial test and FDR adjustments. Genes containing related regions were significantly different (adjusted p-value < 0.05) were defined as DMGs between any two genotypes.

RNA-Seq Library Construction and Sequencing. Two independent libraries of the 11-day-old shoots were constructed for each genotype of *OsMET1-2* (+/+, +/- and -/-) and sequenced by the HiSeq2000. Library construction and HiSeq2000 sequencing were carried out with standard protocols. Raw data were cleaned by removing adaptor contamination and low quality reads. For each library, more than 4Gb clean data (20 million pairs of 100bp reads,Q20 > 90%) were obtained. Clean data have been deposited at the SRA database (http://www.ncbi.nlm.nih.gov/sra/) with accession number SRPXXXXXX.

RNA-seq Data Processing and Designation of Differentially Expressed Genes. Raw RNA-seq data was first cleaned by removing contaminations and low quality reads by

Fastx-tools (http://hannonlab.cshl.edu/fastx_toolkit/). The clean data were mapped against the MSU7.0 rice genome with corresponding annotation by Tophat2.0 (http://tophat.cbcb.umd.edu/) using default parameters. The aligned results were then used to assess the FPKM (Fragments Per Kilobase of transcript per Million mapped reads) and expression differentiation by Cuffdiff v2.0.2(7).

Gene Ontology (GO) Enrichment Analysis. To assay for possible functional relevance of differentially expressed genes of each types between WT and mutant, we performed GO enrichment analysis with GO Slim assignments from MSU7.0 (http://rice.plantbiology.msu.edu/). All GO terms containing differentially expressed genes in each comparison were tested by hypergeometric test(8) and raw *p*-values were adjusted by FDR method. Only GO terms with *q*-value < 0.05 were regarded as significantly enriched.

Small RNA Sequencing, Data Processing and Designation of DSRs. smRNA libraries for the 11-day-old shoot of the three genotypes of *OsMET1-2* (+/+, +/- and -/-) were constructed according to standard workflow of Illumina TrueSeq Small RNA Preparation Kits (http://www.illumina.com/products/truseq_small_rna_sample_prep_kit.ilmn). Raw data were cleaned by removing adaptor contamination and low quality reads.miRNA(http://www.mirbase.org/ftp.shtml), tRNA (http://rice.plantbiology.msu.edu/analyses_search_tRNA.shtml), rRNA (Rice Genome Annotation MSU7.0) and snRNA (http://www.arb-silva.de/no_cache/download/archive/release_111/Exports/) were also removed. Clean reads were aligned to the rice MSU7.0 reference genome by Bowtie2 (9) allowing 0 mismatches. Two types of small RNA abundance was calculated by 21 nt or 24 nt reads number falling into 1 kb windows/regions throughout the whole genomes and windows containing more than 3 reads were tested by binominial test and FDR adjustment between samples. Windows with corrected *p*-value ≤0.05, were defined as differential smRNA regions (DSRs).

1. Zhang H, *et al.* (2013) Persistent whole-chromosome aneuploidy is generally associated with nascent allohexaploid wheat. *Proc Nat Acad Sci USA* 110(9):3447-3452.
2. Pavlopoulou A & Kossida S (2007) Plant cytosine-5 DNA methyltransferases: structure, function, and molecular evolution. *Genomics* 90(4):530-541.
3. Yamauchi T, *et al.* (2008) Alternative splicing of the rice OsMET1 genes encoding maintenance DNA methyltransferase. *J Plant Physiol* 165(17):1774-1782.
4. La H, *et al.* (2011) A 5-methylcytosine DNA glycosylase/lyase demethylates the retrotransposon *Tos17* and promotes its transposition in rice. *Proc Nat Acad Sci USA* 108(37):15498-15503.
5. Li X, *et al.* (2012) Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC genomics* 13(1):300.
6. Li R, *et al.* (2009) SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25(15):1966-1967.
7. Trapnell C, *et al.* (2012) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol* 31(1):46-53.
8. Shi Y-H, *et al.* (2006) Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. *Plant Cell* 18(3):651-664.
9. Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2.*Nat Methods* 9(4):357-359.
10. Rigal M, Kevei Z, Pelissier T, & Mathieu O (2012) DNA methylation in an intron of the IBM1 histone demethylase gene stabilizes chromatin modification patterns. *EMBO J* 31(13):2981-2993.
11. Lister R, *et al.* (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* 133(3):523-536.

12. Cokus SJ, *et al.* (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452(7184):215-219.

Table S1 Summary of differences between *OsMET1-2*-/- and *Arabidopsis met1* in terms of genome-wide DNA methylation, gene expression and siRNA (including 21nt and 24nt) profiles

	OsMET1-2-/- (this study)	*Arabidopsis met1*(refs.10-12)
^{m}CG reduced (relative to WT)	75.7%	98.3%
^{m}CHG reduced (relative to WT)	6.6%	38.0%
^{m}CHH reduced (relative to WT)	43.0%	34.7%
^{m}C retention	(a) RdDM pathway activity; (b) *Os*MET1-1 (the other copy of MET1) and *Os*CMT3 maintaining the ^{m}CG and ^{m}CHG templates, respectively, with possible ectopic regain of some ^{m}CG along with ^{m}CHH mediated by RdDM; (c) The inhibition of rice DNA glycosylase.	(a) RdDM pathway activity; (b) The inhibition of IBM1,leading to specific ^{m}CHG accumulation in gene bodies; (c) The inhibition of DNA glycosylase.
% of total No. of differentially expressed genes	14.13%	3.04%
21ntsiRNA	Increased	Increased
24ntsiRNA	Decrease	Decrease

Table S2 Rate of sodium bisulfite conversion, sequenced and mapped data, and cytosine coverage per strand

Genotype	Conversion rate (%)	Mapped reads(M)	Mapped data (Gb)	Average map rate (%)	Whole genome average coverage depth (x)
OsMET1-2+/+	99.66	104.95	9.45	90.27	25.22
OsMET1-2+/-	99.66	103.46	9.31	88.25	24.87
OsMET1-2-/-	99.65	103.90	9.35	83.79	24.97

Table S3 Average cytosine methylation level of genes and TEs in the three genotypes of *OsMET1-2*(+/+, +/- and -/-). TEs were classified into four types, LTR, LINE, SINE and MITE

