GENETIC AND EXPRESSIONAL EVIDENCE DISPLAYS FUNCTIONAL DISTINCTION OF COTTON BASIC HELIX-LOOP-HELIX PROTEINS IN ARABIDOPSIS TRICHOME INITIATION

ANH PHU NAM BUI

Institute of Applied Technology, Thu Dau Mot University, Binh Duong Province, Vietnam
Corresponding email: buiphunamanh@yahoo.com

Abstract

The cultivated tetraploid cotton species (AD genomes) were originated from two ancestral diploid species (A and D genomes). While the ancestral A-genome species produce spinnable fibers, the D- genome species do not. Cotton fibers are unicellular trichomes originating from seed coat epidermal cells, and currently there is an immense interest in understanding the process of fiber initiation and development. Current knowledge demonstrates that there is a great of deal of resemblance in initiation mechanism between Arabidopsis trichome and cotton fiber. In this study, comparative functional studies between A genome and D-genome species in cotton by using Arabidopsis trichome initiation as a model was performed. Four cotton genes TTG3, MYB2, DEL61 and DEL65 were amplified from A-genome and D-genome species and transformed into their homolog trichomeless mutants Arabidopsis *ttg1*, *gl1*, and *gl3egl3*, respectively. Our data showed that the transgenic plants expressing TTG3 and MYB2 genes from A genome and D-genome species complement the *ttg1* and *gl1* mutants, respectively. It was also discovered that complete absences of two functional basic helix loop helix (bHLH) proteins (DEL65/DEL61) in D- diploid species and one (DEL65) that is functional in A-genome species, but not from D-genome species. This observation is consistent with the natural phenomenon of spinnable fiber production in A- genome species and absence in D-genome species. These results suggested that MYB2, TTG3 and DEL65, when expressed in Arabidopsis, regulated the regulatory network genes during the trichome initiation process.

Key words: Arabidopsis trichome, MYB2, TTG3, DEL65, DEL61.

Introduction

Position-dependent cell fate determination and pattern formation in Arabidopsis trichomes have been well-studied. Trichomes are unicellular single-celled structures emerging and differentiating in the leaf epidermal cells. After cell divisions, these structures subsequently become independent of each other (Hulskamp, 2004). The mature Arabidopsis leaf trichome consists of a stalk and three to four branches and its function can vary from trapping herbivorous insects, dispersing seeds, reducing transpiration, to protecting the plants from ultraviolet radiation. The morphogenesis of a trichome is characterized by a series of six phases (Szymanski *et al*., 1998), starting with the introduction of trichome initial followed by subsequent radial expansion and completing with the promotion of completely developing trichome. Other cellular activities connected with the trichome maturation from phase one to phase six include endo-reduplication of the nuclear DNA to an average of 32–64C (the ploidy level of original un-replicated cells is 2C), vacuolization during the transition, and the development of surface papillae during phase one through four, phase four to five, phase five and six, respectively (Hulskamp, 2004).

The initiation of Arabidopsis single-celled trichome from leaf epidermal cells presents a useful tool to study the genetic pathways and regulatory signals in cell fate regulation (Larkin *et al*., 2003; Serna & Martin, 2006). Genetic and molecular research have elucidated trichome development by a transcriptional and regulatory network controlled by trichome activating and suppressing genes. Over thirty genes have been isolated accounting for diverse aspects in trichome formation including trichome initiation, spacing, size, and morphology. Three groups of proteins

have been showed to participate in a trimeric complex to promote trichome initiation, including the WD40 protein TRANSPARENT TESTA GLABRA1 (TTG1) (Glover, 2000; Ramsay & Glover, 2005; Schiefelbein, 2003; Serna & Martin, 2006; Walker *et al*., 1999), the R2R3 MYB-related transcription factor GLABRA1 (GL1) (Oppenheimer *et al*., 1991), and the basic helix-loop-helix (bHLH)-like transcription factors GL3 and their functionally redundant ENHANCER OF GL3 (EGL3) (Payne *et al*., 2000; Zhang *et al*., 2003). Upon the assembly of these proteins, the trichome trimeric complex activates transcription of its direct downstream gene GLABROUS2 (GL2) encoding a homeodomain-leucine zipper protein. GL2 is documented as the primary target gene of the trichome patterning machinery and is accountable for regulating trichome initiation on Arabidopsis leaves (Hulskamp, 2004; Larkin *et al*., 2003; Serna, 2004). Together with GL2, WRKY transcription factor TTG2 and cell cycle gene SIAMESE (SIM) are also upregulated. TTG2 is strongly expressed during trichome patterning and differentiation, while SIM controls endo-replication, a process essential for trichome development.

The GL1-GL3/EGL3-TTG1 complex also upregulates a number of homologous R3 single repeat MYB genes that partially redundantly function as trichome initiation suppressors. These include TRIPTYCHON (TRY) (Schellmann *et al*., 2002), CAPRICE (CPC), ENHANCER OF TRY AND CPC1 (ETC1), ETC2, ETC3 (Tominaga *et al*., 2008), TRICHOMELESS1 (TCL1), and TCL2 (Wang *et al*., 2007). It has been proposed that these repressors render the trimeric complex inactive by the competition with GL1 for binding site with GL3, forming the inert complex R3 MYB INHIBITOR-GL3/EGL3-TTG1. The binding and competition with GL1 differ substantially among

these repressors. Binding assays suggest that TRY shows the strongest binding affinity, while CPC is the most dominant competitor for binding of GL1 to GL3 (Wang *et al*., 2008; Wester *et al*., 2009).

Cotton (Gossypium spp.) is regarded as one of the most influential crop plants widely cultivated for textile production. Of fifty-two members in Gossypium genus, there are 46 diploids (2n = 2X = 26), five well-established tetraploids, and one purported tetraploid species (2n = 4X = 52). It has been proposed that diploid cotton species may have originated from a common ancestor that subsequently evolved and diversified into eight monophyletic groups denoted as A–G, and K (Fryxell *et al*., 1992; Wendel & Cronn, 2003; Wendel *et al*., 2012). Approximately 1–2 million years ago, there was a spontaneous and interspecific hybridization event between A and D diploids and subsequent polyploidization that introduced a new allotetraploid (AD) lineage. Two diploid *G. herbaceum* (A$_1$), *G. arboretum* (A$_2$) and two allotetraploid Upland cotton species, *Gossypium hirsutum* L (AD$_1$), and Sealand cotton, *G. barbadense* (AD$_2$), are dominating in more than 70 countries and have had significant influence on global economic development (Stephens & Moseley, 1974; Zhang & Feng, 2000) . Interestingly, *G. raimondii* (D$_5$) contributes the D-genome of the allotetraploid cottons, yet it does not confer spinnable fibers production as the A-genome donors do (*G. arboreum*, *G. herbaceum*) (Wendel & Cronn, 2003).

