RESEARCH ARTICLE

Identification of the group IIa WRKY subfamily and the functional analysis of *GhWRKY17* in upland cotton (*Gossypium hirsutum* L.)

Lijiao Gu[1], Libei Li[1], Hengling Wei[1], Hantao Wang[1], Junji Su[1], Yaning Guo[1,2], Shuxun Yu[1] *

1 State Key Laboratory of Cotton Biology, Institute of Cotton Research of CAAS, Anyang, Henan, China,
2 College of Agronomy, Northwest A&F University, Yangling, Shanxi, China

* Ysx195311@163.com

Abstract

WRKY transcription factors play important roles in plant defense, stress response, leaf senescence, and plant growth and development. Previous studies have revealed the important roles of the group IIa *GhWRKY* genes in cotton. To comprehensively analyze the group IIa *GhWRKY* genes in upland cotton, we identified 15 candidate group IIa *GhWRKY* genes in the *Gossypium hirsutum* genome. The phylogenetic tree, intron-exon structure, motif prediction and Ka/Ks analyses indicated that most group IIa *GhWRKY* genes shared high similarity and conservation and underwent purifying selection during evolution. In addition, we detected the expression patterns of several group IIa *GhWRKY* genes in individual tissues as well as during leaf senescence using public RNA sequencing data and real-time quantitative PCR. To better understand the functions of group IIa *GhWRKYs* in cotton, *GhWRKY17* (KF669857) was isolated from upland cotton, and its sequence alignment, promoter *cis*-acting elements and subcellular localization were characterized. Moreover, the over-expression of *GhWRKY17* in *Arabidopsis* up-regulated the senescence-associated genes *AtWRKY53*, *AtSAG12* and *AtSAG13*, enhancing the plant's susceptibility to leaf senescence. These findings lay the foundation for further analysis and study of the functions of WRKY genes in cotton.

Introduction

Throughout their life cycle, plants exhibit a set of complex adjustment mechanisms to perceive and respond to various physiological and developmental signals. Among these diverse adjustment mechanisms, transcriptional regulation mechanisms, which are mainly executed by transcription factors (TFs), play important roles [1]. For example, WRKY proteins constitute one of the largest TF families in plants [2]. Since the first WRKY gene, *SPF1*, was reported in sweet potato [3], increasing numbers of WRKY genes have been reported in various species, including *Arabidopsis thaliana*, *Gossypium hirsutum*, *Oryza sativa*, *Ricinus communis*, *Manihot esculenta* and *Cucumis sativus* [4–9].

WRKY TFs were named for their conserved WRKY domain, which consists of approximately 60 amino acids containing a conserved WRKYGQK core sequence and a zinc finger-

like motif [2]. The WRKY TF family is divided into three main groups according to the number of WRKY domains and the pattern of zinc finger motifs: group I contains two WRKY domains each with a C2H2 zinc finger motif, whereas group II and group III each contain a single WRKY domain with either a C2H2 zinc finger motif or a C2HC zinc finger motif, respectively. However, group II can be further divided into five subgroups (IIa, IIb, IIc, IId, and IIe) according to the amino acid motifs outside the WRKY domain [2,10–13]. The group IIa WRKY domain possesses a conserved VQR intron in the zinc finger motif nearer to the C-terminus [14].

Previous studies have reported 3 group IIa WRKY genes in *Arabidopsis thaliana*, 4 in *Oryza sativa*, 6 in *Gossypium hirsutum*, 5 in *Populus trichocarpa* and 4 in *Cucumis sativus* [9,15,5]. Group IIa WRKY genes appear to include a small number of members, but they participate widely in the regulation of diverse physiological processes, such as defense, trichome development and leaf senescence [16,17,13,18]. In *Arabidopsis*, the genes *AtWRKY18*, *AtWRKY40* and *AtWRKY60* represent the group IIa WRKY subfamily [2]. These three homologs exhibit a complex pattern of physical and functional interactions in response to the microbial pathogens *Pseudomonas syringae* and *Botrytis cinerea* [19]. In addition, the over-expression of *AtWRKY18* in *Arabidopsis* can delay leaf senescence, but *AtWRKY18* T-DNA insertion lines show an early leaf senescence phenotype. An *AtWRKY18-AtWRKY53*-mediated signaling pathway is involved in the senescence process [20]. However, more work has focused on the role of these three genes in the abscisic acid (ABA) signaling pathway. For example, the three WRKY genes were identified as negative regulators of ABA signaling [21,22] and can bind to W-box elements in the promoter regions of *ABI4* and *ABI5* to inhibit the expression of these two genes [23]. In rice, *OsWRKY28*, *OsWRKY62*, *OsWRKY71* and *OsWRKY76* are four members of the *OsWRKY* group IIa subfamily and are involved in modulating plant innate immunity [24]. *OsWRKY28* plays a negative regulatory role in the resistance to rice blast fungus Ina86-137, as determined by phenotypic analysis of an *oswrky28* mutant [25]. *OsWRKY71* was found to be induced by salicylic acid (SA), methyl jasmonic acid (MeJA) and pathogen infection [26]. *OsWRKY62* responds negatively to innate immunity in terms of susceptibility to pathogens [27]. Additionally, inducible alternative splicing of the genes *OsWRKY62* and *OsWRKY76* participates in pathogen defense, as found through the analysis of over-expression and loss-of-function knockout rice plants [28]. Furthermore, the group IIa WRKY genes have also been studied in other species. For example, the over-expression of *PtrWRKY40* can enhance resistance to the necrotrophic fungus *B. cinerea* in *Arabidopsis* and susceptibility to *Dothiorella gregaria* in poplar [29], and *TaWRKY71-1* presents a hyponastic leaf phenotype by altering IAA levels in transgenic *Arabidopsis* [30].

Since the release of abundant genome sequences and publicly available transcriptome data for cotton [31–34], preliminary analyses of the group IIa WRKY subfamily genes have been performed. The three *Gossypium aridum* group IIa genes *GarWRKY28*, *GarWRKY51* and *GarWRKY52* are salt-responsive genes that display distinct expression levels in response to salt treatment [35]. Gene sequence analysis of the group IIa WRKY genes in *Gossypium raimondii* and *Gossypium arboreum* revealed a higher number of SNPs in intron regions than in exon regions. Based on protein sequence analysis, the WRKY domain regions were found to be more conserved than the regions outside the WRKY domain [36]. Cai *et al*. isolated 7 group IIa genes from *Gossypium raimondii*, and expression profiling showed that *GrWRKY24* and *GrWRKY40* are significantly induced by salt, drought and disease treatments [37]. The results of Dou *et al*. indicated that group IIa WRKY genes play important roles in leaf senescence, anther development, fiber growth, and abiotic and biotic stresses responses [5]. Moreover, several reports have addressed the functional and mechanistic details of group IIa WRKY genes in cotton. The stress–induced gene *GhWRKY40* enhances wounding tolerance and sensitivity

to *Ralstonia solanacearum* infection in transgenic tobacco [38]. The transgenic *Arabidopsis* lines of *GarWRKY17* and *GarWRKY104* can enhance salt tolerance during different developmental stages [35]. Above all, most group IIa WRKYs have significant functions in the regulation of stress response and plant growth development.

Cotton is an important economic crop and textile material that plays important roles in the development of the national economy [39]. Previous findings have shown that stresses and senescence are important factors restricting cotton growth, fiber quality and yield. To date, the functional analysis of group IIa *GhWRKYs* has mainly focused on stress conditions, with limited analysis of their possible roles in leaf senescence. Here, we report an analysis of group IIa *GhWRKY* TFs in upland cotton pertaining to phylogeny, intron-exon structure, motif composition, Ka/Ks ratios, and expression patterns in different tissues and during leaf senescence. To further analyze the function of group IIa WRKY genes, *GhWRKY17* was isolated and characterized. *GhWRKY17* expression could be induced during the leaf senescence process and over-expression of *GhWRKY17* resulted in an early aging phenotype in transgenic *Arabidopsis*.

Materials and methods

Characterization of putative WRKY genes in cotton

The genome and protein sequences of *Gossypium hirsutum* were downloaded from the CottonGene database (http://www.cottongen.org) [40]. The Hidden Markov Model (HMM) profile of the WRKY domain (PF03106, WRKY.hmm) was obtained from the Pfam database (http://pfam.janelia.org). WRKY.hmm was then used to search the database for all candidate genes. The SMART program (http://smart.embl-heidelberg.de/) was employed to confirm the potential proteins according to the structural features of WRKY [41,42].

Sequence analysis

For convenience, the gene names *GhWRKY1* to *GhWRKY239* were given based on their gene IDs in the genome database. To better clarify the evolutionary relationship and the classification of different clades, a phylogenetic tree was constructed with the MEGA 7 program [43] using the maximum likelihood method [44] or the neighbor-joining method [45]. MapChart was used to draft a chromosomal distribution sketch map of WRKY genes [46]. Diverse exon-intron structures were identified by inputting GFF3 format data on the group IIa WRKY genes into the Gene Structure Display Server (GSDS2.0) (http://gsds.cbi.pku.edu.cn/) [47]. The Multiple Em for Motif Elicitation (MEME) online software (http://meme-suite.org/tools/meme) was used to predict conserved motifs in group IIa WRKY proteins. To analyze the selective pressures among group IIa WRKY genes, Ka/Ks (nonsynonymous/synonymous) ratios were computed using the PAL2NAL web server (http://www.bork.embl.de/pal2nal/#RunP2N) [48].

Expression pattern analysis

Public cotton RNA-Seq data were obtained from the high-throughput DNA and RNA Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/). The accession numbers of the different tissues of *Gossypium hirsutum* L. acc. TM-1 were SRX797899-SRX797920 (SRA: PRJNA248163) [40]. The RNA-Seq data from *Gossypium hirsutum* during the process of leaf senescence were also analyzed. The accession numbers of new/old (new-1 and old-1) leaves from the three-leaf stage and of new/old (new-2 and old-2) leaves from the maturation/senescence stages were SRX1075619, SRX1075620, SRX1075623 and SRX1075624 [49].

Plant materials and stress treatments

To evaluate the gene expression patterns in different tissues and during natural leaf senescence, the early-aging variety CCRI74 and the non-early-aging variety Liao4086 were planted in the cotton field at the Institute of Cotton Research of CAAS (Anyang, Henan, China). For tissue expression analysis, tissues including root, stem, leaf, petal, pistil, stamen, ovule and fiber were collected from CCRI74. To detect gene expression during natural leaf senescence, two methods were used to collect the samples. To examine different leaf development stages, the top fourth leaves from two varieties were sampled every 10 days from the 80th day after sowing. To examine different leaf senescence areas in one leaf, five different stages of leaves were harvested from CCRI74: Stage 1, an expanded new leaf; Stage 2, a mature but non-senescent leaf; Stage 3, a leaf with 25% senescence area; Stage 4, a leaf with 50% senescence area; and Stage 5, a leaf with at least 75% senescence area.

For stress treatments, healthy and plump seeds of CCRI74 were planted in pots in the greenhouse at 28˚C under a 16 h light/8 h dark cycle. When the seedlings had grown to the cotyledon stage, healthy and uniform 10-day-old seedlings were used for different treatments. In the exogenous hormone treatments, the seedlings were sprayed with 100 μM MeJA, 2 mM SA, 200 μM ABA and 0.5 mM ethylene released from ethephon (ETH). In the abiotic stress treatments, the seedlings were irrigated with 15% polyethylene glycol 6000 (PEG6000) and 200 mM sodium chloride (NaCl). The cotyledons were sampled at 0 h, 2 h, 4 h, 8 h, and 12 h. The samples were frozen in liquid nitrogen and used in the subsequent experiments. The experiments were repeated at least three times.

