Accessible chromatin regions and their functional interrelations with gene transcription and epigenetic modifications in sorghum genome

Chao Zhou1,*, Zhu Yuan1, Xueping Ma2, Huilan Yang3, Ping Wang3, Lanlan Zheng2, Yonghong Zhang2,4,* and Xiaoyun Liu3,*

1Key Laboratory of Three Gorges Regional Plant Genetics and Germplasm Enhancement (CTGU), Biotechnology Research Center, Yichang Key Laboratory of Omics-Based Breeding for Chinese Medicines, China Three Gorges University, Yichang 443002, China
2Laboratory of Medicinal Plants, Institute of Basic Medical Sciences, School of Basic Medicine, Biomedical Research Institute, Hubel Key Laboratory of Embryonic Stem Cell Research, Hubel University of Medicine, Shiyan 442000, China
3Institute for Interdisciplinary Research, Jianghan University, Wuhan 430056, China
4Hubei Key Laboratory of Wudang Local Chinese Medicine Research, Hubei University of Medicine, Shiyan 442000, China

*Correspondence: Chao Zhou (zhouchao@ctgu.edu.cn), Yonghong Zhang (zhangyh@hbmu.edu.cn), Xiaoyun Liu (liuxiaoyun@jhun.edu.cn)

https://doi.org/10.1016/j.xplc.2020.100140

ABSTRACT

Accessible chromatin regions (ACRs) provide physical scaffolds to recruit transcriptional co-regulators and displace their nearby nucleosomes in multiple plant species. Characterization of ACRs and investigation of their biological effects in Sorghum bicolor has lagged behind. Regulation of gene expression relies on the transcriptional co-regulators that are recruited to ACRs to affect epigenomic modifications of surrounding nucleosomes. In this study, we employed transposase-accessible chromatin sequencing to identify ACRs and decipher how the presence of ACRs affects gene expression and epigenetic signatures in the Sorghum genome. As a result, 21 077 ACRs, which are mapped to 22.9% of genes and 2.7% of repeats, were identified. The profiling of ACRs on gene structures reveals a narrow and sharp peak around the transcription start site, with relatively weak and broad signals covering the entire gene body and an explicit but wide peak from the transcription termination site to its downstream regions. We discovered that the correlations between gene expression levels and profiled ACR densities are dependent on the positions of ACRs. The occurrence of genic ACRs cumulatively enhances the transcriptional activity of intergenic ACR-associated genes. In addition, an intricate crosstalk among ACRs, gene expression, and epigenetic marks has been unveiled by integrating multiple-omics analyses of whole-genome bisulfite sequencing, 6mA immunoprecipitation followed by sequencing, RNA sequencing, chromatin immunoprecipitation sequencing, and DNase I hypersensitive sites sequencing datasets. Our study provides a genome-wide landscape of ACRs in sorghum, decrypts their interrelations with various epigenetic marks, and sheds new light on their roles in transcriptional regulation.

Key words: Sorghum bicolor, accessible chromatin regions, transcriptional regulation, epigenetic mark

Zhou C., Yuan Z., Ma X., Yang H., Wang P., Zheng L., Zhang Y., and Liu X. (2021). Accessible chromatin regions and their functional interrelations with gene transcription and epigenetic modifications in sorghum genome. Plant Comm. 2, 100140.

INTRODUCTION

Chromatin accessibility represents the degree to which nuclear macromolecules physically contact chromatinized DNA and are topologically organized by nucleosomes and other chromatin-binding factors (Klemm et al., 2019). The topological organization of nucleosomes across the genome is not uniform: while densely wrapping within facultative and constitutive heterochromatin, histones are depleted at regulatory loci, including within intergenic regions and transcribed gene bodies (Thurman et al., 2012). In addition, the landscape of chromatin accessibility broadly reflects regulatory capacity and a non-static biophysical state, which

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.
Plant Communications

is reversely a critical determinant of chromatin organization and function (Klemm et al., 2019).

Most recently, transposase-accessible chromatin sequencing (ATAC-seq) has been developed as a reliable tool for profiling of accessible chromatin regions (ACRs) across multiple plant species and cell types, requiring less labor and smaller amounts of starting nuclei than DNase I treatment of nuclei coupled with high-throughput sequencing (DNase-seq) (Buenrostro et al., 2015). In ATAC-seq, isolated nuclei were treated with an engineered Tn5 transposase to cleave accessible DNA and insert adapters for high-throughput sequencing (Buenrostro et al., 2013). Applications of ATAC-seq have been successfully utilized in *Oryza sativa* (Wilkins et al., 2016), *Arabidopsis thaliana* (Lu et al., 2017; Sijacic et al., 2018), *Medicago truncatula* (Maher et al., 2018), *Solanum lycopersicum* (Maher et al., 2018), *Sorghum bicolor* (Lu et al., 2019), and others, demonstrating that the highest abundance of open chromatin regions lies in the promoter in plants. For instance, ATAC-seq profiling in *Arabidopsis* has supported that ACRs are highly correlated with DNase hypersensitive sites (DHSs) (Freirichs et al., 2019). Moreover, it has been shown that ACRs are associated with H3K9ac and H3K27ac, but not with H3K4me1, in maize (Lu et al., 2019; Ricci et al., 2019), suggesting that the interplay of ACRs and histone marks is worth unveiling in plants.

Lately, a substantial number of cis-regulatory elements (CREs) has been identified through a set of comparative genomic and epigenomic analyses in 13 plant species, including sorghum (Lu et al., 2019), revealing the association of ACRs not only with genome size, sequence conservation, and histone modifications, but also with putative distal CREs. Intriguingly, the results suggested that, distinct from genic ACRs (gACRs) and proximal ACRs (pACRs), which enrich particular active transcription-associated histone modifications, including H3K4me3, H3K36ac, and H3K36me3, distal ACRs (dACRs) can pose an unmodified state, an H3K56ac-modified state that is associated with active transcription of nearby genes as an enhancer, or an H3K27me3-modified state that is probably involved in Polycomb silencing pathways as a repressor (Lu et al., 2019). In accordance, enhancers harboring dACRs in rice are always characterized by H3K27me3 and H3K4me3 and/or H3K27ac (Sun et al., 2019). In maize, however, long-range chromatin loops accumulating between dACRs or between dACRs and genes function to control transcription activity (Ricci et al., 2019). Thus, the identification of distinct chromatin features using ATAC-seq-coupled strategies will provide novel genomic and epigenomic insights into how chromatin pathways regulate plant gene expression and development.