Cotton fibers are seed trichomes. Since both cotton fiber and Arabidopsis trichome are single-celled structures differentiated from the ovule and leaf epidermal cells, respectively, it is suggested that these two species could share analogous mechanisms for mediating cell fate determination in trichomes. Compared with the Arabidopsis trichome, the underlying mechanism of cotton fiber initiation formation remains elusive. Most of the recent research on cotton fiber development focus on genomic and transcriptomic profiles during the cell elongation stage and secondary wall biosynthesis stage (Arpat *et al*., 2004; Yang *et al*., 2006; Shi *et al*., 2006; Udall *et al*., 2006; Wu *et al*., 2006). However, the mechanism controlling these pathways still needs to be elucidated.

Previous reports also characterized the importance of transcription factors in cotton fiber developmental pathways. So far, dozens of cotton genes encoding numerous classes of transcription factors have been characterized and found to be upregulated in developing fiber cells. Additionally, many of these cotton genes exhibited high protein sequence similarities to Arabidopsis trichome regulators (Guan *et al*., 2007; Serna, 2004). Ectopic expression of *Ga*MYB2 from *G. arboreum*, which is homologous to *At*GL1, rescues the trichomeless phenotype of the Arabidopsis *gl1* T-DNA mutant and induces a single trichome from the epidermis of Arabidopsis seeds, suggesting that *Ga*MYB2 is a functional homolog of GL1 (Guan *et al*., 2011; Guan *et al*., 2014; Wang *et al*., 2004). Additionally, homologs of Arabidopsis GL3, TTG1, CPC, TRY and GL2 (*Ga*DEL65, *Ga*TTG1, *Ga*CPC, *Ga*TRY, and *Ga*HOX1,

respectively) were also isolated in *G. arboreum* and functionally characterized using the Arabidopsis trichome model system (Guan *et al*., 2008; Guan *et al*., 2014; Wang *et al*., 2013). The four WD-repeat AtTTG1-like genes *Gh*TTG1–*Gh*TTG4 from the D$_t$ subgenome of the upland cotton *G. hirsutum* have been identified to be constantly expressed in some tissues, such as ovules and fibers (Humphries *et al*., 2005).

In this paper, it was hypothesized that the cotton genes activate the Arabidopsis trimeric complex similarly to the Arabidopsis genes in initiating the trichome. The results indicated that transgenic lines with MYB2, TTG3 and DEL65 from diploid genomes A (accession GCA_000612285.2) and D (accession PRJNA82769) complemented the trichomeless phenotype of *gl1-1*, *ttg1-1* and *gl3-1 egl3-77439*, respectively. The gene expression of the downstream targets of the trichome initiation complex in three different trichomeless mutants, *ttg1-1*, *gl1-1* and *gl3-1 egl3-77439* and complemented lines with their cotton homologs TTG3, MYB2 and DEL65, respectively, were also analyzed. The quantitative PCR showed that in transgenic lines with MYB2, TTG3 and DEL65, trichome-positive regulators GL2, TTG2, SIM, and HDG11 were up-regulated while the regulation of trichome suppressors TRY, TCL1, ETC1, and CPC were downregulated with the over-production of trichomes on leaves. These results represented a similar regulatory network in trichome formation in Arabidopsis transgenic lines complemented with homologous cotton genes.

Materials and Methods

Plant materials and growth conditions: The trichomeless Arabidopsis *gl3-1 egl3-77439* (Kanamycin resistant) double mutant (CS6516), and two single EMS mutants *ttg1-1* (CS89) and *gl1-1* (CS1644) were previously described by Esch *et al*. (2003), Humphries *et al*., (2005), and Guan *et al*., (2014), respectively (Esch *et al*., 2003; Guan *et al*., 2014; Humphries *et al*., 2005). All the seeds were obtained from Arabidopsis Biological Resource Center (Ohio State University, Columbus, OH). Arabidopsis seeds were surface- sterilized by the vapor-phase sterilization method described by Clough & Bent (1998) (Clough & Bent, 1998). Seeds were transferred into 1.5µl tubes, which were subsequently placed in a desiccator jar. Prior to sealing the desiccator, a beaker containing 250ml bleach and 5ml HCl was positioned in the desiccator. Sterilization was carried out 12 hours in the fume hood. Once seeds were collected, they were plated on Murashige and Skoog (MS) medium containing 0.8% phytogel. After vernalization for 2 days by placing in the dark at 4°C, seeds were finally transferred to a growth chamber with the following environmental conditions: 22°C, light intensity of 130-150 Em^{-2s-1}, 16:8h, light: dark photoperiod and relative humidity of 80% as described previously (Junker *et al*., 2015). Seven days after germination, seedlings were transplanted to soil and grown until maturity in the same temperature and light conditions. Antibiotic selections were performed by supplementing the MS medium with Kanamycin (50mg.ml-1) or Hygromycin (50mg.ml-1) or Basta (50mg.ml-1).

Cloning of DEL65, TTG3 and MYB2: To prepare the 35S::DEL65, 35S::DEL61 and 35S::TTG3 genomic constructs from cotton genomes A (A1, A2) and D (D1, D2, D9), the entire genomic DNA regions of DEL65, DEL61 and TTG3 were amplified by PCR and then cloned in pMDC32 vectors. For cloning DEL65 and DEL61, the forward and reverse primers were engineered with AscI and PacI restriction sites with the following sequences (Table 1): *Ga*DEL65-F/ *Ga*DEL65-R; *Ga*DEL61-F/ *Ga*DEL61-R. For cloning TTG3, the forward and reverse primers were engineered with *Xho*I and *Not*I restriction sites with the following sequences: *Ga*TTG3-F/ *Ga*TTG3-R: The PCR products were subsequently inserted into CaMV 35S expression cassette of pMDC32 vectors. The MYB2 sequence from A and D cotton diploid species was isolated from cotton genomic DNA by the primer pair of *Xho*I-*Gh*MYB2-F/ *Not*I-*Gh*MYB2-R. The amplified products were subsequently ligated into the pBARN vector and sequenced.

Genetic complementation: The pMDC32 and pBARN vectors harboring DEL61, DEL65, TTG3 and MYB2 genes were electroporated into *Agrobacterium tumefaciens* GV3101, respectively. Prior to plant transformation, these constructs were verified by sequencing. *A. tumefaciens*-mediated transformation of Arabidopsis plants homozygous for double mutant for *gl3-1 egl3-77439*, homozygous single mutant for gl1-1 and homozygous single mutant for ttg1-1 were performed by the floral dip method with constructs 35S::DEL65 or 35S::DEL61, 35S:: MYB2, and 35S:: TTG3, respectively (Clough & Bent, 1998). The transgenic seeds were screened on plates containing both Hygromycin and Cefotaxime for selection. For 35S::MYB2 construct transformation, the transgenic seeds were screened on plates containing Basta for selection. The resistant seedlings were transplanted to soil and phenotypically analyzed.

