Plant Cell Physiol. 47(1): 107–127 (2006)
doi:10.1093/pcp/pci228, available online at www.pcp.oupjournals.org

Expression Profiling Identifies Genes Expressed Early During Lint Fibre Initiation in Cotton

Yingru Wu, Adriane C. Machado, Rosemary G. White, Danny J. Llewellyn * and Elizabeth S. Dennis
CSIRO Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia

Cotton fibres are a subset of single epidermal cells that elongate from the seed coat to produce the long cellulose strands or lint used for spinning into yarn. To identify genes that might regulate lint fibre initiation, expression profiles of 0 days post-anthesis (dpa) whole ovules from six reduced fibre or fibreless mutants were compared with wild-type linted cotton using cDNA microarrays. Numerous clones were differentially expressed, but when only those genes that are normally expressed in the ovule outer integument (where fibres develop) were considered, just 13 different cDNA clones were down-regulated in some or all of the mutants. These included: a Myb transcription factor (GhMyb25) similar to the *Antirrhinum* Myb AmMIXTA, a putative homeodomain protein (related to *Arabidopsis* ATML1), a cyclin D gene, some previously identified fibre-expressed structural and metabolic genes, such as lipid transfer protein, α-expansin and sucrose synthase, as well as some unknown genes. Laser capture microdissection and reverse transcription–PCR were used to show that both the GhMyb25 and the homeodomain gene were predominantly ovule specific and were up-regulated on the day of anthesis in fibre initials relative to adjacent non-fibre ovule epidermal cells. Their spatial and temporal expression pattern therefore coincided with the time and location of fibre initiation. Constitutive overexpression of GhMyb25 in transgenic tobacco resulted in an increase in branched long-stalked leaf trichomes. The involvement of cell cycle genes prompted DNA content measurements that indicated that fibre initials, like leaf trichomes, undergo DNA endoreduplication. Cotton fibre initiation therefore has some parallels with leaf trichome development, although the detailed molecular mechanisms are clearly different.

Keywords: Cotton — Expression profiling — Fibre development — *Gossypium hirsutum* — Microarray — Trichome.

Abbreviations: DAPI, 4′,6-diamidino-2-phenylindole; dpa, days post-anthesis; DP16, cultivar Deltapine 16; EST, expressed sequence tag; GL1, Glabrous 1; GL2, Glabrous 2; GUS, β-glucuronidase; LCM, laser capture microdissection; LTP, lipid transfer protein; SEM, scanning electron microscopy.

Introduction

Lint fibres of cotton (*Gossypium hirsutum* L. and other *Gossypium* species) are extremely long single epidermal cells that develop on the outer surface of ovules, reaching upwards of 5 cm in some species (Stewart 1975). Fibres initiate between 1 day pre- and 1 day-post anthesis (dpa), and the fibre initials begin to elongate rapidly immediately after fertilization, extending out from the surface of the seed coat epidermis. After a period of elongation lasting 15–25 d, secondary cell wall thickening fills the fibre with cellulose and the fibre dies and collapses to form the mature lint fibre that is harvested from the seeds and used in textile manufacture (Basra and Malik 1984, Berlin 1986, Ryser 1999). Between 5 and 10 dpa, a second round of fibre initiates, but these fibres fail to elongate appreciably and form the 'fuzz' fibre that remains very short (a few millimetres) at maturity and is still attached to the seed after the lint is removed in ginning. These two types of fibres, the long lint fibres and the short fuzz fibres, probably share common developmental pathways at least early during differentiation. However, the fuzz fibre appears to be under separate genetic control as a number of genetic loci specifying absence of fuzz fibre, but with normal lint, have been identified (e.g. Turley and Kloth 2002). Lintless mutants, however, only occur in conjunction with lack of fuzz fibre, so are essentially fibreless (Turley and Kloth 2002). Most biochemical and molecular studies have focused on the lint fibres because of their economic significance and ease of isolation.

Cotton fibres are an excellent single cell model system for cell growth and cellulose biosynthesis (reviewed by Basra and Malik 1984). Various fibre-specific and fibre-enriched genes involved in cell elongation and cell wall biogenesis, including cellulose synthesis (reviewed by Wilkins and Jernstedt 1999), have been cloned. Studies of genes involved in fibre growth, such as phosphoenopyruvate carboxylase (Vojdani et al. 1997); two membrane-bound electrogenic proton pumps, the vacuolar H^+-ATPase (Hasenfratz et al. 1995) and the plasma membrane H^+-ATPase (Smart et al. 1998); sucrose synthase (Ruan and Chourey 1998); and sucrose and K^+ transporters (Ruan et al. 2001) suggest that cell turgor is the primary driving force behind fibre cell elongation. Cytoskeleton genes, such as α-tubulins (Whittaker and Triplett 1999) and β-tubulin (X.B. Li et al. 2002), may play a dual role in cell elongation and secondary cell wall synthesis. Cell wall-modifying genes, such as

* Corresponding author: E-mail: Danny.Llewellyn@csiro.au; Fax, +61-2-62465000.

endo-1,4-β-glucanase and expansin (Shimisu et al. 1997), act in concert with cell turgor through the relaxation of the primary cell wall structure allowing cell expansion. Other fibre-specific or fibre-enriched genes have been characterized including: lipid transfer proteins (LTPs) (Ma et al. 1995, Ma et al. 1997) presumed to function in cutin synthesis; E6 (John 1996) and Fb-B6 (John 1995), probable cell wall structural proteins; and cellulose synthesis genes cesA1 and cesA2 (Pear et al. 1996). Recent studies on the comparison of gene expression profiles of 5–10 dpa ovules from a *fuzzless-lintless* mutant (*fl*) and its parental wild type (Xu-142) using cDNA macroarrays identified new genes including an auxin-binding protein, a mitogen-activated protein kinase and an RD22-like protein that are all preferentially expressed in elongating fibres (C.H. Li et al. 2002). Ji et al. (2003) identified a large number of elongation phase (10 dpa) cDNA clones (mainly enzymes and cell wall proteins) by subtractive hybridization and cDNA macroarrays comparing wild-type and fibreless mutants, but the physiological or biochemical roles in fibre growth of any of these have yet to be verified. Kim and Triplett (2004) used differential display with a similar mutant/wild type comparison at 5 dpa to identify a germin-like gene with reduced expression in the mutant, but this was shown to be most abundantly expressed in the elongating fibres. Arpat et al. (2004) were the first to take a genomic approach to fibre development examining the growing number of cotton expressed sequence tags (ESTs) from elongating cotton fibres. They confirmed the genetic complexity of this stage of fibre growth and, using an oligonucleotide array of >12,000 non-redundant sequences, they identified 2,500 elongation stage-specific genes that were down-regulated in 24 dpa fibres undergoing secondary cell wall thickening and 81 secondary cell wall stage-specific genes that were up-regulated relative to 10 dpa fibres. This study clearly delineated the two stages of transcriptional activity defining the different physiological and biochemical processes known to be occurring in these cells at these times and will provide interesting targets for further characterization and functional verification. Not surprisingly, many of the genes identified in the microarrays are involved in energy metabolism, cell turgor generation, and primary and secondary cell wall biogenesis, but many are possibly also of novel function.

Few genes, however, have been identified that are associated with the early events in fibre cell initiation and development. As all hair-like or bristle-like outgrowths from the epidermis of plants are called trichomes, cotton fibres are botanically seed trichomes, and numerous authors have drawn parallels between these two cell types (e.g. Suo et al. 2003, Wang et al. 2004, Humphries et al. 2005). At a molecular level, the understanding of plant trichome initiation and development is greatest for *Arabidopsis*, where at least 20 genes are required (Hülskamp et al. 1994). Normal trichome development requires a trichome-promoting protein complex comprised of GLABROUS1 (GL1, a Myb protein), TRANSPARENT TESTA GLABRA1 (TTG1, a WD40 protein) and GLABRA3 (GL3, a basic helix–loop–helix protein). GLABRA2 (GL2, a homeodomain protein) regulates trichome morphology and spacing, while TRIPTYCHON (TRY, another Myb-like protein) mediates lateral inhibition of trichome development in adjacent cells (Rerie et al. 1994, Walker et al. 1999, Szymanski et al. 2000, Ohashi et al. 2002, Schellmann et al. 2002). Potential homologues of these genes have been reported to be expressed in cotton fibres (Loguerico et al. 1999, Suo et al. 2003, Wang et al. 2004, Humphries et al. 2005) and some demonstrated to complement *Arabidopsis glabrous* mutants, but their roles in leaf trichome and/or fibre development remain to be confirmed in transgenic cotton. Circumstantial evidence does suggest that different trichome types in different species may be specified by distinct developmental programmes (Noda et al. 1994, Glover et al. 1998, Payne et al. 1999). Suo et al. (2003) identified a cotton Myb gene, GhMYB109, that is structurally related to both GL1 and another Myb regulating root hair development, WEREWOLF (WER). GhMYB109 is expressed early, but at much higher levels in the elongating fibres, suggesting a role in elongation rather than initiation.

To identify genes that regulate fibre initiation, we have used mRNA from early stage fertilized ovules of wild type and six reduced fibre or fibreless mutants of cotton (that produce few if any lint fibres and no fuzz fibres) to probe a cotton cDNA microarray containing about 10,000 ovule cDNA clones expressed around the time of fibre cell differentiation. Using this strategy, we have identified 13 candidate genes that are differentially expressed in many of the fibreless mutants including two types of transcription factors (Myb and homeodomain genes), a cell cycle gene and other genes with known or presumed roles in cell growth. The spatial and temporal expression of the Myb gene, in particular, coincides with the location and time of fibre initiation and, when expressed in transgenic tobacco, resulted in an increase in number and branching of leaf trichomes. The putative homeodomain gene is expressed mainly in ovules and fibre initials, and to a lesser extent in leaves. Nuclear DNA content measurements also indicate that the majority of fibre cells undergo at least one round of DNA endoreduplication during initiation. These results suggest some similarity between the types of genes and molecular events involved in cotton fibre and *Arabidopsis* trichome initiation, but not a direct conservation of the two developmental systems.

Results

Reduced fibre mutants show a much reduced but varied number of fibre initials with delayed development

A number of naturally reduced fibre and fibreless mutants, lacking both the long lint fibres and the short fuzz fibres, have been found fortuitously in fields of commercial cotton, and these provide unique genetic material to study the molecular events specific for fibre development. The six mutants examined here were of varying provenance and relatively uncharacterized genetically. The four lines Lintless 1A, Lint-

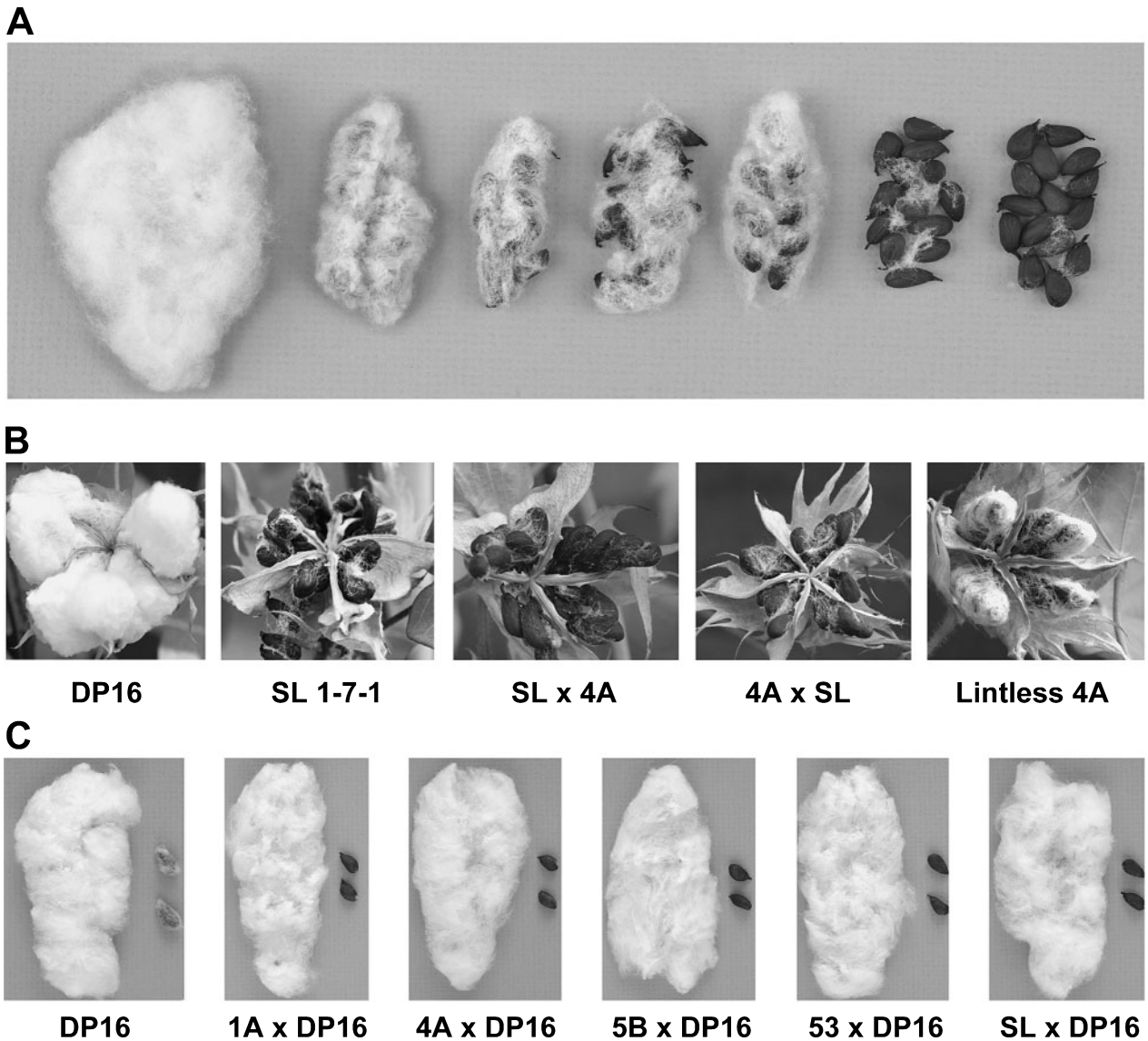


Fig. 1 Phenotypes of linted and lintless cotton mutant lines and crosses. (A) Mature seed phenotypes of cotton lines showing a range of lint levels. From left to right: DP16 (wild type); Lintless 5B; Lintless 53; Lintless 4A; Lintless 1A; SL1-7-1; Xu-142 *fl*. Note: all mutants lack fuzz fibre. (B) Reciprocal crosses of lintless mutants SL1-7-1 and Lintless 4A. Mature seed phenotypes photographed in situ; from left to right: wild-type DP16; SL1-7-1; F2 seed of SL1-7-1 (female) crossed with Lintless 4A; F2 seed of Lintless 4A (female) crossed with SL1-7-1; Lintless 4A. (C) F1 lint and fuzz phenotypes shown on F2 seeds from crosses of five mutants with wild-type DP16 with DP16 phenotype as a control, from left to right: DP16; Lintless 5B (female) crossed with DP16; Lintless 53 (female) crossed with DP16; Lintless 4A (female) crossed with DP16; Lintless 1A (female) crossed with DP16; SL1-7-1 (female) crossed with DP16. Fuzzy seed is grey after removal of the lint, while fuzzless seed is black.