It has been reported that 5-methylcytosine (5mC), which represents the most abundant type of DNA methylation, is depleted in ACRs in plants (Burgess et al., 2019; Lu et al., 2019; Ricci et al., 2019). N^6^-methyladenine (6mA), a non-canonical DNA methylation, preferentially locates at linker DNA between two adjacent nucleosomes (Fu et al., 2015) and decorates ACRs, which can be recognized by Tn5 transposase cleavage (Buenrostro et al., 2013). Moreover, 6mA has been identified as a unique DNA mark that contributes to nucleosome positioning and transcription initiating in unicellular green algae (Fu et al., 2015). In our previous study, by using 6mA immunoprecipitation followed by sequencing (6mA-IP-seq) coupled with micrococcal nuclease sequencing, we found that 6mA-marked genes displayed much higher phasing of nucleosome arrays downstream of the transcription start site (TSS), suggesting that 6mA may have a function in phasing the nucleosome arrays to promote gene expression in rice (Zhou et al., 2018). 6mA is involved in gene expression regulation by either inactivating (Zemach et al., 2010b) or activating (Wang et al., 2019), when it locates in either a promoter or a gene body, respectively. In addition, 6mA and CG methylation at moderate levels in gene bodies have additive effects on gene activity (Zhou et al., 2018). Interestingly, the coordination of 6mA and the histone modification H3K4me2 is related to trans-generationally epigenetic controls in worms (Greer et al., 2015). Although emerging evidence has hinted that 5mC and 6mA are both involved in controlling gene expression activities and are combined with other histone modifications (Greer et al., 2015; Zhou et al., 2018), it is yet to discover whether and to what extent ACRs function in plant transcription regulation.

*S. bicolor* L. is the fifth most important cereal crop in the world. It provides a genetic model for C4 grasses with advantages of a relatively small genome (~800 Mb), diploid genetics, diverse germplasm, and collinearity with other C4 grass genomes (Paterson et al., 2009). The predicted nucleosome occupancy likelihoods in sorghum are similar to those in maize, suggesting that the distributions may vary across each chromosome in both genomes. Interestingly, nucleosomes positioned immediately downstream of the TSS present at different densities across chromosomes (McCormick et al., 2018). However, despite sorghum having a relatively uniform pattern of nucleosome organization (McCormick et al., 2018), we still lack evidence to explain variations in gene or repeat density across each chromosome in sorghum.

With the advance of plant epigenomics, genome-wide characterization of chromatin accessibility and chromatin modifications has been lately probed in sorghum (Burgess et al., 2019; Lu et al., 2019), unveiling evolutionarily conserved functions of ACRs and some chromatin marks in plants. However, specific chromatin regulatory patterns and mechanisms in the unique sorghum genome need to be addressed. Furthermore, it is ambiguous whether and to what extent the involvement of dynamic interactions of ACRs and epigenetic modifications takes responsibility for chromatin accessibility in sorghum.

Here, we employed ATAC-seq-coupled strategies to identify ACRs from aboveground tissues of sorghum seedlings to characterize their genomic distributions and organizations. We studied the correlations between different types of ACRs and gene expression activity. We describe the prevalence and cross talk of ACRs and various epigenetic signatures to unveil distinct chromatin pathways of gene expression regulation. In total, our data provide a genome-wide landscape of ACRs, decrypt their interrelations with epigenetic modifications, and shed light on their transcription regulation roles in sorghum.
Identification of ACRs in the *Sorghum* genome

To enable genome-wide identification of ACRs in sorghum, we applied ATAC-seq in aboveground tissues containing young leaves and stems of *S. bicolor* (L.) cultivar BTx623, a model cultivar. As a result, 40 201, 36 087, 39 423, and 37 214 peaks were called, respectively, from the four replicates of ATAC-seq that showed strong Pearson correlation coefficients (>0.9) (Supplemental Table 1; Supplemental Figure 1A), suggesting the data outputs are reliable, with a high reproducibility. Subsequently, we identified 21 077 ACRs derived from these peaks (Supplemental Table 2). To validate the reliability of the ACRs identified from ATAC-seq, we gradually increased the P threshold (from >e^{-2} to >e^{-20}) to call DHSs from published DNase-seq (GSE97369) (Burgess et al., 2019) to make comparison with ACRs called from our ATAC-seq. We found that the overlapping ratios increased with P value, which was far greater than that at random (Supplemental Figure 2A). In addition, we introduced the Genrich software download from Github (https://github.com/jsh58/Genrich), an alternative analysis tool for ATAC-seq, to reevaluate the ACRs recognized by our ATAC-seq datasets. Consistently, more than 90% of peaks (11 294 of 12 257) called by Genrich were overlapped with ACRs (Supplemental Figure 2B). Moreover, about 10 000 ACRs were also represented in the latest published ATAC-seq dataset from leaves of 7-day-old sorghum (GSE128434) (Supplemental Figure 2C) (Lu et al., 2019). Together, our data provide an overview of repeatable and reliable ACRs in the sorghum genome.

In sorghum, the majority of heterochromatin appeared in the pericentromeric regions, with the exception of chromosome 6, which has a distinct pattern in which the entire right (short) arm is highly heterochromatinized (Paterson et al., 2009). We observed that ACRs tend to be enriched in euchromatic regions (Figure 1A). Next, we made a comparative analysis by including a genome-wide H3K27me3 profiling dataset (a euchromatic gene mark; Figure 1A; GSE128434) and a genome-wide 5mC profiling dataset (a heterochromatic gene mark; Figure 1A; Supplemental Table 1) of the same sorghum tissues. The cumulative distribution of 5mC was significantly close to the heterochromatic and pericentromeric regions, whereas peaks of H3K27me3 were largely detected in euchromatic regions that locate in distal ends of the chromosomes (Figure 1A), which is consistent with our previous observation in rice (Tan et al., 2016). In detail, ACRs occurred in 22.9% of genes and 2.7% of repeats (Figure 1B). ACRs were located in much higher proportions at regions of promoter, intergenic, and 3’UTR than at other regions in gene bodies (Figure 1B). Metaplots of ACR density showed that ACRs over gene-coding regions were not distributed evenly: while merely weak signals cover the entire gene body, a narrow and sharp peak presents around the TSS and an explicit but wide curve appears from the transcription termination site (TTS) to its downstream regions (Figure 1C), which is in line with the observation that merely 3.9% of ACRs occurred in the 3’UTR (Figure 1B).
The genomic distribution of ACRs at downstream transcribed ends is distinct in sorghum compared with *Phaseolus vulgaris* and *Populus trichocarpa*, in which the regions were seen as apparent peaks of ACRs (Lu et al., 2019). In Glycine max, however, ACRs were absent at TTSs (Lu et al., 2019). Altogether, the genomic distribution of ACRs at TSSs is conserved, but the profiling of ACRs at the downstream transcribed end of genes appears to be largely divergent among plant species.