Quantitative real-time PCR (qRT-PCR)

The total RNA was isolated using an RNAprep PurePlant Kit (Polysaccharides & Polyphenolics-rich) (Tiangen, China). The cDNA was synthesized using a PrimeScript™ RT Reagent Kit with gDNA Eraser (Perfect Real Time) (TaKaRa, Japan). qRT-PCR was performed to identify transcript levels using a 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) and UltraSYBR Mixture (With ROXI) (CWBIO, China) in a 20 μl volume: 2×UltraSYBR Mixture (With ROXI) 10 μl, forward primer 0.5 μl, reverse primer 0.5 μl, template cDNA 1 μl, and RNase-free water 9 μl. The PCR procedure was as follows: a pre-denaturation step at 95˚C for 10 min; 40 cycles of 95˚C for 10 s, 60˚C for 30 s and 72˚C for 32 s; and a melting curve of 95˚C for 15 s, 60˚C for 1 min, 95˚C for 15 s and 60˚C for 15 s. The $2^{-\Delta\Delta CT}$ method was applied to calculate the relative expression of genes [50]. Figures for qRT-PCR were drawn using GraphPad Prism software [51].

Promoter cloning

The upstream 1500 bp sequence of *GhWRKY17* was obtained from the genome sequence data of upland cotton (http://www.cottongen.org) [36], and the promoter fragment was obtained from DNA by homology-based cloning [52]. Cis-acting elements in the promoter region were predicted by the online software PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [53].

Subcellular location

GhWRKY17 protein subcellular localization was predicted by WoLF PSORT (http://www.genscript.com/wolf-psort.html). In addition, the coding region of *GhWRKY17* without the termination codon was cloned into the *pBI121-GFP* vector to construct the plasmid *pBI121-GhWRKY17-GFP* driven by the *CaMV35S* promoter. Both the recombinant plasmid

pBI121-GhWRKY17-GFP and the empty vector *pBI121-GFP* wrapped with gold powder were transferred into onion epidermal cells cultivated on MS plates using a desk-type particle gun PDS-1000/He system (Bio-Rad) with the following parameters: particle bombardment running distance 9 cm, rupture disk pressure 1300 psi and vacuum degree 28 mmHg. After bombardment, the onion tissues were transferred onto new MS agar medium and incubated at 25°C for 12 h in the dark, and the green fluorescence of the cells was observed using a confocal laser scanning microscope (ZEISS LSM 700).

Transformation of *Arabidopsis* and phenotype observation

The coding fragment of *GhWRKY17* was inserted into the *BamHI/EcoRI* sites of the binary vector pBI121 to generate the *pBI121-GhWRKY17* recombinant plasmid. The recombinant binary vector was transferred into *Agrobacterium tumefaciens* strain LBA4404, and the positive clones were screened with kanamycin (50 mg/ml). *Columbia* ecotype *Arabidopsis thaliana* plants at the initial fruiting period were transformed by floral dipping [54]. The seeds we obtained were called the T_0 generation. From the beginning of the T_0 generation, we screened the seeds on 1/2 MS containing kanamycin (50 mg/ml), and PCR was performed to eliminate false positive plants until the T_3 transgenic homozygous generation. Wild-type (WT) and T_3 transgenic lines were planted in the greenhouse at 22°C under a 16 h light/8 h dark cycle to observe the natural senescence phenotype. The rosette leaves were harvested for gene expression analysis.

For the ABA treatment, WT and three transgenic lines were germinated on 1/2 MS solid medium for three days. They were then transferred to MS solid medium containing 10 μM ABA for seven days in the vertical position. The phenotype of the seedlings was recorded, and the root length was measured.

All primers used in this paper are listed in S1 Table and were designed using the OLIGO 7 software [55].

Results

Identification and characterization of group IIa WRKY genes in *Gossypium hirsutum*

All the candidate *GhWRKY* proteins were identified using HMMER searching and the SMART program. We obtained 239 *GhWRKYs* in upland cotton, and 3 genes with incomplete WRKY domain structures, *GhWRKY27*, *GhWRKY238* and *GhWRKY239* were classified as group IV (S2 Table). Because of the different names of WRKY genes in different publications, we collected the data, performed a similarity comparison and organized the list in S2 Table for easy reference. The *GhWRKYs* were mapped to different *Gossypium hirsutum* genome chromosomes and showed a heterogeneous distribution (S1 Fig). The proteins that clustered with group IIa *AtWRKYs* were considered to be group IIa *GhWRKYs* (S2 Fig). We identified 15 group IIa *GhWRKY* members (S2 and S3 Tables). The 15 group IIa *GhWRKYs* were mainly scattered on chromosomes 5, 6 and 7 and ranged from 140 to 1442 amino acids in protein length. However, 9 genes contained 5 exons and 4 introns, accounting for more than half of the total of 15 group IIa *GhWRKY* genes. Moreover, six genes (*GhWRKY17, 18, 20, 44, 49 and 155*) were found by Dou *et al.* [5] to have similarity exceeding 97% (Table 1).

Phylogenetic tree, intron–exon structure and motif composition in group IIa *GhWRKY* genes

The 15 candidate group IIa *GhWRKYs* were subjected to phylogenetic analysis using the MEGA 7 program with the neighbor-joining method. The phylogenetic tree showed that

Table 1. General group IIa *GhWRKY* gene information.

Gene name[a]	Gene ID[b]	Start	End	Protein length	Exon	Intron	Gene name[c]	Accession no.[d]	Identity(%)
GhWRKY17	Gh_A05G0483	5264475	5265826	313	5	4	*GhWRKY71*	KF669857	97.44
GhWRKY18	Gh_A05G0484	5267711	5268953	252	4	3	*GhWRKY70*	KF669834	99.6
GhWRKY20	Gh_A05G1019	10243660	10244007	314	5	4	*GhWRKY73*	KF669835	99.35
GhWRKY39	Gh_A06G0917	36697073	36697435	313	5	4	—		—
GhWRKY44	Gh_A06G1923	148517	149874	326	4	3	*GhWRKY1*	KF669831	100
GhWRKY49	Gh_A07G0261	3249562	3250831	305	5	4	*GhWRKY11*	KF669832	98.69
GhWRKY50	Gh_A07G0263	3282938	3298611	1442	12	11	—		—
GhWRKY140	Gh_D05G0600	4880919	4882276	313	5	4	—		—
GhWRKY141	Gh_D05G0601	4884105	4885375	252	4	3	—		—
GhWRKY144	Gh_D05G1137	9746886	9747233	316	5	4	—		—
GhWRKY155	Gh_D06G1078	24005778	24006137	312	5	4	*GhWRKY83*	KF669802	99.68
GhWRKY156	Gh_D06G1082	24288383	24289803	140	3	2	—		—
GhWRKY163	Gh_D06G1966	61082887	61083234	326	4	3	—		—
GhWRKY168	Gh_D07G0317	3360767	3362036	305	5	4	—		—
GhWRKY169	Gh_D07G0318	3381256	3383011	309	5	4	—		—

We name all the *GhWRKY* genes based on their gene IDs in the *Gossypium hirsutum* genome sequence database.

Gene name[a] shows the group IIa *GhWRKY* genes among all the *GhWRKYs* identified.

Gene ID[b] is the ID no. in the *Gossypium hirsutum* genome sequence database.

Gene name[c] and Accession no.[d] indicate the group IIa *GhWRKY* genes identified by Dou *et al.* [5].

"—" represents no results.

GhWRKY17 and *GhWRKY140* were clustered together, followed by *GhWRKY39* and *GhWRKY155*, *GhWRKY49* and *GhWRKY168*, *GhWRKY20* and *GhWRKY144*, *GhWRKY44* and *GhWRKY163*, *GhWRKY18* and *GhWRKY141*, and *GhWRKY50* and *GhWRKY169* (Fig 1A). Among the 15 *GhWRKY* genes, 7 paralogs (*GhWRKY17/140*, *GhWRKY18/141*, *GhWRKY20/144*, *GhWRKY39/155*, *GhWRKY44/163*, *GhWRKY49/168* and *GhWRKY50/169*)
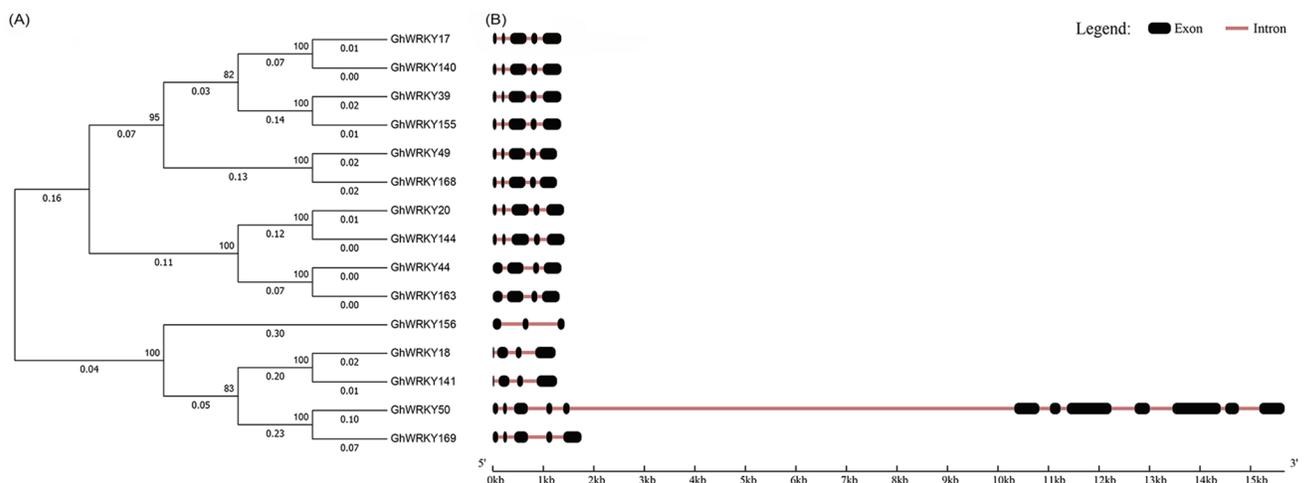

Fig 1. Phylogenetic analysis and exon-intron structures of group IIa *GhWRKY* genes. (A) Phylogenetic analysis among group IIa *GhWRKY* genes. The phylogenetic tree was constructed based on the protein sequences using the MEGA 7 program. The neighbor-joining method was used, and bootstrap analysis was performed with 1000 replications. (B) Exon-intron composition of group IIa *GhWRKY* genes. Exons and introns are represented by black boxes and dark red lines, respectively.

Table 2. The Ka/Ks ratios for duplicate group IIa *GhWRKY* genes.

