Identification of the WRKY Group I Genes and the Functional Analysis of GhWRKY138 in Cotton Pigment Gland Development

Da Qu
Zhengzhou University
Yuanli Dai
Zhengzhou University
Shang Liu
CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute
Fanjia Peng
Hunan Cotton Science Institute
Limin Lv
CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute
Youping Zhang
CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute
Qiaolian Wang
CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute
Dongyun Zuo
CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute
Guoli Song
CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute
Hailiang Cheng (pser2010@163.com)
ICR: Cotton Research Institute

Research Article

Keywords: WRKY group I, GhMYC2-like, cotton pigment gland, quantitative RT–PCR, RNAi, virus-induced gene silencing

Posted Date: March 24th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1460184/v1

Abstract

Background

WRKY transcription factors have various functions in leaf senescence, plant maturation, adaptation to stress and regulation of secondary metabolism. However, the relationship between *WRKY* family and cotton pigment gland development remains unclear.

Result

In this study, 36 *WRKY* group I genes were identified in *G. hirsutum L*. Expression pattern analysis showed that *GhWRKY138* of *WRKY* group I had a large difference in expression between glanded and glandless plants. Virus-induced gene silencing experiments showed that the number of glands decreased significantly with gene silencing. At the same time, we found that the expression of *GhMYC2-like*, a key gene regulating pigment gland development, also decreased with the decrease of *GhWRKY138*.

Conclusion

GhWRKY138 may involved in the formation and development of cotton glands. This study explored the possible relationship between the *WRKY* transcription factor family and gland development, which further enriched the connotation of the biological function of *WRKY* genes. These studies provide new insights into the mechanism of gland development.

Background

Cotton is a very critical resource in the world, and it has a very pivotal strategic significance. It has been cultivated around the world for more than 7,000 years. Cotton can not only be used as an important raw material for textile industry, but also a nutrition resource. For every kilogram of cotton fiber produced, about 1.65kg of seeds will be produced. And cottonseed is not only rich in oil, but also contains 23% protein[1]. But common cotton varieties contain pigment glands, which contain a toxic yellow dimeric sesquiterpene compound called gossypol, monogastric animals and humans are very sensitive to this toxicity[2, 3]. Besides, it should not be overlooked that gossypol acts as a natural plant defender to reduce the damage of pests and diseases and external abiotic stresses[4–7]. According to the characteristics of the two-sidedness of gossypol, it is of great significance to improve the economic value of cotton to cultivate cotton varieties with high content of gossypol in plants and low content of gossypol in cottonseeds.

Pigment glands, also known as gossypol glands, are unique structure on the surface of Gossypium. Cotton glands consist of a cavity and secretory cell mass formed through programmed cell death (PCD) [8, 9]. Gossypol is stored in the mature pigment glands as main component[10]. Previous studies have

shown that the number of pigment glands and gossypol content are closely related. The synthetic pathway of gossypol is relatively complete, but the mechanism of gland formation is still unclear. Previous studies have identified 6 loci involved in cotton gland formation: Gl_1, Gl_2, Gl_3, Gl_4, Gl_5 and Gl_6. McMichael first discovered the recessive gene gl_1, which controls the trait that almost no glands in cotton bolls and stalks but a normal number of glands in cotton leaves[11]. Subsequently, gl_2gl_3 was reported to be a double recessive gene that controls the absence of glands in the whole plant[12, 13]. gl_4 and gl_5 were found to regulate the number and density of glands, but not the complete absence of glands[14]. gl_6 is a gene with a similar but weak function to gl_1[15]. Then, a whole glandless plant was isolated from "Giza45" (G. barbadense L.). The glandless phenotype was controlled by a dominant gene Gl_2^e, which is an allele of Gl_2[16, 17]. Since then, efforts were made to map and isolate the genes involved in gland formation. In 2016, Cheng mapped Gl_2^e to a 15KB region on chromosome A12, in which a transcription factor called *GhMYC2-like* was identified and the presence of an amino acid substitution in the conserved domain may be associated with glandless phenotype[18]. The function of the transcription factor was verified by VIGS and it was named as Gossypium Gland Formation (*GoPGF*)[19].

Gl_2 and Gl_3 genes are located on chromosomes A12 and D12, respectively. gl_2 and gl_3 were found to be derived from premature translation termination and monomer mutation of Gl_2 and Gl_3. Subsequently, three *CGF* genes were identified by transcriptome analysis of gland and glandless cotton embryos and function verification by virus-induced gene silencing (VIGS). Gene knockout mediated by CRISPR/ Cas9 further confirmed the role of *CGF2* and *CGF3*（synonymous with *GhMYC2-like/GoPGF*） in gland formation[20].

WRKY proteins are a family of transcription factors unique to plants. *WRKY* transcription factors generally contain one or two conserve *WRKY* domains, with about 60 amino acid residues including WRKYGQK sequence and a C2H2 or C2HC zinc finger motif[21, 22]. With the cloning of the first *WRKY* gene SPF1 in sweet potatoes[23], *WRKY* genes have been studied and reported in more and more species, including Arabidopsis thaliana[24], maize[25], wheat[26], tomato[27], cotton[28], rice[29], castorbean[30], cassava[31], cucumber[32], pineapple[33],etc. *WRKY* transcription factors play important roles in a variety of complex physiological and biochemical processes, especially in the field of plant resistance to abiotic and biological stresses[34]. *BcWRKY46* could reduce the sensitivity of tobacco to freezing stress, ABA stress, salt stress and dehydration stress[35]. *SbWRKY50* can be involved in plant salt response by regulating ion homeostasis[36]. *WRKY71* were found to be involved in ethylene mediated regulation of leaf senescence[37]. Recently, *WRKY2* and *WRKY10* were reported that involved in molecular mechanisms that regulate plant light signal transduction[38]. In addition, *WRKY* is involved in the accumulation of secondary metabolites in a variety of plants. Maize terpenoid phytoalexins (MTPs) biosynthesis is regulated by *ZmWRKY79* which is highly correlated with expression of MTPs[39]. *GaWRKY1* was found to be involved in the regulation of gossypol biosynthesis in cotton[40]. All these indicate that *WRKY* is widely involved in a variety of plant biological processes, and the study of *WRKY* family has special production significance and application value.

