Cooperation between the H3K27me3 Chromatin Mark and Non-CG Methylation in Epigenetic Regulation

Shaoli Zhou1, Xiaoyun Liu2, Chao Zhou2, Qiangwei Zhou, Yu Zhao, Guoliang Li, and Dao-Xiu Zhou*

National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, 430070 Wuhan, China (S.Z., X.L., C.Z., Q.Z., Y.Z., G.L., D.-X.Z.); Institute for Interdisciplinary Research, Jianghan University, 430056 Wuhan, China (X.L.); and Institute of Plant Sciences of Paris-Saclay (IPS2), Université Paris-Saclay, Université Paris-sud 11, 91405 Orsay, France (D.-X.Z.).

ORCID IDs: 0000-0001-9111-1598 (S.Z.); 0000-0002-5740-9777 (X.L.); 0000-0003-4451-9907 (Y.Z.); 0000-0002-1540-0598 (D.-X.Z.).

H3K27me3 is a repressive chromatin mark of genes and is catalyzed by homologs of Enhancer of zeste [E(z)], a component of Polycomb-repressive complex 2 (PRC2), while DNA methylation that occurs in CG and non-CG (CHG and CHH, where H is A, C, or T) contexts is a hallmark of transposon silencing in plants. However, the relationship between H3K27me3 and DNA methylation in gene repression remains unclear. In addition, the mechanism of PRC2 recruitment to specific genes is not known in plants. Here, we show that SDG711, a rice (Oryza sativa) E(z) homolog, is required to maintain H3K27me3 of many developmental genes after shoot meristem to leaf transition and that many H3K27me3-marked developmental genes are also methylated at non-CG sites in the body regions. SDG711-binding and SDG711-mediated ectopic H3K27me3 also target genes methylated at non-CG sites. Conversely, mutation of OsDRM2, a major rice CHH methyltransferase, resulted in loss of SDG711-binding and H3K27me3 from many genes and their de-repression. Furthermore, we show that SDG711 physically interacts with OsDRM2 and a putative CHG methylation-binding protein. These results together suggest that the repression of many developmental genes may involve both DRM2-mediated non-CG methylation and PRC2-mediated H3K27me3 and that the two marks are not generally mutually exclusive but may cooperate in repression of developmentally regulated genes in rice.

Histone Lys methylation is an important epigenetic modification for regulating genome activity and gene expression. Histone Lys residues can be mono-, di-, or trimethylated and may have either activating or repressive functions in gene expression, depending on the position and methylation level of lysines. For instance, in plants, dimethylation of histone H3 Lys 9 (H3K9me2) is almost exclusively associated with heterochromatin regions and is required for repression of repetitive sequences, whereas trimethylation of H3 Lys 27 (H3K27me3) preferentially marks repressed or lowly expressed genes (Liu et al., 2010, Chen and Zhou, 2013). H3K27me3 has been implicated in gene regulation of many plant developmental pathways and H3K27me3-marked genes often exhibit a high degree of tissue specificity (Zhang et al., 2007; Kohler and Villar, 2008; Zheng and Chen, 2011; Makarevitch et al., 2013), consistent with the function of H3K27me3 in maintaining gene repression during growth. Whole genome chromatin immunoprecipitation (ChIP) experiments indicate that about 20% of plant genes are marked by H3K27me3 (Wang et al., 2009; He et al., 2010; Lafos et al., 2011; Liu et al., 2015). There is a large tissue-specific variation of H3K27me3, with many genes shown to gain or lose H3K27me3 during cell differentiation, demonstrating a dynamic regulation of this epigenetic modification in responding to both developmental and environmental signals in plants (Wang et al., 2009; He et al., 2010; Lafos et al., 2011; Li et al., 2013; Liu et al., 2015).

Polycomb group (PcG) proteins, including Enhancer of Zeste [E(z)], Extra sex combs (Esc), Su(z)12, and Nurf55 in Drosophila (Drosophila melanogaster) form the Polycomb-repressive complex 2 (PRC2), which catalyzes H3K27me3 (Schwartz et al., 2006). Plant homologs of the four core components of PRC2 have been identified. The rice genome contains two genes encoding homologs of E(z) (i.e. CLF or SDG711 and EZ1 or SDG718), Su(Z)12 (i.e. EMF2a and EMF2b), and ESC (i.e. FIE1 and FIE2). The function of rice PRC2 homologs have been studied (Luo et al., 2009; Zhang et al., 2012a; Liu et al., 2014, 2015). For instance, down-regulation of the rice E(z) homolog SDG711 results in
genome-wide loss of H3K27me3 and affects many plant developmental processes (Liu et al., 2014, 2015). It remains unclear how plant PR2 targets specific genes to mediate gene repression.

