



Profiling of H3K4me3 and H3K27me3 and Their Roles in Gene Subfunctionalization in Allotetraploid Cotton

Aicen Zhang^{1†}, Yangyang Wei^{2†}, Yining Shi^{1†}, Xiaojuan Deng^{3†}, Jingjing Gao^{1†}, Yilong Feng¹, Dongyang Zheng¹, Xuejiao Cheng¹, Zhaoguo Li², Tao Wang², Kunbo Wang⁴, Fang Liu^{4*}, Renhai Peng^{2*} and Wenli Zhang^{1*}

¹ State Key Laboratory for Crop Genetics and Germplasm Enhancement, JCIC-MCP, CIC-MCP, Nanjing Agricultural University, Nanjing, China, ² Biological and Food Engineering, Anyang Institute of Technology, Anyang, China, ³ College of Agronomy, Xinjiang Agricultural University, Ürümqi, China, ⁴ Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, China

OPEN ACCESS

Edited by:

Jie Song,
Imperial College London,
United Kingdom

Reviewed by:

Kai Wang,
Nantong University, China
Daisuke Miki,
Shanghai Center for Plant Stress
Biology, Shanghai Institutes
for Biological Sciences, Chinese
Academy of Sciences (CAS), China

*Correspondence:

Wenli Zhang
wzhang25@njau.edu.cn
Fang Liu
liufcri@163.com
Renhai Peng
aydxprh@163.com

[†] These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Plant Development and EvoDevo,
a section of the journal
Frontiers in Plant Science

Received: 19 August 2021

Accepted: 17 November 2021

Published: 15 December 2021

Citation:

Zhang A, Wei Y, Shi Y, Deng X,
Gao J, Feng Y, Zheng D, Cheng X,
Li Z, Wang T, Wang K, Liu F, Peng R
and Zhang W (2021) Profiling
of H3K4me3 and H3K27me3
and Their Roles in Gene
Subfunctionalization in Allotetraploid
Cotton. *Front. Plant Sci.* 12:761059.
doi: 10.3389/fpls.2021.761059

Cotton is an excellent model for studying crop polyploidization and domestication. Chromatin profiling helps to reveal how histone modifications are involved in controlling differential gene expression between A and D subgenomes in allotetraploid cotton. However, the detailed profiling and functional characterization of broad H3K4me3 and H3K27me3 are still understudied in cotton. In this study, we conducted H3K4me3- and H3K27me3-related ChIP-seq followed by comprehensively characterizing their roles in regulating gene transcription in cotton. We found that H3K4me3 and H3K27me3 exhibited active and repressive roles in regulating the expression of genes between A and D subgenomes, respectively. More importantly, H3K4me3 exhibited enrichment level-, position-, and distance-related impacts on expression levels of related genes. Distinct GO term enrichment occurred between A/D-specific and homeologous genes with broad H3K4me3 enrichment in promoters and gene bodies, suggesting that broad H3K4me3-marked genes might have some unique biological functions between A and D subgenome. An anticorrelation between H3K27me3 enrichment and expression levels of homeologous genes was more pronounced in the A subgenome relative to the D subgenome, reflecting distinct enrichment of H3K27me3 in homeologous genes between A and D subgenome. In addition, H3K4me3 and H3K27me3 marks can indirectly influence gene expression through regulatory networks with TF mediation. Thus, our study provides detailed insights into functions of H3K4me3 and H3K27me3 in regulating differential gene expression and subfunctionalization of homeologous genes, therefore serving as a driving force for polyploidization and domestication in cotton.

Keywords: H3K4me3, H3K27me3, ChIP-seq, gene expression, subfunctionalization, regulatory network, cotton

INTRODUCTION

The core nucleosome is the basic unit of chromatin. It comprises a core histone octamer, consisting of two copies of each histone H3, H4, H2A, and H2B, wrapped by approximately 147 bp of double-helix DNA (Kornberg, 1974). The exposed N-terminal tails of core histones, especially for histone H3, are subjected to various covalent modifications like methylation and acetylation with distinct

biological consequences (Fuchs et al., 2006; Kouzarides, 2007; Bannister and Kouzarides, 2011). In plants, individual histone modifications, especially combinatorial actions of multiple modifications, play essential roles during entire processes of normal growth and development (Berr et al., 2011; Deal and Henikoff, 2011; Boycheva et al., 2014; Han et al., 2019) and in response to diverse environmental cues (Kim J. M. et al., 2015; Ramirez-Prado et al., 2018; Alonso et al., 2019; Ueda and Seki, 2020). In particular, aberrant histone modifications frequently cause severe developmental defects in plants (Tian and Chen, 2001; Sanders et al., 2017; Deevy and Bracken, 2019), indicating histone modifications are indispensable for normal plant growth and development.

Histone methylation exhibits lysine site- and methylation degree-dependent effects on chromatin structure and gene transcription. Selective lysine residues can be mono-, di-, and trimethylated with distinct biological outcomes (Sims et al., 2003; Tariq and Paszkowski, 2004; Zhang et al., 2009; Liu et al., 2010; Stillman, 2018). H3K4me3 is a major active histone mark in eukaryotes (Santos-Rosa et al., 2002; Zhang et al., 2015a). By contrast, H3K27me3 is a repressive mark with significant biological relevance in eukaryotes, including plants (Zheng and Chen, 2011; Wiles and Selker, 2017).

So far, H3K4me3 and H3K27me3 are well-studied marks in plants (Deal and Henikoff, 2011). They play crucial roles in regulating gene expression during normal growth and development and responses to environmental cues in plants (Probst and Mittelsten Scheid, 2015; Chen D. H. et al., 2018). In general, H3K4me3 primarily distributes around the TSS or gene bodies of expressed genes, which is directly related to gene activation (Wang et al., 2009; Zhang et al., 2009; He et al., 2010). In contrast, H3K27me3 is primarily enriched in the promoters and gene bodies of repressed genes responsible for silencing conditionally expressed genes (Zhang et al., 2007; Wang et al., 2009; Deal and Henikoff, 2010; He et al., 2010). Moreover, the H3K27me3 mark has been recently found to be more associated with homeologs less expressed in polyploidy wheat (Ramirez-Gonzalez et al., 2018), indicating involvement of this mark in the expression bias of homeologs in the polyploidy genome.

Cotton (*Gossypium* spp.) is one of the major suppliers of natural textile fibers and oilseeds around the world. Allotetraploid cotton is an excellent model system for studying crop polyploidization and domestication. The availability of high-quality genome sequences for cultivated allotetraploid species and their wild relatives (Wang et al., 2012, 2019; Li F. et al., 2015; Zhang et al., 2015b; Chen et al., 2020) has promoted evolutionary and functional genomics studies (Zaidi et al., 2018), therefore, deepening our understanding of cotton biology and benefiting yield and fiber improvement in cotton breeding. However, when compared with tremendous progress in other model plants, like *Arabidopsis*, rice, and maize, histone modification-related epigenomic studies are largely understudied in cotton (Wang et al., 2016, 2017; Zheng et al., 2016; You et al., 2017; Tao et al., 2020), especially H3K27me3, a repressive mark for regulating genes involved in development and stress responses.

Here, we conducted H3K4me3 and H3K27me3 ChIP-seq assays using leaf tissue from the allotetraploid cotton cultivar

Gossypium hirsutum (TM-1). We examined subgenomic distribution of both marks. In particular, we comprehensively investigated the roles of H3K4me3 and H3K27me3 marks in regulating the differential expression of genes and homeologous genes between the A and D subgenomes.

MATERIALS AND METHODS

Plant Material and Growth Conditions

The allotetraploid cotton cultivar *G. hirsutum* was used in this study. Cotton seeds were soaked in water for 24 h, then transferred to soil in pots and continued to grow in a greenhouse with 60% humidity at 28°C/25°C and 16/8-h light/dark cycle. Two and three true leaves collected from 20-day-old plants were ground into a fine powder using liquid nitrogen. The ground powder can be used immediately or kept at -80°C for later use.

RNA-Seq

Total RNA was extracted from ground powder using TRIzol (Thermo Fisher Scientific). Total RNA was treated with DNaseI to completely remove contamination of genomic DNA. mRNA was enriched from DNaseI-treated RNA for library preparation and sequenced on Illumina HiSeq4000.