Phenotypic analyses and microscopy: Arabidopsis wild type and transgenic leaves were collected at the 15-day rosette stage and examined under an Olympus SZ61 industrial microscope. Images were taken by 5-megapixel digital color camera Olympus UC50 (Japan).

RNA extraction, cDNA synthesis and quantitative PCR: Transgenic plants with trichome recovery phenotype were subjected to RNA extraction. Total RNA from Arabidopsis wild type, mutants and transgenic plants were extracted from 100 mg three-week-old leaf tissues using SpectrumTM Plant Total RNA Kit (Sigma-USA) in accordance with the manufacturer's instructions. For synthesis of the first strand cDNA, RNA was treated with RNase-free DNase I (Sigma, USA) to eliminate genomic DNA, and two μg of total DNA-free RNA were used to synthesize first strand cDNA with iScriptTM Reverse Transcription Supermix RT-qPCR (Bio-Rad, USA) in accordance with the manufacturer's instructions.

Table 1. Primers for cloning and qPCR.

| Primer	Sequence
*Ga*DEL65-F	GCTTGGCGCGCCATGTCTACTGGAGTTCAACATCAAG
*Ga*DEL65-R	GGCCTTAATTAATCAACACTTGCCAGCAATTCTTTGC
*Ga*DEL61-F	ATCGCTCGAGATGGCTACTACCGGGGTTCAAAATCAAG
*Ga*DEL61-R	ATCGGCGGCCGCCTAAAAAGATTGTTTTACCCTTGATTTTATAGTCACAG
*Ga*TTG3-F	GCCGCTCGAGATGGAGAATTCAACTCAAGAATCCCACCTG
*Ga*TTG3-R	ATCGGCGGCCGCTCAAACTTTGAGAAGCTGCAATTTGTTGG
*Xho*I-*Gh*MYB2-F	ACTGGCGGCCGCATGGCTCCAAAGAAGGATGGAGT
*Not*I-*Gh*MYB2-R	ACTGCTCGAGTTATACCATTGCTAATGGATCC
Q9*At-Actin*-F	GCACCCTGTTCTTCTTACCG
Q10*At-Actin*-R	AACCCTCGTAGATTGGCACA
Q180_*CPC*_F	CAAGGCTTCTTGTTCCGAAG
Q181_*CPC*_R	GCCGTGTTTCATAAGCCAAT
Q182_*TRY*_F	TGTCGGTGATAGGTGGGATT
Q183_*TRY*_R	GACGGTGAGGCTTGGTATGT
Q184_*ETC1*_F	CCAACCATTGTTGCCTCTTC
Q185_*ETC1*_R	TCATCACCCAAAACCTCTCA
Q186_*GL2*_F	CCCCTCTGGATTCTCAATCA
Q187_*GL2*_R	GACGAGGTTTGTCACGGATT
Q188_*TTG2*_F	GAAGCAGGAGTATCGCAAGG
Q189_*TTG2*_R	GATCATCACTCGCTCGTTCA
Q200_*TCL1*_F	AAGAAGAGTGGTGGGACGTG
Q201_*TCL1*_R	TGATGAGGAGAACCCCACTC
Q202_*SIM*_F	CTTTACACGTCGACCCACTC
Q203_*SIM*_R	CATACTTGTGCATGTGCCTCT
Q204_*HDG11*_F	ATATGGAGTCGGTGGAAACG
Q205_*HDG11*_R	GCATTGAAGGCAAAAGAAGG
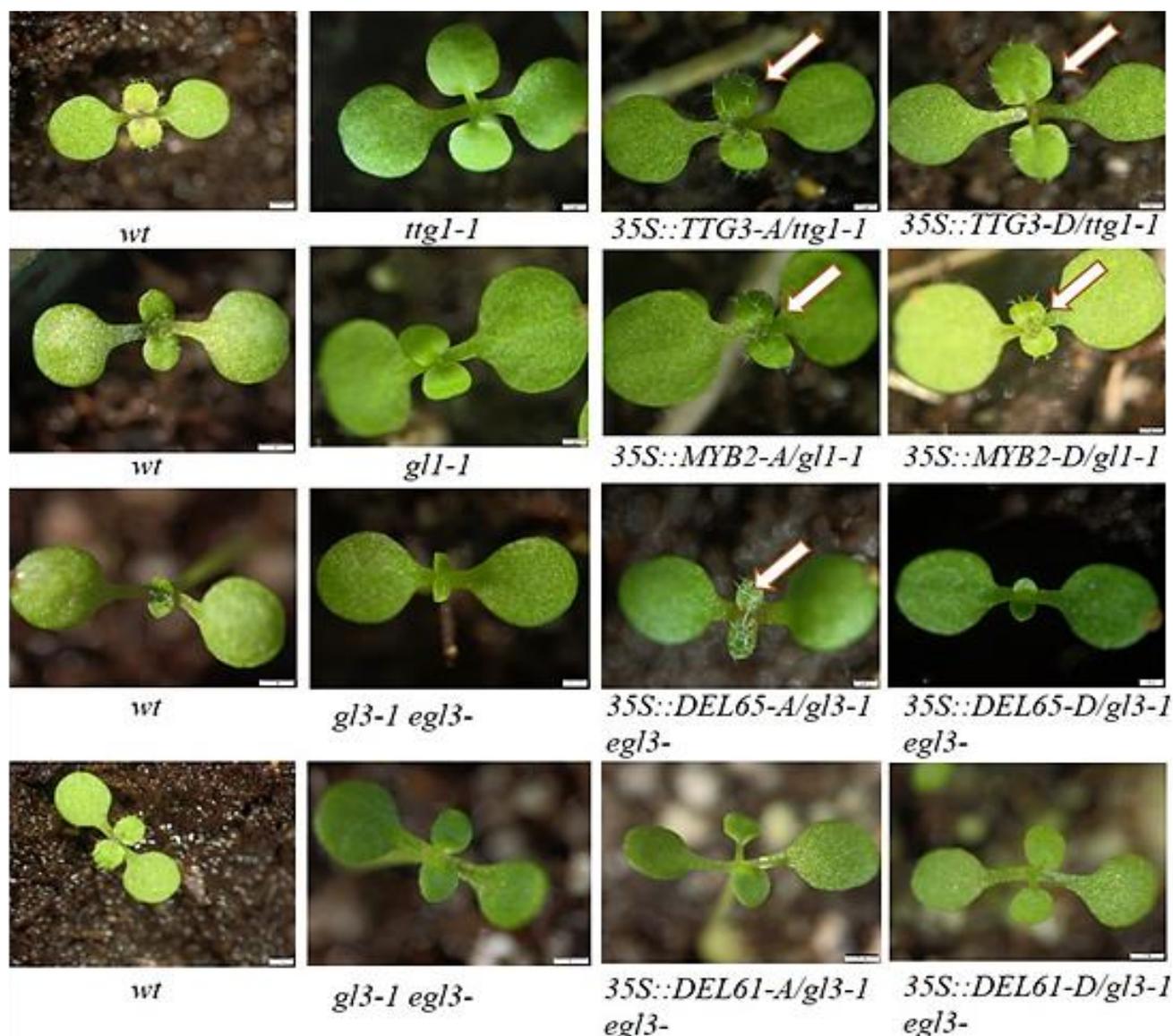
Fig. 1. From top to bottom: Complementing assays on trichome phenotype were performed on *ttg1-1* single mutant, *gl1-1* single mutant, and *gl3-1 egl3-77439* double with TTG3, MYB2, and DEL61, DEL65, respectively. Arrows indicate trichome initiation on first true leaves of transgenic plants