The mechanism of gland development is very important for the application of cotton. *GhMYC2-like* has been studied as a star gene in gland development. As one of the largest transcription factor families in plants, *WRKY* plays an important role in various plant stress resistance processes. However, the study of the WRKY family in gland development is not clear and it is unknown whether *WRKY* genes participate in pigment gland development regulation.

In this study, we identified 36 *WRKY* group I genes and screened out one gene- *GhWRKY138* from the transcriptome. *GhWRKY138* expression was significantly different in glanded and glandless strains. We performed a virus-induced gene silencing assay on *GhWRKY138* to obtain a phenotype with reduced number of pigment glands in cotton. Further analysis showed that the expression of *GhMYC2-like* decreased with the decrease of the gene expression level.

This study provides important information for improving the network between the *WRKY* family and gland development, and also extends the key role of *WRKY* in stress tolerance due to the important role of the gland in cotton stress tolerance.

Results

Result 1

Identification and characteristic information of WRKY **group I genes in** G. hirsutum L.

Previous studies on *WRKY* transcription factors have shown that there are 239 *WRKY* family members in upland cotton, of which 36 are group I family members[41]. This conclusion was verified by bioinformatics analysis and analyzed for 36 members of *WRKY* Group I. The characteristic information of 36 genes of the group I were listed in Supplementary Table S1, including genomic DNA length (bp), coding sequence CDS length (bp), GC content (%), coding protein length (aa), exon number, gene start and end sites on chromosomes, isoelectric point (PI), protein molecular weight (MW). The genomic DNA length of these *WRKY* group I genes ranged from 1,518 bp (*GhWRKY211*) to 4,926 bp (*GhWRKY232*). The range of CDS coding sequences is from 1215bp (*GhWRKY211*) to 2292 bp (*GhWRKY232*). The GC content of CDS sequences varies from 41.8% (*GhWRKY54*) to 50.1% (*GhWRKY229*). The number of exons in candidate genes varies from 4 to 6. In addition, the isoelectric point of them was between 5.671 and 9.308, and the molecular weight ranged from 44.261 to 82.043. *GhWRKY138* was located in the region from 48,694,888 to 48696787 of the D04 chromosome. The genomic DNA length was 1,900bp, which had 5 exons and 43.4% GC content. Its predicted molecular weight and isoelectric point are 55.929kDa and 6.706, respectively (Supplementary Table S1).

Result 2

Phylogenetic and structure analysis of WRKY group **I protieins in** G. hirsutum L.

In order to further analyze the evolutionary relationship of *WRKY* group I genes, we constructed a phylogenetic tree of 36 group I genes in *G. hirsutum L.* that have been identified (Fig. 1A). The

phylogenetic tree is divided into three clades. There are 28 putative group I genes in the first clade, and 4 genes in the second and third clade respectively. Among them, the putative subgenome A genes and its corresponding homologous subgenome D genes can be well clustered in the same fine clade.

According to the results of sequence alignment, it can be found that similar to *WRKY* genes of other groups, the *WRKY* group I has one or two highly conserved structure (WRKYGQK) and C_2H_2 zinc finger motif (Fig. 1B). Of these, the first clade and the second clade have two *WRKY* domains (WRKYGQK), but the third clade has only one domain. We analyzed the gene structure and conserved domain of the genes in group I (Fig. 1C). The members within each clade have similar intron and exon structure and gene length. They have four to seven exons and three to six introns. *GhWRKY138* belongs to the first clade, it contains two WRKY conserved domains and five exons.

Result 3

Expression pattern analysis of WRKY **group I genes in glanded and glandless cotton.**

To explore the role of the *WRKY* group I genes in gland development, the transcriptome of glanded and glandless cotton were used to investigate the gene expression. The expression heatmap of 36 genes was drawn (Fig. 2A). Based on transcriptome data, We found significant differences in the expression levels of some genes in glanded and glandless materials. These genes were selected for real-time fluorescence quantification PCR (qRT-PCR) to further verify the accuracy of transcriptome (Fig. 2B). The expression levels of these genes were consistent with transcriptome data.

We screened the differentially expressed genes and obtained a gene named *GhWRKY138*. As shown in Fig. 2B, *GhWRKY138* was highly expressed in L7, but low in Dgl-L7. We suggest that this interesting phenomenon may be related to gossypol synthesis and gland development. As we know, transcription factors usually function in the nucleus, and here we validated the subcellular localization of *GhWRKY138*. The fluorescence was only visible in the nucleus of tobacco mesophyll cells (Fig. 2C). This result is consistent with the properties of transcription factors.