**The involvement of ACRs in gene transcription regulation in sorghum**

The promotional and repressive roles of ACRs in gene expression have been raised through either the binding of nucleosome-displacing transcription factors and/or chromatin remodelers (Klemm et al., 2019) or the presence of DNA sequences that are resistant to nucleosome assembly (Segal et al., 2006). The mechanism of ACRs in gene transcription regulation in sorghum, however, remains unclear. To address this issue, a set of RNA-sequencing (RNA-seq) data from three replicates of aboveground tissues at the same age ($R^2 > 0.9$; Supplemental Figure 1C and Table 1) were obtained. It was revealed that active genes that showed higher expression levels (top 20%) exhibited typical sharp peaks of ACRs at the TSS in comparison with inactive genes, which showed weak expression levels (bottom 20%); however, no obvious peaks were observed at regions of TTSs (Figure 2A). These results suggested that the accumulation of ACRs at the gene TSS contributed to positively controlling the gene transcription in sorghum.

Next, we classified ACRs based on their proximity to the nearest annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as gACRs when they overlapped the nearby annotated genes for at least 1 bp, and 11 633 (55.19%) were designated as intergenic regions (iACRs) because they were spaced at least 1 bp from nearby genes (Figure 2B; Supplemental Table 3). For iACRs, a total number of 6657 (31.58%) that showed distances over 2 kb from their nearest gene were designated as dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as iACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we dACRs, and 4976 (23.61%) that showed distances less than 2 kb from their nearest gene were designated as gACRs when they overlapped the nearby annotated genes. As a consequence, 9444 (44.81%) of the ACRs were designated as pACRs (Figure 2B). Accordingly, we...
Figure 2. ACRs are associated with gene transcription regulation.

(A) ACR density plots and heatmaps of ACR read counts across genomic regions ranging between 2 kb upstream of the TSS and 2 kb downstream of the TTS. The regions in the heatmap are ranked from highest FPKM values (top) to the lowest (bottom) of RNA-seq. Active genes (top 20%; blue) and inactive genes (bottom 20%; green) are denoted.

(B) Numbers of different types of ACRs. The ACRs are categorized on the basis of their proximity to the nearest annotated genes. gACR, genic ACRs that are overlapping with the nearby annotated genes for at least 1 bp. iACR, intergenic ACRs that are spaced at least 1 bp from the nearby genes; iACRs consist of proximal ACRs and distal ACRs.

(C) Boxplots of expression levels of genes associated with different types of ACRs. **P < 0.05, two-sided Wilcoxon rank sum test. NS, no significance. The five statistical values of the boxplot from the top to the bottom are the maximum, the third quartile, the median, the first quartile, and the minimum. The center line is the median, the box limits are the upper and lower quartiles, and the whiskers are the 1.5 times interquartile ranges. Gene numbers of different groups are shown. **P < 0.01; igACR, both intergenic and genic ACR.

(D) Snapshots representing gene expression of different types of ACR-associated genes (scaffold: chr1, 1190–1260 kb). Sb.001G012800 is an igACR-related gene; Sb.001G12900 and Sb.001G13400 are gACR-related genes; Sb.001G13200 is a non-ACR-related gene; Sb.001G13500 and Sb.001G13800 are iACRs. Their genomic regions are enclosed in different-colored squares. The gene expression levels of transcriptome reads are scaled on the y axis (values ranging from 0 to 100). ACRs are graphed as red boxes on top of their corresponding gene structural models (blue).
Figure 3. Crosstalk of ACRs and DNA cytosine methylation.

(A) Average methylation levels at CG, CHG, and CHH sites across genomic regions ranging between 2 kb upstream of the TSS and 2 kb downstream of the TTS. Blue and green lines are genes associated with ACRs or not, respectively.

(B) Average methylation levels of CG, CHG, and CHH sites of different types of ACR-associated genes.

(C) Average methylation levels in all sequence contexts over four groups of genes with varying expression levels. “1st” to “4th” represents the gene expression levels from the highest to the lowest. Genomic regions range between 2 kb upstream of the TSS and 2 kb downstream of the TTS.

(D and E) Boxplots showing gene expression levels (FPKM) of ACR-associated genes with CG methylation (CG+; D) and CHH methylation (CHH+; E). CG–, without CG. CHH–, without CHH. **P < 0.01, *P < 0.05, two-sided Wilcoxon rank sum tests. NS, no significance. The five statistical values of the boxplot from the top to the bottom are the maximum, the third quartile, the median, the first quartile, and the minimum. The center line is the median, the box limits are the upper and lower quartiles, and the whiskers are the 1.5 times interquartile ranges. Gene numbers of different groups are shown.
The coordination of sorghum ACRs and epigenetic marks

Zhou et al., 2016), probably by affecting the binding affinity of different regulators (Zhang et al., 2018a). As anticipated, overall, genes with high CHH methylation in flanking regions were coupled with significantly higher gene expression levels (Figure 3C). It is well known that CHH methylation mainly targets gene-associated short transposable elements (TEs), thereby affecting the nearby gene transcription level (Tan et al., 2016). It was clear that short TEs were located at higher frequencies at the 5’ and 3’ ends of genes in sorghum compared with long TEs (Supplemental Figure 5B), which was in line with the distribution of TEs in rice (Zemach et al., 2010a; Tan et al., 2016). Boxplot analysis revealed that genes covered by CHH-methylated TEs were more active compared with those without CHH methylation (Supplemental Figure 5C; Supplemental Table 5). In addition, the occurrence of short TEs was sharply depleting within but peaked at both ends of ACRs (Supplemental Figure 5D), presenting a similar association pattern of CHH methylation around ACRs (Supplemental Figure 4A). Our observation reinforces the previous finding that short TEs are highly methylated at CHH and their distribution closely parallels that of CHH methylation (Zemach et al., 2010a).