For Real time quantitative PCR, double-strand cDNA samples were diluted with water to 0.025 to 0.005 times depending on the concentration of the first-strand cDNA samples. Eight downstream target genes- GL2 (Q186_GL2_F/Q186_GL2_R), HDG11 (Q204_HDG11_F/ Q205_HDG11_R), SIM (Q202_SIM_F/Q203_SIM_R), TTG2 (Q188_TTG2_F/Q189_TTG2_R), CPC (Q180 _CPC_F/Q181_CPC_R), ETC1 (Q184_ETC1_F/Q185 _ETC1_R), TCL1 (Q200_TCL1_F/Q201_TCL1_R), and TRY (Q182_TCL1_F/Q201_TCL1_R) were amplified. Primers of target gene and control gene ACTIN (Q9At- Actin-F/Q10At-Actin-R) were listed (Table 1). Quantitative PCR was conducted with FastStart DNA Green Master (Roche-USA) in accordance with the manufacturer's instructions, and a Roche Real-time detection system Light Cycler 96 was used to detect and determine the differential expression of the studied genes. Quantitative PCR data were analyzed by using ΔΔCt method (Yuan *et al*., 2016).

Results

Complementation studies: Transgenic plants expressing TTG3 and MYB2 genes from A and D diploids complemented the *ttg1-1* and *gl1-1* mutants, respectively. However, ectopic expression of DEL65 from the A diploid rescued the *gl3-1 egl3-77439* mutant, while over expression of DEL65 from the D- diploid did not recover the double mutant phenotype (Fig. 1). This observation was consistent with the natural phenomenon of spinnable fiber production in the A genome species and absence in the D genome species. Since these results were from one each of these species, DEL65 from other available A and D diploid species were characterized. Genomic DNA of DEL65 was cloned from A1 (*G. herbaceum*), A2 (*G. arboreum*), D1 (*G. thurberi*), D2 (*G. armourianum*), and D9 (G. laxum) species and transformed into the *gl3-1 egl3-77439* double mutant. Conclusively, TTG3 and MYB2 from both A and D genomes species, and only DEL65 from A diploid species complemented the trichomeless phenotype.

Since Arabidopsis has functionally redundant bHLH proteins (GL3 and EGL3) involved in trichome initiation, the cotton diploids might contain functionally redundant proteins contributing to fiber formation. Genome wide analysis for the presence of DEL65 homolog in A and D diploid was performed and it was found the presence of a closely-related protein, a single copy DEL61 in both genomes (A genome: from position 95853928 bp to position 95856825 bp and D genome: from position 1682788 bp to position 1685967 bp). The DEL61 was amplified from A and D diploid species independently, and subsequently transformed into gl3-1 egl3-77439 double mutants. Interestingly, the DEL61 from both diploids did not complement the trichomeless phenotype of the double mutant (Fig. 1). Taken together, the complementing assays illustrated that the lack of spinnable fiber production in the D- diploids could be attributed to the complete absence of functional DEL65 and DEL61. The non-functionality of these two bHLH proteins may not be the only reason to explain why D genome does not confer spinnable fiber production, but it could be one of the key factors missing in D diploid species.

Conclusively, based on the Arabidopsis trichome model system, it demonstrated the functional differences in DEL65 between A and D diploid species. Sequence comparison demonstrated that there is 96.7% homology in DEL65-A and DEL65-D at DNA level and yet DEL65-D is not functional, therefore, it is highly fascinating to investigate the molecular basis for functional differences of DEL65 in future studies.

Gene expression analysis: Trichomes are well patterned on Arabidopsis leaves due to the lateral inhibition mechanism (Schellmann *et al*., 2002) while there is no apparent pattern in fiber formation on cotton seeds. Our complementing assays illustrated that transgenic lines with MYB2, TTG3 from genome A and D, DEL65 from genome A rescued the trichomeless phenotype of *gl1-1*, *ttg1-1*, *gl3-1 egl3-77439*, respectively. However, trichome initiation in transgenic lines with one cotton gene in the Arabidopsis trimeric complex still reflected defined pattern on Arabidopsis leaves (Fig. 1).

In order to answer the question of pattern difference, quantitative PCR was conducted to observe if there is any difference in gene expression of downstream target genes responsible for trichome promoting regulated by trimeric complex in wild type, trichomeless mutants, and transgenic lines from cotton genomes A and D. Eight candidate target genes were chosen from published data (Morohashi & Grotewold, 2009).

Figure 2 shows differential expression of four positive and four negative regulators in Arabidopsis leaves trichome initiation in wild type, glabrous mutant *gl1- 1* and its respective transgenic lines complemented from diploid A and D cotton genomes. The expression of four positive regulators were significantly elevated in transgenic *Ga*MYB2/ *gl1-1* and *Gr*MYB2/*gl1-1*, which could be attributed to the trichome initiation phenotype compared to the mutant *gl1-1*. It has been documented that GL2 is required for Arabidopsis trichome initiation of leaves while TTG2 is predominantly up-regulated in

trichomes throughout their development. Two other positive regulators, i.e. HDG11 and SIM are responsible for normal trichome branching and development, respectively. The down-regulated expression of R3 MYB repressor genes, namely CPC, ETC1, TCL1 and TRY, was consistent with the over production of trichomes on leaves of 35S:: MYB2-A/*gl1-1* and 35S:: MYB2-D/*gl1-1* compared to *gl1-1*. Similar regulation was observed in eight target genes stimulated by trimeric complex in transgenic plants with TTG3 from cotton diploids A and D genomes compared with wild type and *ttg1-1* mutant (Fig. 3).