In plants, DNA cytosine methylation that occurs in CG, CHG, and CHH contexts (where H is A, C, or T) is a hallmark of inactivation of transposable elements (TE) and silent sequences (Zhanga et al., 2006). CG methylation is common in gene bodies of expressed genes, while methylation at all sequence contexts in the promoter region represses gene transcription (Zemach et al., 2010). In Arabidopsis (Arabidopsis thaliana), CG methylation is maintained by METHYLTRANSFERASE1 (MET1) and non-CG (CHG and CHH) methylation is maintained by the redundant activities of additional DNA methyltransferases such as DOMAINS REARRANGED METHYLASE 2 (DRM2), CHROMOMETHYLASES2 and 3 (CMT2 and CMT3; Stroud et al., 2013, 2014; Zemach et al., 2013). DRM2 is guided by small interfering RNAs (siRNAs) to methylate homologous DNA regions through the process of RNA-dependent DNA methylation (RdDM; Law et al., 2010). Recent data indicated that RdDM is responsible for only a small fraction of CHH methylation, while CMT2 is the major CHH methyltransferase in Arabidopsis (Zemach et al., 2013; Stroud et al., 2014). CMT3 catalyzes CHG methylation in cooperation with H3K9me2 (Du et al., 2012; Stroud et al., 2014). It is established that CHG methylation and H3K9me2 form a positive feedback loop: CHG methylation by CMT3 requires H3K9me2, to which SUVH proteins bind via the SRA (SET and Ring Associated) domain (Law and Jacobsen, 2010; Du et al., 2012). H3K9 dimethylation by SUVH (Suppressor of variegation 3-9 homolog) proteins requires CHG methylation, to which SUVH proteins bind via the SRA (SET and Ring Associated) domain (Johnson et al., 2007; Rajakumara et al., 2011).

Compared to Arabidopsis, the rice genome displays a different global DNA methylation landscape and has a much higher level of genome-wide DNA methylation (Li et al., 2012; Tan et al., 2016). In addition, CHH methylation is enriched in euchromatic regions of the rice genome and plays an important role in gene regulation (Tan et al., 2016). However, the relationship between H3K27me3 and DNA methylation and the mechanism by which PR2 targets to specific loci in rice remain unclear. In this work, we show that many functional genes methylated at non-CG sites are targeted by SDG711-mediated H3K27me3. Further analysis suggests that SDG711 may be recruited to subset of non-CG methylated genes by interaction with OsDMR2 and/or SDG703, a member of SUVH proteins that recognize methylated CHG (Qin et al., 2010; Zhou and Hu, 2010; Zhao and Zhou, 2012). Our results indicate that non-CG methylation and H3K27me3 are not generally exclusive for epigenetic modification of rice genes and that DNA methylation and H3K27me3 may cooperate to repress a set of developmental genes in rice.

RESULTS

**SDG711 Is Involved in H3K27me3 Reprogramming during SAM to Leaf Transition**

We have previously determined genome-wide H3K27me3 in rice shoot apical meristems (SAMS; Liu et al., 2015). To study changes of H3K27me3 after SAM to leaf transition, we analyzed the genomic distributions of the mark in 12-d-old seedling leaves of wildtype and SDG711 over-expression (711OX) plants by ChIP-seq (Fig. S1A; Supplemental Table S1) (Liu et al., 2014). To minimize effect of callus culture during transgenic plant production, callus-regenerated wild-type seedling leaves were used. About 20–40 million clean sequence reads, after filtered by the Trimmomatic software (version 0.32), were aligned to the rice genome (RGAP version 7.0). Threshold of read coverage defined by randomization (Q value < 0.001) was used to identify H3K27me3-marked genes. About 78.1% (7,113 out of 9,106) of the H3K27me3-marked genes overlapped with those identified in rice leaves in a previous study (Guo et al., 2015; Supplemental Fig. S1B). In addition, contour plots of TPKM of the two ChIP-seq datasets revealed a correlation coefficient equal to 0.669 (Supplemental Fig. S1C), indicating the reliability of the ChIP-seq data. When compared with shoot apical meristem (Liu et al., 2015), we observed that 3,479 genes lost H3K27me3 while 1,361 genes gained the mark in leaves (Fig. 1A). By contrast, in 711OX leaves, 2,428 genes gained H3K27me3 and 1,049 genes lost the mark compared to wild-type leaves (Fig. 1A). Most (1,549/2,428) of the genes that gained H3K27me3 in 711OX leaves corresponded to those that lost the mark after SAM to leaf transition in wild type (Fig. S2A; Supplemental Fig. S2), suggesting that loss of H3K27me3 in leaves might be due to reduced expression of SDG711 observed in leaves compared to SAM (Liu et al., 2015).

