Genome-wide identification of WRKY in cotton and the positive role of GhWRKY31 in response to salt and drought stress

Tianyu Dong
Henan Normal University
Haoyuan Li
Henan Normal University
Yajie Du
Henan Normal University
Ying Wang
Henan Normal University
Peilei Chen
Henan Normal University
Jiuchang Su
Henan Normal University
Xiaoyang Ge
Chinese Academy of Agricultural Sciences Cotton Research Institute
Hongying Duan
duanhy536@163.com

Henan Normal University https://orcid.org/0009-0009-1490-8603

Research Article

Keywords: WRKY, Evolutionary analysis, Gene silencing, Molecular docking, Yeast one-hybrid

Posted Date: December 7th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3675629/v1

Abstract

The WRKY gene family is widely distributed in plants and is known to play essential roles in stress response. However, the underlying structure and evolution of WRKY in cotton remains elusive. Herein, 112, 119, 217, and 222 WRKY genes were identified in G. arboreum, G. raimondii, G. hirsutum, and G. barbadense, respectively. The 670 WRKYs were divided into seven subgroups and unevenly mapped to chromosomes. Analysis of motifs, domains, cis-acting elements, and gene structure collectively revealed the evolutionary conservation and diversity of WRKYs in cotton. Synteny and collinearity confirmed the expansion, duplication, and deletion of WRKYs during the evolution of cotton. Further, GhWRKY31 was induced by osmotic and salt stress. Ectopic expression of GhWRKY31 improved osmotic and salt tolerance in Arabidopsis, while silencing GhWRKY31 in cotton increased sensitivity to drought and salinity. This was observed through higher germination rates and root length in GhWRKY31 transgenic Arabidopsis, as well as lower levels of ABA, proline, POD, and SOD in GhWRKY31-VIGS cotton plants under stress. Additionally, silencing of GhWRKY31 reduced the expression levels of drought- and salt-related genes, including GhRD29, GhNAC4, GhABF1, GhABF2, GhDREB2, GhP5CS, and GhSOS1. Yeast one-hybrid and molecular docking experiments confirmed that the GhWRKY31 domain binds to the W box of GhABF1, GhDREB2, and GhRD29, and is connected by hydrogen bonds. Collectively, the results provide a systematic and comprehensive understanding of the evolution of cotton WRKYs, and suggest an appropriate regulatory network for breeding cotton varieties with improved drought and salinity tolerance.

Key Message

The evolution and expression patterns of WRKY genes in cotton were elucidated. *GhWRKY31* positively regulates cotton resistance to drought and salt stress by controlling the expression of *GhABF1*, *GhDREB2*, and *GhRD29*.

Introduction

As global temperatures rise, the growth and development of higher plants are challenged by variable abiotic stresses, such as extreme temperatures, drought, and high salinity, of which not only inhibit plant growth but also lead to a gradual decline in crop production globally (Peck and Mittler 2020; Gao et al. 2016; Zandalinas et al. 2020). To cope with diverse abiotic stress and adapt to adverse growth environments, plants have evolved a series of intricate regulatory mechanisms (Devireddy et al. 2021; Yoon et al. 2020), in which the stress receptor genes, stress-related transcription factors (TFs) and upstream/downstream target genes together form a complex interconnected network (Agarwal et al. 2019; Chen et al. 2018; Liu and Zhang 2017; Li et al. 2020).

The TFs, which play vital roles in the regulation of gene expression by binding to *cis*-acting elements upstream of the transcription start site (TSS), are critical in the molecule network of stress response (Tariq et al. 2022; Rehman et al. 2021; Zhang et al. 2020). WRKY is one of the largest TF families in plants, and its members are involved in various responses to abiotic stress (Ye et al. 2021; Wei et al.

2022). WRKY family members have a highly conserved WRKY domain, which is composed of the WRKYGQK motif and a $CX_{4-5}CX_{22-23}HXH$ zinc-finger motif (Bakshi and Oelmüller 2014). The WRKYGQK motif binds to the W box (TTGACC/T) on the promoter of downstream gene and regulates its expression under various abiotic stresses (Rinerson et al. 2015; Chen et al. 2009). WRKY TFs were typically classified into seven groups based on the number of WRKY domains (two domains in Group I proteins and one in the others) and the primary amino acid sequence (C_2-H_2 structure in Group IIa-e proteins and C2HC structure in Group III proteins) (Bakshi and Oelmüller 2014; Rinerson et al. 2015).

During the development of higher plants, stress is inevitable and requires an effective defense mechanism to cope with it, wherein WRKY TFs related to plant resistance to abiotic stress have been widely reported. For instance, the *AtWRKY46* was rapidly induced upon water stress. The genes involving in ROS scavenging and cellular osmoprotection were regulated by *AtWRKY46*. Overexpression of *AtWRKY46* resulted in hypersensitivity in soil-grown *Arabidopsis* under osmotic stress (Ding et al. 2014). The silence of *AtWRKY63* inhibited the expression of *RD29A* and *COR47*, thereby reducing drought tolerance in *Arabidopsis* seedlings (Ren et al. 2010). In wheat, the heterogeneous expression of *TaWRKY146* in *Arabidopsis* was shown to enhance drought resistance, which was accomplished by promoting stomatal closure, increasing proline content, and reducing the accumulation of malondialdehyde (MDA) (Ma et al. 2017). The expression of *CsWRKY2* was increased by exogenous ABA and drought stress, and it enhanced drought tolerance in *Camellia sinensis* by regulating downstream genes of the ABA signaling pathway (Wang et al. 2016). The *AtWRKY25* and *AtWRKY33* double mutants *Arabidopsis* showed increased susceptibility to salt stress, while overexpression of *AtWRKY25* or *AtWRKY33* resulted in enhanced tolerance to salt stress (Li et al. 2011). In *Dendranthema grandiflorum*, the *DgWRKY2/3/4* significantly increased the germination rate and root length in both soybean and *Arabidopsis* seedlings under high salinity treatment (Liu et al. 2013; Wang et al. 2018; He et al. 2018). Collectively, the reports demonstrate that under drought and salt stress, WRKY family members have emerged as pivotal regulators in plants (Jiang et al. 2017).

Cotton, the most vital oilseed and fiber crop, accounts for 35% of the total fiber used worldwide. Previous studies have suggested that all diploid and tetraploid cotton species have evolved from a common ancestor, which subsequently diversified to produce 9 groups, including A-G, K, and AD genomes (Zhang et al. 2008; Wendel et al. 2010). Cultivated cotton mainly consists of four cotton subspecies, including *G. arboretum* (A_2-genome species), *G. raimondii* (D_5-genome species), *G. hirsutum* (AD_1-genome species), and *G. barbadense* (AD_2-genome species). Encouragingly, due to the release of high-quality whole-genome sequences of the 4 cultivated species (Du et al. 2010; Paterson et al. 2012; Yang et al. 2019; Wang et al. 2019), the genome-wide analysis of WRKY gene family is feasible and will help to elucidate its regulatory functions in stimuli responses such as osmotic, drought, and salt. However, the accurate and systematic analysis of WRKY genes in cultivated cotton remains largely unexplored, and the functional validation of *WRKYs* under osmotic, drought, and salt stress is still incompletely known (Abdelraheem et al. 2019).

In our study, the phylogenetic tree, chromosomal distribution, *cis*-acting elements, conserved motifs and domains, and collinearity relationship of *WRKYs* were performed in cotton. We next confirmed that *GhWRKY31* was up-regulated under salt and osmotic stress. Heterologous expression and VIGS assay revealed that *GhWRKY31* conferred salt and drought tolerance in *Arabidopsis* and cotton. The dehydration- and salt-induced genes expression, such as *GhABF1*, *GhABF2*, *GhDREB2*, *GhRD29*, *GhNAC4*, *GhP5CS*, and *GhSOS1*, was inhibited in *GhWRKY31* silencing seedlings. Finally, the YIH assay was employed to further verify the binding of *GhWRKY31* to the W box of *GhABF1*, *GhDREB2*, and *GhRD29*. Our results not only provide a new overview of cotton *WRKY* gene family, but also serve as a foundation for cotton breeding under abiotic stress conditions.

Materials and methods

Identification of WRKY family members

Four cotton genome assembly file (FASTA format) and genome annotation file (GFF3 format), including *G. arboreum* (CRI version, strain SXY1) (Du et al. 2010), *G. raimondii* (JGI version, strain Ulbr.) (Paterson et al. 2012), G. *hirsutum* (CRI version, strain Tm-1), (Yang et al. 2019)d *barbadense* (HAU version, strain 3–79) (Wang et al. 2019), were downloaded from Cotton FGD (https://cottonfgd.net/) (Zhu et al. 2017). Hidden Markov Model (HMM) file (PF03106) was downloaded from Pfam database (http://pfam.xfam.org/) (Finn et al. 2014). The HMMER 3.0 (Mistry et al. 2013) was used to screen the potential WRKY proteins, and the key parameters were set as default (1e-5). Next, WRKY proteins were manually screened using SMART (http://smart.emblheidelberg.de/) and NCBI CDD (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi/). Finally, non-WRKY domain, incorrect, and repetitive family members were deleted.

Multiple sequence alignment and phylogenetic tree construction

The full-length amino acid sequences of WRKY proteins were aligned by the ClustalW program. Based on the alignments provided, a maximum likelihood tree was constructed using the MEGA 7.0 program (http://www.megasoftware.net/) (Kumar et al. 2016), and the bootstrap test was carried out with 1000 iterations. Finally, the phylogenetic tree was plotted by interactive tree of life v5.0 (iTOL) (https://itol.embl.de/) (Letunic and Bork 2021).

Chromosomal locations, gene structure, conserved motifs and domains, and cis-acting elements of WRKY proteins

To map the chromosomal distribution of WRKY genes in 4 cotton species. The above reference genomes and annotation files, and WRKY protein IDs were incorporated into the gene location visualization toolkit of TBtools (Chen et al. 2020a).

For the analysis of WRKY gene structure, we extracted information on WRKY gene structures using reference genomes and annotation files. We then visualized the WRKY gene structures using gene structure toolkit of TBtools.

The conserved motifs of WRKY proteins were analyzed using the MEME database (http://meme-suite.org/), and the conserved domains were obtained through the NCBI CD-Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The acquired data was then plotted using TBtools.

To investigate the *cis*-acting elements of *WRKYs* promoter, the 5'-upstream regions of 2000 bp were downloaded from Cotton FGD. Subsequently, the sequences were analyzed using PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Lescot et al. 2002) and visualized with TBtools.

Duplication and collinearity analysis of WRKY proteins

The MCScan program (Tang et al. 2008) was used to detect gene pairs with a BLASTp search (e-value < 10^{-5}). Next, the chromosome length file, gene density file, and WRKY ID highlighting file were created from the reference genomes and annotation files, respectively. The prepared files were separately placed into the multiple collinear scanning toolkit, dual synteny plotter toolkit, and advance circos toolkit of TBtools (Chen et al. 2022) for analysis of gene collinearity relationships and duplication events among the *WRKYs* in four cotton species.

Cotton materials and stress tests

Upland cotton *G. hirsutum* (Tm-1) was obtained from Cotton Research of the Chinese Academy of Agricultural Sciences (Anyang, Henan Province, China). The cotton seeds were sterilized with 3% H_2O_2 for 12 h and then washed with distilled water. Subsequently, the seedlings were grown in a greenhouse at 28℃ with a 16 h light/8 h dark photoperiod until the second true leaf expanded.

For the validation of expression levels for *GhWRKYs* under drought and salt stress treatments, two true leaves of cotton seedlings were subjected to 20% (w/v) PEG6000 (drought-mimicking) and 200 mM NaCl solution, respectively. Seedlings with water were used as the control group. All leaves were collected at 0, 1, 3, 6, 12, and 24 h, and stored at -80℃ for further experiments.

Transcriptome analysis and qRT-PCR verification of WRKY genes

The RNA-seq raw data of PEG and NaCl-treated *G. hirsutum* were downloaded from NCBI (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA490626) (Hu et al. 2019). Firstly, we downloaded the raw sequencing data. Following the removal of adapters by Fastq and Trim Galore, the sequencing reads were aligned to the genome of *G. hirsutum* using STAR (Dobin et al. 2013). Next, RSEM (Li and

Dewey 2011) was employed to obtain the expression quantification (FPKM value) of *GhWRKYs*. The FPKM values of WRKY genes were log2-transformed and plotted using TBtools heatmap.

The total RNA of *G. hirsutum* was extracted using a FastPure Universal Plant Total RNA Isolation Kit (Vazyme, Nanjing, China). The first-strand cDNA was synthesized by TaKaRa kit (TaKaRa, Japan). The Universal SYBR qPCR Master Mix kit (Vazyme, Nanjing, China) was used for RT-qPCR, and the primers were designed using Primer Premier 5.0 (Supplementary Table 1). The *GhActin7* was used as an internal control. A total volume of 10 μL was carried out in the Light Cycler® 96 fluorescence quantitative PCR instrument (ABI7500; Applied Biosystems, America). The expression levels of WRKY genes were calculated by $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001).

Heterologous transformation and stress tolerance assay in Arabidopsis

The wild-type (WT) *Arabidopsis* (Ecotype Col-0) was used as the receptor for *GhWRKY31* genetic transformation. The seeds of WT *Arabidopsis* were surface-sterilized with 5% sodium hypochlorite and washed with sterile water 5 times. These seeds were then stored at 4℃ for 24 h. Next, the seeds were evenly sown on 1/2 MS solid mediums and cultured in a greenhouse (16 h light/8 h dark cycle, 22℃) for 7 days. Next, these seedlings were replanted in a 3:1 mixture of vermiculite and nutrient soil.

The *GhWRKY31* CDS was inserted into the XbaI/KpnI restriction enzyme sites of the Super1300 plasmid. The Super1300:*GhWRKY31* vector was transformed into GV3101 (*A. tumefaciens* strain), and full-flowering *Arabidopsis* seedlings were used for genetic transformation by floral dip method (Clough and Bent 1998). The *GhWRKY31* overexpression (OE) lines were selected using hygromycin. Here, the 4-week-old OE lines of *GhWRKY31* were identified by RT-qPCR. We finally obtained 5 independent OE lines of *GhWRKY31* and named them OE1, 2, 3, 4 and 5.