The regulation of downstream target genes in transgenic lines *gl3-1 egl3-77439* transformed with DEL65 from cotton A and D genome species further confirmed the consistency from our complementing assays (Fig. 4). The transcript levels of positive trichome regulators including GL2, TTG2, HDG11, and SIM were increased, and trichome suppressors such as TRY, ETC1 and TCL1 were and decreased in 35S:: DEL65-A/gl3-1 egl3-77439 line than in 35S:: DEL65-D/*gl3-1 egl3-77439*. Surprisingly, it was noticed that a significant increase (at least three fold) in one R3 MYB transcripts, CPC in 35S:: DEL65-A/*gl3-1 egl3-77439* line, which were not in agreement with leaf trichome phenotype (Fig. 4). It was speculated that the increasing level of CPC functioning as a counteractor to suppress the exceedingly high elevation of DEL65 transgene, thus developing an equilibrium feedback control to prevent supernumerary trichomes.

Discussion

In Arabidopsis, previous genetic analysis has revealed that trichome initiation is positively mediated by a trimeric activation complex comprised of GL1, GL3 which acts redundantly with its close homolog EGL3, and TTG1 (Sato *et al*., 2019; Wang *et al*., 2019). The functionality of four individual cotton genes MYB2, TTG3, DEL65 and DEL61 from diploid cotton A and D genomes was studied, which show high similarity in sequence with GL1, TTG1, GL3 and EGL3, respectively, in their respective Arabidopsis glabrous mutants. Our complementation assays proved that transgenic lines with MYB2, TTG3 from diploid genomes A and D could rescue the trichomeless phenotype of *gl1-1* and *ttg1-1* respectively however, DEL61 from both the species could not rescue this phenotype of *gl3-1 egl3-77439* double mutant. Interestingly, the DEL65 from A species rescued *gl3-1 egl3-77439* double mutant but not from D diploid species. Comparative quantitative PCR analysis of the downstream regulatory network genes showed a similar pattern for MYB2, TTG3 complemented lines from A- and D- diploid species. Comparative analysis of the DEL65 from A- (rescued the trichomeless phenotype) and D- (did not rescued the trichomeless phenotype) showed differential expression of regulatory network genes between these two lines. These results showed that the cotton genes, when expressed in Arabidopsis, regulate the regulatory network genes during the trichome initiation process.

Avarge fold change of positive regulators in wild type, gl1-1 mutant and complemented plants with
MYB2 from cotton diploid A and D genomes

Avarge fold change of negative regulators in wild type, gl1-1 mutant and complemented plants with
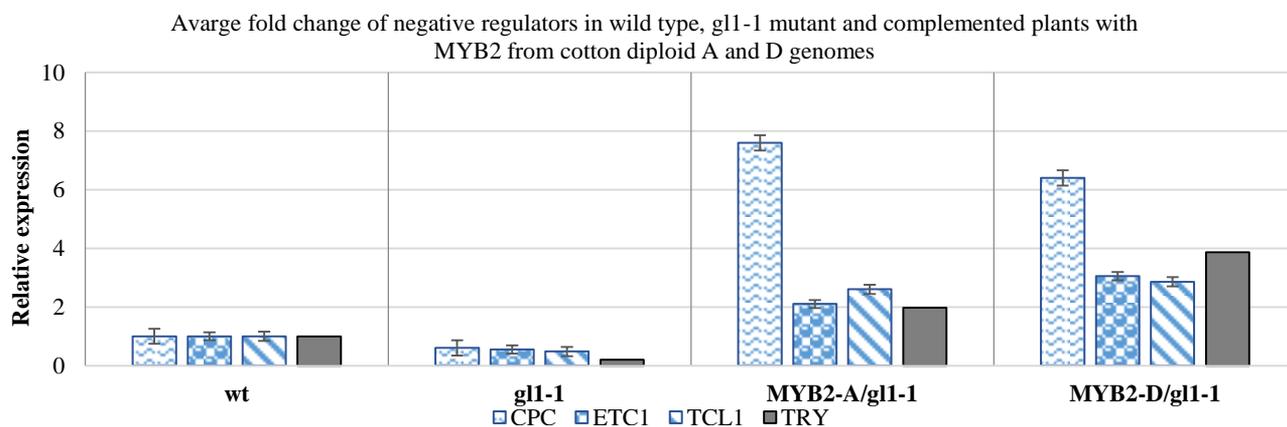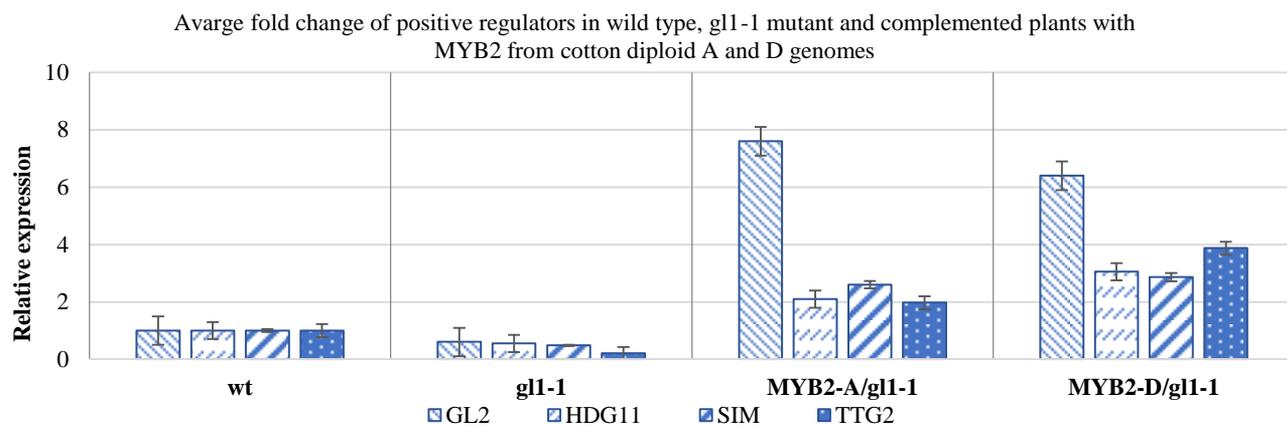MYB2 from cotton diploid A and D genomes

Fig. 2. Transcript levels for individual gene copies of the four trichome negative genes and four positive regulatory genes in *Arabidopsis* wild type, *gl1-1*, 35S::*MYB2*-A/ *gl1-1* and 35S::*MYB2*-A/ *gl1-1*

