Plant Molecular Biology (2005) 57:67–81
DOI 10.1007/s11103-004-6768-1

Two WD-repeat genes from cotton are functional homologues of the *Arabidopsis thaliana TRANSPARENT TESTA GLABRA1* (*TTG1*) gene

John A. Humphries[1], Amanda R. Walker[2], Jeremy N. Timmis[1,*] and Sharon J. Orford[1]
[1]*School of Molecular and Biomedical Science, The University of Adelaide, Adelaide SA 5005, Australia (*author for correspondence; e-mail jeremy.timmis@adelaide.edu.au);* [2]*CSIRO Plant Industry, P.O. Box 350, Glen Osmond SA 5064, Australia*

Received 19 July 2004; accepted in revised form 25 November 2004

Key words: *Arabidopsis thaliana*, complementation, cotton fibre, trichome initiation, WD-repeat

Abstract

Cotton fibres are single, highly elongated cells derived from the outer epidermis of ovules, and are developmentally similar to the trichomes of *Arabidopsis thaliana*. To identify genes involved in the molecular control of cotton fibre initiation, we isolated four putative homologues of the *Arabidopsis* trichome-associated gene *TRANSPARENT TESTA GLABRA1* (*TTG1*). All four WD-repeat genes are derived from the ancestral D diploid genome of tetraploid cotton and are expressed in many tissues throughout the plant, including ovules and growing fibres. Two of the cotton genes were able to restore trichome formation in *ttg1* mutant *Arabidopsis* plants. Both these genes also complemented the anthocyanin defect in a white-flowered *Matthiola incana ttg1* mutant. These results demonstrate parallels in differentiation between trichomes in cotton and *Arabidopsis*, and indicate that these cotton genes may be functional homologues of *AtTTG1*.

Introduction

Cotton fibres (lint) are single, highly elongated cells derived from the ovule epidermis of plants of the genus *Gossypium*. Although all ovule epidermal cells are competent to differentiate into fibre cells for several days before flowering (Graves and Stewart, 1988), only about one in four becomes a fibre (Stewart, 1975). The signals which control the initiation of fibre growth and their distribution are poorly understood at the molecular level, but it is thought that differentiation is triggered by plant growth substances, including auxins and gibberellic acids (Beasley and Ting, 1974). A second type of differentiated epidermal cell (fuzz) initiates growth up to several days after the lint fibres and, being very short, is of little economic value. The timing of initiation of cotton fibres is therefore an important step in obtaining a commercially useful product.

The single-celled nature of cotton fibres and their synchronous growth has made them an attractive system for the study of plant cell elongation and cell wall biogenesis (Basra and Malik, 1984). A number of structural proteins which appear to play a role in fibre elongation or secondary cell wall synthesis have been isolated (reviewed in Wilkins *et al.*, 2000) and fundamental studies on these genes have led to the first molecular and cellular model of fibre development (Wilkins and Jernstedt, 1999). Some success in the genetic alteration of fibre growth, based on the model, has been attained (Wilkins *et al.*, 2000), providing opportunities to manipulate fibre traits *via* the biosynthetic and/or catabolic pathways. However, genes involved in the molecular programming of fibre initiation have not yet been identified.

Arabidopsis and cotton are close relatives (Soltis *et al.*, 1999) and a number of investigators

have proposed that *Arabidopsis* leaf trichomes could serve as a model for elucidating the genetic mechanisms controlling cotton fibre differentiation. Both structures are single-celled, non-glandular hairs of epidermal origin. The initial stages of development, that proceed across the epidermal surface in a synchronous wave, appear similar in that both involve the outward expansion of the cells and the migration of nuclei into the elongated stalks. A second round of trichome initiation which occurs on *Arabidopsis* leaves as the epidermal cells divide (Larkin *et al*., 1996) may be analogous to the fuzz fibres of cotton. There is evidence of limited DNA endoreplication during the early stages of cotton fibre growth (Van't Hof, 1999) resembling that which occurs prior to *Arabidopsis* trichome differentiation.

Arabidopsis trichomes cover most aerial plant surfaces including the leaves, stems, and sepals. Their cytological simplicity and visibility make them an attractive system for the study of plant cell differentiation and morphogenesis. Genetic screens have identified numerous mutants of trichome formation and spatial distribution, and all the principal genes involved have been cloned, and an integrated picture of the interactions between the components is emerging (Szymanski *et al*., 2000; Walker and Marks, 2000; Schellmann *et al*., 2002).

Mutations in only two genes result in a complete loss of trichomes. The first of these, *GLABRA1* (*GL1*), encodes an R2R3-type MYB-related transcription factor (Oppenheimer *et al*., 1991) and the second is *TRANSPARENT TESTA GLABRA1* (*TTG1*). In *Arabidopsis* the *TTG1* gene product is involved in several pathways in addition to its function in trichome formation. These pleiotropic effects include the control of synthesis of anthocyanin pigments, the production of seed coat mucilage and the development of root hairs (Koornneef, 1981; Galway *et al*., 1994). The *TTG1* locus was mapped to chromosome five in *Arabidopsis,* and the positionally cloned gene encodes a protein of 341 amino acid residues containing four WD-40 repeats (Walker *et al*., 1999). Structural modelling identified over 30 functional families of WD-repeat-containing proteins (Yu *et al*., 2000) which are implicated in a diverse range of cellular processes including signal transduction, RNA processing, gene regulation, vesicular traffic and regulation of the cell cycle (Neer *et al*., 1994).

The identification of TTG1 as a WD-repeat protein suggests that it interacts with other proteins to control trichome initiation. In *Arabidopsis*, its physical interaction with GLABRA3 (GL3), a bHLH transcription factor, suggests that an initiation complex consisting of GL1, GL3 and TTG1 activates trichome cell fate (Payne *et al*., 2000; Szymanski *et al*., 2000; Schellmann *et al*., 2002). Several weak mutant *ttg1* alleles cause clustered trichomes (Larkin *et al*., 1994), suggesting that TTG1 also plays a role in trichome spacing *via* the lateral inhibition of differentiation in neighbouring cells.

To identify genes involved in the control of fibre initiation, we have targeted cotton homologues of *AtTTG1*. We describe here the isolation and characterisation of four WD-repeat transcription factor genes. Two of these, *GhTTG1* and *GhTTG3*, showed close sequence similarity to each other and to known anthocyanin regulators, particularly *TTG1* from *Arabidopsis*, *an11* from petunia and PFWD from *Perilla frutescens*. In functional assays, both were able to restore trichome development in *Arabidopsis ttg1* mutants, and one, *GhTTG3*, was shown to complement the full range of *ttg1* mutant phenotypes. Both genes complemented the white-petal phenotype in *Matthiola incana ttg1* mutants. This paper describes the first functional characterisation of a regulatory gene from cotton which may be pivotal in fibre differentiation, and provides a molecular link between the development of *Arabidopsis* trichomes and cotton fibres.

Materials and methods

Plant material and growth conditions

Seeds of *G. hirsutum* ($[AD]_1$ genome) and *G. barbadense* ($[AD]_2$ genome) varieties were propagated and material collected as previously described (Orford and Timmis, 2000). Material from *G. herbaceum* (A_1 genome) and *G. raimondii* (D_5 genome) was obtained from Dr C. Brubaker (CSIRO Division of Plant Industry, Canberra). The *Matthiola incana* mutants were line 17 (*ttg1*) and line 18 (*chalcone synthase; chs*) (Kappert, 1949; Epping *et al*., 1990), both of which are

defective in the anthocyanin pathway and have a white petal phenotype. Individual *Matthiola* plants were grown in 15 cm pots at 25 °C (maximum) and 18 °C (minimum) in glasshouses. A selection of flowers was removed from each of the two lines, and placed onto Petri dishes containing Murashige and Skoog (MS) media and 0.75% (w/v) agar for bombardment. *Arabidopsis thaliana ttg1-1* (L*er*) *mutant* seeds were mixed with sand and sprinkled onto 'Arabidopsis soil mix' (SARDI, Glen Osmond, South Australia) in 3-in pots. Plantlets were grown at 23 °C (light/dark cycle of 8/16 h) and thinned to 5–10 plants per pot at 2–3 weeks. Plants were allowed to grow for several weeks before transferral to flower-inducing conditions (23 °C, light/dark cycle of 18/6 h). The first flowering bolts (10 cm tall) were cut back once, and new bolts allowed to emerge for several days before infiltration.