To obtain a comprehensive and figurative view, we have included a genomic snapshot to visualize CHH methylation, ACR status, and gene expression over a 200 kb region at chromosome 2 (scaffold: 6600–6900 kb), which contains five ACR-associated genes (Sb002G067050, Sb002G067400, Sb002G068300, Sb002G068701, and Sb002G069000) and another five genes (Sb002G066800, Sb002G067200, Sb002G067800, Sb002G069100, and Sb002G069166) that contain no ACRs (Supplemental Figure 5E). It is shown that the former five genes associating with at least one ACR were decorated to a greater extent with CHH methylation than the latter five ACR-free genes (Supplemental Figure 5E). The coincident observation that ACR-associated genes (ACR+) were accompanied by higher CHH levels than ACR− genes is shown as well in Figure 3A and 3B. In all, both CG methylation at gene bodies and CHH methylation within flanking transcription regions (where ACRs are associated with short TEs) were positively correlated with gene transcription.

To determine whether position effects exist for ACRs that are marked by CG and CHH methylation on controlling gene expression, we examined transcription levels of categories of gACR- and iACR-associated genes that were targeted by CG methylation and found that they all accumulated at higher mRNA levels in comparison with the non-CG-methylation control. Moreover, upon CG methylation within gene bodies, gACRs were prone to promoting expression of these iACR-characterized genes (Figure 3D). By contrast, upon CHH methylation in the flanking regions of genes, iACRs would promote expression of these gACR-characterized genes (Figure 3E). These analyses provide clues that the regulatory action of gene-body CG and flanking transcription-region CHG methylation contributing to the gene expression may rely on positional effects of ACRs.

DNA N6-methyladenine affects ACRs

A non-canonical DNA methylation, 6mA, has been identified to play important roles in regulating gene expression in plants (Zhou et al., 2018). Meanwhile, it has been reported that about 0.35% of all adenines in the sorghum genome are 6mA methylated (Zhang et al., 2018b). To explore potential interactions of 6mA and ACRs, we sampled leaves at the same stage and performed 6mA-IP-seq to conduct comparative analysis (Supplemental Table 1). At first, 36 858, 31 562, 24 061, and 25 249 peaks were identified in the four replicates of 6mA-IP-seq (Supplemental Table 1). The replicates show high Pearson correlation coefficients (>0.8) (Supplemental Figure 1B). About 25.8% of annotated genes and 5.7% of repeats were marked by 6mA (Supplemental Figure 6A). The 6mA occurred most frequently at GAGG motifs (Supplemental Figure 6B). The 6mA level was higher in promoters and intergenic regions than in gene bodies (Supplemental Figure 6A), absent at TSSs, but sharply increasing at TTSs (Supplemental Figure 6C). The results also indicated that 6mA in the gene body is correlated with gene expression activity (Supplemental Figure 6C). Collectively, our data suggest that the role of 6mA in regulating gene expression in sorghum resembles that in rice (Zhou et al., 2018).

Then, we examined how and to what extent ACRs affected 6mA enrichment. The presence of 6mA was attributed to higher densities of both gACRs (Figure 4A) and iACRs (Figure 4B). Next, we investigated how differently positioned ACRs are correlated with the enrichment of 6mA. Interestingly, we found that the enrichment of 6mA was present at a significantly greater extent at the TTS of iACR-associated genes compared with other types (Figure 4C). By contrast, there was no difference in the enrichment of 6mA at either the TSS or the gene bodies between genes with or without any type of ACRs (Figure 4C). It was demonstrated, however, that the combined effects of 6mA and all types of ACRs promoted gene expression significantly, while different ACRs may contribute at varying degrees (Figure 4D).

Given that 6mA and CG methylation at moderate levels in gene bodies have positive effects on the gene activity regulation in rice (Zhou et al., 2018), we asked what the cross talk was like between 6mA and two types of DNA methylation in sorghum. It turned out that the 6mA-marked genes apparently showed a dramatic increase in CG methylation and CHG methylation within the gene body (Supplemental Figure 7A), which represented the highest gene expression levels in comparison with only CG methylation or only 6mA-marked genes (Supplemental Figure 7B; Supplemental Table 6). In particular, it was suggested that the co-occupancy of CG methylation and 6mA within the gene body of each of 7491 genes drastically elevated the peak of ACR densities (Supplemental Figure 7C and 7D). For instance, the expression levels of these ACR-characterized genes were elevated (Supplemental Figure 7E). Moreover, we targeted Sobic.005G053200, a gene encoding an NB-ARC domain-containing disease-resistance protein, to monitor its DNA methylation (5mC and 6mA) levels and ACR densities in leaves and roots (Supplemental Figure 8) to discriminate potential differences in their patterns in distinct tissues. As shown by the Genome Browser (Supplemental Figure 8A and 8B), the previous BS-seq data (GSE70903) suggested that the levels of CG methylation and CHH methylation were higher in roots compared with leaves, which is evidently confirmed by the McrBC- and HaeIII-digested PCR tests (Supplemental Figure 8C). Meanwhile, DNase-digested PCR analysis showed...
that the target gene had many more open chromatin regions in roots than in leaves (Supplemental Figure 8D). Also, we used 6mA-IP-qPCR to detect 6mA enrichment in separate tissues. As a result, we observed that the 6mA levels in roots were significantly increased in three regions of Sobic.005G053200 (Supplemental Figure 8E). Accordingly, the quantitative real-time PCR result showed that the expression level of Sobic.005G053200 was higher in roots than in leaves (Supplemental Figure 8F). Collectively, our data supported the idea that different ACR densities are correlated with distinct gene expression activities of a decorated gene. The presence of both CG methylation and 6mA within gene bodies is associated with highly expressed genes, the process of which also involves ACR-associated transcription regulation.