Paralogous	Identities (%)	Ka	Ks	Ka/Ks	Purifying selection
GhWRKY17/140	98.09	0.0173	0.0274	0.6287	YES
GhWRKY18/141	98.29	0.0128	0.0315	0.406	YES
GhWRKY20/144	97.58	0.9186	0.8633	1.064	NO
GhWRKY44/163	98.57	0.0042	0.0481	0.0863	YES
GhWRKY49/168	97.6	0.02099	0.0344	0.6066	YES
GhWRKY39/155	97.66	—	—	—	—

"—" represents no results.

were found in the phylogenetic tree (Fig 1A). Among these 7 paralogs, one gene of each paralog comes from the A genome and the other one from the D genome (Table 1). Usually, the criteria for inferring a gene duplication event are that the length of the alignment sequence covers at least 80% of the longest gene and that the similarity of the aligned regions exceeds 70% [56]. By calculating the sequence coverage and similarity, we identified 6 of the 7 paralogs as having undergone gene duplication, while the genes *GhWRKY50* and *GhWRKY169* exhibited less than 70% similarity (18.08%) (Table 2). For convenience, we defined the 6 paralogs as gene pairs here, and a total of 6 gene pairs occurred among the group IIa *GhWRKY* genes. As shown in Fig 1B, the two genes in each gene pair shared a similar intron-exon structure with the same number of introns and exons and similar length at the nucleic acid and amino acid level (Fig 1B). Furthermore, most of the group IIa WRKY genes had 2 to 4 introns, including 1 gene containing 2 introns, 4 genes containing 3 introns and 9 genes containing 4 introns. Interestingly, the gene *GhWRKY50* contained 11 introns (Fig 1B and Table 1).

The MEME software identified six conserved motifs among the group IIa WRKY proteins. The results showed high similarity in a conserved sequence frame within the group IIa WRKY members. All the members contained six motifs except for *GhWRKY50*, which lacked motif 3, and *GhWRKY156*, which lacked motifs 2, 3 and 6 (Fig 2). Moreover, the two genes in each gene pair possessed the same motif composition (Fig 2).
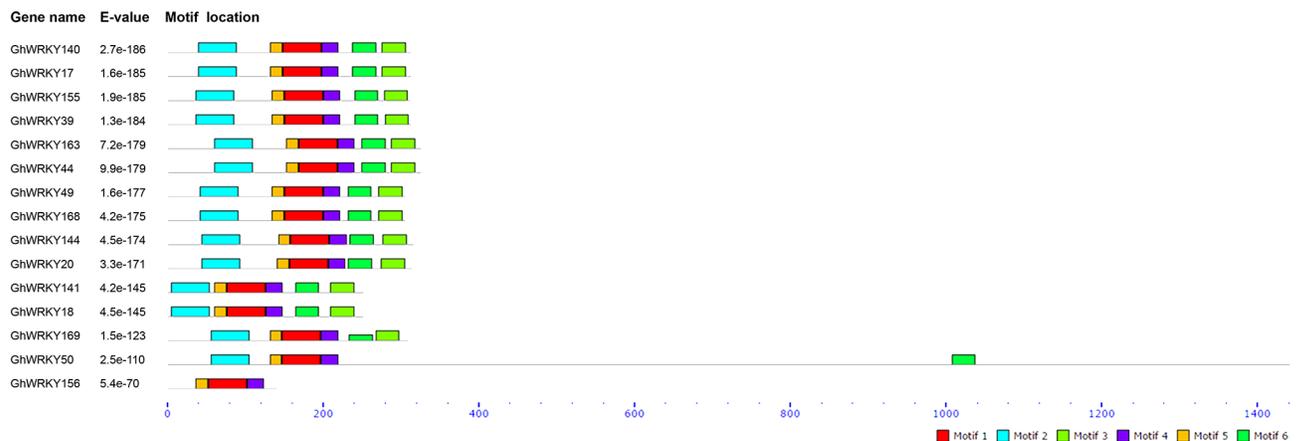

Fig 2. Motif composition in group IIa *GhWRKY* subfamily proteins. Each colored box indicates a different putative motif. The scale plate indicates the protein length. The combined E-value was calculated by the MEME online software.

Ka/Ks analysis

To characterize the evolutionary history of the group IIa *GhWRKY* genes, Ka, Ks and the Ka/Ks ratio were calculated. Apart from the unidentified pair *GhWRKY39/155*, the Ka/Ks ratios of five pairs (*GhWRKY17/140*, *GhWRKY18/141*, *GhWRKY20/144*, *GhWRKY44/163* and *GhWRKY49/168*) were calculated. The similarity of the 5 homologous protein-coding gene pairs exceeded 97% (Table 2). The Ka and Ks values of all 5 gene pairs were lower than 1. For the 4 gene pairs *GhWRKY17/140*, *GhWRKY18/141*, *GhWRKY44/163* and *GhWRKY49/168*, the Ka values were lower than their Ks values, with Ka less than 0.03, and their Ka/Ks ratios were therefore less than 1. However, for the remaining pair, *GhWRKY20/144*, the Ka/Ks ratio was greater than 1 (Table 2).

Distinct expression profiles of group IIa *GhWRKY* genes in various tissues and at different leaf senescence stages

To further study the function of group IIa *GhWRKYs* in cotton, RNA-Seq data from a public database were used to detect group IIa *GhWRKY* gene expression in twelve tissues and at different leaf senescence stages. According to the RNA-Seq data, no transcript expression was detected for *GhWRKY156*, but the remaining 14 genes showed differential expression levels.

As shown in Fig 3A, *GhWRKY17*, *GhWRKY39* and *GhWRKY140* had relatively high expression levels and showed differential expression levels in almost all tissues. The expression of *GhWRKY17* and *GhWRKY140* was the highest in the stem (FPKM>180), followed by the calycle (FPKM>140) and fiber 25 (FPKM>80). The expression of *GhWRKY39* was also the highest in the stem (FPKM>280), but followed by the torus (FPKM>100) and fiber 25 (FPKM>50). However, the remaining genes showed relatively low expression levels overall, with some genes presenting high expression levels in specific tissues. For example, *GhWRKY49*, *GhWRKY168* and *GhWRKY169* were highly expressed specifically in the torus with FPKM values greater than 130 but were hardly expressed in other tissues (Fig 3A).

To detect the gene expression patterns during leaf senescence, we analyzed the RNA-Seq data on new and old leaves from both the three-leaf stage and the maturation/senescence stages. Fig 3B shows that as observed in the tissue expression, the three genes *GhWRKY17*, *GhWRKY39* and *GhWRKY140* also exhibited relatively high expression levels and showed differential expression in new and old leaves. These three genes were specifically highly expressed in old-1 leaves from the three-leaf stage with FPKM values more than 110, and the FPKM value of *GhWRKY39* was near 200. However, the remaining genes in both new and old leaves had FPKM values no greater than 50 (Fig 3B). Taken together, *GhWRKY17*, *GhWRKY39* and *GhWRKY140* may be preferable genes for follow-up studies.

Tissue-specific and senescence expression patterns of *GhWRKY17*, *GhWRKY39* and *GhWRKY140* analyzed by qRT-PCR

To further discover the functions of group IIa *GhWRKY* genes in different tissues and different stages of the leaf senescence process, according to the expression profiles above, we selected three genes (*GhWRKY17*, *GhWRKY39* and *GhWRKY140*) with high expression abundance and differential expression levels to conduct qRT-PCR using CCRI74 material. The tissue expression results showed that the three genes were highly expressed in the stem and floral organs, and *GhWRKY39* and *GhWRKY140* presented similar expression patterns (Fig 4A).

To determine the involvement of the three genes in leaf senescence, we focused on their expression levels at five development phases of blades with different aging of the leaf area (Fig
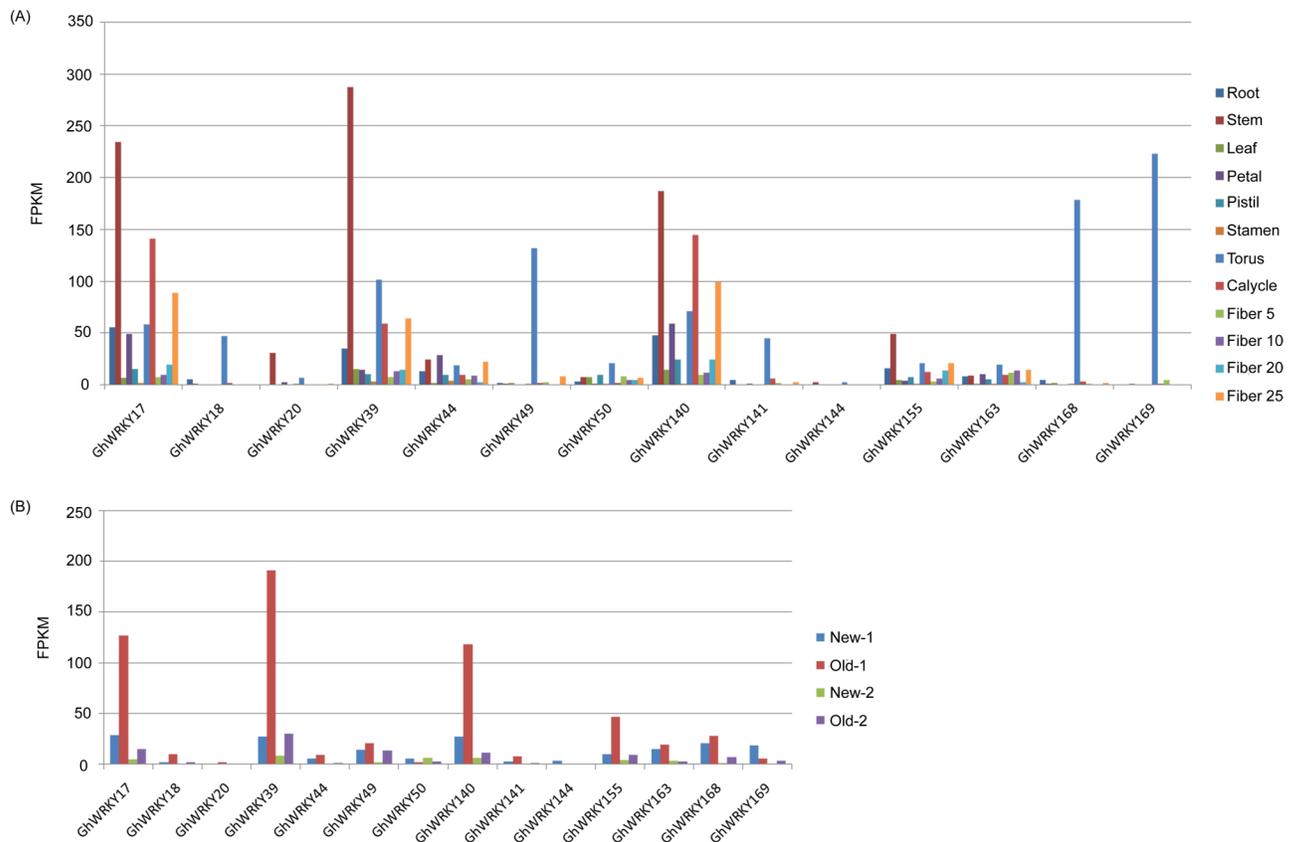
Fig 3. Expression patterns of group IIa *GhWRKY* genes in various tissues and at different leaf senescence stages. (A) Expression patterns of group IIa *GhWRKY* genes in root, stem, leaf, petal, pistil, stamen, torus, calycle, fiber 5, fiber 10, fiber 20 and fiber 25 tissues. Fiber 5, fiber 10, fiber 20 and fiber 25 indicate fiber development stages at 5, 10, 20, and 25 days after anthesis. (B) Expression patterns of group IIa *GhWRKY* genes in leaf senescence. New-1 and old-1 indicate leaves from the three-leaf stage. New-2 and old-2 indicate leaves from the maturation/senescence stages.