Library screening, cloning and sequencing

Two degenerate primers (TL3, 5′-TTYGAICAYCCITAYCCICCIACIAARYIIAGTTT-3′ and TL4, 5′-CATIGGGTCRATICCRTTIGGICCIGCIACIGTNGG-3′) were designed from *AtTTG1* to amplify 700 bp from putative cotton homologues. Amplified fragments were ligated to pGEM-®T Easy (Promega) and used to transform *E. coli* DH5-α F′ cells (Promega). The insert from one clone was released by *Eco*RI digestion, oligolabelled (Hodgson and Fisk, 1987) and used to probe a 12 DPA (days post-anthesis) cotton fibre cDNA library.

Double-stranded cDNA was synthesised from fibre mRNA (Orford and Timmis, 1997) using Superscript™ II (Life Technologies) and DNA Polymerase I (Pharmacia) for the first and second strand synthesis respectively, following manufacturer's instructions. cDNA was blunt-ended using Vent^{R} DNA Polymerase (New England Biolabs) and adapted as described previously (Orford and Timmis, 1997). cDNA was ligated to *Eco*RI-digested and dephosphorylated λZAP®II vector (Stratagene), packaged using *in vitro* packaging extracts (Promega) and plated according to Stratagene instructions. Phage were lifted in duplicate onto Hybond-N^{+} membranes (Amersham, UK) and screened as described in Orford and Timmis (1997).

The same PCR clone was used to probe a *G. hirsutum* cv. Siokra 1-4 genomic library in λGEM12 (Promega), consisting of 8×10^5 clones and obtained from Dr D. Llewellyn (CSIRO Division of Plant Industry, Canberra, ACT). DNA was prepared from five positive plaques (Sambrook *et al.*, 1989). Hybridising regions from the three distinct clones, as determined by Southern analysis, were subcloned into pBluescript®SK(-) (Stratagene) and sequenced.

Sequence analysis

Automated sequencing was carried out on DNA purified with the Quantum Prep Plasmid Miniprep Kit (BioRad) or the QIAquick PCR Purification Kit (QIAGEN), using a Prism BigDye Terminator kit (Applied Biosystems–Perkin Elmer). The nucleotide sequence of both strands of DNA was determined using vector primers and primers designed from previously obtained sequence. Products were analysed at the Institute of Medical and Veterinary Science, Adelaide, South Australia, using an ABI PRISM® 377 DNA sequencer. Sequence analysis utilised the GCG Sequence Analysis Software Package version 8 (Genetics Computer Group, Madison, WI 1984) available online from the Australian National Genomic Information Service, University of Sydney (http://www.angis.org.au). FASTA (Pearson and Lipman, 1988) and BLAST (Altschul *et al.*, 1990) programs were used to search non-redundant nucleotide and protein sequence databases, respectively. Peptide sequences were searched for WD repeats using the protein prediction program available at the Biomolecular Engineering Research Centre (Boston University; http://bmerc-www.bu.edu/). Nucleotide sequence data were deposited in the DDBJ/EMBL/GenBank database with accession numbers AF530907 (GhTTG1 cDNA), AF530908 (*GhTTG1*), AF530909 (GhTTG2 cDNA), AF530910 (*GhTTG2*), AF530911 (*GhTTG3*) and AF530912 (*GhTTG4*).

PCR of cotton TTG1-like sequences

Gene-specific primers were designed as follows: PTTG1, 5′-GACGGTTGGCAACTCCCAAATAAG-3′ and PTTG3, 5′-TCCCCTTTCTCTTCCTCCTCCGAC-3′ (*GhTTG1*); PTTG5, 5′-AGTGTCCGCCGCGACAAGAAATACCGTC-3′ and

PTTG9, 5′-GAGGAGTGCTCCTTGTCGCGC-AGG-3′ (*GhTTG2*); PTTG1 and PTTG2, 5′-CG-CTTCATCCTCTTCTTCTTGTTC-3′ (*GhTTG3*); PTTG10, 5′-CTCCGCAACTCAACCACCGG-CAAC-3′ and PTTG11, 5′-TACCGACTCGCC-ATCGCCAGCCTC-3′ (*GhTTG4*). Genomic DNA and clone DNA templates were amplified using standard reactions containing approximately 100 ng of genomic DNA or 20 pg of cloned DNA, 0.4 mM each dNTP, 1 μM each primer, 2 mM $MgCl_2$ and 1.5 units *Taq* DNA polymerase (GeneWorks) in a final volume of 25 μl of 1 × PCR reaction buffer. Initial template denaturation was at 94 °C for 4 min followed by 30 PCR cycles of denaturation at 94 °C for 30 s, primer annealing for 30 s and extension at 72 °C for 1 min. The annealing temperatures were 70, 75, 68 and 75 °C for *GhTTG1-4* respectively.

Southern analyses

Southern blots were generated and hybridised with radioactively-labelled probe as described previously (Harmer *et al*., 2002). The probes for genomic Southern blots were the full-length cDNA inserts for *GhTTG1* and *GhTTG2*, a genomic PCR clone for *GhTTG3*, and a fragment from a genomic subclone for *GhTTG4*. Prehybridisation, hybridisation and washing of Southern blots were performed as recommended by the manufacturer.

Reverse transcriptase polymerase chain reaction

RT-PCR amplification was as described previously (Orford and Timmis, 2000) except that the RNA was first treated with RNase-free DNase (Roche) to remove any contaminating genomic DNA. A 1/30 sample of RT products was used in all cases, with primers as above. Actin transcripts were detected using the PCR primers of Shimizu *et al*. (1997).

GenomeWalker™ PCR

The *GhTTG2* and *GhTTG3* genes were obtained using a PCR-based technique. Nested, gene-specific primers were designed from PCR clone sequence (*GhTTG3*) or cDNA clone sequence (*GhTTG2*) and applied to total genomic DNA according to the manufacturer's protocol for the Universal GenomeWalker PCR kit (Clontech). Resultant PCR products of appropriate size were ligated to pGEM®-T Easy (Promega) and sequenced. A further walk was performed and the complete gene sequence compiled from overlapping fragments. PCR primers were designed to flank each gene and were as follows: PTTG2G1, 5′-TTTCACGTGAATGGAAAATACTAATGC-G-3′ and PTTG2G2, 5′-TTGGCCATTGATGG-CATTGCTATATTGC-3′ (*GhTTG2*); PTTG3G1, 5′-AATGTATAAAGCTTGTGGTTCTGGGTT-C-3′ and PTTG3G2, 5′-GTTGGTAAAAGTAT-CATGAAGGTCCCTG-3′ (*GhTTG3*). These were used to amplify the entire *GhTTG2* and *GhTTG3* genes directly from cotton genomic DNA, and at least three independent products were sequenced in order to obtain a consensus sequence for each gene.

Gene constructs for complementation experiments

The coding regions of the *Arabidopsis* TTG1 cDNA, the GhTTG1 and GhTTG2 cDNAs and the *GhTTG3* and *GhTTG4* genes were directionally subcloned into appropriately digested pART7 vector (Gleave, 1992), downstream of the cauliflower mosaic virus (CaMV) 35S constitutive promoter. Following transformation into *E. coli* DH5αF′ cells (Sambrook *et al*., 1989), the orientation and identity of the insert were verified by sequencing from primers in pART7.

Complementation of Arabidopsis *ttg1 phenotypes*

The expression cartridges of the *TTG1*, *GhTTG1*, *GhTTG2*, *GhTTG3* and *GhTTG4* pART7 constructs were removed by *Not*I restriction and inserted into similarly-cut binary vector, pART27 (Gleave, 1992). Plasmid DNA was used to transform *Agrobacterium tumefaciens* AGL1 (Lazo *et al*., 1991) by electroporation (Walkerpeach and Velten, 1994) and colonies selected on plates containing kanamycin. Mutant *ttg1-1* plants were transformed by floral dip infiltration (Clough and Bent, 1998) using a solution of *Agrobacterium* at OD_{600}=0.8 in 1.5% (w/v) sucrose and 0.05% (w/v) Silwet L-77. Plants were covered overnight to maintain humidity and grown for 4 weeks to produce seed. Selection of transformants was conducted on 0.8% agar containing Murashige and Skoog (MS) salts (2.2 g/l) and kanamycin (35 μg/ml) for 7 days. Between 8 000 and 16 000

seeds were screened for each experiment, with transformation rates varying between 0% and 0.5%. Kanamycin-resistant seedlings were moved onto fresh kanamycin plates for 10 days before transfer to soil to set seed, and were inspected microscopically for restoration of trichome formation. Progeny from self fertilised primary transformants were examined for complementation of other *ttg1* mutant phenotypes. Three-day old seedlings grown on 0.8% agar containing MS salts and 5% sucrose were inspected for anthocyanin production in cotyledons and hypocotyls. Root hair development was studied in seedlings grown vertically on plates containing MS salts and 1.2% phytagel. Seeds for scanning electron microscope (SEM) analysis of seed coat structure were mounted, sputter-coated with carbon and gold, and examined using a Philips XL20 SEM at 10 kV. Ruthenium red staining of seed coat mucilage was performed as described by Burn *et al.* (2002).