**Interrelationships between ACRs and chromatin signatures**

To investigate the chromatin epigenome signatures in relation to ACRs in the sorghum genome, we integrated our ATAC-seq data with published data, including DNase-seq (GSE97369) (Burgess et al., 2019), chromatin immunoprecipitation sequencing (ChiP-seq) (H2A.Z, H3K4me1, H3K4me3, H3K27me3, H3K36me3, and H3K56ac) (GSE128434) (Lu et al., 2019), and RNA-seq. The profiling was executed after the designation of three ACR categories, gACRs (5,208), iACRs (4,373), and igACRs (2,209), as judged above. The heatmap diagrams supported again that ACRs called from our ATAC-seq were highly consistent with those identified from published DNase-seq in sorghum (Figure 5A). In contrast to the enhancer-specific distribution of H3K4me1 in mammals (Heintzman et al., 2007), our result showed that H3K4me1 was scarcely present in these regions where any types of ACRs were enriched in sorghum (Figure 5A). As well-known active gene markers, H3K4me3 and H3K56ac were enriched at transcriptional sites, whereas H3K36me3 was seen in the gene body (Figure 5A). Conclusively, these histone covalent modifications, most of which are conserved active markers in plant species, prefer to be positively correlated with ACRs in control of gene expression.

By contrast, inactive genes associated with all kinds of ACRs are characterized by H3K27me3 and the histone variant H2A.Z (Figure 5A; Supplemental Figure 9). The expression analysis of gACR-associated genes revealed that H3K27me3 plays a predominant role in restraining gene activities (Figure 5B). Consistently, gene ontology (GO) analysis agreed that these gACR-associated genes characterized by H3K27me3 and H2A.Z were enriched in the categories of antioxidant activity, transcription regulator activity, and catalytic process ($P < 0.01$), and these were characterized by only H3K27me3, which was also involved in the metabolic process (Figure 5C; Supplemental Table 4). By contrast, the gACR-associated genes characterized by only H2A.Z represent no significant GO enrichment (Figure 5C). Altogether, the analysis supports that H3K27me3 may act as a major inactive epigenetic mark in the ACR-related gene transcription regulation process.

**DISCUSSION**

In the present study, we began with ATAC-seq to characterize distribution patterns of ACRs in the sorghum genome, and then dissected the chromatin roles of ACRs by including genome-wide transcriptional profiling as well as characterization of multiple epigenetic modifications. Our data revealed that the ACRs were widely distributed across all chromosomes and were relatively enriched in the euchromatin of the sorghum genome (Figure 1A). The results indicate that nucleosome
Figure 5. Interrelationships between ACRs and chromatin signatures.

(A) Average plots and heatmaps showing associations of gene expression with different categories of ACRs with various chromatin signatures (DHS, H2A.Z, H3K4me1, H3K4me3, H3K27me3, H3K36me3, and H3K56ac). The genomic regions range between 1 kb upstream (−1 kb) of TSS and 1 kb downstream (+1 kb) of TTS. Read densities (normalized read counts) were calculated with a sliding window using our ATAC-seq dataset and published sequencing data of DNase-seq (GSE97369), ChIP-seq (GSE128434), and RNA-seq. The blue bar denotes levels of ACR enrichment in each category. The red bar denotes levels of enrichment of each chromatin mark or RNA-seq read. Active genes (top 20%; blue) and inactive genes (bottom 20%; green) are denoted.

(B) Boxplots showing expression levels of gACR-associated genes that are enriched by different combinations of H3K27me3 and H2A.Z. The five statistical values of the boxplot from the top to the bottom are the maximum, the third quartile, the median, the first quartile, and the minimum. The center

(legend continued on next page)
positions are not the only determinants of ACRs, which are influenced by epigenome marks. Notably, ACR density was sharply increased at TSSs, which nucleosomes occupy frequently (McCormick et al., 2018), but was mildly increased at TTSs (Figure 1C). This pattern is in accordance with the previous study by Lu and colleagues (Lu et al., 2019). Also, we noticed that in some reported plant species, A. thaliana, Asparagus officinalis, and Eutrema salsugineum, for example, ACRs prefer to locate upstream of TSSs (Lu et al., 2019). We still lack evidence to clarify what produces these differences. One explanation is that some unique features of the genome structure and chromatin environment of the cereal plant sorghum might cause distinctions between it and others. Despite the fact that accessible chromatin sequences undergo prevalence and evolutionary dynamics in plants, the distribution of ACRs at the transcriptional ends was largely divergent from some eukaryotic species, in contrast to the ACRs at TSSs (Lu et al., 2019).

Our results showed that ACRs were associated with 22.9% of genes and 2.7% of repeats, as they preferred to be enriched in promoters and intergenic regions in the sorghum genome (Figure 1B). Thus, we concluded that the mechanism of action of ACRs ought to be highly associated with transcriptions regulatory behaviors, which is in line with a recent report (Ricci et al., 2019). Our data indicated that the accumulation of ACRs at gene TSSs contributed to positively controlling the gene transcription in sorghum, whereas the correlations were not observed for those genes whose ACRs were located at TTSs (Figure 2A). In addition, it has been suggested that dACRs, which are mostly enriched in euchromatin, act as remote CREs in physically mediating the expression activities of their supposed targets through the formation of chromatin loops (Ricci et al., 2019). Not surprisingly, our findings pointed out that when iACRs activated the expression of their associated genes, the inclusion of a nearby gACR would even amplify the effects (Figure 2C), highlighting that positional effects of ACRs contribute to gene activity in sorghum.

In agreement with previous reports (Lu et al., 2019; Oka et al., 2017; Rodgers-Melnick et al., 2016), our investigations in sorghum provided clues that epigenomic modifications, DNA cytosine methylation in particular, can be recruited as transcriptional co-regulators to nucleosomes that are adjacent to nearby ACRs across plant species, which suggests novel regulatory mechanisms of ACRs. Although recent studies have claimed that ACRs are depleted from 5mC around TSSs in plants (Lu et al., 2019), how different types of cytosine methylation interact with ACRs remains questionable. To this end, we introduced the profiling of DNA methylations (5mC and 6mA) in sorghum to explore their dynamic cross talk through comparative analyses. In our results, a positive correlation was found, that the ACR densities partially rely on cytosine methylation, especially gene-body CG methylations or gene-flanking CHH methylation (Figure 3A and Supplemental Figure 4); however, the enrichment of 5mC was opposite to that of ACRs at TSSs (Figure 3). Moreover, short TEs were found to assist their associated CG methylation and CHH methylation, in gene bodies and gene flanking regions, respectively, to activate gene expression (Supplemental Figure 5B). One possible explanation is that gene-body CG methylation is capable of preventing aberrant gene transcription initiation, thus facilitating pre-mRNA splicing (Zilberman et al., 2008; Zhang et al., 2018a), and CHH methylation within the transcriptional flanking regions, where the gene-associated short TEs are mainly located (Tan et al., 2018), affects the binding affinity of different regulators (Zhou et al., 2016; Lang et al., 2017; Zhang et al., 2018a).