Result 4

GhWRKY138 **-silencing results in gland number reduction**

To further assess the connection between *GhWRKY138* and cotton glands, virus induced gene silencing (VIGS) which have a silencing effect on the target gene was performed. A 245 bp fragment was selected and designed with low similarity between *GhWRKY138* and the other group I gene sequences. Furthermore, the *PDS* gene was silenced with the *GhWRKY138* as a specific phenotypic marker for VIGS. The transcripts of *GhWRKY138* in the *GhWRKY138* -silenced leaves were significantly reduced compared to the untreated leaves, manifesting that *GhWRKY138* was effectively silenced in VIGS plants. As shown in Fig. 3A, silencing *GhWRKY138* also reduced the number of glands in leaves greatly. It is evident that the gland has decreased as the expression level has decreased (Fig. 3B). This result suggests that there is

a link between the gene and the number of glands, and the specific mechanism needs to be further verified.

Result 5

The expression of GhMYC2-like **was significantly decreased in the** GhWRKY138 **-silenced plants.**

In recent years, *GhMYC2-like* has become the most important gene in the study of glandless cotton. Since *GhMYC2-like* plays a key role in gland development, it is of great significance to explore the relationship between *GhWRKY138* and *GhMYC2-like*. We found an interesting phenomenon in *GhWRKY138*-silenced leaves, with the decrease of *GhWRKY138* expression, the expression of *GhMYC2-like* also decreased significantly (Fig. 4A). Similarly, we also obtained RNAi plants of *GhMYC2-like*. To further explore this phenomenon in expression, we performed qRT-PCR experiments on RNAi plants of *GhMYC2-like* gene. Along with suppressed expression of *GhMYC2-like*, *GhWRKY138* was also decreased dramatically. (Fig. 4B)

Similarly, since gossypol is an important component of cotton glands, genes associated with gossypol was also studied. We also performed qRT-PCR analysis on 6 key enzyme genes in the gossypol synthesis pathway[42]. In *GhWRKY138*-silenced leaves, the expression levels of 4 genes decreased and 2 genes increased conversely (Fig.S1).

Discussion

Pigment glands are the storage organs of gossypol and there is an inseparable and close relationship between the glands and gossypol. It is of great significance to study the genes related to gland development and formation in low gossypol cotton breeding. As one of the largest transcription factor families, *WRKY* proteins plays an important role in various physiological processes in plants, such as stress resistance and regulation of secondary metabolite accumulation. However, the effect of *WRKY* gene on cotton gland development has not been reported. Based on previous research, the cotton gland also plays an important role in plant resistance to natural stress such as insect pests. It is therefore not difficult to link *WRKY* transcription factors to glands. However, the effect of *WRKY* gene on cotton gland development has not been reported.

With advances in next-generation sequencing technology, more accurate genomic information has been annotated, providing a better basis for determining the function of genes. In this study, we identified the *WRKY* group I genes in G. hirsutum L. and constructed the phylogenetic tree of *WRKY* group I. The group has a total of 36 members and can be divided into three clades. These *WRKY* group I genes are predicted to be localized in the nucleus, which is consistent with the properties of transcription factors. *GhWRKY138* is a target gene screened by transcriptome expression analysis. *GhWRKY138* belongs to *WRKY* group I, and its expression level is high in glanded cotton, but low in glandless cotton. Subsequently, we verified that the experimental results were consistent with transcriptome data by qRT-PCR. Based on these results, we performed a virus-induced gene silencing (VIGS) experiment. In the

GhWRKY138-silenced leaves, we found that the number of pigment glands was significantly reduced, and there was a significant positive correlation between the expression level and the number of glands. Thus, we speculated that *GhWRKY138* was involved in the pathway of gland development and it has a positive regulatory effect. Subsequently, we found that the expression of *GhMYC2-like* (*GoPGF/CGF3*) gene decreased in the leaves with reduced glands by qRT-PCR. At the same time, the gene expression was also decreased in *GhMYC2-like* knockout plants, suggesting that there may be a feedback mechanism between *GhWRKY138* and *GhMYC2-like*. This suggests that *GhWRKY138* may involved in the formation and development of cotton glands or interact directly with *GhMYC2-like*. These speculations need to be further verified by other experiments.

Due to the special relationship between glands and gossypol, we found that the expression levels of some key enzymes in the gossypol synthesis pathway also showed different trends. Therefore, *GhWRKY138* may also affect the synthesis of gossypol. Whether this means that *GhWRKY138* is also involved in the gossypol synthesis pathway needs further verification. Similarly, whether *WRKY* transcription factors bind the action element of relevant enzyme target gene remains to be investigated. These evidences further confirms that *GhWRKY138* may be involved in the formation and development of cotton glands, but the relationship with gossypol is unclear.

Through this experiment, the biological function of *GhWRKY138* in cotton pigment gland was verified, and the content of WRKY transcription factor family for stress was further enriched. These results may help to further complete the glandular development network and increase the economic value of cotton production.

Methods

Plant Materials and Strains

The plant materials we used were from National Medium-term Gene Bank of Cotton (Anyang, Henan, China) and maintained by selfing in our lab, including glanded and glandless cotton such as L7, CCRI12 (glanded cotton), Dgl − L7(dominant glandless cotton) ,Rgl − CCRI12(dominant glandless cotton) RNAi transgenic plants. Among them, RNAi transgenic plants interfere with the accumulation of *GhMYC2-like* transcripts.

Identification and nature analysis of WRKY **group I genes in** G. hirsutum L.