Avarge fold change of positive regulators in wild type, ttg1-1 mutant and complemented plants with
TTG3 from cotton diploid A and D genomes

Avarge fold change of negative regulators in wild type, gl1-1 mutant and complemented plants with
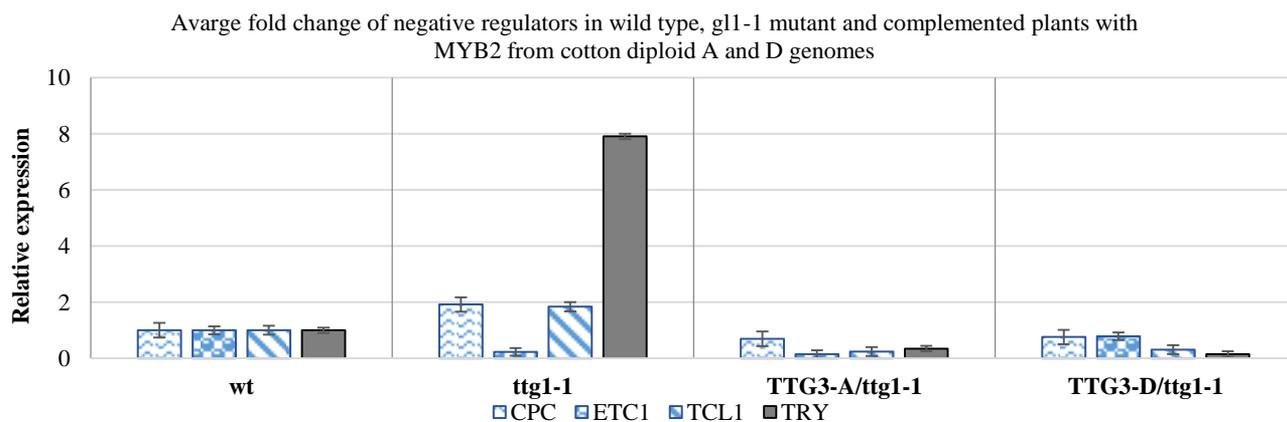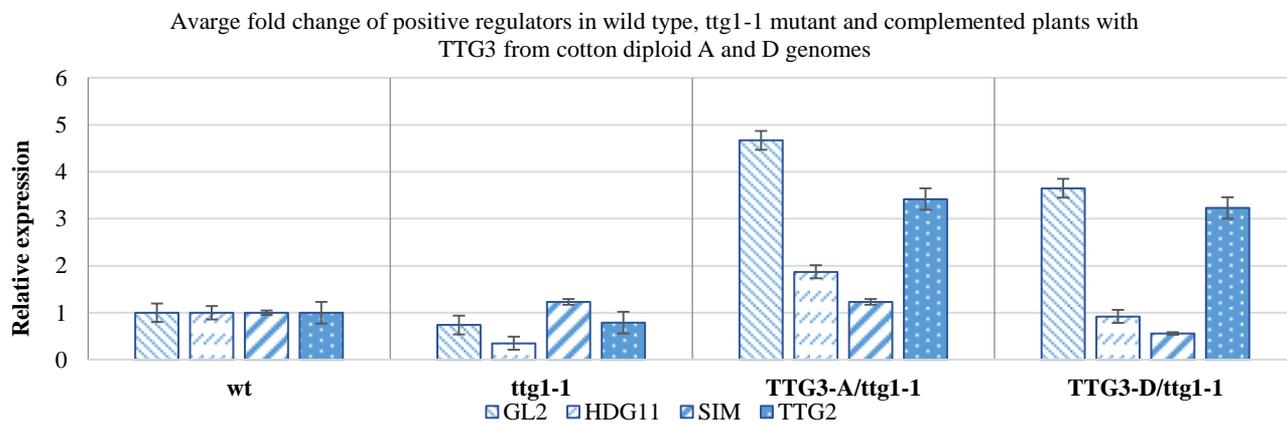MYB2 from cotton diploid A and D genomes

Fig. 3. Transcript levels for individual gene copies of the four trichome negative regulatory genes and four positive regulatory genes in *Arabidopsis* wild type, double mutant *ttg1-1,* 35S::*TTG3*-A/ *ttg1-1* and 35S::*TTG3*-D/ *ttg1-1*

Avarge fold change of positive regulators in wild type, ttg1-1 mutant and complemented plants with TTG3 from cotton diploid A and D genomes

Avarge fold change of negative regulators in wild type, gl1-1 mutant and complemented plants with MYB2 from cotton diploid A and D genomes
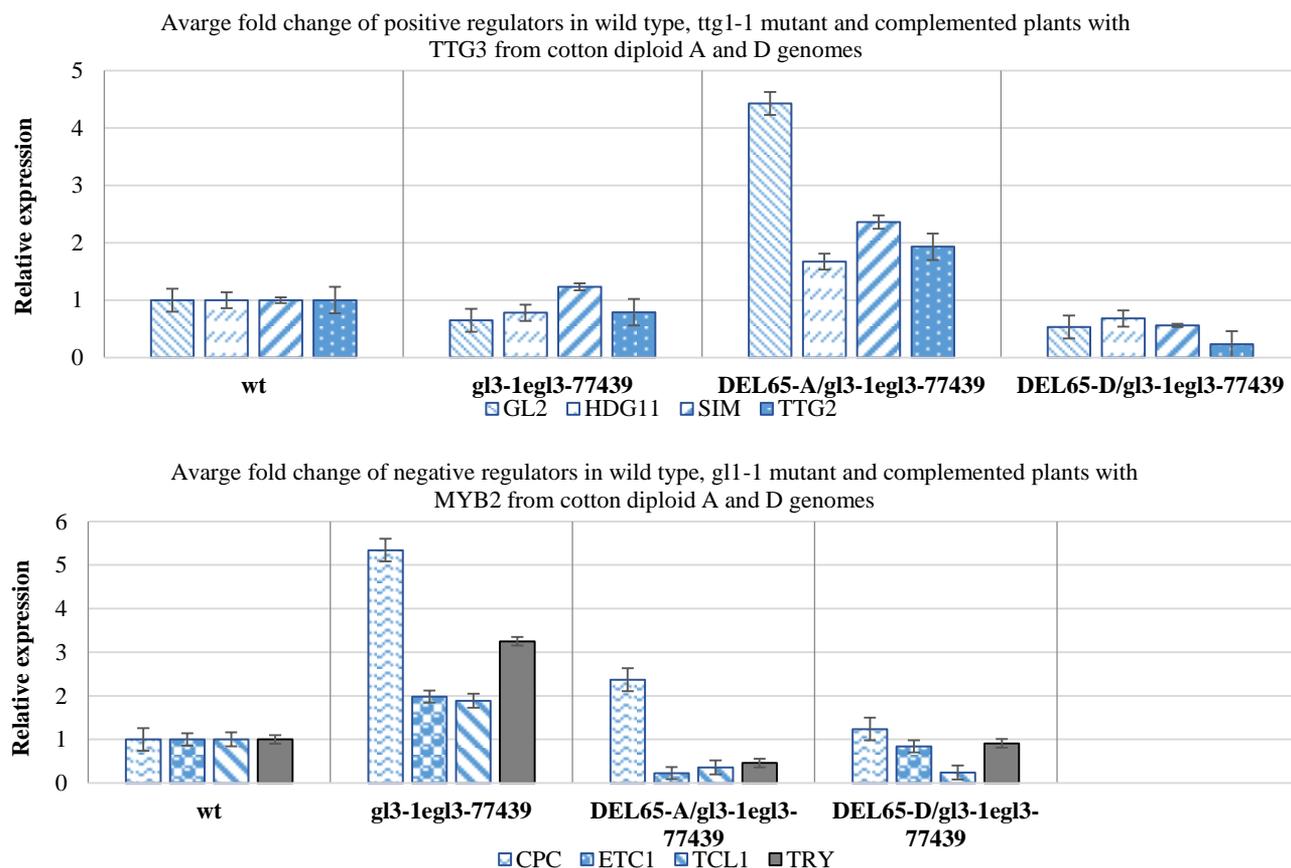

