Planta (2022) 255:23
https://doi.org/10.1007/s00425-021-03737-7

ORIGINAL ARTICLE

# Genome-wide identification and expression analysis of GL2-interacting-repressor (GIR) genes during cotton fiber and fuzz development

Xiaoxu Feng[1,2] · Hailiang Cheng[1] · Dongyun Zuo[1] · Youping Zhang[1] · Qiaolian Wang[1] · Limin Lv[1] · Shuyan Li[1] · John Z. Yu[3] · Guoli Song[1]

Received: 30 March 2021 / Accepted: 20 September 2021


## Abstract

***Main conclusion*** **GL2-interacting-repressor (GIR) family members may contribute to fiber/fuzz formation via a newly discovered unique pathway in *Gossypium arboreum*.**

**Abstract** There are similarities between cotton fiber development and the formation of trichomes and root hairs. The GL2-interacting-repressors (GIRs) are crucial regulators of root hair and trichome formation. The *GaFzl* gene, annotated as *GaGIR1*, is negatively associated with trichome development and fuzz initiation. However, there is relatively little available information regarding the other *GIR* genes in cotton, especially regarding their effects on cotton fiber development. In this study, 21 *GIR* family genes were identified in the diploid cotton species *Gossypium arboreum*; these genes were divided into three groups. The *GIR* genes were characterized in terms of their phylogenetic relationships, structures, chromosomal distribution and evolutionary dynamics. These *GIR* genes were revealed to be unequally distributed on 12 chromosomes in the diploid cotton genome, with no *GIR* gene detected on Ga06. The *cis*-acting elements in the promoter regions were predicted to be responsive to light, phytohormones, defense activities and stress. The transcriptomic data and qRT-PCR results revealed that most *GIR* genes were not differentially expressed between the wild-type control and the fuzzless mutant line. Moreover, 14 of 21 family genes were expressed at high levels, indicating these genes may play important roles during fiber development and fuzz formation. Furthermore, *Ga01G0231* was predominantly expressed in root samples, suggestive of a role in root hair formation rather than in fuzz initiation and development. The results of this study have enhanced our understanding of the *GIR* genes and their potential utility for improving cotton fiber through breeding.

**Keywords** *Cis*-acting elements · Expression patterns · Family genes · *Gossypium arboreum* · *GaGIR*

Communicated by Dorothea Bartels.

✉ John Z. Yu
john.yu@usda.gov

✉ Guoli Song
songguoli@caas.cn

[1] State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang 455000, Henan, China

[2] Plant Genetics, Gembloux Agro Bio-Tech, University of Liège, 5030 Gembloux, Belgium

[3] Southern Plains Agricultural Research Center, USDA-ARS, Crop Germplasm Research Unit, 2881 F&B Road, College Station, Texas 77845, USA

## Abbreviations

| | |
|---|---|
| DPA | Day post anthesis |
| GEB | Green fluorescence protein |
| GL2 | GLABRA2 |
| GIR | GL2-interacting-repressor |
| SCW | Secondary cell wall |
| SD | Synthetic dropout |
| TF | Transcription factort |

## Introduction

Cotton is the most important natural fiber crop because it provides the textile industry with raw materials (Fang et al. 2017). Cotton fibers, which differentiate from the epidermal cells of the ovule, undergo the following four distinct,

Published online: 19 December 2021

but overlapping, stages to reach final maturity: initiation, elongation, secondary cell wall (SCW) synthesis, and dehydrated maturation (Padmalatha et al. 2012; Wang et al. 2015; Hu et al. 2016; Sun et al. 2017). The initiation and early elongation stages are essential and highly correlated with fiber characteristics related to yield and quality, including fiber density, length, and uniformity (Rong et al. 2005; Kim et al. 2015; Zhu et al. 2018). Environmental conditions and genetic factors significantly influence these developmental stages, thereby affecting the final fiber quality (Hinchliffe et al. 2011; Gilbert et al. 2014; Liang et al. 2015; Chen et al. 2019). Therefore, the genetic mechanisms underlying fiber initiation should be investigated so that the associated genes may be used to improve cotton fiber quality via molecular breeding.

The recent development and application of high-throughput DNA sequencing technology has resulted in the publication of increasing amounts of cotton genome data, which have been compiled in databases that used as resources for predicting and screening functional genes (Wan et al. 2016; Cheng et al. 2016; Zhu et al. 2017; Wu et al. 2018; Fang et al. 2020). Many genes involved in fiber development have been identified, cloned and functionally characterized (Jiang et al. 2015; Wan et al. 2016; Hu et al. 2016; Thyssen et al. 2017; Wu et al. 2018; Patel et al. 2020; Sun et al. 2020a). For example, *GhbHLH18* is negatively associated with fiber quality because it encodes a protein that strongly binds to the E-box element of the *GhPER8* promoter to activate expression and modulate peroxidase-mediated lignin metabolism during the fiber elongating stage (Gao et al. 2019). Additionally, *GhFSN5*, which is a NAC domain transcription factor (TF) gene, is reportedly preferentially expressed during the SCW synthesis stage to negatively regulate the expression patterns of SCW-associated genes related to cellulose, xylan, lignin and several TFs mediating SCW formation (Sun et al. 2020b). In earlier studies, *GhPIN3a* expression is down-regulated by cytokinin, and the resulting change to the polar distribution of *GhPIN3a* disrupts the asymmetric accumulation of auxin in the ovule epidermis to inhibit cotton fiber initiation (Mei and Zhang 2019; Zeng et al. 2019). Some loci or candidate genomic regions associated with fuzz fiber initiation and formation in tetraploid cotton lines were detected by multiple combined analyses, including *GhMML3_A12* for the $N_1$ locus (Wan et al. 2016), *GhMML3_D12* for the $n_2$ locus (Zhu et al. 2018; Chen et al. 2020), the *GhMML3_A12* allele for the $n_3$ locus (Chen et al. 2020), the 411-kb genomic interval on chromosome D04 for the $n_4^t$ locus (Naoumkina et al. 2021a, 2021b), and the 250-kb candidate region on chromosome D13 for the $N_5$ locus (Zhu et al. 2021).

Cotton fibers share many similarities with *Arabidopsis thaliana* trichomes and root hairs (Lee et al. 2007). Studies on the regulatory network involved in determining Arabidopsis cell fates have provided researchers with a framework for investigating cotton fiber initiation and elongation (Balkunde et al. 2010; Yang and Ye 2013). The multimeric complex that activates trichome initiation and development includes the R2R3 MYB protein GLABROUS1 (GL1), the WD40 repeat-containing protein TRANSPARENT TESTA GLABRA 1 (TTG1), and the bHLH proteins GLABRA3 (GL3) and ENHANCER OF GLABRA3 (EGL3) (Pattanaik et al. 2014; Dai et al. 2016; Matías-Hernández et al. 2016). Researchers have identified and verified the homolog of the complex-encoding genes as well as other regulatory genes in cotton (e.g., *GhMYB25-like*, *GhTTG1/4*, *GhHOX3* and *GhHD-1*) and subsequently confirmed their effects on fiber development (Walford et al. 2012; Huang et al. 2013; Shan et al. 2014; Liu et al. 2015; Ioannidi et al. 2016; Wan et al. 2016; Ding et al. 2020).

Most genes contributing to trichome development are also involved in root hair patterning, suggesting the underlying models are similar, but the resulting phenotypes are different (Matías-Hernández et al. 2016; Wei and Li 2018). The overexpression of *GLABRA2* (*GL2*) in the shoot eventually activates trichome formation. Root epidermal cells expressing *GL2* are prevented from forming root hairs (Rerie et al. 1994; Di Cristina et al. 1996; Ohashi et al. 2002, 2003; Hülskamp 2004). The GL2-interacting-repressors (GIRs) negatively regulate root hair development by interacting with GL2 and controlling the interaction network (Wu and Citovsky 2017a, b). Several recent studies indicated that *GaFzl*, which is annotated as *GIR1* in diploid cotton, adversely affects fuzz development (Du et al. 2018; Feng et al. 2019; Liu et al. 2020; Wang et al. 2020c). However, the characteristics and underlying mechanisms of the GIRs functions and regulatory network in cotton remain unknown.

In this study, we revealed the basic features of *GaFzl* by determining the subcellular localization and analyzing its transcriptional activation of the encoded protein. Furthermore, on the basis of the published cotton genomes, we screened and identified 21, 21 and 40 *GIR* gene family members in *Gossypium arboreum*, *Gossypium raimondii* and *Gossypium hirsutum*, respectively, and then characterized their phylogenetic relationships, gene structures, chromosomal distribution and evolutionary dynamics. We also investigated the expression patterns of *GaGIR* genes and other fiber-related genes in cotton fiber. The results presented herein will be useful for clarifying the molecular mechanisms associated with the *GIR* genes in diploid cotton.

## Materials and methods

### Plant materials and bacterial strains

The plant materials used in this study included *G. arboreum* lines DPL971 (wild-type), DPL972 (near isogenic fuzzless

mutant), *G. hirsutum* lines XZ142 (wild-type), fuzzless mutant $N_1$, and fuzzless-lintlesss mutant XZ142FLM, which were obtained from Germplasm Repository of Institute of Cotton Research, Chinese Academy of Agricultural Sciences (CRI of CAAS, Anyang, Henan province, China) only for scientific research purpose. Plants were grown and self-pollinated for conservation annually in accordance with standard agronomic practices at the farm of CRI. *Nicotiana benthamiana* was grown in a 26 °C incubator with 16 h light/ 8 h dark.

*Escherichia coli DH5α* competent cells were used for gene cloning. *Agrobacterium tumefaciens* strain GV3101 was used for the transformation of tobacco (*Nicotiana benthamiana)*.

## Subcellular localization

Subcellular localization was performed using tobacco mesophyll cells. The CDS of *GaFzl* was cloned into binary vector pBI121-GFP to construct the GaFzl-GFP fusion protein downstream of CaMV 35S promoter. The recombinant constructs, 35S: GaFzl-GFP and the control vector pBI121- GFP, were respectively transformed into the *A. tumefaciens* strain GV3101 and inoculated into young leaves of tobacco. About 48 h later, the injected tobacco leaves were stained with 1-[3-(triethylaminio)propyl]-4-[6-[4-(diethylamino)phenyl]-1,3,5-hexatrienyl]pyridinium (FM4-64, a membrane-specific dye) and 4′6-diamidino-2-phenylindole (DAPI, a nucleus-specific dye) for 5 min, respectively, and afterward the fluorescence signals of fusion proteins in tobacco mesophyll cells were observed using a confocal fluorescence microscope (Leica). Primers used for gene cloning and vector construction are listed in Supplemental Table S1.

## The transcriptional activation analysis

The transcriptional activation assay of GaFzl was performed in yeast. The CDS of *GaFzl* from DPL971 and DPL972 were cloned into the bait vector pGBKT7 to construct the GaFzl_BD fusion protein, respectively. Both the recombinant constructs were transformed into yeast strain Y2HGold. The transformants were further cultivated on Synthetic dropout (SD) medium (SD/-Trp, SD/-Trp/X-α-Gal, SD/-Trp/X-α-Gal/AbA and SD/-Trp/-Ade/-His/X-α-Gal /AbA).

