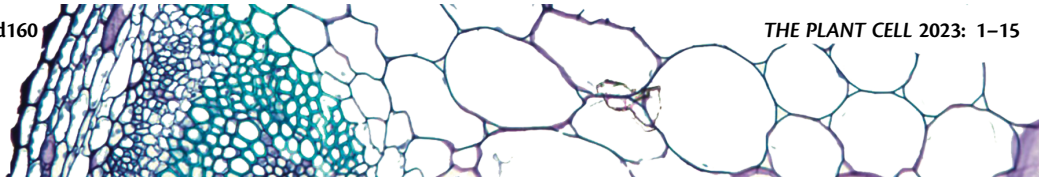
https://doi.org/10.1093/plcell/koad160 THE PLANT CELL 2023: 1–15



Large-Scale Biology

Asymmetric variation in DNA methylation during domestication and de-domestication of rice

Shuai Cao [1,2] Kai Chen [3] Kening Lu [1] Shiting Chen [1] Xiyu Zhang [1] Congcong Shen [3] Shuangbin Zhu [3] Yanan Niu [4] Longjiang Fan [5] Z. Jeffrey Chen [6,*] Jianlong Xu [4,*] and Qingxin Song [1,*]

1 State Key Laboratory of Crop Genetics and Germplasm Enhancement, Jiangsu Collaborative Innovation Center for Modern Crop Production, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China
2 Temasek Life Sciences Laboratory, 1 Research Link, National University of Singapore, Singapore 117604, Singapore
3 Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, Guangdong 518120, China
4 National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China
5 Institute of Crop Science & Institute of Bioinformatics, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, Zhejiang 310058, China
6 Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX 78712, USA

*Author for correspondence: qxsong@njau.edu.cn (Q.S.), xujianlong@caas.cn (J.X.), zjchen@austin.utexas.edu (Z.J.C.)
The author(s) responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell/pages/General-Instructions) is: Qingxin Song (qxsong@njau.edu.cn).

Abstract

Hundreds of plant species have been domesticated to feed human civilization, while some crops have undergone de-domestication into agricultural weeds, threatening global food security. To understand the genetic and epigenetic basis of crop domestication and de-domestication, we generated DNA methylomes from 95 accessions of wild rice (*Oryza rufipogon* L.), cultivated rice (*Oryza sativa* L.) and weedy rice (*O. sativa* f. *spontanea*). We detected a significant decrease in DNA methylation over the course of rice domestication but observed an unexpected increase in DNA methylation through de-domestication. Notably, DNA methylation changes occurred in distinct genomic regions for these 2 opposite stages. Variation in DNA methylation altered the expression of nearby and distal genes through affecting chromatin accessibility, histone modifications, transcription factor binding, and the formation of chromatin loops, which may contribute to morphological changes during domestication and de-domestication of rice. These insights into population epigenomics underlying rice domestication and de-domestication provide resources and tools for epigenetic breeding and sustainable agriculture.

Introduction

Cultivated rice (*Oryza sativa* L.) is 1 of most important staple crops worldwide, feeding more than a third of the global human population. It has broad geographical distribution with adaptability to diverse agroecological niches. There are several genetically distinct rice populations, including the major cultivated *japonica* (*CJ*) and cultivated *indica* (*CI*) groups, as well as the more geographically restricted Aus and aromatic (Aro) rice (Molina et al. 2011; Huang et al. 2012; Choi et al. 2017). Cultivated rice is believed to have been domesticated from its wild progenitor *Oryza rufipogon* ~9,000 yr ago (Xu et al. 2011; Huang et al. 2012; Meyer et al. 2016; Win et al. 2017). Although the domestication history of rice remains controversial, there is a popular belief that ancient *japonica* rice was first domesticated from *O. rufipogon* in southern China and then crossed to local wild rice in South East and South Asia to generate *indica* rice (Huang et al. 2012). Dramatic changes in morphological and physiological traits have occurred during rice domestication, including loss of seed shattering and increased grain size (Huang et al. 2012;

Received April 17, 2023. Accepted May 15, 2023. Advance access publication June 5, 2023

IN A NUTSHELL

Background: Humans have domesticated hundreds of plant and animal species as sources of food and materials. For example, rice (*Oryza sativa*) feeds more than a third of the global population. Cultivated rice was domesticated from a wild progenitor ~9,000 yr ago, which was accompanied by dramatic changes in morphological and physiological traits. Importantly, cultivated rice can turn back into "wild-like" plants under natural selection, by de-domestication. DNA methylation is a conserved epigenetic mark in most eukaryotes that plays important roles in plant development and environmental responses. However, whether epigenetics contributes to domestication and de-domestication of rice is largely unknown.

Question: We wished to know whether and how variation in DNA methylation affected gene expression during rice domestication (cultivars vs wild rice) and de-domestication (weedy rice vs cultivars).

Findings: We generated single-base resolution DNA methylomes from 95 accessions of wild, cultivated, and weedy rice. We detected a significant decrease in DNA methylation during rice domestication. By contrast, DNA methylation dramatically increased following rice de-domestication. Notably, hypomethylated sites from wild to cultivated rice and hypermethylated sites from cultivated to weedy rice were largely not shared. We determined that variation in DNA methylation affected the binding of transcription factors and chromatin accessibility to regulate expression of nearby and distal genes, which would be expected to cause morphological changes during domestication and de-domestication of rice.

Next steps: Future efforts will focus on identification of DNA methylation variants associated with changes for agronomic traits during rice domestication, which could help breeders facilitate rice improvement through epigenetic engineering and breeding.

Wang et al. 2018). These domestication traits involve a number of genes such as *PROSTRATE GROWTH1* (*PROG1*) for erect plant architecture (Wu et al. 2018), *shattering4* (*sh4*) for reduced seed shattering (Li et al. 2006) and *LONG AND BARBED AWN1* (*LABA1*) for barbless awns (Hua et al. 2015). Artificial selection tends to elevate crop fitness under human-manipulated environments. However, cultivated crops can revert back into self-sustainable "wild-like" plants under natural selection, which is called de-domestication (Li et al. 2017; Qiu et al. 2017). Weedy rice (*O. sativa* f. *spontanea*) is a pernicious weed in rice fields that resembles cultivated rice morphologically but exhibits weedy characteristics, including heavy seed shattering and strong seed dormancy (Guo et al. 2018; Wu et al. 2021). Weedy rice is a conspecific relative of cultivated rice and was proposed to have partially evolved from cultivated rice via de-domestication (Guo et al. 2018; Wu et al. 2021). Although weedy rice and wild rice share several morphological characteristics, accumulating evidence suggests that de-domestication is not a mere reversal of domestication (Kaut et al. 2017; Qiu et al. 2020). Indeed, large-scale whole-genome sequencing of wild, cultivated, and weedy rice has revealed many genomic regions underlying morphological changes and environmental adaptation during rice domestication and de-domestication (Huang et al. 2012; Li et al. 2017; Wang et al. 2018). However, epigenetic contributions to domestication and de-domestication of rice are largely unknown.

DNA methylation is a conserved epigenetic mark in most eukaryotes. In plants, DNA methylation occurs in CG, CHG, and CHH (H = A, T, or C) contexts (Zhang et al. 2018). Symmetric CG and CHG methylation can be maintained through cell division by the DNA methyltransferases DNA METHYLTRANSFERASE 1 (MET1) (Hu et al. 2014) and CHROMOMETHYLASE 3 (CMT3) (Cheng et al. 2015), respectively. CHH methylation is mainly established de novo through the CHROMOMETHYLASE 2 (CMT2) and RNA-directed DNA methylation (RdDM) pathways (Chinnusamy and Zhu 2009). While CHH methylation shows obvious distribution variation throughout plant development (Song et al. 2015), CG and CHG methylation, especially CG methylation, is generally stable across plant vegetative tissues (Song et al. 2015; Crisp et al. 2020). DNA methylation participates in transposon silencing to maintain genome stability and transcription regulation to orchestrate morphological development (Lim et al. 2007; Lisch 2009; Cui and Cao 2014). Methylation defects induce developmental abnormalities in *Arabidopsis* (*Arabidopsis thaliana*) and lead to death in rice (Ono et al. 2012; Hu et al. 2014; Cheng et al. 2015; Zhang et al. 2020). During evolution, the rate of changes in DNA methylation is faster than that of sequence substitution (Becker et al. 2011; Song et al. 2017). For example, the epi-mutation rate in the CG context (2.56×10^{-4}) was 5 orders of magnitude higher than the corresponding genetic mutation rate (7×10^{-9}) in *Arabidopsis* (Becker et al. 2011). DNA methylation changes can produce meiotically stable epialleles, which are inherited by the offspring through natural evolution and artificial selection (Richards 2006). Stable changes in DNA methylation at promoters confer symmetric flower development in yellow toadflax (*Linaria vulgaris*) (Cubas et al. 1999). Genome-wide methylation changes during the evolution and domestication of polyploid

cotton (*Gossypium hirsutum* L.) uncovered several hundreds of epialleles including at *CONSTANS-LIKE 2D*, which contributed to the loss of photoperiod sensitivity during cotton domestication (Song et al. 2017).

Here we generated single-base resolution DNA methylomes of 95 wild, cultivated, and weedy rice accessions and report numerous examples of variation in DNA methylation among these species. By integrating MethylC-seq, transcriptome deep sequencing (RNA-seq), assay for transposase-accessible chromatin with sequencing (ATAC-seq) and DNA affinity purification and sequencing (DAP-seq) data, we examined the effects of variation in DNA methylation on chromatin accessibility, transcription factor binding, and gene expression during domestication and de-domestication of rice. This research will provide epigenomic resources for rice breeders to facilitate improvement of crop production through epigenetic engineering and breeding.

Results

Variation in DNA methylation among wild, cultivated, and weedy rice

To uncover the role of DNA methylation during domestication and de-domestication of rice, we generated single-base resolution DNA methylomes of 95 wild, cultivated, and weedy rice accessions using third leaves of rice seedling at 30 d after planting (DAP). We collected leaves from 20 wild rice (*O. rufipogon*), 18 *CJ* (*O. sativa*), 20 *CI* (*O. sativa*), 9 *aromatic* Aro (*O. sativa*), 10 Aus (*O. sativa*), 10 weedy *japonica*-type (*WJ*) (*O. sativa* f. *spontanea*) and 8 weedy *indica*-type rice (*WI*) (*O. sativa* f. *spontanea*) (Fig. 1A and Supplemental Data Sets 1 and 2). We mapped reads produced by whole-genome bisulfite sequencing (WGBS) to the reference genome of the rice cultivar Nipponbare (Nip). On average, we obtained 26 million WGBS paired-end reads (>7 Gb) per accession, resulting in >78% of the genome being covered by WGBS reads. We removed cytosines at single nucleotide polymorphisms (SNPs) between accessions to exclude the effect of nucleotide variation on DNA methylation analysis. We mainly focused on variation and evolution of CG methylation during rice domestication and de-domestication, as the majority (>52.9% for CHG and >89.3% for CHH) of non-CG sites were unmethylated (Supplemental Fig. S1A) and non-CG methylation was much less stable across vegetative tissues than CG methylation (Supplemental Fig. S1, B and C) (Song et al. 2015; Crisp et al. 2020). Pairwise kinship calculated using SNPs was highly correlated with that based on CG methylation levels (Fig. 1B), suggesting that variation in DNA methylation can recapitulate genetic relationships among different rice accessions. Moreover, principal component analysis (PCA) using the variation in CG methylation successfully divided rice accessions into different subgroups (Fig. 1C), consistent with classifications based on SNPs