The seeds of WT and *GhWRKY31* OE lines of *Arabidopsis* were evenly planted on 1/2 MS solid media containing different concentrations of mannitol (0, 100, 200, and 300 mM) and NaCl (0, 50, 100, and 150 mM). The germination rate was recorded for 7 consecutive days using a magnifier. To measure root length, seedlings were initially grown upright on fresh 1/2 MS solid medium for 3 days, and then transferred to 1/2 MS solid medium supplemented with mannitol (0, 100, 200, and 300 mM) and NaCl (0, 50, and 100 mM) for a period of 5 to 7 days.

Virus-induced gene silencing (VIGS), stress treatments, and determination of biochemical indexes

As previous study described by Gu et al., the CDS of *GhWRKY31* was amplified from *G. hirsutum* using RT-PCR. The CDS of *GhWRKY31* was inserted into the pYL156 vector to construct pYL156: *GhWRKY31* fusion vector. Subsequently, pYL156: 00 (empty vector) and pYL156: *GhWRKY31* vector were severally transformed into *A. tumefaciens* strain GV3101. The bacterial fluid of recombinant GV3101 was used to infect the cotyledons of *G. hirsutum* seedlings through injection. The leaves were collected for RNA extraction and to detect interference efficiency using qRT-PCR.

For the drought and salt tolerance assay, seedlings of 'TM1' (no injection), 'TM1'+pYL156: 00 (empty vector injection), and 'TM1'+pYL156: *GhWRKY31* (*GhWRKY31* injection) were treated for 14 days with water, water-deficit, and 200 mM NaCl solution. Additionally, the content of ABA, MDA, and proline, as well as the activity of POD and SOD, were evaluated using the respective kits (ABA Elisa kit: SIONBESTBIO, Shanghai; MDA, proline, POD, and SOD kits: Solarbio, Beijing).

Molecular docking simulation

The interaction between the GhWRKY31 protein and the W box of stress-related genes was investigated using HDOCK v1.1 software. The nucleotide sequence of stress-related genes and the amino acid sequence of GhWRKY31 were introduced into receptors and ligands modules in HDOCK (http://hdock.phys.hust.edu.cn/) (Yan et al. 2020). The output interaction model files were imported into PyMOL 2.5.0. The center of the docking boxes, which were based on the position of the crystal ligand, were constructed minutely. Next, the atoms for polarity docking were selected, and the docking relationship was plotted. In addition, the confidence score (CS) indicates the likelihood of binding between two molecules ($CS = 1.0 / [1.0 + e^{0.02*(\text{docking score}+150)}]$). The two molecules would be very likely to bind if $CS > 0.7$.

Yeast one-hybrid (Y1H) assay

The Y1H assay was performed as described by Yu et al., the full-length sequence of *GhWRKY31* was cloned into the pGADT7 vector between the EcoRI and BamHI sites. The recombinant plasmid was co-transformed into yeast Y1HGold with pAbAi-*GhP5CS*, pAbAi-*GhABF1*, pAbAi-*GhABF2*, pAbAi-*GhDREB2*, and pAbAi-*GhRD29*. The pGADT-53 was used as a positive control, and all transformed candidates were grown on SD/-Ura/-Leu medium with 0 or 100 ng/mL of Aureobasidin A (AbA) for 3–5 days.

Statistical analysis

The data were statistically analyzed using SPSS 10.0 software and plotted by GraphPad Prism 5.0. There were 3 biological replicates for each experiment, and the data were presented as means ± SD of three independent experiments. Experimental data were analyzed by one-way ANOVA test, and the bars with different letters indicate significant differences ($P < 0.05$).

Results

Identification and phylogenetic analysis of WRKY gene family

Based on Pfam, Hmmer search, NCBI-CDD, and SMART databases, a total of 112, 119, 217, and 222 WRKY genes were identified in *G. arboreum*, *G. raimondii*, *G. hirsutum*, and *G. barbadense*, respectively. The shortest length of the WRKY protein (Gbar_D06G009260.1) was 144 aa, while the longest proteins (Gbar_D12G019910.1 and Gorai.008G200800.1) were 1340 aa. The isoelectric point (pI) ranged from 4.72 (Gbar_D11G016820.1) to 9.98 (Gorai.004G219300.1 and Gh_D08G210300.1). The maximum value

of molecular weight (MW) was 151,574.34 (Gorai.008G200800.1), and the minimum was 16,630.56 (Gorai.011G114200.1). The subcellular location of WRKY proteins was predicted using Plant-PLoc database, and it was found that WRKY proteins are located in the nucleus (Supplementary Table 2). The phylogenetic tree was constructed using maximum likelihood method to determine the evolutionary relationship of 670 WRKY proteins in cotton (Fig. 1). The WRKY proteins, including Group I (107), Group IIa-IIe (134, 74, 122, 86, and 75), and Group III (63), were classified into 7 clades. The WRKY proteins are unevenly distributed among 7 subgroups.

Chromosome location of WRKY genes

To further analyze the distribution of WRKY genes, the chromosome location was mapped (Fig. 2). In *G. arboretum* and *G. raimondii*, 109 and 119 *WRKYs* were located in At or Dt sub-genome, respectively. The *Ga14G1656*, *Ga14G1714*, and *Ga14G1560* were positioned in contigs (Fig. 2A, B). The number of WRKY genes in A07 chromosome was the highest (13), while the A03 chromosome had the lowest number (4) in *G. arboreum* (Fig. 2A). In *G. raimondii*, the D01 and D09 chromosomes contained the highest number of *GrWRKYs* (13), while the D02 and D05 chromosomes had the lowest number of *GrWRKYs* (4) (Fig. 2B). In *G. hirsutum*, the chrA05 had 16 WRKY members, which was the highest number among all 26 chromosomes, while the chrA03 contained 3 *GhWRKYs*. The *Gh_Contig00579G000600*, *Gh_Contig00383G000300*, and *Gh_Contig01109G001300* were located in contigs that were not observed in chromosomes (Fig. 2C). In *G. barbadense*, 4, 5, 3, 3, 17, 10, 12, 11, 6, 7, 14, 9, and 5 *WRKYs* were mapped on chrA01-chrA13, and 3, 4, 5, 11, 12, 10, 13, 11, 5, 7, 13, 12, and 5 WRKY genes were located on chrD01 to chrD13 (Fig. 2D). Therefore, the members of the WRKY family were unevenly distributed across the cotton chromosomes.

Conserved motifs and domains, cis-acting elements, and gene structure of WRKYs

To further clarify the detailed characteristics of *WRKYs*, gene structure, conserved motifs and domains, and *cis*-acting element were analyzed. 10 conserved motifs of WRKY proteins were identified in 4 cotton strains. We noted that most WRKY proteins contained more than two motifs, except for 5 GaWRKY, 15 GrWRKY, and 30 GbWRKY proteins. Moreover, motifs 1 and 2 were present in each WRKY member (Supplementary Table 3). Subsequently, we observed that WRKY proteins contain at least one WRKYGQK domain, and 107 WRKY proteins belonging to Group I have two WRKYGQK domains (Fig. 3). Additionally, the basic region-leucine zipper (bZIP) domains (PF00170) were found in 20 WRKY proteins, and the plant_zn_clust (PF10533) structures were located at the N-terminus of 80 WRKY members (Fig. 3). The results suggest that the WRKY proteins are evolutionarily stable and diverse in cotton.

To further investigate the biological function of *WRKYs*, we identified *cis*-acting elements in the 5'-upstream regions of 2000 bp (Supplementary Table 4). 12 different functions of *cis*-elements were identified, and these *cis*-acting elements related to stress response were found abundantly in the promoter of *WRKYs* (Fig. 3). The *cis*-acting elements can be divided into 3 categories: hormones responsive sites (auxin responsive element, gibberellin responsive element and MeJA responsive

element), transcription factor binding sites (MYB binding site, MYBHv1 binding site, and WRKY binding site), and growth and development sites (MYB binding site involved in drought-inducibility, light responsiveness, flavonoid biosynthetic, low-temperature responsive, and defense and stress responsive) (Fig. 3).

Duplication and collinearity of GaWRKYs, GrWRKYs, GhWRKYs, and GbWRKYs

To gain a better understanding of the expansion pattern of *WRKYs*, the duplication circos was plotted. In the diploid genomes of *G. raimondii* and *G. arboretum*, 88 and 102 *WRKYs*, respectively, occurred whole genome duplication (WGD) or segmental duplication events (Fig. 4A, B). The *Ga05G0631*, *Ga08G2219*, *Gorai.001G037800*, *Gorai.004G219400*, and *Gorai.009G062400* were tandem duplications and were distributed on chromosomes A05, A08, D01, D04, and D09, respectively (Fig. 4A, B). Furthermore, 17 and 14 WRKY genes were dispersed in At or Dt sub-genome (Supplementary Table 5). In *G. hirsutum*, 97.66% of *WRKYs* occurred through WGD or segmental duplication events. The 3 WRKY genes (*Gh_A08G214800*, *Gh_D05G062100*, *Gh_D08G210400*) were tandem duplications, and 2 *WRKYs* (*Gh_D04G011700* and *Gh_D07G055500*) were dispersed. These 5 *GhWRKYs* were located on chromosomes A08, D04, D05, D07, and D08, respectively (Fig. 4C) (Supplementary Table 5). A total of 213 *WRKYs* had undergone whole-genome duplication (WGD) or segmental duplication events in *G. barbadense*. The *Gbar_A08G020300* (ChrA08), *Gbar_D08G021260* (ChrD08), *Gbar_A03G013230* (ChrA03), and *Gbar_A11G020420* (ChrA11) appeared as tandem duplications or dispersion (Fig. 4D) (Supplementary Table 5).

It is widely known that *G. hirsutum* and *G. barbadense* evolved through hybridization between *G. arboreum* (A-genome species) and *G. raimondii* (D-genome species). Herein, a relative syntenic map was drawn to analyze the evolutionary relationships of *WRKYs* between *G. hirsutum* and 3 other species (Supplementary Table 6). According to MCScan analysis, 571, 621, and 1044 duplicated gene pairs were found between *G. hirsutum* and *G. arboreum*, *G. hirsutum* and *G. raimondii*, and *G. hirsutum* and *G. barbadense* (Fig. 5). In *G. arboreum* and *G. raimondii*, the highest number of collinear relationships was found on ChrA11 (85) and ChrD07 (98). The minimum collinear relationships, which were mapped to ChrA09 and ChrD05, were 17 and 21, respectively (Fig. 5A, B). Meanwhile, there are 87, 75, 73, 65, and 78 collinear relationships located on A05, A11, D05, D07, and D11 of *G. barbadense* among the 1044 gene pairs (Fig. 5C). In short, the aforementioned results confirmed that the collinear relationships were unevenly distributed across each chromosome, and that deletion and duplication events occurred among WRKY family members.

Expression profiling and qRT-PCR verification of WRKYs response to salt and osmotic stress in G. hirsutum

The transcriptome data of *G. hirsutum* was used to detect the expression profile of *GhWRKYs* under salt and osmotic stress. We found that 3 h of salt treatment, and 3 and 6 h of PEG treatment were grouped into one clade. The WRKY genes, including *Gh_A05G156700.1*, *Gh_D03G026500.1*, *Gh_A05G368400.1*, and *Gh_A06G109400.1*, exhibited higher expression levels in this clade. Furthermore, the expression levels of *Gh_A05G156700.1*, *Gh_D02G067800.1*, *Gh_A08G149000.1*, *Gh_D08G210300.1*, and

Gh_D08G191400.1 reached peak values after 1 h and 6 h salt treatment or 1 h PEG treatment. Additionally, the *Gh_A08G031700.1*, *Gh_A09G013200.1*, and *Gh_D03G050200.1* were expressed at the highest levels after 12 h salt treatment or 12 h PEG treatment. The results showed that partial *GhWRKYs* were regulated under salt and osmotic stress (Fig. 6).

To further verify the above results, qRT-PCR was employed to examine the changes in *GhWRKY*s expression levels. The *Gh_A08G031700.1*, *Gh_D02G067800.1*, and *Gh_A05G156700.1* were selected, and their expression levels were differentially induced by PEG and NaCl solution (Fig. 7). The expression values of *Gh_A08G031700.1* and *Gh_D02G067800.1* were variously under PEG and NaCl treatment. Concretely, the *Gh_A08G031700.1* expression was sensitive to both PEG and NaCl treatment, and its expression level was up-regulated at 3 and 12 h under PEG treatment, and at 1, 12, and 24 h under NaCl treatment ($P < 0.05$) (Fig. 7A). The *Gh_D02G067800.1* was regulated by salt stress, and its expression value significantly increased at 6 and 12 h under NaCl treatment ($P < 0.05$) (Fig. 7B). Additionally, the expression level of *Gh_A05G156700.1* was promoted at 1, 3, and 6 h ($P < 0.05$) after PEG treatment, and it was also upregulated after NaCl treatment at 3, 6, and 12 h (Fig. 7C). Therefore, the *Gh_A05G156700.1* was selected for the next functional verification under salt and osmotic stress.

GhWRKY31 improved the tolerance of transgenic Arabidopsis to osmotic and salt stress

The qRT-PCR result verified that *Gh_A05G156700.1* was positively induced by both osmotic and salt stress. To further investigate the function of *GhWRKY31* (*Gh_A05G156700.1*), the tolerance to osmotic and salt stress of *GhWRKY31* transgenic *Arabidopsis* was carried out after homozygote molecular identification (Supplementary Figure S1). The seed germination and root length both were inhibited by mannitol and NaCl treatment in WT *Arabidopsis*. The germination rates of WT were severely suppressed under 100 mM (83%), 200 mM (76%), and 300 mM (58%) mannitol treatment (Fig. 8A, B), and were suppressed to 56% and 36% under 100 mM and 150 mM NaCl treatment, respectively (Fig. 8E, F). Meanwhile, the root length of WT was also inhibited under 100 mM (2.84 cm), 200 mM (2.23 cm), and 300 mM (1.62 cm) mannitol (Fig. 8C, D), and was suppressed to 2.24, and 1.84 cm under 50 mM, and 100 mM salt conditions, respectively (Fig. 8G, H). As expected, the germination rates and root length of *GhWRKY31* OE lines were significantly higher than those of WT. The germination rates were nearly 100%, 100%, and 90% under 100 mM, 200 mM, and 300 mM mannitol. Under 50 mM, 100 mM, and 150 mM NaCl solution, the germination rates of *GhWRKY31* OE lines were almost up to 100% (Fig. 8B, F). In addition, the root length of OE lines was 3.36 cm, 3.14 cm, 2.33 cm, 3.24 cm, and 2.37 cm under 100 mM, 200 mM, 300 mM mannitol, 50 mM, and 100 mM NaCl treatments, respectively. These measurements were significantly longer than those of WT ($P < 0.05$) (Fig. 8D, H). Hence, the heterologous expression of *GhWRKY31* in *Arabidopsis* significantly improved osmotic and salt tolerance.