## Identification and characterization of GaGIR genes

The protein sequence of AtGIR1 was downloaded from NCBI and subsequently used as a query to identify the GIR members in cotton genomes and search the hidden Markov model (HMM) profile from the Phytozome v12.1 database (https://phytozome.jgi.doe.gov/pz/portal.html#) and the PANTHER Classfication System (http://pantherdb.org/). We downloaded the latest versions of genome annotation data of Arabidopsis from the Arabidopsis Information Resource (TAIR; available online: https://www.arabidopsis.org/index.jsp). The genome data of *G. hirsutum* (version: AD1_ZJU v2.1), *G. aboreum* (version: A2_CRI) and *G. raimondii* (version: D5_JGI) were downloaded from Cotton Functional Genomic Database (CottonFGD; https://cottonfgd.org/). The HMM profile was used to screen and identify the *GIR* genes in *Gossypium* by the HMMSEARCH program from HMMER 3 software. Furthermore, BLAST was also used to confirm the family members.

## Phylogenetic, structure and conserved motif analysis

The conserved region of the 21 GIR protein sequences from *G. arboreum* and 6 from *A. thaliana* were used for phylogenetic tree construction. Multiple sequence alignments were conducted using ClustalX program with default parameters. The phylogenetic tree was constructed in MEGA10 using the neighbor-jointing algorithm with the *P*-distance model and the pairwise deletion option with 1000 bootstrap replicates.

The Gene Structure Display Server 2.0 (http://gsds.cbi.pku.edu.cn/) was employed to graphically visualize the exon–intron structure of *GIR* genes. The MEME (Multiple Expectation Maximization for Motif Elicitation) (http://meme-suite.org/tools/meme) program was used to predict both conserved and potential motifs of GaGIR protein sequences using the parameter settings: the minimum motif width = 6, the maximum motif width = 50, and the maximum number of motifs = 10.

## Chromosomal locations and synteny analysis

MapChart software (http://www.earthatlas.mapchart.com/) was used to visualize the distribution of *GIR* genes on cotton chromosomes. MCScan was utilized to analyze the synteny pairs of *GIR* genes of different cotton species. The synonymous substitution (Ks) and nonsynonymous substitution (Ka) rate of *GIR* gene pairs were calculated using the DnaSP software.

## Cis-acting element analysis and identification of transcription factor (TF) binding sites

The 2000 bp sequence upstream of the start codons of *GaGIR* genes were downloaded from Cotton Functional Genomics Database (https://cottonfgd.org/). The cis-acting elements were screened and predicted using the Plant *Cis*-Acting Regulatory Element website (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) database, and the diagram was generated by TBtools. The TF binding sites in the promoter regions of *GaGIR* genes were analyzed

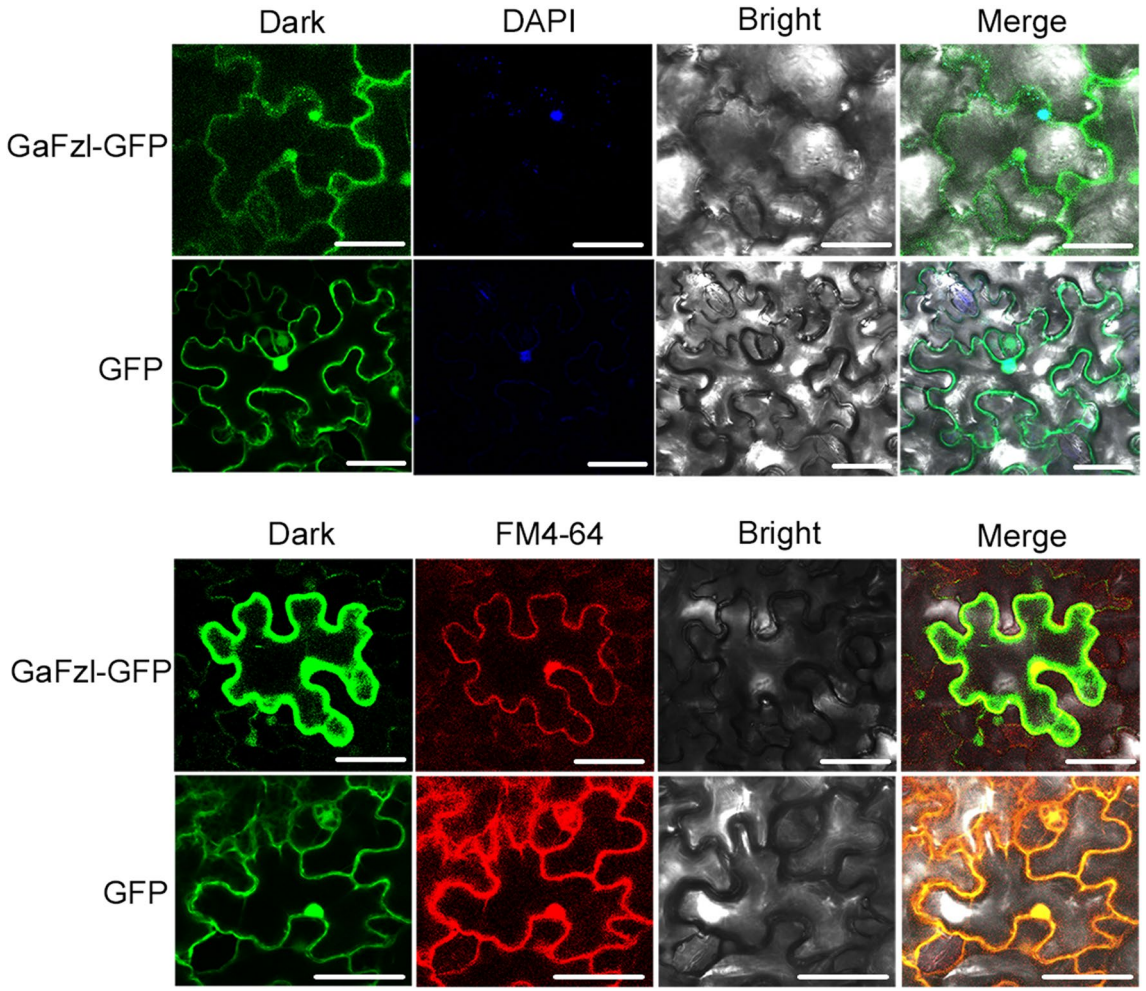



**Fig. 1** Subcellular localization of *GaFzl*. The names of constructs are shown on the left. The scale bar is 50 μm

and identified using Plant Transcription Factor Database (http://planttfdb.gao-lab.org/).

### RNA isolation and qRT-PCR analysis

For samples collection, cotton lines DPL971, DPL972, XZ142, XZ142FLM and $N_1$ were grown in the farm of CRI of CAAS with standard field management in Anyang, Henan province (China). Fiber-bearing ovules were harvested at − 1, 0, 1, 3 and 5 days post-anthesis (DPA). The ovule samples were wrapped in foil and frozen directly in liquid nitrogen. Total RNA was extracted using RNA Prep Pure Plant kit (Tiangen) according to the kit manuals and cDNA was reverse-transcribed from 1 μg total RNA using TransScript all-in-one first-strand cDNA Synthesis SuperMix (TransGen Biotech). The qRT-PCR amplifications were conducted using TransStart TOP Green qPCR SuperMix (TransGen Biotech) on ABI Prism7500 Fast Real-time PCR System (Applied Biosystems). Gene expression levels were normalized and calculated using *GaHis3* as a reference gene. Relative expression levels were calculated using the$2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001). Primers used in this study were designed using NCBI Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and qPrimerDB (https://biodb.swu.edu.cn/qprimerdb/) and synthesized by Sangon Biotech company. Primer sequences are provided in Suppl. Table S1.

## Results

### Subcellular localization analysis of GaFzl

In our previous study, *GaFzl* was fine-mapped and identified as the candidate gene controlling fuzz development in

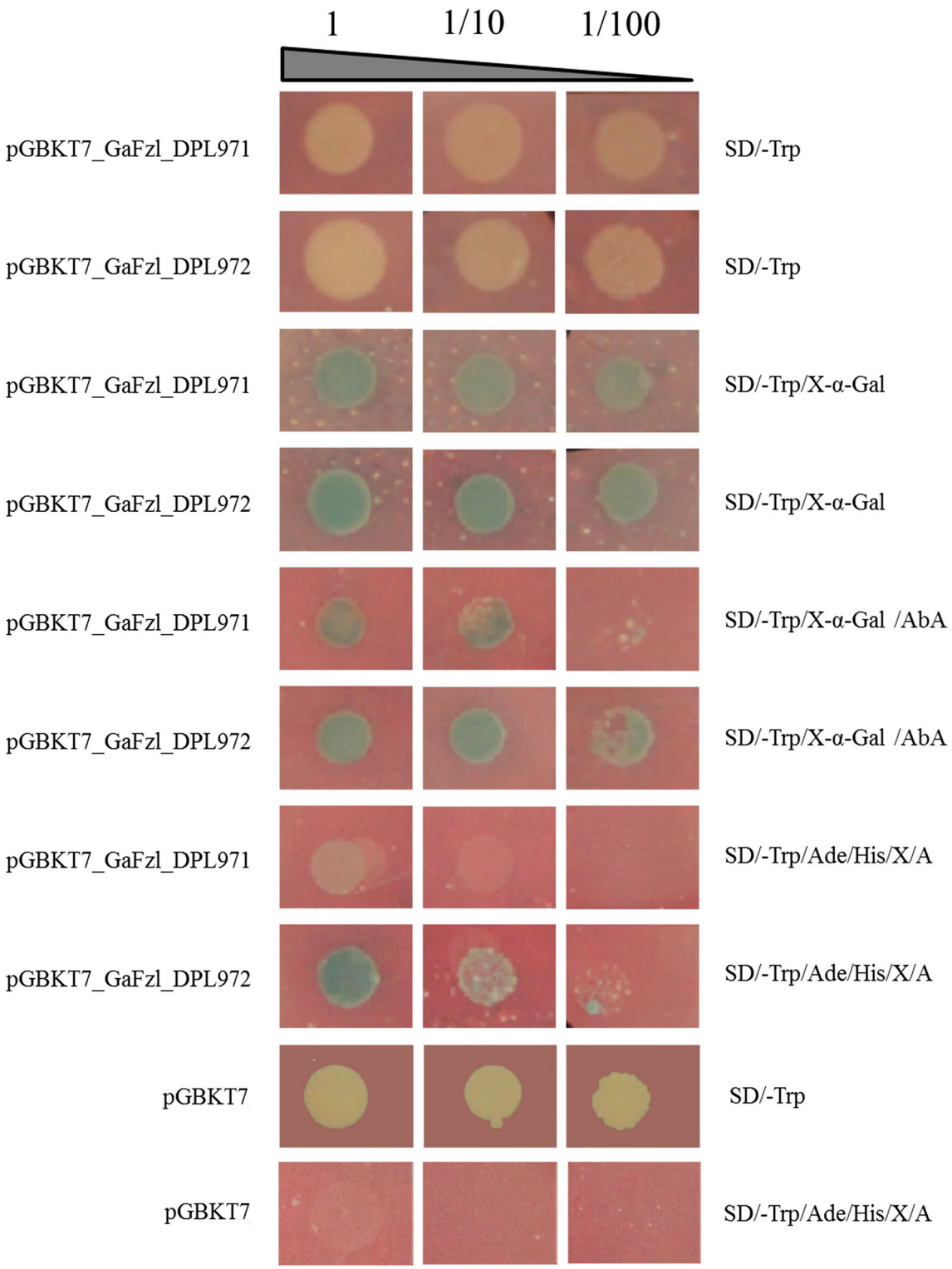



**Fig. 2** Transcriptional autoactivation of GaFzl was detected in yeast. *GaFzl* amplified from DPL971 and DPL972 was inserted into pGBKT7 to construct expression vectors indicated on the left; Y2Hgold stains containing the expression constructs were diluted and inoculated on the SD medium indicated on the right. The number of 1, 1/10 and 1/100 indicate yeast stain without dilution, tenfold dilution and 100-fold dilution