We found that many genes encoding transcription factors and growth hormone (cytokinin and auxin) signaling proteins gained or lost H3K27me3 after SAM to leaf transition (Fig. 1B; Supplemental Table S2). Among those transcription factor genes that gained H3K27me3 in leaves, some have been reported to be expressed in SAM and to regulate meristem functions (e.g. OSH6, OsKN2, OsMADS22, OsMDP1; Postma-Haarsma et al., 1999; Sentoku et al., 1999, 2000, 2005; Duan et al., 2006; Lee et al., 2008). In addition, we found that most of the transcription factors (59/81), cytokinin (8/9), and auxin (5/5)-related genes that gained the mark in 711OX leaves corresponded to those that lost the mark after SAM to leaf transition in wild type (Fig. 1B; Supplemental Table S2). Quantitative ChIP-PCR and RT-PCR analysis of a total of 26 genes validated the ChIP-seq data and indicated that higher levels of H3K27me3 in SAM or in 711OX leaves repressed the genes (Fig. 1C and D). The analysis revealed that the expression level of SDG711 was important for reprogramming of H3K27me3 and developmental gene expression during SAM to leaf transition.
H3K27me3 Marks Genes Methylated at Non-CG Sites

To study genic DNA methylation in rice leaves, we analyzed the genome-wide cytosine methylation of the same leaf materials as for H3K27me3 analysis by BS-Seq. We obtained about 50 million clean reads (125 bp per read) per sample. About 47%–60% of the read tags were uniquely mapped to the genome using the BatMeth software. After removing PCR-amplified redundancy, 65% and 82% of total reads in 711OX and wild type, respectively, were kept for further analysis (Supplemental Table S1). The genome coverage was, respectively, 16.4 and 11.04 for wild type and 711OX from the redundancy-removed reads (Supplemental Table S1), indicating a deep sequencing and a high quality of the methylome. The error conversion rates of the BS sequences were below 1.75% for CG, 1.23% for CHG, and 0.73 for CHH, comparable to previous data (Stroud et al., 2013). We found that about 25% of the rice protein-coding genes were methylated at CHG or CHH sites in the body regions. The non-CG methylated genes had almost no transcript detected in leaves. We could not detect any difference with those marked by both non-CG and H3K27 methylation (Fig. 2A). When we examined the expression of the methylated genes in different rice organs, we found that the percentage of the double-marked genes was lower than those marked by either non-CG methylation or H3K27me3 (Fig. 2B). Importantly, many transcription factors (267) and growth hormone signaling protein (99) genes were methylated at non-CG sites in the body regions (Supplemental Table S3). In wild type, about 37% (3,365/9,016) of H3K27me3-marked genes were methylated at non-CG sites (Fig. 2C). Analysis of the high-throughput data indicated that genes with high H3K27me3 levels displayed significantly higher CHG and CHH methylation in the transcribed or body regions ($P$ values, 0.001, $t$ tests; Supplemental Fig. S3). Many TF and hormone-related genes were enriched from the double-marked genes (Fig. 2D), suggesting that the double marking may be particularly involved in repression of developmental genes in rice leaves. Analysis of SDG711 overexpression plants revealed that about 30% of the genes with ectopic or higher H3K27me3 in 711OX leaves showed methylation at non-CG sites in the wild type (Fig. 3, A and B). The
ectopically methylated genes displayed much higher than genome-wide levels of non-CG methylation (Fig. 3, C and D). The data suggested that non-CG methylation was not inhibitory for SDG711-mediated H3K27me3 and that the two marks were not mutually exclusive in epigenetic modification of genes in rice.