VIGS of GhWRKY31 reduced drought and salt tolerance in G. hirsutum

To further elucidate the function of *GhWRKY31* in *G. hirsutum* Tm-1, VIGS was employed to decrease the transcription level of *GhWRKY31*. The qRT-PCR was used to evaluate the silencing efficiency of *GhWRKY31*. The expression level of *GhWRKY31* was reduced by approximately 75% in pYL156:

GhWRKY31 plants (Supplementary Figure S2). As expected, no stress-related phenotype was observed in the seedlings of 'TM1', 'TM1 + pYL156: 00', and 'TM1 + pYL156: *GhWRKY31*' under water conditions. Nevertheless, the leaves of 'TM1 + pYL156: *GhWRKY31*' seedlings exhibited shrinkage and yellowing characteristics compared with 'TM1' (WT seedlings) and 'TM1 + pYL156: 00' (empty vector seedlings) under 200 mM NaCl treatment (Fig. 9A). Meanwhile, after a 14-day water-deficit treatment, the leaves of 'TM1' and 'TM1 + pYL156: 00' showed a healthier phenotype compared to 'TM1 + pYL156: GhWRKY31' seedlings. The latter exhibited symptoms such as shrinkage, rolling, wilting, and death (Fig. 9A). Additionally, the ABA and proline contents accumulated less in 'TM1 + pYL156: GhWRKY31' seedlings than in the control group seedlings under drought and salt stress. Meanwhile, the MDA content was higher in 'TM1 + pYL156: GhWRKY31' seedlings compared to the seedlings in the control group. Moreover, the activities of POD and SOD were higher in plants of the control group than in 'TM1 + pYL156: GhWRKY31' plants under drought and salt stress (Fig. 9B-F).

GhWRKY31 regulates the expression of salt- and dehydration-induced genes

The *GhWRKY31*-VIGS cotton seedlings showed greater sensitivity to drought and salt stress. To elucidate the target genes of *GhWRKY31* in the cotton drought and salt response, we conducted qRT-PCR analysis to determine whether *GhWRKY31* is essential for the expression of ABA-, dehydration-, and salt-induced genes. The expression levels of *GhRD29*, *GhNAC4*, *GhABF1*, *GhABF2*, *GhDREB2*, *GhP5CS*, and *GhSOS1* were induced in the control group under drought and NaCl stress. However, silencing *GhWRKY31* resulted in a decrease in the induction of the 7 genes under dehydration and salt stress. Specifically, the expression levels of *GhABF1*, *GhABF2*, *GhP5CS*, and *GhSOS1* were suppressed to levels lower than those observed in the control group (Fig. 10A).

The W box (TTGACC/T), which is the minimal sequence required for specific DNA binding of WRKY family members, was explored in the above 7 genes. We found that 1, 2, 2, 1, and 3 W boxes (TTGACC) were located in the promoter regions of *GhP5CS*, *GhABF1*, *GhRD29*, *GhABF2*, and *GhDREB2*, respectively (Fig. 10B). Additionally, to explore the potential interaction sites between the GhWRKY31 protein and the W box of these 5 genes, molecular docking was performed using HDOCK and PyMOL software. The confidence scores for the combination of GhWRKY31 and *GhP5CS*, *GhABF1*, *GhRD29*, *GhABF2*, and *GhDREB2* were 0.8611, 0.9525/0.9050, 0.7619/0.8815, 0.8930, and 0.8576/0.8654/0.9492, respectively. Stable complexes could be formed between the WRKY domain of GhWRKY31 and the adjacent W box sequence. These complexes were maintained by strong hydrogen bonds (Fig. 10C).

GhWRKY31 binds to the promoter regions of GhABF1, GhDREB2, and GhRD29

The Yeast one-hybrid (Y1H) assay was employed to further investigate the binding affinity of GhWRKY31 protein to *GhP5CS*, *GhABF1*, *GhABF2*, *GhDREB2*, and *GhRD29*. Firstly, we confirmed that 100 ng/ml of AbA could inhibit the self-activation of pAbAi-bait. The results showed that the transformation yeast containing the combination of *GhWRKY31* with the W box (TTGACC/T) of *GhABF1*, *GhDREB2*, *GhRD29, GhP5CS*, and *GhABF2* grew on SD/-Leu medium. The *GhWRKY31* protein specifically bound to the fragment that contained the core TTGACC/T motif of *GhABF1*, *GhDREB2*, and *GhRD29* on the SD/-Leu +

AbA (100 ng/ml) medium. These findings support the conclusion that *GhWRKY31* directly binds to the promoter regions of *GhABF1*, *GhDREB2*, and *GhRD29* (Fig. 11).

Discussion

WRKY TFs exist across the entire plant kingdom as a highly conserved protein family (Rushton et al. 2010). Currently, Genome-wide of WRKY gene family has been widely identified (Ye et al. 2021; Wei et al. 2022), and the family members have been shown to play the pivotal roles in responding to abiotic stress (Li et al. 2020, Jiang et al. 2017). Cotton, one of the most important economic crops, has remained relatively scarce research on WRKY gene family. Hence, the study investigates the evolution and function of WRKY genes in cotton development based on analysis of genome-wide duplication, qRT-PCR, heterogenous expression in *Arabidopsis*, VIGS in *G. hirsutum* 'TM1', molecular docking, and Y1H.

In our study, 112 *GaWRKYs*, 119 *GrWRKYs*, 217 *GhWRKYs*, and 222 *GbWRKYs* were identified. Since cotton underwent hybridization and polyploidization 1.5 Mya, the number of WRKY genes in tetraploid cotton has increased ~ 2-fold greater than that of diploid cotton (Chen et al. 2020b). Next, the 670 *WRKYs* were divided into 7 subgroups (Fig. 1) based purely on phylogenetic data (Rushton et al. 2008, Zhang and Wang 2005) and were unevenly distributed among different subfamilies. The chromosomal location indicated that WRKY genes on chrD01, chrD09, and chrD10 were lost, while those on chrD05, chrD06, and chrD11 were added during the formation of tetraploid cotton (Fig. 2). The results reveal the evolutionary information of cotton WRKY gene family.

To further elucidate the evolutionary relationship of WRKY TFs in cotton, the arrangement of *WRKYs* conserved motifs and domains were analyzed. Obviously, each WRKY gene had one or two conserved WRKYGQK domain(s) and an atypical zinc-finger structure at the C-terminus, which were composed of successive conserved motifs (Fig. 3). Almost every WRKY domain is located in the central position of the protein sequences, and the conserved motifs and domains are similar within the same subgroup for 4 cotton species. The above results suggest that WRKY genes were conserved highly during the evolution of cotton. Nevertheless, the similarity of sequences in other regions of *WRKYs* was lower compared to motif and domain regions, which indicated that the WRKY genes of cotton exhibit complexity and diversity during the evolutionary process. The *cis*-acting elements in the promoter region are known to play an important role in regulating gene expression and can provide insight into gene function. In the study, the *cis*-acting elements can be divided into three categories: phytohormone response elements, development and stress related elements, and TFs binding site elements (Fig. 3). These elements may have a significant impact on hormonal response, abiotic stress responses, and TF interactions. Analogously, these *cis*-acting elements, including the growth and development response element, hormone response element, and stress response element, have also been widely observed in *Vitis vinifera* (Huang et al. 2021a), *Calohypnum plumiforme* (Wang et al. 2022), and *Chrysanthemum lavandulifolium* (Muhammad et al. 2022). In addition, structural analysis indicated that the proportions of UTR and CDS were significantly different in each cotton species. This difference may be attributed to the homologous recombination caused by artificial domestication of cotton (Du et al. 2018; Huang et al. 2020).

In general, WGD, segmental duplication, tandem duplication, and transposon-induced duplication are major mechanisms that can alter the function, evolution, and structure of TFs and give rise to new subfamilies (De Smet et al. 2017; Lv et al. 2020; Zhou et al. 2020). In our research, the rates of WGD or segmental duplication were higher than those of tandem duplication, which clarified that the expansion and evolution of WRKY genes are primarily due to the WGD or segmental duplication, while tandem duplication plays a minor role (Fig. 4). Likewise, the evolutionary events were also found in mung bean (Tariq et al. 2022), wheat (Ye et al. 2021), and cherry (Ji et al. 2023), and the WGD and segmental duplication events played the primary-effect. Subsequently, to further explore the underlying evolutionary of *WRKYs*, the syntenic relationships were displayed (Fig. 5). Numerous collinear WRKY gene pairs were found between *G. hirsutum* and 3 other species, that may be attributed to the absence of significant changes in the number and arrangement of genes during the 1.5 million years of hybridization, polyploidy, and evolution in cotton. However, partial WRKY genes were lost during the process of evolution, possibly due to the artificial domestication process that lasted for 8000 years. These preserved WRKY genes may contribute significantly to improving the survival and adaptability of cotton, as well as the length and quality of cotton fibers (Du et al. 2018; Wang et al. 2019).

To date, WRKY TFs have been elucidated to be actively involved in the regulation of drought and salt stress in plants (Ren et al. 2010; Wang et al. 2016; Liu et al. 2013; He et al. 2018). Our research suggested as well that a group of WRKY genes were up- or down-regulated by PEG or NaCl treatment (Fig. 6). The *GhWRKY31* was identified as a candidate gene involved in the response to salt and drought stress in *G. hirsutum* (Fig. 7). Next, in order to further investigate the function of *GhWRKY31*, we constructed the Super1300: *WRKY31* vector and obtained homozygous *GhWRKY31* transgenic *Arabidopsis* lines. Phenotype analysis revealed that the germination rate and root length of WT were significantly lower than those of OE lines under stress conditions (Fig. 8). Similarly, overexpression of *GhWRKY39-1* in *Nicotiana benthamiana* not only enhanced tolerance to salt stress but also conferred greater resistance to bacterial pathogen infection (Shi et al. 2014). *GhWRKY25* transgenic tobacco enhanced the seedlings tolerance to salt stress, but the resistance to mannitol-induced osmotic and drought stress decreased (Liu et al. 2016). Additionally, the WRKY genes have been shown to respond actively to osmotic, drought, and salt stress in other plants. For instance, overexpressed *MbWRKY5* (Han et al. 2019), *MfWRKY40* (Huang et al. 2022), *CmWRKY10* (Jaffar et al. 2016), and *TaWRKY93* (Qin et al. 2015) in *Arabidopsis* or tobacco has been shown to confer increased resistance to osmotic and high salinity stress compared to the WT. On the other hand, overexpression of *ZmWRKY17* (Cai et al. 2017), *CdWRKY50* (Huang et al. 2021b), and *VvWRKY50* (Zhang et al. 2022) could result in a susceptible phenotype under PEG, mannitol, or NaCl treatment. The above results confirm that the heterologous expression of WRKY genes plays a crucial role in diverse abiotic stress responses in *Arabidopsis* and tobacco. Additionally, Gh*WRKY31* conferred dual resistance to both salt and osmotic stress in *Arabidopsis*.

To better evaluate the function of *GhWRKY31*, drought and salt tolerance tests were performed in *G. hirsutum* by VIGS technology. The *GhWRKY31*-VIGS cotton seedling leaves were more sensitive to water-deficit and NaCl conditions. The lower levels of ABA and proline content, along with the higher level of MDA content, indicated that VIGS seedlings had lower resistance to drought and salt stress. POD and

SOD activities results confirmed that the ROS-scavenging capability in WT plants was higher than that of *GhWRKY31*-VIGS seedlings (Fig. 9). Notably, the *GhABF1/2*, *GhDREB2*, and *GhRD29* were found to be involved in ABA-dependent or ABA-independent responses to drought and salt stress, and these genes play a positive regulatory role in both dehydration and NaCl conditions. The *GhP5CS*, which is induced by dehydration and high salt, catalyzes the rate-limiting enzyme in proline biosynthesis. Our study demonstrated that the expression levels of *GhABF1*, *GhABF2*, *GhDREB2*, *GhRD29*, and *GhP5CS* were suppressed in *GhWRKY31*-VIGS cotton leaves. Subsequently, the molecular docking analysis for the binding of the WRKYGQK domain and the W boxes of these genes indicated the presence of numerous hydrogen bonds, which contribute to the formation of stable complexes (Fig. 10). Analogously, the virus-induced gene silencing of *GhWRKY46* (Li et al. 2021) and *XsWRKY20* (Xiong et al. 2020) also resulted in drought or salt sensitivity. This was evident through a weak physiological phenotype, increased MDA content, reduced proline accumulation, and a significant inhibition of stress-related gene expression levels, such as *ABI3*, *ABF2*, *DREB1*, *DREB2*, *RD22*, *LEA5*, and *P5CS* in WRKY-silenced seedlings. Meanwhile, the highly conserved WRKYGQK domain of SlWRKY3/4, CcWRKY1/51/70, and HvWRKY46 could also form hydrogen bonds with the W box of stress-related, and these WRKY subfamily members appeared various bonding strength (Aamir et al. 2017; Pandey et al. 2018; Singh et al. 2019). In addition, the Y1H assay confirmed that *GhABF1*, *GhDREB2*, and *GhRD29* are direct targets of GhWRKY31 protein (Fig. 11). Therefore, *GhWRKY31* positively regulates the induction of downstream stress response genes (*GhABF1*, *GhRD29*, and *GhDREB2*), thereby participating in controlling the cotton drought and salt response network, and endows *Arabidopsis* and cotton with salt and drought resistance (Fig. 12). Collectively, the results not only facilitate future studies on WRKY genes involved in stress tolerance in cotton crops but also provide a foundation for further exploration of the underlying roles of TFs in plants.