Fig. 4. Transcript levels for individual gene copies of the four trichome negative regulatory genes and four positive regulatory genes in *Arabidopsis* wild type, *gl3-1 egl3-77439*, 35S::*DEL65*-A/ *gl3-1 egl3-77439* and 35S::*DEL65*-D/ *gl3-1 egl3-77439*

Conclusion

In this paper, the Arabidopsis trichome model system was employed to examine the mechanism of spinnable fiber production trait. The core trichome initiation complex consisting of bHLH, WD40 and R2R3-MYB proteins was tested for functional differences between A and D genomes, the parental species of the cultivated tetraploid species. The only discrepancy in our complement assays was DEL65 from A- genome species, not D- genome species, complemented the *gl3-1 egl3-77439* trichomeless phenotype, reflecting the consistency of the observation that spinnable fiber production trait is absent in D-diploid species. We also tested the regulation of eight target genes stimulated by the trimeric complex in six different transgenic lines transformed with three cotton genes. The expression of trichome positive regulator was significantly elevated which could be attributed to the trichome initiation phenotype, whereas the transcripts level of trichome suppressors were down regulated. However, CPC levels in 35S:: DEL65-A/*gl3-1 egl3-77439* line were increased, possibly demonstrating a feedback loop to avoid extreme number of trichome initiation. Conclusively, by implementing trichome initiation in Arabidopsis as a model, this study provided functional characterization of four cotton important genes from two diploid cotton genomes A and D in fiber initiation. However, more studies should be conducted in other diploid cotton genomes to elucidate the understanding the mechanism of fiber initiation mechanism in cotton.

References

Arpat, A., M. Waugh, J. Sullivan, M. Gonzales, D. Frisch, D. Main, and T. Wilkins. 2004. Functional genomics of cell elongation in developing cotton fibers. *Plant Mol. Biol.,* 54(6): 911-929. doi:10.1007/s11103-004-0392-y

Clough, S.J. and A.F. Bent. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *The Plant J.,* 16(6): 735-743. doi:10.1046/j.1365-313x.1998.00343.x

Esch, J.J., M. Chen, M. Sanders, M. Hillestad, S. Ndkium, B. Idelkope and M.D. Marks. 2003. A contradictory GLABRA3 allele helps define gene interactions controlling trichome development in Arabidopsis. *Development,* 130(24): 5885-5894. doi:10.1242/dev.00812

Fryxell, P.A., L.A. Craven and J.M. Stewart. 1992. A revision of gossypium sect. grandicalyx (Malvaceae), including the description of six new species. *Syst. Bot.,* 17(1): 91-114. doi:10.2307/2419068

Glover, B.J. 2000. Differentiation in plant epidermal cells. *J. Exp. Bot.,* 51(344): 497-505. doi:10.1093/jexbot/51.344.497

Guan, X.Y., Q.J. Li , C.M. Shan, S. Wang, Y.B. Mao, L.J. Wang and X.Y. Chen. 2008. The HD-Zip IV gene GaHOX1 from cotton is a functional homologue of the Arabidopsis GLABRA2. *Physiol. Plant.,* 134(1): 174-182. doi:10.1111/j.1399-3054.2008.01115.x

Guan, X., J.J. Lee, M. Pang, X. Shi, D.M. Stelly and Z.J. Chen. 2011. Activation of arabidopsis seed hair development by cotton fiber-related genes. *PLoS One,* 6(7): e21301. doi:10.1371/journal.pone.0021301

Guan, X., M. Pang, G. Nah, X. Shi, W. Ye, D.M. Stelly and Z.J. Chen. 2014. miR828 and miR858 regulate homoeologous MYB2 gene functions in Arabidopsis trichome and cotton

fibre development. 5: 3050. doi:10.1038/ncomms4050 https://www.nature.com/articles/ncomms4050#supplementary-information

Guan, X., N. Yu, X. Shangguan, S. Wang, S. Lu, L. Wang and X. Chen. 2007. Arabidopsis trichome research sheds light on cotton fiber development mechanisms. *Chinese Sci. Bull.,* 52(13): 1734-1741. doi:10.1007/s11434-007-0273-2

Hulskamp, M. 2004. Plant trichomes: a model for cell differentiation. *Nat. Rev. Mol. Cell Biol.,* 5(6): 471-480. doi:http://www.nature.com/nrm/journal/v5/n6/suppinfo/nrm1404_S1.html

Humphries, J., A. Walker, J. Timmis and S. Orford. 2005. Two WD-repeat genes from cotton are functional homologues of the Arabidopsis thaliana TRANSPARENT TESTA GLABRA1 (TTG1) gene. *Plant Mol. Biol.,* 57(1): 67-81. doi:10.1007/s11103-004-6768-1

Junker, A., M.M. Muraya, K. Weigelt-Fischer, F. Arana-Ceballos, C. Klukas, A.E. Melchinger and T. Altmann. 2015. Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems. *Front. Plant Sci.,* 5: 770-770. doi:10.3389/fpls.2014.00770

Larkin, J.C., M.L. Brown and J. Schiefelbein. 2003. How do cells know what they want to be when they grow up? Lessons from Epidermal Patterning in Arabidopsis. *Ann. Rev. Plant Biol.,* 54(1): 403-430.