The upland cotton genomic database *(G. hirsutum L*) (NAU version) we need was downloaded from the Cotton Functional Genomics Database (CottonFGD) (https://cottonfgd.org/). The hidden Markov model (HMM) profile of the conserved *WRKY* domain (Pfam: PF03106) was downloaded from the Pfam database (http://pfam.xfam.org). HMMER 3.0 and BLASTP were used to search for candidate genes of *WRKY* family in upland cotton. Then, through comparison, we artificially remove duplicate genes. The remaining genes were further searched and confirmed on SMART (http://smart.embl.de/), NCBI Conserved Domain Search database(http://smart.embl.de/), and Pfam(http://pfam.xfam.org/).The

protein length and isoelectric point of the *WRKY* group I members were all retrieved from the Cotton Functional Genomics Database (CottonFGD) (https://cottonfgd.org/).

Phylogenetic and structure analysis of WRKY **group I genes in** G. hirsutum L.

The specific phylogenetic tree of the *WRKY* Genome of upland cotton is constructed through the MEGA-X program adopting neighbor-joining method. Then, we analysed the members of the *WRKY* group I selected from the phylogenetic tree. The structure information of introns and exons of group I was obtained from CottonFGD, and the gene structures were graphically visualized using the GSDS2.0 web server(http://gsds.gao-lab.org/). The conserved motifs of the *WRKY* group I genes in upland cotton were analyzed using the MEME program(https://meme-suite.org/meme/).

Expression pattern analysis

To investigate the relationship between *WRKY* group I and gland development, a gland-related G. hirustum transcriptome data was analyzed (determined by our group and unpublished now). Multiple species with and without glands are included in the transcriptome data. The heatmap of the transcriptome was drawn by using the Lianchuan Biocloud platform (https://www.omicstudio.cn/).

Subcellular Localization

Predictive analysis of subcellular location of *GhWRKY138* by using ProtComp tool and Wolf-PSORT. To examine the subcellular localization of *GhWRKY138* in cells, the PCR fragment amplified from cDNA from L7 and inserted into transient expression vector pBI121-GFP and generated the constructs 35S:*GhWRKY138*-GFP. The constructs were introduced into Agrobacterium tumefaciens strain GV3101. The recombinant plasmid was introduced into *N. tabacum* blade cells through acupuncture, Green fluorescent signals were detected under aconfocal microscope after 72 hours.

Virus Induced Gene Silencing (VIGS)

A 245-bp specific fragment of *GhWRKY138* and a 327-bp fragment of *PDS* from L7 was PCR-amplified and inserted to pTRV2 named pTRV2- *GhWRKY138* -*PDS*. The vectors pTRV1 and pTRV2- *GhWRKY138* were introduced into the Agrobacterium strain GV3101 by electroporation. Transformants were selected on LB plate containing kanamycin (50 mg/L), rifampicin (50 mg/L). Silencing the *PDS* gene can lead to albino effect on leaves, which can be used as a marker for the phenotype of VIGS plants. The CCRI12 plants grew in a constant temperature growth room(23℃) under a 16-h light per 8-h dark cycle.

RNA isolation and Quantitative Real-time PCR (qRT-PCR)

Total RNA was extracted from leaves using RNAprep Pure Plant Kit (Tiangen), followed by reverse transcription using PrimeScript II 1st Strand cDNA Synthesis Kit (TAKARA). Real-time PCR was performed using gene-specific primers and 2X RealStar Green Fast Mixture with ROX II (GenStar). The qRT-PCRs were performed in triplicate for each sample and performed on ABI QuantStudio5 RT-PCR system (USA).

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All experimental materials, equipment and transcriptome data were provided by functional Genes Research Group of Cotton Research Institute, Chinese Academy of Agricultural Sciences.

Competing interests

The authors declare that they have no competing interests.

Funding

We are grateful to the State Key Laboratory of Cotton Biology, Institute of Cotton research, Chinese Academy of Agricultural Sciences for providing the seeds of *G. hirsutum.* This research is supported by the National Natural Science Foundation of China (No. 31901581), and Central Public-interest Scientific Institution Basal Research Fund (No.1610162021045 and No.1610162022016).

Authors contributions

DQ, YLD,HLC, DYZ and GLS conceived and designed the research, interpreted the results, and wrote the manuscript; DQ,YLD prepared the materials and conducted the experiments; SL contributed to the data analysis and preparations of figures; DYZ, FJP, QLW, LML and YPZ provided technical assistance and research input. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Sunilkumar, G., et al., *Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol.* 2006. **103**(48): p. 18054-18059.
2. Zhang, W.-J., et al., *Advances in gossypol toxicity and processing effects of whole cottonseed in dairy cows feeding.* 2007. **111**(1-2): p. 1-9.
3. Gadelha, I.C.N., et al., *Gossypol toxicity from cottonseed products.* 2014. **2014**.
4. Tholl, D.J.B.o.i., *Biosynthesis and biological functions of terpenoids in plants.* 2015: p. 63-106.