Conclusions

In this study, a total of 670 *WRKYs* were identified in *G. arboreum*, *G. raimondii*, *G. hirsutum* and *G. barbadense*. The *WRKYs* were classified into 7 groups through phylogenetic analysis. Subsequently, the evolutionary relationship and diversification of WRKY genes were in-depth investigated. Finally, the positive function of *GhWRKY31* was demonstrated under osmotic, drought, and salt stress conditions in *Arabidopsis* and cotton Tm-1. *GhWRKY31* was found to directly bind to the W box (TTGACC/T) of *GhABF1*, *GhDREB2*, and *GhRD29*. These results not only suggest a precious resource for gaining a better understanding of *WRKYs* in cotton, but also provide a theoretical basis for the growth of plants in drought and high salinity soil.

Declarations

Authorship contribution statement

Tianyu Dong: Conceptualization, Methodology, Formal analysis, Software, Validation, Writing – original draft. **Haoyuan Li:** Resources, Investigation. **Yajie Du:** Software, Investigation. **Ying Wang:** Software, Formal analysis. **Peilei Chen:** Conceptualization, Supervision. **Jiuchang Su:** Visualization,

Supervision. **Xiaoyang Ge:** Methodology, Writing – review & editing, Supervision, Project administration. **Hongying Duan:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition, Project administration.

Declaration of competing 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.

Data availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Acknowledgements

The study was supported by “The High-Performance Computing Center of Henan Normal University”. The cotton seeds were kindly provided by Xi Wei in Institute of Cotton Research, Chinese Academy of Agricultural Sciences.

Funding

This research was supported by the financial aid of Program for Innovative Research Team (in Science and Technology) in University of Henan Province (No. 23IRTSTHN022); Science and Technology R&D Program of Henan Province (No. 222301420097).

References

1. Aamir M, Singh VK, Meena M, Upadhyay RS, Gupta VK, Singh S (2017) Structural and functional insights into *WRKY3* and *WRKY4* transcription factors to unravel the WRKY-DNA (W-Box) complex interaction in tomato (*Solanum lycopersicum* L.). A computational approach. Front Plant Sci 8:819. https://doi.org/10.3389/fpls.2017.00819
2. Abdelraheem A, Esmaeili N, O’Connell M, Zhang J (2019) Progress and perspective on drought and salt stress tolerance in cotton. Ind Crops Prod 130:118–129. https://doi.org/10.1016/j.indcrop.2018.12.070
3. Agarwal P, Baranwal VK, Khurana P (2019) Genome-wide analysis of bZIP transcription factors in wheat and functional characterization of a TabZIP under abiotic stress. Sci Rep 9:1–18. https://doi.org/10.1038/s41598-019-40659-7
4. Bakshi M, Oelmüller R (2015) WRKY transcription factors. Plant Signal Behav 9:e27700. https://doi.org/10.4161/psb.27700
5. Cai R, Dai W, Zhang C, Wang Y, Wu M, Zhao Y, Ma Q, Xiang Y, Cheng B (2017) The maize WRKY transcription factor. negatively regulates salt stress tolerance in transgenic Arabidopsis plants Planta

246:1215–1231. https://doi.org/10.1007/s00425-017-2766-9

6. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R (2020a) TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant 13:1194–1202. https://doi.org/10.1016/j.molp.2020.06.009
7. Chen C, Wu Y, Xia R (2022) A painless way to customize Circos plot: From data preparation to visualization using TBtools. iMeta 1:e35. https://doi.org/10.1002/imt2.35
8. Chen Y, Li L, Xu Q, Kong Y, Wang H, Wu W (2009) The *WRKY6* transcription factor modulates PHOSPHATE1 expression in response to low Pi stress in *Arabidopsis*. Plant Cell 21:3554–3566. https://doi.org/10.1105/tpc.108.064980
9. Chen Y, Cao Y, Wang L, Li L, Yang J, Zou M (2018) Identification of MYB transcription factor genes and their expression during abiotic stresses in maize. Biol Plant 62:222–230. https://doi.org/10.1007/s10535-017-0756-1
10. Chen Z, Sreedasyam A, Ando A, Song Q, De Santiago LM, Hulse-Kemp AM, Ding M, Ye W, Kirkbride RC, Jenkins J, Plott C, Lovell J, Lin YM, Vaughn R, Liu B, Simpson S, Scheffler BE, Wen L, Saski CA, Grover CE, Hu G, Conover JL, Carlson JW, Shu S, Boston LB, Williams M, Peterson DG, McGee K, Jones DC, Wendel JF, Stelly DM, Grimwood J, Schmutz J (2020b) Genomic diversifications of five Gossypium allopolyploid species and their impact on cotton improvement. Nat Genet 52:525–533. https://doi.org/10.1038/s41588-020-0614-5
11. Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J 16:735–743. https://doi.org/10.1046/j.1365-313x.1998.00343.x
12. De Smet R, Sabaghian E, Li Z, Saeys Y, Van de Peer Y (2017) Coordinated functional divergence of genes after genome duplication in *Arabidopsis thaliana*. Plant Cell 29:2786–2800. https://doi.org/10.1105/tpc.17.00531
13. Devireddy AR, Zandalinas SI, Fichman Y, Mittler R (2021) Integration of reactive oxygen species and hormone signaling during abiotic stress. Plant J 105:459–476. https://doi.org/10.1111/tpj.15010
14. Ding Z, Yan J, Xu X, Yu D, Li G, Zhang S, Zheng S (2014) Transcription factor *WRKY46* regulates osmotic stress responses and stomatal movement independently in *Arabidopsis*. Plant J 79:13–27. https://doi.org/10.1111/tpj.12538
15. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21. https://doi.org/10.1093/bioinformatics/bts635
16. Du X, Huang G, He S, Yang Z, Sun G, Ma X, Li N, Zhang X, Sun J, Liu M, Jia Y, Pan Z, Gong W, Liu Z, Zhu H, Ma L, Liu F, Yang D, Wang F, Fan W, Gong Q, Peng Z, Wang L, Wang X, Xu S, Shang H, Lu C, Zheng H, Huang S, Lin T, Zhu Y, Li F (2018) Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. Nat Genet 50:796–802. https://doi.org/10.1038/s41588-018-0116-x

17. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer EL, Tate J, Punta M (2014) PFAM: the protein families database. Nucleic Acids Res 42:222–230. https://doi.org/10.1093/nar/gkt1223
18. Gao Y, Lu Y, Wu M, Liang E, Li Y, Zhang D, Yin Z, Ren X, Dai Y, Deng D, Chen J (2016) Ability to remove Na and retain K correlates with salt tolerance in two maize inbred lines seedlings. Front Plant Sci 7:1716. https://doi.org/10.3389/fpls.2016.01716
19. Gu Z, Huang C, Li F, Zhou X (2014) A versatile system for functional analysis of genes and microRNAs in cotton. Plant Biotechnol J 12:638–649. https://doi.org/10.1111/pbi.12169
20. Han D, Hou Y, Wang Y, Ni B, Li Z, Yang G (2019) Overexpression of a *Malus baccata* WRKY transcription factor gene (*MbWRKY5*) increases drought and salt tolerance in transgenic tobacco. Can J Plant Sci 99:2. https://doi.org/10.1139/CJPS-2018-0053
21. He L, Wu Y, Zhao Q, Wang B, Liu Q, Zhang L (2018) Chrysanthemum *DgWRKY2* gene enhances tolerance to salt stress in transgenic Chrysanthemum. Intern J Mol Sci 19:2062. https://doi.org/10.3390/ijms19072062
22. Hu Y, Chen J, Fang L, Zhang Z, Ma W, Niu Y, Ju L, Deng J, Zhao T, Lian J, Baruch K, Fang D, Liu X, Ruan YL, Rahman MU, Han J, Wang K, Wang Q, Wu H, Mei G, Zang Y, Han Z, Xu C, Shen W, Yang D, Si Z, Dai F, Zou L, Huang F, Bai Y, Zhang Y, Brodt A, Ben-Hamo H, Zhu X, Zhou B, Guan X, Zhu S, Chen X, Zhang T (2019) *Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. Nat Genet 51:739–748. https://doi.org/10.1038/s41588-019-0371-5
23. Huang G, Wu Z, Percy RG, Bai M, Li Y, Frelichowski JE, Hu J, Wang K, Yu J, Zhu Y (2020) Genome sequence of *Gossypium herbaceum* and genome updates of *Gossypium arboreum* and *Gossypium hirsutum* provide insights into cotton A-genome evolution. Nat Genet 52:516–524. https://doi.org/10.1038/s41588-020-0607-4
24. Huang T, Yang J, Yu D, Han X, Wang X (2021a) Bioinformatics analysis of WRKY transcription factors in grape and their potential roles prediction in sugar and abscisic acid signaling pathway. J Plant Biochem Biot 30:67–80. https://doi.org/10.1007/s13562-020-00571-y
25. Huang X, Amee M, Chen L (2021b) Bermudagrass. response to salt stress Environ Exp Bot 188:104513. https://doi.org/10.1016/j.envexpbot.2021.104513
26. Huang Z, Wang J, Li Y, Song L, Chen D, Liu L, Jiang C (2022) A WRKY Protein, MfWRKY40, of resurrection plant *Myrothamnus flabellifolia* plays a positive role in regulating tolerance to drought and salinity stresses of *Arabidopsis*. Int J Mol Sci 23:8145. https://doi.org/10.3390/ijms23158145
27. Jaffar MA, Song A, Faheem M, Chen S, Jiang J, Liu C, Fan Q, Chen F (2016) Involvement of *CmWRKY10* in drought tolerance of *Chrysanthemum* through the ABA-signaling pathway. Int J Mol Sci 17:693. https://doi.org/10.3390/ijms17050693
28. Ji X, Zhang M, Wang D, Li Z, Lang S, Song X (2023) Genome-wide identification of WD40 superfamily in *Cerasus humilis* and functional characteristics of *ChTTG1*. Int J Biol Macromol 225:376–388. https://doi.org/10.1016/j.ijbiomac.2022.11.074

29. Jiang J, Ma S, Ye N, Jiang M, Cao J, Zhang J (2017) WRKY transcription factors in plant responses to stresses. J Integr Plant Biol 59:86–101. https://doi.org/10.1111/jipb.12513
30. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054
31. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res 30:325–327. https://doi.org/10.1093/nar/30.1.325
32. Letunic I, Bork P (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49:293–296. https://doi.org/10.1093/nar/gkab301
33. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323. https://doi.org/10.1186/1471-2105-12-323
34. Li S, Fu Q, Chen L, Huang W, Yu D (2011) *Arabidopsis thaliana WRKY25*, *WRKY26*, and *WRKY33* coordinate induction of plant thermotolerance. Planta 233:1237–1252. https://doi.org/10.1007/s00425-011-1375-2
35. Li W, Pang S, Lu Z, Jin B (2020) Function and mechanism of WRKY transcription factors in abiotic stress responses of plants. Plants 9:1515. https://doi.org/10.3390/plants9111515
36. Li Y, Chen H, Li S, Yang C, Ding Q, Song C, Wang D (2021) *GhWRKY46* from upland cotton positively regulates the drought and salt stress responses in plant. Environ Exp Bot 186:104438. https://doi.org/10.1016/j.envexpbot.2021.104438
37. Liu C, Zhang T (2017) Expansion and stress responses of the AP2/EREBP superfamily in cotton. BMC Genomics 18:1–16. https://doi.org/10.1186/s12864-017-3517-9
38. Liu Q, Zhong M, Li S, Pan Y, Jiang B, Jia Y, Zhang H (2013) Overexpression of a chrysanthemum transcription factor gene, *DgWRKY3*, in tobacco enhances tolerance to salt stress. Plant Physiol Biochem 69:27–33. https://doi.org/10.1016/j.plaphy.2013.04.016
39. Liu X, Song Y, Xing F, Wang N, Wen F, Zhu C (2016) *GhWRKY25*, a group I WRKY gene from cotton, confers differential tolerance to abiotic and biotic stresses in transgenic *Nicotiana benthamiana*. Protoplasma 253:1265–8121. https://doi.org/10.1007/s00709-015-0885-3
40. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(– Delta Delta C(T)) method. Methods 25:402–408. https://doi.org/10.1006/meth.2001.1262
41. Lv Q, Li W, Sun Z, Ouyang N, Jing X, He Q, Wu J, Zheng J, Zheng J, Tang S, Zhu R, Tian Y, Duan M, Tan Y, Yu D, Sheng X, Sun X, Jia G, Gao H, Zeng Q, Li Y, Tang L, Xu Q, Zhao B, Huang Z, Lu H, Li N, Zhao J, Zhu L, Li D, Yuan L, Yuan D (2020) Resequencing of 1,143 indica rice accessions reveals important genetic variations and different heterosis patterns. Nat Commun 11:4778. https://doi.org/10.1038/s41467-020-18608-0
42. Ma J, Gao X, Liu Q, Shao Y, Zhang D, Jiang L, Li C (2017) Overexpression of *TaWRKY146* increases drought tolerance through inducing stomatal closure in *Arabidopsis thaliana*. Front Plant Sci 8:2036.