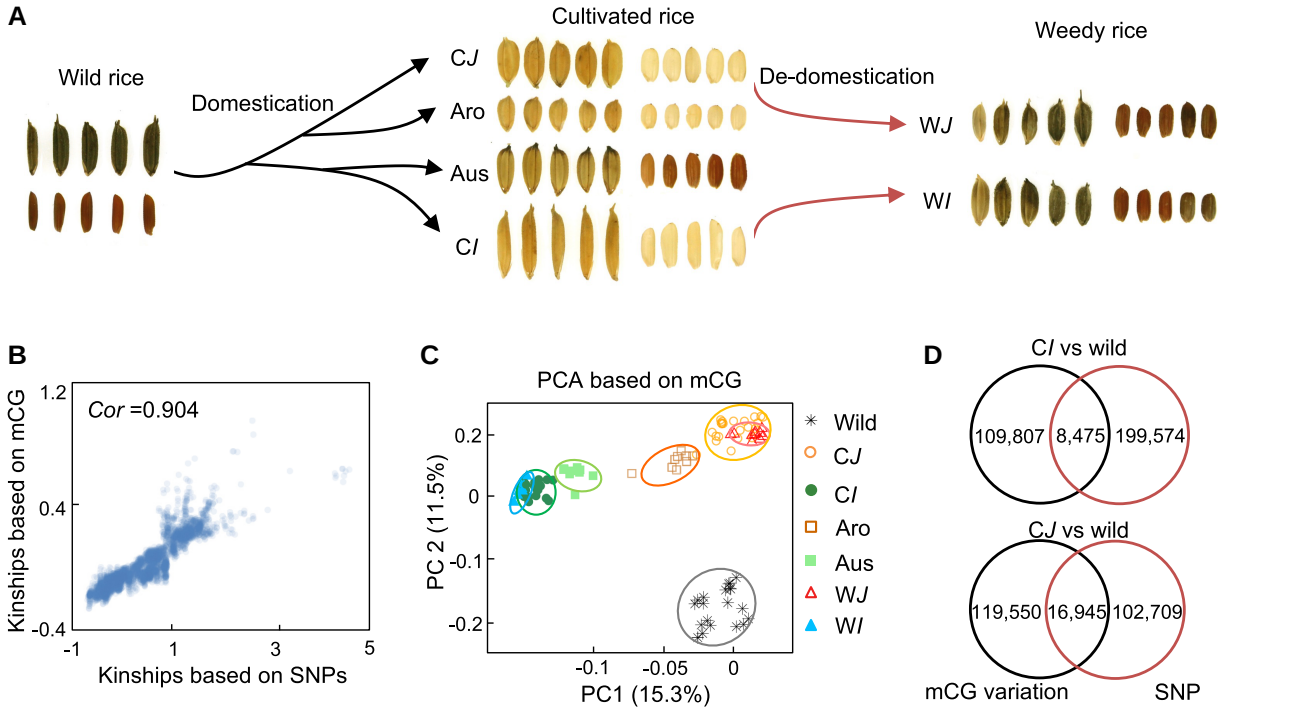


Figure 1. Correlation between variation in DNA methylation and nucleotide diversity during domestication and de-domestication. **A)** Seed morphology of wild rice, *CJ*, *CI*, Aro, Aus, *WJ*-type and *WI*-type rice. **B)** Correlation of kinships between any 2 rice accessions, calculated by SNPs or mCG levels. *Cor* = Pearson's correlation coefficient. **C)** PCA of all rice accessions based on CG methylation levels. **D)** Numbers of 100-bp windows containing SNPs and CG methylation variation.

(Supplemental Fig. S1D). Wild rice was separated from cultivated and weedy rice, to which *CJ* and *CI* rice respectively clustered together with the corresponding weedy rice *WJ* and *WI* (Fig. 1C), which was also similar to the phylogenetic tree of these rice accessions (Supplemental Fig. S2). These results indicate fewer genetic and epigenetic changes during de-domestication compared to domestication. In addition, we identified 118,282 and 136,495 regions containing variation in CG methylation in *CJ* and *CI*, respectively, relative to wild rice using 100-bp windows, of which about 7% to 12% contained SNPs (Fig. 1D). To explore whether genomic variation in core DNA methylation plays important roles in variation in DNA methylation among rice accessions, we performed genome-wide association studies (GWAS) using genome-wide average CG methylation levels as a phenotype but failed to detect significant peaks (Supplemental Fig. S3).

Global decrease of DNA methylation during rice domestication

Compared to wild rice, global CG methylation levels were significantly lower in all cultivated rice (*CJ*, *CI*, Aro, and Aus; $P < 0.01$, Wilcoxon signed-rank test) (Fig. 2A), although we cannot totally exclude potential bias resulting from the reference genome on our analysis of DNA methylation levels. To test the role of DNA methylation during rice domestication, we dissected differentially methylated regions (DMRs) between wild rice (20 accessions) and cultivated rice (*CJ* and *CI* accessions). We identified 48,244 hypo- and 30,084 hyper-DMRs in the *CJ* group (covering a total of 6.6 Mb, or 1.96% of the genome) and 41,368 hypo- and 32,157 hyper-DMRs in the *CI* group (for a total of 6.4 Mb, or 1.9% of genome), relative to wild rice (Fig. 2B; Supplemental Data Set 3). It is worth noting that 19,134 (39.7% to 46.1%) of all hypo-DMRs and 12,526 (39.0% to 40.8%) of all hyper-DMRs are shared by the *CJ* and *CI* groups (Fig. 2B), which accounted for 4.3% (15 Mb) of rice genomes and was significantly more than expected by chance ($P = 0$, hypergeometric test). These results are not surprising, as some domestication genes have been demonstrated to be shared by the *CJ* and *CI* groups (Huang et al. 2012). For example, the rice homolog *OsORM1* (*LOC_Os04g47970*) of the *Arabidopsis* gene *OROSOMUCOID PROTEIN1* (*ORM1*) showed a decrease in DNA methylation along its promoter and 3′ downstream sequences in cultivated rice compared to wild rice (Fig. 2C). Accordingly, we detected higher expression levels for *OsORM1* in cultivated rice relative to wild rice (Fig. 2C).

As most accessible chromatin regions are located in unmethylated regions in plants (Crisp et al. 2020), we hypothesized that changes in DNA methylation might have affected chromatin accessibility during rice domestication. To test this idea, we performed ATAC-seq to examine chromatin accessibility of wild (GW5 and GW7), *CJ* (DongJin [DJ] and Nip) and *CI* (R284 and R61) rice using third leaves at the same developmental stage as used for MethylC-seq (Supplemental Data Set 4). Interestingly, we determined that chromatin accessibility significantly increased at shared hypo-DMRs but significantly decreased at shared hyper-DMRs in the *CJ* and *CI* groups ($P < 0.01$, Wilcoxon signed-rank test) (Figs. 2D; Supplemental Fig. S4A). Moreover, these hypo-DMRs, but not hyper-DMRs, were associated with active histone modifications H3K4me3 (trimethylation of lysine 4 on histone H3) and H3K27ac (acetylation of lysine 27 on histone H3) (Supplemental Fig. S4B). To examine if changes in DNA methylation affected gene expression, we identified 10,647 proximal genes with shared DMRs located in 1-kb flanking sequences and 1,820 Ds interacting with shared DMRs via long-range chromatin loops through reanalysis of published Hi-C data sets (Liu et al. 2017) (Supplemental Data Set 5) (Fig. 2E). Indeed, changes in DNA methylation significantly induced changes in the expression of DMR-associated proximal genes (DPGs) and distal genes (DDGs) in the *CJ* and *CI* groups relative to wild rice (Figs. 2E and Supplemental Fig. S4, C and D). For example, 1 shared hypo-DMR on chromosome 5 (Chr 05) may be involved in elevated expression of the neighboring gene *OsNDPK1* (*NUCLEOSIDE DIPHOSPHATE KINASE 1*, LOC_Os05g51700) (Cho et al. 2004) and the distal gene *OsCAX1b* (*Cation/proton exchanger 1a*, LOC_Os05g51610), located ~70 kb downstream of the DMR ($q < 1.8 \times 10^{-2}$ for interaction loops) in the *CJ* and *CI* groups (Fig. 2F). These results suggest that changes in DNA methylation may have affected chromatin accessibility and histone modifications and hence alter expression levels of proximal and distal genes during rice domestication.

In addition to shared DMRs, the *CJ* and *CI* groups possessed 46,474 and 41,671 group-specific DMRs relative to wild rice, respectively (Figs. 2B and 3A), supporting the hypothesis that the *CJ* and *CI* groups evolved independently after a shared domestication event (Molina et al. 2011; Wang et al. 2018). Shared DMRs, but not group-specific DMRs, tended to occur in the sequences flanking genes and intergenic regions (Supplemental Fig. S5). As with shared DMRs, group-specific DMRs, with the exception of *CI*-specific hypo-DMRs, were associated with significant changes in chromatin accessibility ($P < 0.01$, Wilcoxon signed-rank test) (Fig. 3A) and the histone modifications H3K4me3 and H3K27ac (Supplemental Fig. S4B). *CJ*- and *CI*-specific changes in DNA methylation also significantly induced changes in the expression of genes overlapping with DMRs in flanking sequences (Supplemental Data Set 5) (Supplemental Fig. S4, E and F), which may contribute to the morphological diversification of the *CJ* and *CI* groups. The genes associated with DMRs induced by domestication were significantly enriched in several important processes, including stress response and cell growth (Supplemental Data Set 6). In general, *CJ* rice has wider leaves and shorter grains than *CI* rice. Many genes have been reported to participate in regulating organ development in rice, including *SHORT GRAIN1* (*SG1*) (Nakagawa et al. 2012). Elevated expression of *SG1* leads to wider leaves and shorter grains (Nakagawa et al. 2012). Notably, we detected a massive decrease in DNA methylation in a region