	OsMET1-2	^{m}CG (%)	^{m}CHG (%)	^{m}CHH (%)
Gene body	-/-	3.95	7.62	0.59
	+/+	27.35	10.14	1.45
	+/-	26.63	10.46	1.60
Gene Upstream2kb	-/-	8.07	17.87	2.00
	+/+	37.58	21.98	4.46
	+/-	37.83	23.00	5.14
Gene Downstream2kb	-/-	7.53	15.83	1.54
	+/+	37.18	19.20	3.42
	+/-	37.25	20.12	3.86
TE body	-/-	19.01	42.63	5.88
	+/+	82.56	53.40	20.70
	+/-	80.12	54.07	22.87
TE Upstream2kb	-/-	11.95	25.77	1.98
	+/+	52.48	31.73	4.30
	+/-	51.87	32.32	4.83
TE Downstream2kb	-/-	11.88	25.63	1.96
	+/+	52.33	31.62	4.25
	+/-	51.70	32.14	4.75
SINE body	-/-	20.08	42.29	4.85
	+/+	86.91	53.84	18.10
	+/-	84.32	54.47	20.66
SINE Upstream2kb	-/-	8.98	18.25	1.75
	+/+	43.05	23.23	4.40
	+/-	42.68	23.90	5.05
SINE Downstream2kb	-/-	8.93	18.14	1.86
	+/+	40.55	21.89	4.33
	+/-	40.49	22.63	4.94
LINE body	-/-	14.25	38.26	1.72
	+/+	69.89	41.91	3.90
	+/-	67.85	42.18	4.21
LINE Upstream2kb	-/-	10.42	24.41	1.39
	+/+	50.70	27.94	2.78
	+/-	49.97	28.54	3.14
LINE Downstream2kb	-/-	10.87	25.18	1.40
	+/+	51.34	28.45	2.78
	+/-	50.62	28.97	3.13
LTR body	-/-	21.78	50.46	2.93
	+/+	88.97	63.89	4.18
	+/-	87.38	64.18	4.57
LTR Upstream2kb	-/-	18.81	42.62	2.18
	+/+	77.04	53.54	3.72
	+/-	76.08	54.14	4.12
LTR Downstream2kb	-/-	18.47	42.22	2.11
	+/+	76.93	53.68	3.68
	+/-	75.97	54.14	4.05
MITE body	-/-	18.30	38.41	8.93
	+/+	81.13	50.26	45.92
	+/-	77.92	51.64	51.42

MITE Upstream2kb	-/-	7.46	14.52	1.74
	+/+	36.01	17.22	4.54
	+/-	35.56	17.77	5.13
MITE Downstream2kb	-/-	7.53	14.85	1.81
	+/+	36.28	17.45	4.72
	+/-	35.83	17.98	5.38

Table S4 Total numbers of DMCs and DMRs of *OsMET1-2*-/- vs. WT in genes and TEs

		C	CG	CHG	CHH
DMCs	hyper	7226884	619252	1004378	5780381
	hypo	9267495	8256013	739239	805072
All DMRs	hyper	3344	208	14912	21086
	hypo	296899	274393	114273	165102
Gene-DMRs	hyper	1320	99	4205	2980
	hypo	106527	89612	22161	44882
TE-DMRs	hyper	2402	164	12367	19145
	hypo	232443	220580	105737	151887

Table S5 Differentially expressed genes and TEs of *OsMET1-2*+/- vs. +/+ and of *OsMET1-2*-/- vs. +/+ [The *Arabidopsis met1* data was included as a reference (11)]

	OsMET1-2 +/- vs.+/+	*OsMET1-2* -/- vs.+/+	*Arabidopsis met1* vs. WT
Total number of differentially expressed genes	572	4,338	972
Up-regulation	135	2,269	612
Protein-coding genes	121	1,780	331
TEs	14	489	281
Down-regulation	437	2,069	360
Protein-coding genes	432	1,964	357
TEs	5	105	3
% of total number of differentially expressed genes	1.86%	14.13%	3.04%
Up-regulation	0.44%	7.39%	1.91%
Protein-coding genes	89.63%	78.45%	54.08%
TEs	10.37%	21.55%	45.92%
Down-regulation	1.42%	6.74%	1.13%
Protein-coding genes	98.86%	94.93%	
TEs	1.14%	5.07%	0.83%
Total genes	55,986		
Expressed genes	30,709		
Expressed protein-coding genes	27,993		
Expressed TEs	2,716		

Table S6 Primers used for pyrosequencing-based identification of heterozygotes and homozygotes of *OsMET1-2* mutant, and primers used for RT-PCR detection or qRT-PCR quantification of expression of a set of DNA methylation/chromatin-related genes in the *OsMET1-2* mutant vs. WT

P1L	5'-acttctgaggctgctgaac	For sequencing (*OsMET1-2*)
P1R	5'-ccaccatttatgaattctacttc	Biotin-labeled (*OsMET1-2*)
P2L	5'-atgtggagatgctgatgattg	*OsMET1-2* (for RT-PCR)
P2R	5'-cctctaatcttcttctggaacc	
P3L	5'-cgtgtgcgataatggaactg	*OsActin* (for RT-PCR)
P3R	5'-tctgggtcatcttctcacga	
P4L	5'-gcatgtgcttccatcctgag	*OsMET1-1* (for qRT-PCR)
P4R	5'-atctgcctgtgcttgttctg	
P5L	5'-atccgaatcaggcgagagtt	*OsCMT3* (for qRT-PCR)
P5R	5'-gcacgagcaacaggtacag	
P6L	5'-ttgctgttgaagagtgcttga	*OsCMT2* (for qRT-PCR)
P6R	5'-cttgaggcgtcttgctgac	
P7L	5'-gacacctacattcctaacattgg	*OsDNMT2* (for qRT-PCR)
P7R	5'-tcagcgacattcagacttattg	
P8L	5'-cgtgcggcatcttactactga	*OsDRM2* (for qRT-PCR)
P8R	5'-atctcggtgatggcggttg	
P9L	5'-tgacttcctagagaatccacagat	*Putative OsVIM* (for qRT-PCR)
P9R	5'-ccagatgactcagcaccattg	
P10L	5'-ggaggaggcagatagcagaat	*OsDDM1a* (for qRT-PCR)
P10R	5'-ccgctgaagtaagattgatacca	
P11L	5'-ggaggaggcagatagcagaat	*OsDDM1b* (for qRT-PCR)
P11R	5'-acaggtatcagccgaagtaaga	
P12L	5'-cctggacctgaagagaagagtt	*DNG702* (for qRT-PCR)
P12R	5'-ggctggcgattgttcatacc	
P13L	5'-ctgttgctgcttggtggtg	*OsHDAC1* (for qRT-PCR)
P13R	5'-tggaggcatcttgtcagtgag	
P14L	5'-aatcactcgtcgaatcctaactg	*OsOSiEZ1* (for qRT-PCR)
P14R	5'-tcaggtccatagcggtaatcat	
P15L	5'-gctcggctcaataatccacaa	*OsFIE1* (for qRT-PCR)
P15R	5'-cacttcatcccaacgccatat	
P16L	5'-agatggcttggtgacttcattc	*OsFIE2* (for qRT-PCR)
P16R	5'-cacattctggcacaggatactt	
P17L	5'-tacagcagcacaggcatca	*DNG704* (for qRT-PCR)
P17R	5'-tcaacagaatctcttcggtcaa	
P18L	5'-cgagaacgaggaatgaataatgtc	*DNG701* (for qRT-PCR)
P18R	5'-ttgctgagtctggaggaaca	
P19L	5'-actgctactcctggaagactc	*DNG712* (for qRT-PCR)
P19R	5'-ggctcacacttatagatggattct	
P20L	5'-accacttcgaccgccactact	*OsUBQ5* (for qRT-PCR)
P20R	5'-acgcctaagcctgctggtt	