ChIP-Seq Assay

Twenty-day-old cotton leaves were used for nuclei preparation. The nuclei were extracted using a nuclei isolation buffer (NIB, pH 5.0, 1.0 M glucose, 0.1 M citric acid, 80 mM KCl, 10 mM EDTA, 1% Triton X-100 prepared fresh just before use). The nuclei were purified using 1 × nuclei washing buffer (NWB, pH 5.0, 1 M glucose, 0.1 M Na-citrate, 1% Triton X-100 prepared fresh just before use). The purified nuclei were resuspended using 600 μl MNB buffer (50%, w/v sucrose, 50 mM Tris-HCl, pH 7.5, 4 mM MgCl₂, 1 mM CaCl₂) for MNase digestion at 37°C for 10 min. A digestion mix was pelleted at 13,000 rpm for 10 min at 4°C and the supernatant was transferred into a new 1.5 ml tube. The digested nuclei pellet was resuspended using lysis buffer (1 mM Tris-HCl, pH 7.5, 0.1 mM PMSF, 2%, v/v Complete Mini) and left on ice for 1 h. After centrifugation, the supernatant was transferred into the 1.5 ml tube containing digested chromatin. ChIP incubation buffer was added to the digested chromatin to make a total of 1.7 ml. The remaining steps were conducted following the published procedures (Zhang et al., 2012), namely, antibody incubation followed by adding protein A-sepharose beads, bead washing, elution, and purification of ChIPed DNA for library preparation. The prepared libraries were finally sequenced on the Illumina platform (Illumina HiSeq4000).

Processing of Sequencing Data

Raw reads of all sequencing data were trimmed using fastp (Chen S. et al., 2018) based on the quality value ($Q \geq 25$) and read length (≥ 20 bp). The trimmed reads from RNA-seq and ChIP-seq were mapped to the *G. hirsutum* reference genome (Zhang et al., 2015b) with Hisat2 (Kim D. et al., 2015) and Bowtie2 aligner (Langmead and Salzberg, 2012), respectively. The remaining RNA-seq data of different tissues were obtained from

the previously published data (Zhang et al., 2015b), and processed according to the procedures described above. For ChIP-seq data, any PCR duplicates were removed using Picard. Aligned reads with mapping quality (MapQ) less than 30 were removed using samtools (Li et al., 2009). The bam files were converted to bigwig file and normalized by RPKM (Reads Per Kilobase per Million mapped reads) using deeptools (Ramirez et al., 2016), then Integrative Genomics Viewer (IGV) (Thorvaldsdottir et al., 2013) was used to visualize read distribution across the genome.

Identification of Differentially Expressed Homeologous Genes

Fragments Per Kilobase of exon model per Million mapped fragments (FPKM) computed from StringTie (Pertea et al., 2015) was used to measure the expression level of each gene. Gene annotations were obtained from CottonGen (Yu et al., 2014). Only genes with FPKM ≥ 1 were considered as expressed ones. Homeologous gene pairs with FPKM ≥ 1 in at least one of the subgenomes were used for further analyses. The homeologous gene pairs (Supplementary Table 1) were identified using reciprocal BLAST hits between A and D homeologs as previously reported (Zhang et al., 2015b). Differentially expressed homeologous genes were analyzed using the limma R package. The *P* values were adjusted using the BH method at $\alpha = 0.05$ (Benjamini and Hochberg, 1995). Corrected *P* values of 0.05 and \log_2 (fold-change) values of 1 were set as the threshold for assessing differential expression levels.

ChIP-Seq Data Analyses

In this study, we only used rep I data of H3K4me3 and H3K27me3 for peak calling and subsequent analyses, and we used rep II data to validate the reproducibility of peaks for each mark. H3K27me3 peaks were called with “--broad” parameter on (-f BAM -g 2.3e+9 --nomodel -q 0.01 --broad --broad-cutoff 0.1) and H3K4me3 peaks were called with off (-f BAM -g 2.3e+9 --nomodel -q 0.01) using the MACS2 software (Zhang et al., 2008). The ChIPpeakAnno package (Zhu et al., 2010) was used for peak annotation. Bedtools (Quinlan and Hall, 2010) were used to correlate peaks with genome loci, genes (including 1,000 bp region upstream of the TSS and 1,000 bp region downstream of the TTS) overlapping H3K4me3 or H3K27me3 peaks were considered as peak-related genes. A custom script was used to calculate normalized read counts of related genes.

Regulatory Network Analyses

The biased genes between the A and D subgenomes were chosen to construct gene co-expression networks using the R package WGCNA (Langfelder and Horvath, 2008). The blockwiseModules function was used for network construction with the following parameters: power = 16, minModuleSize = 30, mergeCutHeight = 0.25, corType = “pearson.” The same expression matrix used for WGCNA analyses was used for regulatory network analyses with the R package GENIE3 (Huynh-Thu et al., 2010). All TF annotations were obtained from PlnTFDB version 3.0 (Jin et al., 2014) and CottonFGD¹.

¹<https://cottonfgd.org/>

Gene Ontology Enrichment Analyses

Gene Ontology enrichment analyses were conducted using agriGO v2.0 (Tian et al., 2017). GO terms with an FDR less than 0.05 were considered as being significantly enriched.

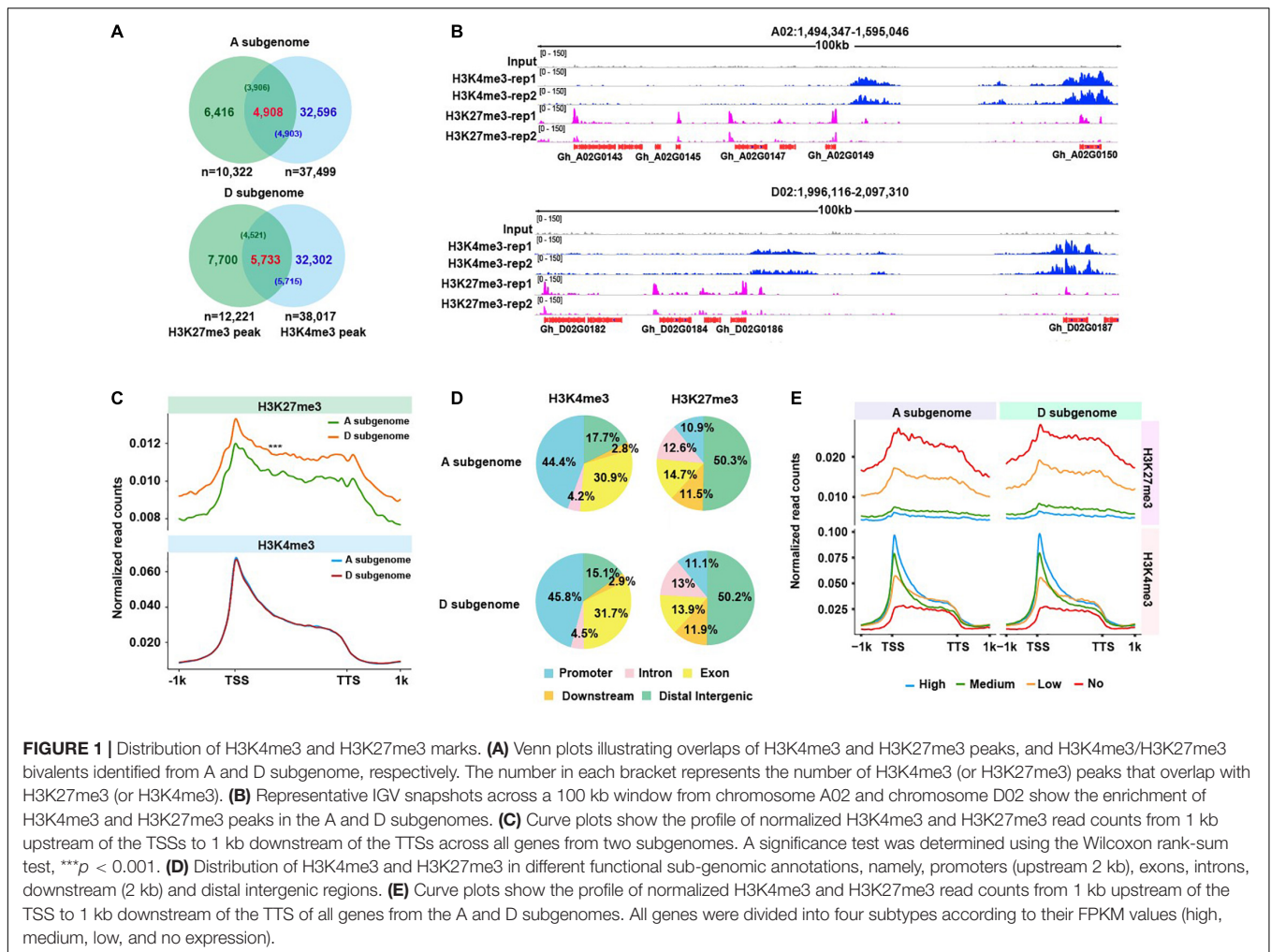
Bimolecular Fluorescent Complimentary (BiFC) Assay