4B). *GhNAP* [57] is an up-regulated senescence marker gene in cotton. During leaf development from stage 1 (an expanded new leaf) to stage 5 (a leaf with at least 75% senescence area), the expression level of *GhNAP* was gradually up-regulated, which confirmed the phenotype above (Fig 4B and 4C). However, *GhWRKY17*, *GhWRKY39* and *GhWRKY140* all showed first an increasing and then a descending tendency. They peaked at stage 3, which is the starting point of aging (Fig 4C). Given the similar expression patterns of the three genes, *GhWRKY17* was selected to identify the differential expression at different developmental stages of leaves from the early-aging variety CCRI74 and the non-early-aging variety Liao4086. The results showed that the expression of *GhWRKY17* was higher in CCRI74 than in Liao4086 (Fig 5). As the leaves developed, the expression level of *GhWRKY17* in CCRI74 declined gradually, but little change was observed in Liao4086 (Fig 5). Therefore, the three genes may participate in the initiation of leaf senescence, and *GhWRKY17* plays an important role in leaf senescence.

Isolation and sequence analysis of *GhWRKY17*

We successfully isolated the *GhWRKY17* gene from upland cotton. Sequence analysis revealed that the complete open reading frame (ORF) of *GhWRKY17* was 942 bp in length. A predicted protein with 313 amino acids was encoded by the ORF with a predicted isoelectric point (pI) of 5.08 and molecular weight (Mw) of 77.58 kDa. Multiple sequence alignment analysis of the
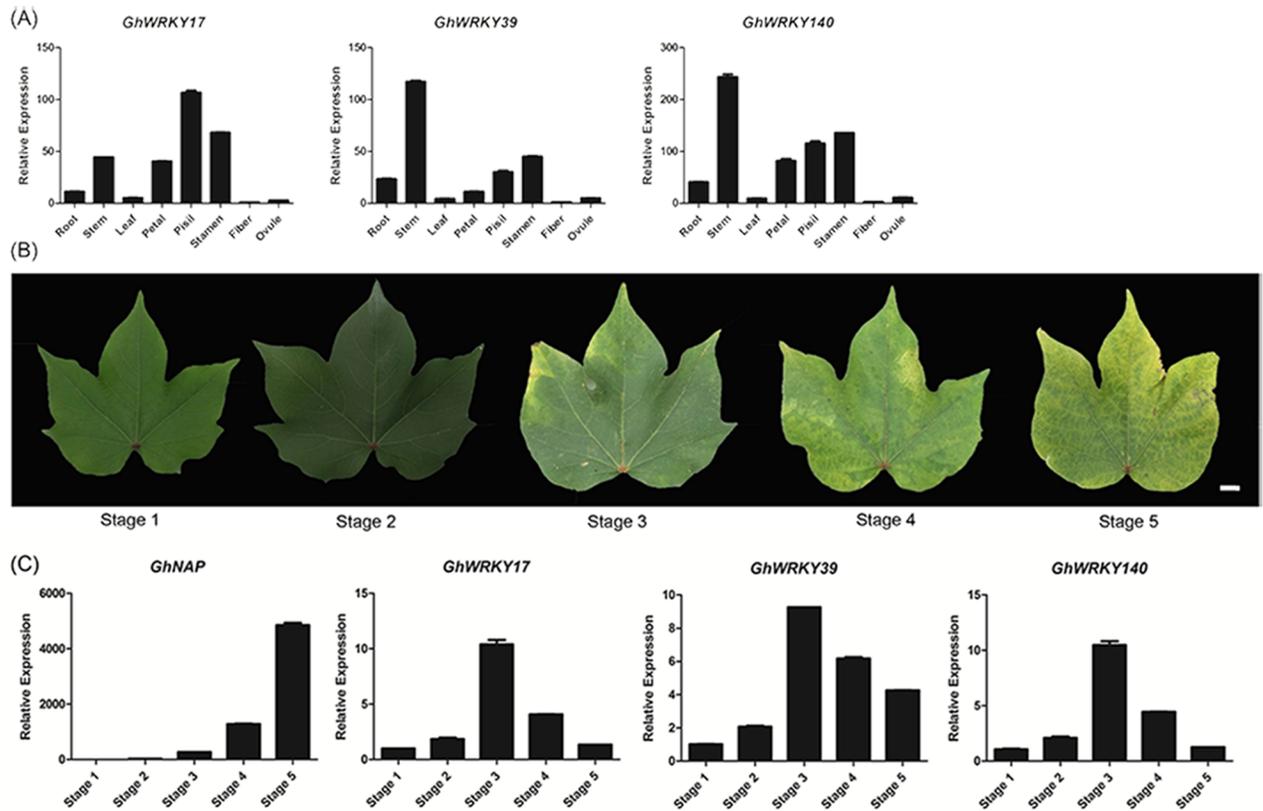
Fig 4. Tissue-specific and senescence expression patterns of *GhWRKY17*, *GhWRKY39* and *GhWRKY140*. (A) Relative expression of *GhWRKY17/39/140* in root, stem, leaf, petal, pistil, stamen, fiber and ovule tissues. The root and stem were sampled from two-week-old seedlings. The leaf, petal, pistil and stamen were sampled at the full flower period. The fiber and ovule were sampled 10 days after anthesis. (B) Five different senescence degrees of true leaves in CCRI74. Stage 1, an expanded new leaf; Stage 2, a mature but non-senescent leaf; Stage 3, a leaf with 25% senescence area; Stage 4, a leaf with 50% senescence area; and Stage 5, a leaf with at least 75% senescence area. (C) Relative expression of *GhNAP* and *GhWRKY17/39/140* in cotton. When qRT-PCR was performed, *GhActin* was used as a reference gene. The data are presented as the means±standard error.

GhWRKY17 protein with its homologs *AtWRKY60*, *AtWRKY40*, *OsWRKY71* and *NtWRKY40* revealed that the deduced protein has a nuclear localization signal (NLS) in the N-terminus together with a single WRKY domain in the C-terminus, which contains a conserved WRKYGQK core sequence and a C2H2 zinc finger-like motif (Fig 6A).

To understand the evolutionary relationship, a phylogenetic tree was constructed using the protein sequences of *GhWRKY17* and group IIa *WRKY* genes from other species such as *Arabidopsis* [58] and rice [59]. In the phylogenetic tree, *GhWRKY17* was mainly clustered with *Arabidopsis* genes, and *GhWRKY17* exhibited the highest homology with *AtWRKY18* and *AtWRKY60* (Fig 6B). Therefore, the evolutionary analysis of *GhWRKY17* can provide a reference for its functional study.

Promoter analysis of *GhWRKY17*

To determine whether *GhWRKY17* is induced by stress and to elucidate the underlying biological mechanism of the gene, a 1500 bp fragment from the upstream region was cloned. *Cis*-acting elements in the promoter region of *GhWRKY17* were predicted using the online *cis*-element prediction software PlantCARE. Many environmental response elements for abiotic stress (anaerobic, heat, drought, and defense elements), hormone stress (ABA, MeJA and gibberellic acid (GA)), light, metabolism and plant development were identified (Table 3). In
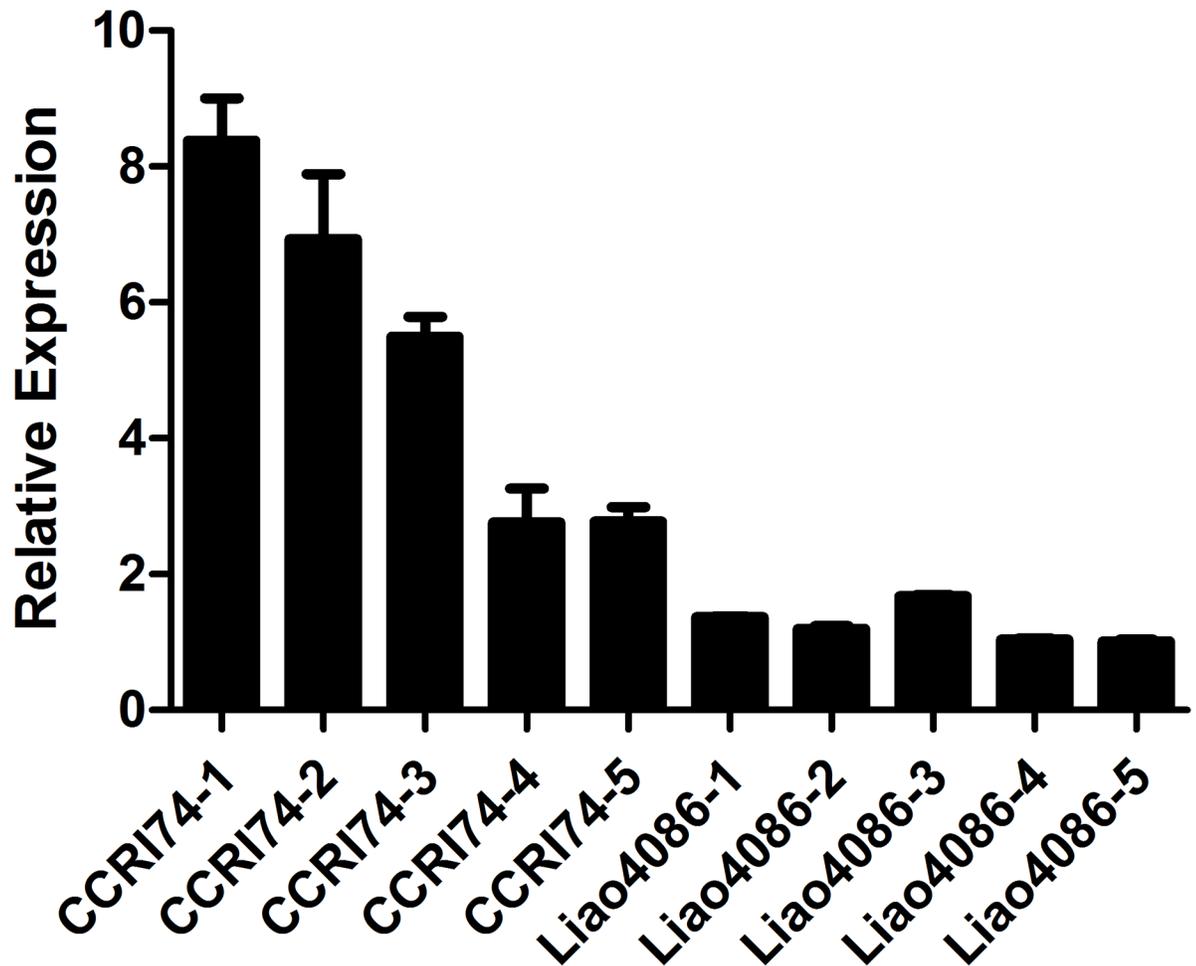
Fig 5. Relative expression of *GhWRKY17* in the early-aging variety CCRI74 and the non-early-aging variety Liao4086 in true leaves at five developmental stages. The fourth leaves from the top of CCRI74 and Liao4086 plants were marked every 10 days from the 80th day after sowing. CCRI 74–1, -2, -3, -4, -5 and Liao 4086–1, -2, -3, -4, -5 represent the true leaves collected 10 days, 20 days, 30 days, 40 days and 50 days after the leaves were marked. The data are presented as the means±standard error. *GhActin* was used as a reference gene.

addition, *GhWRKY17* could be induced by various stresses, including MeJA, ABA, SA, ETH, PEG6000 and NaCl (S3 Fig). Moreover, ABA treatment of transgenic *Arabidopsis* resulted in significantly reduced root length compared with that in the WT (S4 Fig). Our results indicated that *GhWRKY17* might be involved in multiple signaling pathways in plant growth and development and in stress responses.