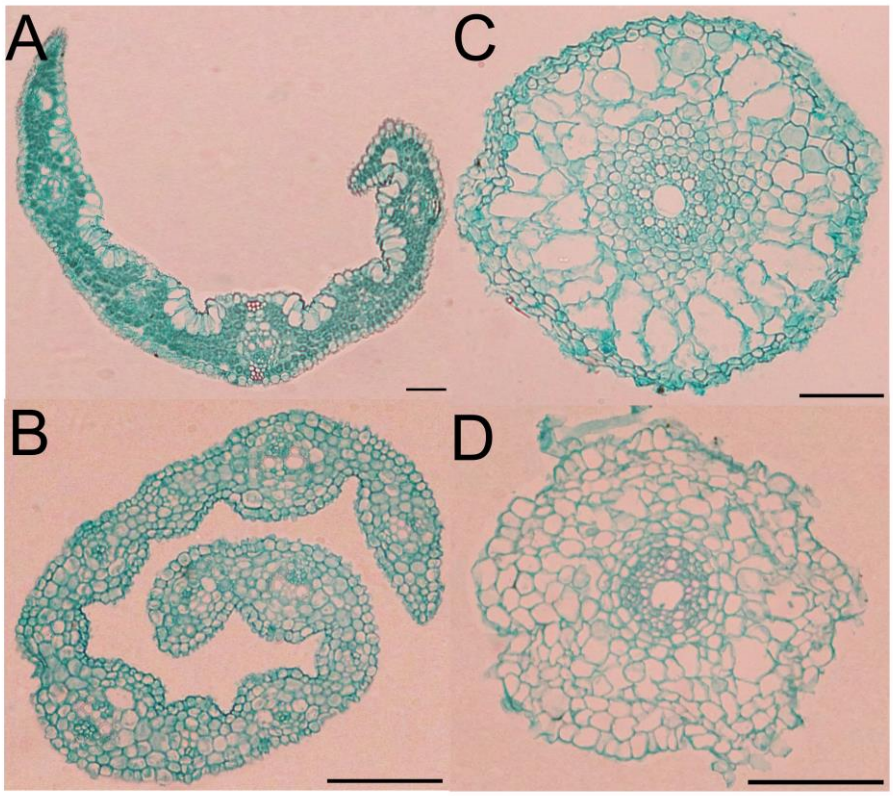

Fig. S1. Abnormal cell differentiation and specification in the *OsMET1-2* mutant (-/-) compared with WT. (*A*) and (*B*) are 8-μmcross sections of paraffin-embedded leaves of *OsMET1-2*-/- and WT, respectively. (*C*) and (*D*) are 8-μm cross sections of paraffin-embedded roots of *OsMET1-2*-/- and WT respectively. Scale bar = 50μm.

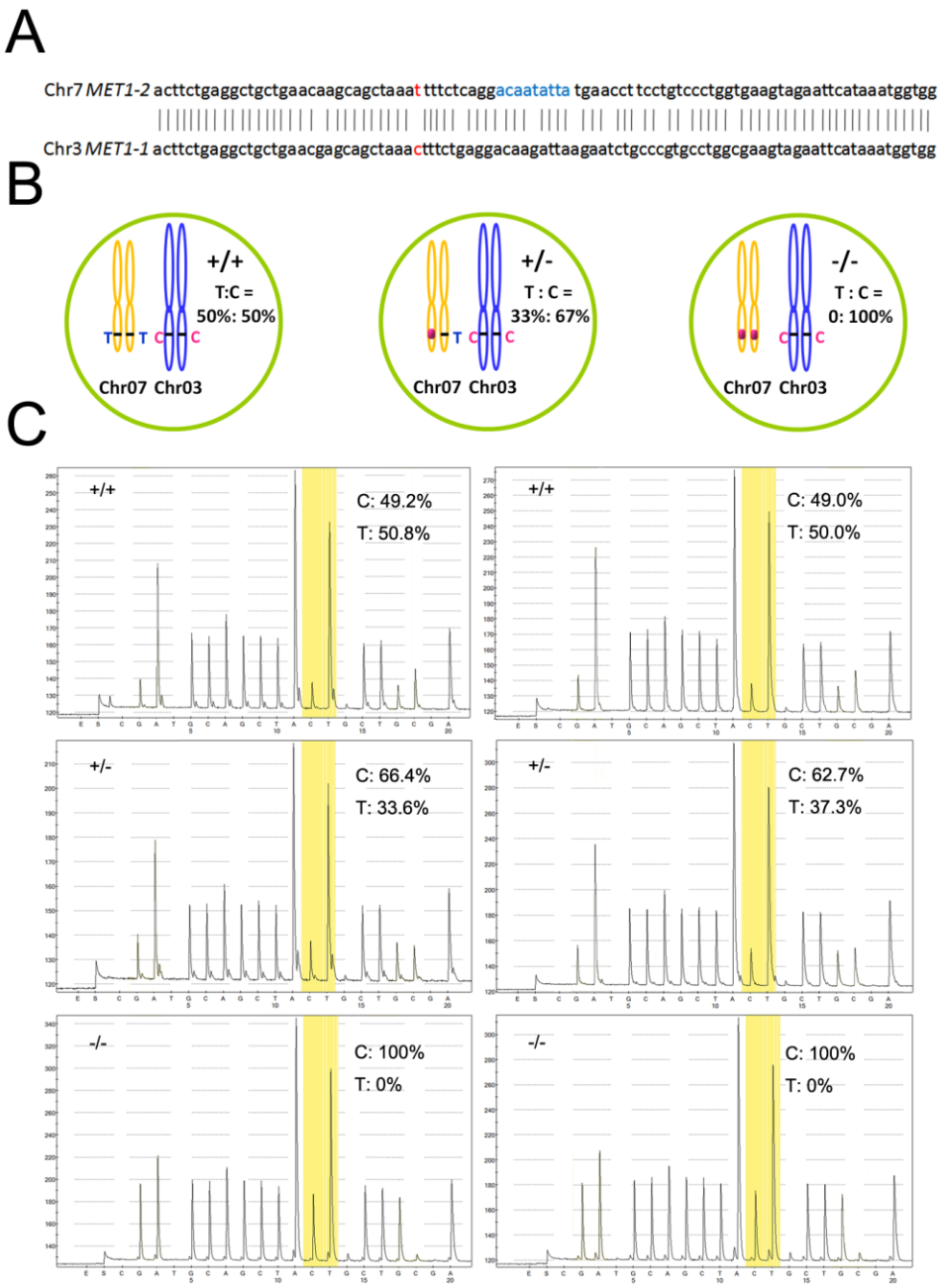

Fig. S2. Discrimination of the three genotypes of *OsMET1-2*, WT (+/+), heterozygote (+/-) and homozygote (-/-), by locus-specific pyrosequencing based on a single nucleotide polymorphism (SNP) existing between the two *MET1* genes of rice (cv. Nipponbare), *OSMET1-1* and *OsMET1-2*. (*A*) A region adjacent to the *Tos17* insertion site (blue font) in *OsMET1-2* (Fig. 1*A*), which harbors a SNP (T/C, red font) upon which the two genes could be distinguished by pyrosequencing. (*B*) The expected ratios of T: C in pyrosequencing results of the three *OsMET1-2* genotypes. (*C*) Pyrosequencing-based genotyping of the three types of plants *OsMET1-2* (+/+, +/- and -/-).