Transient expression of cotton WD-repeat genes in Matthiola incana

Particle bombardment of *Matthiola* flowers was carried out using a particle inflow gun which was built to specifications and operated as described by Vain *et al.* (1993). Each of the gene constructs in pART7 was purified using the QIAGEN midi-plasmid kit and used to coat gold microcarrier particles (Franks *et al.*, 1998). The bombardment parameters (90 PSI pressure, 20 cm distance between target and filter, helium injection for 0.12 ms) were optimised for *Matthiola* flowers, which were then placed on MS plates at 25 °C for approximately 40 h before examination for purple spots of anthocyanin. Each construct was used to bombard at least four *Matthiola* flowers.

Results

Molecular cloning and structural analysis of TTG1-like WD-repeat genes from cotton

A combination of library screening and PCR-based techniques was used to isolate four putative homologues of *Arabidopsis TTG1*, designated *GhTTG1–GhTTG4*, from cotton. Initially, two degenerate *TTG1* PCR primers were used to amplify a 700 bp fragment from cotton genomic DNA. Nucleotide sequencing of four clones revealed two distinct, but very similar, sequences. Probing of a 12 DPA cotton fibre cDNA library with one of the inserts detected two cDNA clones, one of which, GhTTG1, corresponded exactly in sequence to the genomic PCR clone used as probe (also called GhTTG1). The other cDNA, GhTTG2, was a new *TTG1*-like isolate. The second genomic PCR sequence was designated GhTTG3. Probing a cotton genomic DNA library with genomic *GhTTG1* identified three additional clones, one of which corresponded exactly in sequence to GhTTG1, whilst the remaining two clones, which overlapped with each other, contained a fourth *TTG1*-like gene, *GhTTG4*. The *GhTTG2* and *GhTTG3* genes were obtained using PCR-based genome walking, as detailed in the 'Materials and methods', providing genomic clones for all four genes.

In the cases of *GhTTG1* and *GhTTG2*, only minor sequence differences were observed between the gene and a corresponding cDNA, and it was concluded that both genes and their corresponding transcripts had been cloned. The sequence of the *GhTTG1* gene was partial, with the subcloning site located 2 bp downstream from the presumed initiation codon (data not shown). Therefore, cDNA sequences were used in all sequence comparisons which involved GhTTG1 (1253 bp) and GhTTG2 (1694 bp), and gene sequences were used in the cases of *GhTTG3* (2008 bp) and *GhTTG4* (1837 bp). Comparison between the gene and cDNA sequences showed that all four genes lack introns in their protein coding regions, as is the case for *Arabidopsis TTG1* (Walker *et al.*, 1999) and other WD-repeat genes (de Vetten *et al.*, 1997).

Sequence comparisons between the four deduced proteins and *Arabidopsis* TTG1 showed that they form two groups, with GhTTG1 (343 amino acids) and GhTTG3 (345 amino acids) being closely related to each other (87% identical and 93% similar) and to TTG1 (79% and 80% amino acid identity, respectively). GhTTG2 and GhTTG4 (both 346 amino acids) formed the second group, with 95% amino acid identity to each other and only 62–63% identity to TTG1 (Table 1).

Analysis of the four amino acid sequences showed that GhTTG1-4 are basic proteins with

Table 1. Amino acid identity (similarity) between cotton and *Arabidopsis* TTG sequences.

	AtTTG1	GhTTG1	GhTTG2	GhTTG3	GhTTG4
AtTTG1	–				
GhTTG1	79% (89%)	–			
GhTTG2	63% (79%)	62% (79%)	–		
GhTTG3	80% (89%)	87% (93%)	63% (78%)	–	
GhTTG4	62% (79%)	60% (78%)	95% (97%)	61% (77%)	–

molecular weights of 38–39 kDa. Like TTG1, all appear to have a seven-bladed repeat structure with four clearly defined WD repeats (Yu *et al.*, 2000), as indicated in Figure 1, and are therefore members of the large and diverse WD-repeat protein family. The sequence similarity between the proteins extends across the WD-repeats and to the carboxyl terminus (Figure 1).

None of the four cotton TTG1-like proteins contain nuclear localisation signals, suggesting that the proteins reside in the cytosol, or require accessory proteins to move into the nucleus and carry out their presumed role as regulators of transcription. Database searches using the four sequences confirmed the GhTTG1/GhTTG3 and GhTTG2/GhTTG4 groupings and identified similarities to WD-repeat proteins from cotton and other plants. GhTTG2 was virtually identical to a 1401 bp cDNA, wd1522 (AF336287), isolated from ovules of *G. hirsutum* cv. Acala Maxxa on the day of flowering (Matz and Burr, unpublished). The protein sequences differed by one residue, though the conceptual translation of GhTTG2 was 32 amino acids longer than that of wd1522, due to

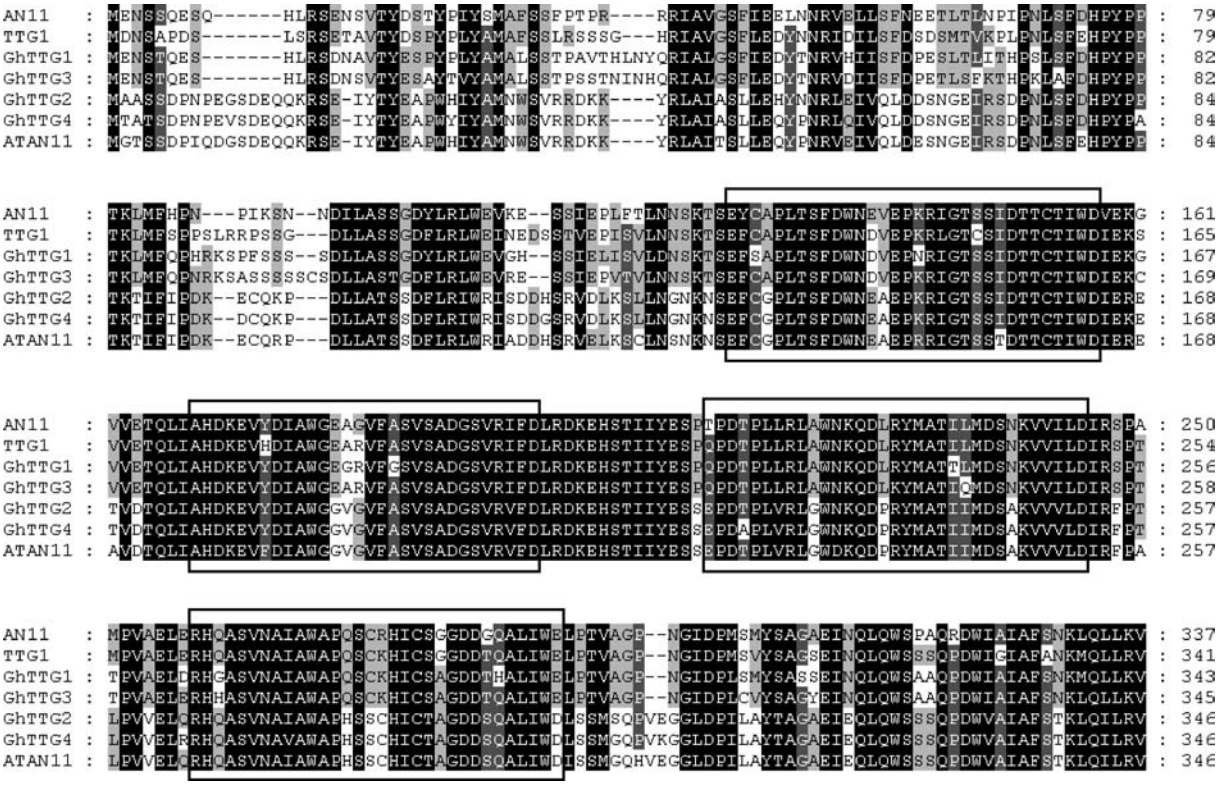

Figure 1. Deduced amino acid sequences of four WD40 genes isolated from cotton, compared with *Arabidopsis* TTG1 and ATAN11-A and *Petunia* AN11. Dark shading indicates positions where all seven sequences are identical or chemically conserved, while lighter shading represents positions where there is a lower level of conservation. The positions of the WD repeats, as determined by the BMERC Protein Structure Prediction Server (http://bmerc-www.bu.edu/) are indicated with open boxes.