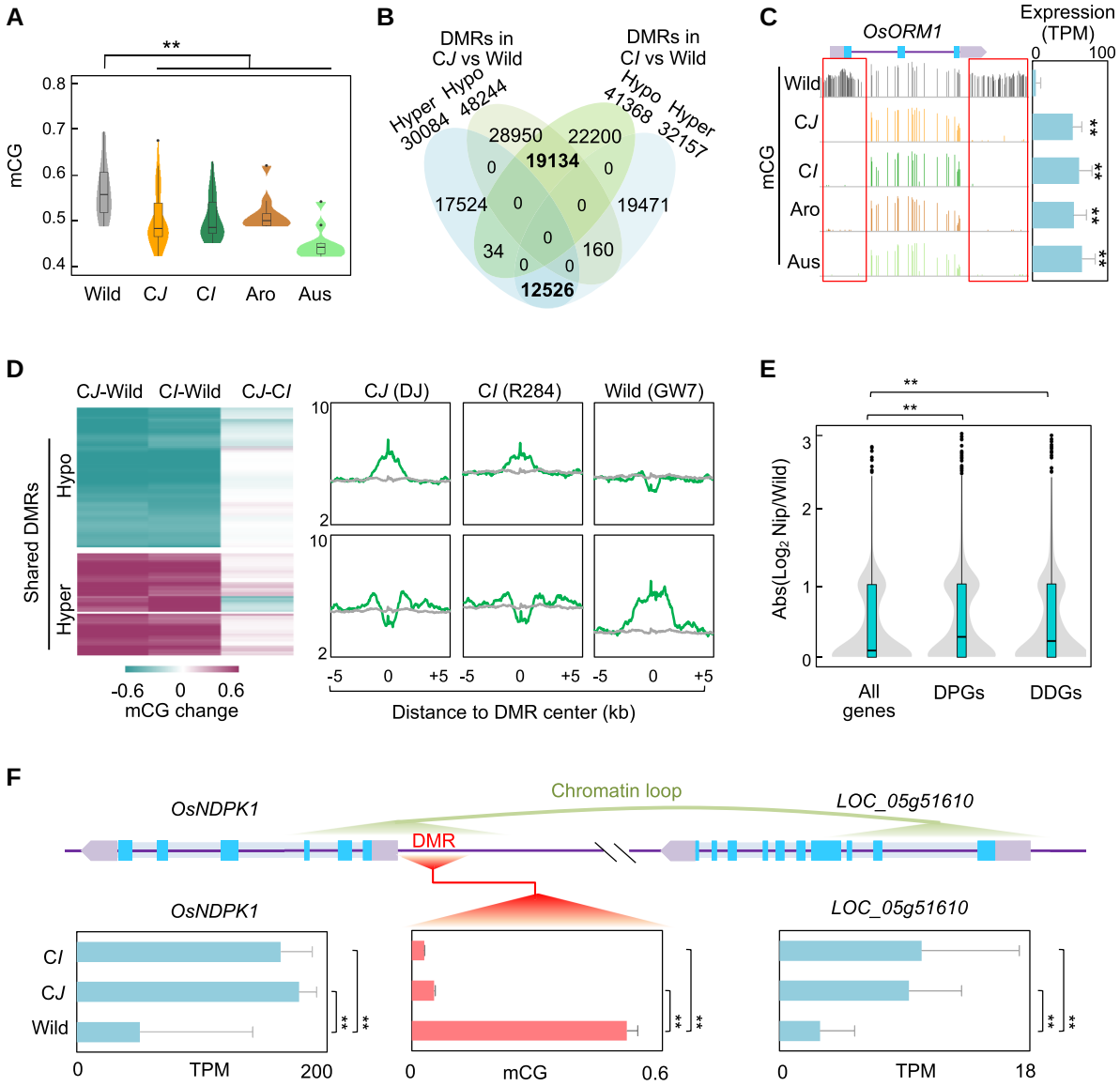


Figure 2. Domestication-induced demethylation of proximal and distal regulatory elements. **A)** Decreased CG methylation levels in cultivated rice compared to wild rice. Significant differences are indicated by ** ($P < 0.01$; see Supplemental Data Set 12), according to a Wilcoxon signed-rank test. **B)** Shared DMRs in the *CJ* and *CI* groups, relative to wild rice. **C)** The example of *OsORM1* (LOC_Os04g47970) showing lower CG methylation levels in its promoter and downstream sequences but higher expression levels in cultivated rice compared to wild rice. Significant differences are indicated by ** ($P < 0.01$; see Supplemental Data Set 12), according to a 2-tailed Student's *t*-test. Scale: 0 to 1 for methylation level. **D)** Methylation changes (left panel) and chromatin accessibility variation (right panel) in shared DMRs between the *CJ* and *CI* groups relative to wild rice. The chromatin accessibility in DMRs and random genomic regions from *CJ* (DJ), *CI* (R284), and wild (GW7) rice was normalized to RPM and scaled from 2 to 10 for each 50-bp bin. Green and gray lines indicate DMRs and random genomic regions, respectively. **E)** Elevated expression between *CJ* (Nip) and wild rice for DPGs and DDGs compared to all genes. Significant differences are indicated by ** ($P < 0.01$; see Supplemental Data Set 12), according to a Wilcoxon signed-rank test. **F)** An example of shared hypo-DMR raising expression of the neighboring gene (*OsNDPK1*) and distal gene (LOC_Os05g51610) in the *CJ* and *CI* groups relative to wild rice. Significant differences are indicated by ** ($P < 0.01$; see Supplemental Data Set 12), according to a 2-tailed Student's *t*-test.

~1.4-kb upstream of *SG1* in the *CJ* and Aro groups, which have shorter grains and wider leaves (Fig. 3, B and C). By contrast, common *CI* and Aus groups with longer grains and thinner leaves had much higher DNA methylation levels across the same region (Fig. 3, B and C). Consequently, *SG1* was silenced in most examined *CI* and Aus accessions but expressed in most examined *CJ* and Aro accessions (Fig. 3C). These results indicate a correlation between variation in DNA methylation and morphological changes among rice varieties.

DNA methylation valleys in the rice genome

There are many large regions that are devoid of DNA methylation in animal and plant genomes; these regions are called

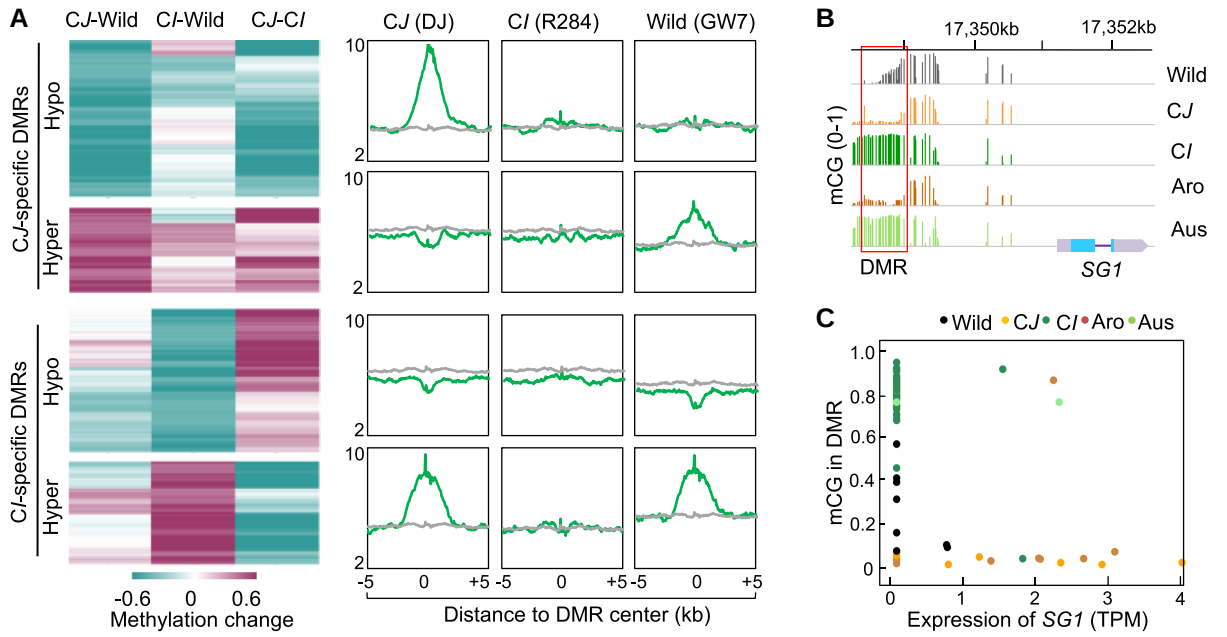


Figure 3. Independent altered methylation during *CJ* and *CI* domestication. **A)** Methylation changes (left panel) and chromatin accessibility variation (right panel) in *CJ*- and *CI*- specific DMRs compared to wild rice. The chromatin accessibility in DMRs and random genomic regions was normalized to RPM and scaled from 2 to 10 for each 50-bp bin. Green and gray lines indicate DMRs and random genomic regions, respectively. **B and C)** An example of *CJ*-specific hypo-DMR (**B**) at the promoter of *SG1* (LOC_Os09g28520) regulating its expression (**C**). Scale: 0 to 1 for methylation level.

DNA methylation valleys (DMVs) (Xie et al. 2013). DMVs were reported to be involved in cell differentiation and cancer pathways in animals (Xie et al. 2013) and associated with transcriptional regulation during soybean (*Glycine max* L.) seed development (Chen et al. 2018). We identified 29,585 CG DMVs of at least 1 kb in length in wild and cultivated rice, spanning over 13% of the rice genome. About 80% (23,673/29,585) of these DMVs were conserved among the 5 rice groups (Fig. 4A). We also observed low non-CG (mCHG and mCHH) methylation levels in these CG DMVs (Supplemental Fig. S6A). These conserved DMVs were enriched near TSSs (Fig. 4B, left) with higher levels of DNA accessibility (Fig. 4B, right). Genes overlapping with DMVs (DMV-genes) showed significantly higher expression than genes without DMVs (non-DMV-genes) ($P < 0.01$, Wilcoxon signed-rank test) (Fig. 4C). These DMV-genes exhibited a lower tissue specificity for their expression than non-DMV-genes (Fig. 4D). In addition, we detected significantly higher levels of H3K4me3 and H3K27ac in DMVs compared to their flanking regions (Supplemental Fig. S6B). Using a published data set of T-DNA insertions in rice (Sallaud et al. 2004), we identified fewer T-DNA insertions mapping to proximal DMVs (pDMVs) overlapping with genes and to distal DMVs (dDMVs) away from genes (Fig. 4E), suggesting a potential role for intact DMVs in plant survival.

Interestingly, 171 (or 0.6%) of conserved DMVs in cultivated rice showed extensive DNA CG methylation (mCG > 0.4) in wild rice, which may have contributed to domestication traits. For example, a selfish genetic element consisting of *OPEN READING FRAME2* (*ORF2*) and *ORF3* was shown to facilitate reproductive isolation in domesticated rice (Yu et al. 2018). *ORF2* encodes a toxin affecting pollen development, while *ORF3* encodes the antidote for pollen viability. The protein encoded by *ORF2* is thought to represent gain-of-function, probably by artificial selection of a non-functional allele from wild rice (Yu et al. 2018). We identified a DMV spanning the promoter of *ORF2* and the neighboring gene *ORF1* in all cultivated rice (Fig. 4F), while this DMV was highly methylated in wild rice (Fig. 4F). As a result, *ORF1* and *ORF2* were expressed at high levels in cultivated and weedy rice accessions but at low levels in wild rice accessions (Fig. 4G). These results reveal a potential role for epigenetic regulation in functionalization of plant development during rice domestication.