GhWRKY17 is located in the nucleus

The *GhWRKY17* protein was predicted to be located in the cell nucleus by the online software WoLF PSORT. To determine the subcellular distribution of the *GhWRKY17* protein, the ORF without its termination codon was fused to a GFP gene driven by the *CaMV35S* promoter (Fig 7A). The control *35S-GFP* and experimental *35S-GhWRKY17:GFP* plasmids were transformed into cultivated onion epidermal cells using particle bombardment. Microscopic observation showed that green fluorescence in the control was distributed in the nucleus and cytoplasm, but the fusion protein was localized in the nucleus (Fig 7B), indicating that the *GhWRKY17* protein is a nucleoprotein.
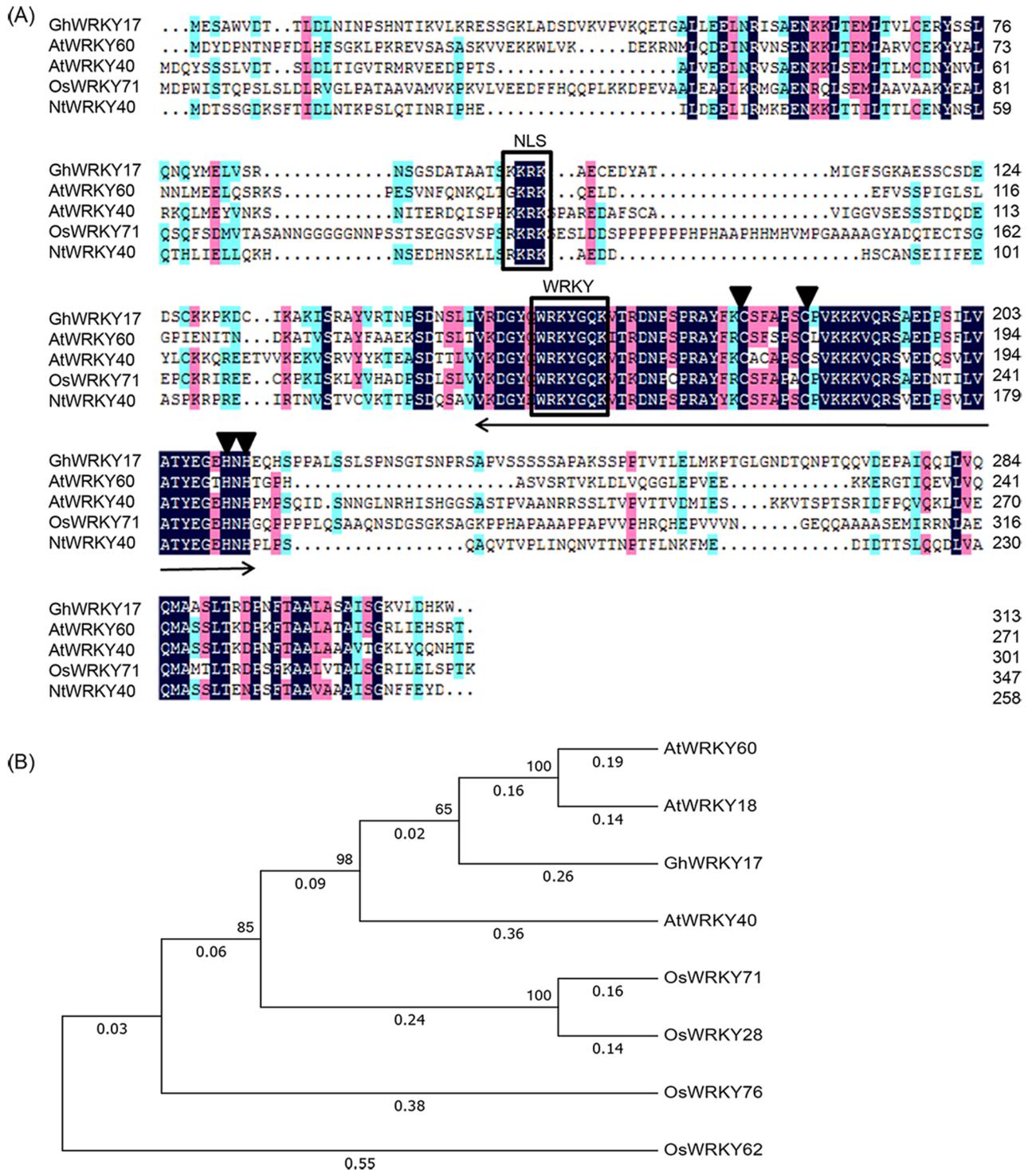
Fig 6. Multiple sequence alignment and phylogenetic analysis of *GhWRKY17*. (A) Multiple sequence alignment of *GhWRKY17* protein with its homologs from different species. The WRKY domain is indicated by a double-headed arrow. The putative NLS and WRKY core sequence are boxed. The zinc finger motif is marked with a downward-pointing triangle. (B) Phylogenetic tree of *GhWRKY17* protein with its homologs from *Arabidopsis thaliana* and *Oryza sativa*. The phylogenetic tree was constructed based on the protein sequences using the MEGA 7 program. The neighbor-joining method was used, and bootstrap analysis was performed with 1000 replications. At, *Arabidopsis thaliana*; Gh, *Gossypium hirsutum*; Nt, *Nicotiana tabacum*; Os, *Oryza sativa*.

Table 3. The predicted *cis*-acting elements in the promoter of *GhWRKY17*.

Cis-element	Position	Sequence (5'-3')	Function
Stress-responsive elements			
ABRE	-927(+),-930(-),-928(+),-931(+)	GGACACGTGGC	Cis-acting element involved in ABA response
ARE	-1137(-),-1431(-),-1254(-)	TGGTTT	Cis-acting regulatory element essential for the anaerobic induction
CGTCA motif	-264(+)	CGTCA	Cis-acting regulatory element involved in MeJA response
HSE	-605(+)	AAAAAATTTC	Cis-acting element involved in heat stress response
MBS	-621(+)	TAACTG	MYB binding site involved in drought-inducibility
P-box	-1440(+)	CCTTTTG	Gibberellin-responsive element
TC-rich repeats	-227(-)	ATTTTCTTCA	Cis-acting element involved in defense and stress response
TGACG motif	-264(-)	TGACG	Cis-acting regulatory element involved in MeJA response
Light-responsive elements			
3-AF1 binding site	-35(+)	TAAGAGAGGAA	Light-responsive element
AT1 motif	-543(+)	AATTATTTTTATT	Part of a light-responsive module
ATCT motif	-417(+)	AATCTAATCC	Part of a conserved DNA module involved in light response
Box 4	-62(+),-628(+),-540(+),-877(-)	ATTAAT	Part of a conserved DNA module involved in light response
G-Box	-45(-),-930(-),-921(+),-940(-)	CACGT (T/G)	Cis-acting regulatory element involved in light response
GATA motif	-38(+),-106(+)	GATAGGA	Part of a light-responsive element
I-box	-1402(-)	AAGATAAGGCT	Part of a light-responsive element
L-box	-1032(-),-1378(-),-1034(-)	TCTCACCTACCAA	Part of a light-responsive element
LAMP-element	-5(-),-1251(+)	CCAAAACCA	Part of a light-responsive element
Box I	-1443(-)	TTTCAA	Light-responsive element
MNF1	-925(-)	GTGCCC (A/T) (A/T)	Light-responsive element
chs-Unit 1 m1	-1030(-)	ACCTACCACAC	Part of a light-responsive element
Metabolism and development-related elements			
O2-site	-254(+),-390(-)	GATGA (C/T) ATGG	Cis-acting regulatory element involved in zein metabolism regulation
Skn-1 motif	-1153(+),-255(-)	GTCAT	Cis-acting regulatory element required for endosperm expression

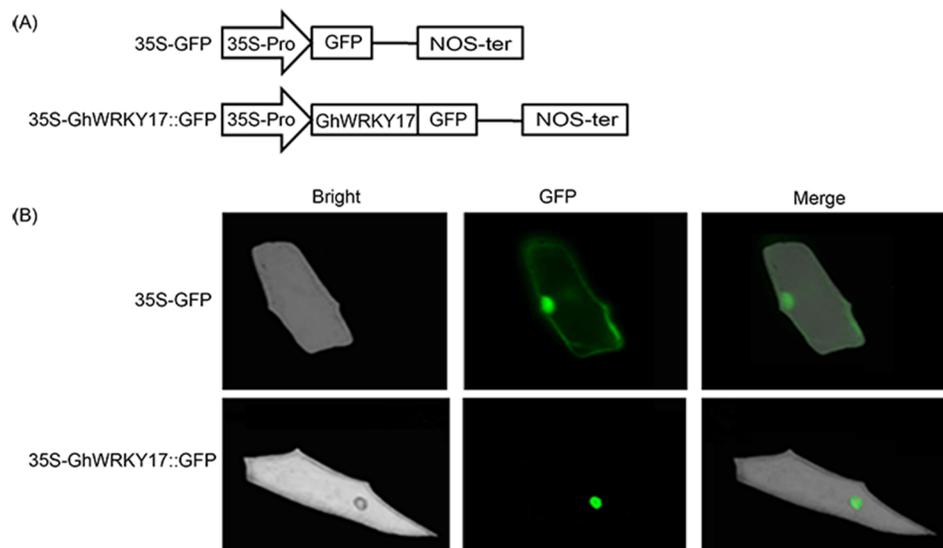https://doi.org/10.1371/journal.pone.0191681.t003

Fig 7. Nuclear localization of *GhWRKY17* in onion epidermal cells. (A) Plasmid sketch of the *35S-GFP* empty vector and *35S-GhWRKY17::GFP* fusion construct. *GhWRKY17* was fused to the N-terminus of GFP driven by the *CaMV 35S* promoter. (B) Transient expression of both *35S-GFP* and *35S-GhWRKY17::GFP* fusion proteins in onion epidermal cells in bright field (Bright), dim field (GFP), and overlapped field (Merge).