less 4A, Lintless 5B and Lintless 53 were originally separate selections from a fully linted and fuzzy seeded cultivar B1278 (A. Low, unpublished) isolated as presumed spontaneous mutants that showed low but differing levels of lint production and no fuzz fibre (so are not completely lintless as mature seeds, despite their name). SL1-7-1 (Mississippi Obsolete Variety Collection Number 0504) was also a naturally occurring variant (mentioned in Turley and Ferguson 1996) that is fuzzless and produces very little lint at maturity (less than the most lintless of the B1278 selections). Xu-142 *fl* is a profoundly lintless and fuzzless mutant of cultivar Xu-142 and produces only a few lint fibres on some seeds (Zhang and Pan 1991). The mature seed phenotype of the different mutants is shown in Fig. 1A. Except for Xu-142 *fl*, all the mutant lines, but particularly Lintless 5B (which had the most lint of any of the lines), showed a variable, but low level of leakiness, but never produced as much lint as a normally linted cultivar. The growth rate and general vegetative and floral development of the mutants were all similar to the wild type except that SL1-7-1 (SL) exhibited higher levels of red anthocyanin pigments in

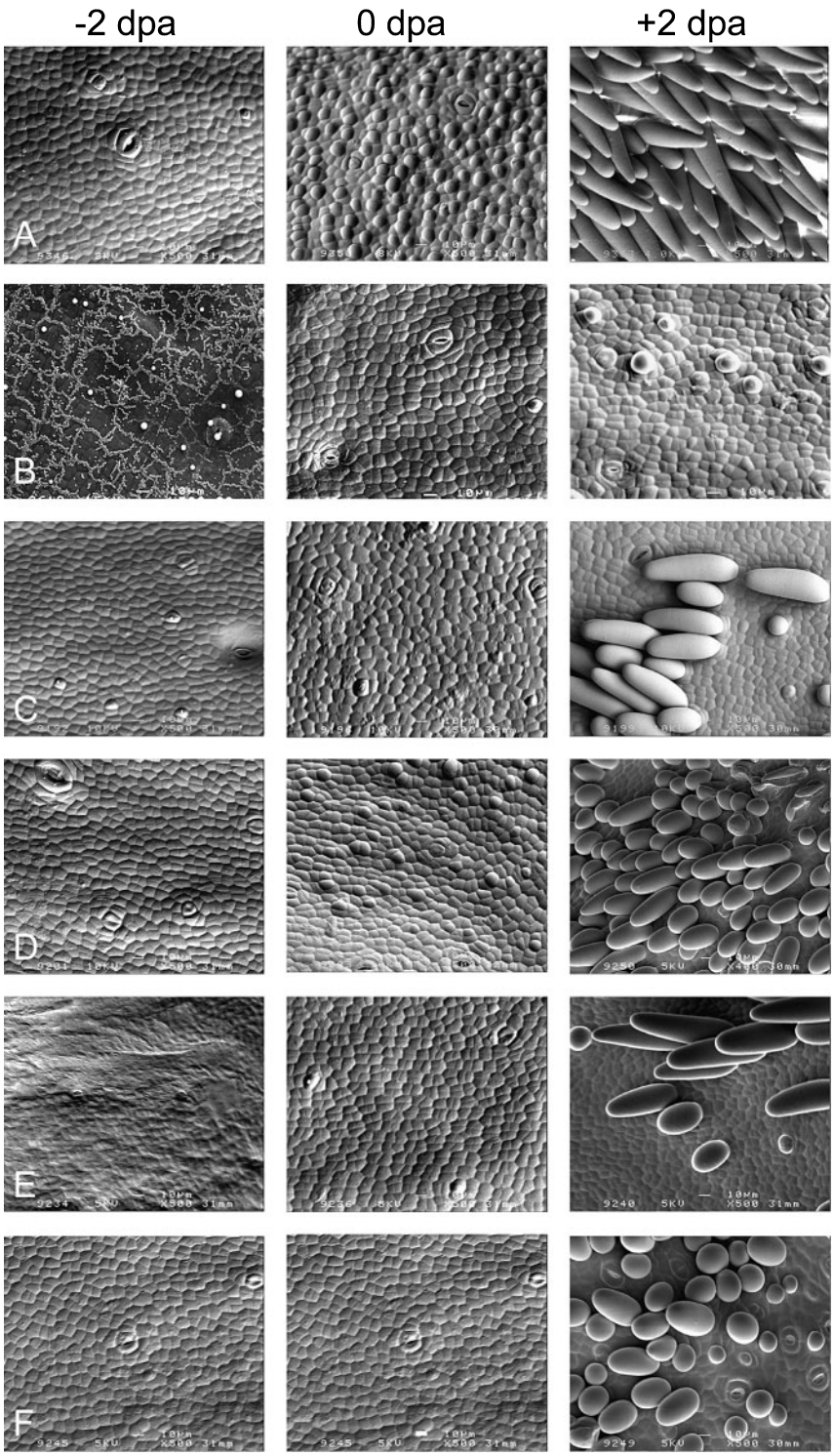


Fig. 2 SEM images of the surface of –2, 0 and +2 dpa cotton ovules from wild type and reduced lint or lintless cotton mutants. (A) Wild-type cultivar DP16; (B) Lintless 4A (note the mucous covering on the surface of –2 dpa ovules that disappears by 0 dpa); (C) Lintless 41; (D) Lintless 53 (note the few fibre initials occurring on 0 dpa ovules); (E) Lintless 5B; (F) SL1-7-1. All of the mutants had some level of fibre initiation by +2 dpa, but mostly towards the chalazal end in patches. The scale bar is 10 μm for all images.

most plant parts and probably carries the red plant mutation known in *G. hirsutum*. All lines produced normal levels of trichomes on their stems and leaves (not shown).

The genetics of lint and fuzz fibre production in cotton is complex and still contentious. Of the lines examined here, only Xu-142 *fl* has been characterized in any detail, and its fibreless (i.e. lintless-fuzzless) phenotype is reportedly conferred by two recessive genes (Zhang and Pan 1991), although Du et al. (2001) had a more complex interpretation, invoking four genes. Turley and Kloth (2002) have put forward a different explanation for the genetic control of lint and fuzz fibre, postulating three sets of dominant and recessive fuzz-determining genes

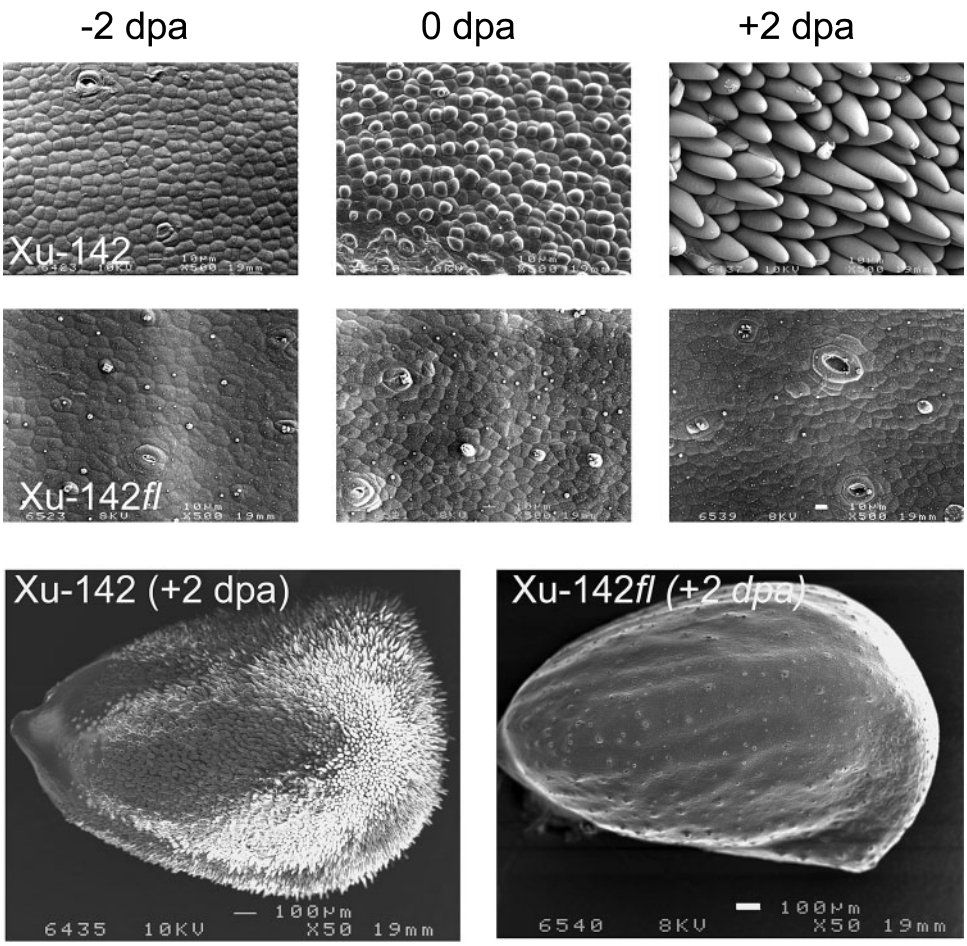


Fig. 3 SEM images of the surface of –2, 0 and +2 dpa cotton ovules from the Chinese cultivar Xu-142 (upper panel) and its completely fibreless derivative Xu-142 *fl* (middle panel). The lower panel shows the corresponding whole ovules at +2 dpa when the mutant still shows no sign of fibre initiation. The scale bar in the top two rows is 10 µm and in the bottom 100 µm.

with epistatic interactions to generate the fibreless trait. SL1-7-1 has at least one dominant gene for fuzzless seed and a second and possibly a third recessive fuzzless gene that together also inhibit lint production (R. Turley, personal communication). We crossed all six of the mutants in a reciprocal manner. F1 seed in each case was identical to the female parent as the seed coat and fibre are maternal tissues. Four F1 plants from each cross were grown and allowed to self. The phenotype of the F1 seed coat on the F2 seed in each cross was fuzzless and predominantly lintless (e.g. Fig. 1B) (although some had a similar amounts of lint to the parental mutants), but none had as much lint as the wild type, so there was no apparent complementation of the defects in any combination of two lintless parents. All crosses to Xu-142 *fl* had as much or more lint than their paternal mutants, while all crosses to SL1-7-1 tended to have as much or less lint than their paternal mutants. Crosses of the mutant lines to wild-type linted and fuzzy cotton [cultivar Deltapine 16 (DP16)] produced F2 seed that was linted, but fuzzless for all the mutants except for Xu-142 *fl* (Fig. 1C), confirming that these Australian and American lines carry a dominant allele for lack of fuzz fibre [probably N1 as suggested for SL1-7-1 by Turley and Kloth (2002)]. Xu-142 *fl*×wild-type F2 seed, on the other hand, was linted and fuzzy (i.e. wild type), confirming the results of Zhang and Pan (1991) that it carried one or more recessive fuzzless seed alleles some of which must be shared with the Australian and US mutants. Because the genetics is complex and still unresolved, it is difficult to assess the relatedness of the loci in the lines without more extensive crosses and progeny testing that is beyond our limited glasshouse resources. The data, while incomplete, suggest that most of the lines (except Xu-142 *fl*) carry a recessive mutation in the same gene that produces lintless seed as well as a dominant mutation conferring fuzzless seed, while Xu-142 *fl* carries a recessive mutation in this same gene for fuzz fibre production and possibly other genes. The different subtle fibre phenotypes of the mutants may be due to other loci that modify the degree of lint production.

Scanning electron microscopy (SEM) was used to examine the fibre development of the six cotton mutants (Lintless 1A, 4A, 5B, 53, SL1-7-1 and Xu-142 *fl*) at 2 d before anthesis (–2 dpa), the day of anthesis (0 dpa) and 2 d after anthesis (+2 dpa) (Fig. 2, 3). There were no obvious differences between the mutants and the wild type (DP16) at –2 dpa except that Lintless 4A ovules were covered with a mucous-like substance (Fig. 2B). The ovule surfaces were flat and epidermal cells were interspersed with stomata. On the day of anthesis,

Table 1 Number of cDNA clones that are up- or down-regulated in the whole 0 dpa ovules of each of the cotton mutants as compared with the wild-type DP16 cultivar

Comparison	No. of genes differentially expressed	No. of genes up-regulated in the mutant	No. of genes down-regulated in the mutant	No. of replicates [a]	Minimal reproducibility [b]
4A/DP16	60	9 (15%)	51 (85%)	8	75%
1A/DP16	67	48 (72%)	19 (28%)	8	75%
SL/DP16	102	88 (86%)	14 (14%)	6	83%
5B/DP16	144	91(63%)	53 (37%)	4	75%
53/DP16	243	199 (82%)	44 (18%)	4	75%

[a] Replicates consist of both biological and technical replications. Each biological replicate (a separate RNA isolation) comprises two dye-swapped technical replicates. The number of biological replicates = total number of replicates/2.