**SDG711 Binding Sites Are Enriched for Non-CG Methylation**

To study genome-wide SDG711-binding sites and their relationship with DNA methylation, we performed ChIP-Seq analysis of rice wild-type 12-d-old seedling leaves by using anti-SDG711 antibodies (Liu et al., 2014, 2015). More than 20 million clean reads were obtained, most of which (96%) were aligned to the rice reference genome (Supplemental Table S1). More than 11 thousands of peaks (P value < 1e-9) were identified, which corresponded to 9,640 protein genes. Analysis of the ChIP-Seq reads revealed that, like H3K27me3, SDG711-binding was enriched in gene bodies (Fig. 4A). In addition, SDG711-associated genes displayed higher levels of H3K27me3 than the genome-wide average (Fig. 4A). Pairwise scatter plots indicated H3K27me3 was positively, although not perfectly, correlated with SDG711-binding (R²=0.665; Fig. 4B). This nonperfect correlation was similar to what was observed in Arabidopsis plants expressing FIE-HA fusion protein by anti-HA ChIP and in Drosophila (Schwartz et al., 2006;
Deng et al., 2013), suggesting that PRC2-binding was not always correlated with H3K27me3. About 41.3% (3,985 out of 9,640) of genes bound by SDG711 were methylated at non-CG sites (Fig. 4, C and E). The data supported the above observations and suggested that SDG711-mediated H3K27me3 targeted many genes methylated at non-CG sites.

To study if there was any consensus sequences among SDG711-binding sites, we used MEME software and identified a CG-rich and a GAGA-like motifs (P value = 6.5e-59 and 1.3e-9, respectively; Fig. 4D). The GAGA motif is also enriched in Arabidopsis PRC2 binding sites (Deng et al., 2013).

Many transcription factors genes and several cytokinin and auxin-related genes were targeted by SDG711 (Supplemental Fig. S4A; Supplemental Table S2). Among these genes, some are key developmental regulators (e.g. OsCKX1, OsCKX2/Gn1a, OsPIN3t, Crl-5, RFL, and LC1; Ashikari et al., 2005; Rao et al., 2008; Kitomi et al., 2011; Zhang et al., 2012b; Zhao et al., 2013; Supplemental Fig. S4B). These genes were found to be enriched for non-CG methylation (Supplemental Fig. S4C). SDG711-binding and H3K27me3 of these genes were validated by ChIP-PCR (Supplemental Fig. S4D). The data indicated that important developmental genes were epigenetically targeted by both SDG711-mediated H3K27me3 and non-CG methylation in rice.

**Loss of CHH Methylation in the osdrm2 Mutant Reduced H3K27me3 of Developmental Genes**

The osdrm2 mutation resulted in a loss of more than 85% of CHH and 23% of CHG methylation genome-wide (Tan et al., 2016). In total, there were 9,175 genes that were differentially methylated (DMG) at CHH sites (Tan et al., 2016). To study whether reduced non-CG methylation affected H3K27me3 in the rice genome, we analyzed H3K27me3 in osdrm2 by ChIP-seq and found a general decrease of H3K27me3 and that 1,865 genes lost the mark in the mutant (Fig. 5A), among which 432 lost CHH methylation in the mutant (Fig. 5B). In addition, we observed that the osdrm2 mutation also reduced H3K27me3 from many genes that were not methylated at CHH sites. This might be, at least partially, due to the activation of a large number of genes in the mutant (Tan et al., 2016), as it is shown that gene up-regulation is associated with reduced H3K27me3 and increased H3K4me3 in rice (Liu et al., 2015).

Among the 432 genes that lost both non-CG methylation and H3K27me3, some of which were found to be involved in SAM function (e.g. APG, TDR; Heang and Sassa, 2012; Niu et al., 2013). We selected eight of the genes (including OsSAUR11 and OsWRKY45) that met the following criteria: with loss of both non-CG methylation and H3K27me3 in osdrm2, up-regulated in osdrm2, and bound by SDG711 in wild type for experimental analysis. Analysis by McrBC-digestion genes confirmed loss of DNA methylation from the loci in the mutant (Fig. 5C). ChIP-PCR analysis with anti-H3K27me3 of the genes confirmed the ChIP-seq data of osdrm2 (Fig. 5D). RT-PCR analysis indicated that these genes were de-repressed in the mutant (Fig. 5D). Interestingly, SDG711-binding to the genes was also reduced in the mutant (Fig. 5D), suggesting that OsDRM2 might be involved in SDG711-mediated H3K27me3 at these loci.
**SDG711 Physically Interacts with OsDRM2 and SDG703**

