RESEARCH PAPER

The transcription factor MML4_D12 regulates fiber development through interplay with the WD40-repeat protein WDR in cotton

Yue Tian1,2, Jingjing Du1,2, Huaitong Wu1, Xueying Guan1,2, Weihang Chen1, Yan Hu1,2, Lei Fang1,2, Linyun Ding1, Menglin Li1, Duofeng Yang1, Qinli Yang1 and Tianzhen Zhang1,2, *

1National Key Laboratory of Crop Genetics and Germplasm Enhancement, Cotton Research Institute, Nanjing Agricultural University, Nanjing 210095, P. R. China
2Zhejiang Provincial Key Laboratory of Crop Genetic Resources, Institute of Crop Science, College of Agriculture and Biotechnology, Zhejiang University, 310029 Zhejiang, China

* Correspondence: cotton@njau.edu.cn

Received 17 August 2019; Editorial decision 14 February 2020; Accepted 28 February 2020

Abstract

In planta, a vital regulatory complex, MYB–basic helix–loop–helix (bHLH)–WD40 (MBW), is involved in trichome development and synthesis of anthocyanin and proanthocyanin in Arabidopsis. Usually, WD40 proteins provide a scaffold for protein–protein interaction between MYB and bHLH proteins. Members of subgroup 9 of the R2R3 MYB transcription factors, which includes MYBMIXTA-Like (MML) genes important for plant cell differentiation, are unable to interact with bHLH. In this study, we report that a cotton (Gossypium hirsutum) seed trichome or lint fiber-related GhMML factor, GhMML4_D12, interacts with a diverged WD40 protein (GhWDR) in a process similar to but different from that of the MBW ternary complex involved in Arabidopsis trichome development. Amino acids 250–267 of GhMML4_D12 and the first and third WD40 repeat domains of GhWDR determine their interaction. GhWDR could rescue Arabidopsis ttg1 to its wild type, confirming its orthologous function in trichome development. Our findings shed more light towards understanding the key role of the MML and WD40 families in plants and in the improvement of cotton fiber production.

Keywords: Complex, cotton, fiber development, GhMML4_D12, GhWDR, trichome.

Introduction

MYB transcription factors, especially R2R3 MYB, have a diverse range of functions in different plant species (Stracke et al., 2001). Among them, subgroup 9, which includes those encoded by MIXTA genes, are important for the specification and regulation of plant cellular differentiation (Perez-Rodriguez et al., 2005; Scoville et al., 2011; Oshima et al., 2013; Wu et al., 2018; Yan et al., 2018). The first MIXTA gene, characterized in snapdragon (Antirrhinum majus), was found to control the development of conical cell shape in the petal epidermis (Noda et al., 1994). It also appears to be capable of driving the initiation of conical epidermal cells from flat epidermal cells (Glover et al., 1998). The MYBMIXTA-Like (MML) transcription factors...
MYB106 and MYB16 regulate epidermal cell morphology and cuticle development in Arabidopsis and Torenia fournieri (Baumann et al., 2007; Oshima et al., 2013). In Artemisia annua, interaction between AaHD8 and AaMIXTA1 has a role in the formation of glandular trichomes and is involved in cuticle development (Yan et al., 2017; Shi et al., 2018). The R2R3 MYB transcription factors can form an MYB–basic helix–loop–helix (bHLH)–WD40 (MBW) complex to modulate diverse biological processes such as trichome development, cell death, cell wall synthesis, hormone signaling, stamen development, and seed production (Ramsay and Glover, 2005; Schaar et al., 2013; Xie et al., 2016). However, the MML transcription factors lack the core amino acids that have been reported to play a decisive role in interactions with bHLH to form the MBW complex. How MML genes regulate the specification of plant cellular differentiation remains to be explored.

In allotetraploid cotton, transcription factors encoded by 10 GhMML homologous genes contain the signature protein motif AQWESARxxAExRLxRES. They are all expressed during fiber initiation in cultivated cotton, but have lower expression in fiberless mutants (Zhang et al., 2015). Using a map-based cloning strategy, we isolated and identified GhMML3 on chromosome A12 (GhMML3_A12), responsible for fuzz fiber production (Wan et al., 2016), and GhMML4 on chromosome D12 (GhMML4_D12), responsible for development of lint fiber (Li et al.), the largest source of natural textile in the world (Wu et al., 2018). A mutation that occurred at the 500 bp site from cytosine (C) in the linted–fuzzless or naked seed mutant (n2NSM) to adenine (A) in the Xuzhou142 fuzzless–lintless mutant (XZ142FLM) has resulted in the production of a TAA stop codon and early termination of GhMML4_D12 translation (Wu et al., 2018). However, the mechanisms of how these GhMML genes control cotton fiber development are still unknown.

In this study, through a yeast two-hybrid (Y2H) assay, a new kind of WD40 repeat protein (GhWDR) was identified as an interacting partner of GhMML4_D12 from the naked seed mutant n2NSM, but not the mutant GhMML4_D12 (GhMML4_D12m) derived from XZ142FLM (Fig. 1). We found for the first time a direct interaction between the MYB and WD40 transcription factors, and the GhMML4–GhWDR complex showed similarity to, but is different from, an MYB and WD40 transcription factors, and the GhMML4–GhWDR complex showed similarity to, but is different from, an MYB–basic helix–loop–helix (bHLH)–WD40 (MBW) complex based on its bHLH domain similarity, and a WD40 repeat domain of GhWDR is important for its function in Arabidopsis. GhMML4_D12 could enhance its own transcriptional activity in n2NSM, but not in XZ142FLM.

Methods

Plant materials

Gossypium hirsutum acc. Texas Marker-1 (TM-1), the fuzzless or naked seed mutant n2NSM, and the fuzzless–lintless mutant XZ142FLM were used to investigate fiber development. All plants were grown at the Dangtu Breeding Station (DBS/NAU) with normal field practices, and in the glasshouses of Nanjing Agriculture University. Arabidopsis seeds were sterilized with 20% bleach, plated on Murashige and Skoog medium (MS; Sigma-Aldrich), chilled at 4 °C for 3 d, and transferred to a growth room with a 16 h (22–24 °C)/8 h (19 °C) light/dark photoperiod.