**Table 1** Genome-wide identification of *GIR* family genes in *G. arboreum*

| No | Gene ID | Chr | Strand | Start | End | AA | kDa | Isoelectric point | Grand average of hydropathy |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Ga01G1554 | 1 | – | 5,80,27,883 | 5,80,28,242 | 119 | 12.989 | 8.209 | – 0.516 |
| 2 | Ga02G1535 | 2 | – | 9,56,60,733 | 9,56,61,080 | 115 | 12.651 | 8.785 | – 0.463 |
| 3 | Ga11G0145 | 11 | + | 12,72,908 | 12,73,225 | 105 | 11.146 | 4.552 | – 0.242 |
| 4 | Ga13G0044 | 13 | – | 4,60,349 | 4,60,690 | 113 | 12.482 | 6.441 | – 0.581 |
| 5 | Ga05G0849 | 5 | – | 74,01,337 | 74,01,666 | 109 | 11.739 | 8.495 | – 0.341 |
| 6 | Ga07G1212 | 7 | + | 1,94,10,537 | 1,94,10,872 | 111 | 12.345 | 5.828 | – 0.51 |
| 7 | Ga09G1549 | 9 | – | 7,26,39,829 | 7,26,41,385 | 113 | 12.266 | 8.114 | – 0.181 |
| 8 | Ga08G0121 | 8 | + | 9,05,296 | 9,05,559 | 87 | 9.834 | 4.041 | – 0.28 |
| 9 | Ga10G2569 | 10 | – | 12,46,89,192 | 12,46,90,513 | 237 | 25.162 | 9.781 | – 0.435 |
| 10 | Ga07G1259 | 7 | – | 2,08,94,052 | 2,08,95,016 | 243 | 26.101 | 8.577 | – 0.472 |
| 11 | Ga13G1545 | 13 | – | 9,94,50,521 | 9,94,52,131 | 313 | 34.876 | 10.271 | – 0.493 |
| 12 | Ga04G1030 | 4 | + | 4,18,54,005 | 4,18,54,336 | 96 | 10.241 | 8.332 | – 0.406 |
| 13 | Ga13G0100 | 13 | – | 10,04,877 | 10,05,634 | 175 | 19.189 | 6.118 | – 0.714 |
| 14 | Ga13G0104 | 13 | – | 10,77,386 | 10,78,137 | 171 | 19.158 | 5.058 | – 0.494 |
| 15 | Ga13G0101 | 13 | – | 10,13,106 | 10,13,903 | 184 | 20.43 | 4.635 | – 0.719 |
| 16 | Ga01G0231 | 1 | – | 17,53,020 | 17,53,839 | 181 | 19.811 | 4.318 | – 0.872 |
| 17 | Ga12G1438 | 12 | – | 2,07,00,695 | 2,07,02,020 | 203 | 22.877 | 4.237 | – 1.005 |
| 18 | Ga13G0103 | 13 | – | 10,75,955 | 10,76,629 | 135 | 15.119 | 8.121 | – 0.479 |
| 19 | Ga05G2409 | 5 | – | 2,26,55,865 | 2,26,58,507 | 172 | 19.27 | 9.172 | – 0.653 |
| 20 | Ga03G0898 | 3 | + | 1,75,60,035 | 1,75,60,783 | 209 | 23.684 | 7.319 | – 0.669 |
| 21 | Ga03G0159 | 3 | – | 13,84,990 | 13,85,763 | 193 | 21.889 | 6.495 | – 0.594 |

*G. arboreum*. To further explore the *GaFzl* function related to cotton fuzz initiation, its subcellular localization was analyzed using tobacco leaves (*Nicotiana benthamiana*). The GaFzl-GFP fusion protein was transiently expressed in young tobacco leaves, which were then stained using a membrane-specific dye (FM4-64) and a nucleus-specific dye (DAPI) to ascertain the localization of the GaFzl-GFP signal by confocal microscopy. Consistent with the findings of an earlier investigation by Wang et al. (2020c), the strong fluorescence of the GaFzl-GFP fusion protein was detected in the membrane and nucleus, which were stained with FM4-64 (red fluorescence) and DAPI (blue fluorescence), respectively (Fig. 1). Accordingly, GaFzl appears to be co-localized in the nuclear and membrane of cells.

## Transcriptional activation analysis of GaFzl

Because AtGIR1 is a TF that regulates root hair formation as, we assessed whether GaFzl has transcriptional activation activity by conducting an autoactivation analysis of GaFzl in yeast. The GaFzl- BD (GAL4 binding domain) fusion protein in yeast activated the expression of the reporter genes MEL1(X-α-Gal) and AUR1-C (AbA), suggesting that GaFzl has transcriptional activation activity and may function as a transcriptional regulator (Fig. 2). To examine whether the transcriptional activation activity of the fusion protein could be repressed, SD agar medium (SD/-Trp/-Ade/-His/X-α-Gal /AbA) was inoculated with Y2HGold cells expressing the fusion proteins. In yeast cells, the fusion protein with GaFzl from DPL972 (fuzzless isogenic line) had a higher transcriptional activation activity level than the fusion protein with GaFzl from DPL971 (wild-type), and it activated the expression of all four reporter genes. The observed difference in the transcriptional activation activity may be related to the GaFzl sequence variation between DPL971 and DPL972.

## Genome-wide identification of GIR genes in *G. arboreum*

To thoroughly characterize the *GIR* gene family members in cotton, the whole-genome sequences of cotton species (*G. arboreum* (version: A2_CRI), *G. hirsutum* (version: AD1_ZJU v2.1) and *G. raimondii* (version: D5_JGI)) were analyzed to identify *GIR* genes. The Arabidopsis GIR1 protein sequence was used as the query to screen three cotton reference genomes for candidate GIR proteins, with an E-value threshold of 0.01. On the basis of the hidden Markov model (HMM) of PTHR33177, the HMMER 3 program was used to identify 21 GIR proteins in *G. arboreum*. Additionally, 21, 40 and 6 GIR proteins were detected in *G. raimondii*, *G. hirsutum*, and *A. thaliana*, respectively (Suppl. Table S2,

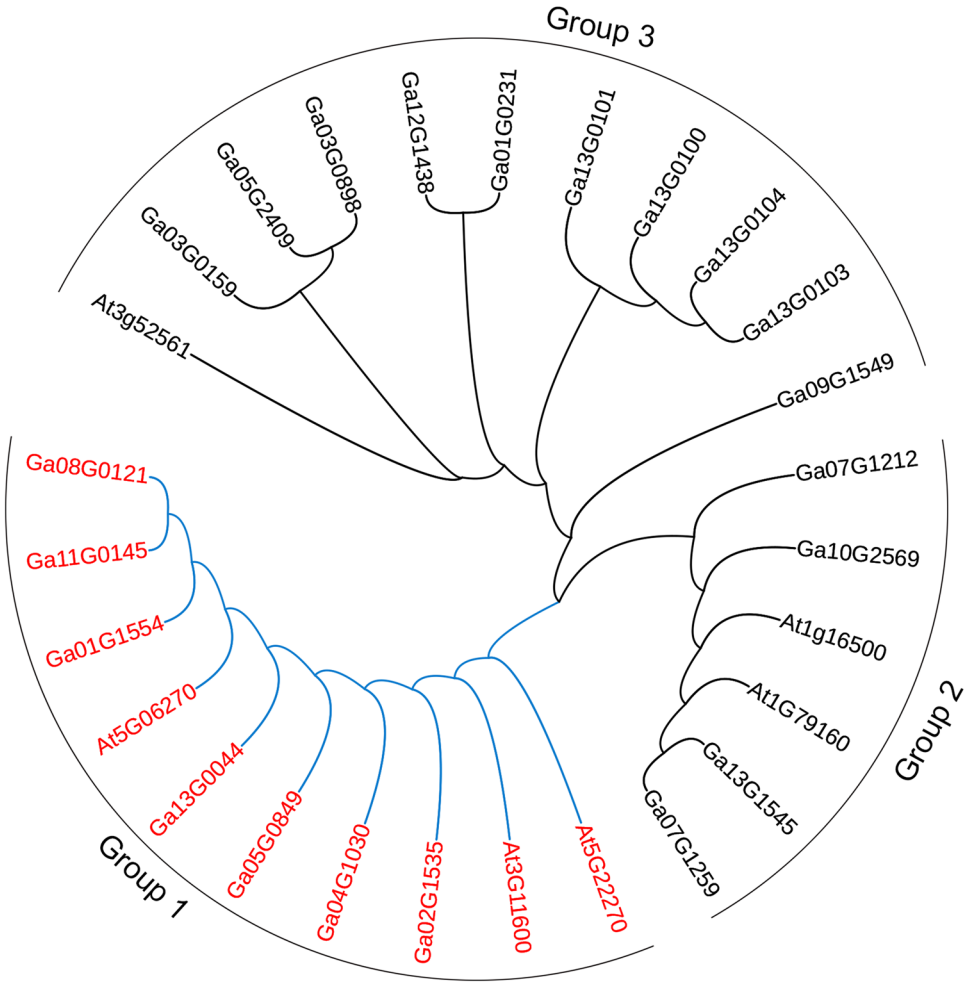



**Fig. 3** Phylogenetic analysis of the *GIR* gene family. The phylogenetic tree was constructed using the full length GIR protein amino acid sequences from *G. arboreum* and *A. thaliana*. The genes highlighted in red fonts and blue lines were classified and named as Group 1. At, *Arabidopsis thaliana*; Ga, *Gossypium arboreum*

S3). Specific information of *GaGIR* genes, such as gene ID, chromosomal location, protein size (AA), molecular weight (kDa) were listed in Table 1. The encoded proteins comprised 87 amino acids (Ga08G0121) to 243 amino acids (Ga07G1259), suggesting these are small proteins that may function as adapters or small molecular peptides that interact with other proteins to perform their functions.

## Phylogenetic, gene structure, and motif analyses of GIR proteins in *G. arboreum*

To classify the GIR proteins and clarify their evolutionary relationships, a multiple sequence alignment analysis (Suppl. Fig. S1) was performed using 21 and 6 homologous protein sequences from *G. arboreum* and *A. thaliana*, respectively. Additionally, an unrooted phylogenetic tree was constructed according to the neighbor-joining algorithm of the MEGA10.0 software, with 1000 bootstraps replicates. In the constructed phylogenetic tree (Fig. 3), the 27 proteins were divided into three groups. Within the GaGIR family, 7 members were clustered in Group 1, 4 members belonged to Group 2 and 10 members were included in Group 3. Group 1 included GaGIR1 and two AtGIR proteins (AtGIR1 and AtGIR2), implying that these members may be associated with similar functions.

Gene structures and organization may provide additional evidence of the evolutionary relationships among *GIR* family members. A comparative analysis of gene structures and motifs detected 0–2 introns in the *GaGIR* genes. More specifically, seven genes had no intron, whereas one gene had two introns, and the remaining genes comprised only one intron (Fig. 4). Ten conserved motifs in the GaGIR proteins were identified using the MEME online software (i.e., motifs 1 to 10). Almost all GaGIR proteins had motifs 1, 2 and 3, implying these motifs may function as a highly conserved domain. Most of the Group 1 GaGIR proteins contained motif 4. In contrast, most of the Group 2 proteins had motifs 5, 6, 8, 9, and 10, whereas the Group 3 proteins had motifs 3, 5, and 7.