different choice of start codons. The GhTTG2 and GhTTG4 deduced protein sequences both showed a striking degree of similarity to *Arabidopsis* ATAN11-A (de Vetten *et al*., 1997), with 92% and 90% amino acid identity, respectively. Lesser similarity was observed to AN11, a WD repeat-containing protein required for anthocyanin production in the flowers of *Petunia hybrida* (de Vetten *et al*., 1997), a TTG1-like protein from apple (Casas-Mollano and Destefano-Beltrán, 2000) and a number of uncharacterised, putative WD-repeat-containing proteins revealed by the *Arabidopsis* genome sequencing project.

The GhTTG1 and GhTTG3 protein sequences both showed a high degree of similarity to WD-repeat proteins which have known or implied roles in anthocyanin biosynthesis, with 79–80% amino acid identity to *Arabidopsis* TTG1 (Table 1; Walker *et al*., 1999), 79% amino acid identity to AN11 (de Vetten *et al*., 1997), 82% identity to the deduced product of a *TTG1*-like gene from apple (Casas-Mollano and Destefano-Beltrán, 2000) and 77–78% to PFWD from *Perilla frutescens* (Sompornpailin *et al*., 2002). *TTG1* and *PFWD* are thought to be orthologues of the Petunia *an11* gene. These sequence comparisons are in agreement with the extensive phylogenetic analysis of Carey *et al*. (2004), in which GhTTG1 (Cotton1) and GhTTG3 (Cotton2) grouped with TTG1, and GhTTG2 (Cotton4) and GhTTG4 (close relative of Cotton5) with ATAN11A and ATAN11B. The *GhTTG3* sequence overlapped completely with a cDNA termed ghttg1 (AF336281), also isolated from ovules of *G. hirsutum* cv. Acala Maxxa on the day of flowering (Matz and Burr, unpublished). Since only three (non-synonymous) single nucleotide substitutions were observed in the 1341 bp of overlap, *GhTTG3* is almost certainly the gene which is transcribed to give the ghttg1 mRNA.

Three of the four nucleotide sequences were also virtually identical to cotton ESTs recently sequenced as part of the International Cotton Genome Initiative. The GhTTG1 cDNA appears to represent a full-length version of an EST sequence (BF271126) from a *Gossypium arboreum* 7–10 DPA fibre cDNA library (Wing *et al*., unpublished), whilst both *GhTTG2* and *GhTTG3* correspond to ESTs (AI731000 and AI729712, respectively) from a *G. hirsutum* cv. Acala Maxxa 6 DPA fibre cDNA library (Blewitt *et al*., unpublished). The fourth gene, *GhTTG4*, is a novel isolate.

The genomic organisation of cotton TTG1-like genes

Commercial *G. hirsutum* and *G. barbadense* are 1–2 million year-old tetraploids, with genetic contribution from two distinct diploid subgenomes, A and D (Wendel, 1989). To determine the subgenomic origin of each cotton *TTG1*-like gene, a PCR-based approach was employed, using genomic DNA from three AD allotetraploid cotton varieties, namely *G. hirsutum* cultivars Siokra 1-2 and Siokra 1-4 and *G. barbadense* cv Pima S-6, and two extant species which best represent the putative diploid ancestors, *G. raimondii* Ulbr. (D_5 genome) and *G. herbaceum* L. (A_1 genome). Gene-specific primers, positioned outside of the WD-repeat regions, were designed from sequence alignments of the four genes (see 'Materials and methods').

Using appropriate clones as templates, PCR was used to demonstrate that each primer pair resulted in amplification of a single band of predicted size only from template DNA of the expected clone (Figure 2a). In control experiments, the primers did not amplify non-specific PCR products. In all four cases, the primers yielded DNA fragments from each of the three AD tetraploid species and from the D genome diploid *G. raimondii* (Figure 2a). The size of each amplicon verified results from genomic DNA sequence, indicating that all four genes are intronless, as is the coding region of *TTG1* (Walker *et al*., 1999). Despite the similarity of GhTTG1 to an EST from *G. arboreum*, an A-genome diploid, no amplification was observed from the A genome donor diploid used in this experiment, *G. herbaceum*, suggesting that all four genes are derived from the D subgenome. DNA sequencing of the PCR products from one of the tetraploids, *G. hirsutum* cv. Siokra 1–4, and from *G. raimondii,* verified that in each case a single gene of the expected sequence had been amplified (result not shown).

In order to examine the organisation of *TTG1*-like genes in the cotton genome, the four isolates were used in turn to probe Southern blots of cotton genomic DNA restricted with various endonucleases. DNA from the diploids *G. herbaceum* and *G. raimondii* was included in the experiment to investigate further the genomic origins of WD-repeat genes in cotton. The genes within each pair cross-hybridised strongly to each other,

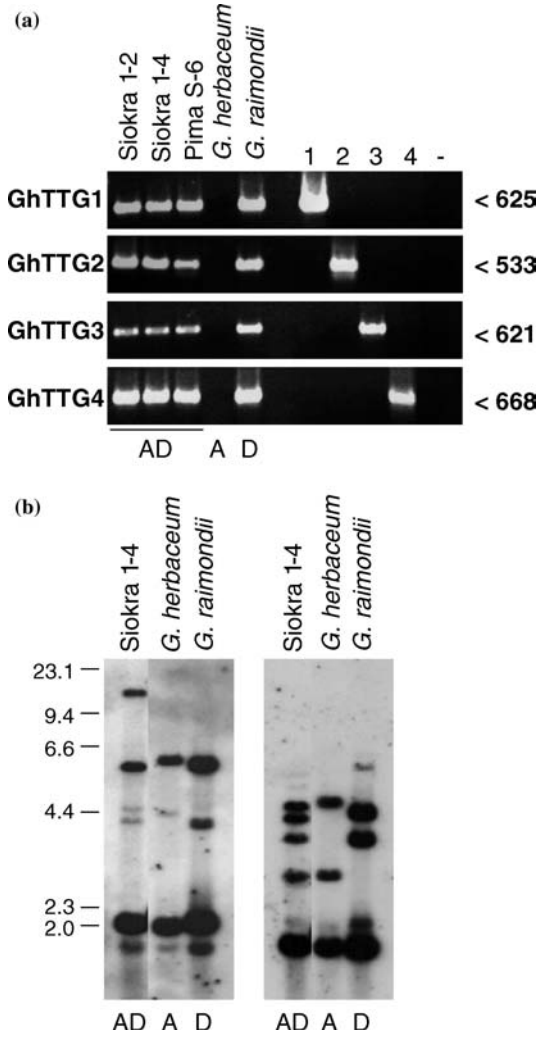

Figure 2. Genomic organisation of WD-repeat genes in cotton. (a) Genomic origins of *TTG1*-like genes in cotton. Each lane contains an equivalent loading of PCR products obtained from genomic DNA templates or clone DNAs as indicated, with 1–4 representing clones of genes *GhTTG1–GhTTG4*, respectively. Numbers on the right indicate the size of the PCR product in bp, and the genomic constitution of each *G. hirsutum* cultivar or diploid species is denoted below the gel. (b) Southern analysis of *TTG1*-like genes in tetraploid and diploid cotton. Each lane contains 10 μg of genomic DNA restricted with *Eco*RI and blots are probed with the GhTTG1 cDNA on the left, to represent the results with both *GhTTG1* and *GhTTG3*, and GhTTG2 cDNA on the right, to represent the result with both *GhTTG2* and *GhTTG4*. Numbers on the left indicate molecular weight markers with sizes in kb, with the genomic constitution of each species denoted below each figure.

producing identical Southern hybridisation patterns, and some cross-hybridisation was also observed between the different sets of probes (Figure 2b). The hybridisation patterns to cotton genomic DNA restricted with a range of restriction enzymes are indicative of only a small number of closely related genes in tetraploid cotton, rather than two large gene families. The two sets of probes also hybridised strongly to sequences present in *G. herbaceum* and *G. raimondii* (Figure 2b) indicating that, whilst the four particular genes described here are derived from the D diploid genome, the A genome also contains closely related WD-repeat genes.