Nicotiana benthamiana was grown under a 16 h (28 °C)/8 h (22 °C) light/dark photoperiod.

Quantitative reverse transcription polymerase chain reaction analysis

Total RNA was isolated from cotton tissues and purified using a Qiagen RNeasy kit (Qiagen) according to the manufacturer’s instructions. Expression profiling of genes in cotton tissues was carried out by quantitative reverse transcription polymerase chain reaction (qRT-PCR) using a cotton polyubiquitin gene, GhUBI1 (EU604080), as a standard control according to a previously described method (Li et al., 2005). In brief, cDNA was synthesized from total RNA and used as a template in qRT-PCR with gene-specific primers (see Supplementary Table S1 at JXB online). The PCR was performed using SYBR–Green Real-time PCR Master Mix according to the manufacturer’s instructions (Toyobo Co. Ltd, Osaka, Japan), and the relative quantity of the target gene expression level was determined. The mean value and standard deviation of three biological replicates were calculated.

Subcellular localization of fusion protein

The full lengths of GhMML4_D12, GhMML4_D12m, GhWDR, N-terminal-truncated domain (GhWDR-N), WD40 repeat domain (GhWDR-R), and C-terminal-truncated domain (GhWDR-C) without a stop codon were inserted into the binary vector pBINGFP4 (Hellens et al., 2000), upstream of the green fluorescent protein (GFP) sequence, to produce the constructs 35S_GhMML4_D12, 35S_GhMML4_D12m, 35S_GhWDR, 35S_GhWDR-N, 35S_GhWDR-R, and 35S_GhWDR-C. The constructs were introduced into Agrobacterium tumefaciens strain GV3101 by electroporation. Subcellular localization of

![Fig. 1. Phenotypic appearance of cotton fiber and structure of the GhMML4_D12 gene between n2NSM and XZ142FLM. GhMML4_D12 is from the linted–fuzzless accession (n2NSM) and the mutant GhMML4_D12 gene (GhMML4_D12m) is derived from our Xuzhou142 fuzzless–lintless mutant (XZ142FLM). A mutation occurred at the 500 bp site, from cytosine (C) in n2NSM to adenine (A) in XZ142FLM, which resulted in the production of a TAA stop codon and early termination of GhMML4_D12 translation. Scale bar: 1.0 cm.](image-url)
all these fusion proteins was determined in tobacco (Nicotiana benthamiana) leaf epidermal cells by A. tumefaciens infiltration (Liu et al., 2014). GFP fluorescence in the tobacco epidermal cell was observed using a Zeiss LSM 710 confocal microscope (Zeiss Microsystems) with a filter set of 466 nm for excitation and 506–538 nm for emission. Zen 2009 software (Zeiss) was used to record and process the digital images taken.

Transcription activation analysis
To investigate the transcriptional activity of the GhWDR protein, the coding sequences (CDS) of GhWDR, GhWDR-N, GhWDR-R, and GhWDR-C were cloned into pGBK7 vectors (Clontech, Palo Alto, CA, USA) and transformed into yeast strain AH109 using the high-efficiency lithium acetate transformation procedure. Yeast transformants were streaked on selective medium lacking tryptophan and adenine added to assay transcriptional activity. β-Galactosidase activity was also assayed by colony-lift filter assay using 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal) as the substrate.

Phylogenetic analysis
Phylogenetic analysis was conducted using two methods: neighbor-joining (NJ) and maximum likelihood (ML). NJ trees were constructed using MEGA 5.1 (Tamura et al., 2011) with 1000 bootstrap resampling and the Poisson correction model, plus the pairwise deletion option. RAXML v.8.2.4 was used to construct ML trees, with the Jones, Taylor, and Thornton (JTT) model, GTR+GAMMA distribution option and 200 non-parametric bootstrap replicates (Stamatakis, 2014).

Yeast two-hybrid screening and assays
A Y2H assay was performed using the Matchmaker GAL4 Two-Hybrid System following the manufacturer’s protocol (Clontech). Full-length cDNA of GhMML4 D12 was fused to the GAL4-DNA-binding domain of the bait vector pGBK7 and transformed into yeast strain Y2H. A cDNA library from cotton ovules and fibers (−3 to ~25 d post-anthesis (DPA)) was constructed by transforming yeast strain AH109 with ds-cDNA and the pGADT7-Rec vector according to the manufacturer’s instructions. The library host strain was mated with bait strain Y2H, and the mating mixture was then spread onto SD/-Trp/-Leu/-His medium and incubated at 30 °C for 3–4 d. Positive clones were isolated and retransformed into bait strains to test their interaction using pGBK7-p53/pGADT7 as the positive control and pGBK7-Lamin c/ pGADT7 as the negative control.

For the Y2H assay, all the GhMML4 D12 CDSs and their domain derivatives were cloned into pGBK7, and the full-length or domain deletion forms of GhWDR were cloned into pGADT7. Primers used for the vector construction were presented in Supplementary Table S1. All the constructs were sequence verified. Self-activation of all the pGBK7 fusion constructs was suppressed by different concentrations of Auroebacinidin A (AbA). The indicated construct pairs were transformed into yeast strain Y2H and cultured at 30 °C. Successful yeast transformants were cultured in liquid SD minimal medium (-Trp/-Leu/-His) at 250 rpm and 30 °C for ~24 h to the final concentration (OD600=1.2–1.5). Then 1 ml of yeast strain was centrifuged and resuspended in distilled water to the final concentration (OD600=0.4–0.6). The mating mixture was diluted in multiples of 10 and 100, and then spread onto SD/-Trp/-Leu/-His medium with different concentrations of AbA and incubated at 30 °C for 3–4 d.