## Chromosomal distribution of GaGIR genes

The *GaGIR* genes were mapped on cotton chromosomes to gain new insights into the organization of *GaGIR* genes in

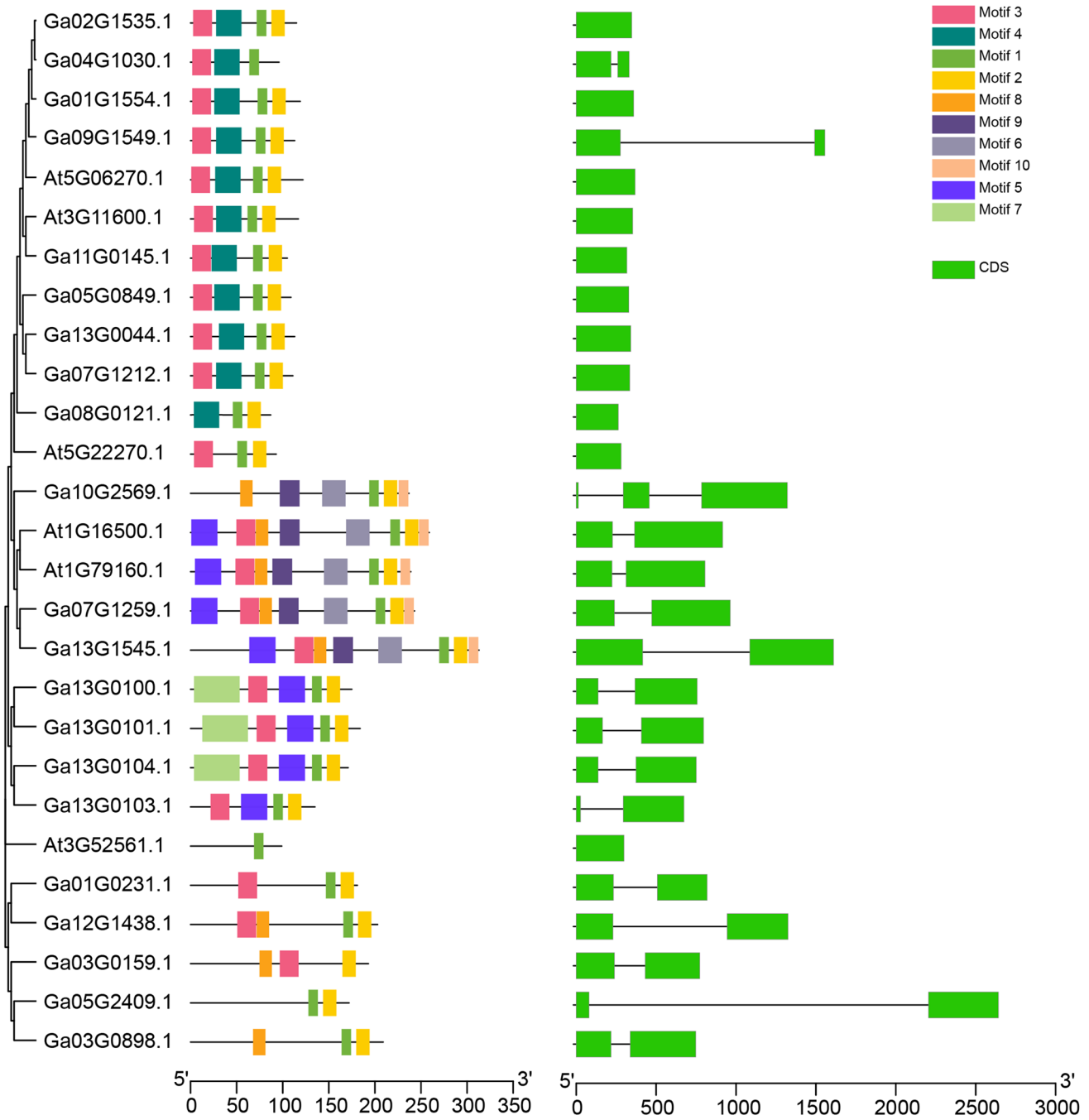



**Fig. 4** Phylogenetic relationship, conserved motif and gene structure analysis of GaGIR proteins. The left part shows the conserved motifs involved in GaGIR proteins. The number and order of motifs in each GaGIR proteins are presented. Gene structure (exon–intron organization) of *GaGIRs* is displayed on the right. Exons and introns are represented by green boxes and black lines, respectively. The scale bar is shown at the bottom

the cotton genome. Using the available *G. arboreum* genome as a reference, we determined that the 21 *GaGIR* genes were unequally distributed among the chromosomes (Fig. 5). Seven chromosomes (chr. 2, 4, 8, 9, 10 and 12) contained only one *GaGIR* gene, whereas four chromosomes (chr. 1, 3, 5 and 7) had two *GaGIR* genes. Chromosome 13 included six *GaGIR* genes, four of which were tandemly distributed. However, the chromosomal distribution of the *GIR* genes varied among *G. arboreum*, *G. raimondii*, and *G. hirsutum* (At and Dt-subgenomes), indicative of gain and loss events during whole-genome duplications (Suppl. Fig. S2, S3, S4). To clarify the collinearity of the *GIR* gene families between

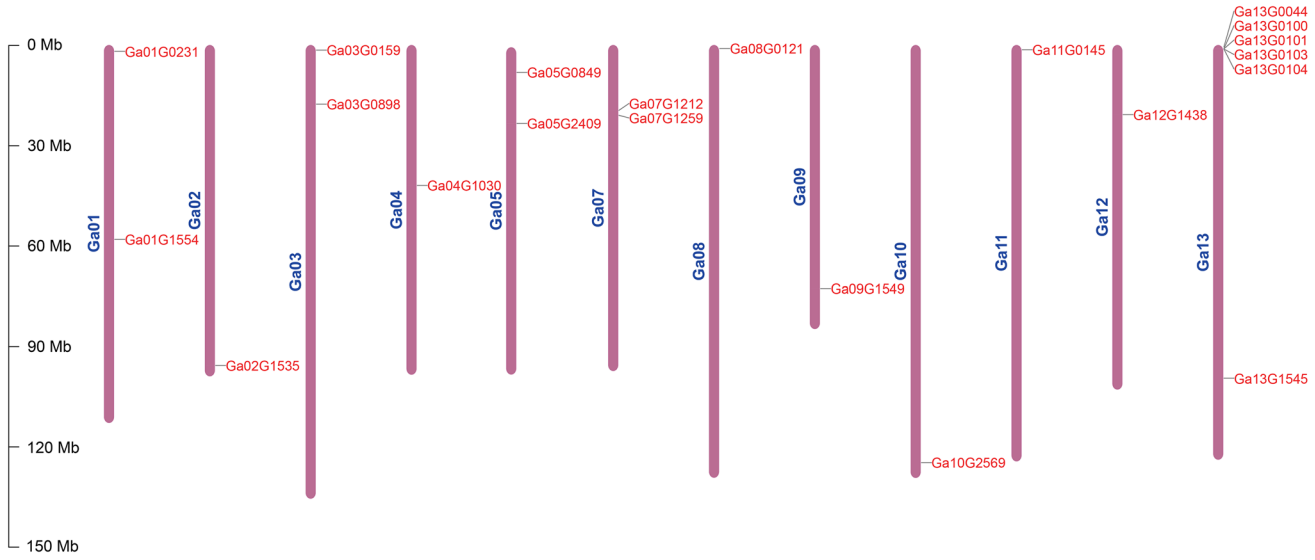



**Fig. 5** Chromosome distribution of *GIR* family genes in *G.arboreum*. The chromosome name is presented on the left side of the graph, and the gene ID is on the right. The vertical scale shows the size of chromosomes and black lines indicate the corresponding position of the genes. The scale bar indicates the chromosome length in base pair (bp)

two diploid cotton ancestors and *G. hirsutum*, linked gene pairs were identified using MCScan. A total of 34 collinear gene pairs were identified between *G. hirsutum* and *G. arboreum*, 18 of which involved the A-subgenome of *G. hirsutum* (Fig. 6). Additionally, 39 collinear gene pairs were detected between *G. hirsutum* and *G. raimondii*, with 20 genes from the D-subgenome. To elucidate the evolutionary dynamics and selection pressures, we calculated the nonsynonymous (Ka) and synonymous (Ks) substitution rates of the *GIR* gene pairs. Most of the Ka/Ks ratios were less than 1.0 in the intergenomic and intragenomic analyses, implying that most *GIR* genes underwent a purifying selection in the diploid and allotetraploid cotton species (Suppl. Table S4, S5, S6).

## Cis-acting element analysis

Because most *GIRs* remain unannotated, the putative functions of *GIRs* were first assessed by identifying the *cis*-acting elements in the 2-kb promoter region of the corresponding genes. In addition to the general elements (i.e., TATA-box and CAAT-box), the detected *cis*-acting elements were mainly enhancer activity elements and elements responsive to light, phytohormones, defense activities and abiotic stress (i.e., drought, low-temperature) (Fig. 7). Interestingly, 20 of the 21 *GaGIR* genes had *cis*-acting elements responsive to 3–6 phytohormones (ABA, followed by ethylene, MeJA, SA, GA and auxin). An analysis of the transcription factor-binding site in the promoter regions indicated the binding sites of ERF, BBR-BPC, MYB, AP2/B3, DOF and C2H2 transcription factors were extensively distributed in the *GaGIR* promoters (Suppl. Table S7). These results suggest that *GaGIR* genes likely contribute to plant growth and development through various signal transductions pathways, especially those related to phytohormone, light, and abiotic stress responses.

## Expression patterns of GaGIR genes at different stages of fiber development

To investigate the potential functions of *GIR* genes in *G. arboreum*, we analyzed their expression patterns using the transcriptome data for different stages of fiber development (1, 3 and 5 DPA). According to their expression levels, the *GaGIR* genes were roughly classified as strongly expressed or barely expressed. The heatmap representation of expression revealed that the *GaGIRs* genes were not significantly differentially expressed between DPL971 (wild-type) and DPL972 (fuzzless isogenic line), with the exception of *GaFzl* (Fig. 8). Consistent with previous studies, the *GaFzl* expression level was much higher at 1, 3 and 5 DPA in DPL972 than in DPL971, implying that the up-regulated expression of *GaFzl* likely suppresses fuzz development. This gene is reportedly significantly up-regulated by a 6.2 kb insertion in fuzzless mutant and thus leads to a fuzzless phenotype (Wang et al. 2020c). Hence, other genes which normally expressed at low levels in wild-type cotton might also alter fuzz development if they are expressed at higher levels. To verify the accuracy and reliability of the RNA-seq results, we also examined the relative expression levels of *GIR* family genes in ovules at different stages by qRT-PCR. The results were consistent with the RNA-seq data,