To verify protein-protein interactions among PRE6, Rf2b, and bHLH63 as shown in Figure 4B, we performed transient transformation related Bimolecular Fluorescent Complimentary (BiFC) assays using *Nicotiana benthamiana*. To generate vectors of pXY104 with the expression cassette of 35S: PRE6-cYFP, Rf2b-cYFP and bHLH63-cYFP, we fused the coding sequences of PRE6, Rf2b, and bHLH63 to the C-terminal half of YFP. Meanwhile, we tagged each gene with the N-terminal half of YFP to generate vectors of pXY104 with the expression cassette of 35S:PRE6-nYFP, Rf2b-nYFP and bHLH63-nYFP (Supplementary Table 2). All plasmids were individually transferred into *Agrobacterium tumefaciens* (GV3101), and cultured in LB medium at 28°C until OD600 reached 1.2–1.5. After collecting the strains by centrifugation at 5,000 rpm at 4°C, the strains were individually suspended in injection buffer (10 mM MES-KOH, 10 mM MgCl₂, 40 μ M AS, pH = 5.7). After incubation at room temperature for 2 h, the bacteria containing cYFP/nYFP, Rf2b-cYFP/Rf2b-nYFP, Rf2b-cYFP/PRE6-nYFP, Rf2b-cYFP/bHLH63-nYFP, PRE6-cYFP/PRE6-nYFP, PRE6-cYFP/bHLH63-nYFP, and bHLH63-cYFP/bHLH63-nYFP were combined in 1:1 ratio, and were then individually co-injected into young and healthy *N. benthamiana* leaves for transient expression. The fluorescent signal was monitored and recorded using the Leica DMi8 confocal laser scanning microscope at 72 h post-infiltration, at least three images per injection. The functional validation for each combination was repeated by at least three separate injections.

RESULTS

Global Distribution of H3K4me3 and H3K27me3 in Cotton

To examine global distributions of H3K4me3 and H3K27me3 in the allotetraploid cotton, we conducted each mark-related ChIP-seq (Supplementary Table 3). We found that biological replicates of each mark were well correlated (Supplementary Figure 1). After data processing and peak calling, we identified total 75,516 H3K4me3 peaks (37,499 peaks in A subgenome and 38,017 peaks in D subgenome), 22,543 H3K27me3 peaks (10,322 peaks in A subgenome and 12,221 peaks in D subgenome) and 10,641 loci with at least 1bp overlap between H3K4me3 and H3K27me3 peaks (4,908 loci in A subgenome and 5,733 loci in D subgenome) (Figures 1A,B). The representative IGV snapshots (Figure 1B) show reproducible distribution of each mark in the A and D subgenomes. We then plotted normalized H3K4me3 and H3K27me3 ChIP-seq read counts across ± 1 kb of the TSS and the TTS of all genes in the A and D subgenomes. We observed a distinct genic distribution of each mark, H3K4me3 was highly



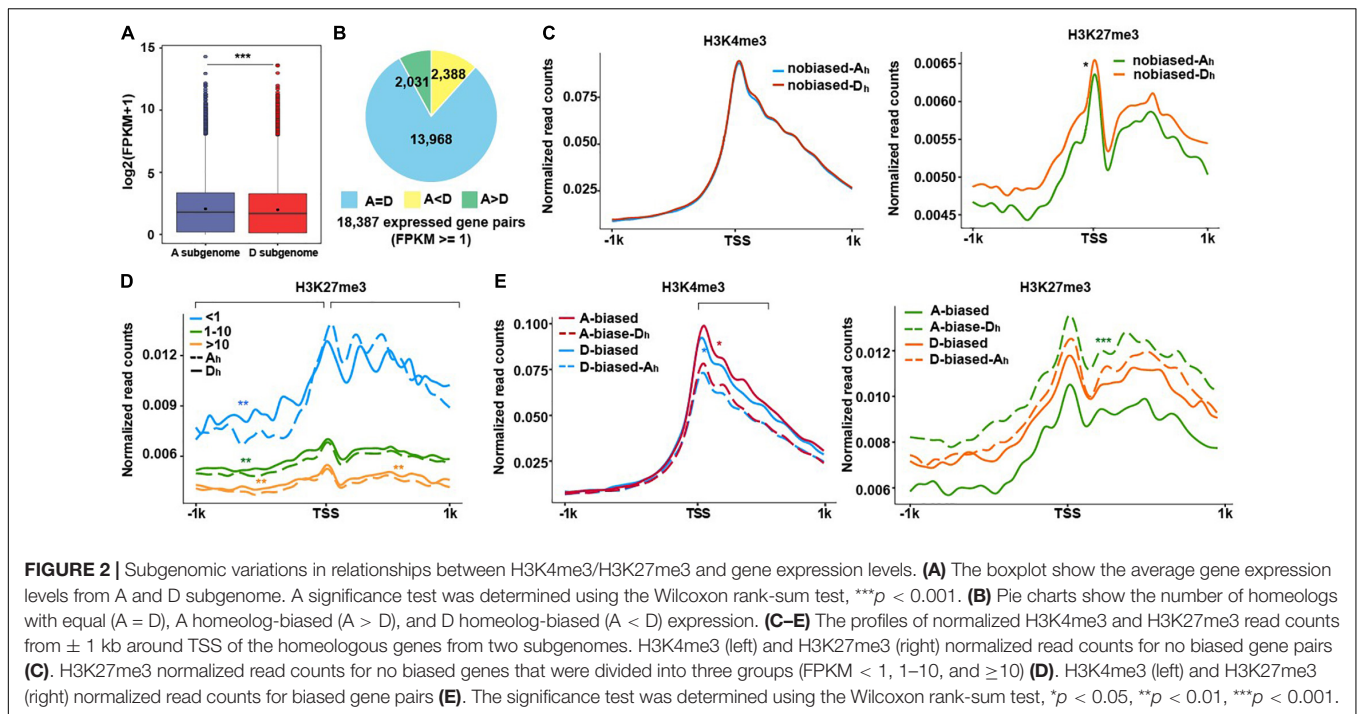
enriched immediately downstream of the TSS and extended to the whole gene body, by contrast, H3K27me3 primarily covered the whole gene body (Figure 1C). Strikingly, we found that distributions of normalized read counts for H3K4me3 were similar between A and D subgenome, whereas distributions of normalized read counts for H3K27me3 were higher in D subgenome relative to A subgenome.

To visualize the distribution of each mark in the A and D subgenomes, we partitioned each subgenome into five functionally annotated subregions, namely, promoters, exons, introns, downstream of TTS, and distal intergenic regions. We observed subtle differences in H3K4me3 but similar distributions for H3K27me3 between A and the D subgenomes (Figure 1D). Compared to the D subgenome, A subgenome had approximately 2% more H3K4me3 distributed in distal intergenic regions, and 1.4% less H3K4me3 distributed in promoters. Moreover, a distinct subgenomic distribution was observed for each mark, H3K4me3 exhibited the highest distribution in promoters, while H3K27me3 had the highest distribution in distal intergenic regions, suggesting a potential mark-dependent functional divergence. To assess an association between H3K4me3 or H3K27me3 enrichment levels and gene

expression levels in the A and D subgenome, we classified all genes in the A and D subgenome into four subtypes (high, medium, low, and no expression) according to FPKM values. We then plotted normalized H3K4me3 or H3K27me3 read counts around ± 1 kb of the TSS and the TTS of genes. We found that H3K4me3 enrichment levels, indicative of normalized read counts, exhibited a positive correlation with gene expression levels, whereas H3K27me3 enrichment levels were anti-correlated with gene expression levels in A and D subgenome (Figure 1E). Consistent with previous reports in other plant species, H3K4me3 and H3K27me3 also have contrasting roles in regulating gene expression in cotton, the former can facilitate gene expression whereas the latter usually suppresses gene expression.