Also, we found that all types of ACRs in sorghum were depleted of DNA cytosine methylation compared with their surrounding regions (Supplemental Figure 4), which is in line with previous reports in other plant species (Lu et al., 2019; Ricci et al., 2019). In addition, our data showed that CG and CHH methylation levels of genes with ACRs were higher than those of genes without ACRs (Figure 3A and 3B), suggesting that the presence of an ACR may contribute to the enrichment of CG and CHH methylation in ACR-associated genes. It has been known that some transcription co-regulators, Polycomb-repressive complex 2 (PRC2), for instance, will be recruited to nucleosomes to incorporate epigenome modifications, such as H3K27me3 in the case of PRC2 (Liu et al., 2015). In our previous study, we identified that SDG711, one homologous member of Enhancer of zeste, acting as a co-player of PRC2, binds directly to chromatin and interacts physically with the DNA methyltransferase DRM2 to regulate DNA methylation (Liu et al., 2014; Zhou et al., 2016). Therefore, we assumed that ACRs may provide chromatin environments for recruiting DNA methylation transferases and/or related co-regulators to establish cytosine methylation in their flanking or associated gene regions.

Previous investigation has identified 6mA, a non-canonical DNA modification, which functions as a gene-expression-associated epigenomic mark in higher plants (Zhou et al., 2016). Herein, our findings furthermore suggested that the model of actions of 6mA in sorghum appeared to resemble those in the rice plant (Supplemental Figure 6) (Zhou et al., 2016). Actually, we found that the addition of all types of ACRs displayed positive associations with gene expression (Figure 4F), suggesting that they are both beneficial for activating gene expression. In summary, both CG methylation and 6mA, present within gene bodies, are associated with ACRs in transcription regulation of active genes (Figures 3D, 3F, and 4F). In other words, we propose an intensive chromatin pathway in which DNA methylation (5mC and 6mA) acts in addition to ACRs to control gene transcription in sorghum.

In addition, histone modifications and histone variants have been said to be indispensable for the formation of accessible chromatin (Klemm et al., 2019; Lu et al., 2019). For instance, ACRs have been shown to be linked to H3K9ac and H3K27ac, but not H3K4me1, in maize (Oka et al., 2017; Ricci et al., 2019). The
The coordination of sorghum ACRs and epigenetic marks

latest comprehensive study that includes multiple plant species has demonstrated that H3K56ac-associated dACRs, regarded as active enhancers, are often related to the active transcription status of nearby genes, whereas H3K27me3-associated dACRs are related to Polycomb silencing (Lu et al., 2019). Using self-transcribing active regulatory region sequencing, researchers found in rice that enhancers are also labeled by H3K27me3 and H3K4me3 and/or H3K27ac (Sun et al., 2019). Likewise, histone modifications, including H3K4me3, H3K36me3, and H3K56ac, serving as conserved active marks in plant species, were found to be in positive relation to ACRs in controlling gene expression (Figure 5A and Supplemental Figure 9). On the other hand, H3K27me3 was the predominant repressive mark in controlling gene expression associated with ACRs (Figure 5B and Supplemental Figure 9). In conclusion, our study sheds light on the characterization of the ACRs in the sorghum genome, unveils their interplay with multiple epigenetic markers, and suggests their positive roles in the regulation of gene transcription and expression in higher plants.

METHODS

Plant materials and growing conditions

The background of S. bicolor (L.) Moench plants used in this study was BTx623. Seedlings were grown in one-half-strength Murashige and Skoog medium under 16 h light/8 h dark cycles at 30°C. Leaves of 12-day-old seedlings were harvested and frozen immediately in liquid nitrogen before follow-up experiments.

ATAC-seq and analysis

ATAC of frozen samples was performed according to the ATAC-seq protocol from Shanghai Jiayin Biotechnology Ltd (Shanghai). In brief, native nuclei were purified from frozen samples as previously described (Corces et al., 2017). The Nextera DNA Library Preparation Kit (Illumina) was used to perform the transposition according to the manufacturer’s manual. Fifty thousand nuclei were pelleted and resuspended with transposase for 30 min at 37°C. The transposed DNA fragments were purified immediately using a MinElute PCR Purification Kit (Qiagen). After that, samples were PCR-amplified using 1 x NEBNext High-Fidelity PCR Master Mix (New England Biolabs, MA). Subsequent libraries were purified with the MinElute PCR Purification Kit (Qiagen) and subjected to sequencing on an Illumina Novaseq 6000 using PE150.

Raw reads were trimmed using Trim_galore with filter parameters (–q 25 –phred33 –length 35 -e 0.1 –stringency 4) (https://github.com/FelixKrueger/TrimGalore). The reference genomes were obtained from JGI (https://genome.jgi.doe.gov/portal/). Clean reads were aligned to the reference genome using Bowtie v.1.1.2 (Langmead et al., 2009) using the following parameters: "bowtie -X 1000 -v 1 -d 3 -q 1 -v 2 -b -strata." Aligned reads were sorted using Samtools v.1.3.1 (Li et al., 2009). Clonal duplicates were removed using Picard v.2.16.0 (http://broadinstitute.github.io/picard/).

RNA-seq and analysis

Total RNA (10 mg) was used to purify poly(A) mRNA. mRNA was used for the synthesis and amplification of complementary DNA. The RNA-seq libraries were prepared using the TruSeq RNA Sample Preparation Kit from Illumina. Libraries were sequenced on an Illumina HiSeq. The experiments, including library construction, and sequencing were performed at Annoroad Gene Technology Co. Ltd (Beijing). RNA-seq data were first cleaned by removing contaminations and low-quality reads by Trimmomatic (Bolger et al., 2014). The clean data were aligned with TopHat 2.0 (Kim et al., 2013) using the JGI S. bicolor (BTx623) genome as the reference under default parameters. Then, the gene counts were calculated with Cufflink (Trapnell et al., 2010).