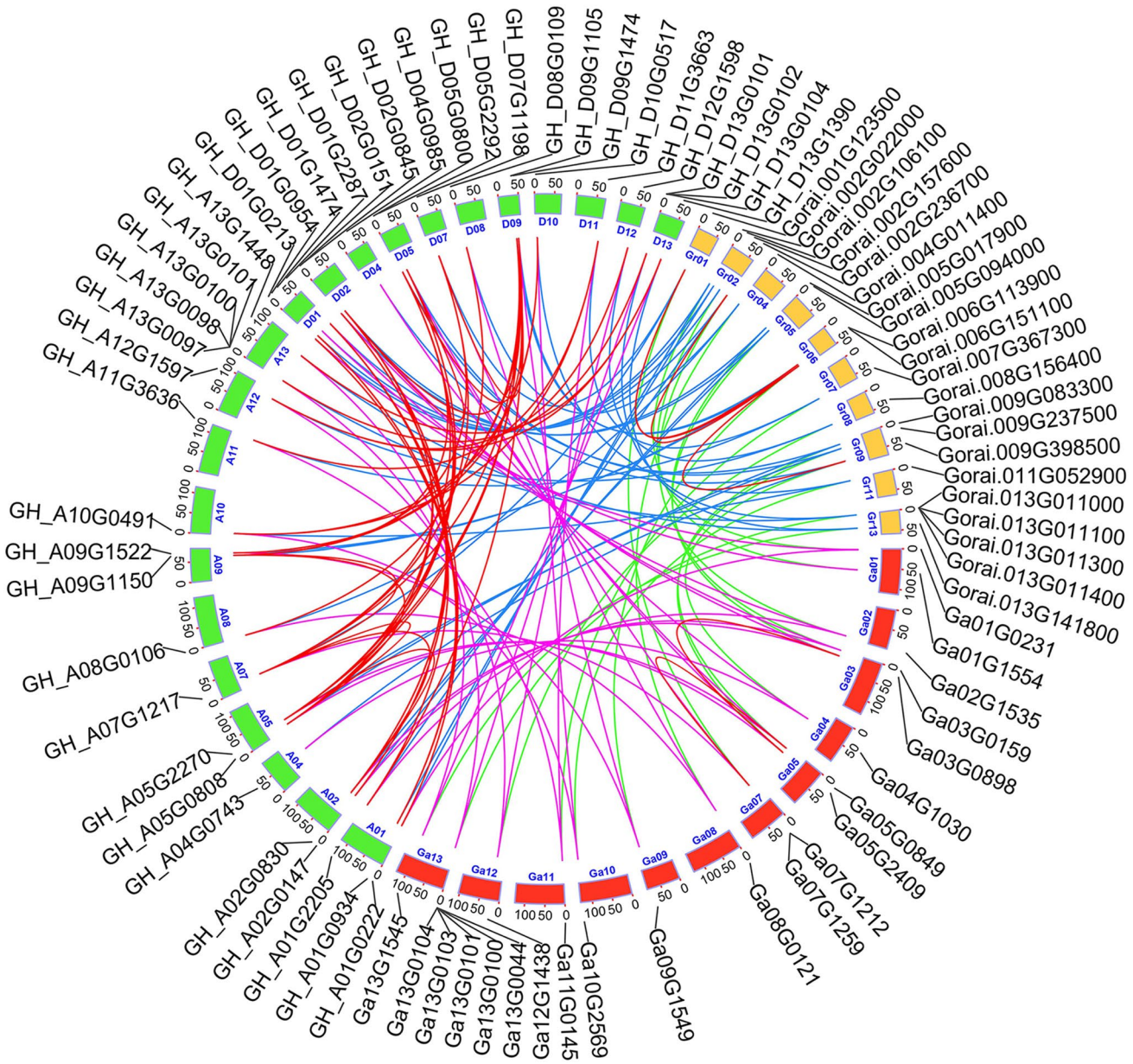



**Fig. 6** Collinearity analyses of *GIR* genes among *G. hirsutum*, *G. arboreum* and *G. raimondii*. Red lines indicate the intergenomic collinearity; other three lines represent the intragenomic collinearity

confirming the transcriptome data were reliable (Fig. 9). Previous studies proved AtGIRs are involved in Arabidopsis root hair formation (Wu and Citovsky 2017a, b). To explore the functions of *GaGIR* genes, we examined the root phenotypes of two diploid cotton species and then compared the *GaGIR* expression patterns between the two materials by qRT-PCR. There were no obvious differences in the root hair phenotypes (Suppl. Fig. S5) and gene expression levels (Fig. 9). Moreover, most of the *GIR* genes were expressed at much lower levels in the root than in the fiber. However, *Ga01G0231* was highly expressed in the root (Fig. 9), indicative of an important role in root hair formation and development.

## Regulatory relationship between GaFzl and other fiber-associated genes

To further investigate the potential regulatory network of *GaGIR* genes involved in fuzz development, we analyzed the expression level of *GaFzl* in various fiber mutants as well as the expression of other fiber-related genes in *G. arboreum* by qRT-PCR. Because gene expression

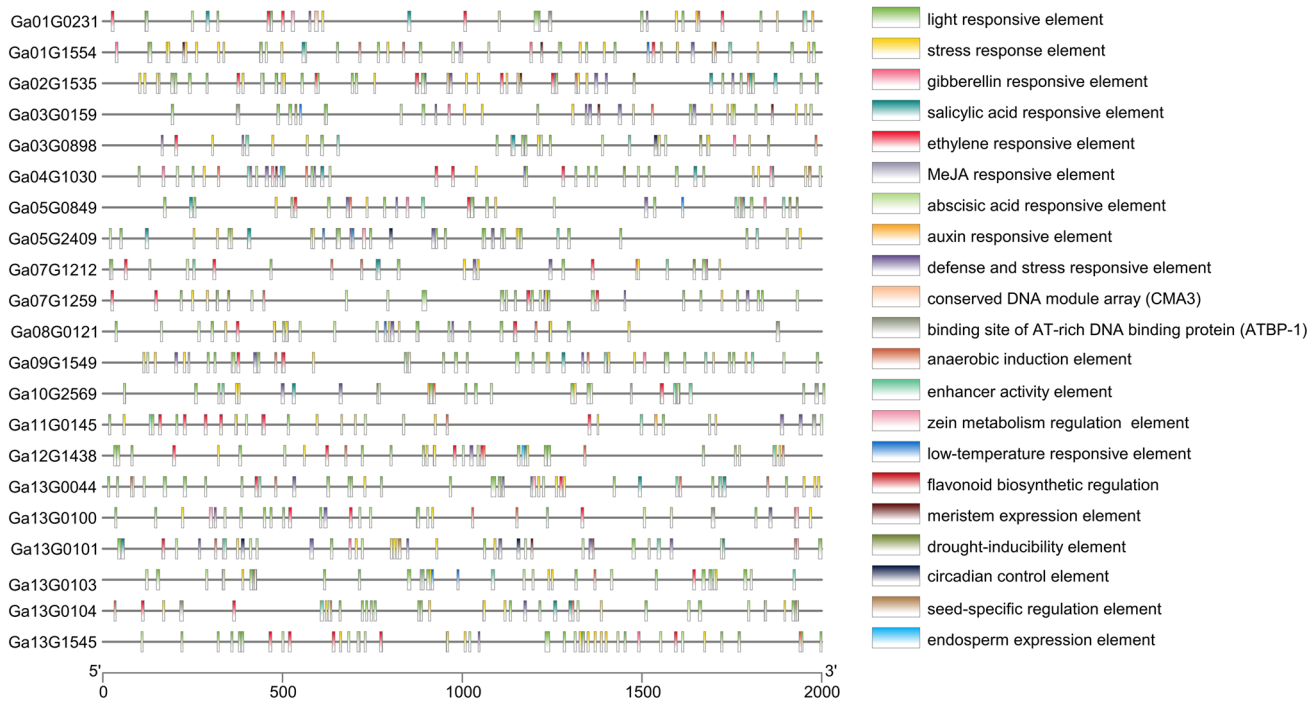



**Fig. 7** Predicted *cis*-acting elements in promoter regions of *GaGIRs*

levels may be influenced by allele dosage effects in polyploids, we designed sequence-specific primers to examine the expression level of the *GaFzl* homolog (*GhFzl*) in tetraploid cotton lines. Unlike the expression pattern in diploid cotton, *GhFzl_A* and *GhFzl_D* were expressed at very low levels (Fig. 10), with no major differences among wild-type XZ142, fuzzless mutant $N_1$, and fiberless mutant XZ142FLM. Additionally, the genes encoding well-known fiber-related regulatory factors were similarly expressed in DPL971 and DPL972 (Fig. 11). For example, the *MML3*, *TTG1*, and *GL2* expression levels were unaffected by changes to *GaFzl* expression, indicating that the altered expression of *GaFzl* had no effect on the regulatory network of the MIXTA-WD40 complex and other fiber-related genes.

## Discussion

Arabidopsis is a model plant system for studying cell fate determination and differentiation (Balkunde et al. 2010; Yang and Ye 2013). Because of the considerable similarity in the underlying process, studies on Arabidopsis trichome and root hair formation may provide useful insights into cotton fiber initiation and elongation (Lee et al. 2007; Ding et al. 2014; Wan et al. 2014; Wang et al. 2019). The *GIR* genes have been identified in diverse monocotyledonous and dicotyledonous plant species. The *GIR* genes in Arabidopsis are involved in the GL2-mediated control of root hair development (Wu and Citovsky 2017a). In tomato, *SlCycB2* encodes a GIR1/2, homolog to GL2, reportedly interacting with Wo, to regulate trichome formation and development (Yang et al. 2011). Several recent studies demonstrated that *GaFzl,* annotated as *GaGIR1*, is closely associated with the fuzzless phenotype of diploid cotton. It encodes a negative regulator of trichome formation and fuzz development, but it does not affect root hair formation (Du et al. 2018; Feng et al. 2019; Liu et al. 2020; Wang et al. 2020c). These studies suggest that the functions of these homologs may have diverged slightly as the different species evolved. However, very few of the *GIR* genes in cotton have been identified and functionally annotated. Thus, it is important that these genes in cotton are identified so their expression patterns and functions related to cotton fiber development may be determined.

On the basis of a whole-genome analysis, we identified 21, 21, and 40 *GIR* family genes in *G. arboreum*, *G. raimondii* and *G. hirsutum*, respectively. The *GaGIRs* genes, which were divided into three groups, were unevenly distributed on 12 of 13 chromosomes. Additionally, the encoded proteins varied substantially in terms of amino acid sequence length and molecular weight. Analyses of the phylogenetic relationships, structures, and motifs revealed that most of the GIR proteins in the same subgroup were similar regarding the organization of exons and conserved motifs, whereas obvious differences were detected between subgroups. These differences may reflect some functional diversification during