https://doi.org/10.3389/fpls.2017.02036

43. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M (2013) Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res 41:e121. https://doi.org/10.1093/nar/gkt263
44. Muhammad AK, Kang D, Wu Y, Wang Y, Ai P, Wang Z (2022) Characterization of WRKY gene family in Whole-Genome and exploration of flowering improvement genes in *Chrysanthemum lavandulifolium*. Front Plant Sci 26:861193. https://doi.org/10.3389/fpls.2022.861193
45. Pandey B, Grover A, Sharma P (2018) Molecular dynamics simulations revealed structural differences among WRKY domain-DNA interaction in barley (*Hordeum vulgare*). BMC Genomics 19:132. https://doi.org/10.1186/s12864-018-4506-3
46. Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J, Yoo MJ, Byers R, Chen W, Doron-Faigenboim A, Duke MV, Gong L, Grimwood J, Grover C, Grupp K, Hu G, Lee TH, Li J, Lin L, Liu T, Marler BS, Page JT, Roberts AW, Romanel E, Sanders WS, Szadkowski E, Tan X, Tang H, Xu C, Wang J, Wang Z, Zhang D, Zhang L, Ashrafi H, Bedon F, Bowers JE, Brubaker CL, Chee PW, Das S, Gingle AR, Haigler CH, Harker D, Hoffmann LV, Hovav R, Jones DC, Lemke C, Mansoor S, ur Rahman M, Rainville LN, Rambani A, Reddy UK, Rong JK, Saranga Y, Scheffler BE, Scheffler JA, Stelly DM, Triplett BA, Van Deynze A, Vaslin MF, Waghmare VN, Walford SA, Wright RJ, Zaki EA, Zhang T, Dennis ES, Mayer KF, Peterson DG, Rokhsar DS, Wang X, Schmutz J (2012) Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature 492:423–427. https://doi.org/10.1038/nature11798
47. Peck S, Mittler R (2020) Plant signaling in biotic and abiotic stress. J Exp Bot 71:1649–1651. https://doi.org/10.1093/jxb/eraa051
48. Qin Y, Tian Y, Liu X (2015) A wheat salinity-induced WRKY transcription factor *TaWRKY93* confers multiple abiotic stress tolerance in *Arabidopsis thaliana*. Biochem Biophys Res Commun 464:428–433. https://doi.org/10.1016/j.bbrc.2015.06.128
49. Rehman A, Atif RM, Azhar MT, Peng Z, Li H, Qin G, Jia Y, Pan Z, He S, Qayyum A, Du X (2021) Genome wide identification, classification and functional characterization of heat shock transcription factors in cultivated and ancestral cottons (*Gossypium spp*). Int J Biol Macromol 182:1507–1527. https://doi.org/10.1016/j.ijbiomac.2021.05.016
50. Ren X, Chen Z, Liu Y, Zhang H, Zhang M, Liu Q, Hong X, Zhu JK, Gong Z (2010) ABO3, a WRKY transcription factor, mediates plant responses to abscisic acid and drought tolerance in *Arabidopsis*. Plant J 63:417–429. https://doi.org/10.1111/j.1365-313X.2010.04248.x
51. Rinerson CI, Rabara RC, Tripathi P, Shen QJ, Rushton PJ (2015) The evolution of WRKY transcription factors. BMC Plant Biol 15:66. https://doi.org/10.1186/s12870-015-0456-y
52. Rushton PJ, Bokowiec MT, Han S, Zhang H, Brannock JF, Chen X, Laudeman TW, Timko MP (2008) Tobacco transcription factors: novel insights into transcriptional regulation in the Solanaceae. Plant Physiol 147:280–295. https://doi.org/10.1104/pp.107.114041

53. Shi W, Hao L, Li J, Liu D, Guo X, Li H (2014) The *Gossypium hirsutum* WRKY gene *GhWRKY39-1* promotes pathogen infection defense responses and mediates salt stress tolerance in transgenic *Nicotiana benthamiana*. Plant Cell 33:483–498. https://doi.org/10.1007/s00299-013-1548-5
54. Singh A, Sharma AK, Singh NK, Sonah H, Deshmukh R, Sharma TR (2019) Understanding the effect of structural diversity in WRKY transcription factors on DNA binding efficiency through molecular dynamics simulation. Biology (Basel) 8:83. https://doi.org/10.3390/biology8040083
55. Song Y, Chen L, Zhang L, Yu D (2010) Overexpression of *OsWRKY72* gene interferes in the abscisic acid signal and auxin transport pathway of *Arabidopsis*. J Biosci 35:459–471. https://doi.org/10.1007/s12038-010-0051-1
56. Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH (2008) Synteny and collinearity in plant genomes. Science 320:486–488. https://doi.org/10.1126/science.1153917
57. Tariq R, Hussain A, Tariq A, Khalid MHB, Khan I, Basim H, Ingvarsson PK (2022) Genome-wide analyses of the mung bean NAC gene family reveals orthologs, co-expression networking and expression profiling under abiotic and biotic stresses. BMC Plant Biol 22:343. https://doi.org/10.1186/s12870-022-03716-4
58. Wang K, Bai ZY, Liang QY, Liu QL, Zhang L, Pan YZ, Liu GL, Jiang BB, Zhang F, Jia Y (2018) Transcriptome analysis of chrysanthemum (*Dendranthema grandiflorum*) in response to low temperature stress. BMC Genom 19:319. https://doi.org/10.1186/s12864-018-4706-x
59. Wang M, Tu L, Yuan D, Zhu D, Shen C, Li J, Liu F, Pei L, Wang P, Zhao G, Ye Z, Huang H, Yan F, Ma Y, Zhang L, Liu M, You J, Yang Y, Liu Z, Huang F, Li B, Qiu P, Zhang Q, Zhu L, Jin S, Yang X, Min L, Li G, Chen LL, Zheng H, Lindsey K, Lin Z, Udall JA, Zhang X (2019) Reference genome sequences of two cultivated allotetraploid cottons, *Gossypium hirsutum* and *Gossypium barbadense*. Nat Genet 51:224–229. https://doi.org/10.1038/s41588-018-0282-x
60. Wang Y, Shu Z, Wang W, Jiang X, Li D, Pan J, Li X (2016) *CsWRKY2*, a novel WRKY gene from *Camellia sinensis*, is involved in cold and drought stress responses. Biol Plant 60:443–451. https://doi.org/10.1007/s10535-016-0618-2
61. Wang Y, Zhu R, Shi M, Huang Q, Zhang S, Kai G, Guo S (2022) Genome-Wide identification and comparative analysis of WRKY transcription factors related to momilactone biosynthesis in *Calohypnum plumiforme*. Front Ecol Evol 9:809729. https://doi.org/10.3389/fevo.2021.809729
62. Wei Y, Jin J, Liang D, Gao J, Li J, Xie Q, Lu C, Yang F, Zhu G (2022) Genome-wide identification of *Cymbidium sinense* WRKY gene family and the importance of its Group III members in response to abiotic stress. Front Plant Sci 13:969010. https://doi.org/10.3389/fpls.2022.969010
63. Wendel JF, Brubaker CL, Seelanan T (2010) The origin and evolution of *Gossypium*. Springer, Dordrecht
64. Xiong C, Zhao S, Yu X, Sun Y, Li H, Ruan C, Li J (2020) Yellowhorn drought-induced transcription factor. acts as a positive regulator in drought stress through ROS homeostasis and ABA signaling pathway Plant Physiol Biochem 155:187–195. https://doi.org/10.1016/j.plaphy.2020.06.037

65. Yan Y, Tao H, He J, Huang SY (2020) The HDOCK server for integrated protein-protein docking. Nat Protoc 15:1829–1852. https://doi.org/10.1038/s41596-020-0312-x
66. Yang Z, Ge X, Yang Z, Qin W, Sun G, Wang Z, Li Z, Liu J, Wu J, Wang Y, Lu L, Wang P, Mo H, Zhang X, Li F (2019) Extensive intraspecific gene order and gene structural variations in upland cotton cultivars. Nat Commun 10:1–13. https://doi.org/10.1038/s41467-019-10820-x
67. Ye H, Qiao L, Guo H, Guo L, Ren F, Bai J, Wang Y (2021) Genome-Wide identification of wheat WRKY gene family reveals that *TaWRKY75-A* is referred to drought and salt resistances. Front Plant Sci 12:663118. https://doi.org/10.3389/fpls.2021.663118
68. Yoon Y, Seo DH, Shin H, Kim HJ, Kim CM, Jang G (2020) The role of stress-responsive transcription factors in modulating abiotic stress tolerance in plants. Agronomy 10:788. https://doi.org/10.3390/agronomy10060788
69. Yu M, Liu J, Du B, Zhang M, Wang A, Zhang L (2021) NAC transcription factor. activates ERD1 by interaction with ABF3 and DREB2a to enhance drought tolerance in transgenic Arabidopsis Int J Mol Sci 22:6952. https://doi.org/10.3390/ijms22136952
70. Zandalinas SI, Fichman Y, Devireddy AR, Sengupta S, Azad RK, Mittler R (2020) Systemic signaling during abiotic stress combination in plants. Proc Natl Acad Sci USA 17:13810–13820. https://doi.org/10.1073/pnas.2005077117
71. Zhang H, Li Y, Wang B, Chee P (2008) Recent advances in cotton genomics. Int J Plant Genomics 2008:742304. https://doi.org/10.1155/2008/742304
72. Zhang L, Zhang R, Ye X, Zheng X, Tan B, Wang W, Li Z, Li J, Cheng J, Feng J (2022) Overexpressing. from grapevine reduces the drought tolerance in Arabidopsis by increasing leaf stomatal density J Plant Physiol 275:153741. https://doi.org/10.1016/j.jplph.2022.153741
73. Zhang Q, Zhang J, Wei H, Fu X, Ma L, Lu J, Wang H, Yu S (2020) Genome-wide identification of NF-YA gene family in cotton and the positive role of *GhNF-YA10* and *GhNF-YA23* in salt tolerance. Int J Biol Macromol 165:2103–2115. https://doi.org/10.1016/j.ijbiomac.2020.10.064
74. Zhang Y, Wang L (2005) The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. BMC Evol Biol 5:1. https://doi.org/10.1186/1471-2148-5-1
75. Zhou Y, Zhao X, Li Y, Xu J, Bi A, Kang L, Xu D, Chen H, Wang Y, Wang Y, Liu S, Jiao C, Lu H, Wang J, Yin C, Jiao Y, Lu F (2020) Triticum population sequencing provides insights into wheat adaptation. Nat Genet 52:1412–1422. https://doi.org/10.1038/s41588-020-00722-w
76. Zhu T, Liang C, Meng Z, Sun G, Meng Z, Guo S, Zhang R (2017) CottonFGD: an integrated functional genomics database for cotton. BMC Plant Biol 17:101. https://doi.org/10.1186/s12870-017-1039-x

Figures

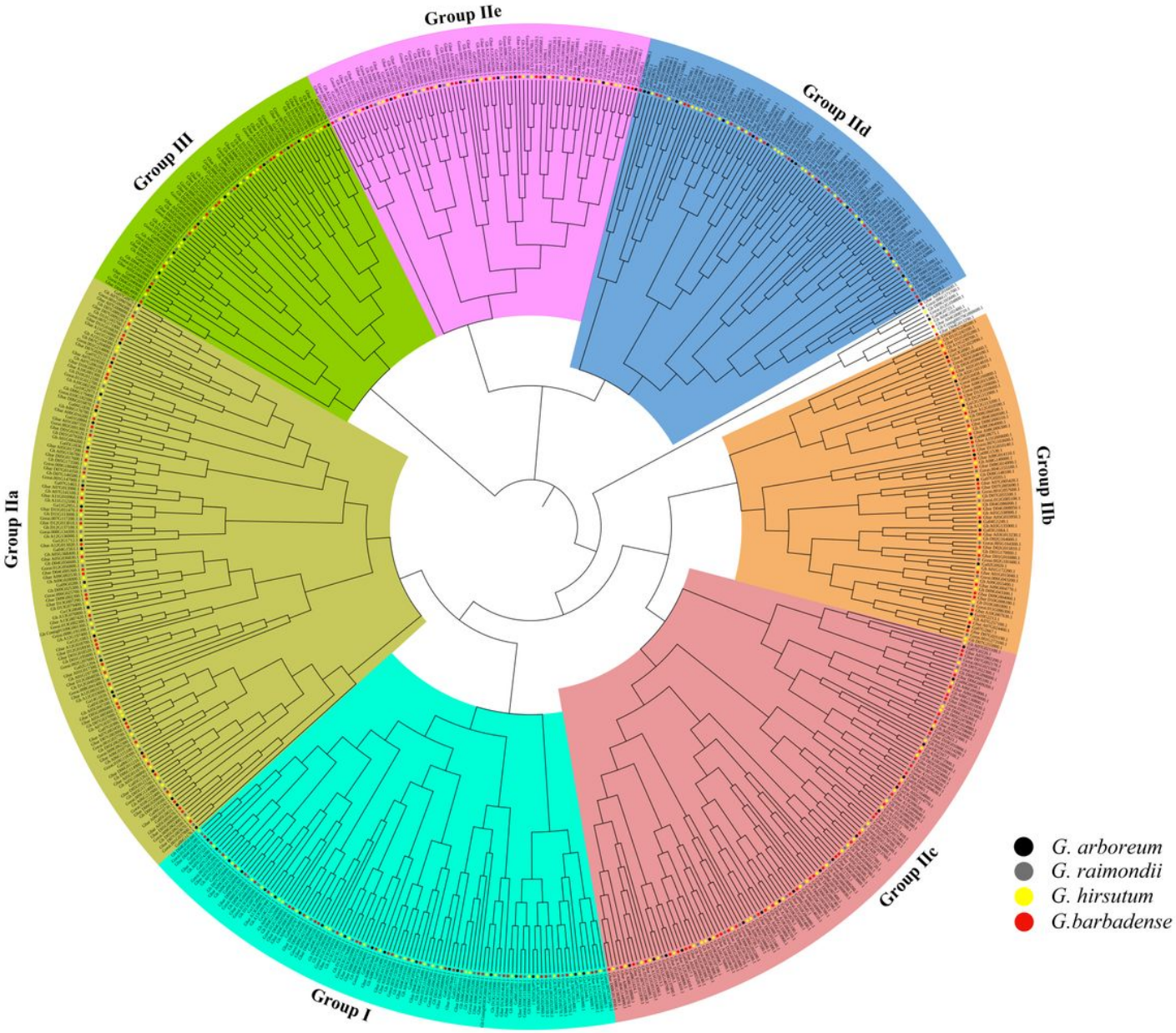

Figure 1

A maximum likelihood (1000 bootstraps) phylogenetic tree of WRKY proteins in *G. arboreum*, *G. raimondii*, *G. hirsutum*, and *G. barbadense*. The 7 color modules represent 7 subfamilies of WRKY proteins, and no background module indicates unclassified WRKY proteins.