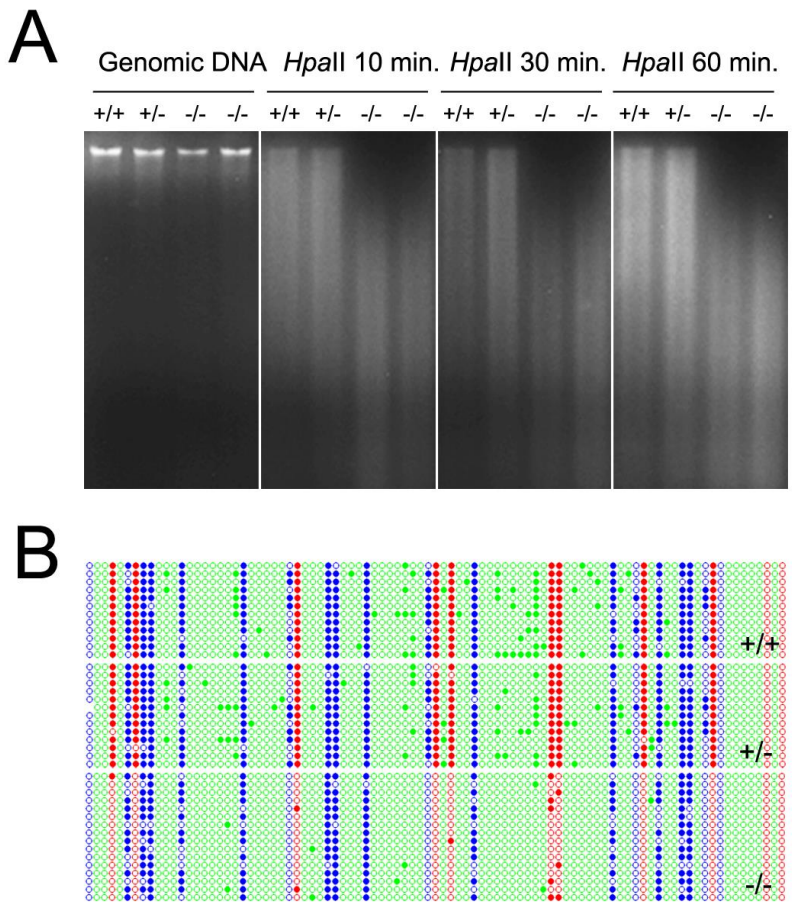

Fig. S3. Sensitivity to methylation-sensitive endonuclease digestion, and bisulfite sequencing analysis of a body region of *Tos17* of the three genotypes of *OsMET1-2* (+/+, +/- and -/-). (*A*)Agarose gel electrophoresis of intact genomic DNA (left-most lane) and DNA digested by *Hpa*II at three time points. (*B*) Bisulfite sequencing of a body region of *Tos17*,which has been established as a reliable marker for assaying cytosine methylation liability in rice (4). Each circle represents a cytosine residue, either methylated (solid) or non-methylated (empty). Cytosines are color-coded by their sequence context: red for CG, blue for CHG and green for CHH (H = C, T or A). Each row represents an independent clone.

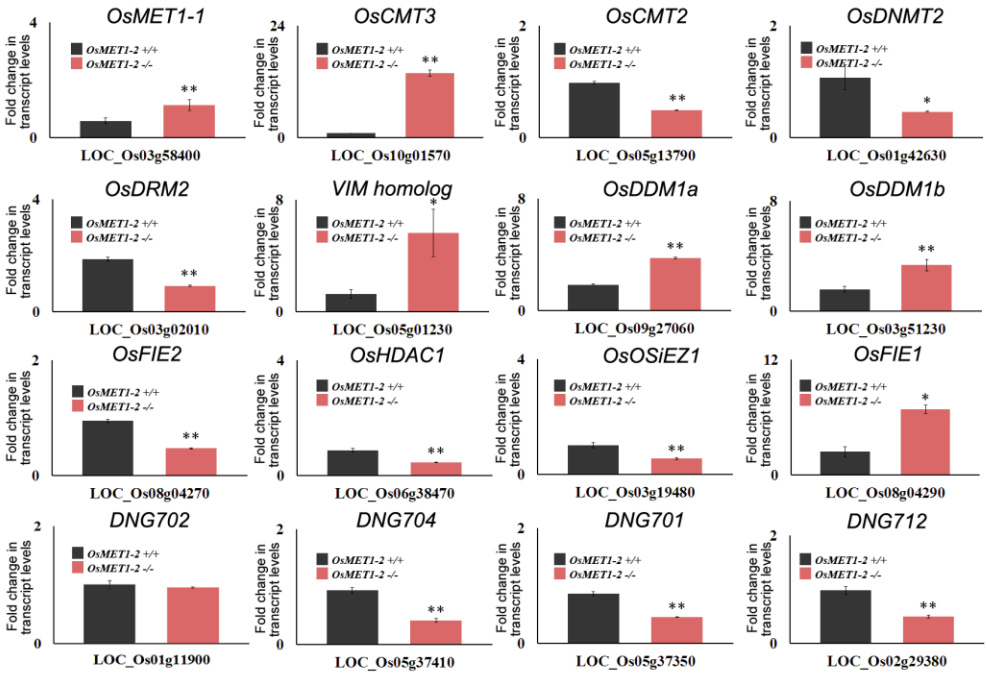

Fig. S4. Fold change in transcript levels of 16 chromatin-related genes known or supposedly related to cytosine methylation, based on qRT-PCR quantifications between the *OsMET1-2*mutant (-/-) and WT. These genes also showed significantly different transcript abundance between the two genotypes based on the RNA-seq data (*Datasets S1* and *S2*).