Morohashi, K. and E. Grotewold. 2009. A systems approach reveals regulatory circuitry for arabidopsis trichome initiation by the GL3 and GL1 selectors. *PLoS Genet,* 5(2): e1000396. doi:10.1371/journal.pgen.1000396

Oppenheimer, D.G., P.L. Herman, S. Sivakumaran, J. Esch and M.D. Marks. 1991. A *MYB* gene required for leaf trichome differentiation in Arabidopsis is expressed in stipules. *Cell,* 67(3): 483-493.

Payne, C.T., F. Zhang and A.M. Lloyd. 2000. GL3 encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. *Genetics,* 156(3): 1349-1362.

Ramsay, N.A. and B.J. Glover. 2005. MYB–bHLH–WD40 protein complex and the evolution of cellular diversity. *Trends in Plant Sci.,* 10(2): 63-70. doi:http://dx.doi.org/10.1016/j.tplants.2004.12.011

Sato, Y., R. Shimizu-Inatsugi, M. Yamazaki, K.K. Shimizu and A.J. Nagano. 2019. Plant trichomes and a single gene GLABRA1 contribute to insect community composition on field-grown Arabidopsis thaliana. *BMC Plant Biol.,* 19(1): 163. doi:10.1186/s12870-019-1705-2

Schellmann, S., A. Schnittger, V. Kirik, T. Wada, K. Okada, A. Beermann and M. Hülskamp. 2002. TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in Arabidopsis. *The EMBO J.,* 21(19): 5036-5046.

Schiefelbein, J. 2003. Cell-fate specification in the epidermis: a common patterning mechanism in the root and shoot. *Curr. Opin. in Plant Biol.,* 6(1): 74-78. doi:http://dx.doi.org/10.1016/S136952660200002X

Serna, L. 2004. A network of interacting factors triggering different cell fates. *The Plant Cell,* 16(9): 2258-2263.

Serna, L. and C. Martin. 2006. Trichomes: different regulatory networks lead to convergent structures. *Trends in Plant Sci.,* 11(6): 274-280.

Shi, Y.H., S.W. Zhu, X.Z. Mao, J.X. Feng, Y.M. Qin, L. Zhang and Y.X. Zhu. 2006. Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. *The Plant Cell,* 18(3): 651-664.

Stephens, S.G. and M.E. Moseley. 1974. early domesticated cottons from archaeological sites in central Coastal Peru. *Amer. Antiq.,* 39(1): 109-122.

Szymanski, D.B., R.A. Jilk, S.M. Pollock and M.D. Marks. 1998. Control of GL2 expression in Arabidopsis leaves and trichomes. *Development,* 125(7): 1161-1171.

Tominaga, R., M. Iwata, R. Sano, K. Inoue, K. Okada and T. Wada. 2008. Arabidopsis CAPRICE-LIKE MYB 3 (CPL3) controls endoreduplication and flowering development in addition to trichome and root hair formation. *Development,* 135(7): 1335-1345.

Udall, J.A., J.M. Swanson, K. Haller, R.A. Rapp, M.E. Sparks, J. Hatfield and J.F. Wendel. 2006. A global assembly of cotton ESTs. *Genome Res.,* 16(3): 441-450.

Walker, A.R., P.A. Davison, A.C. Bolognesi-Winfield, C.M. James, N. Srinivasan, T.L. Blundell and J.C. Gray. 1999. The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. *The Plant Cell,* 11(7): 1337-1350.

Wang, G., G.H. Zhao, Y.H. Jia and X.M. Du. 2013. Identification and characterization of cotton genes involved in Fuzz-fiber development. *J. Integ. Plant Biol.,* 55(7): 619-630.

Wang, S., L. Hubbard, Y. Chang, J. Guo, J. Schiefelbein and J.G. Chen. 2008. Comprehensive analysis of single-repeat R3 MYB proteins in epidermal cell patterning and their transcriptional regulation in Arabidopsis. *BMC Plant Biol.,* 8: 81-81.

Wang, S., S.H. Kwak, Q. Zeng, B.E. Ellis, X.Y. Chen, J. Schiefelbein and J.G. Chen. 2007. TRICHOMELESS1 regulates trichome patterning by suppressing GLABRA1 in Arabidopsis. *Development,* 134(21): 3873-3882.

Wang, S., J.W. Wang, N. Yu, C.H. Li, B. Luo, J.Y. Gou and X.Y. Chen. 2004. Control of plant trichome development by a cotton fiber MYB gene. *The Plant Cell,* 16(9): 2323-2334.

Wang, Z., Z. Yang and F. Li. 2019. Updates on molecular mechanisms in the development of branched trichome in Arabidopsis and nonbranched in cotton. *Plant Biotech. J.,* 17(9): 1706-1722. doi:https://doi.org/10.1111/pbi.13167

Wendel, J.F. and R.C. Cronn. 2003. Polyploidy and the evolutionary history of cotton. In: *Advances in Agronomy*, 78: pp. 139-186): Academic Press.

Wendel, J.F., L.E. Flagel and K.L. Adams. 2012. Jeans, genes, and genomes: Cotton as a model for studying polyploidy. In: *Polyploidy and Genome Evolution.* (Eds.): S.P. Soltis & E.D. Soltis, Berlin, Heidelberg: Springer Berlin Heidelberg. pp. 181-207.

Wester, K., S. Digiuni, F. Geier, J. Timmer, C. Flec and M. Hülskamp. 2009. Functional diversity of R3 single-repeat genes in trichome development. *Development,* 136(9): 1487-1496.

Wu, Y., A.C. Machado, R.G. White, D.J. Llewellyn and E.S. Dennis. 2006. Expression profiling identifies genes expressed early during lint fibre initiation in cotton. *Plant & Cell Physiol.,* 47(1): 107-127.

Yang, S., F. Cheung, J.J. Lee, M. Ha, N.E. Wei, S.H. Sze and Z.J. Chen. 2006. Accumulation of genome-specific transcripts, transcription factors and phytohormonal regulators during early stages of fiber cell development in allotetraploid cotton. *The Plant J.,* 47(5): 761-775. doi:10.1111/j.1365-313X.2006.02829.x

Yuan, J., H. Yue, M. Zhang, J. Luo, L. Liu, W. Wu and R. Chen. 2016. Transcriptional profiling analysis and functional prediction of long noncoding RNAs in cancer. *Oncotarget,* 7(7): 8131-8142.

Zhang, B.H. and R. Feng. 2000. Cotton resistance to insect and pest-resistant cotton. *Beijing: Chinese Agri. Sci. & Technol. Press.*

Zhang, F., A. Gonzalez, M. Zhao, C.T. Payne and A. Lloyd. 2003. A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. *Development,* 130(20): 4859-4869.