[b] Genes identified as being significantly differentially expressed in at least six out of eight or three out of four replicates (75%), or five out of six replicates (83%).

most cells destined to become fibres on the wild type had already started to balloon out from the epidermis, more so from the chalazal end (Fig. 2A, 3). In comparison, in all the mutants, few if any fibre initials had become visible on the seed epidermis (except for Lintless 53 that had a low level of fibre initiation; Fig. 2D) so at this developmental stage all were effectively lintless (fuzz fibres do not form for another 4–5 d). The mucous-like substance covering Lintless 4A ovules had disappeared by 0 dpa (Fig. 2B). The mutant phenotype was more obvious at 2 dpa when the fibres are rapidly elongating in the wild type, but are relatively short and patchy in the mutants (Fig. 2A–F, 3). All the mutants had a much reduced number of fibre initials on the surface of ovules and those fibres that developed appeared to be less synchronized in their elongation than the fibre initials of the wild type. As previously reported (Yu et al. 2000), the Xu-142 *fl* mutant had virtually no fibre initials (Fig. 3). Regardless of the genetic nature of the lesions in each of the mutants, the relative lack of fibre initials at 0 dpa relative to wild-type ovules provided an ideal comparator for the identification of gene expression differences that might correlate with fibre cell initiation, and this stage was chosen for the main microarray comparisons. As noted above, fuzz fibre is not initiated until later so does not complicate the analysis.

Many genes are differentially expressed in the mutant ovules

Gene expression in 0 dpa whole ovules of each of the mutants except Xu-142 *fl* was compared with the wild-type DP16 at the same stage using our 10,000 ovule cDNA microarray. Each experiment was replicated a minimum of four times as two biological replicates (ovules collected from the same batch of plants on a different date) and two dye swap replicates (Table 1). The lists of differentially expressed cDNA clones for all the mutant/wild type microarray experiments are provided as Supplementary material (supplementary Table S1) and the full data sets are available at http://www.pi.csiro.au/gena/ and The Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) as Series Accession No: GSE3545, GSE3546 with Platform Accession No: GPL3035 describing the array. The total number of statistically significant differentially expressed cDNA clones varied considerably, from an average of 60 clones in the Lintless 4A/DP16 comparison up to an average of 243 clones in the Lintless 53/DP16 comparison (the line with the largest number of fibre initials at 0 dpa) (Table 1); however, there was no consistent correlation between the numbers of differentially expressed genes and the number of fibre initials formed in the different mutants at 0 dpa, as most mutants had few if any fibre initials at day zero. Not surprisingly, the number of statistically significant differentially expressed genes tended to drop with increasing levels of replication of the hybridizations.

As an additional control and to validate the statistical robustness of the tRMA package used for the microarray data analysis, we also carried out three self-hybridizations of DP16 0 dpa ovules (three biological replications, i.e. three batches of 0 dpa ovules collected from the same pool of DP16 plants on different dates). Each batch of RNA was labelled with two different fluorescent tags, and hybridized to a separate array for each replicate. In each case, the data transformation and normalization were the same as detailed in Materials and Methods and generated lists of the differentially expressed genes in each of the hybridizations. The 'compare interesting genes' function of tRMA was then used to select those genes that were repeatedly (three out of three replications) identified in all the biological replications in a manner similar to the experimental data sets. This method gave a list of only three genes that were differentially expressed among the 10,000 spotted clones (discounting the flagged spots, see supplementary Table S2), resulting in a false-positive rate of only 0.03% and none of these genes corresponded to clones identified in the experimental data sets.

Among the genes that are differentially expressed in fibreless mutants, only a small number are normally elevated in expression in the seed coat outer integument

A 0 dpa cotton ovule is a rapidly developing complex organ, composed of at least three separable layers of tissues: the outer integument, the inner integument and the nucellus. As the cotton fibre develops only from the epidermal cells of the

Table 2 Genes [a] that are up- or down-regulated in the 0 dpa ovules of the lintless cotton mutants and up-regulated in the outer integument of wild-type cotton ovules

Name	Clone ID	GenBank ID	SL/DP	1A/DP	53/DP	4A/DP	5B/DP	OI/II [b]	Most homologous gene [c]
GhMyb25	ON035F4	AY464054	–5.8	–10.1	–10.8	–10.8	–10.8	4.6	(AF336283) GHMYB25
GhFaEl	ON035N9	AY464065	–5.6	–6.7	–5.7	–7.0	–5.8	4.7	(NP_195909) Transferase
GhFU1	ON035C9	AY464064	–8.5	–11.0	–14.1	–15.2	–13.8	2.7	No match (probably secreted glycine-rich cell wall protein)
GhEX1	pFS14×	AF043284	–6.8	–7.6	–9.1	NS	–9.1	3.3	(AF512539) α-Expansin
GhFU2	ON003F1	AY464061	NS	NS	–5.9	–5.6	–8.2	2.8	No match
GhHD1	ON033M7	AY464063	NS	NS	NS	–5.5	–4.9	2.8	(T05850) Homeobox protein ATML1
GhTMTP	CHX015K18	AY464053	NS	NS	NS	–8.1	–4.4	4.7	(NP_175557) ATP-dependent transmembrane transporter
GhCycD3;1	OCF007F4	AY464058	NS	NS	NS	–5.3	–3.7	4.9	(AAQ19972) Cyclin D3
CHX007D10 [d]	CHX007D10	DT462003	NS	NS	3.2	NS	NS	4.4	(P92983) Proline oxidase/ (AC084282) putative protein phosphatase 2C
GhSus	CHX002C10	AY464055	NS	NS	NS	NS	–3.5	4.6	(AAD28641) Sucrose synthase
GhLTP	ON033M19	AY464062	NS	NS	NS	NS	–5.9	3.2	(AAM62634) Lipid transfer protein

[a] The values presented in the table are the medians of Log_2-transformed, normalized and rescaled ratios of the two compared samples. The rescaling was performed by dividing through by an estimate of the median absolute deviation (Wilson et al. 2003).
[b] OI, outer integument; II, inner integument and nucellus.
[c] Based on the top BlastX hit.
[d] Most probably a chimeric clone.
NS, not significantly different in this comparison.

outer integument, genes that show a higher expression level in this layer should be more relevant to fibre initiation and development than genes that are predominantly expressed in the inner integument and nucellus. Outer integuments were therefore separated from the inner integuments and nucellus of 0 dpa wild-type ovules from DP16 by micro-dissection and labelled cDNA prepared from the partitioned tissues. The gene expression in the outer integument was then compared with that of the inner integument and nucellus by probing the 10,000 ovule cDNA microarray. The results, averaged over four replicates (two biological replicates each with two dye-swapped technical replicates), revealed a total of 121 cDNA clones that were differentially expressed, with 66 clones up-regulated and 55 clones down-regulated in the outer integument of wild-type ovules relative to inner ovule tissues (supplementary Table S3). The list of 66 outer integument up-regulated clones was then used as a filter in silico on the differentially expressed gene lists identified from the fibreless mutant/wild type comparisons to select for cDNA clones whose expression might correlate with fibre initiation in the outer epidermis of linted cultivars. This filtering resulted in the identification of a surprisingly small number of clones: four, seven, six, 10 and four clones from the 1A, 4A, 5B, 53 and SL mutants, respectively, most of which were significantly down-regulated in the fibreless mutants compared with wild-type ovules and also up-regulated in the ovule outer integument relative to inner ovule tissues in the wild-type cultivar.

In all, from this set of five wild type/mutant comparisons, only 10 unique cDNA clones were identified as potentially being involved in the early stages of fibre development. The relative expression changes in each mutant and predicted sequence identities of these candidate genes are summarized in Table 2. Only three clones [ON35F04 (GhMyb25), ON35N09 (GhFaE1, transferase protein) and ON35C09 (GhFU1, unknown glycine-rich protein)] were consistently down-regulated in all five mutants, while the other clones were down-regulated in 2–4 of the mutants. The other less consistent genes included known fibre development-related genes such as α-expansin, sucrose synthase and LTPs, and some regulatory genes that have not been previously identified as important in early fibre development, including a gene encoding a putative homeodomain protein; a cyclin D3 homologue; a transmembrane transporter; and a gene of unknown function. For ease of referring to the different genes, we have assigned them gene names that refer to their presumed functions, such as GhHD1 to refer to the cotton homeodomain protein-like gene represented by ON33M07 (Table 2), until such time as the complete genes are cloned and fully characterized.

The candidate genes show distinctive expression patterns during the time of fibre initiation

RNA from wild-type DP16 ovules of –4, –2 and +2 dpa was compared with RNA from 0 dpa DP16 ovules using the 10,000 ovule cDNA microarray (each as two biological repli-

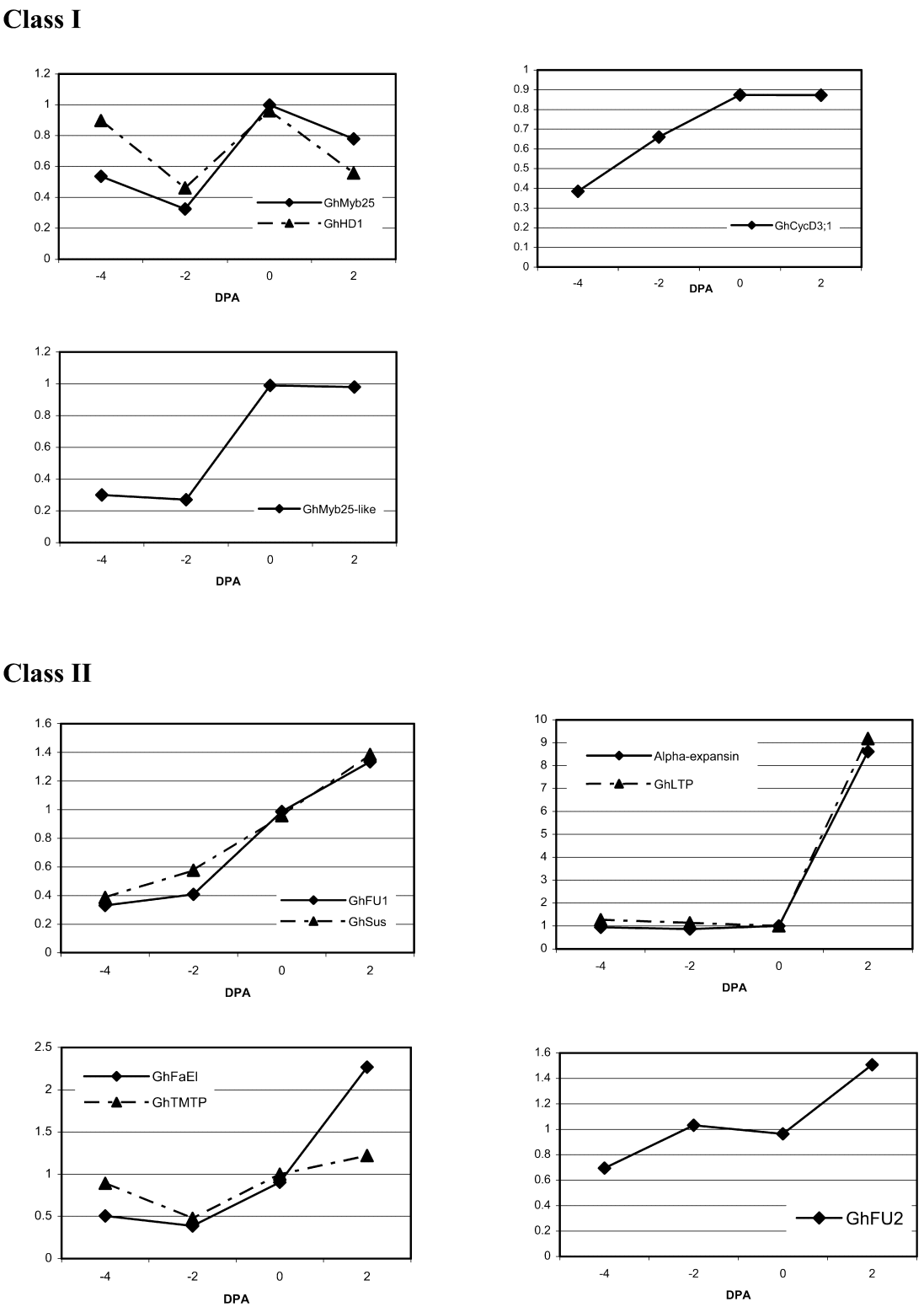


Fig. 4 Expression profiles in cotton ovules of the candidate fibre genes from –4 to 2 dpa. Plots are based on microarray hybridizations of DP16 ovules at –4, –2, 0 and 2 dpa compared with DP16 0 dpa ovules. Each comparison consists of two biological replicates (and two dye-swapped technical replicates) except the 0 dpa self-hybridization which consists of three biological replicates. The data plotted are the back-transformed ratios resulting from tRMA analysis. Genes showing similar expression profiles and functional classifications are plotted together and grouped into class I and II (see Results).

Table 3 Genes differentially expressed in the outer integument of the ovules of the fibreless (*fl*) cotton mutant as compared with its wild-type parent Xu-142

Name	Clone ID	GenBank ID	Fluorescence ratio *fl*/Xu-142 [b]	Most homologous gene [c]
α-Expansin [a]	pFS14×	AF043284	–11.5	(AF512539) α-expansin
GhFU1 [a]	ON35C09	AY464064	–9.8	Unknown
GhMyb25 [a]	ON35F04	AY464054	–9.0	(AF336283) GHMYB25
GhFU2 [a]	ON03F01	AY464061	–7.8	Unknown
GhFaEl [a]	ON35N09	AY464065	–7.4	(NP_195909) Transferase
GhHD1 [a]	ON33M07	AY464063	–6.1	(T05850) Homeobox protein ATML1
GhLTP [a]	ON33M19	AY464062	–4.9	(AAM62634) Lipid transfer protein
GhTMTP [a]	CHX15K18	AY464053	–3.7	(NP_175557) ATP-dependent transmembrane transporter
GhMyb25-like	ON38N08	AY464066	–4.4	(AF336283) GHMYB25
GhRD22	OCF05C10	AY464056	–4.4	(AAL67991) Dehydration-induced protein RD22
GhAsp	OCF08G09	AY464059	–3.4	(BAC66615) L-Aspraginase
GhLTP2	OCF10D08	AY464060	3.4	(CAA65477) Non-specific lipid-transfer protein
GhFU3	OCF06C01	AY464057	5.3	Unknown

[a] Genes in common with the candidate genes from the other five mutants.
[b] The values presented in the table are the medians of Log_2-transformed, normalized and rescaled ratios of the two compared samples. The rescaling was performed by dividing through by an estimate of the median absolute deviation (computed on the final residual mean-difference data) as described by Wilson et al. (2003). Negative numbers indicate lower expression in the mutant and positive numbers higher expression in the mutant compared to wild-type.
[c] Based on the top BlastX hit.

cates and two dye swaps) to profile the temporal changes in expression of genes around the time of fibre initiation. The expression changes for the 10 initial candidate fibre genes relative to their levels in 0 dpa ovules are shown in Fig. 4.