Biomolecular fluorescence complementation
pSPYNE and pSPYCE plasmids have been reported previously (Waadt et al., 2008). Full-length GhMML4 D12, GhMML4 D12m, and GhWD were fused with the N- or C-terminal fragment of yellow fluorescent protein (nYFP or cYFP) to generate GhWDR-nYFP, cYFP-GhMML4 D12, and cYFP-GhMML4 D12m plasmids. Primers used for plasmid construction are presented in Supplementary Table S1.

All constructs were transformed into A. tumefaciens strain GV3101. Agrobacterium strains containing different constructs were resuspended in infiltration buffer (10 mM MES, 0.2 mM acetylsyringone, and 10 mM MgCl2) to a final concentration of OD600=0.6–0.8. Equal volumes and concentrations of different combinations of Agrobacterium strains were co-infiltrated into N. benthamiana leaves. Plants were incubated at 23 °C for 2–3 d. YFP fluorescence was detected with a Zeiss microscope (Zeiss LSM710) and analysed using ZEN software. All experiments comprised three biological replicates.

In vitro pull-down assays
The coding regions of GhMML4 D12 and GhMML4 D12m were cloned into pET-GST, and the CDS of GhWDR was fused with pET-28a. These constructs were transformed into E.coli to express glutathione S-transferase (GST)—GhMML4 D12, GST—GhMML4 D12m, and His—GhWDR proteins, respectively. The pull-down assay was performed as described previously (Messa et al., 2016). Briefly, 2 μg of purified fusion protein GST—GhMML4 D12 or GST—GhMML4 D12m was incubated with immobilized 10 μg His-GhWDR fusion protein at 4 °C for 2 h and then separated by SDS-PAGE and immunoblotted with the corresponding antibody.

Yeast one-hybrid assay
The CDSs of GhMML4 D12, GhMML4 D12m, and GhWD were ligated into the pGADT7 vector (Clontech). GhMML4 D12, GhMML4 D12m, and GhWDR promoters were ligated into the pAbA1 vector (Clontech). All primers used are listed in Supplementary Table S1. The yeast one-hybrid (Y1H) assay was conducted using the Matchmaker Gold Yeast One-Hybrid Library Screening System kit (cat. no. 630491, Clontech, USA).

Transient expression assay
A transient expression assay was performed in N. benthamiana leaves. The nuclear localization site (NLS) was ligated to the 5’ end of GhMML4 D12 and GhMML4 D12m promoters, then GFP was fused with the 3’ end of NLS to construct GhMML4 D12-NLS-GFP and GhMML4 D12m-NLS-GFP, respectively. The full-length CDSs of GhMML4 D12, GhMML4 D12m, GhWD, and GUS were driven by the Cauliflower mosaic virus (CaMV) 35S promoter to construct 35S::GhMML4 D12, 35S::GhMML4 D12m, 35S::GhWD, and 35S::GUS. These constructs were then introduced into A. tumefaciens (strain GV3101). Infected tissues were analysed 48 h after infiltration. The GFP signal was observed under a confocal laser scanning microscope (Zeiss LSM710). All experiments were repeated with three independent biological replicates with similar results.

Dual-luciferase assay
The GhMML4 D12, GhMML4 D12m, and GhWD promoters were inserted into pGreenII 0800-LUC to drive the firefly LUC reporter gene, with Renilla (REN) luciferase controlled by the constitutive 35S promoter on the same plasmid as a reference to normalize infection efficiency. The CDSs of GhMML4 D12, GhMML4 D12m, and GhWD were inserted into the pGreenII62-SK vector under the control of the 35S promoter. All primers used are listed in Supplementary Table S1. The constructs were transformed into A. tumefaciens (strain GV3101). The transformed Agrobacterium cells were mixed with the Agrobacterium strains harboring the effectors and reporters, in a volume ratio of 1:2. Transient transformation was conducted by infiltration of the Agrobacterium mixtures into the abaxial side of N. benthamiana leaves using a syringe. After culturing for 3 d, firefly LUC and REN activities were measured using the Dual-Luciferase Reporter Assay System (E2490, Promega, USA), and the LUC/REN ratio was determined (the value of LUC was normalized to that of REN). Three biological repeats were measured for each combination.
Complementation of Arabidopsis ttg1 phenotypes

The full length CDS of the GhWDR were fused with pBIENGFP4, under cotton native promoter. Binary construct (GhWDR::GhWDR) was introduced into A. tumefaciens strain GV3101 and subsequently transferred into Arabidopsis using the floral dip method (Clough and Bent, 1998). Selection of transformants was conducted on 0.8% agar containing Murashige and Skoog (MS) salts (2.2 g l−1) added with kanamycin (50 µg ml−1) for 7 d. Kanamycin-resistant seedlings were moved onto fresh kanamycin plates for 10 d before transfer to soil. Progeny from self-fertilized primary transformants were examined for complementation of other ttg1 mutant phenotypes. Trichome restoration was examined by growing on plates containing MS salts and 1.2% phytagel for about 2 weeks. Root hair position was studied in seedlings grown vertically on plates containing MS salts and 1.2% phytagel for about 5–6 d. Anthocyanin synthesis was investigated in seedlings grown on plates containing MS salts and 1.2% phytagel for about 10 d. Seed coat color was examined by stereoscope after harvest.

GhWDR-R was fused with GFP driven by CaMV35S (35Spro:GhWDR-R) into ttg1 Arabidopsis mutants. Selection of transformants and phenotype examination were as above.

Results

Molecular characterization of GhWDR interaction with GhMML4_D12

To investigate how GhMML4_D12 regulates lint fiber development, we used a Y2H assay to identify its potential interaction partners. Full-length GhMML4_D12 CDS was isolated and fused with bait vector (pGBK7-GhMML4_D12) to screen the cotton fiber specific yeast library. We did not observe any bHLH genes in the Y2H screening, probably due to both GhMML genes lacking the core amino acids needed in the interaction with bHLH as reported before (Grotewold et al., 2000; Zimmermann et al., 2004). However, a new kind of WD40 repeat protein (Gh_D01G0508) (Zhang et al., 2015) was found by prototrophy (see Supplementary Table S2). Structural analysis showed that the number of WD40 repeat domains in GhWDR differed from that in AtTTG1 and in GhTTG1 to GhTTG4 (Fig. 2A). Phylogenetic analysis showed that they belong to different clades (Fig. 2B), implying functional differentiation among them. We named this as GhWDR.