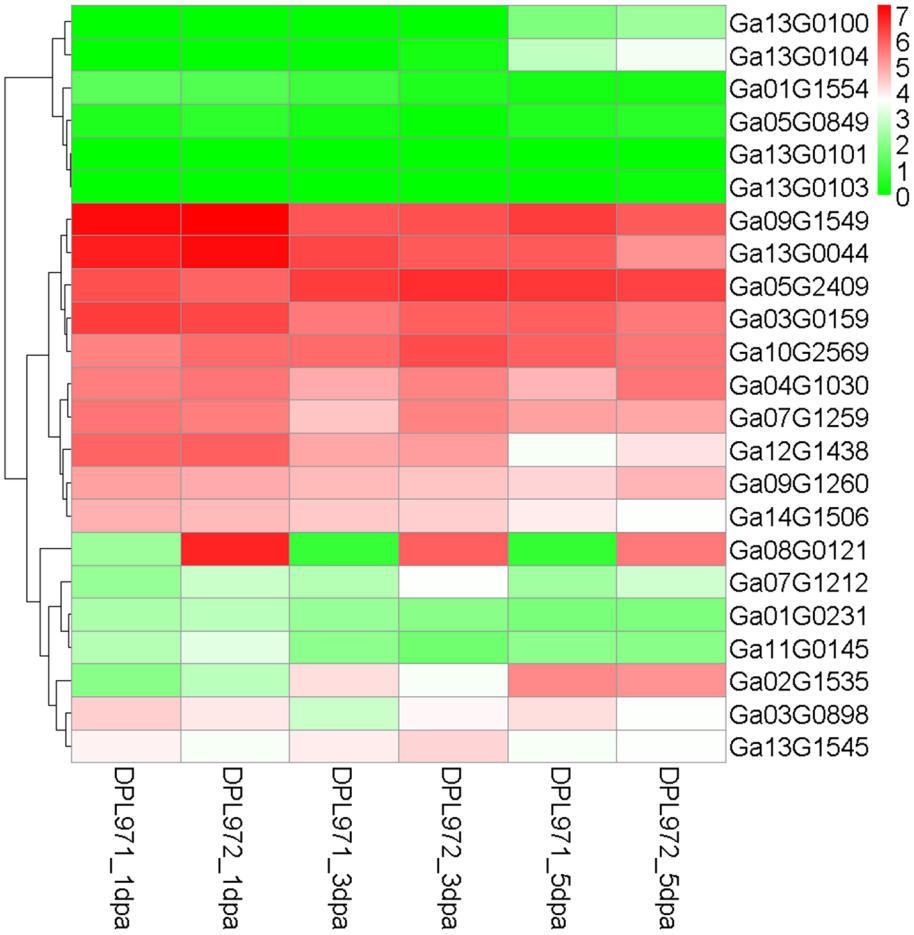
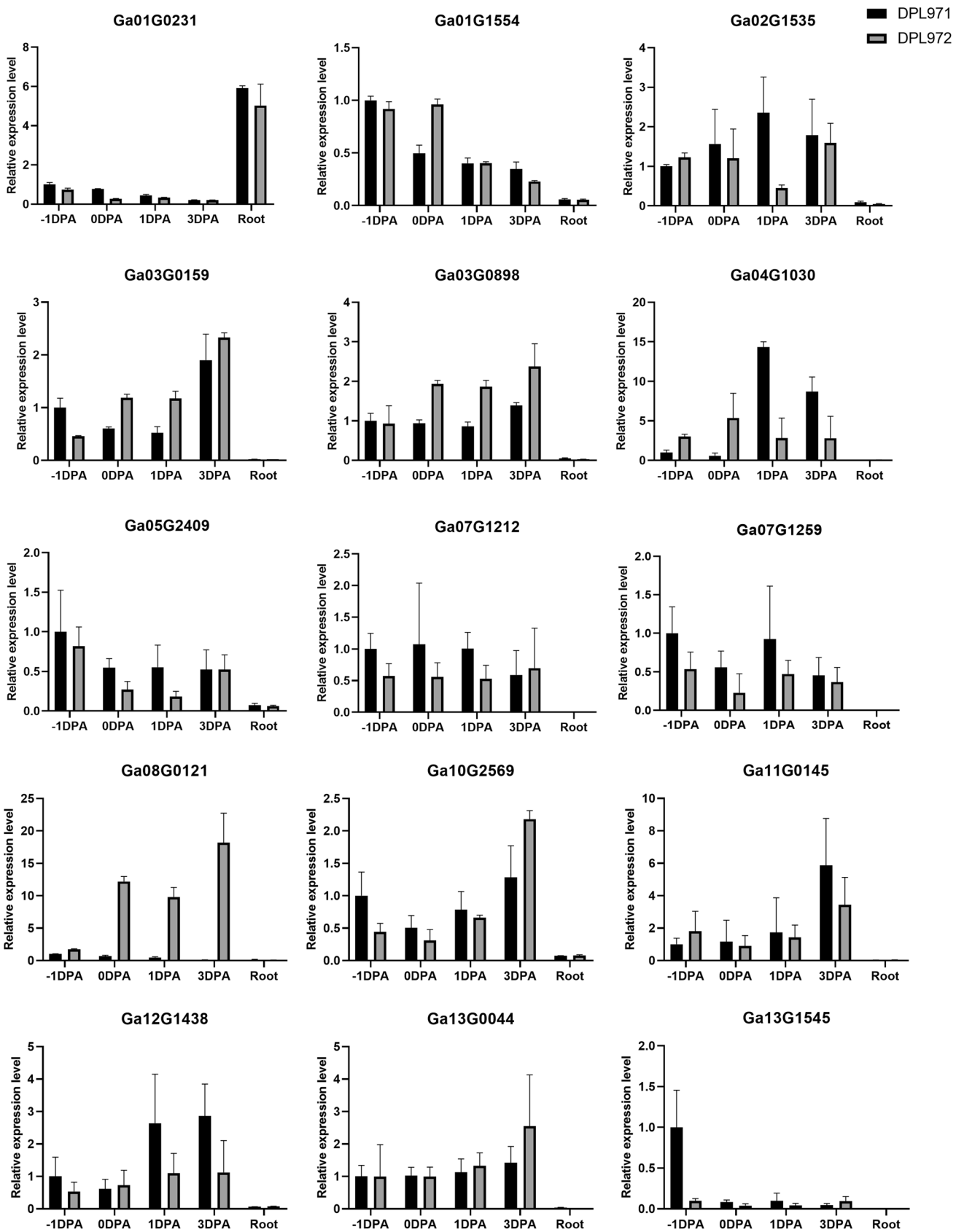



**Fig. 8** RNA-seq data heat map of *GaGIR* gene expression levels at different stages of fiber development. The differences in gene expression are shown in different colors. *DPA* day post anthesis

evolution. Moreover, some conserved motifs were identified in all GIR proteins, whereas others were subgroup-specific, suggesting that motif diversity might be related to functional diversity. The gain or loss of key motifs as species evolved may have resulted in changes to protein functions that altered plant development. Phytohormones and abiotic stresses might affect fiber development through signal transduction pathways (Chen et al. 2019; He et al. 2019; Cheng et al. 2020; Wang et al. 2020a, b; Zhang et al. 2020; Tian and Zhang 2021; Wu et al. 2021). We identified numerous *cis*-acting elements responsive to light, phytohormones, defense processes, and stress as well as multiple TF-binding sites upstream of transcriptional start sites. Accordingly, GIR proteins may participate in signal transduction pathways that help regulate fiber and plant development.

We compared the *GaGIRs* expression patterns between wild-type and fuzzless mutant *G. arboreum* at different stages of fiber development. Some of the genes were highly expressed at multiple fiber developmental stages, indicating they may positively regulate fiber formation. In contrast, some *GaGIR* genes were expressed at relatively low levels, suggesting that they are not directly involved in fiber development or they may negatively regulate fiber formation. There were no obvious differences in the expression of most of these genes between the wild-type and mutant cotton. Interestingly, *GaFzl* was significantly expressed in the fuzzless mutant but not in the wild-type control, during the fiber and fuzz initiation stages (0, 1, 3 DPA) (Fig. 10). This is in accordance with the results of previous studies (Feng et al. 2019; Liu et al. 2020). This gene was annotated as *GaGIR1*, and its upregulated expression is associated with the fuzzless phenotype in *G. arboreum*. The significant difference in the expression of this gene between the wild-type and mutant cotton plants may be caused by a ~ 6.2-kb insertion (Copy Number Variation) in the upstream region. Our analyses of phylogenetic relationships and gene expression patterns indicated that five Group 1 members that were clustered with *GaFzl* were expressed at low levels during the fiber and fuzz initiation stages. We hypothesize that overexpressing

**Fig. 9** Relative expression levels of *GaGIR* genes in *G. arboreum* ▶ wild-type DPL971 and fuzzless isogenic mutant DPL972 at different stages of cotton fiber development. *GhHis3* was applied as the internal control. The expression value of fiber samples in DPL971 at − 1 DPA was set as 1. Data are presented as mean ± standard deviation ($n = 3$)


DPL971
DPL972
Ga01G0231
Ga01G1554
Ga02G1535
Ga03G0159
Ga03G0898
Ga04G1030
Ga05G2409
Ga07G1212
Ga07G1259
Ga08G0121
Ga10G2569
Ga11G0145
Ga12G1438
Ga13G0044
Ga13G1545
Relative expression level
-1DPA
0DPA
1DPA
3DPA
Root

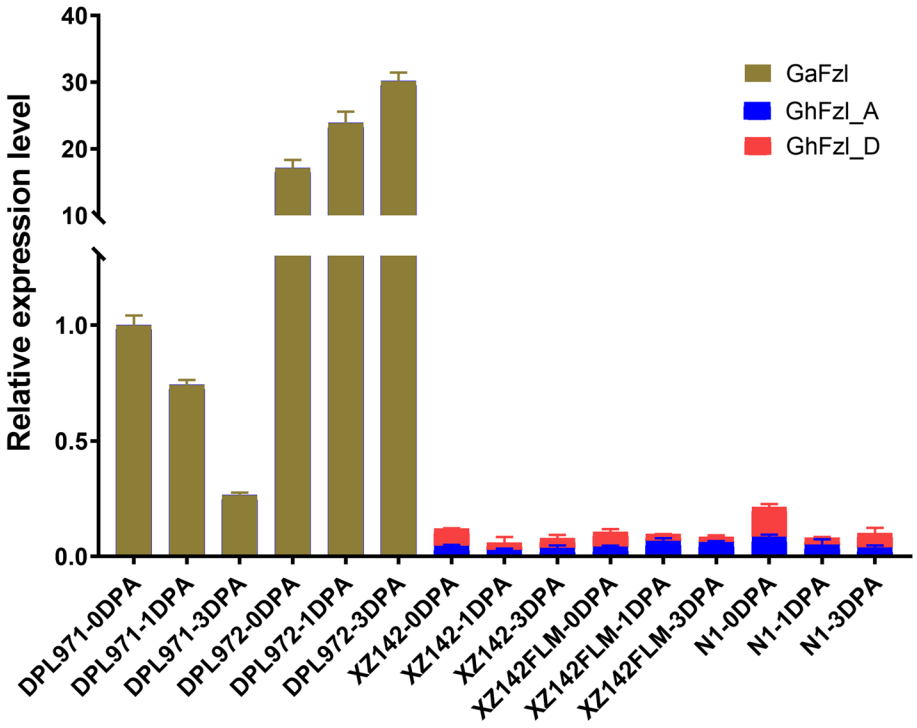



**Fig. 10** Expression patterns of *GaFzl* in *G. arboreum* DPL971 and DPL972, in *G. hirsutum* $N_1$, XZ142 and XZ142FLM. *GhHis3* was applied as the internal control. The expression value of fiber samples in DPL971 at − 1 DPA was set as 1. Data are presented as mean ± standard deviation ($n = 3$)

the genes classified in the same group as *GaFzl* with the same expression patterns may similarly affect the fuzz initiation stage, ultimately resulting in the fuzzless phenotype. Because the functions of most of these genes are unknown, they will need to be more thoroughly investigated in future studies.

There were no obvious differences in *GhFzl* expression among *G. hirsutum* XZ142, the fuzzless mutant $N_1$, and the lintless–fuzzless mutant XZ142FLM, which is in contrast to the corresponding expression patterns in *G. arboreum*. Furthermore, there was no significant difference between *G. arboreum* DPL971 and DPL972 regarding the expression patterns of the core genes involved in fiber and fuzz development (i.e., *MML3*, *TTG1*, and *CPC*). Thus, the mechanism regulating *GaFzl* expression in *G. arboreum* may not be associated with the well-known MYB-bHLH-WD40 complex involved in the fiber/fuzz initiation and development of *G. hirsutum*. This finding may lead to the development of new ways to specifically regulate fiber/fuzz formation in the initial steps of this process in *G. arboreum*.