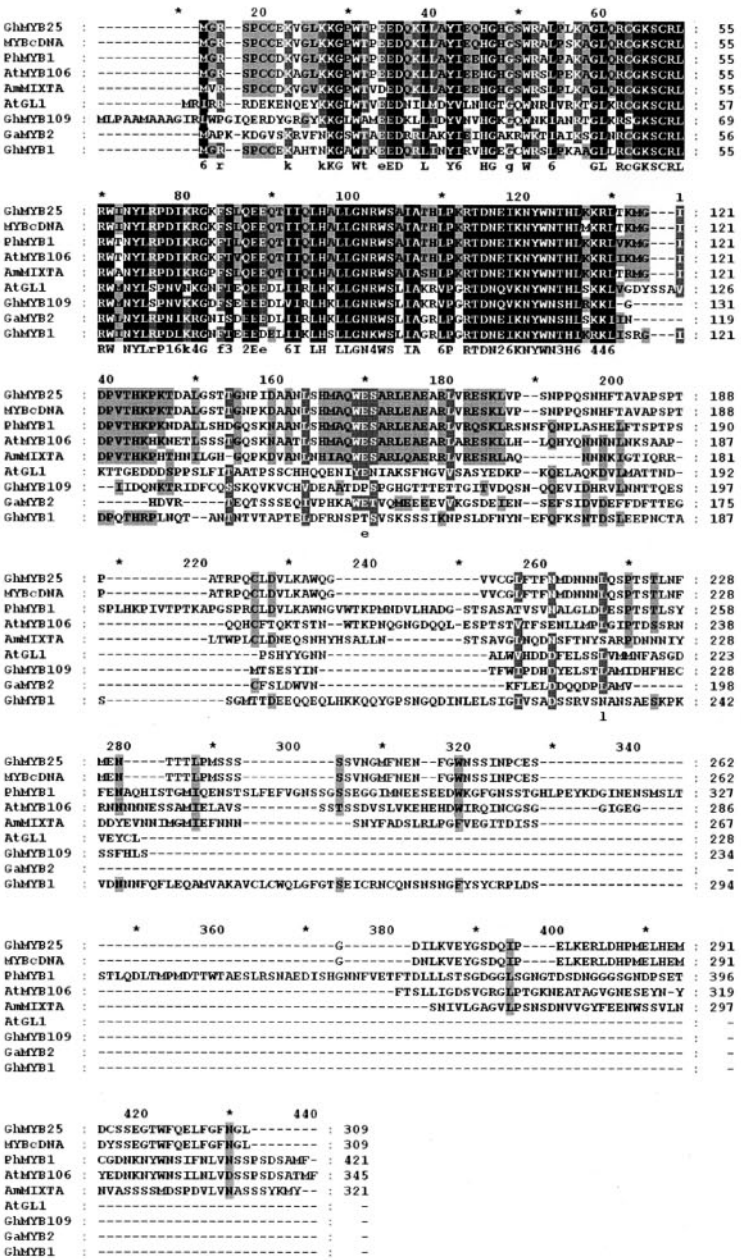
Based on their functional identity in a BlastX search and their patterns of expression, we could divide the genes into two classes: class I genes are regulatory genes that generally show a plateau of expression around 0 dpa; while class II genes tend to have more structural or biochemical functions and when compared with their pre-initiation values exhibit increasing expression towards +2 dpa (Fig. 4). The three genes, GhMyb25, GhHD1 and GhCycD3;1, in class I were considered more likely to have a role in the early events of fibre initiation at anthesis. The expression of GhCycD3;1 increases continually from –4 to 0 dpa and plateaus between 0 and +2 dpa, while GhMyb25 and the GhHD1 exhibit a dip in expression at –2 dpa followed by a peak at 0 dpa and then a slight decline towards +2 dpa. The plateau in expression at 0–2 dpa of these three genes coincides with the time of fibre initiation, although given the level of replication and hence standard errors, these expression patterns can only be taken as a rough guide [see reverse transcription (RT)–PCR results in Fig. 4].

The rest of the genes show a class II expression pattern although the specific details differ among them. The expression of GhFU1 (unknown glycine-rich protein gene) and GhSus (sucrose synthase) increases gradually in the time period examined. GhFaEl (identified as a possible transferase similar to a fatty acid elongase) and GhTMTP (transmembrane transporter) show a slight decrease from –4 to –2 dpa and then increase gradually towards +2 dpa. The third group from this class comprises GhEX1 (α-expansin) and GhLTP which show a distinctly flat profile from –4 to 0 dpa followed by a sharp increase from 0 to +2 dpa. The last member from this class GhFU2, which encodes an unknown protein, exhibits increased expression from –4 to –2 dpa and again from 0 to +2 dpa, while the expression between –2 and 0 dpa remains unchanged.

The identified candidate genes are also down-regulated in a completely independent fuzzless-lintless (fl) mutant

The parental genotypes (B1278 and SL1-7-1) of the Australian and American fibre mutants we analysed are either not available or unknown so could not be used for these experiments. An *fl* mutant has recently been isolated from the Chinese *G. hirsutum* cultivar Xu-142 and used to identify and characterize fibre development-related genes (Yu et al. 2000, C.H. Li et al. 2002, Ji et al. 2003). These two lines provided an isogenic pair with which to validate the genes identified from our other mutants. RNA from 0 dpa ovules of Xu-142 *fl* was compared with that of 0 dpa ovules of Xu-142, and 119 clones were identified as differentially expressed in the four replicates (supplementary Table S4). The same outer/inner integument gene expression filter was applied to the data set and identified 13 differentially expressed genes that were also more highly expressed in the outer integument of the linted DP16 ovules (Table 3). Amongst the 13 genes, eight are the same down-regulated genes identified in the other lintless mutant/DP16 comparisons and five are additional genes (three down-regulated and two up-regulated in the mutant) that had not been noted in

A.



B.

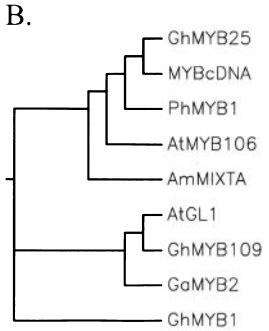


Fig. 5 (A) Alignment of the cotton GhMyb25 protein with other plant Myb proteins. GhMYB25, *Gossypium hirsutum* MYB25 (AF336283); MYBcDNA, the GhMyb25 cDNA clone (ON35F04) described in this work; PhMYB1, *Petunia hybrida* Myb.Ph1 (S26605); AtMYB106, *Arabidopsis thaliana* MYB106 (NP_186763); AmMIXTA, *Antirrhinum* MIXTA (S45338); GhMYB1, *Gossypium hirsutum* MYB1 (AAA33067) encoded by *COTMYBA*; AtGL1, *Arabidopsis thaliana* GLABROUS1 (CAB97485); GhMYB109, *Gossypium hirsutum* MYB109 (CAD71140); GaMYB2, *Gossypium arboreum* MYB2 (AY626160). GenBank accession numbers are in parentheses. Conserved amino acids are shaded. (B) Unrooted phylogenetic tree generated from the alignment of the above-mentioned Myb protein sequences using ClustalW.

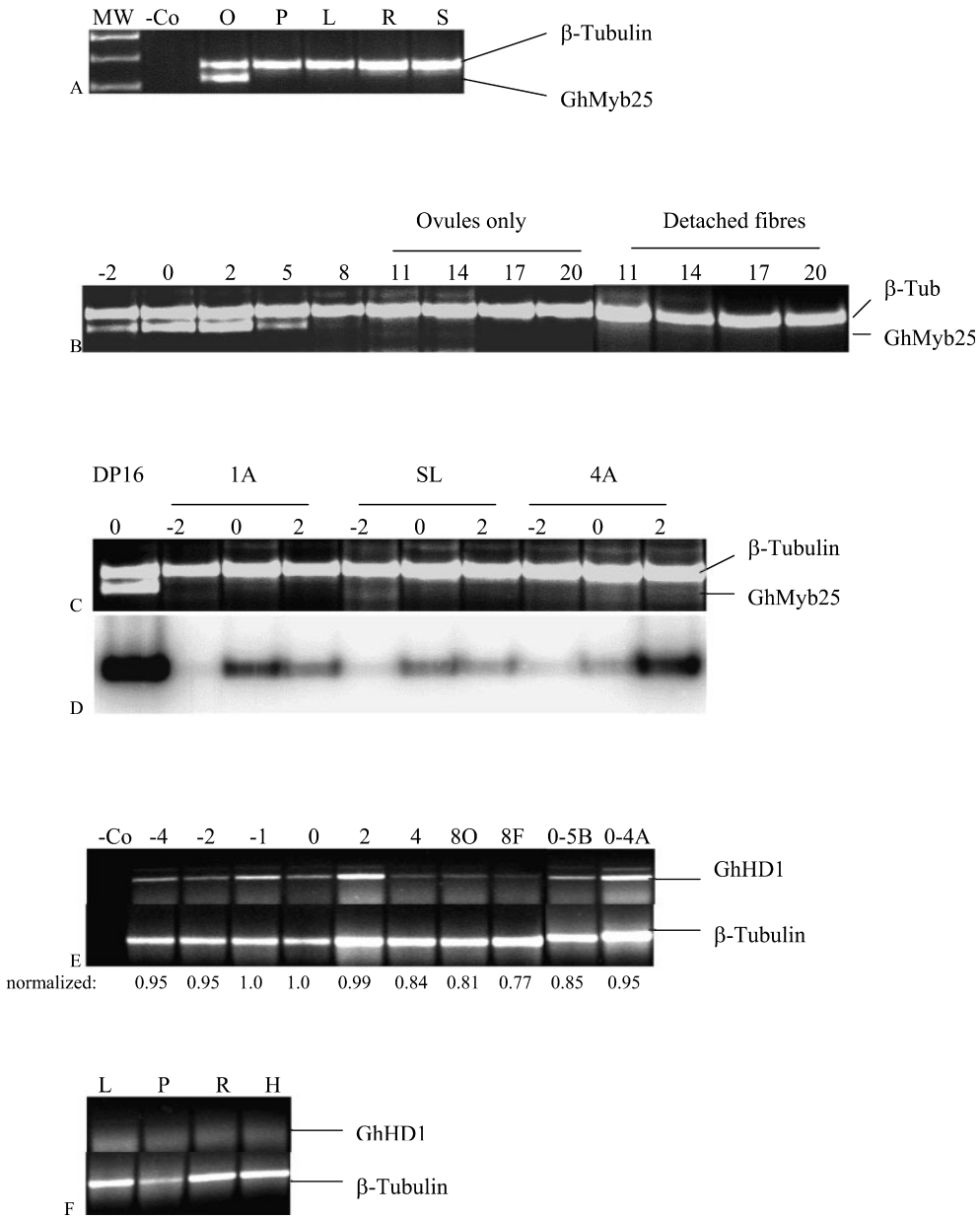


Fig. 6 RT–PCR analysis of GhMyb25 (A–D) and GhHD1 (E and F) expression in different cotton tissues relative to β-tubulin as the reference control. (A) Single reaction GhMyb25 and β-tubulin RT–PCR of RNA from different tissues from the wild-type DP16 cultivar. –Co, negative control, reaction without reverse transcriptase; O, 0 dpa ovule; P, petal; L, leaf; R, root; S, stem. (B) Ovule and detached fibres from wild-type DP16. The numbers indicate the corresponding dpa. (C) Ovules from DP16 and three reduced lint mutants: Lintless 1A, SL1-7-1 and Lintless 4A. The numbers indicated the corresponding dpa. (D) Southern hybridization of (C) hybridized with a GhMyb25 3′ gene-specific probe. (E) Separate GhHD1 and β-tubulin RT–PCR of RNA from ovule and detached fibres from the wild-type DP16, and ovules from two of the fibreless mutants. –Co, negative control, reaction without reverse transcriptase; –4 to 4, DP16 ovules of various stages (dpa) as indicated by the corresponding number; 8O, 8 dpa ovule of DP16 after fibres had been removed; 8F, 8 dpa detached fibres of DP16; 0–5B, 0 dpa ovule from the mutant Lintless 5B; 0–4A, 0 dpa ovule from the mutant Lintless 4A. Values under the panel are GhHD1 expression levels (after normalization against the β-tubulin band) relative to their value in 0 dpa ovules of DP16. (F) Different tissues from DP16. L, leaf; P, petal; R, root; H, hypocotyl. Faint expression was only noted in leaf RNA.

the previous comparisons. One of the new candidate genes, ON38N08 (886 bp), encodes a Myb protein which is 69% similar to the GhMyb25 protein and has a similar expression profile to GhMyb25, consistent with also being a class I gene (Fig. 4). It has been designated as GhMyb25-like. Other genes include an RD22 stress-related gene (down-regulated in the mutant), a non-specific LTP (up-regulated in the mutant), an L-asparaginase (down-regulated in the mutant) and a function-

ally unknown protein (GhFU3, up-regulated in the mutant). RD22 showed a similar expression profile to GhFaEl, while the non-specific LTP was similar to GhFU2, so both could be grouped with the other class II genes. GhFU3 had a relatively flat expression profile in wild-type ovules, while the L-asparanigase had an early peak at –2 dpa, so is different from the other genes (not shown).

GhMyb25 shows homology to AmMIXTA and is ovule specific

The most interesting of the candidate genes, ON35F04, was down-regulated in all six mutants. It is 1,160 bp in length and 98% identical to GhMyb25 (GenBank accession number AF336283), which encodes an R2R3 Myb transcription factor, expressed in 0 dpa ovules of *G. hirsutum* cultivar Acala Maxxa (B. Burr, unpublished). Because of its high similarity to GhMyb25, ON35F04 is presumed to be the same gene, although the GhMyb25 sequence in GenBank (AF336283) contains an unspliced intron not present in our clone. The small differences in sequence between the clones presumably result from polymorphisms between the different cultivars. It is also 96% identical to the *Gossypium arboreum* EST (BE054276), suggesting that it may be encoded by the A-genome present in tetraploid cotton. Outside the R2R3 region, which is highly conserved amongst all Myb transcription factors, GhMyb25 shows highest homology to the *Petunia hybrida* PhMYB1, *Antirrhinum majus* MIXTA (AmMIXTA) and *Arabidopsis* AtMyb106 (Fig. 5). It is more distantly related to *Arabidopsis* GL1 and cotton MYB1, another cotton myb which causes distinct abnormalities when overexpressed in transgenic tobacco including the production of cotyledonary trichomes (Payne et al. 1999).