Global increase of DNA methylation during rice de-domestication

Plant de-domestication is often accompanied with loss of agronomic traits that had been acquired under domestication (Qiu et al. 2017), which may involve genomic and epigenomic changes. We determined that the allele frequencies at 8.1% and 10% of SNPs turned back to allele frequencies similar to those of wild rice in *WJ* and *WI* rice, respectively. In addition to genomic changes, a significant increase in DNA methylation occurred in weedy rice relative to the corresponding cultivated rice (Fig. 5A and Supplemental Fig. S7), which was opposite that of demethylation observed during domestication. We observed higher DNA methylation throughout chromosomes, affecting euchromatin and heterochromatin alike in weedy rice

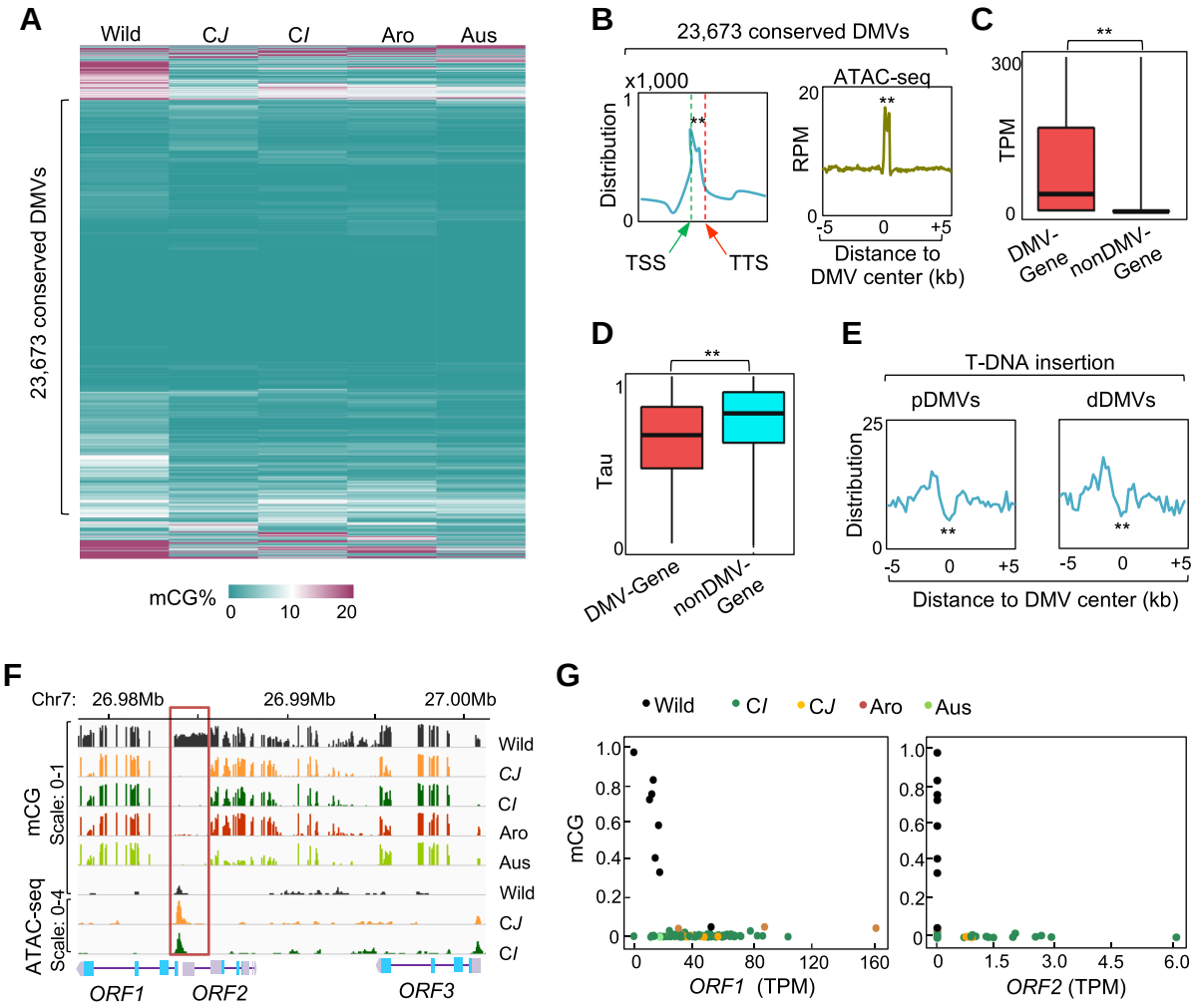


Figure 4. Variation of DMVs in wild and cultivated rice. **A)** Heatmap showing the mean methylation levels of DMVs. **B)** Difference in chromatin accessibility (left panel) and distribution (right panel) in conserved DMVs compared to flanking regions. Significant differences are indicated by ** ($P < 0.01$; see Supplemental Data Set 12), according to a Wilcoxon signed-rank test. **C)** Expression levels of genes with DMVs (DMV-gene) and genes without DMVs (nonDMV-gene). Significant differences are indicated by **($P < 0.01$, see Supplemental Data Set 12), according to a Wilcoxon signed-rank test. **D)** Tissue specificity of genes associated with DMVs or lacking DMVs. Significant differences are indicated by ** ($P < 0.01$; see Supplemental Data Set 12), according to a Wilcoxon signed-rank test. **E)** Density of T-DNA insertions throughout pDMVs overlapping with genes and dDMVs away from genes compared to flanking regions. Significant differences are indicated by ** ($P < 0.01$; see Supplemental Data Set 12), according to a Wilcoxon signed-rank test. **F)** An example of domestication-induced DMV which is indicated by a red box. **G)** Association between DNA methylation levels in the DMV shown in F and expression levels of neighboring genes (*ORF1*, left; *ORF2*, right) in rice.

(Fig. 5B). Notably, the correlation between domestication-induced and de-domestication-induced DNA methylation changes was very low ($Cor = -0.08$, $P = 4.9 \times 10^{-13}$) (Fig. 5C), indicating that most DNA methylation changes occurred independently during rice domestication and de-domestication. We identified 7,271 hypo- and 9,946 hyper-DMRs in the W*J* group relative to the C*J* group and 2,199 hypo- and 2,347 hyper-DMRs in the W*I* group relative to the *CI* group (Supplemental Data Set 7). The hyper-DMRs were slightly overrepresented in the gene body, but the hypo-DMRs were slightly enriched over intergenic regions (Supplemental Fig. S8). Only 15% to 24% of all DMRs overlapped between domestication and de-domestication events (Supplemental Fig. S9A). Interestingly, 12 hypo-DMRs in the C*J* group and 2 hypo-DMRs in the C*I* group showed further DNA methylation decrease during de-domestication. As observed during domestication, these DMRs were negatively associated with changes in chromatin accessibility during rice de-domestication (Fig. 5, D and F). Genes associated with de-domestication-induced DMRs were enriched for several important biological processes, including photosynthesis and defense response (Supplemental Data Sets 8 and 9) (Fig. 5, E and G).

Ethylene response factors (ERFs) are responsible for stress responses by binding to their cognate promoter motifs in plants (Muller and Munne-Bosch 2015). Remarkably, the

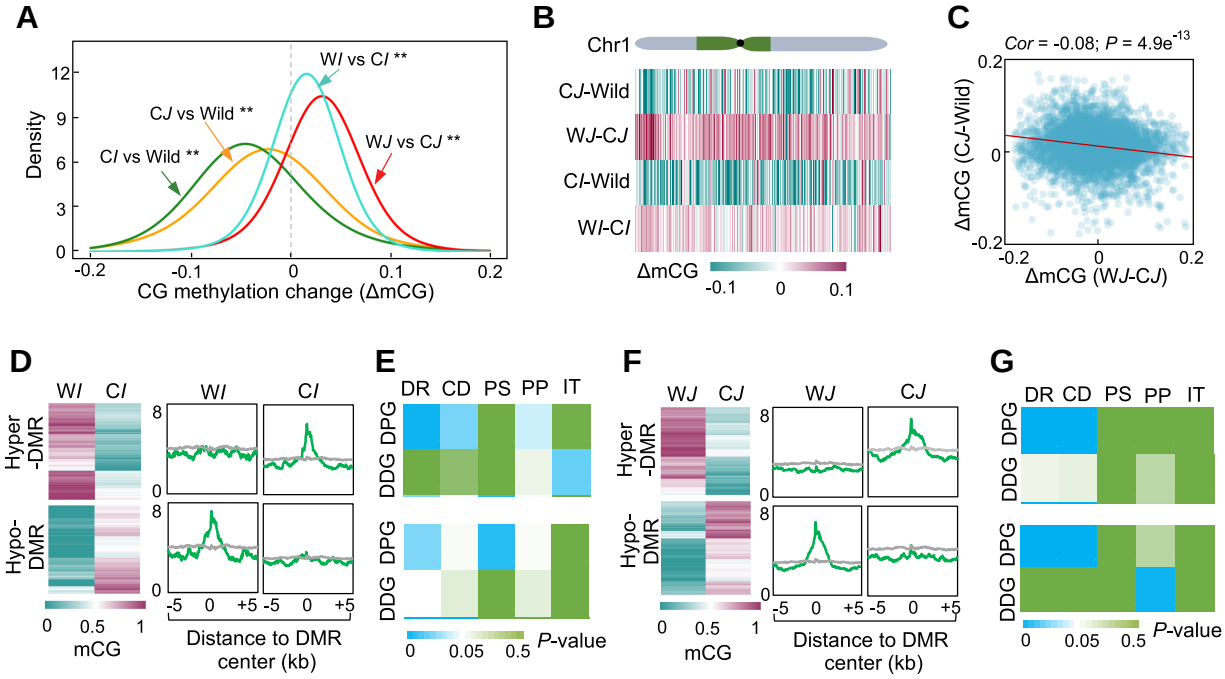


Figure 5. Reverse changes of DNA methylation induced by de-domestication relative to domestication. **A)** Kernel density distribution of methylation changes in *CJ* vs wild, *CI* vs wild, *WI* vs *CI*, and *WJ* vs *CJ*. Significant differences are indicated by ** ($P < 0.01$; see Supplemental Data Set 12), according to a Wilcoxon signed-rank test. Methylation changes were calculated for each 50-kb window. **B)** Heatmap representation of methylation changes along chromosome 1 (Chr 1) during rice domestication and de-domestication. Methylation changes were calculated for each 50-kb window. **C)** Weak correlation between domestication-induced methylation changes (*CJ* vs wild) and de-domestication-induced methylation changes (*WJ* vs *CJ*). *Cor* = Pearson's correlation coefficient. Methylation changes were calculated for each 50-kb window. **D and F)** Methylation state (left panel) and chromatin accessibility variation (right panel) in de-domestication-induced DMRs for the *WI* group (**D**) and the *WJ* group (**F**). Chromatin accessibility in DMRs and random genomic regions were normalized to RPM and scaled from 0 to 8 for each 50-bp bin for *WI* (ZJ2), *CI* (R284), *WJ* (H6), and *CJ* (Nip). Green and gray lines indicate DMRs and random genomic regions, respectively. **E and G)** Gene ontology enrichment of DPGs and DDGs associated with de-domestication-induced DMRs for the *WI* group (**E**) and the *WJ* group (**G**). DR, defense response; CD, cell death; PS, photosynthesis; PP, phosphorylation; IT, ion transport.

short palindromic $(GCC)_5$ motif was enriched in de-domestication-induced DMRs (Fig. 6A), which is similar to the binding motif of ERFs in *Arabidopsis* (Fig. 6A) (O'Malley et al. 2016), including AtERF115 and AtERF2 protein whose homologs in rice are encoded by *OsRAP2.6* (*RELATED TO AP2 6*, LOC_Os04g32620) and *OsERF3* (LOC_Os02g43790), respectively (Supplemental Fig. S9B, Supplemental Data Set 10, and Supplemental File 1). To examine if changes in DNA methylation affected the binding ability of ERFs during rice de-domestication, we profiled the DNA binding sites of OsRAP2.6 and OsERF3 through DAP-seq and amplified DNA affinity purification sequencing (ampDAP-seq) using third leaves at the same developmental stage as that used for MethylC-seq (Supplemental Data Set 4) (see Methods). The $(GCC)_5$ motifs or $(NCC)_5$ motifs were enriched among the identified binding sites for OsRAP2.6 and OsERF3, respectively (Fig. 6B and Supplemental Fig. S9C). In addition, OsRAP2.6 and OsERF3 were sensitive to changes in DNA methylation in their binding motifs (Supplemental Fig. S9, D and E). Consistent with this notion, de-domestication-induced changes in DNA methylation were negatively correlated with the binding activities of OsRAP2.6 and OsERF3 (Fig. 6C and Supplemental Fig. S9F). As the expression of *OsRAP2.6* and *OsERF3* is induced by stress (Cao et al. 2005, 2006; Wamaitha et al. 2012), we propose that changes in DNA methylation may contribute to stress resistance changes through affecting binding activities of related transcription factors such as ERFs during de-domestication. We determined that the target genes of OsRAP2.6 and OsERF3 with hypo-DMRs are significantly more highly expressed following de-domestication, while target genes with hyper-DMRs showed no significant changes in their expression (Supplemental Fig. S9G). For example, *PATHOGENESIS-RELATED5* (*PR5*) is associated with disease resistance in *Arabidopsis* (Bozkurt et al. 2007; Wichmann et al. 2011). The rice homolog *OsPR5* (LOC_Os08g04590) exhibited decreased DNA methylation, increased chromatin accessibility, and higher binding activities of OsRAP2.6 and OsERF3 around its transcription start site (TSS) in *WJ* rice compared to *CJ* rice (Fig. 6D). Consequently, *OsPR5* was expressed in most *WJ* accessions and completely repressed in the *CJ* accessions studied (Fig. 6E).