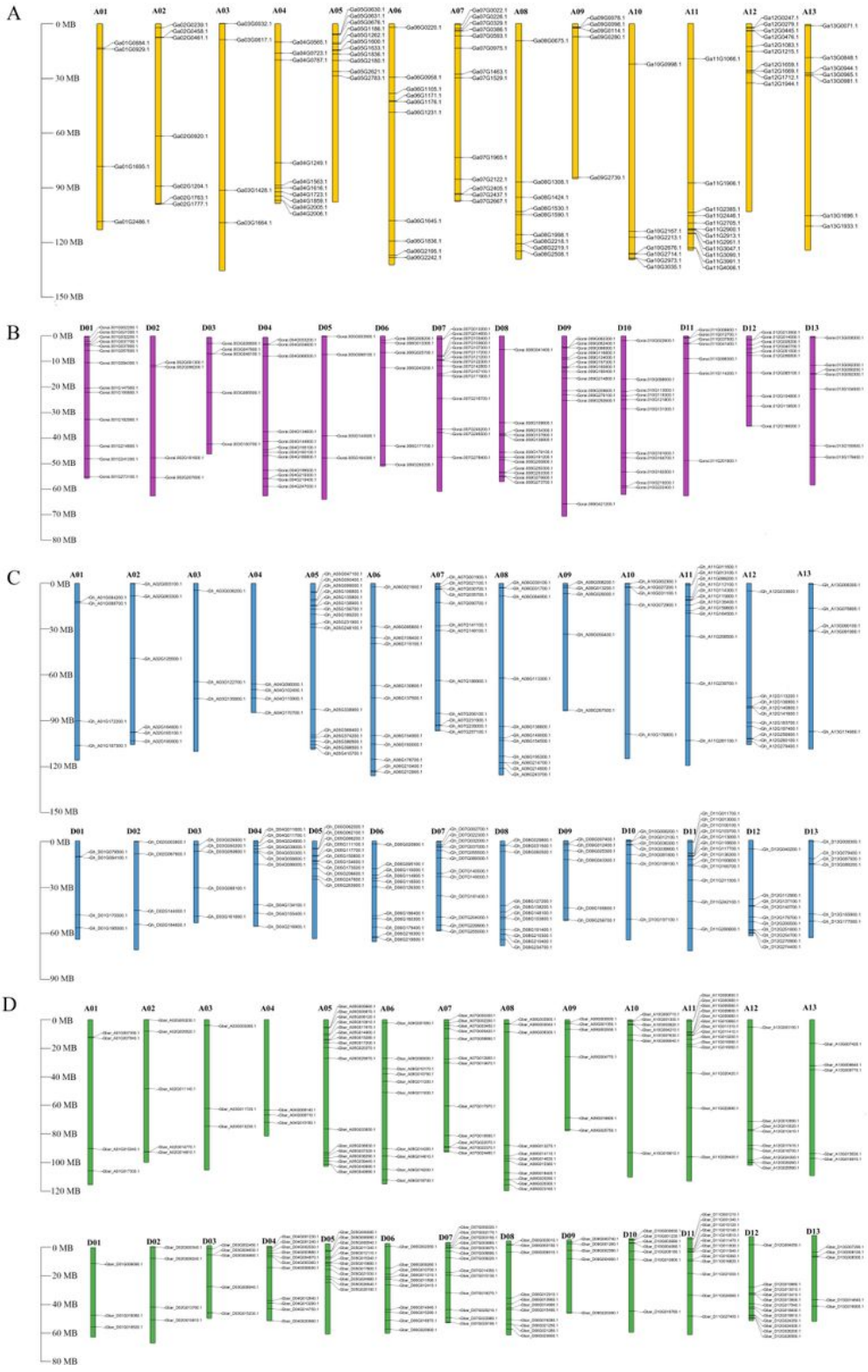

Figure 2

Chromosomal distribution of WRKY genes in *G. arboretum* (**A**), *G. raimondii*(**B**), *G. hirsutum* (**C**), and *G. barbadense* (**D**). The chromosome number is shown at the top of each chromosome, and the scale for the length of chromosome is mega bases (Mb).

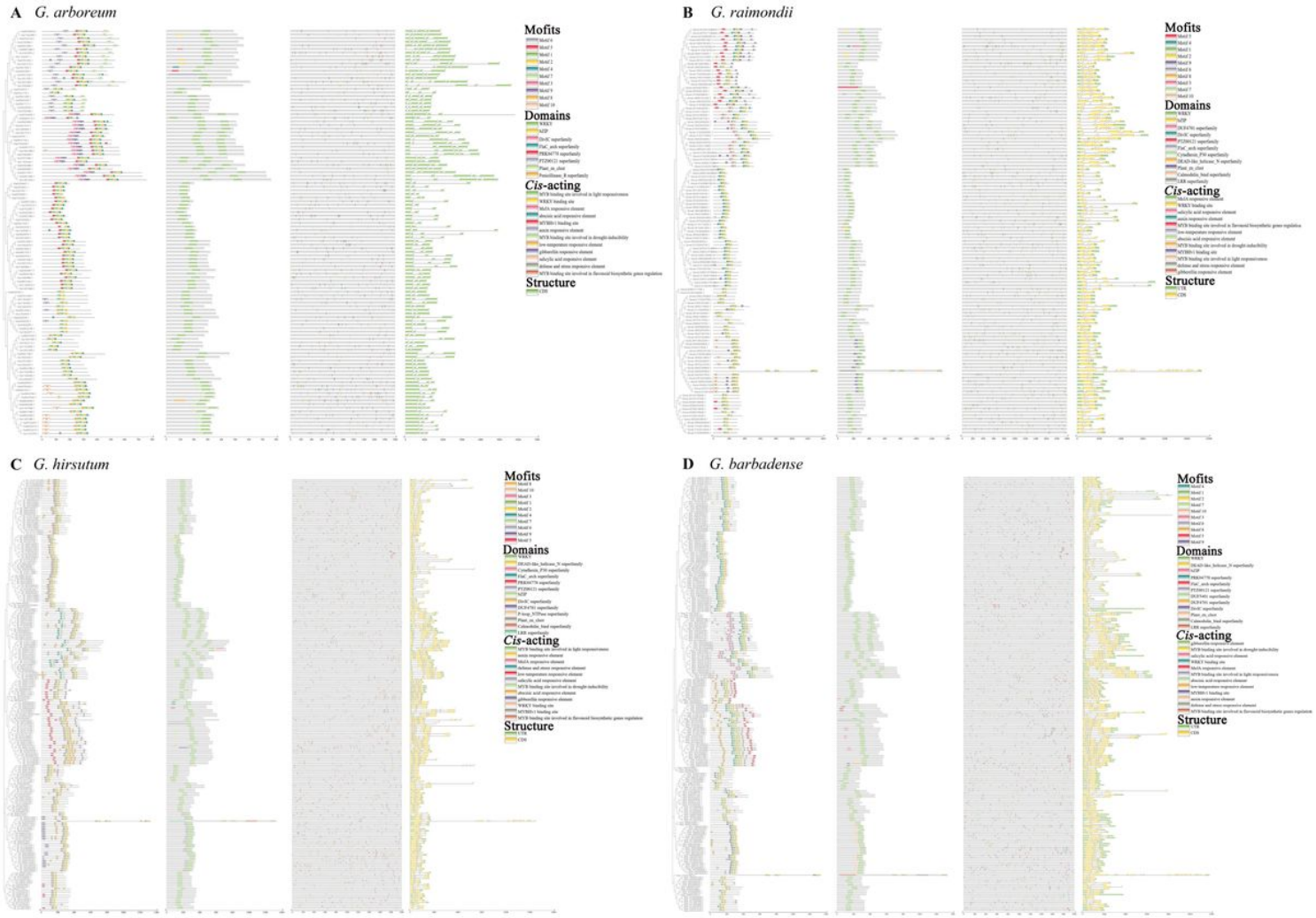

Figure 3

Analysis of conserved motifs and domains, *cis*-acting elements, and structures of WRKY members in *G. arboreum* (**A**), *G. raimondii* (**B**), *G. hirsutum* (**C**), and *G. barbadense* (**D**). The identification elements were represented by distinct colored boxes. The black lines of gene structure indicated non-conserved regions.

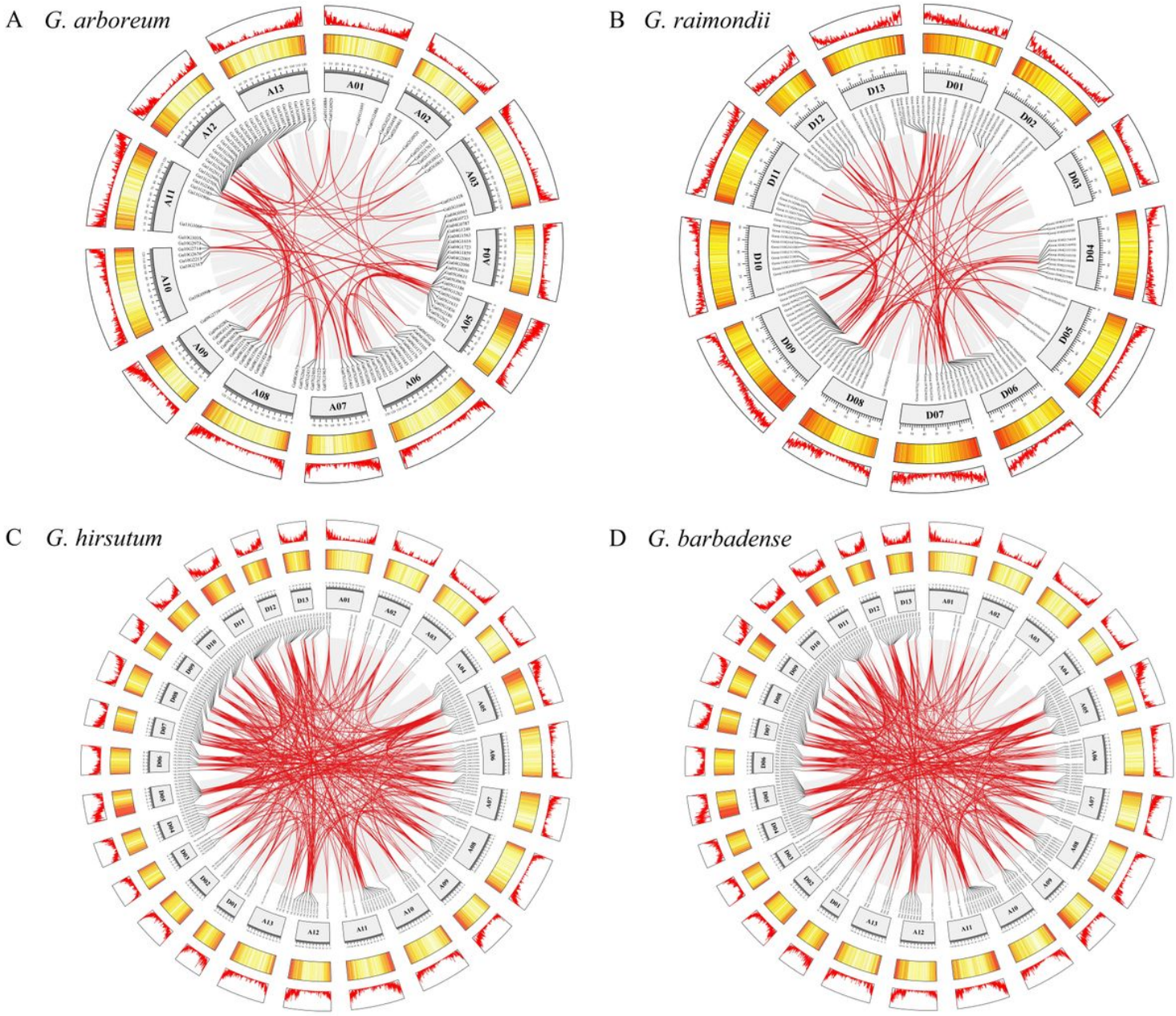

Figure 4

Duplicated WRKY genes based on the collinearity of all chromosomes in *G. arboreum* (**A**), *G. raimondii* (**B**), *G. hirsutum* (**C**), and *G. barbadense* (**D**). The number of genes is presented by a heatmap and a linear map, of which the red presents regions of high gene density, and yellow means low density region. The WRKY gene pairs with a syntenic relationship are linked by red lines, and the scale on the boxes above is in mega bases (Mb).