6mA-IP-seq and analysis

6mA-IP-seq library preparation was constructed as previously described (Fu et al., 2015). Briefly, 5 μg of genomic DNA was fragmented by sonication to a mean size of approximately 200–400 bp, followed by end repair, A-base tailing, and adapter ligation using the Acegen DNA Library Prep Kit (Acegen, Cat. No. AG0810) according to the manufacturer’s protocol. The adaptor-ligated DNA was immunoprecipitated by the anti-6mA antibody (Synaptic Systems, Cat. No. 202011). The captured DNA was then purified and amplified by 14–17 cycles of PCR using Illumina 8 bp dual index primers. Simultaneously, the input library was obtained from 5–10 ng of the ligated DNA before immunoprecipitation. The constructed libraries were then analyzed by an Agilent 2100 Bioanalyzer and finally sequenced on Illumina platforms using a 150 x 2 paired-end sequencing protocol. The experiments, including 6mA IP, library construction, and sequencing, were performed at Shenzhen Aecgen Technology Co., Ltd (Shenzhen).

Low-quality reads were removed from raw data using the fastp package (Chen et al., 2018), and the clean data were mapped to the JGI S. bicolor reference genome using BWA (Li and Durbin, 2009) with default parameters. Then MACS (Zhang et al., 2008) was introduced to locate enriched regions to call 6mA peaks by comparing reads from the IP sample with the input sample.

For 6mA-IP-qPCR, independent sorghum leaf tissues were collected and used. The fold-enrichment of each fragment was determined by quantitative real-time PCR (Supplemental Table 7).

BS-seq and analysis

Whole-genome BS-seq libraries were constructed using the Acegen Bisulfite-Seq Library Prep Kit (Acegen, Cat. No. AG0311). Briefly, 1 μg of genomic DNA spiked with 1 ng unmethylated Lambda DNA was fragmented by sonication to a mean size of approximately 200–500 bp and then end-repaired, 5' phosphorylated, 3'-dA-tailed, and ligated to 5mC-modified adapters. After bisulfite treatment, the DNA was amplified with 10 cycles of PCR using Illumina 8 bp dual index primers. The constructed WGBS libraries were then analyzed by an Agilent 2100 Bioanalyzer and finally sequenced on Illumina platforms using a 150 × 2 paired-end sequencing protocol. The library was prepared and sequenced at Shenzhen Aecgen Technology Co., Ltd. After the removal of raw data by fastp, the clean data were mapped to the rice reference genome by Bismark (Krueger and Andrews, 2011). Only uniquely mapped reads were retained for further analysis. DNA methylation levels in three contexts were calculated by bismark_methylation_extractor.

ChIP-seq and DNase-seq data processing

We downloaded previously published datasets of ChIP-seq concerning H3K4me1, H3K4me2, H3K27me3, H3K36me3, H3K56ac, and H2A.Z (Lu et al., 2019), and DNase-seq in sorghum (Burgess et al., 2019). After cleaning with fastp to remove duplication reads, clean reads were mapped to the JGI S. bicolor reference genome using BWA. The DeepTools 2.0 (Ramirez et al., 2016) software was used to generate a heatmap of different epi-modifications.

Identification of ACRs

Processing of call ACRs was conducted according to protocols described previously, with slight modifications (Lu et al., 2019). MACS2 was used to define ACRs with the “--keepup all” function. To find high-quality ACRs, the following filtering steps were generally performed: (1) peaks called with MACS2 were split into 50 bp windows with 25 bp steps; (2) the Trn integration frequency in each window was calculated and normalized to the average frequency in the total genome; (3) windows passing the integration frequency cut-off were merged together with 150 bp gaps; and (4)
The coordination of sorghum ACRs and epigenetic marks

ACKNOWLEDGMENTS
The authors declare no competing interests.

Received: October 27, 2020
Revised: December 23, 2020
Accepted: December 29, 2020
Published: December 31, 2020

REFERENCES

Bailey, T.L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics 27:1653–1659.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30:2114–2120.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods 10:1213–1218.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: a method for assaying chromatin accessibility genome-wide. Curr. Protoc. Mol. Biol. 109:21.29.21–21.29.29.

Burgess, S.J., Reyna-Llorens, I., Stevenson, S.R., Singh, P., Jaeger, K., and Hibberd, J.M. (2019). Genome-wide transcription factor binding in leaves from C(3) and C(4) grasses. Plant Cell 31:2297–2314.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat. Methods 14:959–962.

Frechichs, A., Engelhorn, J., Aitnulmer, J., Gutierrez-Marcos, J., and Werr, W. (2019). Specific chromatin changes mark lateral organ founder cells in the Arabidopsis inflorescence meristem. J. Exp. Bot. 70:3867–3879.

Fu, Y., Luo, G.-Z., Chen, K., Deng, X., Yu, M., Han, D., Hao, Z., Liu, L., Lu, X., Dore, L.C., et al. (2015). N6-methyldeoxyadenosine marks active transcription start sites in Chlamydomonas. Cell 161:879–892.

Greer, E.L., Blanco, M.A., Gu, L., Sendinc, E., Liu, J., Aristizabal-Corrales, D., Hsu, C.-H., Aravind, L., He, C., and Shi, Y. (2015). DNA methylation on N6-adenine in C. elegans. Cell 161:868–878.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. 39:311–318.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14:R36.

Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. Nat. Rev. Genet. 20:207–220.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics 27:1571–1572.

Lang, Z., Wang, Y., Tang, K., Tang, D., Datsenka, T., Cheng, J., Zhang, Y., Handa, A.K., and Zhu, J.-K. (2017). Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. Proc. Natl. Acad. Sci. U S A 114:E4611.
The coordination of sorghum ACRs and epigenetic marks

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10:R25.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMTools. Bioinformatics 25:2078–2079.

Liu, X., Zhou, C., Zhao, Y., Zhou, S., Wang, W., and Zhou, D.-X. (2014). The rice enhancer of zeste [E(z)] genes SDG711 and SDG718 are respectively involved in long day and short day signaling to mediate the accurate photoperiod control of flowering time. Front. Plant Sci. 5:591.