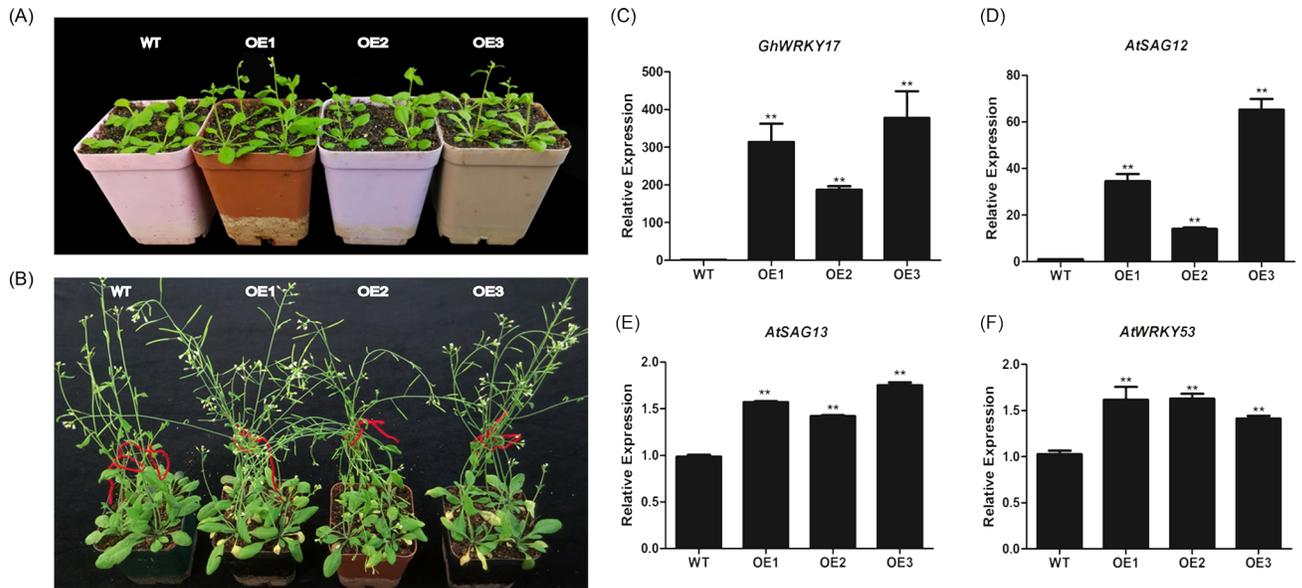https://doi.org/10.1371/journal.pone.0191681.g007

Fig 8. Over-expression of *GhWRKY17* can promote leaf senescence in *Arabidopsis thaliana*. (A) Flowering phenotypes of WT and OE1, OE2 and OE3 transgenic lines. (B) Senescent phenotypes of WT and OE1, OE2 and OE3 transgenic lines. (C) Confirmation of transgenic lines by qRT-PCR. (D-F) Relative expression of the senescence-associated genes *AtSAG12*, *AtSAG13* and *AtWRKY53*. OE1, OE2 and OE3 represent three transgenic lines. When qRT-PCR was performed, *AtUBQ10* was used as a reference gene. The data are presented as the means±standard error. Values significantly different from WT at the 0.01 confidence level.

Over-expression of *GhWRKY17* results in precocious leaf senescence

To assess the function of *GhWRKY17* in the model plant *Arabidopsis*, an expression vector was constructed and transformed into *Arabidopsis* using the floral dipping method [54]. The three transgenic lines were confirmed by qRT-PCR and used to observe the natural growth phenotype (Fig 8A–8C). The transgenic lines showed an early flowering phenotype, while the WT did not (Fig 8A). When the transgenic lines showed an aging phenotype and turned yellow, the WT was still green (Fig 8B). Furthermore, the expression levels of the three senescence marker genes *AtSAG12*, *AtSAG13*, and *AtWRKY53* were significantly higher than in the WT (Fig 8D–8F). Therefore, we concluded that the over-expression of the *GhWRKY17* gene in *Arabidopsis* could result in early senescence.

Discussion

WRKY TFs are widely involved in the processes of defense, trichome development, plant growth and development and leaf senescence. In this study, we comprehensively analyzed the group IIa WRKY genes in *G. hirsutum* using bioinformatics and qRT-PCR analysis. We particularly explored the roles of group IIa WRKY genes in leaf senescence. Thereafter, *GhWRKY17*, one of the group IIa WRKY genes that are differentially expressed in leaf senescence, was isolated and extensively studied. In the functional analysis, the over-expression of *GhWRKY17* in *Arabidopsis* elevated the expression of SAGs and promoted leaf senescence. Our findings extend our knowledge of the functional roles of group IIa WRKY genes in cotton.

The publication of cotton genome sequences for the D genome [60], A genome [61] and AD genome [33,34] provides a basis for the identification and analysis of the group IIa WRKY gene subfamily. Previous studies indentified 7 group IIa WRKY genes in *G. raimondii*, 7 in *G. arboreum* and 6 in *G. hirsutum* [36,37,5]. Here, we identified a total of 15 predicted group IIa *GhWRKY* genes in upland cotton. Among these genes, six (*GhWRKY17*, *18*, *20*, *44*, *49* and

155) were previously identified by Dou *et al.* [5], and the remaining nine genes (*GhWRKY39,
50, 140, 141, 144, 156, 163, 168* and *169*) are newly predicted. *G. hirsutum* was produced by the
interspecific hybridization and polyploidization of *G. raimondii* and *G. arboretum* [62]. Poly-
ploidization results in whole-genome duplication, and the whole-genome sequencing of cotton
revealed massive large-scale gene duplications [63,64,34]. Usually, gene duplication is defined
by the length of the alignment sequence covering at least 80% of the longest gene and the si-
milarity of the aligned regions exceeding 70% [56]. The *AtWRKY* gene family was shown to
exhibit a certain amount of tandem and segmental duplications [65]. Our results showed 7 para-
logs, but only 6 paralogs (*GhWRKY17/140, GhWRKY18/141, GhWRKY20/144, GhWRKY39/155,
GhWRKY44/163* and *GhWRKY49/168*) appeared to have undergone a high gene duplication
event. The 6 paralogs were considered as 6 gene pairs. In these 6 gene pairs, one gene was from At
and the other from Dt, and they were located on different chromosomes, which was confirmed to
represent fragment duplication [66]. The two genes in each gene pair possessed similar and con-
served intron-exon and motif structures, indicating their close evolutionary relationship. Interest-
ingly, although *GhWRKY50* and *GhWRKY169* were clustered together, *GhWRKY50* covered only
the front part of *GhWRKY169* and shared a similar structure in the overlap region, which might
be explained by the insertion of a long terminal repeat and expansion [61].

The Ka/Ks ratio is as a fairly good indicator for the identification of selective pressure acting
on a set of homologous protein-coding genes at the sequence level. As a method of estimating
the gene diversity caused by duplication, Ka/Ks>1 implies positive selection, Ka/Ks = 1 implies
neutral selection and Ka/Ks<1 implies purifying selection. These values indicate the purifying
selection, neutral mutations and beneficial mutations during the evolutionary process [67–69].
The Ka/Ks values of *GhWRKY17/140, GhWRKY18/141, GhWRKY44/163* and *GhWRKY49/168*
were lower than 1, indicating that these 4 gene pairs were under purifying selection and that
these genes tended to eliminate deleterious mutations during evolution [70]. The Ka/Ks of
GhWRKY20/144 was greater than 1, suggesting that its evolution at the protein level was accel-
erating under positive selection.

The publication of a large amount of microarray expression data and expression profiling
data provided a basis for the analysis of gene functions. The expression patterns of group IIa
WRKY members were preliminarily identified using the expression atlas data for different tis-
sues and different senescence stages [5,34,49]. One gene, *GhWRKY156*, showed no detectable
expression in the tissues or different senescent leaves when screening the expression data. The
expression level of this gene may be particularly low, or the gene might be expressed only
under specific conditions [9]. The expression analysis revealed that *GhWRKY49, GhWRKY168*
and *GhWRKY169* were highly expressed in the torus, suggesting a potential role in floral devel-
opment. *GhWRKY17, GhWRKY39* and *GhWRKY140* showed relatively high expression levels
in both different tissues and during the leaf senescence process. The qRT-PCR results showed
that *GhWRKY17, GhWRKY39,* and *GhWRKY140* exhibited high expression levels in the flower
organs and stem and were also highly expressed during the initial stage of leaf senescence. The
WRKY genes are highly expressed in plant organs, indicating that they play an important role
in plant growth and development [71]. The group IIa *GhWRKYs* are aging-related genes
according to the expression profile analysis [5]. Therefore, we speculated that these three
highly expressed genes *GhWRKY17, GhWRKY39* and *GhWRKY140* played a significant role in
regulating cotton development and leaf senescence. However, more research is needed to clar-
ify the specific functions of these genes.

To gain a greater understanding of the functions of group IIa *GhWRKY* TFs in cotton, the
GhWRKY17 gene was isolated from upland cotton. Several putative stress-related *cis*-regulatory
elements were found in the promoter region of *GhWRKY17*. The defense-regulatory elements in
the promoter of *CaWRKY1* indicated its function as a molecular player in the plant defense

machinery [72]. The stress-induced expression of *GhWRKY17* in cotton and the transgenic *Arabidopsis* phenotype under ABA treatment have proven its potential functions in stress signal pathways. A NLS sequence, "KKRK", was identified in the protein sequence of *GhWRKY17*. A subcellular localization experiment verified that *GhWRKY17* was localized in the nucleus, which was consistent with previous research on WRKY TFs in cotton [73]. Therefore, *GhWRKY17* may function as a nucleoprotein in signaling pathways and plant regulatory adaptation ability.

Leaf senescence is a complex process that can be affected by environmental factors and phytohormones and involves a decrease in chlorophyll content, macromolecule degradation, nutrient translocation and yield reduction [74–78]. In addition, leaf senescence is a process under the control of regulatory genes [74]. Previous studies have shown that group IIa *WRKY* TFs play an important role in leaf senescence. Based on microarray data and semi-quantitative RT-PCR analysis, *TaWRKY36* was found to be up-regulated during flag leaf senescence in wheat [79]. The expression level of *GhWRKY17* decreased gradually from the beginning of leaf senescence and revealed a higher expression level in the early-aging cotton cultivar CCRI74 than in the non-early-aging cultivar Liao4086, which suggested that *GhWRKY17* might be involved in leaf senescence. *GhWRKY17* was ectopically expressed in *Arabidopsis* to determine its response to senescence. *AtSAG12*, *AtSAG13* and *AtWRKY53* function as positive senescence regulators and have been used as molecular markers to the study the process of leaf senescence in *Arabidopsis* [80,81]. The over-expression of *GhWRKY17* in *Arabidopsis* resulted in a premature senescence phenotype, as confirmed by the higher expression levels of *AtSAG12*, *AtSAG13* and *AtWRKY53* than in the WT. As shown in the phylogenetic tree, *GhWRKY17* shared higher similarity with *AtWRKY18/40/60* than with *OsWRKY28/62/71/76* and shared the highest homology with *AtWRKY18*, suggesting a function similar to that of *AtWRKY18*. The over-expression of *AtWRKY18* in *Arabidopsis* led to delayed senescence [20]. As we can see, *GhWRKY17* and *AtWRKY18* have counteractive effects on the regulation of leaf senescence. Amino acid sequence alignment showed that only 46% similarity between *GhWRKY17* and *AtWRKY18* at the protein level, which might partially explain the difference in function between *GhWRKY17* and *AtWRKY18*. Therefore, understanding the biological function of group IIa *GhWRKY* genes can enrich our knowledge regarding the functions of WRKY genes in crops. Moreover, our study offers guiding significance for future experimental work.