GhMyb25 expression was detected only in 0 dpa ovules and not in petal, leaf, roots or stem using RT–PCR (Fig. 6A). Expression was detected in –2, 0, 2 and 5 dpa ovules, with the highest expression in 0 and 2 dpa ovules (Fig. 6B), consistent with the microarray data. No apparent expression of GhMyb25 was observed by RT–PCR in ovules of –2, 0 and 2 dpa from three of the mutants, Lintless 1A, SL1-7-1 and Lintless 4A (Fig. 6C), although very low expression was revealed after hybridizing RT–PCR products with an ON35F04 probe (Fig. 6D), consistent with the low levels of lint production that still occur on these ovules. GhMyb25 expression peaked at 0–2 dpa in lines Lintless 1A and SL1-7-1, whereas in Lintless 4A, the highest expression was detected slightly later at 2 dpa.

The putative homeodomain protein gene GhHD1 is expressed mainly in ovules

ON33M07 (GhHD1) was a second gene of interest because of the possible functional similarities between fibre development and *Arabidopsis* leaf trichome development. *Arabidopsis* trichome development has been shown to be regulated by both myb and homeodomain proteins. The evidence for the involvement of this gene in fibre initiation was not as strong as for GhMyb25 since it was differentially expressed in only three of the six mutants. The original clone was a partial cDNA clone of 442 nucleotides, but was extended by RT–PCR to 2,222 nucleotides (GhHD1), and encodes a protein with homology to two homeodomain proteins; protodermal factor 2 (GenBank accession number NP_567274) (507 out of 634 or 79% identical amino acids) and the L1-specific and ovule-specific homeodomain gene ATML1 (GenBank accession number T05850) (518 out of 657 or 78% identical amino acids). ATML1 has previously been grouped with the *Arabidopsis* GLABRA2 in the same HD-GL2 class and they share a common L1 layer-specific or dermal-specific pattern of expression (Lu et al. 1996). GhHD1 is only 43 and 42% identical to the other cotton homeodomain proteins that are present in GenBank, GhHOX1 (AAM97321) and GhHOX2 (AAM97322), respectively.

GhHD1 is mainly expressed in ovules of various developmental stages and at much lower levels in leaves as revealed by RT–PCR (Fig. 6E, F). Quantification relative to β-tubulin indicated that the expression of GhHD1, however, is relatively constant during early fibre development (Fig. 6), being expressed in both detached fibres and ovules. The expression levels of GhHD1 in 0 dpa ovules of mutants Lintless 5B and 4A relative to DP16 0 dpa ovules after normalization with β-tubulin expression were only slightly lower than in similar wild-type ovules (Fig. 6), at least for the two mutants examined, so is not quite consistent with the microarray results that show a convincing down-regulation in some, but not all, of the mutants. The RT–PCR band was confirmed as GhHD1 specific by Southern blot hybridization using the ON33M07 cDNA fragment as probe (data not shown).

GhMyb25 and GhHD1 are up-regulated in the fibre initials compared with epidermal cells

To investigate whether GhMyb25 and GhHD1 are expressed in fibre initials on the day of anthesis, laser capture microdissection (LCM) was used to isolate fibre initial cells and epidermal cells from tissue sections of 0 dpa DP16 ovules. RNA isolated from the captured cells was amplified and used in RT–PCR. Both GhHD1 and GhMyb25 exhibit fibre initial-enriched expression (Fig. 7A) with the expression of GhHD1 being >2-fold in fibre initial cells relative to in non-fibre epidermal cells, and GhMyb25 1.8-fold the expression in fibre initial cells relative to in non-fibre epidermal cells.

Heterologous overexpression of GhMyb25 in tobacco increases the number of branched long-stalked trichomes

An overexpression construct with the subterranean clover stunt virus promoter 7 (Schünmann et al. 2003) driving the expression of the full-length GhMyb25 cDNA clone was introduced into tobacco and *Arabidopsis* by *Agrobacterium*-mediated transformation. In transgenic *Arabidopsis*, the expression of GhMyb25 had no obvious effect on trichomes in several transformants analysed (not shown). No trichomes were observed on the seeds of either species transformed with GhMyb25. Six independent primary transformants (T_0 genera-

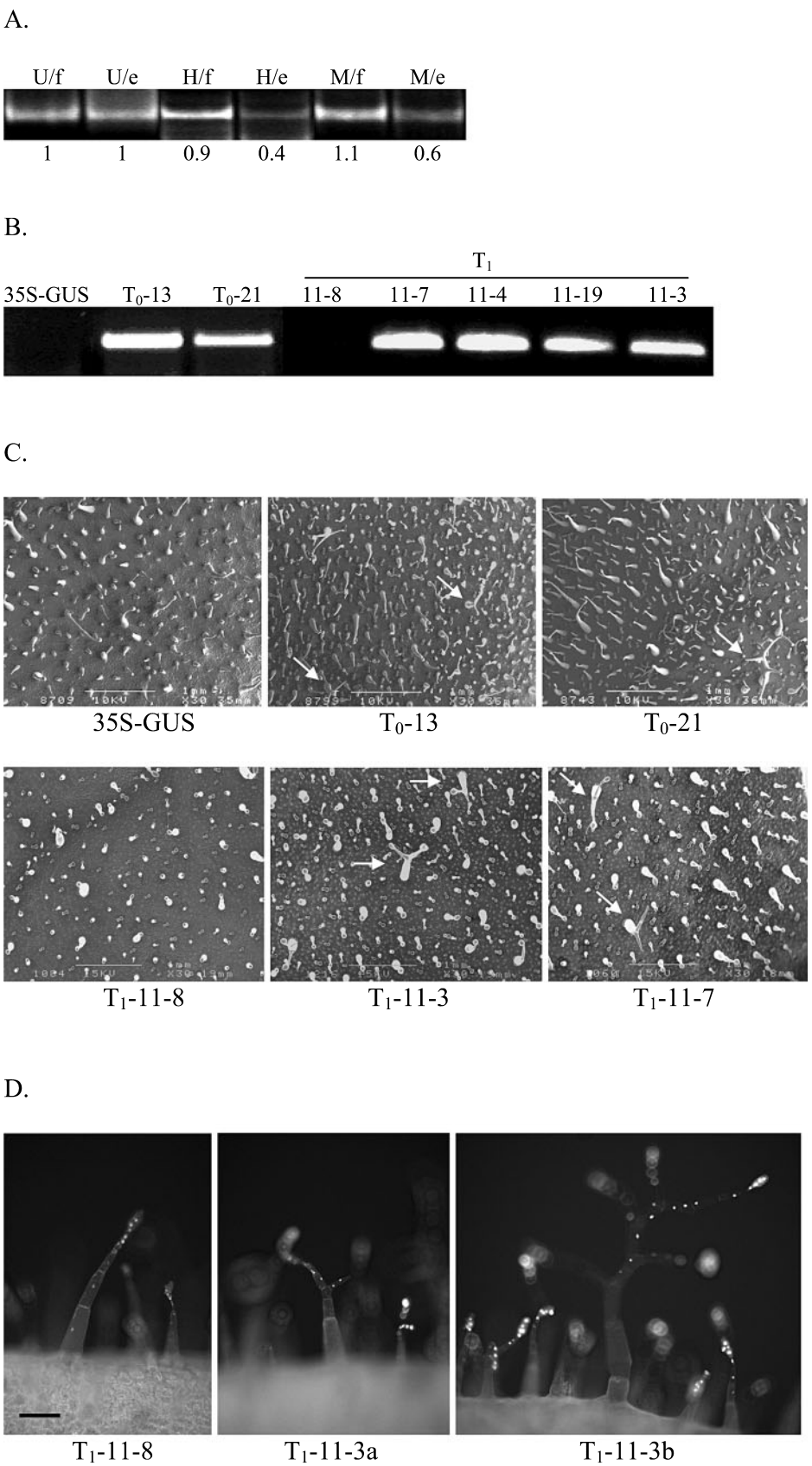


Fig. 7 (A) Expression of GhMyb25 and GhHD1 in amplified RNA isolated from laser-captured cotton fibre initial cells and non-fibre epidermal cells of DP16 0 dpa ovules determined by RT–PCR with the cotton polyubiquitin gene as the reference control. U, H, and M represent polyubiquitin, GhHD1 and GhMyb25, respectively. f and e, fibre initial cells and epidermal cells, respectively. The quantification of the RT–PCR band is normalized using the corresponding polyubiquitin band. The standard deviation for the quantification is estimated at 0.08 based on two replications. (B) Expression of GhMyb25 in two primary transgenic tobacco lines (T_0-13 and T_0-20) revealed by RT–PCR with a 35S–GUS-transformed tobacco as a negative control and five T_1 plants from transgenic line 11 (11–8, 11–7, 11–4, 11–19 and 11–3), with plant 11–8 showing no expression of GhMyb25, as a negative control. (C) Increased number of branched long-stalked trichomes (indicated by an arrow) on the leaf adaxial surface of primary transgenic tobacco lines T_0-13 and T_0-20 as compared with the control tobacco leaf (35S–GUS) and on the adaxial leaf surface of T_1 GhMyb25-expressing plants 11–3 and 11–7 as compared with the negative control plant 11–8. (D) Long-stalked trichome on the leaf adaxial surface of T_1 tobacco plants 11–8 (control) and 11–3. The number of cells per trichome was revealed by DAPI staining and fluorescence microscopy. The scale bar is 200 μm.

Table 4 The significant increase of long-stalked trichome density from the adaxial leaf surface of four T_1 plants of tobacco line 11 expressing GhMyb25 as compared with a negative control tobacco plant (T_1-11–8) that showed no expression of GhMyb25

	T_1-11–8	T_1-11–7	T_1-11–4	T_1-11–19	T_1-11–3
Mean long-stalked trichomes/1,000 epidermal cells [a]	18.5	26.2	21.4	20.3	23.7
SD	1.7	2.6	2.3	1.9	2.9
P-value [b]	–	1.7×10^{-6}	4.4×10^{-4}	0.029	2.5×10^{-4}
Branched trichomes/10,000 epidermal cells	0.6	5.3	4.2	2.3	6.8

[a] Averaged over nine SEM images covering a total leaf area of about 100 mm^2.
[b] Significance level of *t*-test. $P > 0.05$ indicates a significant increase of long-stalked trichome density compared with the null segregant control (T_1-11–8).

tion) of tobacco were selected based on their expression of GhMyb25 by RT–PCR (Fig. 7B, two lines only shown). These lines were examined by SEM and all six showed increased numbers of the branched long-stalked trichomes that are not normally formed on the adaxial leaf surface (e.g. Fig. 7C, two lines shown). Branched long-stalked trichomes were occasionally observed on leaf veins of wild-type or other control tobacco lines, but on the transformants, a 2- to 10-fold increase in these trichomes was observed primarily on parts of the leaf other than on veins and, as indicated below, the trichome phenotype was inherited with the transgene in the T_1 progeny.

Heritability of the expression of the GhMyb25 and its effect on trichome number and branching was examined using genomic PCR, RT–PCR and standard microscopy in the T_1 generation (minimum 20 plants per line) of the six transgenic tobacco lines. A perfect correlation between the increased number of branched long-stalked trichomes and transgene expression was observed in all the six lines. Line 11, for example, showed a 3 : 1 transgene segregation ratio. Five plants from this line were subjected to detailed trichome number analysis using the more sensitive SEM. The number of long-stalked trichomes, short trichomes, stomata and epidermal pavement cells were counted on SEM images taken at the same magnification. The number of the long-stalked trichomes, short-stalked trichomes and stomata per thousand epidermal cells from the four plants expressing the GhMyb25 transgene (11–7, 11–4, 11–19 and 11–3) were compared with that of a null segregant plant, 11–8, that showed no expression of the transgene (Fig. 7B). Amongst the three cell types, only the density of long-stalked trichomes showed a small (about 20%, averaged over the four plants), but significant, increase over that of plant 11–8 (Table 4). The short-stalked trichome and stomata had variable densities, with some plants showing slightly higher and other plants slightly lower numbers when compared with 11–8 (not shown), but none of these changes were statistically significant. The most visible phenotype, however, was the branching of the long-stalked trichomes observed previously on the T_0 plants. On average, branching was seven times more prevalent in the GhMyb25-expressing T_1 lines than in the null segregant controls (Table 4). Examples of the branched trichomes on the non-vein areas of leaves from plants 11–3 and 11–7 are shown in Fig. 7C, with those of the null segregant plant 11–8 as the negative control. A long-stalked trichome generally consists of 8–9 cells (Fig. 5D, T_1-11–8). Branched trichomes have both long branches with a similar number of cells as an unbranched trichome and short branches consisting of 2–3 cells (Fig. 7D, T_1-11-3a and 3b).

Fibre cells undergo DNA endoreduplication during initiation

The down-regulation of expression of a putative cyclin D3 gene encoded by clone OCF07F04 in the ovules of a couple of the fibreless mutants prompted us to investigate the cell division and DNA replication activities of the epidermal layers of DP16 and Lintless 4A ovules.

Since fibre cells are known to cease dividing after differentiation, the ovule epidermal cell division activities of DP16 and 4A-183 were examined at and before anthesis. Over the period examined (–2, –1 and 0 dpa), the extent of cell division in the ovule epidermis of Lintless 4A and DP16 was not significantly different (paired *t*-test, $P = 2.1\%$) although the division rate in Lintless 4A was slightly higher than in DP16 (Table 5). It appears that the ovule epidermis cell division rates are higher at –1 dpa for both DP16 and Lintless 4A, although the significance of this is unclear.

Relative DNA contents of the epidermal cells and fibre cells were also measured, and normalized using DNA contents of nuclei at anaphase or telophase (2C), and the results are presented in Fig. 8. Since the pre-fibre initials and epidermal cell are visually indistinguishable at –2 and –1 dpa, the data for

Table 5 Cell division rates in ovule epidermal cells of the cotton cultivar DP16 and the reduced fibre mutant 4A

dpa	DP16	4A
–2 dpa	2.78%	2.97%
–1 dpa	3.25%	3.5%
0 dpa	2.69%	2.84%

Values are the percentage of epidermal cells undergoing mitotic divisions.