5. Hedin, P.A., J.N. Jenkins, and W.L.J.J.o.c.e. Parrott, *Evaluation of flavonoids inGossypium arboreum (L.) cottons as potential source of resistance to tobacco budworm.* 1992. **18**(2): p. 105-114.
6. Stipanovic, R., A. Bell, and C.J.B.A.N.P.A. Benedict, CRC Press, Boca Raton, Florida, *Cotton pest resistance: the role of pigment gland constituents.* 1999: p. 211-220.
7. Mao, Y.-B., et al., *Silencing a cotton bollworm P450 monooxygenase gene by plant-mediated RNAi impairs larval tolerance of gossypol.* 2007. **25**(11): p. 1307-1313.
8. Liu, W.Z., et al., *Programmed cell death during pigment gland formation in Gossypium hirsutum leaves.* 2010. **12**(6): p. 895-902.
9. Stanford, E.E. and A. Viehoever, *Chemistry and histology of the glands of the cotton plant, with notes on the occurrence of similar glands in related plants.* 1918.
10. Tian, X., et al., *Characterization of gossypol biosynthetic pathway.* 2018. **115**(23): p. E5410-E5418.
11. McMichael, S.J.A.J., *Glandless boll in upland cotton and its use in the study of natural crossing.* 1954. **46**: p. 527-523.
12. McMichael, S.C.J.A.J., *Combined effects of glandless genes gl2 and gl3 on pigment glands in the cotton plant.* 1960. **52**(7): p. 385-386.
13. Miravalle, R.J.J.C.S., *Action of the Genes Controlling the Character Glandless Seed in Cotton 1.* 1962. **2**(5): p. 447-447.
14. Lee, J.A.J.G., *Genetical studies concerning the distribution of pigment glands in the cotyledons and leaves of upland cotton.* 1962. **47**(1): p. 131.
15. MURRAY, J.C.J.J.o.H., *A new locus for glanded stem in tetraploid cotton.* 1965. **56**(1): p. 42-44.
16. Afifi, A., et al., *Bahtim 110, a new strain of Egyptian cotton free from gossypol.* 1966. **43**: p. 112-120.
17. Kohel, R. and J.J.C.s. Lee, *Genetic analysis of Egyptian glandless cotton 1.* 1984. **24**(6): p. 1119-1121.
18. Cheng, H., et al., *Fine mapping and candidate gene analysis of the dominant glandless gene Gl 2 e in cotton (Gossypium spp.).* 2016. **129**(7): p. 1347-1355.
19. Ma, D., et al., *Genetic basis for glandular trichome formation in cotton.* 2016. **7**(1): p. 1-9.
20. Janga, M.R., et al., *Genes regulating gland development in the cotton plant.* 2019. **17**(6): p. 1142-1153.
21. Eulgem, T., et al., *The WRKY superfamily of plant transcription factors.* 2000. **5**(5): p. 199-206.
22. Rushton, P.J., et al., *WRKY transcription factors.* 2010. **15**(5): p. 247-258.
23. Ishiguro, S., K.J.M. Nakamura, and G.G. MGG, *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 β-amylase from sweet potato.* 1994. **244**(6): p. 563-571.
24. Wang, Q., et al., *WRKY gene family evolution in Arabidopsis thaliana.* 2011. **139**(8): p. 973-983.
25. Wei, K.-F., et al., *Molecular phylogenetic and expression analysis of the complete WRKY transcription factor family in maize.* 2012. **19**(2): p. 153-164.

26. Ning, P., et al., *Genome-wide analysis of WRKY transcription factors in wheat (Triticum aestivum L.) and differential expression under water deficit condition.* 2017. **5**: p. e3232.
27. Huang, S., et al., *Genome-wide analysis of WRKY transcription factors in Solanum lycopersicum.* 2012. **287**(6): p. 495-513.
28. Dou, L., et al., *Genome-wide analysis of the WRKY gene family in cotton.* 2014. **289**(6): p. 1103-1121.
29. Ramamoorthy, R., et al., *A comprehensive transcriptional profiling of the WRKY gene family in rice under various abiotic and phytohormone treatments.* 2008. **49**(6): p. 865-879.
30. Zou, Z., et al., *Gene structures, evolution and transcriptional profiling of the WRKY gene family in castor bean (Ricinus communis L.).* 2016. **11**(2): p. e0148243.
31. Wei, Y., et al., *Genome-wide identification and expression analysis of the WRKY gene family in cassava.* 2016. **7**: p. 25.
32. Ling, J., et al., *Genome-wide analysis of WRKY gene family in Cucumis sativus.* 2011. **12**(1): p. 1-20.
33. Xie, T., et al., *Genome-wide investigation of WRKY gene family in pineapple: evolution and expression profiles during development and stress.* 2018. **19**(1): p. 1-18.
34. Xie, Z., et al., *Salicylic acid inhibits gibberellin-induced alpha-amylase expression and seed germination via a pathway involving an abscisic-acid-inducible WRKY gene.* 2007. **64**(3): p. 293-303.
35. Wang, F., et al., *A novel cold-inducible gene from Pak-choi (Brassica campestris ssp. chinensis), BcWRKY46, enhances the cold, salt and dehydration stress tolerance in transgenic tobacco.* 2012. **39**(4): p. 4553-4564.
36. Song, Y., et al., *The sweet sorghum SbWRKY50 is negatively involved in salt response by regulating ion homeostasis.* 2020. **102**(6): p. 603-614.
37. Yu, Y., et al., *Arabidopsis WRKY71 regulates ethylene-mediated leaf senescence by directly activating EIN2, ORE1 and ACS2 genes.* 2021. **107**(6): p. 1819-1836.
38. Wang, S., et al., *WRKY2 and WRKY10 regulate the circadian expression of PIF4 during the day through interactions with CCA1/LHY and phyB.* 2021: p. 100265.
39. Fu, J., et al., *ZmWRKY79 positively regulates maize phytoalexin biosynthetic gene expression and is involved in stress response.* 2018. **69**(3): p. 497-510.
40. Xu, Y.-H., et al., *Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-δ-cadinene synthase-A.* 2004. **135**(1): p. 507-515.
41. Gu, L., et al., *Identification of the group IIa WRKY subfamily and the functional analysis of GhWRKY17 in upland cotton (Gossypium hirsutum L.).* 2018. **13**(1): p. e0191681.
42. Zhao, T., et al., *Cotton roots are the major source of gossypol biosynthesis and accumulation.* 2020. **20**(1): p. 1-11.