To better understand this newly found gene, we cloned its genomic DNA and CDS. The GhWDR gene contained seven exons and six introns in its open reading frame (ORF), but GhWDR-A at the A subgenome had a seven amino acid insertion in its N-terminus (see Supplementary Fig. S1). Both these two genes had three WD40 repeat domains. Both GhWDR and GhTTG1 to GhTTG4 were expressed in whole ovules at the anthesis day as well as throughout the elongation and secondary cell wall synthesis stages of fiber development (Fig. 2C, D; Supplementary Fig. S2). qRT-PCR analysis showed that the expression levels of these WD40 genes were lower in fiberless mutant XZ142FLM and n2NSM plants compared with the normal TM-1 (Supplementary Fig. S2), suggesting that these genes may contribute to fiber development.

To further confirm the interaction between GhMML4_D12 and GhWDR, full-length GhWDR CDS was introduced into the prey vector (pGADT7-GhWDR), and GhMML4_D12 and GhMML4_D12m introduced into bait vectors (pGBK7-GhMML4_D12 and pGBK7-GhMML4_D12m, respectively). Bait and prey vectors were co-transformed into yeast and the GhMML4–GhWDR interaction was reconstructed. However, no interaction was observed between GhMML4_D12m and GhWDR (Fig. 3A).

Validation of interaction between GhWDR and GhMML4_D12

Subcellular localization of GhMML4_D12, GhMML4_D12m, and GhWDR revealed that all these proteins were mainly present in the cell nucleus (see Supplementary Fig. S3). Biomolecular fluorescence complementation (BiFC) assays were adopted to validate the interaction between GhWDR and GhMML4_D12 or GhMML4_D12m in leaf epidermal cells in tobacco. The GhWDR protein was fused with the nYFP, and the cYFP was ligated with GhMML4_D12 and GhMML4_D12m. GhMML4_D12–cYFP and GhMML4_D12m–cYFP were transiently co-expressed with GhWDR–nYFP in tobacco. YFP fluorescence was observed in the nucleus of epidermal cells in tobacco leaves when GhWDR was co-expressed with GhMML4_D12, but no YFP fluorescence was detected when GhWDR was co-expressed with GhMML4_D12m (Fig. 3B).

The interaction between GhWDR and GhMML4_D12 was further confirmed in vitro. Purified GST–GhMML4_D12 and GST–GhMML4_D12m fusion proteins were incubated with His–GhWDR expressed in E. coli. Bound proteins were washed, separated on SDS-PAGE, and immunoblotted with an anti-GST antibody. As shown in Fig. 3C, the negative control (GST resin) and GST–GhMML4_D12m were unable to pull down His–GhWDR, whereas GST–GhMML4_D12 could efficiently pull down His–GhWDR, suggesting that GhMML4_D12 physically interacts with GhWDR in vitro. We have therefore shown for the first time that MML proteins directly interact with WD40 proteins, similar to but different from the MBW complex in regulating Arabidopsis leaf or root hair trichome patterning (Yang and Ye, 2013).

Decisive domains identified for interaction between GhMML4_D12 and GhWDR

R2R3 domains and the C-terminus of MYB genes can interact with other transcription factors (Qi et al., 2011, 2014; Song et al., 2011). Since we found that GhMMLs interact with WD40 genes directly, we further investigated how the interaction between GhMML4_D12 and GhWDR occurs. We divided GhMML4_D12 into two parts, GhMML4_D12-R2R3 and GhMML4_D12-C, as shown in Fig. 4A. GhMML4_D12-C was identified as being responsible for the interaction with GhWDR (Fig. 4A). Further investigation showed that amino acids from 250 to 267 of GhMML4_D12 contributed to the interaction with GhWDR, and amino acids from 259 to 267 may be more important (Fig. 4B); however, these amino acids are not conserved in GhMML genes (see Supplementary Fig. S4).

Since many WD40 proteins have been reported to act as transcriptional regulators through protein–protein interactions...
MML4 interact with WDR in cotton

(Baudry et al., 2004; Zhu et al., 2008), GhWDR was separated into three parts, GhWDR-N, GhWDR-C, and the WD40 repeat domain GhWDR-R. A Y2H assay showed that GhWDR-R was responsible for the interaction with GhMML4_D12, but not GhWDR-N and GhWDR-C (Fig. 4C). GhWDR-R was further divided into three parts based on the number of WD40 repeat domains, namely GhWDR-R1, GhWDR-R2, and GhWDR-R3, as illustrated in Fig. 4D. It was found that GhWDR-R1 and GhWDR-R3 could interact with GhMML4_D12, but the GhWDR-R2 did not (Fig. 4D).

Moreover, we also investigated the interaction between GhMML4_D12 and GhMML4_D12m and other WD40 proteins including GhTTG1 to GhTTG4 and AtTTG1. Interestingly, we found that GhTTG1 to GhTTG4 could interact with both GhMML4_D12 and GhMML4_D12m, and there was no obvious difference in the interaction efficacy with these WD40 genes between GhMML4_D12 and GhMML4_D12m. However, AtTTG1 was unable to interact with them (see Supplementary Fig. S5), implying a functional divergence of WD40 genes between cotton and Arabidopsis in regulating epidermal cell differentiation. A Y2H assay further showed that