Distinct Roles of H3K4me3 and H3K27me3 in Regulating Expression of Homeologous Genes

After comparing expression levels of genes between the A and D subgenomes, we observed that genes in A subgenome generally expressed more than those in D subgenome (Figure 2A),



indicating subgenome-biased expression in cotton. To examine how H3K4me3 and H3K27me3 are involved in regulating the differential expression of homeologous genes, we classified 18,387 expressed homeologous gene pairs into three subtypes: non-biased expression, A = D ($n = 13,968$), representing equal expression levels between A and D; A-biased expression, A > D ($n = 2,031$), representing genes expressed more in A homeologs and D-biased expression, D > A ($n = 2,388$), representing genes expressed more in D homeologs (**Figure 2B**; **Supplementary Figure 2**; **Supplementary Table 4**). After comparing normalized read counts, we found that there was no difference for H3K4me3 for no biased A_h and D_h (A_h and D_h represent A homeologs and D homeologs, respectively), whereas more H3K27me3 occurred in no biased- D_h than no biased- A_h , the counterparts in the A subgenome (**Figure 2C**).

To assess if read density changes of H3K27me3 in no biased homeologous genes are possibly related to gene expression levels, we classified no biased homeologous genes into three subtypes: FPKM < 1, 1–10, and ≥ 10 and conducted similar plotting assays as **Figure 2C**. When compared with genes in the D subgenome, we observed that non-expressed genes (FPKM < 1) in A exhibited less H3K27me3 reads distributed at the upstream of TSSs, but more H3K27me3 reads distributed at the downstream of TSSs. For expressed genes (FPKM ≥ 1), H3K27me3 reads tended to have more in the D subgenome than the A subgenome across all regions examined (**Figure 2D**).

We conducted similar plotting analyses for A- and D-biased expressed genes. As shown in **Figure 2E** (left), H3K4me3 exhibited higher enrichment in A-biased and D-biased genes compared to their respective counterpart of homeologous genes (A-biased- D_h and D-biased- A_h), indicating that H3K4me3

enrichment levels are directly correlated with gene expression levels. By contrast, H3K27me3 was less enriched in A-biased and D-biased genes compared to the corresponding A-biased- D_h and D-biased- A_h , respectively (**Figure 2E** right), exhibiting an anticorrelation between H3K27me3 enrichment levels and gene expression levels. After a careful examination, we found that the difference in H3K27me3 between A-biased and A-biased- D_h was more pronounced than that between D-biased and D-biased- A_h . These results showed that an anti-correlation between H3K27me3 enrichment levels and expression levels of homeologous genes was more pronounced in the A subgenome relative to the D subgenome, reflecting distinct enrichment of H3K27me3 in homeologous genes between A and D subgenome. Similarly, the roles of H3K27me3 in regulating biased expression of homeologs have been investigated in polyploidy wheat (Ramirez-Gonzalez et al., 2018), indicating potential common roles of H3K27me3 in regulating differential expression of homeologs in polyploidy plants.

Biological Implications of Genes With Broad H3K4me3

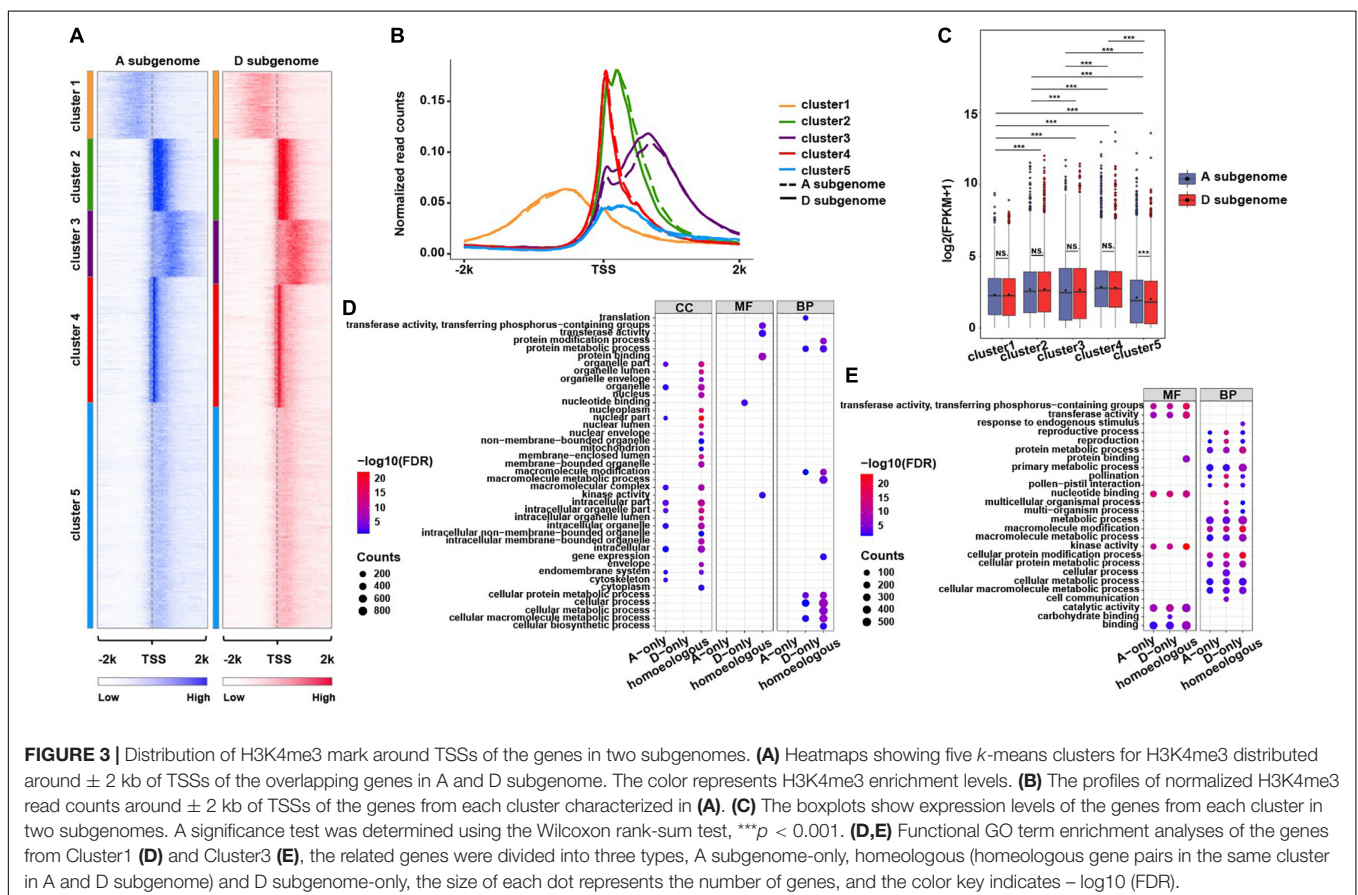
It has been reported that genes with broad H3K4me3 enrichment have some particular biological implications, such as determination of cell identity, regulation of expression of cell-type-specific tumor suppressors, regulation of expression of genes responsible for gamete development, and pre-implantation in mammals (Benayoun et al., 2014; Chen et al., 2015; Dahl et al., 2016; Liu et al., 2016; Lv and Chen, 2016; Zhang et al., 2016) and potential roles in photosynthesis in *Arabidopsis* (Brusslan et al., 2015). To interrogate if genes marked with broad H3K4me3 enrichment have any distinct biological

implications in cotton, we conducted *k*-means clustering assay using H3K4me3 associated genes in the A and D subgenomes. We obtained five clusters of genes with distinct H3K4me3 ChIP-seq read distribution (Figures 3A,B). There were two types of genes (Cluster 1 and 3) with broad H3K4me3 enrichment in promoter and gene body regions, respectively. After comparing gene expression levels and gene length in each cluster, genes in each cluster between the A and D subgenomes had overall similar mean expression levels (Figure 3C) and mean gene length (Supplementary Figure 3A). We found that genes in Cluster 4 had the highest mean expression levels and gene length, whereas genes in Cluster 5 had the highest gene length but the lowest mean expression levels in both subgenomes and genes in Cluster 3 had the shortest gene length (Figure 3C; Supplementary Figure 3A and Supplementary Table 5). These results suggest that the impacts of H3K4me3 on gene expression may depend on its enrichment levels, position, and distance from H3K4me3 to the TSS.

We then conducted GO term enrichment analyses using genes in Clusters 1 and 3 in A and D subgenome. For the genes in Cluster 1, we found that majority of GO terms were common between A and D subgenome despite several distinct GO terms occurred between A and D subgenome. For instance, the genes in D subgenome were more enriched in macromolecular complex/metabolic processes, cellular processes and transducer activities while the genes in A subgenome were

more enriched in protein binding (Supplementary Figure 4). We further divided the genes in Cluster 1 into A or D subgenome-specific and homeologous gene pairs (Supplementary Figure 3B) for re-conducting GO term enrichment assays. We observed distinct GO terms occurred between A- and D-specific genes (Figure 3D). A subgenome-only genes were more enriched in cellular component category but less enriched in molecular function and biological process categories compared to D subgenome-only genes. Homeologous gene pairs were more involved in the cellular component category but less in the molecular function category.