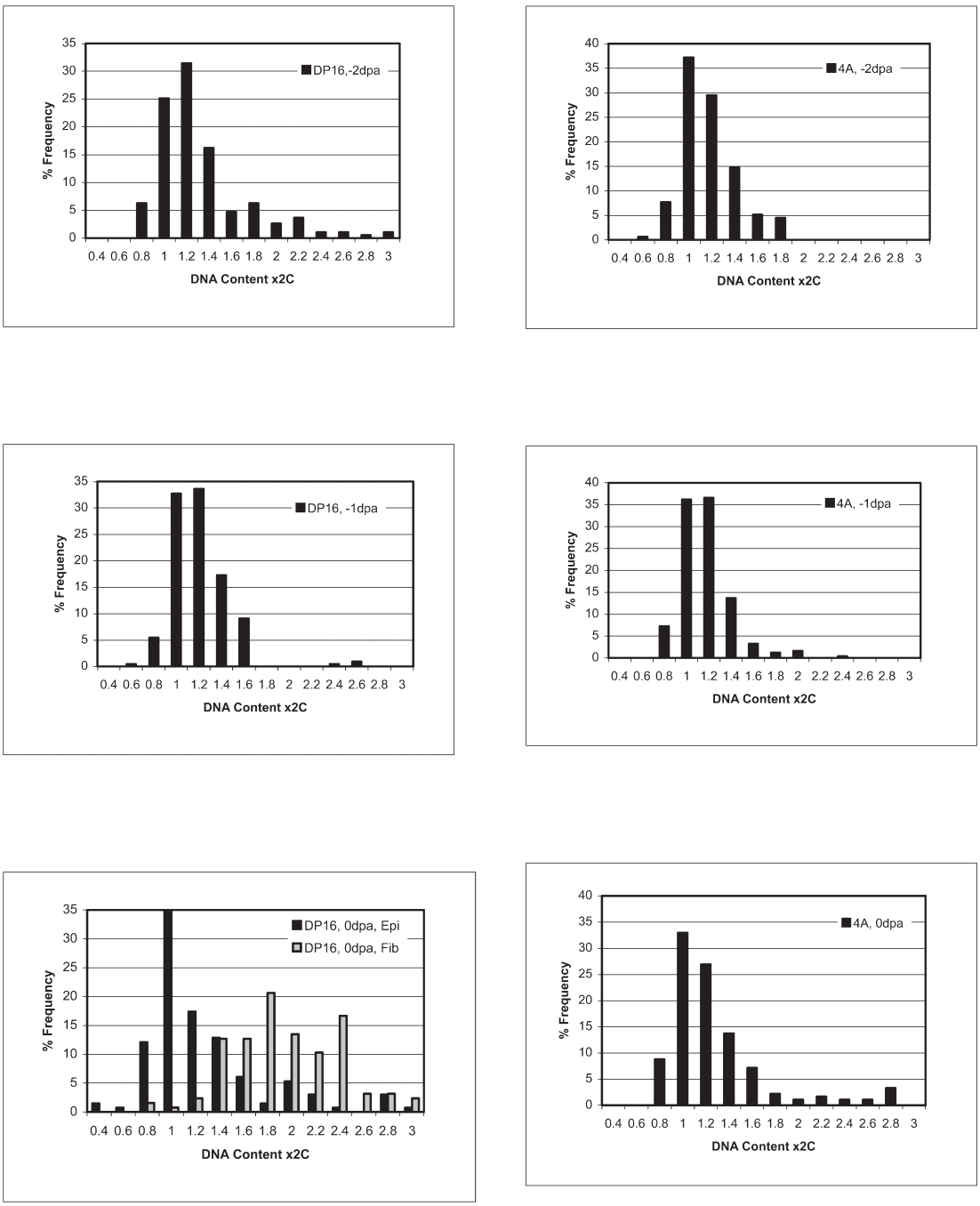


Fig. 8 Histograms of relative DNA contents of cotton ovule epidermal cells and fibre cells relative to 2C content of anaphase or telophase nuclei. Left column: distributions of cells in different DNA content bins at three developmental stages (–2, –1 and 0 dpa) (top to bottom) of DP16 ovule epidermal cells (black columns) and fibre cells (grey columns) presented as a percentage of the total number of measured cells. Fibre cells and epidermal non-fibre cells were only distinguishable at 0 dpa. Right column: the distribution of Lintless 4A ovule epidermal cells (black columns) in different DNA content bins at the corresponding developmental stages.

those times are presented as total epidermal cells. The results indicate that epidermal pavement cells of DP16 and Lintless 4A at these time points have a DNA content peak at around 2–2.4C, while differentiated fibre initials of DP16 when clearly distinguishable at 0 dpa have an increased DNA content, with the majority of cells showing DNA contents between 2.8 and 5.2C. While this result clearly suggests that the majority of fibre cells undergo at least one round of DNA endoreduplication during initiation, the involvement of the cyclin D3 gene in this process still needs to be verified.

Discussion

The complex nature of the regulation of epidermal cell differentiation as illustrated by *Arabidopsis* trichome and root hair development suggests that there may also be a large number of genes controlling cotton fibre initiation and subsequent development. Arpat et al. (2004) noted thousands of elongation phase-specific genes in their analysis of cotton fibre ESTs and from microarray comparisons between fibres in the mid-elongation phase compared with the secondary cell wall phase, but there has been little analysis on the early stages of fibre initiation. We developed an ovule-specific cDNA microarray containing about 10,000 randomly selected cDNA clones to investigate gene expression differences between ovules of linted cotton and a set of lintless mutants on the day when fibre initiation begins. Among the genes down-regulated in the lintless mutants, we identified a number of genes as candidates for having a role in fibre development, including Myb and homeodomain transcription factors that have some similarity to the types of genes known to regulate trichome development in *Arabidopsis thaliana*. We have verified that the cotton Myb has some effect on trichome development in a transgenic tobacco system, although it did not appear to affect *Arabidopsis* trichome development.

The use of multiple mutant lines in these microarray expression studies had a number of advantages, not the least being that it helped to smooth out 'noise' contributed by the biological variability in gene expression during seed and fibre growth and the diverse genetic backgrounds of the mutants. Genetic analysis of the mutants to determine their interrelationships proved inconclusive, but at least two distinct groups of mutants were identified. Despite these difficulties in confirming the genetic nature of the mutants, similar sets of genes were repeatedly identified as underexpressed in the different mutants, attesting to their specific role in fibre initiation and/or early fibre development. Overlaying the expression patterns between the inner and outer layers of wild-type cotton ovules on that of the wild-type/mutant whole ovule comparisons served as a convenient in silico filter to reduce the complexity of the gene lists generated and to eliminate those genes not normally expressed in the outer integument where fibres initiate.

The above filtering strategy led to the detection of only 13 different genes whose normal increase in expression at initiation is prevented or delayed in the fibreless mutants. Not surprisingly, these genes could be classified into two classes; a regulatory class, class I (GhMyb25, GhMyb25-Like, GhHD1 and GhCycD3;1), that had similarities to known transcriptional regulators or cell cycle genes, and a biochemical/ structural class, class II, generally expressed later, suggesting a role in fibre cell expansion and/or elongation. Some of the class II genes (GhSus, α-expansin and GhLTP) were already known and had well defined biochemical or cell biology functions as indicated in the Introduction. New members of this class either encode structural proteins or have been shown to be involved in cell wall biogenesis in other species. The class I genes were of most interest, as similar types (Myb or homeodomain transcription factors) are known to regulate *Arabidopsis* trichome development.

Cotton, like other plants, contains many Myb genes most of whose functions remain unknown. GhMyb25, and the GhMyb25-like gene, show higher sequence similarity (Fig. 5) to the *P. hybrida* PhMyb1 and the *Antirrhinum* MIXTA factors than to *Arabidopsis* GL1 or other cotton Mybs, but are unlikely to be direct homologues as they do not show the same petal epidermis-specific expression of these genes (Noda et al. 1994, Solano et al. 1995). MIXTA is a regulator of the conical shape of the petal epidermal cells (Noda et al. 1994), and overexpression in transgenic tobacco caused the production of supernumerary trichomes on cotyledons, leaves, stems and floral organs, as well as the production of novel conical cells on leaves (Glover et al. 1998, Payne et al. 1999). GhMyb25, on the other hand, only caused an increase in the branching of long-stalked trichomes and a small increase in the total number of long-stalked trichomes in tobacco, but did not have any effect on short trichomes or epidermal cell shape. Furthermore, GhMyb25 is expressed only in the ovules, predominantly in fibre initials and not in later stage fibres, or petals. Unlike cotton fibres, tobacco trichomes, particularly the long-stalked trichomes, are multicellular structures with a long stalk of 3–4 cells supporting a small structure of 5–6 cells. Branching of these trichomes is normally very rare. The primary effect of GhMyb25 expression in tobacco was to increase branching, apparently caused by secondary initiation of new long trichome structures from the cells of the main stalk. This ability to initiate trichome development is clearly weak in tobacco as there is only a small increase in long trichome numbers and the majority of the long trichomes remain unbranched, so it will be interesting to see if overexpression of this gene has a more profound effect on epidermal cell initiation into fibres in transgenic cotton plants.

A large number of Myb genes are expressed in cotton ovules at the time of fibre initiation (Loguerico et al. 1999, Suo et al. 2003), but not all of them are specific to fibres and may be involved in other metabolic or cell cycle processes occurring in the nucellus or subepidermal layers. Some, like GhMYB109 (Suo et al. 2003) that is structurally related to the *Arabidopsis* GL1 and WER genes, are expressed in fibre initials, but have much higher expression in the rapidly elongating fibres so may have more general roles than in controlling initiation. Another pair of Myb genes isolated from fibres (GhMyb7 and GhMyb9, thought to be homeologues) (Hsu et al. 2005) are expressed in flowers and elongating fibres, but when overexpressed in *Arabidopsis* and tobacco had dramatic effects on plant development, suggesting they may have more general roles than fibre development. GaMyb2 (Wang et al. 2004) from *G. arboreum* (GenBank accession number AY626160) clusters with GL1 (Fig. 5) and complements a *gl1*

insertional mutation in transgenic *Arabidopsis* so has been proposed as its functional homologue in cotton. They saw the occasional single very short trichome on seeds of *Arabidopsis* plants constitutively overexpressing GaMyb2, but not on plants overexpressing either GL1 or GhMyb109. The significance of this is unclear and does not provide proof for a role for GaMyb2 in cotton fibre initiation. It should be emphasized that none of the fibreless cotton mutants is affected in leaf trichome development, suggesting distinct genetic programmes controlling the development of these two different types of trichomes in cotton. How all these different Mybs fit into the fibre development pathway with GhMyb25 and GhMyb25-like has yet to be elucidated. Experiments are underway to silence and overexpress our two GhMyb25 genes in transgenic cotton plants, but the long time frames for cotton transformation preclude us including this in this study.

The identification of a second Myb transcription factor in the Chinese linted and lintless lines, the GhMyb25-like gene, suggests that there may be some redundancy in fibre regulation. It too was more like MIXTA (Stracke et al. 2001) than GL1 and had a similar expression pattern to GhMyb25. Nucleotide comparisons between the GhMyb25-like gene and GhMyb25 (probably an A-genome isoform) suggest that it is not the D-genome isoform of GhMyb25 so there may be even more members of this gene family to discover in tetraploid cotton.

The role in fibre development of the putative homeodomain gene identified in this study has not yet been demonstrated. Its expression is predominantly in ovules, with higher expression in the outer integument and in fibre initials during initiation as well as in the elongating fibres. It is expressed at a much lower level in leaves (where trichomes develop). Its expression in the fibreless mutants was, however, not convincingly reduced so it could perhaps have a more general role in epidermal cell differentiation and growth. Functional characterization in transgenic cotton is in progress and will confirm whether GhHD1 has more relevance to leaf and/or seed trichome development. GhHD1 does not appear to be a direct homologue of the Arabidopsis GL2 or the two putative GL2 homologues from cotton, GhHox1 and GhHox2 (AF530913 and AF530914, respectively). It is more similar to the L1 layer-specific homeodomain gene ATML1 [that is classified in the same HD-GL2 class as GL2 based on sequence homology (Lu et al. 1996)] that is proposed to be involved in setting up morphogenetic boundaries of positional information necessary for controlling cell specification and pattern formation, so it is intriguing to speculate that GhHD might play a similar role in fibre initial patterning.

Among the three additional genes identified in the analysis of the Xu-142 pair of lines was an RD22 gene, similar to that expressed during fibre elongation identified by C.H. Li et al. (2002) using cDNA macroarrays. The promoters of other RD22-like and LTP genes have been found to contain Myb (and in the case of RD22-like, homeodomain) binding sites (Wang et al. 2004, Hsu et al. 2005) and can be transactivated by Myb genes, confirming that they are downstream in the regulatory cascades controlling fibre development. The asparaginase down-regulated in Xu-142 *fl* is an asparagine catabolic enzyme (EC 3.5.1.1) that may supply a reduced form of nitrogen to the developing seeds and fibres. Asparagine has been suggested as the major form in which nitrogen is transported in developing cotton seeds (Elmore and Leffler 1976).

DNA endoreduplication, a strategy to amplify nuclear DNA without cell division (reviewed by Joubès and Chevalier 2000), is common in cell types with a high metabolic activity and/or large size. It is well established that *Arabidopsis* trichomes undergo four rounds of endoreduplication during development, leading to large branched cells with nuclei containing about 32C DNA (Schnittger et al. 2002), but it has not been clear whether cotton fibre initials undergo a similar process. Van't Hof (1999) reported that the DNA content of nuclei isolated from developing cotton fibre cells only increased by about 24% from 2 to 5 dpa and suggested that this may be the result of selective amplification of parts of their genome. Using laser confocal microscopy and propidium iodide staining (Fig. 8), we have demonstrated that relative to the adjacent epidermal cells, the majority of the fibre initials on the day of anthesis have higher than 2C DNA content (between 2.8 and 5.2C), suggesting that they are undergoing at least one round of DNA endoreduplication starting at 0 dpa. Assuming endoreduplication involves a modified cell cycle, it may share common determinants with the classic cell cycle and cell cycle genes such as cyclins (Joubès and Chevalier 2000). The *Arabidopsis* CycD3;1, for example, when ectopically expressed, induced not only DNA replication, but also cell division in trichomes (Schnittger et al. 2002). The decreased expression of the GhCycD3;1 in the outer integuments of the Lintless 4A mutant did not appear to affect epidermal cell division rates (Table 5), so it is tempting to speculate that this gene may be involved in the DNA endoreduplication, but this will have to await further analysis.

The identification of 13 candidate genes, in particular the three transcription factors and a cell cycle gene, that appear to be involved in fibre initiation and elongation has provided targets for overexpression and knockout studies in transgenic cotton to elucidate further the molecular mechanisms that control the differentiation of this most elongated plant cell.