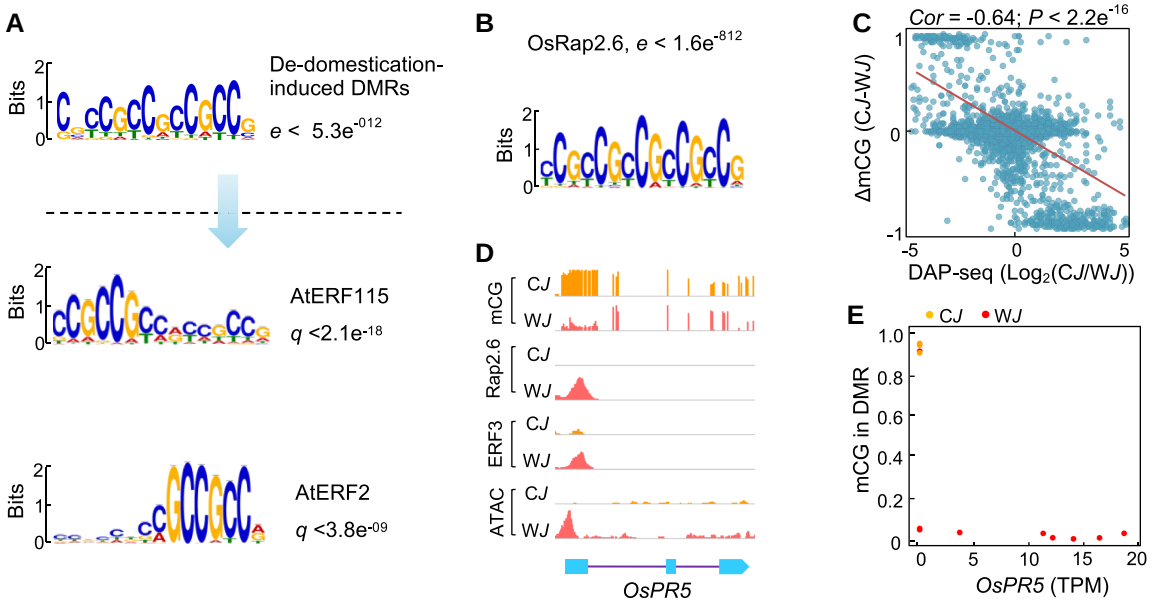


Figure 6. Altered DNA methylation contributed to gene expression changes by affecting the binding ability of ERFs during de-domestication. **A)** DNA motifs enriched in the de-domestication-induced DMRs (upper panel) is similar to the motifs of binding sites for *Arabidopsis* ERF115 and ERF2 in published data sets. The E-value (*e*) indicates an estimate of the expected number of motifs by MEME. The *q*-value (*q*) indicates the probability that a random motif has an optimal alignment to the target motifs. **B)** DNA motifs enriched in peaks called by DAP-seq for OsRap2.6. **C)** Negative correlation between methylation changes and DAP-seq binding changes (RPM) for each peak region called by DAP-seq of OsRap2.6 during de-domestication (*CJ* [Nip] vs *WJ* [H1]). *Cor* = Pearson's correlation coefficient. DAP-seq binding changes were normalized to RPM. **D and E)** An example of de-domestication-induced hypo-DMR associated with increased DAP-seq binding of Rap2.6 and OsERF3 (**D**) on the promoter of LOC_Os08g04590 (*OsPR5*) regulating its expression (**E**) in *WJ* compared to *CJ*. Scale: 0 to 1 for methylation level; 0 to 4 (RPM) for DAP-seq and ATAC-seq data.

Discussion

DNA methylation is essential for genome integrity and affects many biological programs including cancer incidence in humans and organ development in plants (Rakyan et al. 2011; Zhang et al. 2018). The 1,001 epigenomes project for *Arabidopsis* opened the door to understanding population epigenetics during plant evolution (Kawakatsu et al. 2016). However, genome-wide identification of epigenetic variation in natural populations is still limited in crops. In this study, we generated single-base resolution DNA methylomes of 95 wild, cultivated, and weedy rice accessions and uncovered extensive DNA methylation variation among these species.

Genomic variation induced DNA methylation changes during plant evolution (Zhang et al. 2018). In maize, 25.1% to 38.7% of DMRs were associated with SNPs, and an additional 0.3% to 1% of DMRs were associated with structural variants (Xu et al. 2019). In soybean, 22.5% of all DMRs can be explained by local genetic variation (Shen et al. 2018). Although we detected a limited overlap between SNPs and DNA methylation variation in comparisons between cultivated rice and wild rice (Fig. 1D), we cannot exclude *cis* and *trans* effects of structural variation on changes in DNA methylation during rice domestication. The *cis*-regulatory elements in accessible chromatin regions play important roles in regulating gene expression in plant development and environmental responses (Lu et al. 2018). A recent report revealed that most accessible chromatin regions are located in unmethylated genomic regions in plants (Crisp et al. 2020). Although extensive DNA methylation variation occurred in different rice populations, most (80%) DMVs were conserved among the 5 rice groups investigated here (Fig. 4A). Similar to this observation, methylated genes showed a high rate of changes in DNA methylation during long-term evolution of cotton, whereas unmethylated genes stayed unmethylated with little or no change in their DNA methylation (Song et al. 2017). Fewer DNA methylation changes occurred in DMVs during rice domestication, probably because changes in DNA methylation in these DMVs may affect chromatin accessibility and induce significant changes in expression and influence plant development. Indeed, changes in DNA methylation did affect the expression of nearby and distal genes through altering chromatin accessibility and histone modifications during rice domestication and de-domestication (Figs. 2 to 6), although we cannot exclude the potential effects of genome variants in DMRs on gene expression. DNA methylation modifies chromatin accessibility probably through affecting the binding of transcription factors such as ERFs (Fig. 6C). Cytosine methylation in transcription factors binding sites will inhibit their binding (Fig. 6C), as also observed in *Arabidopsis* (O'Malley et al. 2016). These results indicate that variation in DNA methylation may contribute to changes of chromatin accessibility throughout the genome during plant evolution.

De-domestication has led to "wild-like" weedy rice, which is opposite to domestication from wild rice, correlating with extensive changes in DNA methylation among wild, cultivated, and weedy rice. However, hypomethylation sites from wild to cultivated rice and hypermethylation sites from cultivated to weedy rice were largely not shared (Fig. 5C). This observation is consistent with changes in genomic sequences, which differ during rice domestication and de-domestication (Li et al. 2017; Qiu et al. 2020). Interestingly, methylation changes during domestication and de-domestication are both associated with stress-responsive genes, such as *ERF* genes (Fig. 5, E and G). Stress treatment has been shown to induce DNA methylation in plants (Secco et al. 2015; Zheng et al. 2017; Wang et al. 2021). Cultivated crops grow under human-manipulated environments, which will be accompanied by stress resistance trade-offs to favor growth and yield (Li et al. 2017; Wang et al. 2020). DNA methylation related to stress tolerance may passively decrease during rice domestication. By contrast, weedy rice has to survive under adverse stress conditions in natural environment, which may be accompanied by increased DNA methylation related to stress tolerance.

Collectively, our study reveals epigenomic variation among wild, cultivated, and weedy rice and provides insights into epigenetic regulation of gene expression during rice domestication and de-domestication. The identified epialleles underlying agronomic traits are valuable resource for epigenetic breeding to improve rice yield and quality, as well as to prevent and control weedy rice for sustainable agriculture.

Materials and methods

Plant materials and growth conditions

All rice seeds were disinfected by incubation in 10% (v/v) sodium hypochlorite for 10 min before being germinated in a dark chamber at 30 °C for 2 d. Seedlings were transferred to a hydroponics system using 0.5× Kimura B nutrient solution which was changed every 2 d in a growth chamber with 28 °C/24 °C (day/night) and 16-h light/8-h dark (LED light [Guixiang, T20V2]; 400 μmol m^{-2} s^{-1} light intensity) cycles. The third leaf at 30 DAP was collected at ZT4 (Zeitgeber time 4, namely, 4 h after dawn) for isolating DNA and RNA for library construction.

Library construction for MethylC-seq and DNA-seq

Genomic DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method (Allen et al. 2006). About 2 μg genomic DNA was fragmented (300 to 500 bp) with 8 cycles of 30-s on/30-s off using Bioruptor (Diagenode, Denville, NJ, USA), end-repaired and 3′-end adenylated to ligate methylated adapter (GenScript, NJ, China) according to the protocol of the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, Ipswich, MA, USA). For MethylC-seq, half (~0.5 μg) of adapter-ligated DNA fragments was treated with bisulfite using a Zymo EZ DNA Methylation-Gold kit (ZYMO RESEARCH, Irvine, CA, USA), followed by 10 cycles of PCR amplification using KAPA HiFi HotStart ReadyMix (Roche, Basel, Swiss). After purification with VAHTSTM DNA Clean Beads (Vazyme, NJ, China), MethylC-seq libraries were sequenced on a NovaSeq platform (Illumina, San Diego, CA, USA) as 150-bp paired-end reads.

For DNA-seq library, the reminder (~0.5 μg) of the above adapter-ligated DNA fragments was amplified by 6 cycles of PCR using Q5 HiFi HotStart DNA Polymerase (NEB, Ipswich, MA, USA). After purification with VAHTSTM DNA Clean Beads (Vazyme, NJ, China), DNA-seq libraries were sequenced on a NovaSeq platform (Illumina, San Diego, CA, USA) as 150-bp paired-end reads.

SNP calling

After filtering raw sequencing reads by NGSQCtookit v2.3 with default parameters (Patel and Jain 2012), clean reads were mapped against the rice Nip reference genome (MSU7.0) using BWA (v0.7.15) with default parameters (Li and Durbin 2009). Uniquely mapped reads were extracted and potential PCR duplicates were removed using Picard-tools (version 2.0.1). The remaining reads were used for variant detection using the Genome Analysis Toolkit (GATK, version 3.5.0) (McKenna et al. 2010). The SNPs and InDels were separated using "SelectVariants" in GATK. To decrease the variant discovery rate, the SNP calls were filtered according to the following threshold: quality by depth (QD) < 10.0 || mapping quality (MQ) < 20.0 || fisher strand (FS) > 30.0 || symmetric odds ratio (SOR) > 3.0 || rank sum test for MQ (MQRankSum) < −2.5 || rank sum test for read position bias (ReadPosRankSum) < −3.5 || read depth (RD) < 5 || DP > 100.0. The remaining SNPs were combined with published SNPs from the 3 K rice (https://snp-seek.irri.org/_download.zul) for analyses. These SNPs were used to build neighbor-joining trees using MEGAX with 1,000 replicates for bootstrap confidence analysis (Kumar et al. 2018).