As illustrated in Figure 3A (heatmap) and Supplementary Figure 3A (boxplot), the genes in Cluster 3 contain broad H3K4me3 mark covering downstream 2 kb of the TSSs, but have the shortest gene length. To specifically look into biological relevance of genes with broad H3K4me3 enriched in gene body instead of extending to downstream of the genes with length less than 2 kb, we conducted similar GO term analyses using the genes in Cluster 3 with length greater than 2 kb in A and D subgenome (Supplementary Figures 5A,B). We observed subtle differences in GO terms between A and D genome. Compared to the genes in A subgenome, the genes in D subgenome were more enriched in signal transducer activity, carbohydrate binding, cell communication and pollination and pollen-pistil interaction (Supplementary Figure 5C). After dividing the genes specific for A or D subgenome and



homeologous gene pairs (**Supplementary Figure 5D**), we found the genes in D subgenome only were more enriched in functions associated with carbohydrate binding, metabolic processes, pollination, and pollen–pistil interaction as compared to the genes in A subgenome only (**Figure 3E**). To test if genes with broad H3K4me3 mark exhibit tissue-specific differential expression, we conducted *k*-means clustering analyses using RNA-seq data derived from 12 distinct tissues (the public data), we found that the genes with broad H3K4me3 enrichment exhibited tissue-specific expression profiles, including significantly highly expressed genes in stamen and petal (**Supplementary Figures 6A,B**), suggesting that broad H3K4me3-marked genes may function in reproductive stage for flower development. These analyses suggest that broad H3K4me3-marked genes might have some unique biological functions between A and D subgenome.

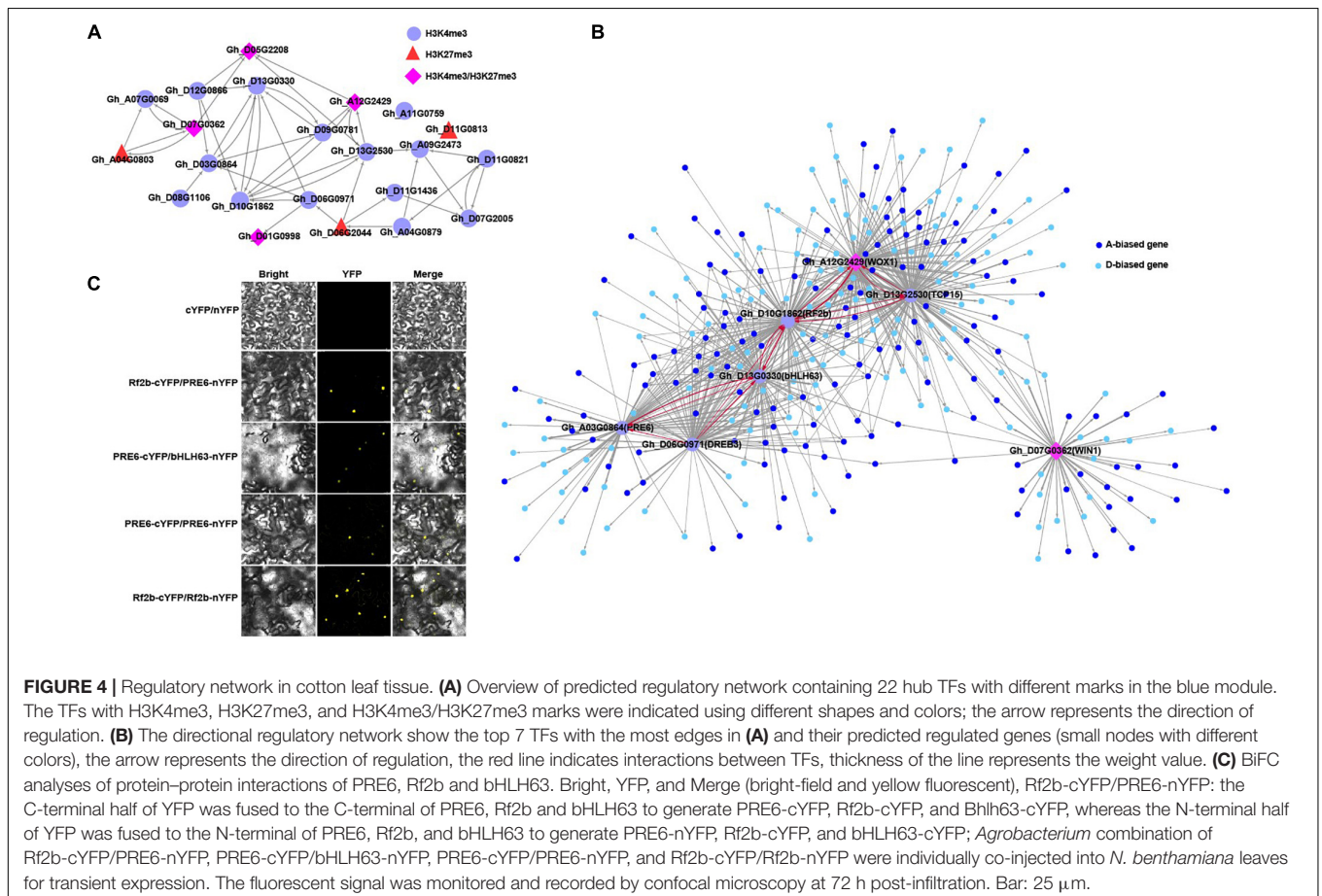
Involvement of H3K4me3 and H3K27me3 in Regulating Gene Expression Through TF-Mediated Regulatory Network

After specifically examining TFs with subgenome-related differential expression, we detected 180 and 204 TFs with biased expression in A and D subgenome, respectively (**Supplementary Figure 7** and **Supplementary Table 6**). A- and

D-biased TFs associated with H3K4me3-only, H3K27me3-only, and H3K4me3/H3K27me3 mark were summarized in **Supplementary Tables 6, 7**. To interrogate if H3K4me3 and H3K27me3 function in regulating gene expression through TF-mediated regulatory networks, we conducted a co-expression assay with the WGCNA R package, an expression matrix of biased genes across 12 tissues was used for the downstream analyses. We obtained 21 co-expression modules (**Supplementary Figure 8A**). Genes in the blue module exhibited a high association in the leaf tissue (**Supplementary Figure 8A**) and eigengenes in the blue module were specifically expressed in the leaf tissue (**Supplementary Figure 8B**).

Hub genes have been reported to be essential for maintaining the structure of the corresponding module and network (Li Y. et al., 2015; van Dam et al., 2018). We found that the blue module contained 22 hub TFs with a module membership value ($|kME|$) > 0.9, which is designed as the Pearson's correlation coefficient between the expression of a gene and a given module eigengene (Langfelder and Horvath, 2008). The 22 hub TFs contained 3 TFs enriched with H3K27me3, 4 TFs associated with H3K4me3/H3K27me3 marks, and the rest (15 TFs) enriched with H3K4me3 (**Supplementary Figure 9** and **Supplementary Table 8**).

To further infer interactions between hub TFs, we built a gene regulatory network with 22 hub TFs as candidate regulators



for the expression of other genes using the GENIE3 R package (Huynh-Thu et al., 2010; **Figure 4A**). To clearly show predicted regulatory relationships between TFs, the edges among other genes involved in the network were not displayed. Subsequently, the top 7 TFs with the most edges as displayed in **Figure 4A** and their regulated genes preferentially expressed in either the A or D subgenome were illustrated in **Figure 4B**. Strong interactions between TFs occurred among PRE6, BHLH63, RF2b, WOX1, and TCP15. PRE6, BHLH63, and RF2b were regulated by DREB3. To validate the accuracy of the network, we conducted a BiFC assay for Rf2b, bHLH63, and PEE6. Protein interactions were detected between Rf2b/bHLH63 and PEE6, and self-interaction was observed for PRE6 and Rf2b proteins (**Figure 4C**).

Functions of each TF ortholog have been documented in other plants such as *Arabidopsis* and rice. For example, it has been documented that TCP15 acts as a repressor of auxin biosynthesis and functions in the regulation of *Arabidopsis* gynoecium development (Lucero et al., 2015). DREB3 is a transcriptional activator functioning in abiotic stress responses in plants (Niu et al., 2020). bHLH63 also known as *CRYPTOCHROME-INTERACTING BHLH1* and bHLH100, is mainly involved in embryo suspensor and postembryonic development in *Arabidopsis* (Liu et al., 2008; Radoeva et al., 2019).