Fig. 2. GhWDR belongs to different clade from GhTTG1 to GhTTG4 and AtTTG1. (A) Schematic diagrams of GhWDR, GhTTG1 to GhTTG4, and AtTTG1 domain constructs. The black box indicates the WD40 repeat domain. Numbers indicate the positions of the first and last amino acids of the domain. (B) GARLI maximum likelihood phylogram of GhWDR, GhTTG1 to GhTTG4 from cotton, and AtTTG1 from Arabidopsis, revealing two different clades. The NJ tree was constructed using the MEGA 6.0 program (http://www.megasoftware.net/). GhWDR belongs to a different clade from GhTTG1 to GhTTG4 and AtTTG1. (C) Fragments per kilobase per million mapped reads (FPKM) value to illustrate the expression pattern of GhWDR in various tissues of G. hirsutum cv. TM-1. R: root; S: stem; L: leaf; To: torus; Pe: petal; St: stamen; Pi: pistil; Ca: calycle; −3D, −1D, 0D, 1D, and 3D: ovules attached with fibers at −3, −1, 0, 1, and 3 DPA; 5O, 10O, 20O, and 25O: 5, 10, 20, and 25 DPA ovules without fibers; 5F, 10F, 20F, 25F, and 35F: 5, 10, 20, 25, and 35 DPA fibers. (D) qRT-PCR expression analysis of GhWDR in various tissues of G. hirsutum cv. TM-1. The error bar represents the standard deviation of the mean values of three biological replicates. The GhUBI1 gene was used as the internal control.
the R2R3 domain of GhMML4_D12 was responsible for the interaction with GhTTG1 to GhTTG4, and the C-terminus also participated in the interaction (Supplementary Fig. S5). These results imply that GhWDR is a unique gene that is important for the function of GhMML4_D12; however, the mutated GhMML4_D12m lost its interaction with GhWDR. These results show that amino acids from 250 to 267 of GhMML4_D12 and the first and third WD40 repeat domain of GhWDR are required for the interaction between these proteins.

**GhWDR acts as a transcriptional activator**

To investigate the subcellular localization of GhWDR, GhWDR, GhWDR-N, GhWDR-C, and GhWDR-R proteins were each fused in-frame to the 5’ terminus of GFP reporter gene under the control of the CaMV35S promoter and transformed into tobacco leaf cells by *Agrobacterium*, as illustrated in Fig. 4C. The GhWDR and GhWDR-C protein were both detected in the cell nucleus and cytomembrane, GhWDR-R was detected only in the cell nucleus, and GhWDR-N was mainly found in the cell cytomembrane (Fig. 5A).

To identify the transcriptional activity of GhWDR, the yeast GAL4-responsive reporter system was employed. Different domains of GhWDR were fused to the pGBKT7 to generate effector constructs and transform them into yeast. P53-pGBKT7 and pGBKT7 were used as positive and negative controls, respectively. As shown in Fig. 5B, the transformed yeast cells containing GhWDR-pGBKT7 could grow on SD/−Trp to which 50 ng ml⁻¹ AbA was added, indicating that the reporter gene, *LacZ*, was activated. Furthermore, the transformed yeast cells containing GhWDR-C-pGBKT7 could grow on SD/−Trp/+100 ng ml⁻¹ AbA, whereas those containing GhWDR-N and GhWDR-R were unable to grow on this medium. These results suggest that the GhWDR protein is present in the nucleus, its WD40 domain and C-terminus are sufficient for nuclear localization, and GhWDR functions as a transcriptional activator; the transcriptional activity of GhWDR mainly depends on GhWDR-C.

**Ectopic expression of GhWDR rescues ttg1 mutant phenotypes in Arabidopsis**

Arabidopsis trichomes and cotton fibers are both unicellular hairs of the epidermis, and may share a similar molecular machinery of regulation (Wang *et al.*, 2004; Humphries *et al.*, 2005; Guan *et al.*, 2008). Functional investigation of cotton
genes in Arabidopsis has been successfully used to elucidate the mechanisms that regulate cotton fiber development. Due to the difficulties of generating transgenic cotton, functional investigation of GhWDR was carried out in Arabidopsis glabrous mutant ttg1. The ttg1 allele results in severe defects in anthocyanin synthesis, trichome development, seed coat pigmentation, and root hair position. To investigate whether GhWDR could complement ttg1 mutant phenotypes, we transformed GhWDR fused with GFP under the control of a native promoter (GhWDR::GhWDR) into the ttg1 mutant. Of the 19 primary kanamycin-resistant T1 transgenic lines, four (L4, L5, L9, and L10) with ectopic expression of GhWDR showed the wild type phenotype, whereas no trichomes were observed in non-transformed (indicating as ttg1) and empty-vector controls (indicating as 35S_prom::GFP) (Fig. 6A). qRT-PCR analysis showed higher levels of ectopic expression of GhWDR in these four independent T2 lines, which could rescue ttg1 to its wild type phenotype (Fig. 6B).

The WD40 repeat domain of GhWDR decided the interaction with GhMML4, and we were interested in whether it could complement ttg1 mutant phenotypes. We transformed GhWDR-R fused to GFP driven by CaMV35S (35S_prom::GhWDR-R) into ttg1 Arabidopsis mutants. Phenotype examination revealed that no T1 GhWDR-R overexpression plants had wild type anthocyanin synthesis or trichome development. After all seeds were harvested, we were surprised to find that some of the mature seeds from GhWDR-R overexpression T1 lines were brown, similar to the wild type (Fig. 6C), indicating that GhWDR could rescue ttg1 to its wild type. Higher expression levels of GhWDR-R were examined in all those T2 lines that also rescue ttg1 to wild type (Fig. 6D).
Together, these results reveal that, similar to Arabidopsis TTG1, GhWDR plays an important orthologous role in anthocyanin synthesis, trichome development, seed coat pigmentation and root hair position, confirming its function in trichome development. Further investigation showed that its WD40 repeat domain is important for the function of GhWDR in ttg1 Arabidopsis.