In previous studies, *GaFzl* was identified as a candidate gene controlling fuzz initiation and development in *G. arboreum*. Liu et al. (2020) suggested that the variability in the sequences of the *GaGIR1* haplotypes in diverse fuzzless mutants may be related to the changes in expression patterns or gene functions, while Wang et al. (2020c) identified an enhancer as the essential element for controlling gene expression and fuzz development. To explore the potential function of *GaFzl* during fuzz initiation, we conducted a transcriptional activation analysis in yeast, which demonstrated that GaFzl has strong transcriptional activation activity and may act as a transcriptional activator. However, the fusion proteins constructed from two materials revealed minor differences in the transcriptional activation activity in yeast cells. Because of the differences in the *GaFzl* coding sequence, we speculated that sequence variations between two parental lines may be responsible for the differences in transcriptional activation activities or even gene functions. This possibility will need to be experimentally verified in future studies.

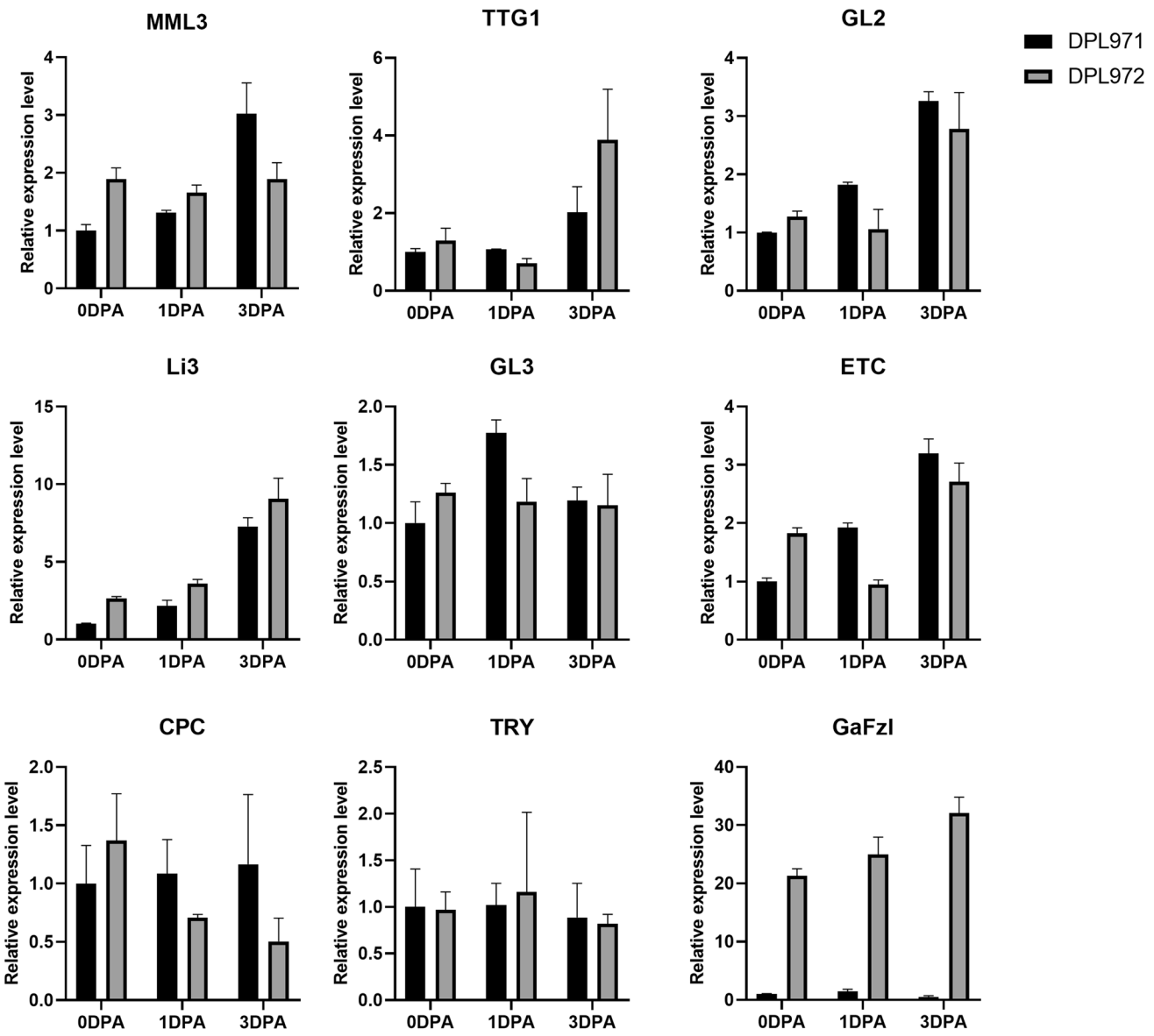



**Fig. 11** Expression analysis of several core genes in *G. arboreum* DPL971 and DPL972 in upland cotton fiber development. *GhHis3* was applied as the internal control. The expression value of fiber samples in DPL971 at − 1 DPA was set as 1. Data are presented as mean ± standard deviation ($n = 3$)

***Author contribution statement*** Research conception and design, GLS and JY; Data collection, XXF; Data analysis, XXF, HLC, DYZ, YPZ, QLW, LML and SYL; Research management, GLS; Writing—initial draft, XXF; Writing—revised manuscript, JY. All authors read and approved the final manuscript.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00425-021-03737-7.

**Acknowledgements** The authors thank Lima Soares Emanoella from Plant Genetics Laboratory of Gembloux Agro Bio-Tech, University of Liège for kindly revising the draft.

**Funding** This work was funded by grants from the National Natural Science Foundation of China (No. 31621005), the National Natural Science Foundation of China (No. 31901581), the National Key R & D Plan of China (No. 2018YFD0100402) and the United States Department of Agriculture—Agricultural Research Service (USDA-ARS Project No. 3091-21000044-00D). The funders had no role in the experimental design, data collection and analysis, or in writing work.

**Data availability statements** All related datasets supporting the results of this study are available within the manuscript and its supplementary files.

## Declarations

**Conflict of interest** The authors declare that they have no conflicts of interest.

## References


Balkunde R, Pesch M, Hülskamp M (2010) Trichome patterning in *Arabidopsis thaliana*: from genetic to molecular models. Curr Top Dev Biol 91:299–321

Chen Y, Chen B, Wang H, Hu W, Wang S, Zhou Z (2019) Combined elevated temperature and soil waterlogging stresses limit fibre biomass accumulation and fibre quality formation by disrupting protein activity during cotton fibre development. Funct Plant Biol 46(8):715–724

Chen W, Li Y, Zhu S, Fang S, Zhao L, Guo Y, Wang J, Yuan L, Lu Y, Liu F, Yao J, Zhang Y (2020) A retrotransposon insertion in *GhMML3_D12* is likely responsible for the lintless locus *li3* of tetraploid cotton. Front Plant Sci 11:593679. https://doi.org/10.3389/fpls.2020.593679

Cheng H, Lu C, John ZY, Zou C, Zhang Y, Wang Q, Huang J, Feng X, Jiang P, Yang W (2016) Fine mapping and candidate gene analysis of the dominant glandless gene Gl2e in cotton (*Gossypium* spp). Theor Appl Genet 129(7):1347–1355

Cheng G, Zhang L, Wei H, Wang H, Lu J, Yu S (2020) Transcriptome analysis reveals a gene expression pattern associated with fuzz fiber initiation induced by high temperature in *Gossypium barbadense*. Genes (basel) 11(9):1066

Dai X, Zhou L, Zhang W, Cai L, Guo H, Tian H, Schiefelbein J, Wang S (2016) A single amino acid substitution in the R3 domain of GLABRA1 leads to inhibition of trichome formation in Arabidopsis without affecting its interaction with GLABRA3. Plant Cell Environ 39(4):897–907

Di Cristina M, Sessa G, Dolan L, Linstead P, Baima S, Ruberti I, Morelli G (1996) The *Arabidopsis* Athb-10 (GLABRA2) is an HD-Zip protein required for regulation of root hair development. Plant J 10(3):393–402

Ding M, Jiang Y, Cao Y, Lin L, He S, Zhou W, Rong J (2014) Gene expression profile analysis of Ligon lintless-1 (*Li1*) mutant reveals important genes and pathways in cotton leaf and fiber development. Gene 535(2):273–285

Ding M, Cao Y, He S, Sun J, Dai H, Zhang H, Sun C, Jiang Y, Paterson AH, Rong J (2020) *GaHD1*, a candidate gene for the *Gossypium arboreum* SMA-4 mutant, promotes trichome and fiber initiation by cellular $H_2O_2$ and $Ca^{2+}$ signals. Plant Mol Biol 103(4–5):409–423

Du X, Huang G, He S, Yang Z, Sun G, Ma X, Li N, Zhang X, Sun J, Liu M, Jia Y, Pan Z, Gong W, Liu Z, Zhu H, Ma L, Liu F, Yang D, Wang F, Fan W, Gong Q, Peng Z, Wang L, Wang X, Xu S, Shang H, Lu C, Zheng H, Huang S, Lin T, Zhu Y, Li F (2018) Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. Nat Genet 50(6):796–802

Fang L, Gong H, Hu Y, Liu C, Zhou B, Huang T, Wang Y, Chen S, Fang DD, Du X (2017) Genomic insights into divergence and dual domestication of cultivated allotetraploid cottons. Genome Biol 18(1):33

Fang DD, Naoumkina M, Thyssen GN, Bechere E, Li P, Florane CB (2020) An EMS-induced mutation in a tetratricopeptide repeat-like superfamily protein gene (*Ghir_A12G008870*) on chromosome A12 is responsible for the *li y* short fiber phenotype in cotton. Theor Appl Genet 133(1):271–282

Feng X, Cheng H, Zuo D, Zhang Y, Wang Q, Liu K, Ashraf J, Yang Q, Li S, Chen X (2019) Fine mapping and identification of the fuzzless gene *GaFzl* in DPL972 (*Gossypium arboreum*). Theor Appl Genet 132(8):2169–2179

Gao Z, Sun W, Wang J, Zhao C, Zuo K (2019) *GhbHLH18* negatively regulates fiber strength and length by enhancing lignin biosynthesis in cotton fibers. Plant Sci 286:7–16

Gilbert MK, Kim HJ, Tang Y, Naoumkina M, Fang DD (2014) Comparative transcriptome analysis of short fiber mutants Ligon-lintless 1 and 2 reveals common mechanisms pertinent to fiber elongation in cotton (*Gossypium hirsutum* L.). PLoS ONE 9(4):95554

He P, Zhang Y, Liu H, Yuan Y, Wang C, Yu J, Xiao G (2019) Comprehensive analysis of *WOX* genes uncovers that *WOX13* is involved in phytohormone-mediated fiber development in cotton. BMC Plant Biol 19(1):312

Hinchliffe DJ, Turley RB, Naoumkina M, Kim HJ, Tang Y, Yeater KM, Li P, Fang DD (2011) A combined functional and structural genomics approach identified an EST-SSR marker with complete linkage to the Ligon lintless-2 genetic locus in cotton (*Gossypium hirsutum* L.). BMC Genom 12(1):445

Hu H, He X, Tu L, Zhu L, Zhu S, Ge Z, Zhang X (2016) GhJAZ2 negatively regulates cotton fiber initiation by interacting with the R2R3-MYB transcription factor GhMYB25-like. Plant J 88(6):921–935

Huang Y, Liu X, Tang K, Zuo K (2013) Functional analysis of the seed coat-specific gene *GbMYB2* from cotton. Plant Physiol Biochem 73:16–22