Spatial and temporal regulation of WD-repeat gene expression in cotton

Since Northern blot analysis produced only weak or undetectable signals (result not shown), we used the more sensitive technique of reverse transcriptase PCR (RT-PCR) to obtain an assessment of transcription levels of the four WD-repeat genes in various tissues of the cotton plant. Amplification of a conserved region of actin (Shimizu *et al.*, 1997) was used as an internal reference.

Using the gene-specific primers described in 'Materials and methods', RT-PCR was carried out on total RNA templates from fibres at 6, 12, 18 and 24 DPA as well as RNAs from ovule (0 DPA), leaf, flower, stem and root tissue. Transcription of the four *GhTTG* genes was detected throughout fibre development and in all organs tested (Figure 3). In all experiments the size of the amplicon was as expected, and no amplification was observed in non-reverse transcribed controls of DNase-treated RNA samples (result not shown). All the genes were expressed in whole ovules on the day of anthesis as well as throughout the expansion and elongation stages of fibre development (−1 to 21 DPA), and transcripts persisted into secondary cell wall synthesis (15–40 DPA).

Complementation of ttg1 mutant phenotypes in Arabidopsis *by cotton WD-repeat genes*

To test whether the observed sequence and structural similarities between the cotton WD-repeat proteins and *Arabidopsis TTG1* reflect functional similarity, the ability of the cotton genes to complement *ttg1* mutant phenotypes was investigated. Of particular interest was whether the *G. hirsutum* proteins could complement the *ttg1* defect in *Arabidopsis* trichome differentiation. We

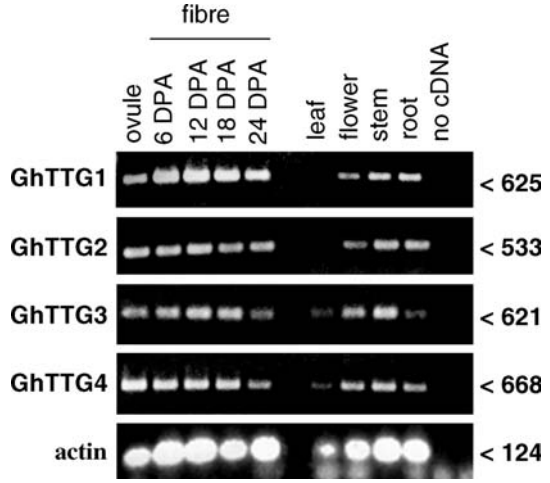

Figure 3. RT-PCR analysis of expression patterns of *TTG1*-like genes in cotton. Each lane contains equivalent loadings of PCR products obtained after initial reverse transcription of cotton tissue RNA, as indicated. Numbers on the right indicate the size of the PCR product in bp.

transformed the cotton WD-repeat genes and *Arabidopsis TTG1*, each under the control of the 35S promoter, into *ttg1-1* mutant *Arabidopsis* plants. The *ttg1-1* allele results in severe defects in anthocyanin pigmentation, seed coat pigmentation, seed coat mucilage, root hair positioning and trichome differentiation (Figure 4, column 2). The molecular basis of the mutant phenotypes is a premature stop codon located 25 amino acids from the C terminus of the TTG1 protein (Walker *et al*., 1999).

To determine if the transgenes complement the *ttg1* defect in trichomes, all primary transformants (kanamycin-resistant) seedlings were monitored for trichome formation. As expected, wild-type *TTG1* was able to complement the *ttg1-1* mutant phenotype, with all transformants bearing normal trichomes (result not shown). In addition, *GhTTG3* (Figure 4D) and *GhTTG1* (result not shown) from cotton were able to rescue the glabrous phenotype, whereas *GhTTG2* was unable to do so, with transformants showing no leaf trichomes (Figure 4C). After several independent attempts and screening of 16 000 seeds, no transformants were obtained for the *35S::GhTTG4* construct. No trichomes were observed in non-transformed and empty-vector controls (result not shown).

Progeny of selfed *35S::GhTTG2* and *35S::GhTTG3 ttg1-1* transformants were further examined for complementation of the full mutant phenotype. The *35S*:*GhTTG1* transformants were only analysed by scoring of trichomes. In the case of *GhTTG3*, seeds from all kanamycin-resistant transformants exhibited wild-type characteristics of brown colour (Figure 4H), indicating complementation of the *ttg1* proanthocyanidin defect in the maternally-derived seed coat, and normal seed coat morphology, with wild-type columellae formation apparent after scanning electron micrograph (SEM) analysis (Figure 4L). In contrast, a collapsed-columellae phenotype was observed in the *GhTTG2* progeny (Figure 4K), resembling that seen in the *ttg1-1* mutant (Figure 4J) and seed pigmentation was not restored by the *GhTTG2* transgene, with yellow seeds being present (Figure 4G). Related to the production of columellae in the seed coat cells, ruthenium red staining of imbibed seeds demonstrated that *GhTTG3* seeds produced wild-type releasable mucilage, whereas *GhTTG2* seeds produced no mucilage, characteristic of the *ttg1-1* mutant (results not shown). In addition, wild-type purple anthocyanins (Figure 4M) were observed in cotyledons and the hypocotyl of *GhTTG3* transformants (Figure 4P), but not *GhTTG2* transformants (Figure 4O).

In wild-type *Arabidopsis* roots, only those cell files that are positioned above the junction of two underlying cortical cells produce root hairs. All root cell files in *ttg1-1* mutants adopt the hair cell fate (Galway *et al*., 1994), resulting in the production of root hairs in adjacent files (Figure 4R), a pattern that is exhibited in *GhTTG2* transformant roots (Figure 4S). *GhTTG3* root hair positioning is wild-type, with each hair cell file being separated by at least one non-hair cell file (Figure 4Q, 4T). SEM analysis confirmed complementation of the root-hair mutant phenotype in *GhTTG3* transformants (result not shown).

Complementation of ttg1 mutations in Matthiola incana *by cotton WD-repeat genes*

We also used a novel transient expression system to examine whether the cotton genes could complement a *ttg1* mutation in *Matthiola incana*. *Matthiola,* like *Arabidopsis*, is a member of the Brassicaceae, and a homologue of *TTG1* exists in *Matthiola* (*MiTTG1*) that affects both trichome and anthocyanin production (Kappert, 1949). The