To further study the mechanism by which non-CG methylation was involved in SDG711-mediated H3K27me3, using yeast (*Saccharomyces cerevisiae*) two-hybrid system, we tested whether SDG711 could interact with rice proteins that are involved in DNA methylation or bind to methylated cytosines. In our mini-screening, we detected two potential interacting candidates: OsDRM2 and SDG703 (Fig. 6A). SDG703 belongs to the SUVH group proteins that contain the SRA domain capable to bind to methylated CHG sites (Johnson et al., 2007; Fig. 6A). Deletion analysis indicated that the five conserved cysteines (C5) domain of SDG711 was necessary and sufficient for the interaction with OsDRM2 and SDG703 (Liu et al., 2014; Fig. 6A). The N-terminal region of OsDRM2 (containing the first UBA, ubiquitin-associated domain; Dangwal et al., 2013), was involved in the interaction with SDG711 (Fig. 6A). The interactions were confirmed by in vitro pull-down (Fig. 6B). Coimmunoprecipitation assays of rice callus nuclear protein extracts with anti-OsDRM2 followed by western blot analysis with anti-SDG711 indicated that the interaction of the proteins occurred in vivo (Fig. 6C). In vivo interaction between SDG711 and SDG703 was demonstrated by coimmunoprecipitation experiments using transfected tobacco (*Nicotiana benthamiana*) protoplasts (Fig. 6D). These data suggested that SDG711 could be recruited to non-CG methylated loci by interacting with DRM2 and/or SDG703.

**DISCUSSION**

The relationship between H3K27me3 and DNA methylation in gene silencing is unclear in plants,
although antagonistic function between the two marks has been detected in imprinted loci in Arabidopsis (Lindroth et al., 2008; Weinhofer et al., 2010). The present data showing that H3K27me3 and non-CG methylation co-occur in body region of many genes suggest that the two marks are not exclusive, but are likely to be mutually reinforced at a set of functional genes in rice. Although not always perfectly aligned at some regions of the gene bodies, the two marks co-occurred at the gene or locus level. The co-occurrence might reflect specific features of the rice genome-wide DNA methylation. For instance, rice non-CG methylation is not peaked close to the centromere regions as found in Arabidopsis and that non-CG, especially CHH, methylation is mainly found in genic regions and a large number of rice genes are marked and repressed by DNA methylation (Feng et al., 2010; Tan et al., 2016). The current study indicates that many genes with non-CG methylation at the body region are also marked by H3K27me3. The even lower expression levels of non-CG methylated genes compared to those only marked by H3K27me3 suggest that DNA methylation plays a more prominent role in gene repression in rice. At this stage, it is not clear whether the comarking by H3K27me3 and non-CG methylation is to reinforce gene repression or to transit to a different chromatin state for chromatin remodeling during the developmental processes, although relatively fewer double marked genes are expressed in different organs (Fig. 2B).

In mammals that have only CG methylation, there exists a close relationship between DNA methylation and H3K27me3. Notably, in cancer, the codependency of marks is largely redistributed with an increase of the dual repressive marks at CpG islands and transcription start sites of silent genes (Statham et al., 2012). In mouse embryonic stem cells, PRC2 and H3K27me3 localize mainly to inactive regions rich in CG dinucleotides or CG islands and almost 97% of PRC2 target genes are associated with annotated CpG islands or similar CG-rich regions (Ku et al., 2008). Recently, it is shown that a cellular signal for recruiting the polycomb machinery is encoded within the target DNA sequences and that CG frequency and positioning are sufficient to recruit PRC2 and establish an H3K27me3 domain in mammals (Jermann et al., 2014).