Liu, X., Zhou, S., Wang, W., Ye, Y., Zhao, Y., Xu, Q., Zhou, C., Tan, F., Cheng, S., and Zhou, D.-X. (2015). Regulation of histone methylation and reprogramming of gene expression in the rice inflorescence meristem. Plant Cell 27:1428–1444.

Lu, Z., Hofmeister, B.T., Vollmers, C., DuBois, R.M., and Schmitz, R.J. (2017). Combining ATAC-seq with nucleotide sorting for discovery of cis-regulatory regions in plant genomes. Nucleic Acids Res. 45:e41.

Lu, Z., Marand, A.P., Ricci, W.A., Ethridge, C.L., Zhang, X., and Schmitz, R.J. (2019). The prevalence, evolution and chromatin signatures of plant regulatory elements. Nat. Plants 5:1250–1259.

Maher, K.A., Bajic, M., Kajala, K., Reynoso, M., Pauluzzi, G., West, D.A., Zumstein, K., Woodhouse, M., Bubb, K., Dörity, M.W., et al. (2018). Profiling of accessible chromatin regions across multiple plant species and cell types reveals common gene regulatory principles and new control modules. Plant Cell 30:15.

McCormick, R.F., Truong, S.K., Sreedasyam, A., Jenkins, J., Shu, S., Sims, D., Kennedy, M., Amirebrahimi, M., Weers, B.D., McKinley, B., et al. (2018). The Sorghum bicolor reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. Plant J. 93:338–354.

Oka, R., Zicola, J., Weber, B., Anderson, S.N., Hodgman, C., Gent, J.I., Wesselinink, J.-G., Springer, N.M., Hoefsloot, H.C.J., Turck, F., et al. (2017). Genome-wide mapping of transcriptional enhancer candidates using DNA and chromatin features in maize. Genome Biol. 18:137.

Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellisten, U., Mitros, T., Poliakov, A., et al. (2009). The Sorghum bicolor genome and the diversification of grasses. Nature 457:551–556.

Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Klipert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44:W160–W165.

Ricci, W.A., Lu, Z., Ji, L., Marand, A.P., Ethridge, C.L., Murphy, N.G., Noshay, J.M., Galli, M., Mejía-Guerra, M.K., Colomé-Tatché, M., et al. (2019). Widespread long-range cis-regulatory elements in the maize genome. Nat. Plants 5:1237–1249.

Rodgers-Melnick, E., Vera, D.L., Bass, H.W., and Buckler, E.S. (2016). Open chromatin reveals the functional maize genome. Proc. Natl. Acad. Sci. U S A 113:E3177.

Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thästrom, A., Field, Y., Moore, I.K., Wang, J.-P.Z., and Widom, J. (2006). A genomic code for nucleosome positioning. Nature 442:772–778.

Sijacic, P., Bajic, M., McKinney, E.C., Meagher, R.B., and Deal, R.B. (2018). Changes in chromatin accessibility between Arabidopsis stem cells and mesophyll cells illuminate cell type-specific transcription factor networks. Plant J. 94:215–231.

Sun, J., He, N., Niu, L., Huang, Y., Shen, W., Zhang, Y., Li, L., and Hou, C. (2019). Global quantitative mapping of enhancers in rice by STARR-seq. Genomics Proteomics Bioinformatics 17:140–153.

Tan, F., Zhou, C., Zhou, Q., Zhou, S., Yang, W., Zhao, Y., Li, G., and Zhou, D.-X. (2016). Analysis of chromatin regulators reveals specific features of rice DNA methylation pathways. Plant Physiol. 171:2041.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature 489:75–82.

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W., and Su, Z. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res. 45:W122–W129.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28:511–515.

Wang, Y., Song, F., Zhu, J., Zhang, S., Yang, Y., Chen, T., Tang, B., Dong, L., Ding, N., Zhang, Q., et al. (2017). GSA: genome sequence archive. Genomics Proteomics Bioinformatics 15:14–18.

Wang, Y., Wang, X., Lee, T.-H., Mansoor, S., and Paterson, A.H. (2013). Gene body methylation shows distinct patterns associated with different gene origins and duplication modes and has a heterogeneous relationship with gene expression in Oryza sativa (rice). New Phytol. 198:274–283.

Wilkins, O., Hafemeister, C., Plessis, A., Holloway-Phillips, M.-M., Pham, G.M., Nicotra, A.B., Gregorio, G.B., Jagadish, S.V.K., Septiningsih, E.M., Bonneau, R., et al. (2016). EGRNs (Environmental Gene Regulatory Influence Networks) in rice that function in the response to water deficit, high temperature, and agricultural environments. Plant Cell 28:2365–2384.

Zemach, A., Kim, M.Y., Silva, P., Rodrigues, J.A., Dotson, B., Brooks, M.D., and Zilberman, D. (2010a). Local DNA hypomethylation activates genes in rice endosperm. Proc. Natl. Acad. Sci. U S A 107:18729.

Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. (2010b). Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science 328:916.

Zhang, H., Lang, Z., and Zhu, J.-K. (2018a). Dynamics and function of DNA methylation in plants. Nat. Rev. Mol. Cell Biol. 19:489–506.

Zhang, Q., Liang, Z., Cui, X., Ji, C., Li, Y., Zhang, P., Liu, J., Riaz, A., Yao, P., Liu, M., et al. (2018b). N6-Methyladenine DNA methylation in Japonica and Indica rice genomes and its association with gene expression, plant development, and stress responses. Mol. Plant 11:1492–1508.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChiP-seq (MACS). Genome Biol. 9:R137.

Zhou, C., Wang, C., Liu, H., Zhou, Q., Liu, Q., Guo, Y., Peng, T., Song, J., Zhang, J., Chen, L., et al. (2018). Identification and analysis of adenine N6-methylation sites in the rice genome. Nat. Plants 4:554–563.

Zhou, S., Liu, X., Zhou, C., Zhou, Q., Zhao, Y., Li, G., and Zhou, D.-X. (2016). Cooperation between the H3K27me3 chromatin mark and non-CG methylation in epigenetic regulation. Plant Physiol. 172:1131.

Zilberman, D., Coleman-Derr, D., Ballinger, T., and Henikoff, S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. Nature 456:125–129.

Plant Communications 2, 100140, January 11 2021 © 2020 The Authors.