Figures

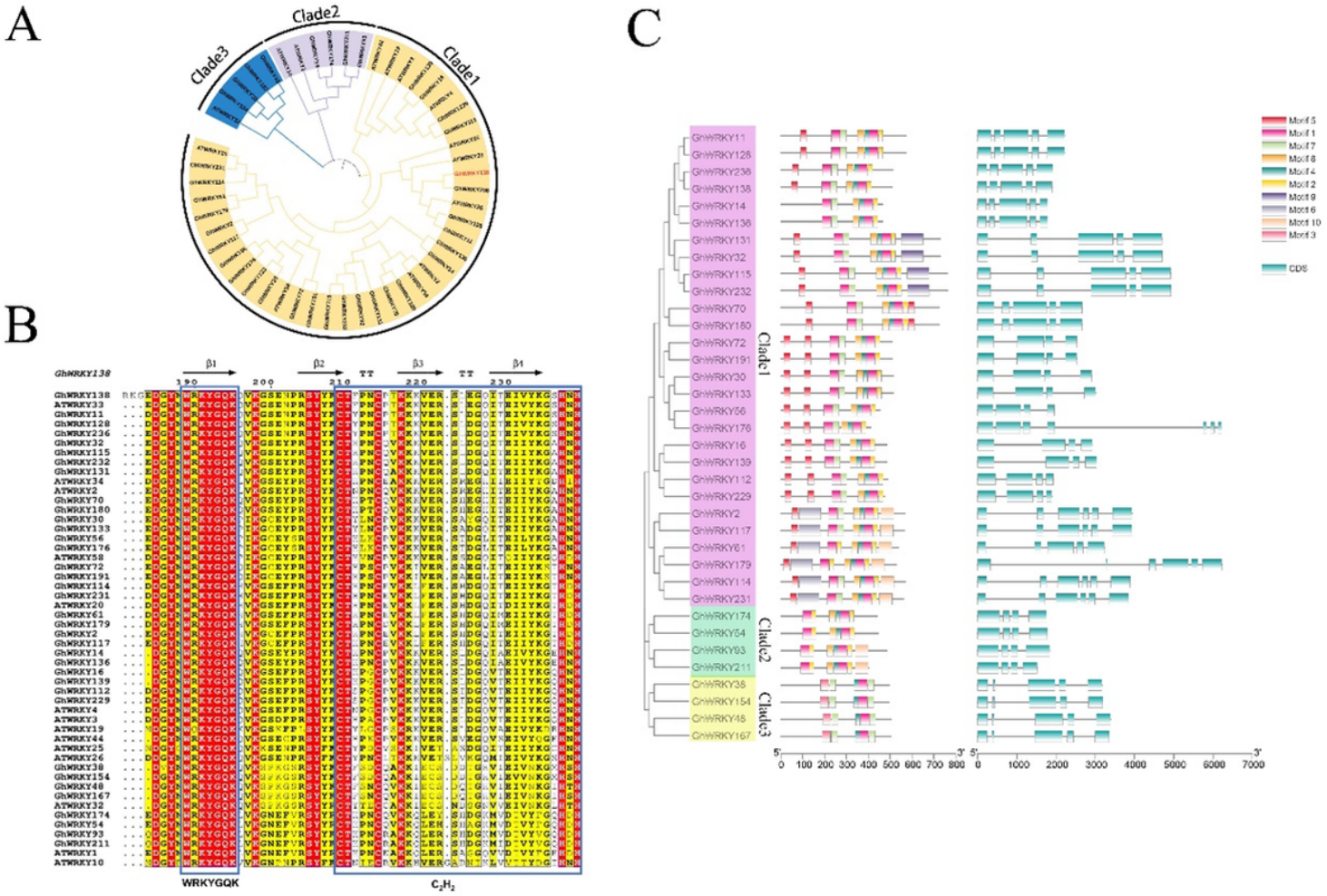

Figure 1

Phylogenetic and structure analysis.(A)The phylogenetic tree of *WRKY* group I proteins in*G. hirsutum L.* and*A. thaliana*.According to phylogenetic tree clustering, it can be divided into three Clades. (B) Multiple alignment of the amino acid sequences of *WRKY* group I in *G. hirsutum L.* and *A. thaliana*. The blue box represents the conserved domain (WRKYGQK) and the C_2H_2 zinc finger motif of *WRKY* group I in *G. hirsutum L.* and*A.thaliana,*respectively. (C) Phylogenetic relationship, motifs, and gene structures of *WRKY* group I in *G. hirsutum L.*