GhMML4_D12 promotes transcriptional activation of itself in n2NSM

An increasing number of studies have shown that transcription factors can bind to the promoter of their target genes to either activate or repress expression (Chen et al., 2017; Lei et al., 2017; Zhang et al., 2018). To understand the relationship between GhMML4_D12 and GhWDR, we employed Y1H analysis to investigate the effect of GhMML4_D12, GhMML4_D12m, and GhWDR on the transcriptional activity of GhMML4_D12. GhMML4_D12, GhMML4_D12m, and GhWDR were fused with pGADT7, and the promoters of GhMML4_D12 from n2NSM and XZ142FLM (GhMML4_ D12m) were ligated into pAbAi. Sequence comparison of these two promoters showed no notable difference between them (see Supplementary Fig. S6). As shown in Fig. 7A, B, GhMML4_D12 could bind to its own promoter in n2NSM,
but GhMML4_D12m could not in XZ142FLM. To further verify the results, 35Spro:GhMML4_D12, 35Spro:GhMML4_D12m, 35Spro:GhWDR, and 35Spro:GUS were used as effector plasmids. NLS was ligated at the 3′ end of GhMML4_D12 and GhMML4_D12m promoters, then fused with GFP at the 3′ end of NLS to construct GhMML4_D12 pro:NLS-GFP and GhMML4_D12m pro:NLS-GFP. These two constructs were used as reporter plasmids (Fig. 7C). Co-transformation of these two reporters with 35Spro:GUS into tobacco leaves resulted in a relatively low level of fluorescence signal (Fig. 7D, E). However, when 35Spro:GhMML4_D12 was infiltrated into tobacco leaves, a much stronger fluorescence signal was observed (Fig. 7D). In contrast, co-infiltration of GhMML4_D12 pro:NLS-GFP with 35Spro:GhWDR generated relatively low levels of fluorescence (Fig. 7D). However, co-infiltration of GhMML4_D12m pro:NLS-GFP with either 35Spro:GhMML4_D12m or 35Spro:GhWDR had no effect (Fig. 7E).

To confirm these results, we further employed a dual-luciferase (Dual-LUC) reporter approach. 35Spro:GhMML4_D12, 35Spro:GhMML4_D12m, 35Spro:GhWDR, and 35Spro:GUS were used as effector plasmids. In addition, the reporters consisted of 35S promoter-driven Renilla luciferase (REN, as internal control) and GhMML4_D12 or GhMML4_D12m promoter-driven firefly luciferase (Fig. 7F). The expression of 35S:GhMML4_D12 increased the LUC/REN ratio (Fig. 7G), while co-expression 35Spro:GhWDR had little influence on GhMML4_D12 activity compared with 35Spro:GUS alone (Fig. 7G). Co-expression of 35Spro:GhMML4_D12m and 35Spro:GhWDR had no effect on GhMML4_D12m activity compared with 35Spro:GUS alone (Fig. 7H). All these results suggest that GhMML4_D12 enhances the transcriptional activity of itself in n2NSM, but not in XZ142FLM, and the transcriptional activity may improve the function of GhMML4_D12 in regulating cotton lint fiber development.

The transcriptional regulation of GhMML4_D12, GhMML4_D12m and GhWDR on GhWDR was also investigated. We found that neither GhMML4_D12 nor GhMML4_D12m was able to activate the expression of GhWDR, and GhWDR was also unable to activate its own transcriptional activity (see Supplementary Fig. S7), implying GhWDR simply serves as a scaffolding molecule in the protein-protein interactions to assist the function of GhMML4_D12.

Discussion

Evolution of trichome differentiation-related genes in plants

The MYB gene family members, which largely function as transcription factors, have several conserved binding domains.
Of these, R2R3 MYB is the largest subgroup. Subgroup 9 of the R2R3 MYBs contains MML genes and subgroup 15 contains GL1-Like genes. These genes play important roles in specification and regulation of plant cellular differentiation (Paterson et al., 2012; Brockington et al., 2013). MML genes mainly control fiber development. The GL1-Like gene GaMYB2 complements Arabidopsis gl1, and its ectopic expression induces a single trichome from the epidermis of Arabidopsis seeds (Wang et al., 2004). Although GaMYB2 is highly expressed in developing fibers, some other genes, GhMYB25, GhMML3_A12, and GhMML4_D12, belong to a novel MIXTA clade of MYB transcription factors involved in cotton seed fiber differentiation (Machado et al., 2009; Walford et al., 2011; Wan et al., 2016; Wu et al., 2018). Expression of these MML genes is restricted to cotton ovule epidermal cells and young fiber cells, not trichomes on other plant parts, such as stems, leaves, and petals (Zhang et al., 2015). These MML genes are more distinct in cotton fibers and have independent ovule-specific functions that directly expand epidermal cells for seed trichome or fiber production. These genes have closer phylogenetic relationships with MML genes expressed in petals than some other MYB genes (GL1 clade) associated with other types of trichome or root hairs. So, these MML transcription factors have likely evolved independent ovule-specific functions to direct the expansion of epidermal cells to produce the cotton fiber. These findings will also be helpful to illustrate the roles of MIXTA genes in regulating epidermal cell differentiation such as production of a ‘pulpy layer’ secreted from the teguments surrounding cacao seeds, and mucilages in other Malvaceae fruit (Abelmoschus (okra), and Cola (kola)).

WD40 repeat proteins are characterized by the presence of 40–60 amino acid peptide motifs, which are usually delimited by the GH dipeptide (Gly-His) and the WD dipeptide (Trp-Asp) at the C- and N-termini (van Nocker and Ludwig, 2003). WD40 proteins are found widely in the plant kingdom and regulate the biosynthesis of anthocyanin, proanthocyanin,
MML4 interact with WDR in cotton

and mucilage in seeds, and the development of trichomes and root hairs. The TTG1 locus regulates several developmental and biochemical pathways in Arabidopsis, including the formation of hairs on leaves, stems, and roots, as well as the production of seed mucilage and anthocyanin pigments (Walker et al., 1999). Two cotton WD40 transcription factors, GhTTG1 and GhTTG3, could rescue ttg1 to its wild type (Humphries et al., 2005). Phylogenetic analysis showed that GhWDR belongs to a clade that is different from that of AtTTG1 and GhTTG1 to GhTTG4, implying functional divergence between them (Fig. 2B). GhWDR could complement ttg1 to its wild type, implying its conservative function in trichome development (Fig. 6A). Further investigation showed that the WD40 repeat domain (from amino acid 454 to 773) of GhWDR could rescue ttg1 (Fig. 6C). Further transgenic analysis of GhWDR in cotton fiber development is needed in future work to help in understanding the key role of the WD40 family in plants.