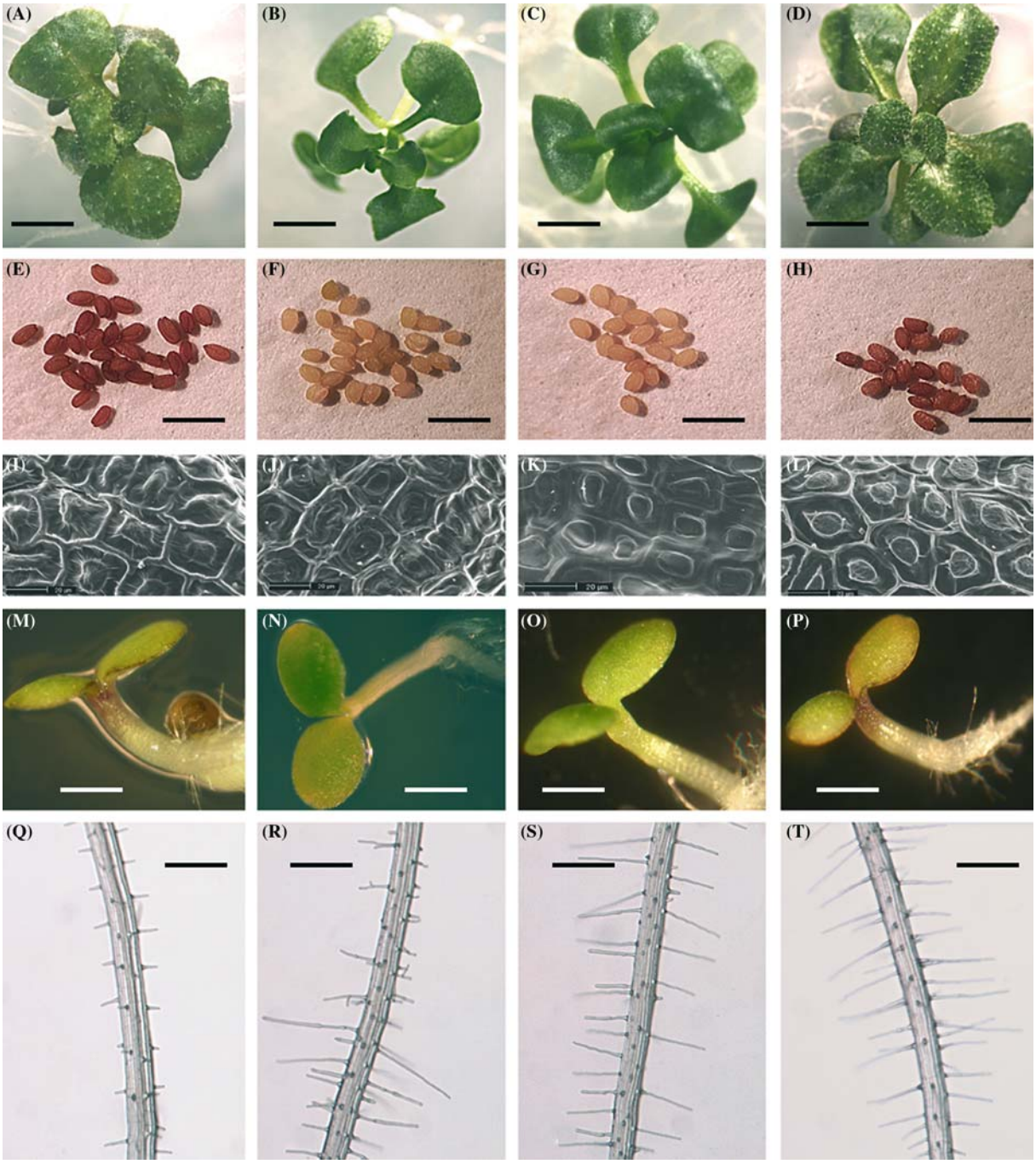

Figure 4. Phenotypes of *Arabidopsis* transformants and controls. Wild-type *Arabidopsis* (column 1), *ttg1-1* mutant (column 2), *ttg1-1* transformed with 35S::*GhTTG2* (column 3) and 35S::*GhTTG3* (column 4). (A–D) Leaf trichomes of seedlings. Bar = 5 mm. (E–H) Seed coat pigmentation. Bar = 2 mm. (I–L) SEM of seed coat structure. Bar = 20 μm. (M–P) Anthocyanin production in seedlings. Bar = 1 mm. (Q–T) Root hair phenotypes. Bar = 0.5 mm.

anthocyanin biosynthetic pathway in *ttg1* mutant *M. incana* is blocked at the dihydroflavanol-4-reductase (DFR) step (Walker, unpublished), which results in a white-petal phenotype as compared to the purple petals of wild-type plants. A second *Matthiola* mutant which has white petals is

deficient for chalcone synthase (CHS) (Epping *et al*., 1990), an enzyme downstream of DFR in the anthocyanin pathway, and this mutant was used as a negative control in the experiment.

Particle bombardment was used to transiently express each of the four cotton genes and *Arabidopsis TTG1*, under the control of the 35S promoter, in petal tissues of the *M. incana ttg1* mutant. Bombardment with *35S*::*TTG1* resulted in purple spots of anthocyanin production, indicating that *Arabidopsis TTG1* complements the mutation (Figure 5A). This result in not surprising, given that 97% amino acid identity exists between *TTG1* and *MiTTG1* (Walker, unpublished). Purple spots of anthocyanin were observed by the naked eye approximately 40 h after bombardment, and the anthocyanin spots spread and increased in intensity until approximately 60 h after transformation.

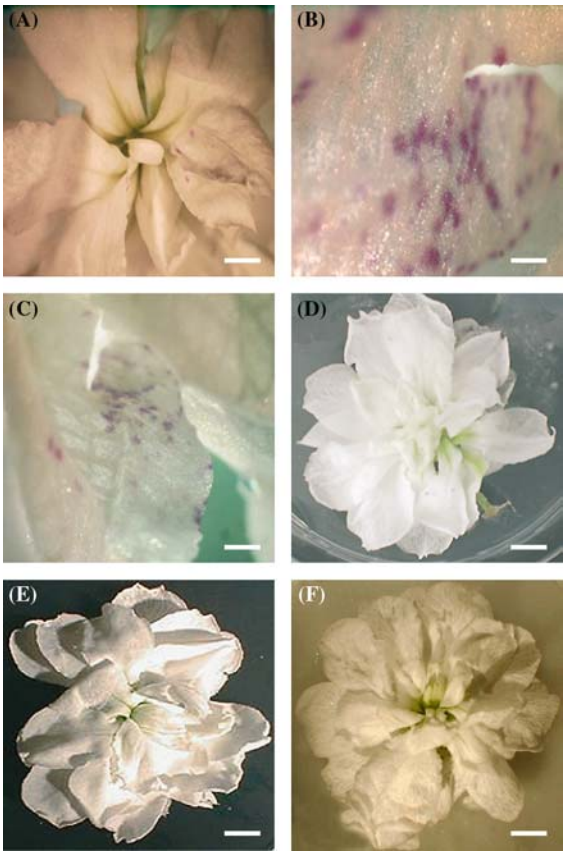

Figure 5. Particle bombardment of *Matthiola* flowers with *TTG1* gene constructs. (A) *M. incana ttg1* mutant petals bombarded with *Arabidopsis TTG1* positive control, showing purple spots of anthocyanin. (B) and (C) Patches of anthocyanin on *M. incana ttg1* petals transiently expressing cotton *TTG1*-like genes *GhTTG1* (B) and *GhTTG3* (C). (D) and (E) *M. incana ttg1* mutant flowers bombarded with cotton *TTG1*-like genes *GhTTG2* (D) and *GhTTG4* (E) , showing no purple spots of anthocyanin. (F) *M. incana chs* mutant flower bombarded with *Arabidopsis TTG1*, showing no purple spots of anthocyanin. Bars = 3 mm in (A), 0.2 mm in (B), 1 mm in (C) and 1 cm in (D), (E) and (F).

Similar purple spots of anthocyanin production were visible on petals transiently expressing *35S*::*GhTTG1* (Figure 5B) and *35S*::*GhTTG3* (Figure 5C). None of the petals transformed with *35S*::*GhTTG2* or *35S*::*GhTTG4* showed any complementation, even after a series of careful searches with a dissection microscope at 40 and 60 h after bombardment (Figure 5D, 5E). In the negative control, transformation of the *M. incana* chalcone synthase (*chs*) mutant with *Arabidopsis TTG1* or any of the cotton constructs did not produce purple spots of anthocyanin (Figure 5F). The failure of *GhTTG1* and *GhTTG3* to complement a mutation in *chs* shows that the result is specific for *MiTTG1*.

Discussion

Apart from its agricultural importance, the study of cotton fibres has resulted in major advances in our understanding of cell growth and cellulose biosynthesis. A large number of genes with fibre-specific expression have been isolated, the majority of which probably function in elongation of the cotton fibre or in secondary cell wall deposition. However, only rarely has a specific function been assigned, and none of the genes appears to be involved in the initiation of fibre development or encodes a transcription factor. The molecular mechanisms which program the differentiation of these highly elongated cells therefore remain unknown. A number of laboratories including our own have suggested that the developmental similarities which exist between cotton seed fibres and *Arabidopsis* trichomes imply conserved regulatory mechanisms between the two systems. Compelling evidence for such a link is provided by our demonstration, in this study, that two cotton genes which are expressed throughout fibre initiation and development, can substitute for *TTG1* in *Arabidopsis* trichome formation. This result shows that *GhTTG1* and *GhTTG3* encode functional

proteins which are clearly active in the context of *Arabidopsis* protein factors, suggesting that similar molecular machinery operates in *Arabidopsis* trichomes and cotton fibres.