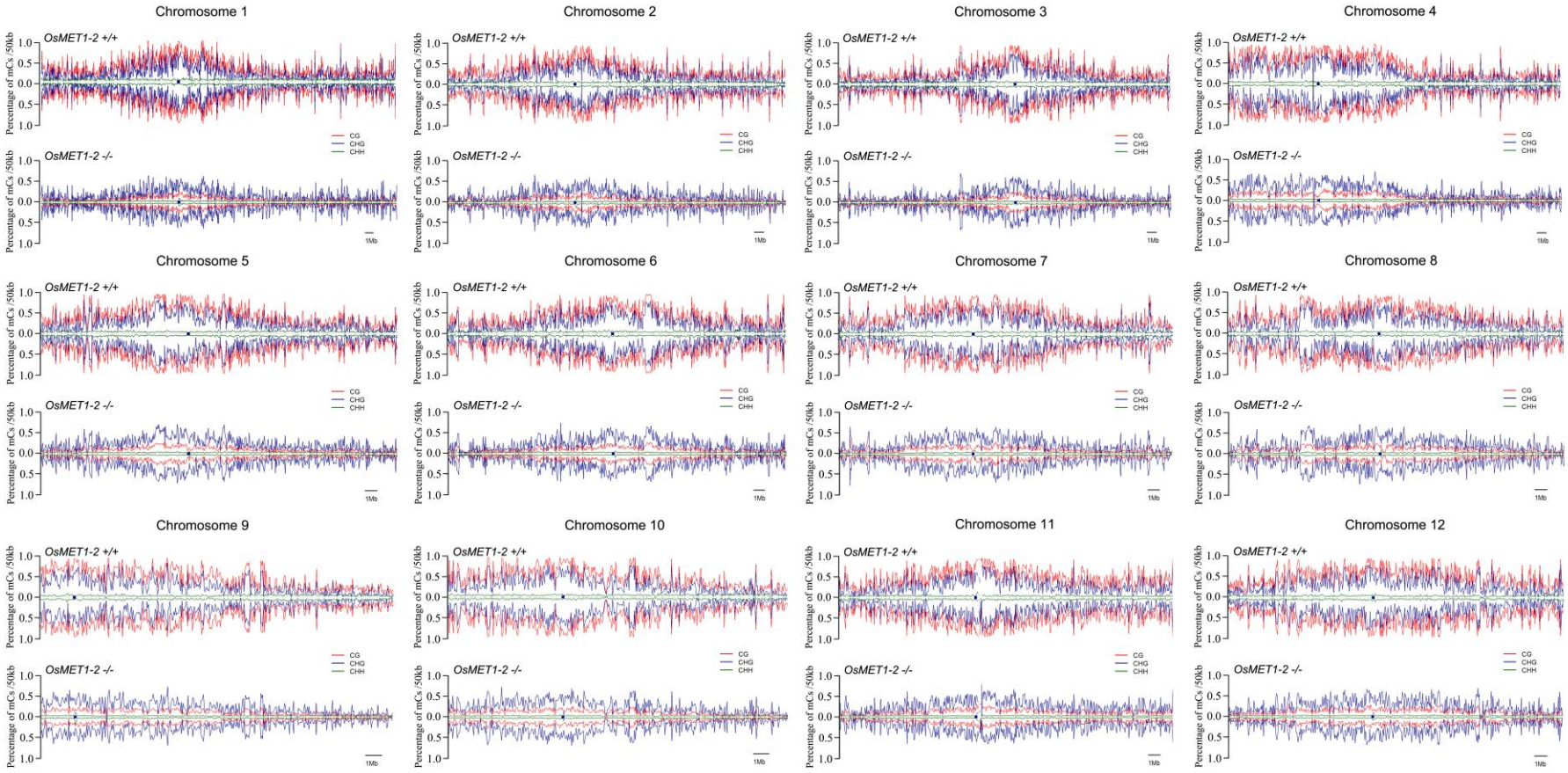

Fig. S5. Chromosomal distribution of ^{m}Cs in each sequence context (CG, CHG and CHH) in WT and the homozygous mutant of *OsMET1-2* along each of the of the 12 rice chromosomes. Percentages of ^{m}Cs out of total Cs (*y*-axis) in each sequence context were computed in 50kb sliding windows.

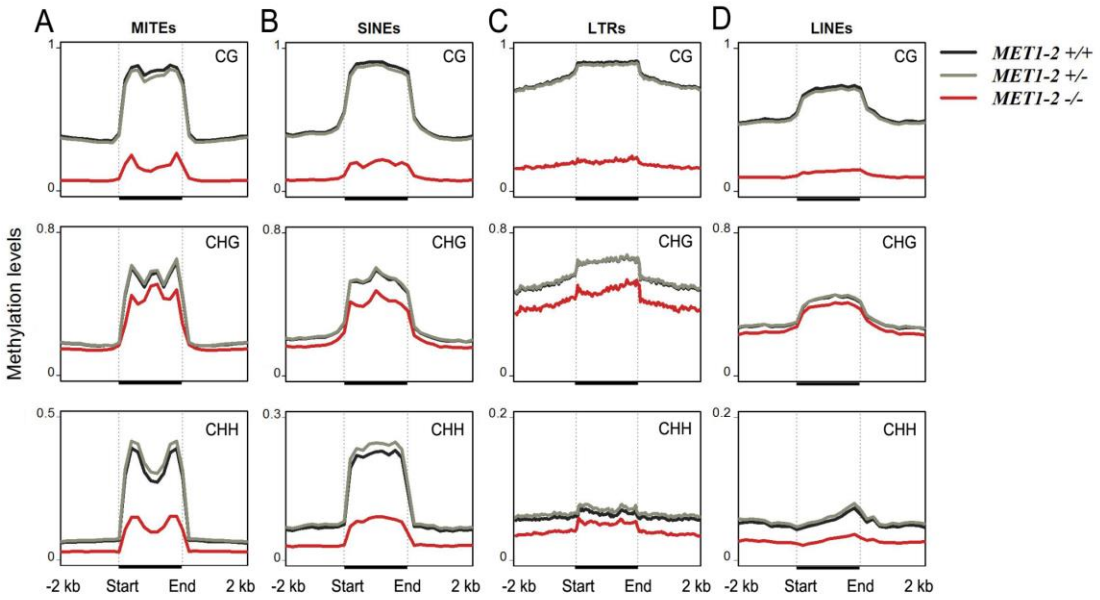

Fig. S6. Genomewide average ^{m}C levels in each sequence context (CG, CHG and CHH) of four types of TEs (MITEs, SINEs, LINEs, and LTR-retrotransposons) in 5% bins, each with their contiguous upstream and downstream 2 kb flanks in the three *MET1-2*genotypes (+/+, +/- and -/-).