Single cytosine methylation calling and population parameter (Pi, π)

After filtering by NGSQCtookit v2.3 default parameters (Patel and Jain 2012), clean MethylC-seq reads were mapped to the Nip genome (MSU7.0) using Bismark (v0.15.0) (Krueger and Andrews 2011; Krueger et al. 2012) with options (–score_min L,0,-0.2 -X 1000). The reads mapped to the same sites were collapsed into a single consensus read to reduce clonal bias. Then, each cytosine site covered by at least 3 uniquely mapped reads was retained. The genomic variants like C->T could be mistakenly interpreted as the conversion of a methylated cytosine to an unmethylated thymine. To reduce confounding effects of genomic variants on DNA methylation analysis, the cytosines affecting SNPs were excluded in the analysis. The mixtools (Benaglia et al. 2009) was used under "normalmixEM" to evaluate thresholds of methylation level for distinguishing unmethylated cytosines in the CG context. The significance thresholds were selected

as follows: unmethylation (≤0.1), heterozygous methylation (>0.1 and <0.8), and methylation (≥0.8). The methylation haplotype (meplotype) was used to calculate the diversity of DNA methylation in the rice population (Xu et al. 2018). In the meplotype analysis, the methylation status of each cytosine was converted to nucleobase based on the threshold of methylation level: <0.1 marked as "TT," homozygous TT representing unmethylated cytosine; 0.1 to 0.8 marked as "CT," heterozygous CT representing partially methylated cytosine; and >0.8 marked as "CC," homozygous CC representing methylated cytosine. The meplotype and above SNPs were used to calculate the population parameter π using VCFtools (0.1.13) (Danecek et al. 2011) in 100-kb windows with a 10-kb step across the genome.

Identification of DMRs and DMVs

To identify DMRs among the 5 rice groups of wild rice (20 accessions), cultivated rice (18 *CJ* accessions and 20 *CI* accessions), and weedy rice (8 *WJ* accession and 10 *WI* accessions), the accessions from each population were considered as repeats, and each CG site detected in more than 3 accessions from each population was considered as effective cytosine sites as previously described (Shen et al. 2018). In brief, the genomic window containing 8 successive effective cytosine sites (<500 bp) were used to identify DMRs. Each CG DMR between 2 populations being compared was determined using the cut-off values of average methylation level differences (>0.4) and the corrected false discovery rate (FDR < 0.05) that were calculated by adjusting *P*-values (from ANOVA test) of the pairwise comparisons using the Benjamini–Hochberg method (Higo et al. 2020).

The identification of each DMV was performed as previously described (Xie et al. 2013). In brief, the genome was divided into 100-bp bins, and the bins with at least 4 CG sites were retained. The remaining bins showing an average methylation level < 0.1 were considered as unmethylated bins. Each window containing at least 10 consecutive unmethylated bins was considered as a DMV.

ATAC-seq library construction and data processing

Construction of ATAC-seq libraries was performed as previous described (Bajic et al. 2018) (Supplemental Data Set 4). Approximately 0.5 g leaves per replicate were cross-linked in 35 mL cross-linking buffer (0.4 M sucrose, 10 mM Tris-HCl [pH = 8.0], 1 mM PMSF, 1 mM EDTA and 1% [w/v] formaldehyde) for 20 min, before stopping the reaction by adding 2.2 mL of 2 M glycine. After wash in sterile water by 3 times, The cross-linked leaves were washed 3 times in sterile water immediately frozen in liquid nitrogen and ground into a fine powder. Nuclei were isolated by mixing the powder to 10 mL nuclei extraction buffer 1 (20 mM MOPS [pH = 8.0], 40 mM NaCl, 90 mM KCl, 2 mM EDTA, 0.5 mM EGTA, 0.5 mM spermidine, 0.2 mM spermine, and 1× Protease Inhibitor Cocktail). The mixture was filtered on a 70-μm strainer and washed with 1 mL nuclei extraction buffer 2 (0.25 M sucrose, 10 mM Tris-HCl [pH = 8.0], 10 mM $MgCl_2$, 1% [v/v] Triton X-100 and 1× Protease Inhibitor Cocktail) and then centrifuged at 12,000 × *g* at 4 °C for 10 min. Following resuspension in 300 μL nuclei extraction buffer 3 (1.7 M sucrose, 10 mM Tris-HCl [pH = 8.0], 2 mM $MgCl_2$, and 0.15% [v/v] Triton X-100 and 1× Protease Inhibitor Cocktail), the nuclei suspension was loaded on the surface of 600 μL nuclei extraction buffer 3 and centrifuged at 16,000 × *g* at 4 °C for 10 min. The purified nuclei were resuspended in 1 mL nuclei extraction buffer 1, stained with 4,6-diamidino-2-phenylindole and loaded into a hemocytometer to count nuclei. About 50,000 nuclei were incubated with Tn5 transposase (Vazyme, NJ, China) at 55 °C for 10 min. Next, cross-linking was reversed by incubation at 65 °C overnight in 50 mM Tris-HCl (pH = 8.0), 1 mM EDTA, 1% (w/v) SDS, 0.2 M NaCl, and 5 ng/mL proteinase K. DNA was recovered with a MinElute PCR Purification Kit (Qiagen) and amplified for 6 cycles. Libraries were sequenced on a NovaSeq platform (Illumina, San Diego, CA, USA) as 150-bp paired-end reads.

After filtering by NGSQCtookit v2.3 with default parameters (Patel and Jain 2012), clean reads were mapped to the Nip reference genome (MSU7.0) using BWA (v0.7.15) (Li and Durbin 2009). Uniquely mapped reads were extracted, and potential PCR duplicates were removed using Picard-tools (version 2.0.1). Read coverage in 50-bp bins across the entire genome was normalized as reads per million mapped reads (RPM) using the remaining paired reads. To show chromatin accessibility changes between DMRs and random genomic regions, the genome was divided into windows with the same average length as the DMRs, and then the same numbers of windows were randomly selected to compare to DMRs.

Library construction and analysis of DAP-seq and ampDAP-seq

DAP-seq libraries were constructed as previously reported (Bartlett et al. 2017) (Supplemental Data Set 4). In brief, about 2 μg genomic DNA was fragmented (300 to 500 bp), end-repaired, and 3′-end adenylated to ligate adapters (NEB, Ipswich, MA, USA). After purification using VAHTS DNA Clean Beads (Vazyme, NJ, China), about 95% of adapter-ligated DNA fragments was used for DAP-seq. The rest (~0.05 μg) of adapter-ligated DNA fragments was amplified to remove DNA methylation information by 12 cycles of PCR using Q5 HiFi HotStart DNA Polymerase (NEB, Ipswich, MA, USA). The DNA was purified with VAHTS DNA Clean Beads (Vazyme, NJ, China) and used for ampDAP-seq as the control of DAP-seq (Bartlett et al. 2017).

The coding sequences of *Rap2.2* and *OSERF3* were individually amplified by PCR using specific primers (Supplemental Data Set 11) and cloned into the pUC18-Halo-ORF vector to generate pUC18-Halo-Rap2.2 and pUC18-Halo-OSERF3. Each recombinant protein was translated from the 2 vectors using TNT SP6 Coupled Wheat Germ Extract System (Promega, Madison, WI, USA). The resulting proteins were immobilized onto Magne HaloTag Beads and incubated

with 100 ng of adapter-ligated DNA fragments for DAP-seq or 100 ng of amplified adapter-ligated DNA fragments for ampDNA-seq. After bead washing with 1× phosphate-buffered saline (pH = 7.4, 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4) containing 0.005% (vol/vol) NP40 solution, DNA binding by proteins immobilized into beads was directly amplified by 20 cycles of PCR using Q5 HiFi HotStart DNA Polymerase (NEB, Ipswich, MA, USA) (Bartlett et al. 2017). After purification, the DAP-seq and ampDAP-seq libraries were sequenced on a NovaSeq platform (Illumina, San Diego, CA, USA) as 150-bp paired-end reads.

After filtering by NGSQCtookit v2.3 with default parameters (Patel and Jain 2012), clean reads were mapped to the Nip reference genome (MSU7.0) using BWA (v0.7.15) (Li and Durbin 2009). Uniquely mapped reads were extracted, and potential PCR duplicates were removed using Picard-tools (version 2.0.1). Read coverage in 50-bp bins along the entire genome was normalized as RPM using the remaining paired reads before being used for peak calling using MACS2 (2.2.6) (Liu 2014). Total peaks were used to analyze DNA motifs using MEME (5.1.0), in which DNA motifs were annotated (Gupta et al. 2007).

Construction and analysis of 3′ RNA-seq libraries

Total RNA was extracted from the third leaf of rice seedling at 30 DAP using TRIzol reagent (Invitrogen, Carlsbad, California, USA), which was collected at ZT4. After DNase I (EN401, Vazyme, China) treatment, aliquots of total RNA from all samples were individually fragmented (100 to 1000 bp) and reverse-transcribed into first-strand cDNA using oligo(dT)-index primers (Supplemental Data Set 11). The first-strand cDNA from each 10 samples were pooled together and subjected to second strand cDNA synthesis, end-repair and adapter ligation. Finally, DNA was amplified by 15 cycles of PCR using Q5 HiFi HotStart DNA Polymerase (NEB, Ipswich, MA, USA). After purification, the 3′ RNA-seq libraries were sequenced on a NovaSeq platform (Illumina, San Diego, CA, USA) as 150-bp paired-end reads.

According to the index of Read 2 from the pair-end reads, Read 1 in each library was divided into corresponding samples and mapped to the Nip reference genome (MSU7.0) using BWA (v0.7.15) (Li and Durbin 2009). After removing potential PCR duplicates, uniquely mapped reads were used to calculate expression levels (in transcripts per million reads [TPM]) for each gene.

Analysis of published MethylC-seq and ChIP-seq data sets

Published epigenetic data sets were downloaded from NCBI under accession numbers PRJNA597475 (Zhao et al. 2020), PRJNA542905 (Zhao et al. 2019), PRJNA597065 (Zhao et al. 2020), and PRJNA122435 (He et al. 2010), covering MethylC-seq of panicles, roots, young leaves and mature leaves (ZS97 and MH63 accessions), and ChIP-seq of histone modification (H3K27ac and H3K4me3) of young leaves (ZS97, MH63, 9311, and Nip accessions) to investigate chromatin state. After filtering by NGSQCtookit v2.3 with default parameters (Patel and Jain 2012), clean MethylC-seq reads were mapped to the Nip reference genome (MSU7.0) using Bismark (v0.15.0) (Krueger and Andrews 2011; Krueger et al. 2012) with options (–score_min L,0,-0.2 -X 1000). The reads mapping to the same sites were collapsed into a single consensus read to reduce clonal bias. Then, each cytosine site covered by at least 3 uniquely mapped reads was retained for analysis.

All raw ChIP-seq reads were parsed using NGSQC Toolkit_v2.3 (Patel and Jain 2012), and then clean reads were mapped to the Nip reference genome (MSU7.0) using BWA (v0.7.15) with default parameters (Li and Durbin 2009). Uniquely mapped reads were extracted to remove potential PCR duplicates using Picard-tools (version 2.0.1). The read coverage was normalized as RPM for each 50-bp bin using the remaining reads.

GWAS for CG methylation variation

A total of 477,318 high-quality SNPs was used after removing variants with >40% missing calls and a minor allele frequency (MAF) <0.05. To control for spurious associations, genetic relatedness was estimated using "GCTA-GRM" (Yang et al. 2011). GWAS was performed using the genome-wide average CG methylation levels as dependent variable-like phenotypes following a mixed linear model-based association analysis (MLMA) in the genome-wide complex trait analysis (GCTA) program (Yang et al. 2011). The significant *P*-value was specified as 1.05×10^{-07} (0.05/SNP number).