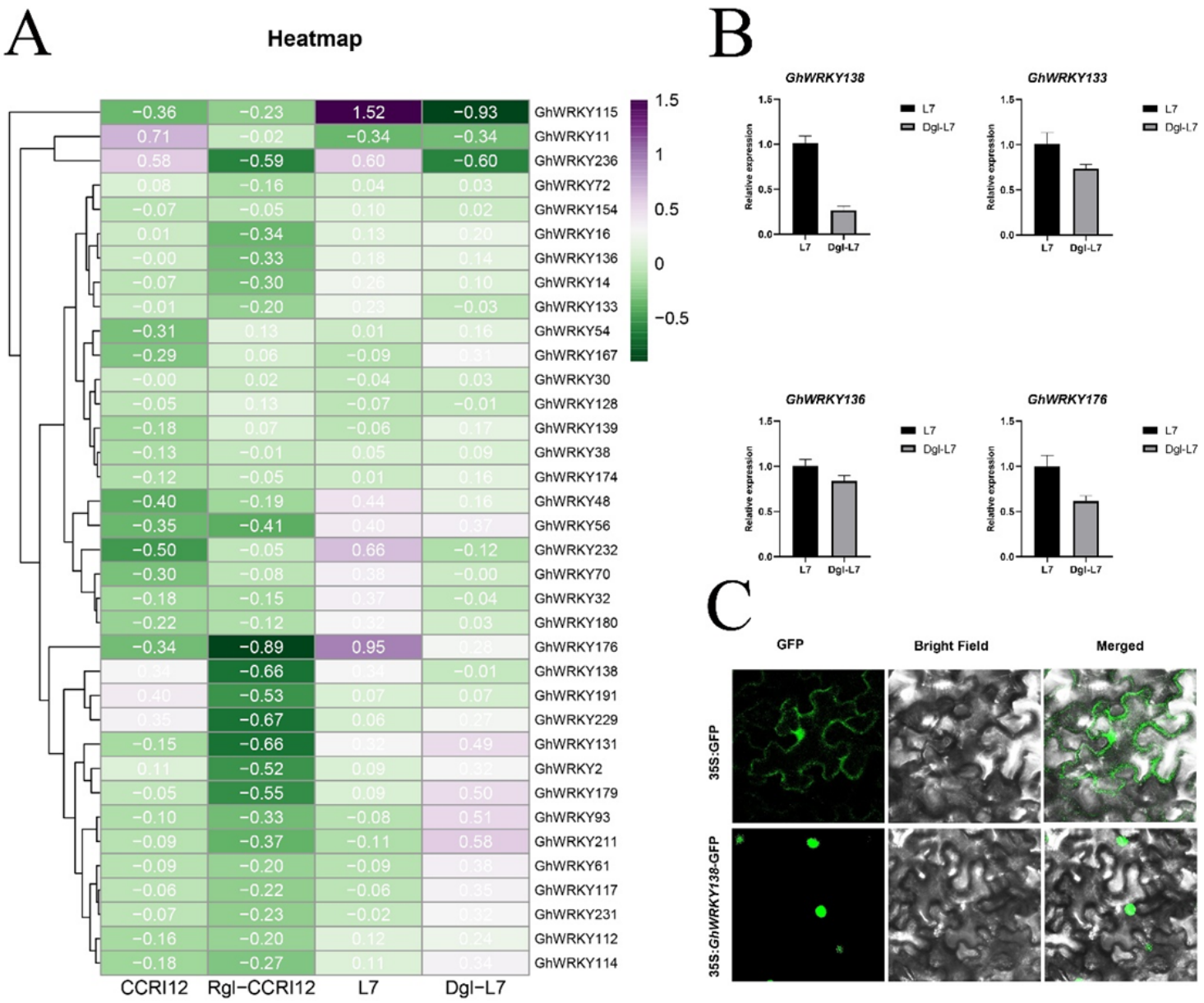

Figure 2

Expression pattern analysis (A) Expression heatmap of *WRKY* group I genes in glanded and glandless cotton. Dgl: dominant glandless, Rgl:recessive glandless, CCRI12 and L7 are glanded cotton (B) Expression analysis of *GhWRKY138* and other 3 genes in glanded and glandless cotton.Dgl: dominant glandless, L7 are glanded cotton. (C) Subcellular localization in tobacco leave cell.