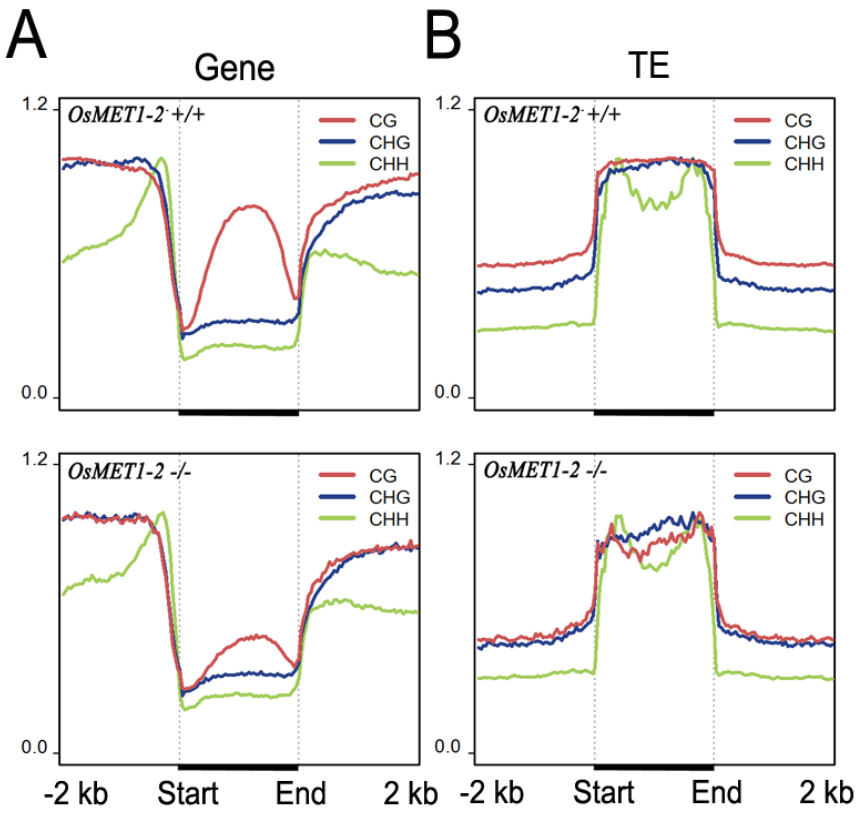

Fig. S7. Genomewide average ^{m}C levels in each sequence context (CG, CHG and CHH) of genes (*A*) and TEs (*B*), each with their contiguous upstream and downstream 2kb flanks in the three *MET1-2* genotypes (+/+, +/- and -/-). Values are normalized against the highest value of each profile.

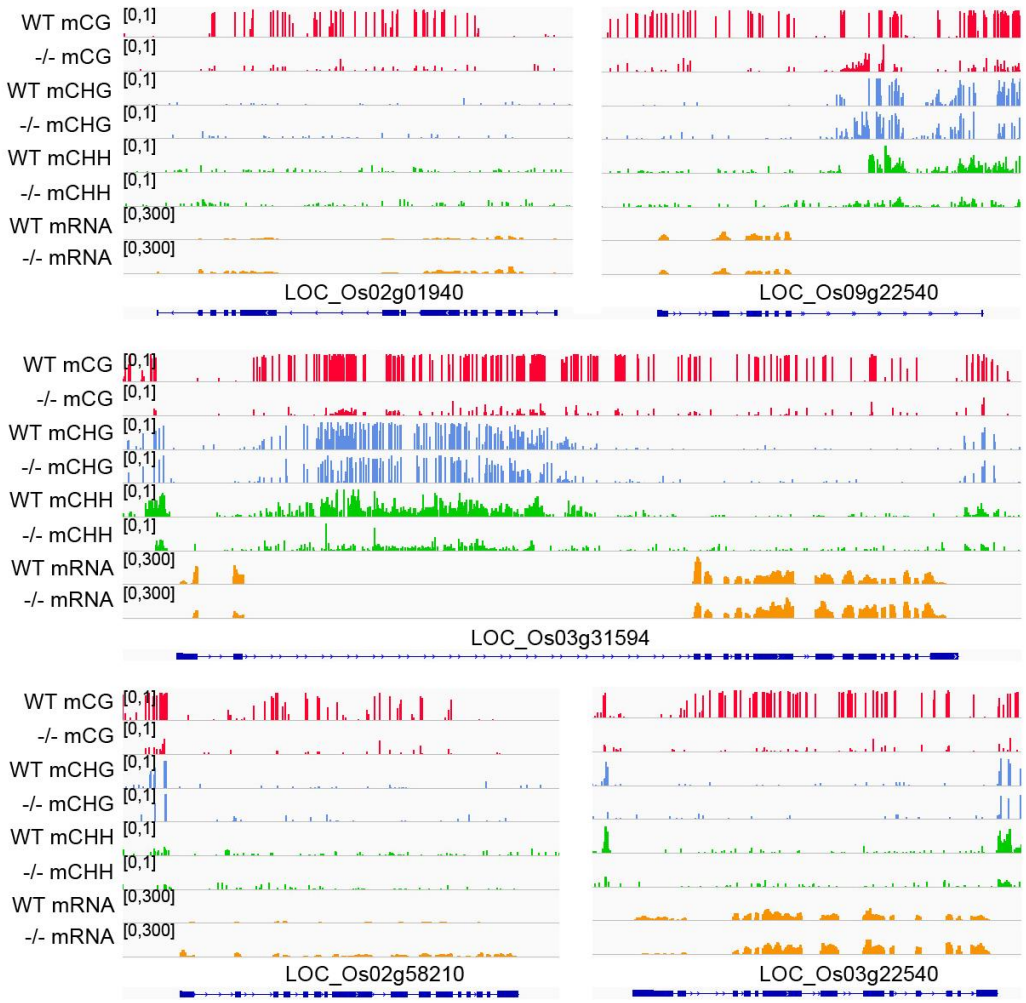

Fig. S8. DNA methylation level (based on methylome) and mRNA abundance (based on RNA-seq) of five JmjC domain-containing genes which were homologous to *Arabidopsis* IBM1in WT and *OsMET1-2*-/-.