Supporting information

S1 Fig. Chromosomal location of *GhWRKY* genes in the *Gossypium hirsutum* genome. (PDF)

S2 Fig. Phylogenetic tree of WRKY proteins of *Gossypium hirsutum* and *Arabidopsis thaliana*. The protein sequences of all *GhWRKYs* and *AtWRKYs* were aligned using Clustal W. The phylogenetic tree was constructed based on the protein sequences using the MEGA 7 program. The maximum likelihood method was used, and bootstrap analysis was performed with 1000 replications. (PDF)

S3 Fig. Expression levels of *GhWRKY17* under various stress conditions. Ten-day-old healthy and uniform seedlings were irrigated with 15% PEG6000 (A), 200 mM NaCl (B), and sprayed with 100 μM MeJA (C), 200 μM ABA (D), 2 mM SA (E) and 0.5 mM ETH (F). The total RNA was extracted from the samples at 0 h, 2 h, 4 h, 8 h and 12 h after stress treatments. *GhActin* was used as an internal reference. The data are presented as the means±standard error. The bars represent the standard error. (PDF)

S4 Fig. Over-expression of *GhWRKY17* led to susceptibility to ABA treatment in *Arabidopsis*. Three-day-old seedlings grown on 1/2 MS medium were transferred to new MS medium containing 10 μM ABA for seven days. (A) Phenotypic characteristics of WT and transgenic plants under 10 μM ABA treatment for seven days. (B) Root length of seedlings grown on MS medium containing 10 μM ABA for seven days. The data are presented as the means±standard error. The bars represent standard error. Values significantly different from WT at the 0.01 confidence level.
(PDF)

S1 Table. Primers used in this study.
(DOCX)

S2 Table. General information on WRKY gene family in cotton. a: We named all the *GhWRKYs* based on their gene IDs in the genome sequence database. b: Subgroups were divided according to the results of evolutionary analysis. The *GhWRKY* proteins that clustered with *AtWRKY* proteins were considered the same subfamily. c: WRKY genes identified by Cai *et al.* [37], using the sequence information of Paterson *et al.* [82]. d: WRKY genes identified by Cai *et al.* [37], using the sequence information of Wang *et al.* [60]. e: WRKY genes identified by Dou *et al.* [5]. "—" represents no results.
(DOCX)

S3 Table. The number of different subfamilies of the *GhWRKY* gene.
(DOCX)

Acknowledgments

Author Contributions

Data curation: Hantao Wang.

Funding acquisition: Shuxun Yu.

Project administration: Shuxun Yu.

Software: Libei Li.

Supervision: Shuxun Yu.

Writing – original draft: Lijiao Gu.

Writing – review & editing: Lijiao Gu, Hengling Wei, Junji Su, Yaning Guo.

References

1. Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, et al. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science. 2000; 290: 2105–2110. PMID: 11118137

2. Eulgem T, Rushton PJ, Robatzek S, Somssich IE. The WRKY superfamily of plant transcription factors. Trends In Plant Science. 2000; 5: 199–206. PMID: 10785665

3. Ishiguro S, Nakamura K. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and beta-amylase from sweet potato. Mol Gen Genet. 1994; 244: 563–571. PMID: 7969025

4. Wang Q, Wang M, Zhang X, Hao B, Kaushik SK, Pan Y. WRKY gene family evolution in Arabidopsis thaliana. Genetica. 2011; 139: 973–983. https://doi.org/10.1007/s10709-011-9599-4 PMID: 21805321

5. Dou L, Zhang X, Pang C, Song M, Wei H, Fan S, et al. Genome-wide analysis of the WRKY gene family in cotton. Mol Genet Genomics. 2014; 289: 1103–1121. https://doi.org/10.1007/s00438-014-0872-y PMID: 24942461

6. Ramamoorthy R, Jiang SY, Kumar N, Venkatesh PN, Ramachandran S. A comprehensive transcriptional profiling of the WRKY gene family in rice under various abiotic and phytohormone treatments. Plant Cell Physiol. 2008; 49: 865–879. https://doi.org/10.1093/pcp/pcn061 PMID: 18413358

7. Zou Z, Yang L, Wang D, Huang Q, Mo Y, Xie G. Gene Structures, Evolution and Transcriptional Profiling of the WRKY Gene Family in Castor Bean (Ricinus communis L.). PLoS One. 2016; 11: e0148243. https://doi.org/10.1371/journal.pone.0148243 PMID: 26849139

8. Wei Y, Shi H, Xia Z, Tie W, Ding Z, Yan Y, et al. Genome-Wide Identification and Expression Analysis of the WRKY Gene Family in Cassava. Front Plant Sci. 2016; 7: 25. https://doi.org/10.3389/fpls.2016.00025 PMID: 26904033

9. Ling J, Jiang W, Zhang Y, Yu H, Mao Z, Gu X, et al. Genome-wide analysis of WRKY gene family in Cucumis sativus. BMC Genomics. 2011; 12: 471. https://doi.org/10.1186/1471-2164-12-471 PMID: 21955985

10. Xie Z, Zhang ZL, Zou XL, Huang J, Ruas P, Thompson D, et al. Annotations and functional analyses of the rice WRKY gene superfamily reveal positive and negative regulators of abscisic acid signaling in aleurone cells. Plant Physiology. 2005; 137: 176–189. https://doi.org/10.1104/pp.104.054312 PMID: 15618416

11. Rinerson CI, Rabara RC, Tripathi P, Shen QXJ, Rushton PJ. The evolution of WRKY transcription factors. Bmc Plant Biology. 2015; 15:

12. Rushton DL, Tripathi P, Rabara RC, Lin J, Ringler P, Boken AK, et al. WRKY transcription factors: key components in abscisic acid signalling. Plant Biotechnology Journal. 2012; 10: 2–11. https://doi.org/10.1111/j.1467-7652.2011.00634.x PMID: 21696534

13. Rushton PJ, Somssich IE, Ringler P, Shen QXJ. WRKY transcription factors. Trends In Plant Science. 2010; 15: 247–258. https://doi.org/10.1016/j.tplants.2010.02.006 PMID: 20304701

14. Rinerson CI, Rabara RC, Tripathi P, Shen QJ, Rushton PJ. The evolution of WRKY transcription factors. BMC Plant Biol. 2015; 15: 66. https://doi.org/10.1186/s12870-015-0456-y PMID: 25849216

15. He H, Dong Q, Shao Y, Jiang H, Zhu S, Cheng B, et al. Genome-wide survey and characterization of the WRKY gene family in Populus trichocarpa. Plant Cell Rep. 2012; 31: 1199–1217. https://doi.org/10.1007/s00299-012-1241-0 PMID: 22371255

16. Xie Z, Zhang ZL, Hanzlik S, Cook E, Shen QJ. Salicylic acid inhibits gibberellin-induced alpha-amylase expression and seed germination via a pathway involving an abscisic-acid-inducible WRKY gene. Plant Mol Biol. 2007; 64: 293–303. https://doi.org/10.1007/s11103-007-9152-0 PMID: 17390108

17. Zhang CQ, Xu Y, Lu Y, Yu HX, Gu MH, Liu QQ. The WRKY transcription factor OsWRKY78 regulates stem elongation and seed development in rice. Planta. 2011; 234: 541–554. https://doi.org/10.1007/s00425-011-1423-y PMID: 21547461

18. Rushton PJ, Somssich IE, Ringler P, Shen QJ. WRKY transcription factors. Trends Plant Sci. 2010; 15: 247–258. https://doi.org/10.1016/j.tplants.2010.02.006 PMID: 20304701

19. Xu X, Chen C, Fan B, Chen Z. Physical and functional interactions between pathogen-induced Arabidopsis WRKY18, WRKY40, and WRKY60 transcription factors. Plant Cell. 2006; 18: 1310–1326. https://doi.org/10.1105/tpc.105.037523 PMID: 16603654

20. Potschin M, Schlienger S, Bieker S, Zentgraf U. Senescence Networking: WRKY18 is an Upstream Regulator, a Downstream Target Gene, and a Protein Interaction Partner of WRKY53. Journal of Plant Growth Regulation. 2013; 33: 106–118.

21. Yan L, Liu ZQ, Xu YH, Lu K, Wang XF, Zhang DP. Auto- and Cross-repression of Three Arabidopsis WRKY Transcription Factors WRKY18, WRKY40, and WRKY60 Negatively Involved in ABA Signaling. Journal Of Plant Growth Regulation. 2013; 32: 399–416.

22. Chen H, Lai Z, Shi J, Xiao Y, Chen Z, Xu X. Roles of arabidopsis WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. BMC Plant Biol. 2010; 10: 281. https://doi.org/10.1186/1471-2229-10-281 PMID: 21167067

23. Liu ZQ, Yan L, Wu Z, Mei C, Lu K, Yu YT, et al. Cooperation of three WRKY-domain transcription factors WRKY18, WRKY40, and WRKY60 in repressing two ABA-responsive genes ABI4 and ABI5 in Arabidopsis. Journal Of Experimental Botany. 2012; 63: 6371–6392. https://doi.org/10.1093/jxb/ers293 PMID: 23095997

24. Peng Y, Bartley LE, Canlas P, Ronald PC. OsWRKY IIa Transcription Factors Modulate Rice Innate Immunity. Rice. 2010; 3: 36–42. https://doi.org/10.1007/s12284-010-9039-6 PMID: 21961049

25. Chujo T, Miyamoto K, Shimogawa T, Shimizu T, Otake Y, Yokotani N, et al. OsWRKY28, a PAMP-responsive transrepressor, negatively regulates innate immune responses in rice against rice blast fungus. Plant Mol Biol. 2013; 82: 23–37. https://doi.org/10.1007/s11103-013-0032-5 PMID: 23462973

26. Liu X, Bai X, Wang X, Chu C. OsWRKY71, a rice transcription factor, is involved in rice defense response. J Plant Physiol. 2007; 164: 969–979. https://doi.org/10.1016/j.jplph.2006.07.006 PMID: 16919842

27. Peng Y, Bartley LE, Chen X, Dardick C, Chern M, Ruan R, et al. OsWRKY62 is a negative regulator of basal and Xa21-mediated defense against Xanthomonas oryzae pv. oryzae in rice. Mol Plant. 2008; 1: 446–458. https://doi.org/10.1093/mp/ssn024 PMID: 19825552

28. Liu J, Chen X, Liang X, Zhou X, Yang F, Liu J, et al. Alternative Splicing of Rice WRKY62 and WRKY76 Transcription Factor Genes in Pathogen Defense. Plant Physiol. 2016; 171: 1427–1442. https://doi.org/10.1104/pp.15.01921 PMID: 27208272

29. Karim A, Jiang Y, Guo L, Ling Z, Ye S, Duan Y, et al. Isolation and characterization of a subgroup IIa WRKY transcription factor PtrWRKY40 from Populus trichocarpa. Tree Physiol. 2015; 35: 1129–1139. https://doi.org/10.1093/treephys/tpv084 PMID: 26423133

30. Qin Z, Lv H, Zhu X, Meng C, Quan T, Wang M, et al. Ectopic expression of a wheat WRKY transcription factor gene TaWRKY71-1 results in hyponastic leaves in Arabidopsis thaliana. PLoS One. 2013; 8: e63033. https://doi.org/10.1371/journal.pone.0063033 PMID: 23671653

31. Wang KB, Wang ZW, Li FG, Ye WW, Wang JY, Song GL, et al. The draft genome of a diploid cotton Gossypium raimondii. Nature Genetics. 2012; 44: 1098-+. https://doi.org/10.1038/ng.2371 PMID: 22922876

32. Li FG, Fan GY, Wang KB, Sun FM, Yuan YL, Song GL, et al. Genome sequence of the cultivated cotton Gossypium arboreum. Nature Genetics. 2014; 46: 567–572. https://doi.org/10.1038/ng.2987 PMID: 24836287

33. Li F, Fan G, Lu C, Xiao G, Zou C, Kohel RJ, et al. Genome sequence of cultivated Upland cotton (Gossypium hirsutum TM-1) provides insights into genome evolution. Nat Biotechnol. 2015; 33: 524–530. https://doi.org/10.1038/nbt.3208 PMID: 25893780

34. Zhang T, Hu Y, Jiang W, Fang L, Guan X, Chen J, et al. Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement. Nat Biotechnol. 2015; 33: 531–537. https://doi.org/10.1038/nbt.3207 PMID: 25893781

35. Fan X, Guo Q, Xu P, Gong Y, Shu H, Yang Y, et al. Transcriptome-wide identification of salt-responsive members of the WRKY gene family in Gossypium aridum. PLoS One. 2015; 10: e0126148. https://doi.org/10.1371/journal.pone.0126148 PMID: 25951083

36. Ding M, Chen J, Jiang Y, Lin L, Cao Y, Wang M, et al. Genome-wide investigation and transcriptome analysis of the WRKY gene family in Gossypium. Mol Genet Genomics. 2015; 290: 151–171. https://doi.org/10.1007/s00438-014-0904-7 PMID: 25190108

37. Cai C, Niu E, Du H, Zhao L, Feng Y, Guo W. Genome-wide analysis of the WRKY transcription factor gene family in Gossypium raimondii and the expression of orthologs in cultivated tetraploid cotton. The Crop Journal. 2014; 2: 87–101.