Given the ability of *GhTTG1* and *GhTTG3* to generate trichomes in a glabrous *Arabidopsis* mutant, it is tempting to speculate that *GhTTG1* and *GhTTG3* play a pivotal role in cotton fibre differentiation. However, it should be noted that heterologous systems may not reliably reveal gene function. In the particular case of TTG-related processes, there are many examples in which the regulatory genes have been shown to be interchangeable between different species. For example, although *TTG1* homologues from petunia and maize (*an11* and *pac1* respectively) complement multiple *ttg1* mutant phenotypes, neither *an11* nor *pac1* mutants have a trichome phenotype (Payne *et al*., 2000; Carey *et al*., 2004). Yeast and human *an11* homologues partially complement the petunia mutant in transient assays, even though the source species do not produce anthocyanins (de Vetten *et al*., 1997). These results indicate a general functional similarity between members of this group of WD-repeat proteins, which does not always reflect the normal function of the endogenous protein.

The four proteins described here each appear to contain four WD repeats, the minimum number required to form a functional β-propeller fold (Yu *et al*., 2000). WD-repeat proteins occur in large gene families, with 136 identified in humans, 98 in *Drosophila melanogaster*, 88 in *C. elegans*, 58 in *S. cerevisiae* and 59 in *Arabidopsis* (Yu *et al*., 2000; Li and Roberts, 2001; http://bmerc-www.bu.edu/wdrepeat). None of the WD-repeat proteins described to date is thought to be an enzyme and all have presumed regulatory roles in a wide range of cellular processes. Their common function appears to be co-ordination of the formation of multiprotein complexes, mediated by the repeating units, which serve as a scaffold for protein-protein interactions. WD-repeat proteins are known to interact with bHLH transcription factors to control a number of diverse plant developmental processes including trichome development, anthocyanin production and root hair differentiation (Galway *et al*., 1994; Payne *et al*., 2000; Sompornpailin *et al*., 2002), but whether cotton WD-repeat proteins interact with similar molecules is currently unknown.

Consistent with the pleiotropic mutant phenotype, Northern blot experiments showed that *TTG1* is expressed in most major plant organs (Walker *et al*., 1999). Transcripts of *TTG1* homologues from other species, namely *an11*, *PFWD* and *pac1*, are also present in both pigmented and unpigmented tissues (de Vetten *et al*., 1997; Sompornpailin *et al*., 2002; Carey *et al*., 2004). The same appears to be true for cotton *TTG1*-like genes, with *GhTTG1-4* exhibiting similar expression profiles to each other and being expressed in all cotton tissues tested, including ovules and elongating fibre cells. The 0 DPA cotton ovule is a complex structure, consisting of the nucellus, inner integument and the outer integument, from which the fibres are derived. Specific expression in fibre initials is an important prerequisite of a gene associated with fibre differentiation, and *in situ* hybridisation experiments are required to determine whether the *GhTTG1–4* transcripts are present in specific cell types or generally throughout the cotton ovule.

Expression of the individual cotton genes is lower than that of *Arabidopsis TTG1*, which is readily detectable by Northern blot (Walker *et al*., 1999), and may be an indication that several cotton genes, with overlapping functions, perform the same role as the single *TTG1* gene in *Arabidopsis* and *Matthiola*. The increased complexity of the allotetraploid cotton genome makes it probable that the family of WD-repeat-containing proteins will contain more members, and functional redundancy is likely. The results of our transcription analysis and the presence of several cotton ESTs in the database with similarity to our isolates suggests that a large number of WD-repeat genes are expressed in cotton tissues.

Southern analysis under high-stringency conditions indicated the presence of three to four WD-repeat family members related to each of the *GhTTG1*/*GhTTG3* and *GhTTG2*/*GhTTG4* groups in tetraploid cotton. Sequence database searches and Southern analysis detected the presence of related genes in both A and D ancestral diploid genomes, although all four *TTG1*-like genes isolated in this study are evolutionarily derived from the D genome of *G. raimondii*, and therefore do not, as suspected from the sequence data, form two homoeologous pairs. It has been observed previously that the majority of loci affecting fibre yield and quality originate from the D genome

rather than the A genome (Jiang *et al*., 1988), and it would be of interest to investigate the expression patterns and functions of WD-repeat genes in *G. raimondii*.

The complementation by *GhTTG3* of multiple *ttg1* mutant phenotypes in *Arabidopsis*, including trichome development, anthocyanin production, seed coat structure and pigmentation and root hair growth, together with its ubiquitous transcription, suggests that the encoded protein has multiple functions in the cotton plant. The comparable expression pattern and high structural similarity between *GhTTG1* and *GhTTG3* implies that the former will have a similar array of functions. In transient assays, both were able to complement the white-petal phenotype of *Matthiola incana ttg1* mutants, suggesting that the cotton *TTG1*-like genes can function like *TTG1* in the production of anthocyanin pigments in both species. Interestingly, the complementation was not confined to single cells, but groups of cells appeared to produce the purple pigment, allowing visualisation with the naked eye. This result contrasts those of Ludwig *et al*. (1990) and Ramsay *et al*. (2003), in which single cells in maize aleurone and *Matthiola*, respectively, turned purple upon transient expression of bHLH genes. The presence of multi-cellular spots in the *Matthiola ttg1* mutants in this experiment therefore suggests that cell–cell transfer of a signalling molecule has occurred. The TTG1 protein itself may move between cells, as a component of the putative diffusible trichome regulatory complexes (Payne *et al*., 2000; Szymanski *et al*., 2000) or it may regulate a biosynthetic intermediate that is transported between cells.

GhTTG2 and *GhTTG4* did not complement the *ttg1* mutation in *M. incana*, and *GhTTG2* was also unable to rescue the trichome or other phenotypes of *Arabidopsis ttg1*. Whether *GhTTG4* can substitute for *Arabidopsis TTG1* is unknown but unlikely, given the 97% amino acid similarity between the GhTTG2 and GhTTG4 protein sequences. The GhTTG2 and GhTTG4 conceptual translation products show high sequence similarity to the *Arabidopsis* protein ATAN11-A, with 92% (GhTTG2) and 90% (GhTTG4) amino acid identity. In phylogenetic analysis, GhTTG4 groups strongly in the clade containing ATAN11, rather than with AN11 and TTG1 (Carey *et al*., 2004). Based on its high amino acid identity with AN11, ATAN11-A is thought to have a regulatory role in floral pigmentation (de Vetten *et al*., 1997), although its exact biological role is unknown. An investigation of ATAN11 mutant phenotypes in *Arabidopsis* may give an insight into possible roles of GhTTG2 and GhTTG4 in cotton.

Due to its role in trichome initiation in *Arabidopsis*, *TTG1* homologues have been sought in cotton, under the expectation that they could regulate fibre cell initiation. The results obtained in this investigation provide evidence for the replacement of *TTG1* function in trichome initiation and in the anthocyanin pathways of heterologous plant species by the cotton WD-repeat genes *GhTTG1* and *GhTTG3*. Their structural similarity to the *Arabidopsis* trichome regulator *TTG1*, their comparable expression patterns, and the cross-species complementation experiments suggest that *GhTTG1* and *GhTTG3* play a pivotal role in cotton fibre initiation. These results provide the first assignment of function to a regulatory gene from cotton and demonstrate that *Arabidopsis* trichomes and cotton fibres are more than simply analogous structures, with shared transcription factors between the two systems.

Late-initiating cotton ovule epidermal cells form 'fuzz' rather than true fibres, suggesting control of differentiation by a threshold effect within a gene regulatory cascade. Therefore it should be possible to achieve stricter control of fibre development by gene manipulation if the various components of the pathway can be elucidated. The identification of transcription factors from cotton provides targets for misexpression studies in transgenic plants, experiments which will clarify their apparent role in fibre initiation. Further characterisation of the genes described here may lead to significant advances in the understanding of cotton fibre initiation, thereby allowing the development of strategies to increase the percentage of epidermal cells that initiate into fibre, and contributing to long-term crop improvement programs.

Acknowledgements

JAH, JNT and SJO are grateful for financial support from the Cotton Research and Development Corporation, Narrabri, Australia. We thank

Justine Chambers for technical support and help with the plant transformations.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403–410.

Basra, A.S. and Malik, C.P. 1984. Development of the cotton fibre. Int. Rev. Cytol. 89: 65–113.

Beasley, C.A. and Ting, I.P. 1974. Effects of plant growth substances on *in vitro* fibre development from unfertilised cotton ovules. Am. J. Bot. 61: 188–194.

Burn, J.E., Hurley, U.A., Birch, R.J., Arioli, T., Cork, A. and Williamson, R.E. 2002. The cellulose-deficient *Arabidopsis* mutant *rsw3* is defective in a gene encoding a putative glucosidase II, an enzyme processing N-glycans during ER quality control. Plant J. 32: 949–960.