In Drosophila, which lacks DNA methylation, polycomb can be recruited to polycomb-responsive elements (PRE) (Ringrose et al., 2003). Binding sites of several transcription factors including GAGA factors are enriched with PRE and might contribute to Polycomb recruitment (Ringrose et al., 2003; Muller and Kassis, 2006). Several mechanisms have been identified to be involved in PRC2 recruitment to specific
loci. In addition to interaction with sequence-specific transcription factors or chromatin proteins, PRC2 can be recruited by long noncoding RNAs (lncRNA) to specific loci and mammalian PRC2 binds thousands of RNAs in vivo (Davidovich and Cech, 2015). In plants, the mechanisms by which PcG proteins are recruited to specific genomic loci remain elusive, although it is shown that the lncRNA COLDAIR recruits a component of PRC2 and targets PRC2 to silence the FLC gene (Heo and Sung, 2011). Identification of CG-rich and GAGA-like sequences as SDG711-binding motifs suggests a partial conservation of PcG recruitment mechanism in different organisms. In addition, CG-rich and GAGA-like sequences are potential cytosine methylation sites. The enrichment of non-CG methylation in the body region of genes targeted by SDG711-binding and H3K27me3 (Fig. 5), and physical interactions between SDG711 and the CHH methyltransferase OsDRM2 and between SDG711 and the putative CHG-binding protein SDG703 (Fig. 6), support the hypothesis that non-CG methylation and/or OsDRM2 may be involved in PRC2 recruitment and H3K27me3 at a subset of genes in rice.

A very high coincidence between H3K9me2 and CHG methylation has been observed mainly in TE and repetitive sequences in Arabidopsis (Bernatavichute et al., 2008). H3K9me2 and CHG methylation form a mutually reinforced loop involving the H3K9 methyltransferase KRYPTONITE (KYP, or SUVH4) and the CHG methyltransferase CMT3 (Johnson et al., 2007). The low frequency of H3K9me2 and CHG methylation occurrence in gene body regions is attributed to activity of the Jumonji-C histone demethylase IBM/JMJ25 that removes H3K9me2 from gene bodies (Saze et al., 2008). Recent results showed that H3K9me2 also recruits CMT2 that methylates both CHG and CHH sites in Arabidopsis (Stroud et al., 2014). The present data showing a high correlation between H3K27me3 and non-CG methylation in body region of many genes raises the possibility that H3K27me3 may replace H3K9me2 to mediate gene repression, which may perhaps prevent activation of genes by an IBM1-like activity. This hypothesis is reminiscent of the results in Arabidopsis met1 mutants showing that H3K27me3 and H3K9me2/CHG methylation can be mutually exclusive and replace one another for transcriptional repression in a locus-specific manner (Deleris et al., 2012). The inhibitory effect of ectopic or high levels of H3K27me3 on CHG methylation is in favor of the hypothesis.
MATERIAL AND METHODS

Plant Materials and Growing Conditions
Rice (Oryza sativa ssp. japonica) plants used in this study for ChiP-Seq and BS-Seq analysis were from the ‘Donglin’ (D) background callus culture-regenerated wild type, SDG711 overexpression lines (lines 2, 4, and 5 combined, 711OX) and odon2 mutant plants as described previously (Liu et al., 2014). Seedling leaves from all genotypes grown in one-half-strength Murashige and Skoog medium under a 16-h light/8-h dark cycle at 30°C for 12 DAG were harvested and frozen in liquid nitrogen for DNA, RNA, or chromatin extraction.

Chromatin Immunoprecipitation
About 2 g of seedling leaves were cross-linked in 1% formaldehyde under vacuum. Chromatin was extracted and fragmented to 200 bp by sonication, and ChiP-Seq was performed as previously described (Sun and Zhou, 2008; Hu et al., 2012). Antibodies of H3K27me3 were purchased from Abclonal (s2363p); and anti-SDG711 were prepared previously (Liu et al., 2014). The precipitated and input DNA samples were analyzed either by high throughput sequencing or by real-time PCR with gene-specific primers listed in Supplemental Table S4. All assays were performed at least three times from two biological replicates.

ChiP-Seq and Data Analysis
DNA from ChiP was used to construct sequencing libraries following the protocol provided by Illumina TruSeq ChiP Sample Prep Set A described in (http://support.illumina.com/content/dam/illumina-support/documents/sequencing cheministry_documentation/samplepreps_truseq/trueseqchip/truechip-chip-sample-prep-guide-15023092-b.pdf). Raw reads were generated by HiSeqEquation 2000. Trimomatic (version 0.32) was used to filter out low-quality reads and crop all reads to a uniform length (45 bp; Bolger et al., 2014). Clean reads were mapped to the rice genome (RGAP version 7) by default allowing up to two mismatches using Bowtie2 (version 2.1.0; Trapnell et al., 2012). Samtools (version 0.1.17) is used to remove potential PCR duplication. Peaks of H3K27me3 were called out by MACS software with the default P value (1e-5), and peaks of anti-SDG711 using a P value of 1e-9 because of the largest percentage overlap of H3K27me3 marked genes and SDG711 bound genes, and the output wig files were used for viewing the data by Cbrowse 2.0.