Collectively, our analyses indicate that, in addition, to directly affecting the expression of overlapping genes, H3K4me3, H3K27me3, and H3K4me3/H3K37me3 marks can indirectly influence gene expression through TF-mediated regulatory networks in the leaf tissue.

DISCUSSION

Similar to previous findings in plants (Zhang et al., 2007, 2009; Wang et al., 2009; He et al., 2010; Deal and Henikoff, 2011), our study indicated that H3K4me3 and H3K27me3 are highly enriched in gene bodies, H3K4me3 is an active mark that directly correlates with expression levels of genes, whereas H3K27me3 is a repressive mark that anti-correlates with expression levels of genes in allotetraploid cotton. It has been documented that H3K4me3 is involved in the biased expression of homeologous genes in allotetraploid cotton root (Zheng et al., 2016), and differential enrichment of H3K4me3 is responsible for transcriptional changes of genes associated with cotton development and evolution (You et al., 2017). However, our study for the first time characterized possible roles of H3K27me3, and broad H3K4me3 in differentially regulating gene expression between A and D subgenome in cotton.

Roles of H3K27me3 and Broad H3K4me3 Enrichment in Gene Transcription in Cotton

Compared with extensive H3K27me3 studies in plant development and stress responses in other plant species (Zheng and Chen, 2011; Gan et al., 2015; Chang et al., 2020), the roles of H3K27me3 in cotton are still much less studied. In addition to the overall repressive role of H3K27me3 in regulating gene expression in cotton, our study showed that distinct enrichment

of H3K27me3 in homeologous genes occurred between the A and D subgenomes (**Figures 2C–E**), since H3K27me3 enrichment in the A subgenome displayed a more pronounced anti-correlation with expression levels of homeologous genes as compared to the D subgenome.

Biological functions of broad H3K4me3 enrichment have been well studied in mammals, namely, cell identity (Benayoun et al., 2014), transcription of tumor suppressor genes (Chen et al., 2015; Dhar et al., 2018), transcription of genes in pre-implantation development, and embryonic stem cell differentiation (Liu et al., 2016), however, it is poorly understood in plants. In addition to the active roles of H3K4me3 in regulating gene expression, our study showed that the effects of H3K4me3 on expression levels of overlapping genes were related to their enrichment levels, relative position, and distance around the TSS in cotton (**Figures 3A–C**). Importantly, we found that A- and D-specific genes and homeologous genes with broad H3K4me3 in promoters and gene bodies were potentially involved in differential biological relevance and tissue-specific expression, suggesting that broad H3K4me3-marked genes might have some unique biological functions between A and D subgenome. For example, A-specific genes with broad H3K4me3 in promoters had more enriched GO terms relative to those D-specific genes, moreover, they had distinct GO terms associated with molecular functions and biological processes as compared to the corresponding genes with broad H3K4me3 enrichment in gene bodies (**Figures 3D,E**). It has been reported that genes with broad H3K4me3 enrichment have enriched GO terms associated with photosynthesis in *Arabidopsis* (Brusslan et al., 2015). Thus, our study provides further insights into the roles of broad H3K4me3 enrichment in subfunctionalization of homeologous genes in cotton.

Impacts of H3K4me3, H3K27me3 and H3K4me3/H3K27me3 Marks on Gene Transcription Through the Regulatory Network

In addition to direct impacts of H3K4me3 and H3K27me3 on transcription of overlapping genes in two subgenomes of cotton, our directional regulatory network related to TF indicated that both marks acted individually or in combined actions to indirectly regulate expression of co-expressed A- or D-biased genes through interacting between H3K4me3 or H3K27me3 mark overlapping TFs. For instance, predicted mutual interactions occurred between BHLH63 and PRE6 or RF2b, and between RF2b and TCP15, some of which were validated using a bimolecular fluorescence complementation assay. TCP 14 and TCP15 can regulate internode length and leaf shape in *Arabidopsis* through modulating cell proliferation (Kieffer et al., 2011). WOX1 is a key regulator during meristem development in *Arabidopsis* (Zhang et al., 2011). RF2b is a bZIP protein functioning in symptom development of rice tungro disease through interacting with RF2a (Dai et al., 2004). Functions of RF2b in leaf and root development in *Arabidopsis* are mediated by bZIP29 (Van Leene et al., 2016). PRE6 is a

paclobutrazol resistance protein belonging to non-DNA binding basic helix–loop–helix transcription factor, which has been reported to be involved in phytohormone signaling, such as GA, auxin, and BR and light responses in *Arabidopsis* (Gommers et al., 2017; Zheng et al., 2017). Gene regulatory network has already been applied to predict key nitrogen regulators, followed by successful experimental validation in rice (Ueda et al., 2020). We further extended the gene regulatory network to infer direct or indirect interactions between genes or TFs, which are responsible for differential expression of subgenome specific genes in allotetraploid cotton.

Collectively, our study provides evidence to indicate direct and indirect impacts of H3K4me3 and H3K27me3 on differential transcription of genes or homeologous gene pairs between A and D subgenome in cotton leaf tissue. In particular, the involvement of typically repressive mark H3K27me3 in expression bias of the homeologs in cotton is still understudied. Therefore, we provide further evidence showing the involvement of H3K27me3 individually or in combination with H3K4me3 in regulating differential expression of homeologous gene pairs between the A and D subgenomes. Thus, H3K27me3 individually or coordinated with H3K4me3 could play important roles in genomic evolution and/or domestication through controlling bias expression of the homeologs with some specific biological relevance in allotetraploid cotton. It has been reported that methylated genes evolve faster than unmethylated genes, and changes in DNA methylation and H3K4me3 enrichment are directly associated with expression bias of the homeologs in allotetraploid cotton and their relatives (Song et al., 2017), thereby both epigenetic marks possibly functioning in gene domestication, including flowering time-related gene *GhCOL2* in cotton. Profiling of H3K4me3 and H3K27me3 marks in ancestors and relatives of allotetraploid cotton could help to address epigenetic regulatory roles underlying polyploidization and domestication in cotton.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available from the corresponding author (WZ, wzhang25@njau.edu.cn), upon request. The raw sequencing data are deposited in the NCBI Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE165245.

AUTHOR CONTRIBUTIONS

WZ conceived and designed the study and wrote the manuscript with contributions from all the authors. YS, JG, YW, and XD

conducted the experiments. AZ analyzed the data. YF, TW, and XC helped with plant growth. RP and ZL helped with the generation of RNA-seq data. FL and KW supervised the experiments. YS, JG, YW, AZ, and WZ interpreted the results. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (32070561, U20A2030, and 31471548), Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX21_0602), the Fundamental Research Funds for the Central Universities (KYYJ 201808), the State Key Laboratory of Cotton Biology Open Fund (CB2019A06), the Postdoctoral Fund of Anyang Institute of Technology (BHJ2019009), and the Innovation Practice Base Postdoctoral Fund of Henan Province (2019, to YW).

ACKNOWLEDGMENTS

The authors thank the Bioinformatics Center in Nanjing Agricultural University for providing facilities to assist sequencing data analyses.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.761059/full#supplementary-material>

Supplementary Table 1 | The homeologous gene pairs between A and D subgenome (in a separate Excel file).

Supplementary Table 2 | Primer information was used in this study.

Supplementary Table 3 | Summary of sequencing data.

Supplementary Table 4 | The expression levels and fold changes of differentially expressed homeologous gene pairs (in a separate Excel file).

Supplementary Table 5 | The gene length and expression levels of 5 broad H3K4me3 marked genes from five clusters (in a separate Excel file).

Supplementary Table 6 | Details of A- or D-biased TFs and 182 biased TFs with H3K4me3, H3K27me3, or H3K4me3/H3K27me3 marks (in a separate Excel file).

Supplementary Table 7 | Summary of A- and D-biased TFs associated with H3K4me3, H3K27me3, and both marks.

Supplementary Table 8 | Genes and their kME values in the leaf-related blue module (in a separate Excel file).