Analysis of T-DNA insertions

The flanking sequences of Genoplante (Sallaud et al. 2004) were downloaded from the URL: (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=nucleotide&cmd=search&term=Rice%5BORGN%5D±Genoplante). All flanking sequences were aligned to the Nip reference genome (MSU7.0) using BLASTN (v2.5.0). Mapped sequences with $E < 10^{-5}$ were extracted to identify breakpoints of insertions. These breakpoints were normalized as breakpoints per million for each 50-bp bin.

Identification of intra-chromosomal interactions

Two biological replicates of Hi-C data were combined from Nip seedlings grown at 30 °C (Liu et al. 2017). High-quality reads were mapped to the Nip reference genome (MSU7.0) using HiC-Pro (2.11.1) (Servant et al. 2015). The mapped Hi-C matrices were used to identify intra-chromosomal interaction loops with the "FitHiC" package (2.0.7) ($q <$ 0.05) (Ay et al. 2014).

Acknowledgments

We thank Dr. Xiaoming Zheng and Dr. Zhijian Xu for providing wild rice materials; Dr. Xiangdong Fu for comments on a draft of the manuscript.

Author contributions

S.C. and Q.S. conceived the research. S.C., K.C., K.L., S.C., X.Z., C.S., S.Z., and Y.N. performed the experiments. S.C. analyzed the data. K.C., L.F., and J.X. provided materials, reagents, and technical assistance. S.C., Z.J.C., and Q.S. wrote the manuscript. S.C. and K.C. contributed equally to this work.

Supplemental data

The following materials are available in the online version of this article.

Supplemental Figure S1. CG methylation is more stable across different tissues than CHG and CHH methylation.

Supplemental Figure S2. The phylogenetic tree of 95 rice accessions using in this study.

Supplemental Figure S3. GWAS analysis of global CG methylation levels.

Supplemental Figure S4. Domestication-induced DMRs with chromatin accessibility and histone modification.

Supplemental Figure S5. Distribution of domestication-induced DMRs in rice genome.

Supplemental Figure S6. DNA methylation and histone modification around DMVs in the rice genome.

Supplemental Figure S7. Increased DNA methylation levels during de-domestication.

Supplemental Figure S8. Distribution of DMRs induced by de-domestication.

Supplemental Figure S9. Altered DNA methylation contributed to gene expression changes by affecting the binding ability of ERFs during de-domestication.

Supplemental Data Set 1. Methylomes for 59 accessions of cultivated rice.

Supplemental Data Set 2. DNA methylomes for wild and weedy rice.

Supplemental Data Set 3. Domestication-induced DMRs.

Supplemental Data Set 4. WGS, ATAC-seq, DAP-seq, ChIP-seq, and 3′RNA-seq used in this study.

Supplemental Data Set 5. Genes associated with DMRs induced by domestication of rice.

Supplemental Data Set 6. Gene ontology enrichment of genes associated with DMRs induced by domestication of rice.

Supplemental Data Set 7. De-domestication-induced DMRs.

Supplemental Data Set 8. Genes associated with DMRs induced by de-domestication of rice.

Supplemental Data Set 9. Enrichment fold changes of genes related to DMRs on the different pathways.

Supplemental Data Set 10. The alignments of ERF proteins in *Arabidopsis* and rice.

Supplemental Data Set 11. Primer used in this study.

Supplemental Data Set 12. Statistical analyses in this study.

Supplemental File 1. Machine-readable tree file of ERF protein phylogenetic analysis with Newick format.

Funding

Funding was provided by Hainan Yazhou Bay Seed Lab (B21HJ0216), the Agricultural Science and Technology Innovation Program and the Cooperation and Innovation Mission (CAAS-ZDXT202001) and the Natural Science Foundation of Jiangsu Province (BK20190510).

Conflict of interest statement. None declared.

Data availability

All high-throughput sequencing data derived from MethylC-seq, DNA-seq, 3′ RNA-seq, ATAC-seq, and DAP-seq were deposited into the Genome Sequence Archive (GSA) in BIG Data Center (https://bigd.big.ac.cn/gsa/index.jsp) under accession numbers (PRJCA003490). All codes are available upon request or at GitHub (https://github.com/EPICS135/).

References

Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nat Protoc. 2006:**1**(5):2320–2325. https://doi.org/10.1038/nprot.2006.384

Ay F, Bailey TL, Noble WS. Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. Genome Res. 2014:**24**(6):999–1011. https://doi.org/10.1101/gr.160374.113

Bajic M, Maher KA, Deal RB. Identification of open chromatin regions in plant genomes using ATAC-seq. Methods Mol Biol. 2018:**1675**:183–201. https://doi.org/10.1007/978-1-4939-7318-7_12

Bartlett A, O'Malley RC, Huang SC, Galli M, Nery JR, Gallavotti A, Ecker JR. Mapping genome-wide transcription-factor binding sites using DAP-seq. Nat Protoc. 2017:**12**(8):1659–1672. https://doi.org/10.1038/nprot.2017.055

Becker C, Hagmann J, Muller J, Koenig D, Stegle O, Borgwardt K, Weigel D. Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. Nature 2011:**480**(7376):245–249. https://doi.org/10.1038/nature10555

Benaglia T, Chauveau D, Hunter DR, Young DS. Mixtools: an R package for analyzing finite mixture models. J Stat Softw. 2009:**32**(6):1–29. https://doi.org/10.18637/jss.v032.i06

Bozkurt O, Unver T, Akkaya MS. Genes associated with resistance to wheat yellow rust disease identified by differential display analysis. Physiol Mol Plant. 2007:**71**(4–6):251–259. https://doi.org/10.1016/j.pmpp.2008.03.002

Cao YF, Song FM, Goodman RM, Zheng Z. Molecular characterization of four rice genes encoding ethylene-responsive transcriptional factors and their expressions in response to biotic and abiotic stress. J Plant Physiol. 2006:**163**(11):1167–1178. https://doi.org/10.1016/j.jplph.2005.11.004

Cao YF, Wu YF, Zheng Z, Song FM. Overexpression of the rice EREBP-like gene OsBIERF3 enhances disease resistance and salt tolerance in transgenic tobacco. Physiol Mol Plant. 2005:**67**(3–5):202–211. https://doi.org/10.1016/j.pmpp.2006.01.004

Chen M, Lin JY, Hur J, Pelletier JM, Baden R, Pellegrini M, Harada JJ, Goldberg RB. Seed genome hypomethylated regions are enriched in transcription factor genes. Proc Natl Acad Sci U S A. 2018:**115**(35):E8315–E8322. https://doi.org/10.1073/pnas.1811017115

Cheng C, Tarutani Y, Miyao A, Ito T, Yamazaki M, Sakai H, Fukai E, Hirochika H. Loss of function mutations in the rice

chromomethylase OsCMT3a cause a burst of transposition. Plant J. 2015:**83**(6):1069–1081. https://doi.org/10.1111/tpj.12952

Chinnusamy V, Zhu JK. RNA-directed DNA methylation and demethylation in plants. Sci China C Life Sci. 2009:**52**(4):331–343. https://doi.org/10.1007/s11427-009-0052-1

Cho SM, Shin SH, Kim KS, Kim YC, Eun MY, Cho BH. Enhanced expression of a gene encoding a nucleoside diphosphate kinase 1 (OsNDPK1) in rice plants upon infection with bacterial pathogens. Mol Cells. 2004:**18**(3):390–395.

Choi JY, Platts AE, Fuller DQ, Hsing YI, Wing RA, Purugganan MD. The rice paradox: multiple origins but single domestication in Asian rice. Mol Biol Evol. 2017:**34**(4):969–979. https://doi.org/10.1093/molbev/msx049

Crisp PA, Marand AP, Noshay JM, Zhou P, Lu Z, Schmitz RJ, Springer NM. Stable unmethylated DNA demarcates expressed genes and their *cis*-regulatory space in plant genomes. Proc Natl Acad Sci U S A. 2020:**117**(38):23991–24000. https://doi.org/10.1073/pnas.2010250117

Cubas P, Vincent C, Coen E. An epigenetic mutation responsible for natural variation in floral symmetry. Nature 1999:**401**(6749):157–161. https://doi.org/10.1038/43657

Cui X, Cao X. Epigenetic regulation and functional exaptation of transposable elements in higher plants. Curr Opin Plant Biol. 2014:**21**:83–88. https://doi.org/10.1016/j.pbi.2014.07.001

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. The variant call format and VCFtools. Bioinformatics 2011:**27**(15):2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Guo LB, Qiu J, Li LF, Lu BR, Olsen K, Fan LJ. Genomic clues for crop-weed interactions and evolution. Trends Plant Sci. 2018:**23**(12):1102–1115. https://doi.org/10.1016/j.tplants.2018.09.009

Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. Genome Biol. 2007:**8**(2):R24. https://doi.org/10.1186/gb-2007-8-2-r24

He GM, Zhu XP, Elling AA, Chen LB, Wang XF, Guo L, Liang MZ, He H, Zhang HY, Chen FF, et al. Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. Plant Cell 2010:**22**(1):17–33. https://doi.org/10.1105/tpc.109.072041

Higo A, Saihara N, Miura F, Higashi Y, Yamada M, Tamaki S, Ito T, Tarutani Y, Sakamoto T, Fujiwara M, et al. DNA methylation is reconfigured at the onset of reproduction in rice shoot apical meristem. Nat Commun. 2020:**11**(1):4079. https://doi.org/10.1038/s41467-020-17963-2

Hu L, Li N, Xu C, Zhong S, Lin X, Yang J, Zhou T, Yuliang A, Wu Y, Chen YR, et al. Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. Proc Natl Acad Sci U S A. 2014:**111**(29):10642–10647. https://doi.org/10.1073/pnas.1410761111

Hua L, Wang DR, Tan LB, Fu YC, Liu FX, Xiao LT, Zhu ZF, Fu Q, Sun XY, Gu P, et al. LABA1, A domestication gene associated with long, barbed awns in wild rice. Plant Cell 2015:**27**(7):1875–1888. https://doi.org/10.1105/tpc.15.00260

Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, et al. A map of rice genome variation reveals the origin of cultivated rice. Nature 2012:**490**(7421):497–501. https://doi.org/10.1038/nature11532

Kaut O, Schmitt I, Tost J, Busato F, Liu Y, Hofmann P, Witt SH, Rietschel M, Frohlich H, Wullner U. Epigenome-wide DNA methylation analysis in siblings and monozygotic twins discordant for sporadic Parkinson's disease revealed different epigenetic patterns in peripheral blood mononuclear cells. Neurogenetics 2017:**18**(1):7–22. https://doi.org/10.1007/s10048-016-0497-x

Kawakatsu T, Huang SSC, Jupe F, Sasaki E, Schmitz RJ, Urich MA, Castanon R, Nery JR, Barragan C, He YP, et al. Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. Cell 2016:**166**(2):492–505. https://doi.org/10.1016/j.cell.2016.06.044

Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. Bioinformatics 2011:**27**(11):1571–1572. https://doi.org/10.1093/bioinformatics/btr167

Krueger F, Kreck B, Franke A, Andrews SR. DNA methylome analysis using short bisulfite sequencing data. Nat Methods. 2012:**9**(2):145–151. https://doi.org/10.1038/nmeth.1828

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018:**35**(6):1547–1549. https://doi.org/10.1093/molbev/msy096