Materials and Methods

Plant material

Eight cotton lines (*Gossypium hirsutum* L.) were used in this study and these include two wild-type cotton cultivars: DP16 and Xuzhou 142 (also called Xu-142 or XZ142), and six lintless lines: Lintless 1A, Lintless 4A, Lintless 5B, Lintless 53, SL1-7-1 and Xu-142 *fl* (also called XZ142w). The 1A, 4A, 5B and 53 lines were obtained from the Queensland Department of Primary Industry Tropical Crops and Pastures Germplasm Collection and were originally selections from a linted cultivar B1278 isolated by Dr. Alistair Low (unpublished, CSIRO Irrigation Research, Griffiths, NSW). SL1-7-1

was obtained from the United States Department of Agriculture-Agricultural Research Service (College Station, TX, USA). The Xu-142*fl* mutant and its parental Xu-142 were provided by Professor Xiao-Ya Chen (Institute of Plant Physiology, Chinese Academy of Science, Shanghai, PR China). All the cotton lines were grown in a glasshouse with a temperature of 30°C/22°C (day/night). Ovules were always collected at a similar time each day (1–3 p.m.) and samples to be compared on a microarray were only used when they were collected on the same date and from the same glasshouse to minimize between time or location variability.

Scanning electron microscopy (SEM)

Cotton ovules were collected at developmental stages of –2, 0 and 2 dpa as described by Hasenfratz et al. (1995). Ovules were frozen in an Oxford CT 1500 cryotrans system, gold coated and observed using a JEOL 6400 scanning electron microscope as described by Craig and Beaton (1996).

Tobacco primary transformants were isolated from kanamycin selection medium and transferred to soil after root formation. A top leaf around 120 mm long was taken for SEM imaging from a control plant transformed with a 35S–β-glucuronidase (GUS) construct and six independent transgenic lines. In order to minimize variations, an adaxial leaf area of 100 mm^2 from a similar position on each leaf was visualized by SEM, photographed and used for the trichome counting. T_1 plants were first germinated in tissue culture, genotyped by PCR for the inserted transgene and transferred to soil after 2 months. Trichomes were viewed and counted as for the T_0 plants. The number of cells in each long-stalked trichome was examined using fluorescence microscopy after 4,6-diamidino-2-phenylindole (DAPI; 1 μg ml^{-1} in water) staining.

The cotton ovule cDNA microarray

The cotton ovule cDNA microarray comprised a total of 10,410 PCR-amplified inserts from cDNA clones printed to glass CMT-GAPS II™-coated microarray slides (Corning, Acton, MA, USA) as described in Wu et al. (2005). Except for 52 clones encoding known cotton genes provided by colleagues, and 13 clones of negative controls (non-plant genes, intron sequences, etc.), the rest of the clones were randomly picked from three cDNA libraries derived from early stage cotton ovules of cultivar DP16. This included 5,496 clones from a library derived from cycloheximide-treated cultured ovules (CHX library), 1,149 clones from 0 dpa whole ovules (OCF library) and 3,700 clones from the OCF library after normalization (ON library). The sequences of all clones and the array layout are available at http://www.pi.csiro.au/gena/ through the link from this manuscript, and all the EST sequences have been lodged with GenBank as accession numbers DT455583–461485 (ON ESTs), DT461486–469116 (CHX ESTs) and DT526800–527666 (OCF ESTs). Sequence data for the ovule ESTs and relative EST abundance data in a variety of cotton tissues for these and 160,000 other cotton ESTs are also available from the Arizona Genomics Institute website http://www.agcol.arizona.edu/pave/cotton/.

RNA isolation and microarray hybridizations

Cotton ovules used for RNA isolation were kept in RNALater (Ambion, Austin, TX, USA) solution at –20°C. Ovules were used whole or the outer integument was separated from the inner tissues under a microscope at room temperature as required. Total RNA isolations and poly(A)$^+$ mRNA purification were as described in Wu et al. (2005). Equal amounts of mRNA (0.5–1 μg) of the two samples to be compared were reverse transcribed using Superscript II reverse transcriptase (Invitrogen Life Technologies, Carlsbad, CA, USA), using 1 μg of oligo(dT)12–18 primer and 6 μg of random primers (Invitrogen Life Technologies) per reaction. The purification and Cy3-dUTP and Cy5-dUTP (Amersham Pharmacia Biotech, Little Chalfont, UK) labelling of the first strand cDNA was as described by Schenk et al. (2000). The labelled probes were combined and purified using a Qiaquick PCR purification kit. Slide hybridization and washing were as described in the manufacturer's instruction manual (Corning, CMT-GAPS coated slides).

Statistic analysis of microarray data

Microarray images were captured using a GenePix 4000A microarray scanner (Axon Instruments, Union, CA, USA) and analysed using the GenePix Pro program (Axon Instruments). Grids were predefined and manually adjusted to ensure optimal spot recognition. Spots were quantified using the GenePix fixed circle method, and medians of the fluorescence intensity of the red and green channels were used to generate the ratio of the two channels.

The statistical analysis of the microarray data was carried out using tRMA (tools for R Microarray Analysis, Version 1.7.0), a suite of statistical functions written in R code (Wilson et al. 2003). The detailed description of tRMA functions is available online (http://www.cmis.csiro.au/iap/tRMA/). Using tRMA, the red/green fluorescence ratio data were $\log_2$ transformed and normalized using a spatial normalization method described by Wilson et al. (2003). The data were then rescaled by dividing by an estimate of the median absolute deviation before running the 'find differentially expressed gene' function of tRMA as described by Klok et al. (2002). The rescaling allows for multiple comparison corrections between different experiment and replications. The 'find differentially expressed gene' function assumes that the non-differentially expressed genes follow a standard normal distribution and many differentially expressed genes will fall outside this distribution, as outliers. The detection of differentially expressed genes was computed by removing the 'outliers' one by one until the remaining data distribution is not different from the standard normal distribution (using the residue sum of squares of fit of data distribution to the standard distribution in a quantile–quantile plot as the optimization criteria). For a typical microarray comparison in this study that consisted of four replications, the 'find differentially expressed gene' function of tRMA was used to select differentially expressed genes from each of the replications separately, and the differentially expressed genes were then compared using the 'compare interesting genes' function of tRMA to refine a new gene list of the differentially expressed genes occurring in at least three of four replications (or five of six, or six of eight replications). All microarray data will be available at http://www.pi.csiro.au/gena/ after publication.

RT–PCR

Total RNA isolated from cotton tissues were DNase (RQ1 RNase-free DNase, Promega, Madison, WI, USA) treated and 0.5 μg of the total RNA was used in an RT–PCR essentially as described by McFadden et al. (2001). Since all β-tubulin cDNA clones on the array showed relatively consistent expression in 0 dpa ovules of all the cotton lines, one β-tubulin cDNA clone (OCF006F9) was used as a control in all the RT–PCRs. The forward and reverse primers used for the β-tubulin are 5′-AGAACATGATGTGTGCTGC-3′ and 5′-AGCTGTGAACTGCTCACTC-3′, respectively, and the resulting cDNA fragment was 300 bp.

The forward and reverse primers used for GhMyb25 RT–PCR were: 5′-TCAAACCCTCCTCAAAGCAACC-3′ and 5′-ATTCCATTACCAGACGATGATGAC-3′, respectively, and this produced a cDNA fragment of 224 bp. The GhMyb25 and β-tubulin RT–PCRs were performed in a one-tube reaction amplified with an initial denaturation cycle at 95°C for 3 min followed by 23 cycles at 95°C for 15 s, 55°C for 15 s, 72°C for 1 min and with a final cycle of 72°C for 2 min. A 5 μl aliquot of the RT–PCR was resolved on a 2% agarose

gel and the gel was Southern blotted to Hybond-N^+ membrane (Amersham Pharmacia Biotech, Little Chalfont, UK) and hybridized with a ^{32}P-labelled probe derived from the cDNA clone of GhMyb25.

The forward and reverse primers used for the GhHD1 RT–PCR were: 5′-GCTTTCTCTTGGATCAG-3′ and 5′-CAATAACACATGAAACCAG-3′, respectively, and these resulted in a cDNA fragment of 384 bp. The GhHD1 and β-tubulin RT–PCRs were performed in separate reactions amplified with an initial denaturation cycle at 95°C for 3 min followed by 20 and 16 cycles, respectively (to keep the amplification in the linear range), at 95°C for 15 s, 57°C for 15 s, 72°C for 1 min and with a final cycle of 72°C for 2 min. RT–PCRs were resolved on 2% agarose gels, Southern blotted to Hybond-N^+ membrane and hybridized with a ^{32}P-labelled probe derived from the cDNA clone of the putative homeodomain gene. Since the expression of β-tubulin appeared to be variable between different cotton tissues, the quantification of GhHD1 expression using the β-tubulin gene as a standard was only performed on the 0 dpa ovules of different cotton lines using Multi Gange V 2.11 (FUJIFILM, San Jose, CA, USA).

Laser capture microdissection (LCM)

Ovaries of DP16 were fixed in 75% (v/v) ethanol and 25% (v/v) acetic acid on ice immediately after the ovaries were dissected from 0 dpa flowers and ovary wall removed. The subsequent infiltration of the fixative, 10% (w/v) sucrose and 15% (w/v) sucrose, was as described by Nakazono et al. (2003). The ovaries were then embedded in TissueTek OCT (Sakura Finetechnical, Tokyo, Japan), frozen immediately on brass stubs and sectioned at 40–50 μm in a cryomicrotome (Model CT1, International Equipment Co., Nedham Heights, MA, USA). The tissue sections were mounted on polylysine-coated slides (Polysine, Biolab Scientific, Clayton, Australia), air dried and then dehydrated for 1 min in each of 70, 95 and 100% ethanol on ice. The slides were then stored at –80°C.

The PALM laser capture system (P.A.L.M. Microlaser Technologies AG Inc., Bernried, Germany) was used for LCM. The slides were removed from the freezer, dehydrated in 100% ethanol for 3 min and air dried before LCM. The LCM was performed according to the manufacturer's instructions. Individual fibre initial cells (a total of approximately 400 cells) or epidermal cells (a total of approximately 100 cells) were catapulted without pre-cutting into 45 μl of RNALater then stored at –20°C. Visual examination of the sections after each collection was used to verify that predominantly only cells of the designated cell type were collected, but some cross-contamination was likely.

RNA was extracted from the captured cells using the method described by Wu et al. (2002) with some modifications. The sample of the captured cells (with 45 μl of RNALater) were homogenized in 500 μl of RNA extraction buffer with 20 ng of carrier RNA (carrier RNA from the Qiagen RNeasy micro kit) using a Ystral homogenizer (HD Scientific, Sydney, Australia). After spinning for 2 min at maximum speed in an Eppendorf microcentrifuge, the supernatant was transferred to a fresh tube. A 250 μl aliquot of ethanol was added to the sample before the sample was loaded onto a Qiagen RNeasy mini column. The column washing and RNA elution were as described in the Qiagen RNeasy mini kit protocol. The eluted RNA was concentrated under vacuum until the remaining volume was about 10 μl.

The isolated RNA was amplified using a MessageAmp aRNA Kit (Ambion, Austin, TX, USA) following the manufacturer's instructions. Two rounds of amplifications were performed and the resulting antisense RNA was quantified at OD_{260} and resuspended at a concentration of 100 ng μl^{-1}.

The RT–PCR of Myb25 and GhHD1 using the amplified RNA was essentially as described earlier except that 100 ng of amplified RNA (instead of 0.5 μg of total RNA) per reaction was used as template. In addition to the β-tubulin-positive control, a cotton polyubiquitin (CK738219 in Dowd et al. 2004) was used as a normalization standard. The forward and reverse primers used for polyubiquitin RT–PCR were: 5′-CAAGACAAGGAAGGCATCCCAC-3′ and 5′-TCGGAACTCTCCACCTCCAAAG-3′, respectively, and these resulted in a cDNA fragment of 200 bp. All RT–PCRs were amplified using the same RT–PCR programme (described earlier) with 28 cycles and the resulting RT–PCR bands were quantified using Multi Gange V 2.11 (FUJIFILM) and then normalized using the corresponding polyubiquitin bands.

Overexpression of Myb25 in tobacco

The coding region of Myb25 cDNA as a 0.9 kb fragment was cloned into the *Eco*RI site of a binary vector pPLEX3003 (GenBank accession number AY159024) expressed from the constitutive subterranean clover stunt virus promoter 7 and linked to NADP malic enzyme terminator Me1 (Schünmann et al. 2003). The pPLEX3003-GhMyb25 construct was then introduced into *A. tumefaciens* AGL1 strain and used to transform *Nicotiana tabacum* L. cv. W38 leaf as described by Horsch et al. (1985) with selection on medium containing kanamycin sulfate.

For RT–PCR verification of GhMyb25 expression in T_0 and T_1 generations of transgenic tobacco, total RNA was isolated from young leaves of the transgenic tobacco lines using the Trizol method (Invitrogen, Carlsbad, CA, USA) and treated with DNase. A 2 μg aliquot of RNA was used in each RT–PCR using the Qiagen one-step RT–PCR Kit and following the manufacturer's protocol. The primers were the same as in the RT–PCR from cotton RNA resulting in a 224 bp cDNA fragment. A total number of 32 cycles at 94°C for 1 min, 50°C for 1 min and 72°C for 1 min was used for the amplification, and 5 μl of the reaction was loaded on a 2% agarose gel to visualize the RT–PCR band.

Genomic DNA was isolated from leaves of transgenic tobacco T_1 plants using half-strength CTAB buffer and a simplified version of the procedure of Paterson et al. (1993). PCRs using genomic DNA of tobacco plants were performed to verify transgene (GhMyb25) segregation in the T_1 population. The PCR contained 2 mM $MgCl_2$, 0.2 mM each of the dNTPs, 0.2 μM each of the same Myb25 forward and reverse primers as for RT–PCR, 1 U of Taq F2 DNA polymerase and 1× F2 buffer (Fisher Biotech Limited, Perth, Australia) in a 50 μl reaction. Twenty plants from each line were analysed, and the presence of the transgene resulted in a 224 bp fragment.