For quantitative analysis of histone modifications, read count for a gene was divided by gene length and multiplied by 1,000 (1 kb) for calculating fold change. For detecting differential (or ectopic) histone modification by quantitative comparison, the following formula, described in (He et al., 2010), was used, in which a Bayesian approach was applied to calculate the P value of differentially marked genes according to observing a given number of tags. Genes with P value < 1e-4 and TPKM fold change > 2 were considered as ectopic marked genes.

\[
p(y|x) = \frac{N_y}{N} \cdot \left( \frac{x+y}{1+y} \right)^y \left( \frac{x}{1+x} \right)^x \left( 1+y \right)^{y+1} \\
\quad P = \min \left\{ \sum_{k=x}^{\infty} p(k|x), \sum_{k=y}^{\infty} p(k|x) \right\}
\]

BS-Seq and Data Analysis
For genomic bisulfitie sequencing, the total genomic DNA of callus culture-regenerated wild type, SDG711 over-expression lines were prepared using the DNeasy plant mini kit (Qiagen). Genomic DNA (5g) was used to generate BS-Seq libraries as previously described (Feng et al., 2011). Sequencing was performed at the Novogene Bioinformatics Technology Co. Ltd, China. Trimomatic (version 0.32) was used to filter out low quality reads and BatMeth (Lim et al., 2012) was used to align clean tags to the rice genome (RGAP version 7) by default parameters followed by removing PCR-amplified redundancy. Methylation levels were calculated by no. C/(no. C + no. T).

For DMR and DMC, the MethySet software was used to identify different methylated regions or genes by using q value (adjusted P value by the BH method) < 0.05 and methylation difference larger than 25, and 2, respectively, for CG, CHG, and CHH (Park et al., 2013). For identification of CG, CHG, and CHH-methylated genes, thresholds of average methylation level of genome-wide CG (53.27%), CHG (8.32%), and CHH (2.21%) sites were used, and genes with a sequenced cytosine number less than 50 were excluded.

Expression Analysis
For rice-seedling expression proles, RNA-Seq data (ID: SRR576931), which come from the similar sample or material with us, were downloaded from the SRA resources (Cui et al., 2013). Clean tags filtered by Trimomatic (version 0.32) were aligned to rice genome (RGAP version 7) by Tophat (version 2.0.13). The suite of Cufflink software (version 2.2.1) was used for splicing transcripts and analyzing expression scores (FPKM, reads per kilobase of exon model per million mapped reads) of each gene.

Motif Search
All sequences of SDG711 binding sites were extracted for analysis by MEME to identify consenus sequence motifs (Bailey and Elkan, 1994).

Yeast Two-Hybrid Assay
Constructs for yeast two-hybrid analysis were generated using the Matchmaker Gold Yeast Two-Hybrid System (Clontech) vectors pGBK7 and pGADT7, which express protein fusions to the GAL4 DNA-binding domain or transcriptional-activation domain, respectively. Full length of cDNA inserts encoding SDG711 and different domains of DNA inserts respectively containing EZD1 domain, five conserved cysteines (C5) domains, and SANT, Cys-rich (CXC), and SET domain of SDG711 were introduced in pGADT7 using the restriction enzymes BamHI and NotI and the Matchmaker Gold Yeast Two-Hybrid System as described in the user manuals. Full length of cDNA inserts encoding SDG711 and DMT706 and SDG703, different domain of cDNA inserts containing one UBA domain, two UBA domains, and a methytransferase domain of DMT706 were introduced in pGBK7 using the restriction enzymes BamHI and Sall and the Matchmaker Gold Yeast Two-Hybrid System as described in the user manuals. The analysis was performed in strain AH109 carrying HIS3 and MEL1 reporters for reconstituted GAL4 activity.

OsDRM2 Polyclonal Antibody Preparation
For preparation of OsDRM2 polyclonal antibody, a 675 bp DNA fragment encoding a 225-amino acid peptide of OsDRM2 (residues 1 to 225) was cloned into pET32a vector (Novagen). The recombinant protein was expressed in E. coli DE3 (Transgen) and purified using nickel nitriotratic acid agarose (Qiagen) to produce rabbit polyclonal antibodies (prepared by Abclonal).