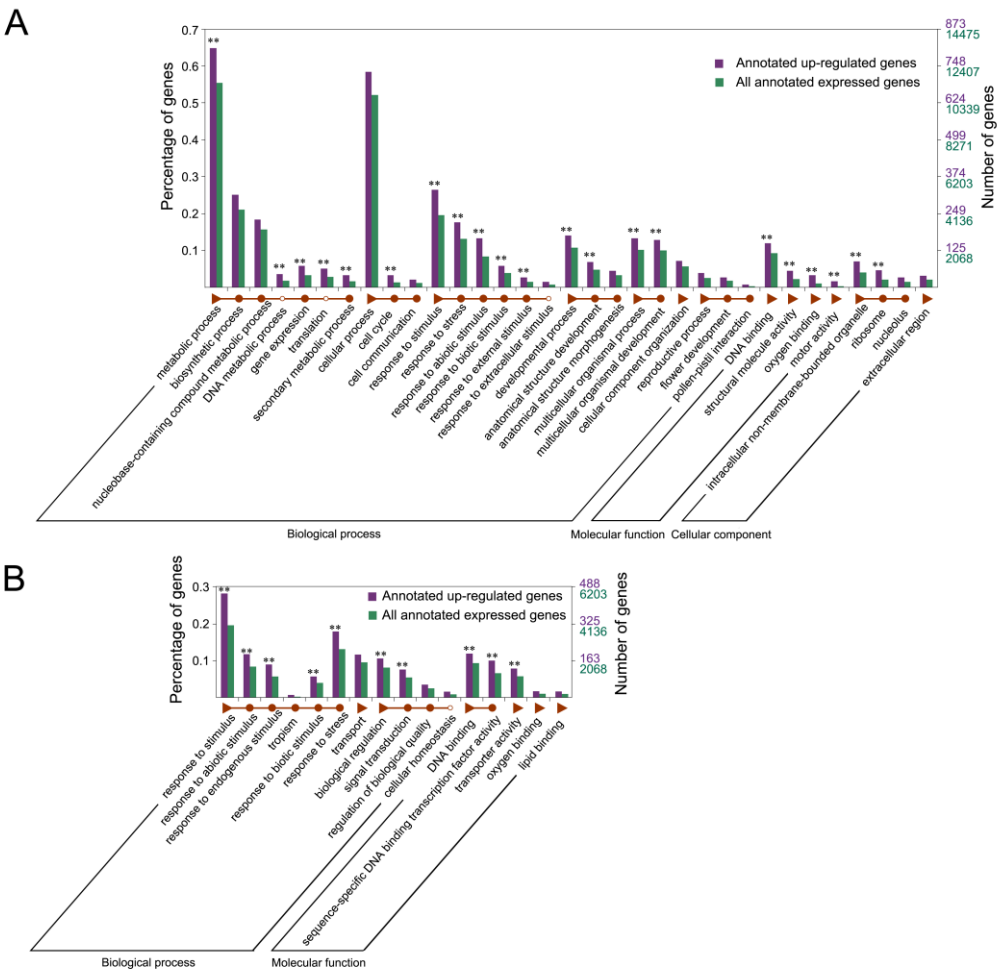

Fig. S9. Gene ontology (GO) analysis of the differential expressed genes (*q*-value < 0.05) in the homozygous mutant (*OsMET1-2*-/-) compared to WT. The *y*-axis is the percentage of genes mapped to GO category terms: the percentages were calculated by the number of genes mapped to a given particular GO category divided by the number of all annotated genes in the category. The *x*-axis is the GO category terms, which were ordered by their affiliation. Triangles, solid circles and empty circles represent the first, second and minimal levels of hierarchy, respectively. Green bars denote percentages for each category of all the annotated genes (out of 20,678 annotated), and purple bars denote percentages of all the annotated differentially expressed genes in each GO category. (*A*) Significantly enriched up-regulated genes (out of 1,247 annotated). (*B*) Significantly enriched down-regulated genes (out of 1,625 annotated). **Significant at *q*-value < 0.01.

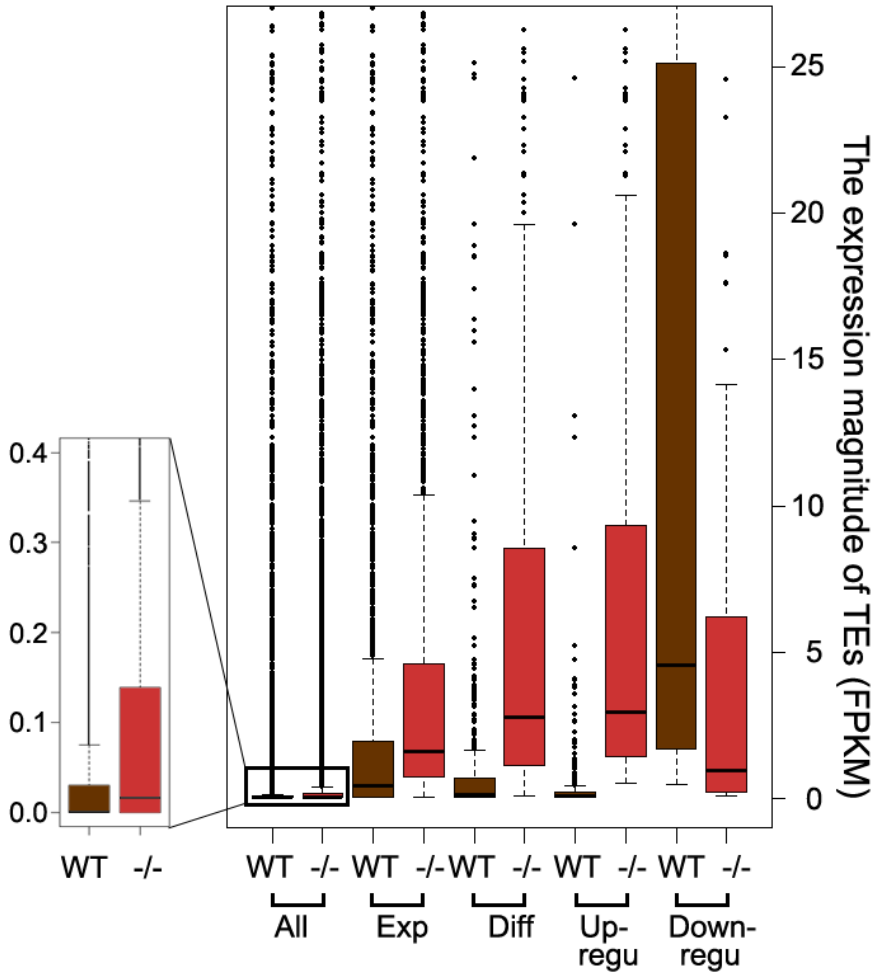

Fig. S10. Box plot showing transcript level of TEs (based on FPKM of RNA-seq data) between WT and *OsMET1-2*-/-. For convenience of observation, the transcript level of all the annotated TEs (according to MSU7.0 rice reference genome) was enlarged and shown in the left panel.