Hülskamp M (2004) Plant trichomes: a model for cell differentiation. Nat Rev Mol Cell Biol 5(6):471–480

Ioannidi E, Rigas S, Tsitsekian D, Daras G, Alatzas A, Makris A, Tanou G, Argiriou A, Alexandrou D, Poethig S (2016) Trichome patterning control involves TTG1 interaction with SPL transcription factors. Plant Mol Biol 92(6):675–687

Jiang Y, Ding M, Cao Y, Yang F, Zhang H, He S, Dai H, Hao H, Rong J (2015) Genetic fine mapping and candidate gene analysis of the *Gossypium hirsutum* Ligon lintless-1 (Li1) mutant on chromosome 22 (D). Mol Genet Genomics 290(6):2199–2211

Kim HJ, Hinchliffe DJ, Triplett BA, Chen ZJ, Stelly DM, Yeater KM, Moon HS, Gilbert MK, Thyssen GN, Turley RB (2015) Phytohormonal networks promote differentiation of fiber initials on preanthesis cotton ovules grown in vitro and in planta. PLoS ONE 10(4):e0125046

Lee JJ, Woodward AW, Chen ZJ (2007) Gene expression changes and early events in cotton fibre development. Ann Bot 100(7):1391–1401

Liang W, Fang L, Xiang D, Hu Y, Feng H, Chang L, Zhang T (2015) Transcriptome analysis of short fiber mutant Ligon lintless-1 (Li1) reveals critical genes and key pathways in cotton fiber elongation and leaf development. PLoS ONE 10(11):e0143503

Liu B, Zhu Y, Zhang T (2015) The R3-MYB gene *GhCPC* negatively regulates cotton fiber elongation. PLoS ONE 10(2):e0116272

Liu X, Moncuquet P, Zhu Q-H, Stiller W, Zhang Z, Wilson I (2020) Genetic identification and transcriptome analysis of lintless and fuzzless traits in *Gossypium arboreum* L. Int J Mol Sci 21(5):1675

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25(4):402–408

Matías-Hernández L, Aguilar-Jaramillo AE, Cigliano RA, Sanseverino W, Pelaz S (2016) Flowering and trichome development share hormonal and transcription factor regulation. J Exp Bot 67(5):1209–1219

Mei G, Zhang Z (2019) Optimization of polar distribution of *GhPIN3a* in the ovule epidermis improves cotton fiber development. J Exp Bot 70(12):3021

Naoumkina M, Thyssen GN, Fang DD, Bechere E, Li P, Florane CB (2021a) Mapping-by-sequencing the locus of EMS-induced mutation responsible for tufted-fuzzless seed phenotype in cotton. Mol Genet Genomics 296(5):1041–1049

Naoumkina M, Thyssen GN, Fang DD, Li P, Florane CB (2021b) Elucidation of sequence polymorphism in fuzzless-seed cotton lines. Mol Genet Genomics 296(1):193–206

Ohashi Y, Oka A, Ruberti I, Morelli G, Aoyama T (2002) Entopically additive expression of *GLABRA2* alters the frequency and spacing of trichome initiation. Plant J 29(3):359–369

Ohashi Y, Oka A, Rodrigues-Pousada R, Possenti M, Ruberti I, Morelli G, Aoyama T (2003) Modulation of phospholipid signaling by GLABRA2 in root-hair pattern formation. Science 300(5624):1427–1430

Padmalatha KV, Patil DP, Kumar K, Dhandapani G, Kanakachari M, Phanindra ML, Kumar S, Mohan T, Jain N, Prakash AH (2012) Functional genomics of fuzzless-lintless mutant of *Gossypium hirsutum* L. cv. MCU5 reveal key genes and pathways involved in cotton fibre initiation and elongation. BMC Genom 13(1):624

Patel JD, Huang X, Lin L, Das S, Chandnani R, Khanal S, Adhikari J, Shehzad T, Guo H, Roy-Zokan EM (2020) The Ligon lintless-2 short fiber mutation is located within a terminal deletion of chromosome 18 in cotton. Plant Physiol 183(1):277–288

Pattanaik S, Patra B, Singh SK, Yuan L (2014) An overview of the gene regulatory network controlling trichome development in the model plant. Arabidopsis Front Plant Sci 5:259

Rerie WG, Feldmann KA, Marks MD (1994) The *GLABRA2* gene encodes a homeo domain protein required for normal trichome development in Arabidopsis. Gene Dev 8(12):1388–1399

Rong J, Pierce GJ, Waghmare VN, Rogers CJ, Desai A, Chee PW, May OL, Gannaway JR, Wendel JF, Wilkins TA (2005) Genetic mapping and comparative analysis of seven mutants related to seed fiber development in cotton. Theor Appl Genet 111(6):1137–1146

Shan C, Shangguan X, Zhao B, Zhang X, Chao L, Yang C, Wang L, Zhu H, Zeng Y, Guo W (2014) Control of cotton fibre elongation by a homeodomain transcription factor GhHOX3. Nat Commun 5:5519

Sun R, Li C, Zhang J, Li F, Ma L, Tan Y, Wang Q, Zhang B (2017) Differential expression of microRNAs during fiber development between fuzzless-lintless mutant and its wild-type allotetraploid cotton. Sci Rep 7(1):3. https://doi.org/10.1038/s41598-017-00038-6

Sun H, Hao P, Gu L, Cheng S, Wang H, Wu A, Ma L, Wei H, Yu S (2020a) Pectate lyase-like gene *GhPEL76* regulates organ elongation in *Arabidopsis* and fiber elongation in cotton. Plant Sci 293:110395

Sun Q, Huang J, Guo Y, Yang M, Guo Y, Li J, Zhang J, Xu W (2020b) A cotton NAC domain transcription factor, GhFSN5, negatively regulates secondary cell wall biosynthesis and anther development in transgenic Arabidopsis. Plant Physiol Biochem 146:303–314

Thyssen GN, Fang DD, Turley RB, Florane CB, Li P, Mattison CP, Naoumkina M (2017) A Gly65Val substitution in an actin, GhACT_LI1, disrupts cell polarity and F-actin organization resulting in dwarf, lintless cotton plants. Plant J 90(1):111–121

Tian Y, Zhang T (2021) MIXTAs and phytohormones orchestrate cotton fiber development. Curr Opin Plant Biol 59:101975

Walford SA, Wu Y, Llewellyn DJ, Dennis ES (2012) Epidermal cell differentiation in cotton mediated by the homeodomain leucine zipper gene, *GhHD-1*. Plant J 71(3):464–478

Wan Q, Zhang H, Ye W, Wu H, Zhang T (2014) Genome-wide transcriptome profiling revealed cotton fuzz fiber development having a similar molecular model as *Arabidopsis* trichome. PLoS ONE 9(5):e97313

Wan Q, Guan X, Yang N, Wu H, Pan M, Liu B, Fang L, Yang S, Hu Y, Ye W (2016) Small interfering RNAs from bidirectional transcripts of *GhMML3_A12* regulate cotton fiber development. New Phytol 210(4):1298–1310

Wang L, Zhu Y, Hu W, Zhang X, Cai C, Guo W (2015) Comparative transcriptomics reveals jasmonic acid-associated metabolism related to cotton fiber initiation. PLoS ONE 10(6):e0129854

Wang Y, Yu Y, Chen Q, Bai G, Gao W, Qu Y, Ni Z (2019) Heterologous expression of *GbTCP4*, a Class II TCP transcription factor, regulates trichome formation and root hair development in Arabidopsis. Genes 10(9):726

Wang L, Cheng H, Xiong F, Ma S, Zheng L, Song Y, Deng K, Wu H, Li F, Yang Z (2020a) Comparative phosphoproteomic analysis of BR-defective mutant reveals a key role of GhSK13 in regulating cotton fiber development. Sci China Life Sci 63(12):1905–1917

Wang L, Wang G, Long L, Altunok S, Feng Z, Wang D, Khawar KM, Mujtaba M (2020b) Understanding the role of phytohormones in cotton fiber development through omic approaches; recent advances and future directions. Int J Biol Macromol 163:1301–1313

Wang X, Miao Y, Cai Y, Sun G, Jia Y, Song S, Pan Z, Zhang Y, Wang L, Fu G, Gao Q, Ji G, Wang P, Chen B, Peng Z, Zhang X, Wang X, Ding Y, Hu D, Geng X, Wang L, Pang B, Gong W, He S, Du X (2020c) Large-fragment insertion activates gene *GaFZ* (*Ga08G0121*) and is associated with the fuzz and trichome reduction in cotton (*Gossypium arboreum*). Plant Biotechnol J 19(6):1110–1124

Wei Z, Li J (2018) Receptor-like protein kinases: Key regulators controlling root hair development in *Arabidopsis thaliana*. J Integr Plant Biol 60(9):841–850

Wu R, Citovsky V (2017a) Adaptor proteins GIR1 and GIR2. I. Interaction with the repressor GLABRA2 and regulation of root hair development. Biochem Biophys Res Commun 488(3):547–553

Wu R, Citovsky V (2017b) Adaptor proteins GIR1 and GIR2. II. Interaction with the co-repressor TOPLESS and promotion of histone deacetylation of target chromatin. Biochem Biophys Res Commun 488(4):609–613

Wu H, Tian Y, Wan Q, Fang L, Guan X, Chen J, Hu Y, Ye W, Zhang H, Guo W (2018) Genetics and evolution of *MIXTA* genes regulating cotton lint fiber development. New Phytol 217(2):883–895

Wu H, Zheng L, Qanmber G, Guo M, Wang Z, Yang Z (2021) Response of phytohormone mediated plant homeodomain (PHD) family to abiotic stress in upland cotton (*Gossypium hirsutum* spp.). BMC Plant Biol 21(1):13

Yang C, Ye Z (2013) Trichomes as models for studying plant cell differentiation. Cell Mol Life Sci 70(11):1937–1948

Yang C, Li H, Zhang J, Luo Z, Gong P, Zhang C, Li J, Wang T, Zhang Y, Ye Z (2011) A regulatory gene induces trichome formation and embryo lethality in tomato. Proc Natl Acad Sci USA 108(29):11836–11841

Zeng J, Zhang M, Hou L, Bai W, Yan X, Hou N, Wang H, Huang J, Zhao J, Pei Y (2019) Cytokinin inhibits cotton fiber initiation by disrupting PIN3a-mediated asymmetric accumulation of auxin in the ovule epidermis. J Exp Bot 70(12):3139–3151

Zhang G, Yue C, Lu T, Sun L, Hao F (2020) Genome-wide identification and expression analysis of NADPH oxidase genes in response to ABA and abiotic stresses, and in fibre formation in *Gossypium*. Peer J 8:e8404

Zhu J, Chen J, Gao F, Xu C, Wu H, Chen K, Si Z, Yan H, Zhang T (2017) Rapid mapping and cloning of the *virescent-1* gene in cotton by bulked segregant analysis–next generation sequencing and virus-induced gene silencing strategies. J Exp Bot 68(15):4125–4135

Zhu Q, Yuan Y, Stiller W, Jia Y, Wang P, Pan Z, Du X, Llewellyn D, Wilson I (2018) Genetic dissection of the fuzzless seed trait in *Gossypium barbadense*. J Exp Bot 69(5):997–1009

Zhu Q, Stiller W, Moncuquet P, Gordon S, Yuan Y, Barnes S, Wilson I (2021) Genetic mapping and transcriptomic characterization of a new fuzzless-tufted cottonseed mutant. G3 (bethesda) 11(1):1–14