**GhMML4–GhWDR complex shows similarity to but is distinct from MBW complex**

The MBW ternary complex consists of MYB, bHLH, and WD40 transcription factors, and has important roles in anthocyanin and proanthocyanin biosynthesis and trichome initiation in a variety of plant species. Usually, WD40 proteins do not interact with MYB proteins directly; they just serve as scaffolding molecules to interact with bHLH factors, and then bHLH factors interact with different kinds of MYB factors to regulate different cell fates (Ramsay and Glover, 2005). The MYB–bHLH complex regulates distant and diverse biological processes such as cell death, cell wall synthesis, hormone signaling, stamen development, and seed production (Stracke et al., 2001). Further investigation of protein–protein interaction specificities of the MYB protein in Arabidopsis revealed a conserved amino acid signature ([DE]Lx2[RK]x3Lx6Lx3R) as the structural basis for the interaction between MYB and R/B-like bHLH proteins (Zimmermann et al., 2004).

In this study, we found a lint fiber-specific MML factor, GhMML4_D12, could directly interact with a newly found WD40 factor, GhWDR, but not with any bHLH genes. And no interaction between GhMML4_D12 and GhWDR was observed (Fig. 3). The newly found GhMML4–GhWDR complex shares some similarity to but is distinct from the MBW ternary complex in regulation of leaf/root hair trichome patterning in Arabidopsis. The similarity between them is that the members of these two complexes both belong to the MYB, bHLH, or WD40 factors, and thus GhMML4_D12 and GhWDR have diverged from R2R3MYB and WD40. Meanwhile, the combining form is different.

We believe cotton fiber (seed trichome) and leaf/root hair trichomes have a similar but different regulatory network. Both Arabidopsis trichomes and cotton fibers are unicellular structures of epidermal origin. And they may share similar molecular machinery of regulation. Several genes implicated in cotton fiber development were shown to functionally substitute for their homologues in Arabidopsis. Ectopic expression of GaMYB2 complements gl1 mutant in Arabidopsis (Wang et al., 2004). Moreover, two cotton WD40 transcription factors, GhTTG1 and GhTTG3, are able to rescue Arabidopsis ttg1 to its wild type (Humphries et al., 2005). Furthermore, one HOX

---

**Fig. 8.** A proposed model of the mechanism by which the GhMML4–GhWDR regulatory complex regulates cotton lint fiber development. In n2NSM, the C-terminal domain of GhMML4_D12 directly interacts with the WD40 repeat domain of GhWDR; GhMML4_D12 bound to its own promoter to enhance its transcriptional activity in regulation of lint fiber development. However, in XZ142FLM, mutated GhMML4_D12m did not interact with GhWDR, and thus GhMML4_D12m could not activate its own transcriptional activity. The green box indicates R2R3 domains; the purple circle indicates C-terminal domain of GhMML4_D12; orange and light green circles indicate N-terminal domain and C-terminal domain of GhWDR, respectively; and the blue box indicates WD40 repeat domain of GhWDR.
gene, GaHOX1, is a functional homologue of Arabidopsis GL2 (Guan et al., 2008). Therefore, it had been supposed for a long time that development of cotton fibers and Arabidopsis trichomes shares a similar mechanism. However, Machado et al. (2009), Walford et al. (2011), and our previous reports (Wan et al., 2016; Wu et al., 2018) all found that GhMML3_A12 and GhMML4_D12 belong to a novel MIXTA clade of MYB transcription factors, involve in cotton fiber differentiation. So, the genes controlling Arabidopsis trichome and cotton fiber development may be different, although unbranched cotton fibers and trichomes share some degree of morphological similarity. For example, Arabidopsis MYBMIXTA-Like (MML) gene AtMYB106 was a negative regulator of branching in Arabidopsis, while its homologue GhMYB25 is reported to be a key regulator for cotton fiber initiation (Machado et al., 2009; Oshima et al., 2013). Furthermore, spontaneous fibreless mutants showed normal trichome development (Ruan et al., 2000; Du et al., 2001; Ruan, 2005), and genes regulating fiber initiation and leaf trichomes have been mapped to different loci (Loguercio et al., 1999). All this evidence confirms that cotton fibers and leaf trichomes may be regulated by different mechanisms.

Due to the similarity and difference between the GhMML4–GhWDR complex and MBW complex, therefore, we report here that cotton fibers and Arabidopsis leaf/ hair trichomes may have similar but different regulatory networks in plants. Further investigation is needed to uncover the mechanisms underlying cotton fiber development and trichome formation in other organs of the plant.

In summary, the C-terminal domain of GhMML4_D12 and WD40 repeat domain of GhWDR determine their interaction in n2NSM. GhWDR simply serves as a scaffolding molecule in protein–protein interactions to assist the function of GhMML4_D12. However, the mutated GhMML4_D12m has lost the C-terminal domain, and therefore could not interact with GhWDR in XZ142FLM. GhMML4_D12 bound to its own promoter to enhance its transcriptional activity, and this may help in the function of GhMML4_D12 in n2NSM, but GhMML4_D12m did not do this in XZ142FLM (Fig. 8). These together provide a new insight into the function of GhMML4_D12 in the regulation of cotton lint fiber development in n2NSM.

Supplementary data

Supplementary data are available at JXB online.

Fig. S1. Cloning and characterization of GhWDR.
Fig. S2. Expression pattern analysis of WD40 genes in fiber mutants.
Fig. S3. Subcellular localization of GhMML4_D12, GhMML4_D12m, and GhWDR protein in leaf cells of tobacco.
Fig. S4. Alignment of amino acids from 10 GhMMLs genes in cotton.
Fig. S5. Interaction between GhMML4_D12, GhMML4_D12m, and GhWDR in yeast.
Fig. S6. Sequence comparison of GhMML4_D12 promoters from TM-1, n2NSM, and XZ142FLM.