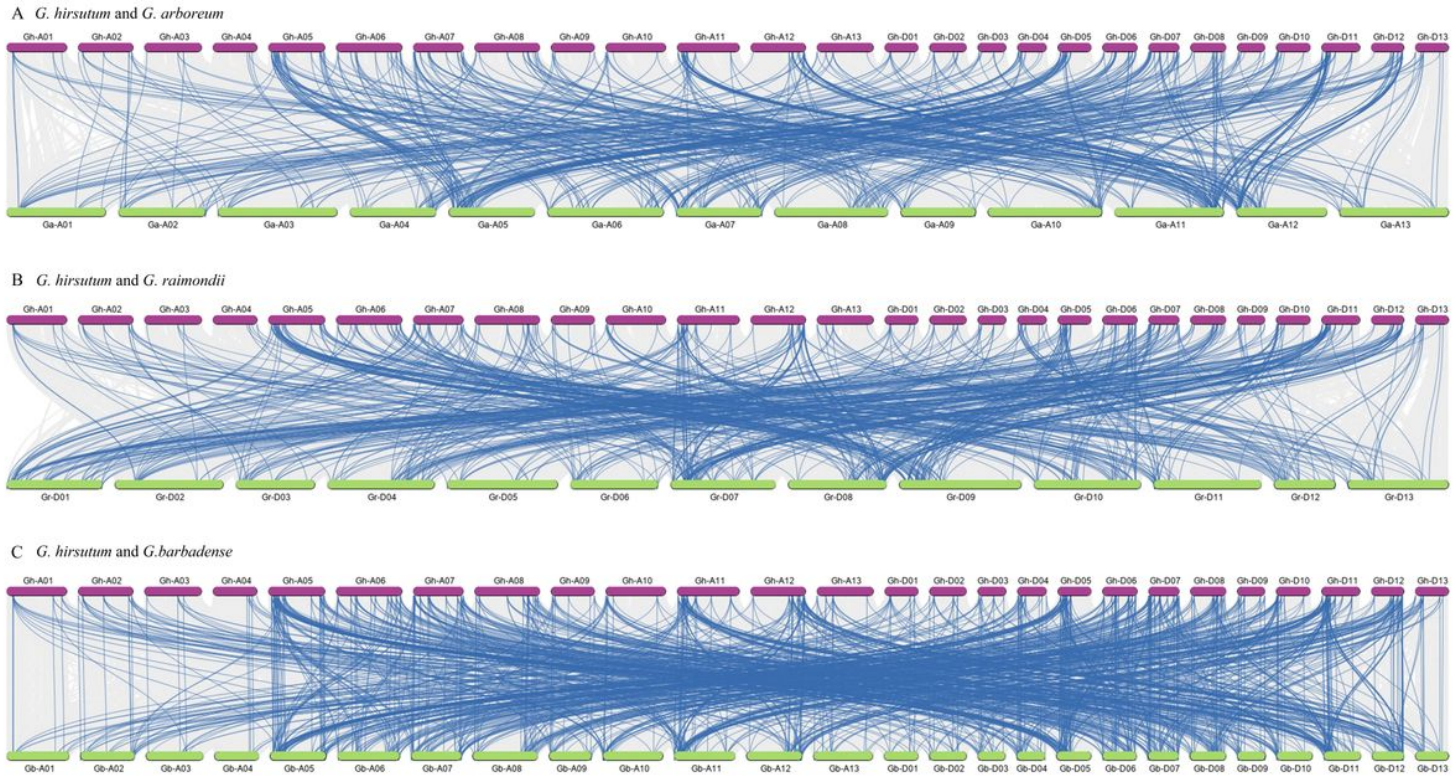

Figure 5

Synteny analysis of WRKY genes. Orthologous relationships between *G. hirsutum* and *G. arboretum* (**A**), *G. hirsutum* and *G. raimondii* (**B**), and *G. hirsutum* and *G. barbadense* (**C**) were investigated. Blue lines highlight duplicated WRKY gene pairs, while the gray lines in the background indicate all collinear relationships.

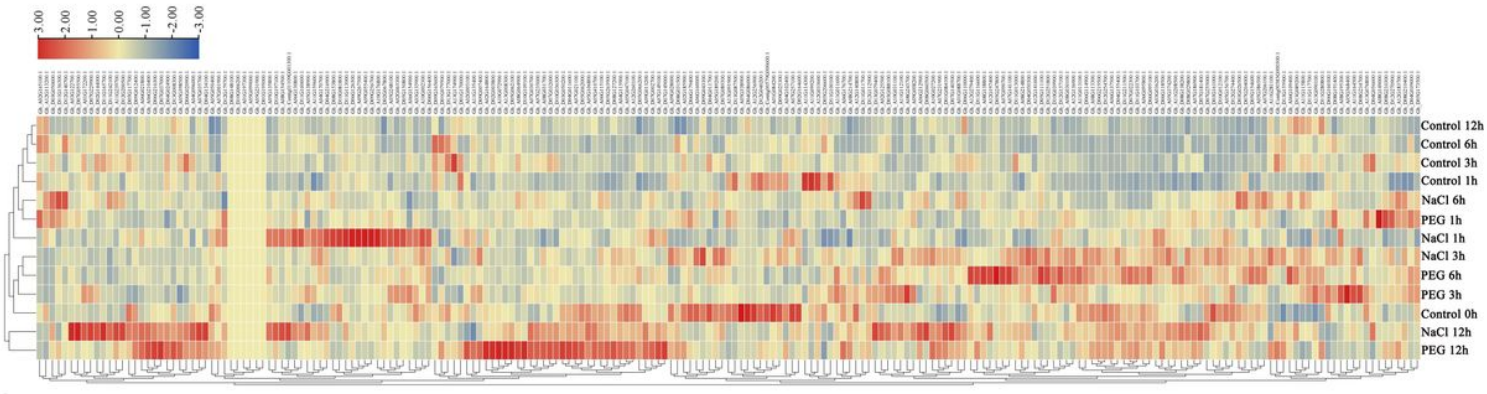

Figure 6

A cluster heatmap of expression patterns of *GhWRKY*genes in response to NaCl and PEG treatment. Each line presents the expression of WRKY gene in different treatments, and the expression values in row scale were normalized. The color scale varies from red to blue, indicating the high or low expression of each WRKY gene.

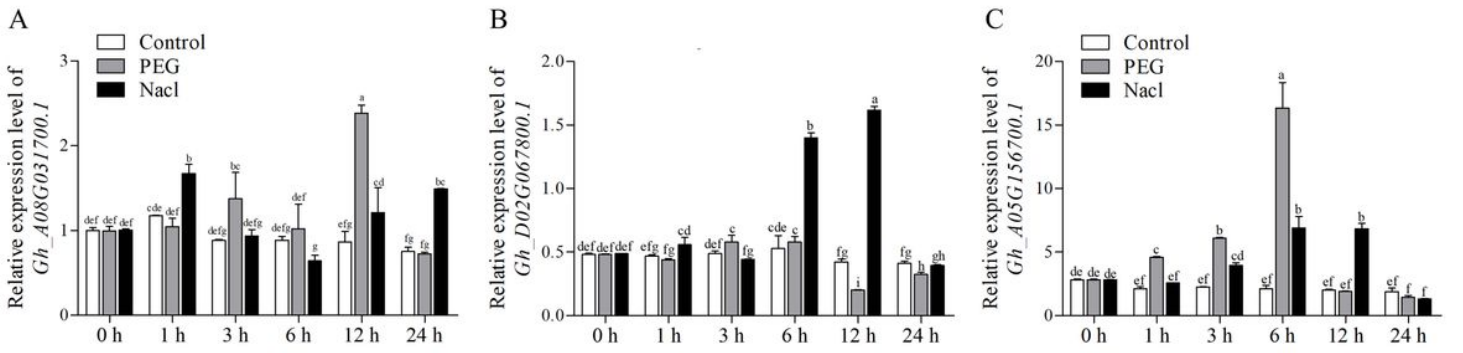

Figure 7

The expression levels of 3 WRKY genes in the leaves of *G. hirsutum* seedlings under PEG and NaCl stress. *Gh_A08G037100.1*expression in the control group was set to 100% at 0 h. The error bar represents the standard error of the mean, and the lower letter above the bar indicates a significant difference ($P < 0.05$).

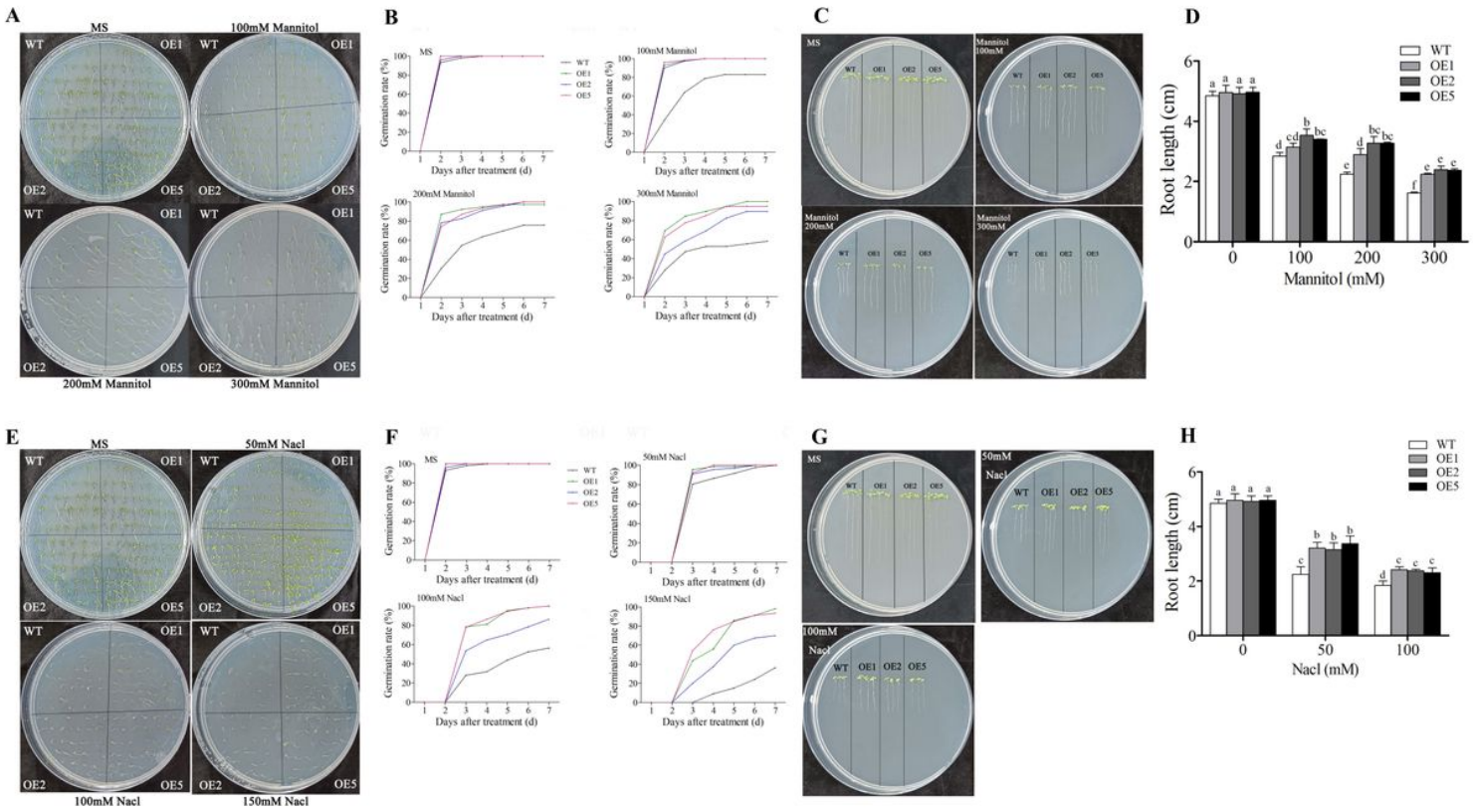

Figure 8

The germination rates and root length of *GhWRKY31* OE lines and WT under mannitol and salt conditions. **A** and **E** Phenotypic comparison of seedlings grown on 1/2 MS with 0 mM, 100 mM, 200 mM and 300 mM mannitol, or 0 mM, 50 mM, 100 mM, and 150 mM NaCl after 7 days. **B** and **F** Germination rates of seedlings grown under the conditions described in (**A**) and (**E**). **C**, **D**, **G,** and **H** Phenotypic comparison and root length of seedlings grown on 1/2 MS with 0 mM, 100 mM, 200 mM, and 300 mM mannitol, or 0 mM, 50 mM and 100 mM NaCl after 7 days. Data presents the means ± SE from three

independent experiments. The lowercase letters above the bar indicate the significant difference ($P < 0.05$).

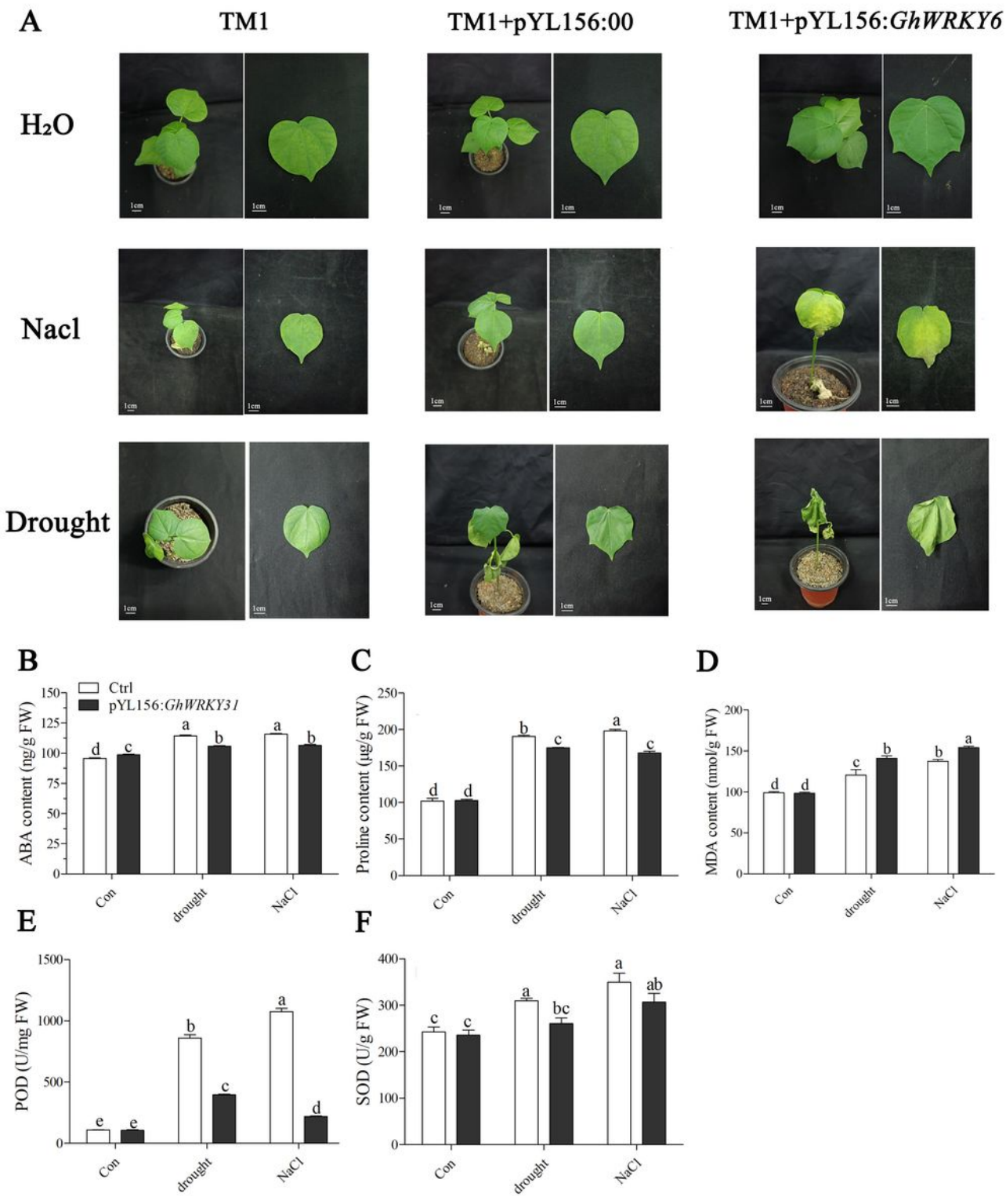

Figure 9

GhWRKY31-VIGS cotton seedlings exhibit increased sensitivity to drought and salt stress. **A** Leaves phenotypes were shrinkage, yellowing, wilting, and death under water deficit and 200 mM NaCl treatment.