Li CB, Zhou AL, Sang T. Rice domestication by reducing shattering. Science 2006:**311**(5769):1936–1939. https://doi.org/10.1126/science.1123604

Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 2009:**25**(14):1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li LF, Li YL, Jia Y, Caicedo AL, Olsen KM. Signatures of adaptation in the weedy rice genome. Nat Genet. 2017:**49**(5):811–814. https://doi.org/10.1038/ng.3825

Lim KY, Kovarik A, Matyasek R, Chase MW, Clarkson JJ, Grandbastien MA, Leitch AR. Sequence of events leading to near-complete genome turnover in allopolyploid Nicotiana within five million years. New Phytol. 2007:**175**(4):756–763. https://doi.org/10.1111/j.1469-8137.2007.02121.x

Lisch D. Epigenetic regulation of transposable elements in plants. Ann Rev Plant Biol. 2009:**60**:43–66. https://doi.org/10.1146/annurev.arplant.59.032607.092744

Liu C, Cheng YJ, Wang JW, Weigel D. Prominent topologically associated domains differentiate global chromatin packing in rice from *Arabidopsis*. Nat Plants. 2017:**3**(9):742–748. https://doi.org/10.1038/s41477-017-0005-9

Liu T. Use model-based analysis of ChIP-seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells. Methods Mol Biol. 2014:**1150**:81–95. https://doi.org/10.1007/978-1-4939-0512-6_4

Lu Z, Ricci WA, Schmitz RJ, Zhang X. Identification of *cis*-regulatory elements by chromatin structure. Curr Opin Plant Biol. 2018:**42**:90–94. https://doi.org/10.1016/j.pbi.2018.04.004

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010:**20**(9):1297–1303. https://doi.org/10.1101/gr.107524.110

Meyer RS, Choi JY, Sanches M, Plessis A, Flowers JM, Amas J, Dorph K, Barretto A, Gross B, Fuller DQ, et al. Domestication history and geographical adaptation inferred from a SNP map of African rice. Nat Genet. 2016:**48**(9):1083–1088. https://doi.org/10.1038/ng.3633

Molina J, Sikora M, Garud N, Flowers JM, Rubinstein S, Reynolds A, Huang P, Jackson S, Schaal BA, Bustamante CD, et al. Molecular evidence for a single evolutionary origin of domesticated rice. Proc Natl Acad Sci U S A. 2011:**108**(20):8351–8356. https://doi.org/10.1073/pnas.1104686108

Muller M, Munne-Bosch S. Ethylene response factors: a key regulatory hub in hormone and stress signaling. Plant Physiol. 2015:**169**(1):32–41. https://doi.org/10.1104/pp.15.00677

Nakagawa H, Tanaka A, Tanabata T, Ohtake M, Fujioka S, Nakamura H, Ichikawa H, Mori M. Short grain1 decreases organ elongation and brassinosteroid response in rice. Plant Physiol Bioch. 2012:**158**(3):1208–1219. https://doi.org/10.1104/pp.111.187567

O'Malley RC, Huang SC, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR. Cistrome and epicistrome features shape the regulatory DNA landscape. Cell 2016:**165**(5):1280–1292. https://doi.org/10.1016/j.cell.2016.04.038

Ono A, Yamaguchi K, Fukada-Tanaka S, Terada R, Mitsui T, Iida S. A null mutation of ROS1a for DNA demethylation in rice is not transmittable to progeny. Plant J. 2012:**71**(4):564–574. https://doi.org/10.1111/j.1365-313X.2012.05009.x

Patel RK, Jain M. NGS QC toolkit: a toolkit for quality control of next generation sequencing data. Plos One 2012:**7**(2):e30619. https://doi.org/10.1371/journal.pone.0030619

Qiu J, Jia L, Wu D, Weng X, Chen L, Sun J, Chen M, Mao L, Jiang B, Ye C, et al. Diverse genetic mechanisms underlie worldwide convergent rice feralization. Genome Biol 2020:**21**(1):70. https://doi.org/10.1186/s13059-020-01980-x

Qiu J, Zhou Y, Mao L, Ye C, Wang W, Zhang J, Yu Y, Fu F, Wang Y, Qian F, et al. Genomic variation associated with local adaptation of weedy rice during de-domestication. Nat Commun. 2017:**8**(1):15323. https://doi.org/10.1038/ncomms15323

Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat Rev Genet. 2011:**12**(8):529–541. https://doi.org/10.1038/nrg3000

Richards EJ. Inherited epigenetic variation–revisiting soft inheritance. Nat Rev Genet. 2006:**7**(5):395–401. https://doi.org/10.1038/nrg1834

Sallaud C, Gay C, Larmande P, Bes M, Piffanelli P, Piegu B, Droc G, Regad F, Bourgeois E, Meynard D, et al. High throughput T-DNA insertion mutagenesis in rice: a first step towards in silico reverse genetics. Plant J. 2004:**39**(3):450–464. https://doi.org/10.1111/j.1365-313X.2004.02145.x

Secco D, Wang C, Shou HX, Schultz MD, Chiarenza S, Nussaume L, Ecker JR, Whelan J, Lister R. Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. Elife 2015:**4**:e09343. https://doi.org/10.7554/eLife.09343

Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, Heard E, Dekker J, Barillot E. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. 2015:**16**(1):259. https://doi.org/10.1186/s13059-015-0831-x

Shen Y, Zhang J, Liu Y, Liu S, Liu Z, Duan Z, Wang Z, Zhu B, Guo YL, Tian Z. DNA Methylation footprints during soybean domestication and improvement. Genome Biol. 2018:**19**(1):128. https://doi.org/10.1186/s13059-018-1516-z

Song Q, Guan X, Chen ZJ. Dynamic roles for small RNAs and DNA methylation during ovule and fiber development in allotetraploid cotton. PLoS Genet. 2015:**11**(12):e1005724. https://doi.org/10.1371/journal.pgen.1005724

Song Q, Zhang T, Stelly DM, Chen ZJ. Epigenomic and functional analyses reveal roles of epialleles in the loss of photoperiod sensitivity during domestication of allotetraploid cottons. Genome Biol. 2017:**18**(1):99. https://doi.org/10.1186/s13059-017-1229-8

Wamaitha MJ, Yamamoto R, Wong HL, Kawasaki T, Kawano Y, Shimamoto K. Osrap2.6 transcription factor contributes to rice innate immunity through its interaction with receptor for activated kinase-C 1 (RACK1). Rice 2012:**5**(1):35. https://doi.org/10.1186/1939-8433-5-35

Wang L, Cao S, Wang P, Lu K, Song Q, Zhao FJ, Chen ZJ. DNA Hypomethylation in tetraploid rice potentiates stress-responsive gene expression for salt tolerance. Proc Natl Acad Sci U S A. 2021:**118**(13):e2023981118. https://doi.org/10.1073/pnas.2023981118

Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z, Li M, Zheng T, Fuentes RR, Zhang F, et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature 2018:**557**(7703):43–49. https://doi.org/10.1038/s41586-018-0063-9

Wang Z, Hong Y, Zhu G, Li Y, Niu Q, Yao J, Hua K, Bai J, Zhu Y, Shi H, et al. Loss of salt tolerance during tomato domestication conferred by variation in a Na(+)/K(+) transporter. EMBO J. 2020:**39**(10):e103256. https://doi.org/10.15252/embj.2019103256

Wichmann F, Asp T, Widmer F, Kolliker R. Transcriptional responses of Italian ryegrass during interaction with Xanthomonas translucens pv. graminis reveal novel candidate genes for bacterial wilt resistance. Theor Appl Genet. 2011:**122**(3):567–579. https://doi.org/10.1007/s00122-010-1470-y

Win KT, Yamagata Y, Doi K, Uyama K, Nagai Y, Toda Y, Kani T, Ashikari M, Yasui H, Yoshimura A. A single base change explains the independent origin of and selection for the nonshattering gene in African rice domestication. New Phytol. 2017:**213**(4):1925–1935. https://doi.org/10.1111/nph.14290

Wu D, Lao S, Fan L. De-domestication: an extension of crop evolution. Trends Plant Sci. 2021:**26**(6):560–574. https://doi.org/10.1016/j.tplants.2021.02.003

Wu Y, Zhao S, Li X, Zhang B, Jiang L, Tang Y, Zhao J, Ma X, Cai H, Sun C, Tan L. Deletions linked to PROG1 gene participate in plant architecture domestication in Asian and African rice. Nat Commun. 2018:**9**:4157.

Xie W, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, Whitaker JW, Tian S, Hawkins RD, Leung D, et al. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. Cell. 2013:**153**(5):1134–1148. https://doi.org/10.1016/j.cell.2013.04.022

Xu J, Chen G, Hermanson PJ, Xu Q, Sun CS, Chen WQ, Kan QX, Li MQ, Crisp PA, Yan JB, et al. Population-level analysis reveals the widespread occurrence and phenotypic consequence of DNA methylation variation not tagged by genetic variation in maize. Genome Biol. 2019:**20**(1):243. https://doi.org/10.1186/s13059-019-1859-0

Xu J, Zhao L, Liu D, Hu S, Song X, Li J, Lv H, Duan L, Zhang M, Jiang Q, et al. EWAS: epigenome-wide association study software 2.0. Bioinformatics 2018:**34**(15):2657–2658. https://doi.org/10.1093/bioinformatics/bty163

Xu X, Liu X, Ge S, Jensen JD, Hu F, Li X, Dong Y, Gutenkunst RN, Fang L, Huang L, et al. Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. Nat Biotechnol. 2011:**30**(1):105–111. https://doi.org/10.1038/nbt.2050

Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011:**88**(1):76–82. https://doi.org/10.1016/j.ajhg.2010.11.011

Yu XW, Zhao ZG, Zheng XM, Zhou JW, Kong WY, Wang PR, Bai WT, Zheng H, Zhang H, Li J, et al. A selfish genetic element confers non-Mendelian inheritance in rice. Science 2018:**360**(6393):1130–1132. https://doi.org/10.1126/science.aar4279

Zhang C, Wei Y, Xu L, Wu KC, Yang L, Shi CN, Yang GY, Chen D, Yu FF, Xie Q, et al. A Bunyavirus-inducible ubiquitin ligase targets RNA polymerase IV for degradation during viral pathogenesis in rice. Mol Plant. 2020:**13**(6):836–850. https://doi.org/10.1016/j.molp.2020.02.010

Zhang H, Lang Z, Zhu JK. Dynamics and function of DNA methylation in plants. Nat Rev Mol Cell Biol. 2018:**19**(8):489–506. https://doi.org/10.1038/s41580-018-0016-z

Zhao L, Wang S, Cao Z, Ouyang W, Zhang Q, Xie L, Zheng R, Guo M, Ma M, Hu Z, et al. Chromatin loops associated with active genes and heterochromatin shape rice genome architecture for transcriptional regulation. Nat Commun. 2019:**10**(1):3640. https://doi.org/10.1038/s41467-019-11535-9

Zhao L, Xie L, Zhang Q, Ouyang W, Deng L, Guan P, Ma M, Li Y, Zhang Y, Xiao Q, et al. Integrative analysis of reference epigenomes in 20 rice varieties. Nat Commun. 2020:**11**(1):2658. https://doi.org/10.1038/s41467-020-16457-5

Zheng XG, Chen L, Xia H, Wei HB, Lou QJ, Li MS, Li TM, Luo LJ. Transgenerational epimutations induced by multi-generation drought imposition mediate rice plant's adaptation to drought condition. Sci Rep-Uk. 2017:**7**:39843. https://doi.org/10.1038/srep39843