Table S1. All primers developed and used in present research.
Table S2. Proteins interacting with GhMML4_D12 identified by Y2H screening.

Acknowledgements

This study was financially supported in part by grants from the National Research and Development Project of Transgenic Crops of China (2015ZX08009-003) to YH, the earmarked fund for the China Agriculture Research System and the Distinguished Discipline Support Program of Zhejiang University to TZ. This study was also financially supported in part by grants from the National Natural Science Foundation of China (31701469; the Fundamental Research Funds for the Central Universities (2019ZZX004-13).

Author contributions

T-ZZ, YT, X-YG, YH, and L-YD designed the research. YT, J-JD, H-TW, W-HC, M-LL, D-FY, and Q-LY performed research. YT analysed data. T-ZZ, and YT wrote the article.

Accession numbers

Sequence data for the genes described in this study can be found in the GenBank database under the following accession numbers: AtTTG1 (NP_001318637), GhTTG1 (AAM95642.1), GhTTG2 (NP_001313775), GhTTG3 (XP_012460128), GhTTG4 (XP_0116693970) and GhWDR (XP_016702485).

Conflict of interest

The authors declare no competing interests.

References

Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L. 2004. TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in Arabidopsis thaliana. The Plant Journal 39, 366–380.
Baumann K, Perez-Rodriguez M, Bradley D, Venail J, Bailey P, Jin H, Koes R, Roberts K, Martin C. 2007. Control of cell and petal morphogenesis by R2R3 MYB transcription factors. Development 134, 1691–1701.
Brockington SF, Alvarez-Fernandez R, Landis JB, Alcorn K, Walker RH, Thomas MM, Hileman LC, Glover BJ. 2013. Evolutionary analysis of the MIXTA gene family highlights potential targets for the study of cellular differentiation. Molecular Biology and Evolution 30, 526–540.
Chen L, Xiang S, Chen Y, Li D, Yu D. 2017. Arabidopsis WRKY45 interacts with the DELLA protein RGL1 to positively regulate age-triggered leaf senescence. Molecular Plant 10, 1174–1189.
Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant Journal 16, 735–743.
Du XM, Pan JJ, Wang RH, Zhang TZ, Shi YZ. 2001. Genetic analysis of presence and absence of lint and fuzz in cotton. Plant Breeding 120, 519–522.
Glover BJ, Perez-Rodriguez M, Martin C. 1998. Development of several epidermal cell types can be specified by the same MYB-related plant transcription factor. Development 125, 3497–3508.
Grotewold E, Sainz MB, Tagliani L, Hernandez JM, Bowen B, Chandler VL. 2000. Identification of the residues in the Myb domain of
maize C1 that specify the interaction with the bHLH cofactor R. Proceedings of the National Academy of Sciences, USA 97, 13579–13584.

Quan XY, Li QJ, Shan CM, Wang S, Mao YB, Wang LJ, Chen XY. 2008. The HD-Zip IV gene GaHox1 from cotton is a functional homologue of the Arabidopsis GLABRA2. Physiologia Plantarum 134, 174–182.

Hellens R, Mullineaux P, Klee H. 2000. Technical focus: A guide to Agrobacterium binary Ti vectors. Trends in Plant Science 5, 446–451.

Humphries JA, Walker AR, Timmins JN, Orford SJ. 2005. Two WD-repeat genes from cotton are functional homologues of the Arabidopsis thaliana TRANSPARENT TESTA GLABRA1 (TTG1) gene. Plant Molecular Biology 57, 67–81.

Lei R, Li X, Ma Z, Lv Y, Hu Y, Yu D. 2017. Arabidopsis WRKY2 and WRKY34 transcription factors interact with VQ20 protein to modulate pollen development and function. The Plant Journal 91, 962–976.

Li XB, Fan XP, Wang XL, Cai L, Yang WC. 2005. The cotton ACT1 gene is functionally expressed in fibers and participates in fiber elongation. The Plant Cell 17, 859–875.

Liu T, Song T, Zhang X, et al. 2014. Unconventionally secreted effectors of two filamentous pathogens target plant salicylate biosynthesis. Nature Communications 5, 4686.

Loguerico LL, Zhang JQ, Wilkins TA. 1999. Differential regulation of six novel MYB-domain genes defines two distinct expression patterns in allotetraploid cotton (Gossypium hirsutum L.). Molecular & General Genetics 260, 661–670.

Machado A, Wu Y, Yang Y, Llewellyn DJ, Dennis ES. 2009. The MYB transcription factor GhMYB25 regulates early fibre and trichome development. The Plant Journal 59, 52–62.

Messina G, Atterato MT, Fanti L, Giordano E, Dimitri P. 2016. Expression of human Ctdp1 gene in Drosophila reveals new insights into the function of the evolutionarily conserved BCNT protein family. Scientific Reports 6, 25511.

Noda K, Glover BJ, Linstead P, Martin C. 1994. Flower colour intensity depends on specialisation of cell shape controlled by a Myb-related transcription factor. Nature 369, 661–664.

Oshima Y, Shikata M, Koyama T, Ohtsubo N, Mitsuda N, Ohme-Takagi M. 2013. MXTA-like transcription factors and WAX INDUCER1/Oshima Y, Shikata M, Koyama T, Ohtsubo N, Mitsuda N, Ohme-Takagi M. 2013. MXTA-like transcription factors and WAX INDUCER1/Oshima Y, Shikata M, Koyama T, Ohtsubo N, Mitsuda N, Ohme-Takagi M. 2013. MXTA-like transcription factors and WAX INDUCER1/3 complex containing CONSTITUTIVELY PHOTOMORPHOGENIC1 and SUPPRESSOR OF PHYA proteins in light control of plant development. The Plant Cell 20, 2307–2323.

Zimmermann IM, Heim MA, Weishhaar B, Uhrig JF. 2004. Comprehensive identification of Arabidopsis thaliana MYB transcription factors interacting with P/B-like BHLH proteins. The Plant Journal 40, 22–34.