In situ relative DNA content measurement

Ovules of cotton line DP16 and 4A at –2, –1 and 0 dpa stages were fixed in 3 : 1 (95% ethanol : acetic acid) for 1 h at room temperature, cleared in 95% ethanol/1 mM $MgCl_2$ overnight at room temperature and rehydrated through an ethanol series to 10 mM Tris/1 mM $MgCl_2$ according to Szymanski and Marks (1998). The ovules were stained in 0.1 μg ml^{-1} aqueous propidium iodide for 30 s and then stored in 10 mM Tris/1 mM $MgCl_2$ at 4°C until observation.

The nuclear DNA content of ovule epidermal and fibre cells at the chalazal end was measured using a Leica SP2 confocal laser scanning microscope (Leica, Wetzler, Germany). At least 200 nuclei were measured from each sample which consisted of at least three ovules. Fluorescence at 600–740 nm was collected after excitation at 488 and 543 nm using a 63× NA 1.25 water-immersion lens. After optically sectioning through the ovule epidermis, the mean fluorescence intensity and dimensions of epidermal and fibre cell nuclei were measured from the maximum projection of the optical stack. The total fluorescence of individual nuclei was calculated by multiplying the nuclear area by the average fluorescence. This value was converted to a ratio by normalizing against total fluorescence of epidermal cell nuclei at telophase or anaphase (2C) within the same image. The normalized

values were then used to construct histograms of epidermal and fibre cell relative nuclear DNA content.

The same maximum projection images were also used to assess the epidermal cell division rates by counting the total and dividing epidermal cells per image, and division rates averaged over at least three ovules were presented.

Supplementary material

Supplementary material mentioned in the article is available to online subscribers at the journal website www.pcp.oupjournals.org.

Acknowledgments

This work was partially funded by the Cotton Research and Development Corporation of Australia and Cotton Seed Distributors, Wee Waa, NSW, Australia. The authors would like to thank Todd Collins and Merran Hunter for excellent technical assistance; Dr. Iain Wilson for help in printing our microarray slides and many helpful suggestions on the microarray experiments; Gavin Kennedy (CSIRO Maths and Information Sciences) for curation of the microarray data and maintenance of the GENA microarray database; Celia Miller for SEM work; Professor Xiao-Ya Chen for providing the *fl* and Xu-142 cotton lines; Dr. Alistair Low for information on the Lintless 1A, 4A, 5B and 53 mutant lines; Dr Sharon Orford for the pFS14× clone of the cotton α-expansin 1; and Drs Qing Liu, Yong-ling Ruan, Catriona Dowd, Lexie Press and Belinda Townsend for supplying specific cDNA clones present on the ovule cDNA microarray.

References

Arpat, A.B., Waugh, M., Sullivan, P.J., Gonzales, M., Frisch, D., Main, D., Wood, T., Leslie, A., Wing, R.A. and Wilkins, T.A. (2004) Functional genomics of cell elongation in developing cotton fibers. *Plant Mol. Biol.* 54: 911–929.

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

Berlin, J.D. (1986) The outer epidermis of the cotton seeds. *In* Cotton Physiology. Edited by Mauney, J.R. and Stewart, J.Mc.D. pp 375–414. The Cotton Foundation, Memphis, Tennessee.

Craig, S. and Beaton, C.D. (1996) A simple cryo-SEM method for delicate plant tissues. *J. Microsc.* 182: 102–105.

Dowd, C., Wilson, I. and McFadden, H. (2004) Different gene expression responses in cotton root and hypocotyl tissues during infection with Fusarium wilt disease. *Mol. Plant Microbe Interact.* 17: 654–667.

Du, X.M., Pan, J.J., Wang, R.H., Zhang, T., Zh. and Shi, Y.Zh. (2001) Genetic analysis of presence and absence of lint and fuzz in cotton. *Plant Breeding* 120: 519–522.

Elmore, C.D. and Leffler, H.R. (1976) Development of cotton fruit. III. Amino acid accumulation in protein and non-protein nitrogen fractions of cottonseed. *Crop Sci.* 16: 867–871.

Glover, B.J., Perez-Rodriguez, M. and Martin, C. (1998) Development of several epidermal cell types can be specified by the same MYB-related plant transcription factor. *Development* 125: 3497–3508.

Hasenfratz, M.P., Tsou, C.L. and Wilkins, T.A. (1995) Expression of two related vacuolar H^+-ATPase 16-kilodalton proteolipid genes is differentially regulated in a tissue-specific manner. *Plant Physiol.* 108: 1395–1404.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichoholtz, D., Rogers, S.G. and Fraley, R.T. (1985) A simple and general method for transferring genes into plants. *Science* 227: 1229–1231.

Hsu, C.-Y., Jenkins, J.N., Saha, S. and Ma, D.-P. (2005) Transcriptional regulation of the lipid transfer protein gene *LTP3* in cotton fibers by a novel MYB protein. *Plant Sci.* 168: 167–181.

Hülskamp, M., Misera, S. and Jürgens, G. (1994) Genetic dissection of trichome cell development in *Arabidopsis*. *Cell* 76: 555–566.

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

Ji, S.J., Lu, Y.C., Feng J.X., Wei, G., Li, J., Shi, Y.H., Fu, Q., Liu, D., Luo, J.C. and Zhu, Y.X. (2003) Isolation and analysis of genes preferentially expressed during early cotton fibre development by subtractive PCR and cDNA array. *Nucleic Acids Res.* 31: 2534–2543.

John, M.E. (1995) Characterization of a cotton (*Gossypium hirsutum* L.) fibre mRNA (Fb-B6). *Plant Physiol.* 107: 1477–1478.

John, M.E. (1996) Structural characterization of genes corresponding to cotton fibre mRNA, E6: reduced E6 protein in transgenic plants by antisense gene. *Plant Mol. Biol.* 30: 297–306.

Joubès, J. and Chevalier, C. (2000) Endoreduplication in higher plants. *Plant Mol. Biol.* 43: 735–745.

Kim, H.-J. and Triplett, B.A. (2004) Cotton fiber germin-like proteins. I. Molecular cloning and gene expression. *Planta* 218: 516–524.

Klok, E.J., Wilson, I.W., Wilson, D., Chapman, S.C., Ewing, R.M., Somerville, S.C., Peacock, W.J., Dolferus, R. and Dennis, E.S. (2002) Expression profile analysis of the low-oxygen response in Arabidopsis root cultures. *Plant Cell* 14: 2481–2494.

Li, C.H., Zhu, Y.Q., Meng, Y.L., Wang, J.W., Xu, K.X., Zhang, T.Z. and Chen, X.Y. (2002) Isolation of genes preferentially expressed in cotton fibres by cDNA filter array and RT–PCR. *Plant Sci.* 163: 1113–1120.

Li, X.B., Cai, L., Cheng, N.H. and Liu, J.W. (2002) Molecular characterization of the cotton *GhTUB1* gene that is preferentially expressed in fibre. *Plant Physiol.* 130: 666–674.

Loguerico, L.L., Zhang, Q.C. and Wilkins, T.A. (1999) Differential regulation of six novel MYB-domain genes defines two distinct expression patterns in allotetraploid cotton (*Gossypium hirsutum* L.). *Mol. Gen. Genet.* 261: 660–671.

Lu, P., Porat, R., Nadeau, J.A. and O'Neill, S.D. (1996) Identification of a meristem L1 layer-specific gene in Arabidopsis that is expressed during embryonic pattern formation and defines a new class of homeobox genes. *Plant Cell* 8: 2155–2168.

Ma, D.P., Liu, H.C., Tan, H., Creech, R.G., Jenkins, J.N. and Chang, Y.F. (1997) Cloning and characterization of a cotton lipid transfer protein gene specifically expressed in fibre cells. *Biochim. Biophys. Acta* 1344: 111–114.

Ma, D.P., Tan, H., Si, Y., Creech, R.G. and Jenkins, J.N. (1995) Differential expression of a lipid transfer protein gene in cotton fibre. *Biochim. Biophys. Acta* 1257: 81–84.

McFadden, H.G., Chapple, R., De Feyter, R. and Dennis, E.S. (2001) Expression of pathogenesis-related genes in cotton stems in response to infection by *Verticillium dahliae*. *Physiol. Mol. Plant Pathol.* 58: 119–131.

Nakazono, M., Qiu, F., Borsuk, L.A. and Schnable, P.S. (2003) Laser-capture microdissection, a tool for the global analysis of gene expression in specific cell types: identification of genes expressed differentially in epidermal cells or vascular tissues of maize. *Plant Cell* 15: 583–596.

Noda, K.I., Glover, B.J., Linstead, P. and Martin, C. (1994) Flower colour intensity depends on specialized cell shape controlled by a Myb-related transcription factor. *Nature* 369: 661–664.

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

Paterson, A.H., Brubaker, C.L. and Wendel, J.F. (1993) A rapid method for extraction of cotton (*Gossypium* ssp) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol. Biol. Rep.* 11: 122–127.

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: 561–682.

Pear, J.R., Kawagoe, Y., Schreckengost, W.E., Delmer, D.P. and Stalker, D.M. (1996) Higher plants contain homologs of the bacterial *celA* genes encoding the catalytic subunit of cellulose synthase. *Proc. Natl Acad. Sci. USA* 93: 12637–12642.

Rerie, W.G., Feldmann, K.A. and Marks, M.D. (1994) The *GLABRA2* gene encodes a homeodomain protein required for normal trichome development in *Arabidopsis*. *Genes Dev.* 8: 1388–1399.

Ruan, Y.L. and Chourey, P.S. (1998) A fibreless seed mutation in cotton is associated with lack of fibre cell initiation in ovule epidermis and alterations in sucrose synthase expression and carbon partitioning in developing seeds. *Plant Physiol.* 118: 399–406.

Ruan, Y.L., Llewellyn, D.J. and Furbank, R.T. (2001) The control of single-celled cotton fibre elongation by developmentally reversible gating of plasmodesmata and co-ordinated expression of sucrose and K^+ transporters and expansin. *Plant Cell* 13: 47–60.

Ryser, U. (1999) Cotton fibre initiation and histodifferentiation. *In* Cotton Fibres. Edited by Basra, A.S. pp 1–45. Haworth Press, New York.

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

Schenk, P.M., Kazan, K., Wilson, I., Anderson, J.P., Richmond, T., Somerville, S.C. and Manners, J.M. (2000) Coordinated plant defence responses in *Arabidopsis* revealed by microarray analysis. *Proc. Natl Acad. Sci. USA* 97: 11655–11660.

Schnittger, A., Schöbinger, U., Bouyer, D., Weinl, C., Stierhof, Y.-D. and Hülskamp, M. (2002) Ectopic D-type cyclin expression induces not only DNA replication but also cell division in *Arabidopsis* trichomes. *Proc. Natl Acad. Sci. USA* 99: 6410–6415.

Schünmann, P.H.D., Llewellyn, D.J., Surin, B., Boevink, P., De Feyter, R.C. and Waterhouse, P.M. (2003) A suite of novel promoters and terminators for plant biotechnology. *Funct. Plant Biol.* 30: 443–452.

Shimisu, 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.

Smart, L.B., Vojdani, F., Maeshima, M. and Wilkins, T.A. (1998) Genes involved in osmoregulation during turgor-driven cell expansion of developing cotton fibres are differentially regulated. *Plant Physiol.* 116: 1539–1549.

Solano, R., Nieto, C., Avila, J., Canas, L., Diaz, I. and Paz-Ares, J. (1995) Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB.Ph3) from *Petunia hybrida*. *EMBO J.* 14: 1773–1784.

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

Stracke, R., Werber, M. and Weisshaar, B. (2001) The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* 4: 447–456.

Suo, J., Liang, X., Pu, L., Zhang, Y. and Xue, Y. (2003) Identification of *GhMYB109* encoding a R2R3 MYB transcription factor that expressed specifically in fibre initials and elongating fibres of cotton (*Gossypium hirsutum* L.). *Biochim. Biophys. Acta* 1630: 25–34.

Szymanski, D.B. and Marks, M.D. (1998) Glabrous 1 overexpression and Triptychon alter the cell cycle and trichome cell fate in Arabidopsis. *Plant Cell* 10: 2047–2062.

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.

Turley, R.B. and Ferguson, D.L. (1996) Changes of ovule protein during early fibre development in a normal and a fibreless line of cotton (*Gossypium hirsutum* L.). *J. Plant Physiol.* 149: 695–702.

Turley, R.B. and Kloth, R.H. (2002) Identification of a third fuzzless seed locus in Upland cotton (*Gossypium hirsutum* L.). *J. Hered.* 93: 359–364.

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

Vojdani, F., Kim, W. and Wilkins, T.A. (1997). Phosphoenolpyruvate carboxylase cDNA from developing cotton (*Gossypium hirsutum* L.) fibres (accession nos.AF0089393 and AF008940) (PGR97–135). *Plant Physiol.* 115: 315.

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 protein. *Plant Cell* 11: 1337–1350.

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

Whittaker, D.J. and Triplett, B.A. (1999) Gene specific changes in α-tubulin transcript accumulation in developing cotton fibres. *Plant Physiol.* 121: 181–188.

Wilkins, T.A. and Jernstedt, J.A. (1999) Molecular genetics of developing cotton fibres. *In* Cotton Fibres. Edited by Basra, A.S. pp. 231–269. Haworth Press, New York.

Wilson, D.L., Buckley, M.J., Helliwell, C.A. and Wilson I.W. (2003) New normalization methods for cDNA microarray data. *Bioinformatics* 19: 1325–1332.

Wu, Y., Llewellyn, D.J. and Dennis, E.S. (2002) A quick and easy method for isolating good-quality RNA from cotton (*Gossypium hirsutum* L.) tissues. *Plant Mol. Biol. Rep.* 20: 213–218.

Wu, Y., Rozenfeld, S., Defferrard, A., Ruggiero, K., Udall, J.A., HyeRan K., Llewellyn, D.J. and Dennis, E.S. (2005) Cycloheximide treatment of cotton ovules alters the abundance of specific classes of mRNAs and generates novel ESTs for microarray expression profiling. *Mol. Genet. Genomics* 274: 477–493.

Yu, X.H., Zhu, Y.Q., Lu, S., Zhang T.Z., Chen X.Y. and Xu, Z.H. (2000) A comparative analysis of a *fuzzless-lintless* mutant of *Gossypium hirsutum* L. cv. Xu-142. *Sci. China Ser. C* 43: 623–630.

Zhang, T. and Pan, J. (1991) Genetic analysis of a fuzzless-lintless mutant in *Gossypium hirsutum* L. *Jiangsu J. Agric. Sci.* 7: 13–16.

(Received August 15, 2005; Accepted November 2, 2005)