Pull-Down Assay
Full cDNA coding sequences of OsDRM2 and SDG703 were cloned into pET28a, while 5s domain coding sequence of SDG711 was cloned into pGEX4T-1. All these recombinant vectors and empty vectors were transformed into E. coli BL21 (DE3) individually, and 0.2 mm isopropyl β-D-thiogalactoside (IPTG) was added to induce the expression of those proteins. The cell lysate containing 25–50 mg His-fused OsDRM2 and SDG703 or His tag were incubated with MagneHis Protein Purification System (V8550, GE Healthcare) at 4°C for 2 h under gentle agitation. After that, the coated His-fused protein beads were washed three times in 1X phosphate-buffer saline (PBS) solution. Then, the cell lysates containing 25–50 mg GST-fused proteins or GST tag were added into the washed His-fused protein beads. After 1 h of incubation at 4°C under gentle agitation, the bound protein–bead complexes were sedimented by centrifugation and washed five times in 1X PBS solution. Then, the beads were resuspended in 50 μL SDS-PAGE loading buffer and heated for 10 min in the boiled water. Protein samples were resolved on 10% SDS-PAGE gels for further blotting analysis using an anti-GST antibody.
Co-IP Assay

For Co-IP assays, nucleoprotein extracts were prepared from the leaves of transgenic plants with 2 volumes of immunoprecipitation buffer (25 mM Tris–HCl, pH 8.0, 1 mM EDTA, 150 mM NaCl, 0.1% NP-40, 1% Triton X-100, 0.2 mM NaVO3, 0.2 mM PMSF and 10% [v/v] glycerol). The extracts were centrifuged at 16,000 rcf for 20 min at 4°C, and the supernatant was incubated with anti-ODRM2 and IgG antibody independently incubated protein A (Invitrogen) beads overnight at 4°C. The resin was washed three times with immunoprecipitation buffer and then eluted in the SD5 sample buffer. After separation on 12% SDS-PAGE gels, the proteins were transferred onto nitrocellulose membranes and probed with anti-SDG711 in a dilution of 1:1000.

Accession Numbers

Sequence data from this article can be found in the Rice Genome Annotation Project Web site (http://rice.plantbiology.msu.edu/) under the following accession numbers: SDG711, Oso06g16390; OsCCX1, Oso19g09260; OsCCX2, Oso1q10110; OsPIN3h, Oso1g45550; Crl-3, Oso1g3250; RFL, Oso1p51000; LC1, Oso1g57610. The ChIP-seq and BS-seq data described in this paper have been deposited into the GEO database (GSE71640, GSE81436).

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Analysis of H3K27me3-marked genes in wild type (callus regenerated), SDG711 over-expression plants by ChIP-seq assays.

Supplemental Figure S2. Genome browser screenshots of randomly selected regions of H3K27me3 in wild type SAM and leaves and 711OX leaves.

Supplemental Figure S3. H3K27me3 levels correlate with non-CG methylation in the rice genome.

Supplemental Figure S4. Key developmental genes that are marked by H3K27me3 and bound by SDG711 show higher levels of non-CG methylation.

Supplemental Table S1. ChIP-seq data of anti-H3K27me3 and anti-SDG711 and BS-seq data.

Supplemental Table S2. Developmental genes with H3K27me3 changes in SAM-leaf transition, 711OX, osdm2, and bound by SDG711.

Supplemental Table S3. Rice developmental genes with CHG and CHH methylation in seedling leaves.

Supplemental Table S4. Primers used for quantitative RT-PCR, ChIP-PCR, and McrBC digestion-PCR analysis.

ACKNOWLEDGMENTS

We thank Dr Qinghua Zhang for high throughput sequencing and Qinglu Zhang and Prof. Xianghua Li for help in field experiments and management.

LITERATURE CITED

Heo JB, Sung S (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science 331: 76–79
Interplay between H3K27me3 and Non-CG Methylation


Statham AL, Robinson MD, Song JZ, Cooen MW, Storrzaker C, Clark SJ (2012) Bisulfite sequencing of chromatin immunoprecipitated DNA (BisChIP-seq) directly informs methylation status of histone-modified DNA. Genome Res 22:1120–1127


Copyright (c) 2020 American Society of Plant Biologists. All rights reserved.