Carey, C.C., Strahle, J.T., Selinger, D.A. and Chandler, V.L. 2004. Mutations in the *pale aleurone color1* regulatory gene of the *Zea mays* anthocyanin pathway have distinct phenotypes relative to the functionally similar *TRANSPARENT TESTA GLABRA1* gene in *Arabidopsis thaliana*. Plant Cell 16: 450–464.

Casas-Mollano, A. and Destefano-Beltrán, N. 2000. Characterisation of a cDNA encoding a TTG1-like protein (Accession No AF220203) from *Malus domestica* fruits. Plant Physiol. 122: 1458.

Clough, S.J. and Bent, A.F. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. 16: 735–743.

de Vetten, N., Quattrocchio, F., Mol, J. and Koes, R. 1997. The *an11* locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. Genes Dev. 11: 1422–1434.

Epping, B., Kittel, M., Ruhnau, B. and Hemleben, V. 1990. Isolation and sequence analysis of a chalcone synthase cDNA of *Matthiola incana* R. Br. (Brassicaceae). Plant Mol. Biol. 14: 1061–1063.

Franks, T., He, D.G. and Thomas, M. 1998. Regeneration of transgenic *Vitis vinifera* L. Sultana plants: genotypic and phenotypic analysis. Mol. Breed. 4: 321–333.

Galway, M.E., Masucci, J.D., Lloyd, A.M., Walbot, V., Davis, R.W. and Schiefelbein, J.W. 1994. The *TTG* gene is required to specify epidermal cell fate and cell patterning in the *Arabidopsis* root. Dev. Biol. 166: 740–754.

Gleave, A.P. 1992. A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. Plant Mol. Biol. 20: 1203–1207.

Graves, D.A. and Stewart, J.M. 1988. Chronology of the differentiation of cotton (*Gossypium hirsutum* L.) fibre cells. Planta 175: 254–258.

Harmer, S.E., Orford, S.J. and Timmis, J.N. 2002. Characterisation of six α-expansin genes in *Gossypium hirsutum* (upland cotton). Mol. Gen. Genom. 268: 1–9.

Hodgson, C.P. and Fisk, R.Z. 1987. Hybridisation probe size control: optimised 'oligolabelling'. Nucl. Acids Res. 15: 6295.

Hsu, C.Y., Creech, R.G., Jenkins, J.N. and Ma, D.P. 1999. Analysis of promoter activity of cotton lipid transfer protein gene LTP6 in transgenic tobacco plants. Plant Sci. 143: 63–70.

Jiang, C., Wright, R., El-Zik, K. and Paterson, A.H. 1988. Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). Proc. Natl. Acad. Sci. USA 95: 4419–4424.

Kappert, H. 1949. Die genetik das incana-charakters und der anthocyanbildung bei der levkoje. Zuchter 19: 289–297.

Koornneef, M. 1981. The complex syndrome of *ttg* mutants. *Arabidopsis* Inform. Serv. 18: 45–51.

Larkin, J.C., Oppenheimer, D.G., Lloyd, A.M., Paparozzi, E.T. and Marks, M.D. 1994. Roles of the GLABROUS1 and TRANSPARENT TESTA GLABRA genes in *Arabidopsis* trichome development. Plant Cell 6: 1065–1076.

Larkin, J.C., Young, N., Prigge, M. and Marks, M.D. 1996. The control of trichome number and spacing in *Arabidopsis*. Development 122: 997–1005.

Lazo, G.R., Stein, P.A. and Ludwig, R.A. 1991. A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. Biotech. 9: 963–967.

Li, D. and Roberts, R. 2001. WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. Cell Mol. Life Sci. 58: 2085–2097.

Liu, H.-C., Creech, R.G., Jenkins, J.N. and Ma, D.-P. 2000. Cloning and promoter analysis of the cotton lipid transfer protein gene *Ltp3*. Biochim. Biophys. Acta 1487: 106–111.

Neer, E.J., Schmidt, C.J., Nambudripad, R. and Smith, T.F. 1994. The ancient regulatory-protein family of WD-repeat proteins. Nature 371: 297–300.

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

Orford, S.J. and Timmis, J.N. 1997. Abundant mRNAs specific to the developing cotton fibre. Theor. Appl. Genet. 94: 909–918.

Orford, S.J. and Timmis, J.N. 2000. Expression of a lipid transfer protein gene family during cotton fibre development. Biochim. Biophys. Acta 1483: 275–284.

Payne, T., Clement, J., Arnold, D. and Lloyd, A. 1999. Heterologous *myb* genes distinct from *GL1* enhance trichome production when overexpressed in *Nicotiana tabacum*. Development 126: 671–682.

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

Pearson, W.R. and Lipman, D.J. 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85: 2444–2448.

Ramsay, N.A., Walker, A.R., Mooney, M. and Gray, J.C. 2003. Two basic-helix–loop–helix genes (*MYC-146* and *GL3*) from *Arabidopsis* can activate anthocyanin biosynthesis in a white-flowered *Matthiola incana* mutant. Plant Mol. Biol. 52: 679–688.

Sambrook, J., Fritsch, E. and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY, USA.

Schellmann, S., Schnittger, A., Kirik, V., Wada, T., Okada, K., Beermann, A., Thumfarht, J., Jürgens, G. and Hülskamp, M. 2002. *TRYPTYCHON* and *CAPRICE* mediate lateral inhibition during trichome and root hair patterning in *Arabidopsis*. EMBO J. 21: 5036–5046.

Shimizu, Y., Aotsuka, S., Hasegawa, O., Kawada, T., Sakuno, T., Sakai, F. and Hayashi, T. 1997. Changes in levels of mRNAs for cell wall-related enzymes in growing cotton fibre cells. Plant Cell Physiol. 38: 375–378.

Soltis, P.S., Soltis, D.E. and Chase, M.W. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature 402: 402–404.

Sompornpailin, K., Makita, Y., Yamazaki, M. and Saito, K. 2002. A WD-repeat-containing putative regulatory protein in anthocyanin biosynthesis in *Perilla frutescens*. Plant Mol. Biol. 50: 485–495.

Stewart, J.M. 1975. Fibre initiation on the cotton ovule (*Gossypium hirsutum*). Am. J. Bot. 62: 723–730.

Szymanski, D.B., Lloyd, A.M. and Marks, M.D. 2000. Progress in the molecular genetic analysis of trichome initiation and morphogenesis in *Arabidopsis*. Trends Plant Sci. 5: 214–219.

Vain, P., Keen, N., Murillo, J., Rathus, C., Nemes, C. and Finer, J.J. 1993. Development of the Particle Inflow Gun. Plant Cell Tissue Organ Cult. 33: 237–246.

Van't Hof, J. 1999. Increased nuclear DNA content in developing cotton fibre cells. Am. J. Bot. 86: 776–779.

Walker, A.R. and Marks, M.D. 2000. Trichome initiation in *Arabidopsis*. Adv. Bot. Res. 31: 219–236.

Walker, A.R., Davison, P.A., Bolognesi-Winfield, A.C., James, C.M., Srinivasan, N., Blundell, T.L., Esch, J.J., Marks, M.D. and Gray, J.C. 1999. The *TRANSPARENT TESTA GLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. Plant Cell 11: 1337–1349.

Walkerpeach, C.R. and Velten, J. 1994. *Agrobacterium*-mediated gene transfer to plant cells: cointegrate and binary vector systems. In: S.B. Gelvin and R.A. Schilperoort (Eds.), Plant Molecular Biology Manual. Kluwer Academic Publishers, Belgium, pp. 1–19.

Wendel, J.F. 1989. New World tetraploid cottons contain Old World cytoplasm. Proc. Natl. Acad. Sci. USA 86: 4132–4136.

Wilkins, T.A. and Jernstedt, J.A. 1999. Molecular genetics of developing cotton fibres. In: A.S. Basra (Ed.), Cotton Fibres: developmental Biology, Quality Improvement, and Textile Processing. The Haworth Press, Inc., USA, pp. 47–63.

Wilkins, T.A., Rajasekaran, K. and Anderson, D.M. 2000. Cotton biotechnology. Crit. Rev. Plant Sci. 19: 511–550.

Yu, L.H., Gaitatzes, C., Neer, E. and Smith, T.F. 2000. Thirty-plus functional families from a single motif. Prot. Sci. 9: 2470–2476.