REFERENCES

- Alonso, C., Ramos-Cruz, D., and Becker, C. (2019). The role of plant epigenetics in biotic interactions. *New Phytol.* 221, 731–737. doi: 10.1111/nph.15408
- Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395. doi: 10.1038/cr.2011.22

- Benayoun, B. A., Pollina, E. A., Ucar, D., Mahmoudi, S., Karra, K., Wong, E. D., et al. (2014). H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* 158, 673–688. doi: 10.1016/j.cell.2014.06.027
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc., Ser. B, Methodol.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

- Berr, A., Shafiq, S., and Shen, W. H. (2011). Histone modifications in transcriptional activation during plant development. *Biochim. Biophys. Acta* 1809, 567–576. doi: 10.1016/j.bbagr.2011.07.001
- Boycheva, I., Vassileva, V., and Iantcheva, A. (2014). Histone acetyltransferases in plant development and plasticity. *Curr. Genomics* 15, 28–37. doi: 10.2174/138920291501140306112742
- Brusslan, J. A., Bonora, G., Rus-Canterbury, A. M., Tariq, F., Jaroszewicz, A., and Pellegrini, M. (2015). A Genome-wide chronological study of gene expression and two histone modifications, H3K4me3 and H3K9ac, during developmental leaf senescence. *Plant Physiol.* 168, 1246–1261. doi: 10.1104/pp.114.252999
- Chang, Y. N., Zhu, C., Jiang, J., Zhang, H., Zhu, J. K., and Duan, C. G. (2020). Epigenetic regulation in plant abiotic stress responses. *J. Integr. Plant. Biol.* 62, 563–580. doi: 10.1111/jipb.12901
- Chen, K., Chen, Z., Wu, D., Zhang, L., Lin, X., Su, J., et al. (2015). Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. *Nat. Genet.* 47, 1149–1157. doi: 10.1038/ng.3385
- Chen, D. H., Huang, Y., Jiang, C., and Si, J. P. (2018). Chromatin-based regulation of plant root development. *Front. Plant Sci.* 9:1509. doi: 10.3389/fpls.2018.01509
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Chen, Z. J., Sreedasyam, A., Ando, A., Song, Q., De Santiago, L. M., Hulse-Kemp, A. M., et al. (2020). Genomic diversifications of five *Gossypium* allopolyploid species and their impact on cotton improvement. *Nat. Genet.* 52, 525–533. doi: 10.1038/s41588-020-0614-5
- Dahl, J. A., Jung, I., Aanes, H., Greggains, G. D., Manaf, A., Lerdrup, M., et al. (2016). Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* 537, 548–552. doi: 10.1038/nature19360
- Dai, S., Zhang, Z., Chen, S., and Beachy, R. N. (2004). RF2b, a rice bZIP transcription activator, interacts with RF2a and is involved in symptom development of rice tungro disease. *Proc. Natl. Acad. Sci. U S A.* 101, 687–692. doi: 10.1073/pnas.0307687100
- Deal, R. B., and Henikoff, S. (2010). A simple method for gene expression and chromatin profiling of individual cell types within a tissue. *Dev. Cell* 18, 1030–1040. doi: 10.1016/j.devcel.2010.05.013
- Deal, R. B., and Henikoff, S. (2011). Histone variants and modifications in plant gene regulation. *Curr. Opin. Plant Biol.* 14, 116–122. doi: 10.1016/j.pbi.2010.11.005
- Deevy, O., and Bracken, A. P. (2019). PRC2 functions in development and congenital disorders. *Development* 146:dev181354. doi: 10.1242/dev.181354
- Dhar, S. S., Zhao, D., Lin, T., Gu, B., Pal, K., Wu, S. J., et al. (2018). MLL4 is required to maintain broad H3K4me3 peaks and super-enhancers at tumor suppressor genes. *Mol. Cell* 70, 825–841.e6. doi: 10.1016/j.molcel.2018.04.028
- Fuchs, J., Demidov, D., Houben, A., and Schubert, I. (2006). Chromosomal histone modification patterns—from conservation to diversity. *Trends Plant Sci.* 11, 199–208. doi: 10.1016/j.tplants.2006.02.008
- Gan, E. S., Xu, Y., and Ito, T. (2015). Dynamics of H3K27me3 methylation and demethylation in plant development. *Plant Signal. Behav.* 10:e1027851. doi: 10.1080/15592324.2015.1027851
- Gommers, C. M., Keuskamp, D. H., Buti, S., van Veen, H., Koevoets, I. T., Reinen, E., et al. (2017). Molecular profiles of contrasting shade response strategies in wild plants: differential control of immunity and shoot elongation. *Plant Cell* 29, 331–344. doi: 10.1105/tpc.16.00790
- Han, Q., Bartels, A., Cheng, X., Meyer, A., An, Y. C., Hsieh, T. F., et al. (2019). Epigenetics regulates reproductive development in plants. *Plants (Basel)* 8:564. doi: 10.3390/plants8120564
- He, G., Zhu, X., Elling, A. A., Chen, L., Wang, X., Guo, L., et al. (2010). Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell* 22, 17–33. doi: 10.1105/tpc.109.072041
- Huynh-Thu, V. A., Irrthum, A., Wehenkel, L., and Geurts, P. (2010). Inferring regulatory networks from expression data using tree-based methods. *PLoS One* 5:e12776. doi: 10.1371/journal.pone.0012776
- Jin, J., Zhang, H., Kong, L., Gao, G., and Luo, J. (2014). PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Res.* 42, D1182–D1187. doi: 10.1093/nar/gkt1016
- Kieffer, M., Master, V., Waites, R., and Davies, B. (2011). TCP14 and TCP15 affect internode length and leaf shape in *Arabidopsis*. *Plant J.* 68, 147–158. doi: 10.1111/j.1365-313X.2011.04674.x
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317
- Kim, J. M., Sasaki, T., Ueda, M., Sako, K., and Seki, M. (2015). Chromatin changes in response to drought, salinity, heat, and cold stresses in plants. *Front. Plant Sci.* 6:114. doi: 10.3389/fpls.2015.00114
- Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* 184, 868–871. doi: 10.1126/science.184.4139.868
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705. doi: 10.1016/j.cell.2007.02.005
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, F., Fan, G., Lu, C., Xiao, G., Zou, C., Kohel, R. J., et al. (2015). Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nat. Biotechnol.* 33, 524–530. doi: 10.1038/nbt.3208
- Li, Y., Pearl, S. A., and Jackson, S. A. (2015). Gene networks in plant biology: approaches in reconstruction and analysis. *Trends Plant Sci.* 20, 664–675. doi: 10.1016/j.tplants.2015.06.013
- Liu, C., Lu, F., Cui, X., and Cao, X. (2010). Histone methylation in higher plants. *Annu. Rev. Plant Biol.* 61, 395–420. doi: 10.1146/annurev.arplant.043008.091939
- Liu, H., Yu, X., Li, K., Klejnot, J., Yang, H., Lisiero, D., et al. (2008). Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in *Arabidopsis*. *Science* 322, 1535–1539. doi: 10.1126/science.1163927
- Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., et al. (2016). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* 537, 558–562. doi: 10.1038/nature19362
- Lucero, L. E., Uberti-Manassero, N. G., Arce, A. L., Colombatti, F., Alemano, S. G., and Gonzalez, D. H. (2015). TCP15 modulates cytokinin and auxin responses during gynoecium development in *Arabidopsis*. *Plant J.* 84, 267–282. doi: 10.1111/tpj.12992
- Lv, J., and Chen, K. (2016). Broad H3K4me3 as a novel epigenetic signature for normal development and disease. *Genom. Proteom. Bioinf.* 14, 262–264. doi: 10.1016/j.gpb.2016.09.001
- Niu, X., Luo, T., Zhao, H., Su, Y., Ji, W., and Li, H. (2020). Identification of wheat DREB genes and functional characterization of TaDREB3 in response to abiotic stresses. *Gene* 740:144514. doi: 10.1016/j.gene.2020.144514
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122
- Probst, A. V., and Mittelsten Scheid, O. (2015). Stress-induced structural changes in plant chromatin. *Curr. Opin. Plant Biol.* 27, 8–16. doi: 10.1016/j.pbi.2015.05.011
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- Radoeva, T., Lokerse, A. S., Llavata-Peris, C. I., Wendrich, J. R., Xiang, D., Liao, C. Y., et al. (2019). A robust auxin response network controls embryo and suspensor development through a basic helix loop helix transcriptional module. *Plant Cell* 31, 52–67. doi: 10.1105/tpc.18.00518
- Ramirez, F., Ryan, D. P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., et al. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160–W165.
- Ramirez-Gonzalez, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., et al. (2018). The transcriptional landscape of polyploid wheat. *Science* 361:eaar6089. doi: 10.1126/science.aar6089
- Ramirez-Prado, J. S., Piquerez, S. J. M., Bendahmane, A., Hirt, H., Raynaud, C., and Benhamed, M. (2018). Modify the histone to win the battle: chromatin dynamics in plant-pathogen interactions. *Front. Plant Sci.* 9:355. doi: 10.3389/fpls.2018.00355