ABA (**B**), proline (**C**), MDA (**D**) content, and POD (**E**), SOD (**F**) activity under water-deficit and 200 mM NaCl treatment. Data presents the means ± SE from three independent experiments. The lowercase letter above the bar indicates the significant difference (P< 0.05).

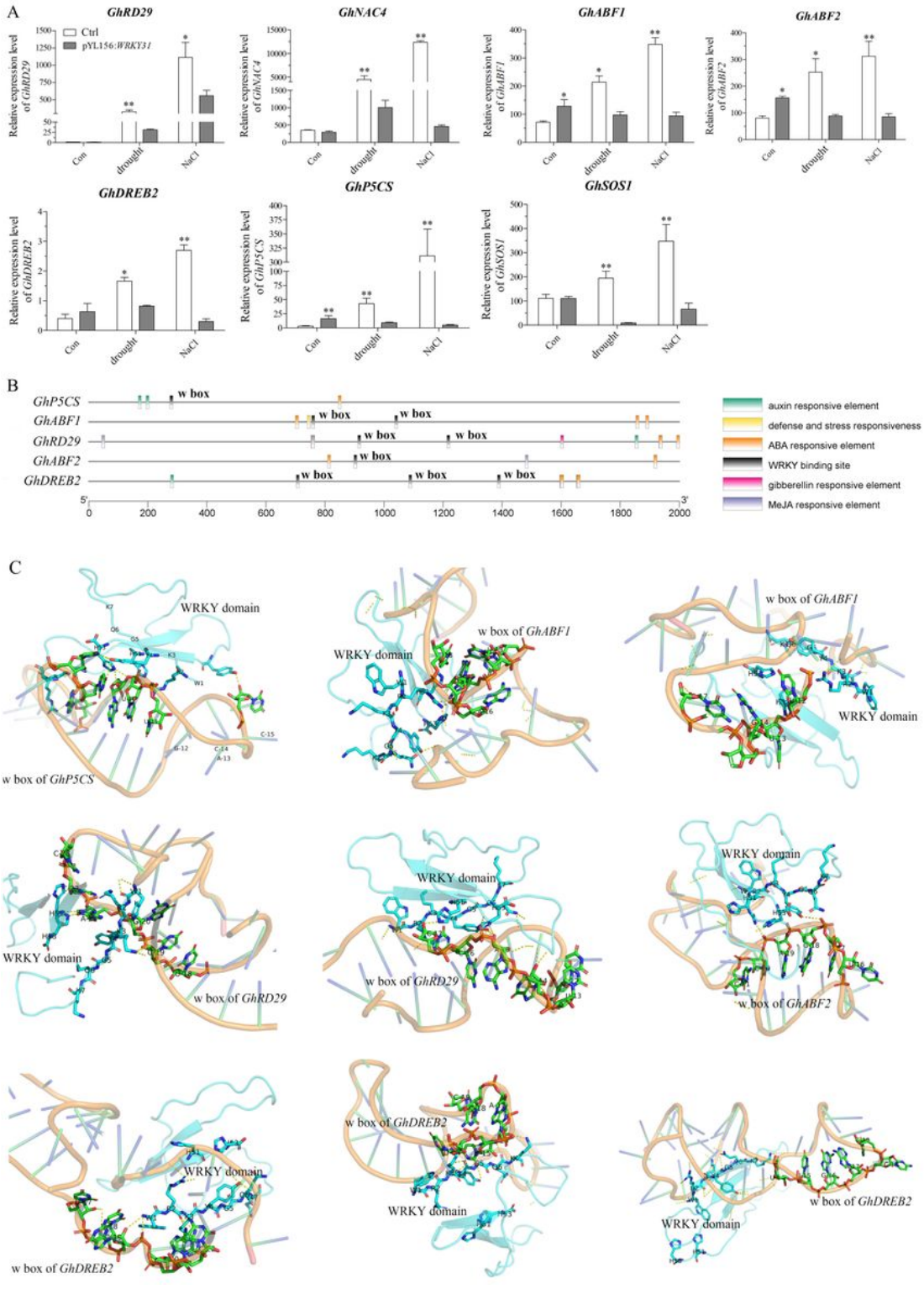

Figure 10

The gene expression levels induced by salt and dehydration were regulated by *GhWRKY31* in *G. hirsutum* leaves. **A** Silencing of *GhWRKY31* inhibits salt- and dehydration-induced gene expression. The data is shown as the mean ± SD from three independent biological replicates. (**, $P < 0.01$; *, $P < 0.05$; Student's *t*-test). **B** The *cis*-acting elements are located 2000bp upstream of the *GhP5CS*, *GhABF1*, *GhRD29*, *GhABF2*, and *GhDREB2* promoters. **C** The 3D structure of molecular docking for the binding of GhWRKY31 protein and the W boxes of *GhP5CS*, *GhABF1*, *GhRD29*, *GhABF2*, and *GhDREB2*. The yellow dashed line represents hydrogen bonding interactions.

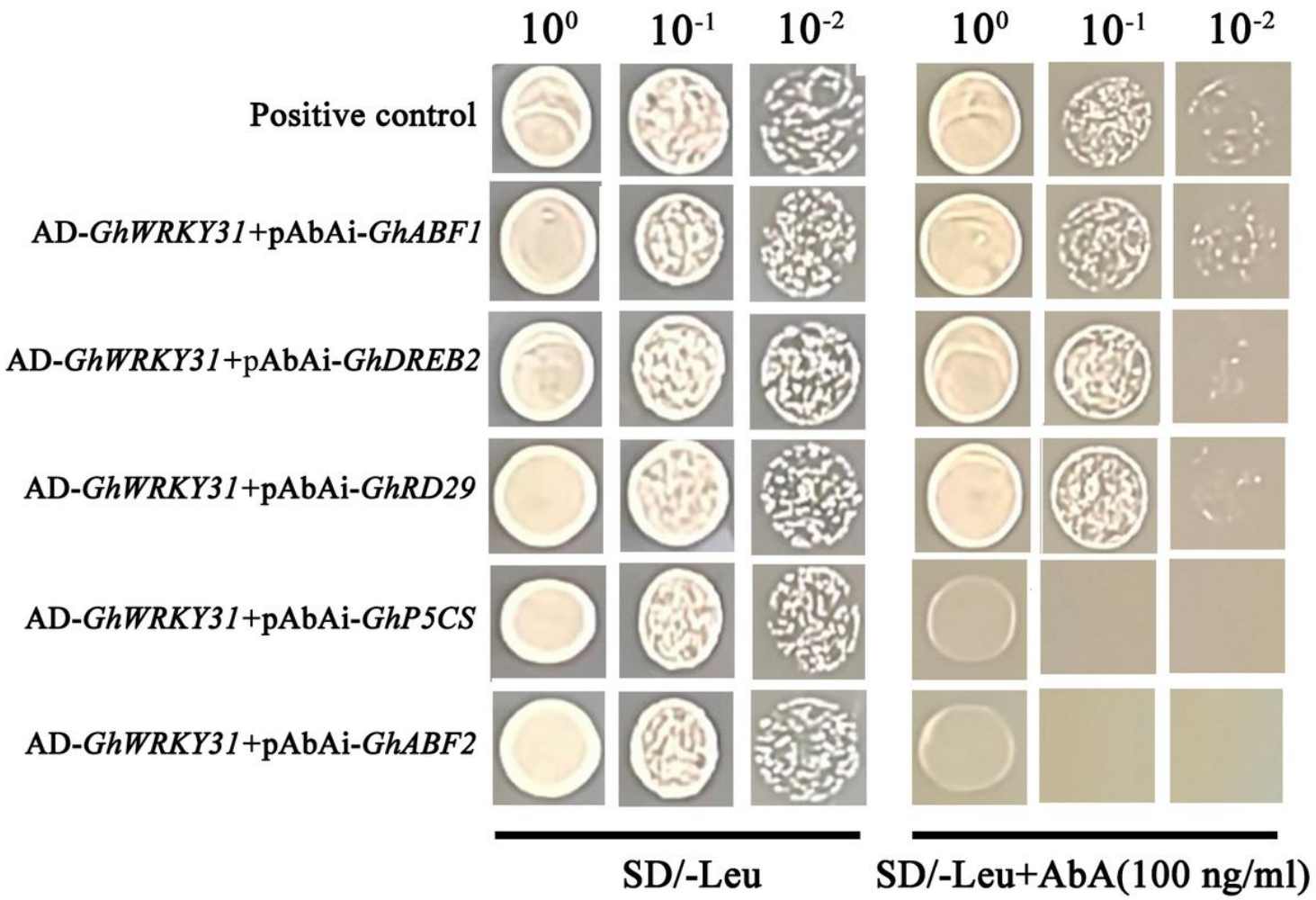

Figure 11

Y1H assay of *GhWRKY31* with *GhABF1*, *GhDREB2*, *GhRD29, GhP5CS*, and *GhABF2*. The promoters of *GhABF1*, *GhDREB2*, *GhRD29, GhP5CS*, and *GhABF2*, which contain the putative TTGACC/T transformation (W box), were constructed in the pAbAi vector. The ORF of *GhWRKY31* was constructed in the pGADT7 vector. Yeast cells were diluted with distilled water (10^0 to 10^{-2}) and cultured on SD/-Leu medium supplemented with 100 ng/mL of Aureobasidin A (AbA).

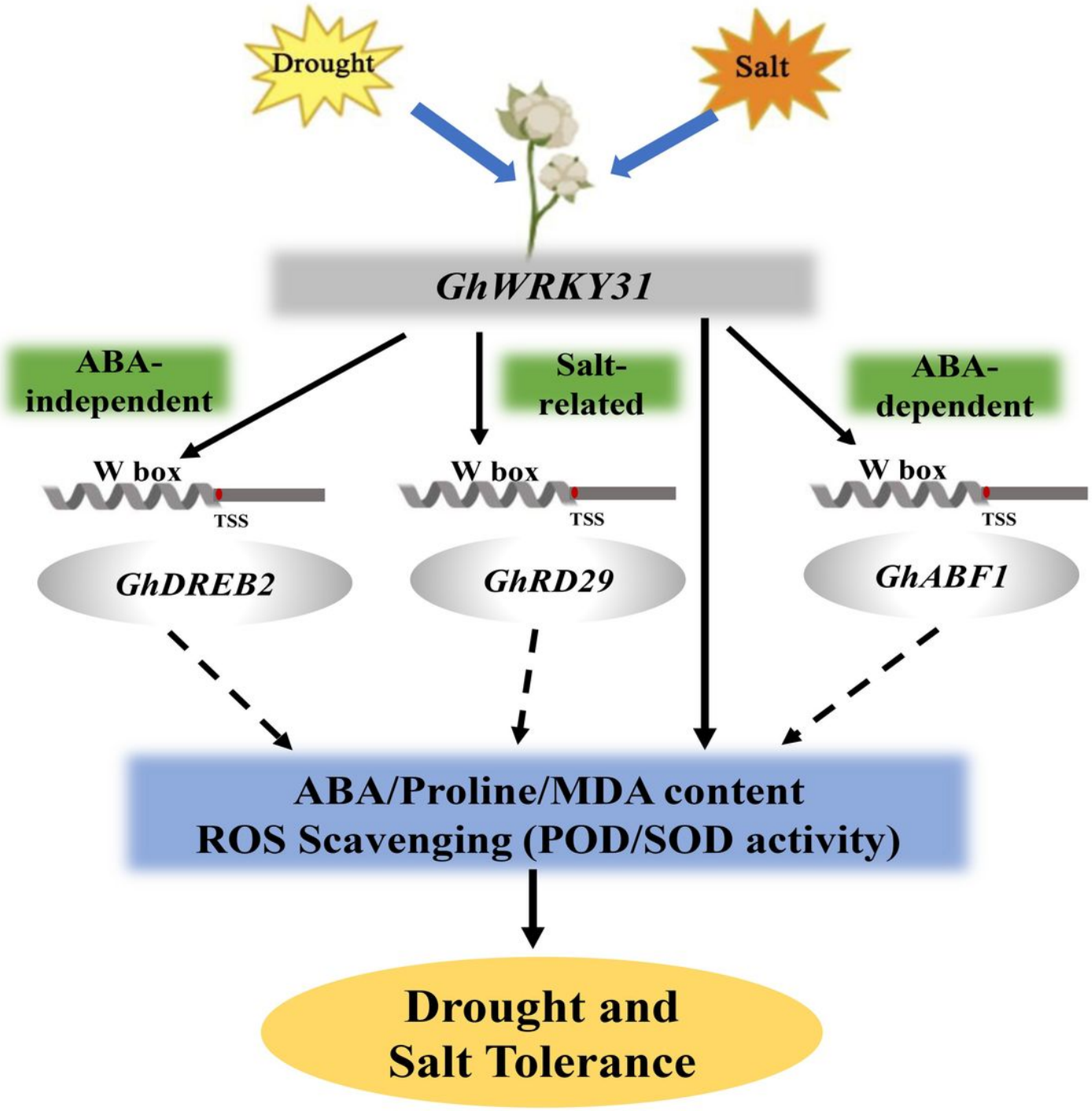

Figure 12

A working model of the role of the *GhWRKY31* module in the drought and salt stress response in cotton.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Highlights.docx
- SupplymentFigures.docx
- SupplymentTables.xlsx