38. Wang X, Yan Y, Li Y, Chu X, Wu C, Guo X. GhWRKY40, a multiple stress-responsive cotton WRKY gene, plays an important role in the wounding response and enhances susceptibility to ralstonia solanacearum infection in transgenic Nicotiana benthamiana. PLoS One. 2014; 9: e93577. https://doi.org/10.1371/journal.pone.0093577 PMID: 24747610

39. Yu SX, Song MZ, Fan SL, Wang W, Yuan RH. Biochemical genetics of short-season cotton cultivars that express early maturity without senescence. Journal Of Integrative Plant Biology. 2005; 47: 334–342.

40. Zhang TZ, Hu Y, Jiang WK, Fang L, Guan XY, Chen JD, et al. Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement. Nature Biotechnology. 2015; 33: 531–U252. https://doi.org/10.1038/nbt.3207 PMID: 25893781

41. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. Nucleic Acids Research. 2015; 43: D257–D260. https://doi.org/10.1093/nar/gku949 PMID: 25300481

42. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu SN, Chitsaz F, Geer LY, et al. CDD: NCBI's conserved domain database. Nucleic Acids Research. 2015; 43: D222–D226. https://doi.org/10.1093/nar/gku1221 PMID: 25414356

43. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology And Evolution. 2016; 33: 1870–1874. https://doi.org/10.1093/molbev/msw054 PMID: 27004904

44. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallographica Section D-Biological Crystallography. 1997; 53: 240–255.

45. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987; 4: 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454 PMID: 3447015

46. Voorrips RE. MapChart: Software for the graphical presentation of linkage maps and QTLs. Journal Of Heredity. 2002; 93: 77–78. PMID: 12011185

47. Hu B, Jin JP, Guo AY, Zhang H, Luo JC, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2015; 31: 1296–1297. https://doi.org/10.1093/bioinformatics/btu817 PMID: 25504850

48. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Research. 2006; 34: W609–W612. https://doi.org/10.1093/nar/gkl315 PMID: 16845082

49. Ma L, Wang Y, Yan G, Wei S, Zhou D, Kuang M, et al. Global analysis of the developmental dynamics of Gossypium hirsutum based on strand-specific transcriptome. Physiol Plant. 2016; 158: 106–121. https://doi.org/10.1111/ppl.12432 PMID: 26892265

50. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. Nature Protocols. 2008; 3: 1101–1108. PMID: 18546601

51. Kalet IJ, Jacky JP, Austin-Seymour MM, Hummel SM, Sullivan KJ, Unger JM. Prism: a new approach to radiotherapy planning software. Int J Radiat Oncol Biol Phys. 1996; 36: 451–461. PMID: 8892471

52. Kapitonov D, Bieberich E, Yu RK. Combinatorial PCR approach to homology-based cloning: cloning and expression of mouse and human GM3-synthase. Glycoconj J. 1999; 16: 337–350. PMID: 10619706

53. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research. 2002; 30: 325–327. PMID: 11752327

54. Narusaka M, Shiraishi T, Iwabuchi M, Narusaka Y. The floral inoculating protocol: a simplified Arabidopsis thaliana transformation method modified from floral dipping. Plant Biotechnology. 2010; 27: 349–351.

55. Rychlik W. OLIGO 7 primer analysis software. Methods Mol Biol. 2007; 402: 35–60. https://doi.org/10.1007/978-1-59745-528-2_2 PMID: 17951789

56. Gu ZL, Cavalcanti A, Chen FC, Bouman P, Li WH. Extent of gene duplication in the genomes of Drosophila, nematode, and yeast. Molecular Biology And Evolution. 2002; 19: 256–262. https://doi.org/10.1093/oxfordjournals.molbev.a004079 PMID: 11861885

57. Fan K, Bibi N, Gan S, Li F, Yuan S, Ni M, et al. A novel NAP member GhNAP is involved in leaf senescence in Gossypium hirsutum. J Exp Bot. 2015; 66: 4669–4682. https://doi.org/10.1093/jxb/erv240 PMID: 25991739

58. Eulgem T, Rushton PJ, Robatzek S, Somssich IE. The WRKY superfamily of plant transcription factors. Trends Plant Sci. 2000; 5: 199–206. PMID: 10785665

59. Ross CA, Liu Y, Shen QXJ. The WRKY gene family in rice (Oryza sativa). Journal Of Integrative Plant Biology. 2007; 49: 827–842.

60. Wang K, Wang Z, Li F, Ye W, Wang J, Song G, et al. The draft genome of a diploid cotton Gossypium raimondii. Nat Genet. 2012; 44: 1098–1103. https://doi.org/10.1038/ng.2371 PMID: 22922876

61. Li F, Fan G, Wang K, Sun F, Yuan Y, Song G, et al. Genome sequence of the cultivated cotton Gossypium arboreum. Nat Genet. 2014; 46: 567–572. https://doi.org/10.1038/ng.2987 PMID: 24836287

62. Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin DC, et al. Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. Nature. 2012; 492: 423-+. https://doi.org/10.1038/nature11798 PMID: 23257886

63. Soltis PS, Soltis DE. The Role of Hybridization in Plant Speciation. Annual Review Of Plant Biology. 2009; 60: 561–588. https://doi.org/10.1146/annurev.arplant.043008.092039 PMID: 19575590

64. Otto SP. The evolutionary consequences of polyploidy. Cell. 2007; 131: 452–462. https://doi.org/10.1016/j.cell.2007.10.022 PMID: 17981114

65. Cannon SB, Mitra A, Baumgarten A, Young ND, May G. The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. BMC Plant Biol. 2004; 4: 10. https://doi.org/10.1186/1471-2229-4-10 PMID: 15171794

66. Liu Y, Jiang HY, Chen WJ, Qian YX, Ma Q, Cheng BJ, et al. Genome-wide analysis of the auxin response factor (ARF) gene family in maize (Zea mays). Plant Growth Regulation. 2011; 63: 225–234.

67. Hurst LD. The Ka/Ks ratio: diagnosing the form of sequence evolution. Trends In Genetics. 2002; 18: 486–487. PMID: 12175810

68. Yang Z, Bielawski JP. Statistical methods for detecting molecular adaptation. Trends Ecol Evol. 2000; 15: 496–503. PMID: 11114436

69. Hurst LD. The Ka/Ks ratio: diagnosing the form of sequence evolution. Trends Genet. 2002; 18: 486. PMID: 12175810

70. Rehmeyer CJ, Li W, Kusaba M, Farman ML. The telomere-linked helicase (TLH) gene family in Magna-porthe oryzae: revised gene structure reveals a novel TLH-specific protein motif. Curr Genet. 2009; 55: 253–262. https://doi.org/10.1007/s00294-009-0240-3 PMID: 19360408

71. Ramamoorthy R, Jiang SY, Kumar N, Venkatesh PN, Ramachandran S. A comprehensive transcriptional profiling of the WRKY gene family in rice under various abiotic and phytohormone treatments. Plant And Cell Physiology. 2008; 49: 865–879. https://doi.org/10.1093/pcp/pcn061 PMID: 18413358

72. Petitot AS, Barsalobres-Cavallari C, Ramiro D, Albuquerque Freire E, Etienne H, Fernandez D. Promoter analysis of the WRKY transcription factors CaWRKY1a and CaWRKY1b homoeologous genes in coffee (Coffea arabica). Plant Cell Rep. 2013; 32: 1263–1276. https://doi.org/10.1007/s00299-013-1440-3 PMID: 23568411

73. Liu X, Song Y, Xing F, Wang N, Wen F, Zhu C. GhWRKY25, a group I WRKY gene from cotton, confers differential tolerance to abiotic and biotic stresses in transgenic Nicotiana benthamiana. Protoplasma. 2016; 253: 1265–1281. https://doi.org/10.1007/s00709-015-0885-3 PMID: 26410829

74. Lim PO, Kim HJ, Nam HG. Leaf senescence. Annu Rev Plant Biol. 2007; 58: 115–136. https://doi.org/10.1146/annurev.arplant.57.032905.105316 PMID: 17177638

75. Fischer AM. The Complex Regulation of Senescence. Critical Reviews in Plant Sciences. 2012; 31: 124–147.

76. Distelfeld A, Avni R, Fischer AM. Senescence, nutrient remobilization, and yield in wheat and barley. J Exp Bot. 2014; 65: 3783–3798. https://doi.org/10.1093/jxb/ert477 PMID: 24470467

77. Liu X, Li Z, Jiang Z, Zhao Y, Peng J, Jin J, et al. LSD: a leaf senescence database. Nucleic Acids Res. 2011; 39: D1103–1107. https://doi.org/10.1093/nar/gkq1169 PMID: 21097471

78. Uzelac B, Janosevic D, Simonovic A, Motyka V, Dobrev PI, Budimir S. Characterization of natural leaf senescence in tobacco (Nicotiana tabacum) plants grown in vitro. Protoplasma. 2016; 253: 259–275. https://doi.org/10.1007/s00709-015-0802-9 PMID: 25837009

79. Zhang H, Zhao M, Song Q, Zhao L, Wang G, Zhou C. Identification and function analyses of senescence-associated WRKYs in wheat. Biochem Biophys Res Commun. 2016; 474: 761–767. https://doi.org/10.1016/j.bbrc.2016.05.034 PMID: 27166153

80. Noh YS, Amasino RM. Identification of a promoter region responsible for the senescence-specific expression of SAG12. Plant Mol Biol. 1999; 41: 181–194. PMID: 10579486

81. Miao Y, Smykowski A, Zentgraf U. A novel upstream regulator of WRKY53 transcription during leaf senescence in Arabidopsis thaliana. Plant Biology. 2008; 10: 110–120. https://doi.org/10.1111/j.1438-8677.2008.00083.x PMID: 18721316

82. Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, et al. Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. Nature. 2012; 492: 423–427. https://doi.org/10.1038/nature11798 PMID: 23257886