- Sanders, D., Qian, S., Fieweger, R., Lu, L., Dowell, J. A., Denu, J. M., et al. (2017). Histone lysine-to-methionine mutations reduce histone methylation and cause developmental pleiotropy. *Plant Physiol.* 173, 2243–2252. doi: 10.1104/pp.16.01499
- Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., et al. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* 419, 407–411. doi: 10.1038/nature01080
- Sims, R. J. III, Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. *Trends Genet.* 19, 629–639. doi: 10.1016/j.tig.2003.09.007
- Song, Q., Zhang, T. Z., David, M. S., and Chen, Z. J. (2017). Epigenomic and functional analyses reveal roles of epialleles in the loss of photoperiod sensitivity during domestication of allotetraploid cottons. *Genome Biol.* 18:99. doi: 10.1186/s13059-017-1229-8
- Stillman, B. (2018). Histone modifications: insights into their influence on gene expression. *Cell* 175, 6–9. doi: 10.1016/j.cell.2018.08.032
- Tao, X., Feng, S., Zhao, T., and Guan, X. (2020). Efficient chromatin profiling of H3K4me3 modification in cotton using CUT&Tag. *Plant Methods* 16:120. doi: 10.1186/s13007-020-00664-8
- Tariq, M., and Paszkowski, J. (2004). DNA and histone methylation in plants. *Trends Genet.* 20, 244–251. doi: 10.1016/j.tig.2004.04.005
- Thorvaldsdottir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* 14, 178–192. doi: 10.1093/bib/bbs017
- Tian, L., and Chen, Z. J. (2001). Blocking histone deacetylation in *Arabidopsis* induces pleiotropic effects on plant gene regulation and development. *Proc. Natl. Acad. Sci. U S A.* 98, 200–205. doi: 10.1073/pnas.011347998
- Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., et al. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* 45, W122–W129.
- Ueda, M., and Seki, M. (2020). Histone modifications form epigenetic regulatory networks to regulate abiotic stress response. *Plant Physiol.* 182, 15–26. doi: 10.1104/pp.19.00988
- Ueda, Y., Ohtsuki, N., Kadota, K., Tezuka, A., Nagano, A. J., Kadowaki, T., et al. (2020). Gene regulatory network and its constituent transcription factors that control nitrogen-deficiency responses in rice. *New Phytol.* 227, 1434–1452. doi: 10.1111/nph.16627
- van Dam, S., Vosa, U., van der Graaf, A., Franke, L., and de Magalhaes, J. P. (2018). Gene co-expression analysis for functional classification and gene-disease predictions. *Brief. Bioinform.* 19, 575–592. doi: 10.1093/bib/bbw139
- Van Leene, J., Blomme, J., Kulkarni, S. R., Cannoot, B., De Winne, N., Eeckhout, D., et al. (2016). Functional characterization of the *Arabidopsis* transcription factor bZIP29 reveals its role in leaf and root development. *J. Exp. Bot.* 67, 5825–5840.
- Wang, K., Wang, Z., Li, F., Ye, W., Wang, J., Song, G., et al. (2012). The draft genome of a diploid cotton *Gossypium raimondii*. *Nat. Genet.* 44, 1098–1103. doi: 10.1038/ng.2371
- Wang, M., Tu, L., Lin, M., Lin, Z., Wang, P., Yang, Q., et al. (2017). Asymmetric subgenome selection and cis-regulatory divergence during cotton domestication. *Nat. Genet.* 49, 579–587. doi: 10.1038/ng.3807
- Wang, M., Tu, L., Yuan, D., Zhu, S., Shen, C., Li, J., et al. (2019). Reference genome sequences of two cultivated allotetraploid cottons, *Gossypium hirsutum* and *Gossypium barbadense*. *Nat. Genet.* 51, 224–229. doi: 10.1038/s41588-018-0282-x
- Wang, M., Wang, P., Tu, L., Zhu, S., Zhang, L., Li, Z., et al. (2016). Multi-omics maps of cotton fibre reveal epigenetic basis for staged single-cell differentiation. *Nucleic Acids Res.* 44, 4067–4079. doi: 10.1093/nar/gkw238
- Wang, X., Elling, A. A., Li, X., Li, N., Peng, Z., He, G., et al. (2009). Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize. *Plant Cell* 21, 1053–1069. doi: 10.1105/tpc.109.065714
- Wiles, E. T., and Selker, E. U. (2017). H3K27 methylation: a promiscuous repressive chromatin mark. *Curr. Opin. Genet. Dev.* 43, 31–37. doi: 10.1016/j.gde.2016.11.001
- You, Q., Yi, X., Zhang, K., Wang, C., Ma, X., Zhang, X., et al. (2017). Genome-wide comparative analysis of H3K4me3 profiles between diploid and allotetraploid cotton to refine genome annotation. *Sci. Rep.* 7:9098. doi: 10.1038/s41598-017-09680-6
- Yu, J., Jung, S., Cheng, C. H., Ficklin, S. P., Lee, T., Zheng, P., et al. (2014). CottonGen: a genomics, genetics and breeding database for cotton research. *Nucleic Acids Res.* 42, D1229–D1236. doi: 10.1093/nar/gkt1064
- Zaidi, S. S., Mansoor, S., and Paterson, A. (2018). The rise of cotton genomics. *Trends Plant Sci.* 23, 953–955. doi: 10.1016/j.tplants.2018.08.009
- Zhang, B., Zheng, H., Huang, B., Li, W., Xiang, Y., Peng, X., et al. (2016). Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* 537, 553–557. doi: 10.1038/nature19361
- Zhang, T., Cooper, S., and Brockdorff, N. (2015a). The interplay of histone modifications - writers that read. *EMBO Rep.* 16, 1467–1481. doi: 10.15252/embr.201540945
- Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., et al. (2015b). Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol.* 33, 531–537. doi: 10.1038/nbt.3207
- Zhang, W., Wu, Y., Schnable, J. C., Zeng, Z., Freeling, M., Crawford, G. E., et al. (2012). High-resolution mapping of open chromatin in the rice genome. *Genome Res.* 22, 151–162. doi: 10.1101/gr.131342.111
- Zhang, X., Bernatavichute, Y. V., Cokus, S., Pellegrini, M., and Jacobsen, S. E. (2009). Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol.* 10:R62. doi: 10.1186/gb-2009-10-6-r62
- Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V., Pellegrini, M., Goodrich, J., et al. (2007). Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol.* 5:e129. doi: 10.1371/journal.pbio.0050129
- Zhang, Y., Liu, T., Meyer, C. A., Eeckhout, J., Johnson, D. S., Bernstein, B. E., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9:R137. doi: 10.1186/gb-2008-9-9-r137
- Zhang, Y., Wu, R., Qin, G., Chen, Z., Gu, H., and Qu, L. J. (2011). Over-expression of WOX1 leads to defects in meristem development and polyamine homeostasis in *Arabidopsis*. *J. Integr. Plant. Biol.* 53, 493–506. doi: 10.1111/j.1744-7909.2011.01054.x
- Zheng, B., and Chen, X. (2011). Dynamics of histone H3 lysine 27 trimethylation in plant development. *Curr. Opin. Plant. Biol.* 14, 123–129. doi: 10.1016/j.pbi.2011.01.001
- Zheng, D., Ye, W., Song, Q., Han, F., Zhang, T., and Chen, Z. J. (2016). Histone modifications define expression bias of homoeologous genomes in allotetraploid cotton. *Plant Physiol.* 172, 1760–1771. doi: 10.1104/pp.16.01210
- Zheng, K., Wang, Y., Zhang, N., Jia, Q., Wang, X., Hou, C., et al. (2017). Involvement of PACLOBUTRAZOL RESISTANCE6/KIDARI, an atypical bHLH transcription factor, in auxin responses in *Arabidopsis*. *Front. Plant Sci.* 8:1813. doi: 10.3389/fpls.2017.01813
- Zhu, L. J., Gazin, C., Lawson, N. D., Pages, H., Lin, S. M., Lapointe, D. S., et al. (2010). ChIPpeakAnno: a bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* 11:237. doi: 10.1186/1471-2105-11-237

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Zhang, Wei, Shi, Deng, Gao, Feng, Zheng, Cheng, Li, Wang, Wang, Liu, Peng and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.