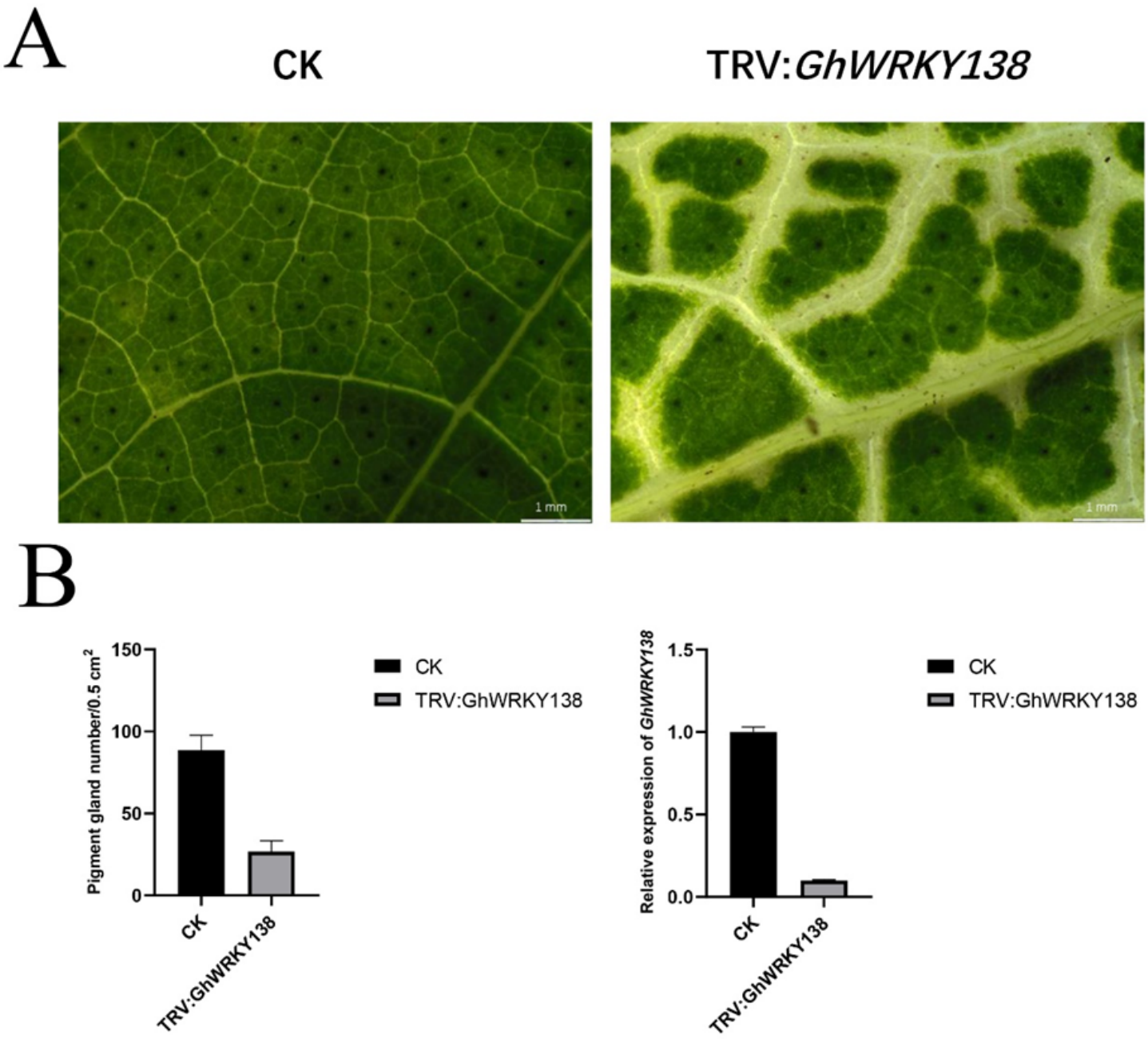

Figure 3

Silencing *GhWRKY138* causes glandless phenotype. (A) Phenotype of *GhWRKY138*-silenced plants (B) The number of pigment glands and expression level in plants before and after VIGS.

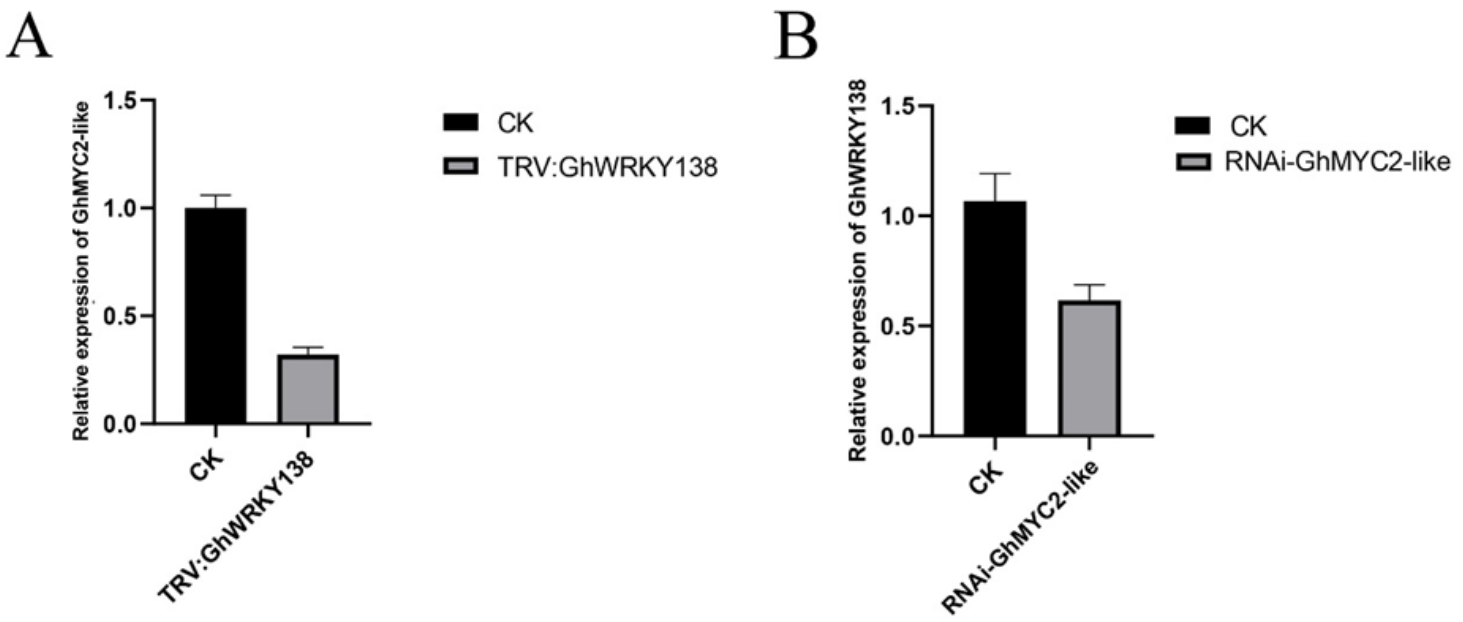

Figure 4

Expression levels of *GhMYC2-like*and *GhWRKY138* in different strains.(a) *GhMYC2-like* expression in *GHWRKY138*-silenced plants. (b)*GhWRKY138* expression in RNAi-*GhMYC2-like* plants.

Supplementary Files

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

- supplementarymaterial.rar