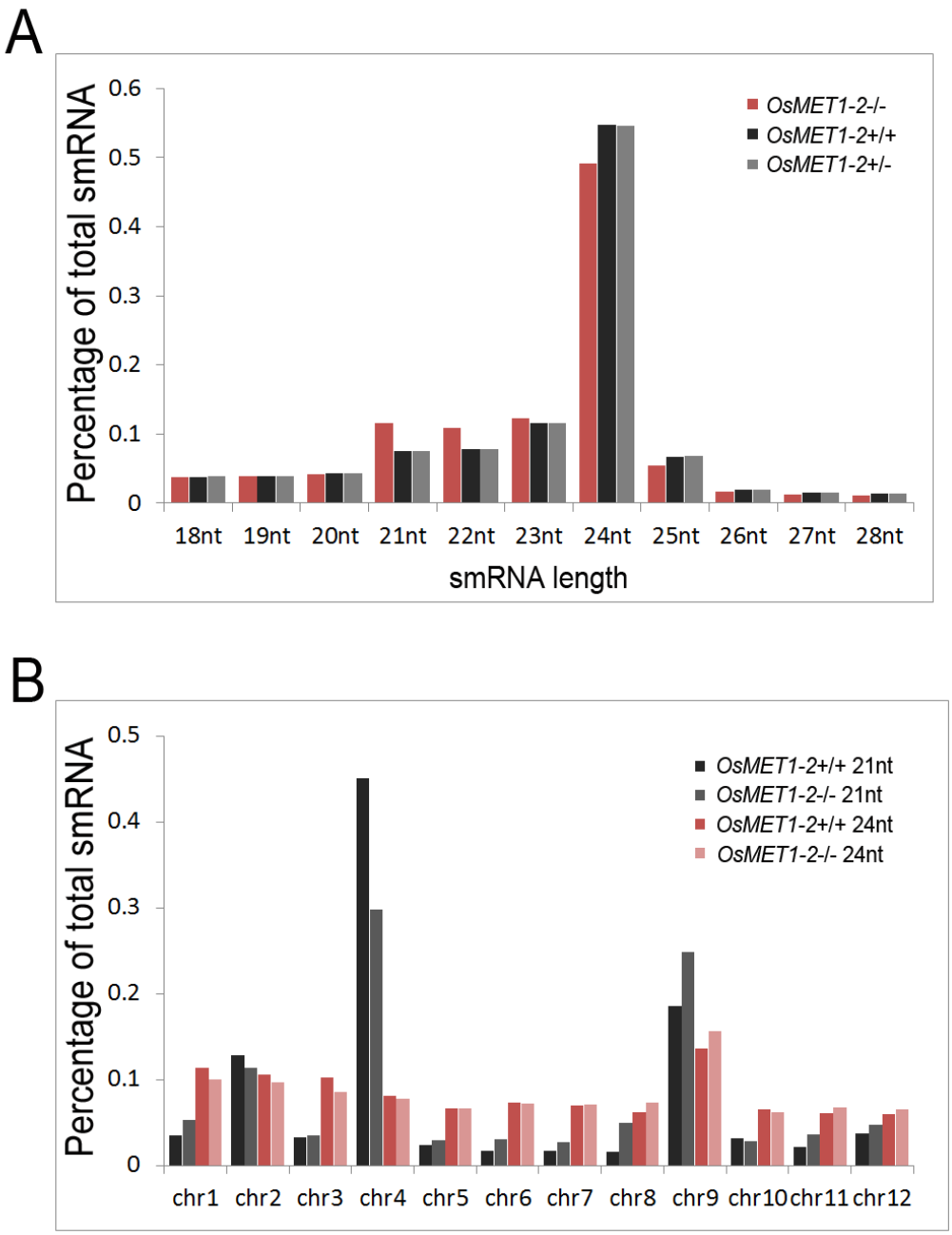

Fig. S11. Global profiles and chromosomal distribution of smRNAs based on smRNA-seq. (*A*) Global profiles of smRNAs of various size groups ranging from 18-28 nt of the three genotypes(+/+, +/- and -/-) of *OsMET1-2*.(*A*) Percentage of smRNAs of each size group(18-28 nt) relative to the total numbers of smRNA reads in the three genotypes of*OsMET1-2*. (*B*) Chromosomal distribution of 21nt and 24ntsmRNAsin each of the 12 rice chromosomes between WT and homozygous *OsMET1-2*mutant (-/-). Significant difference in abundance of both size groups was detected between the two genotypes in each chromosome (prop. test, $P < 0.01$).

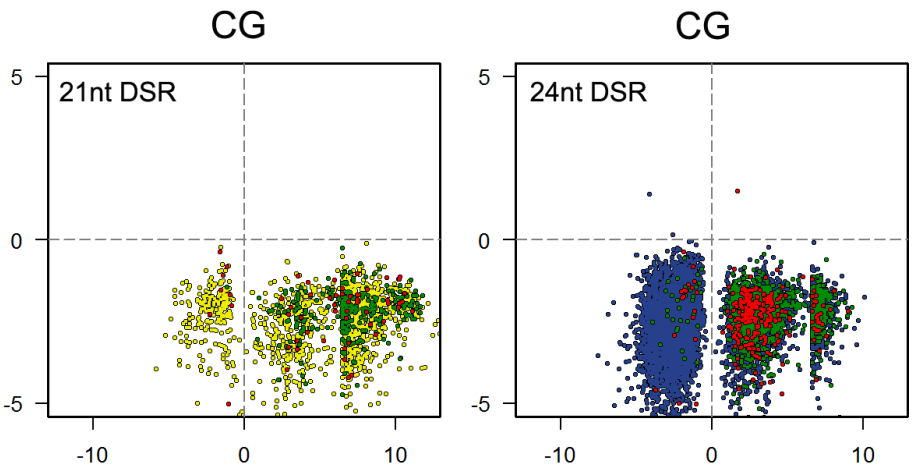

Fig. S12. Co-localization of the two classes (21 nt and 24 nt) of differential smRNA regions (DSRs) that at the same time were also CG DMRs in terms of relative smRNA abundance (*x*-axis, presented as $\log_2$FC of DSRs) and DNA methylation level (*y*-axis, presented as $\log_2$FC of CG DMRs) between mutant and WT. The red and green dotes denote for CHH hyper-DSRs which were also CG DSRs mapped to genes and TEs, respectively.

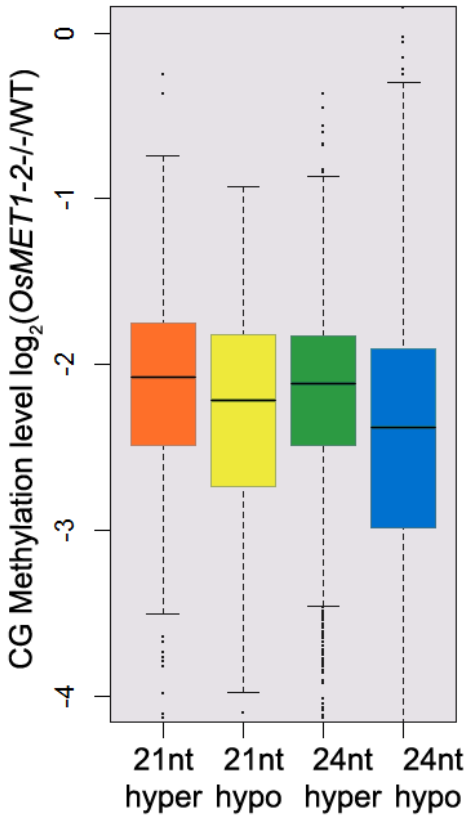

Fig. S13. Box plot showing different CG methylation levels in the 21 nt and 24 nt CHH-DSRs between *OsMET1-2*-/- and WT. The abbreviations "hyper" and "hypo" refer to CHH-hyper DMRs and CHH-hypo DMRs, respectively.