the plant journal

SEB Society for Experimental Biology

The Plant Journal (2011) **65**, 785–797 doi: 10.1111/j.1365-313X.2010.04464.x

GhMYB25-like: a key factor in early cotton fibre development

Sally-Ann Walford, Yingru Wu, Danny J. Llewellyn and Elizabeth S. Dennis*

CSIRO Plant Industry, PO Box 1600, Canberra ACT 2601, Australia

Received 9 August 2010; revised 5 December 2010; accepted 14 December 2010; published online 14 January 2011.
*For correspondence (fax +61 2 62465000; e-mail Liz.Dennis@csiro.au).

SUMMARY

MYB transcription factors have been implicated in regulation of the development of ovule epidermal cells into the elongated seed fibres of cotton. An R2R3 MYB, *GhMYB25-like*, identified from its reduced expression in a fibreless mutant of cotton (Xu142 *fl*), is here shown to play a key role in the very early stages of fibre cell differentiation. A *GhMYB25-like* promoter–GUS construct was expressed predominantly in the epidermal layers of cotton ovules before anthesis (−3 days post-anthesis, dpa), increasing in expression in 0-dpa ovules, primarily in those epidermal cells expanding into fibres, and then in elongating fibres at +3 dpa, declining thereafter. This was consistent with *GhMYB25-like* transcript abundance during fibre development. RNA interference suppression of *GhMYB25-like* resulted in cotton plants with fibreless seeds, but normal trichomes elsewhere, phenocopying the Xu142 *fl* mutant. Like Xu142 *fl* these plants had reduced expression of the fibre-expressed MYBs, *GhMYB25* and *GhMYB109*, indicating that *GhMYB25-like* is upstream from those MYBs. This hierarchy was supported by the absence of any change in transcript level of *GhMYB25-like* in *GhMYB25*- and *GhMYB109*-silenced transgenic lines. Transgenic cotton with an additional copy of the native gene had elevated expression of *GhMYB25-like* in ovules, but no obvious increase in fibre initials, suggesting that there may be other factors that interact with GhMYB25-like to differentiate epidermal cells into fibre cells.

Keywords: cotton fibre, *Gossypium hirsutum*, MYB, transcription factors, transgenic, trichomes.

INTRODUCTION

Although the fibre produced on cotton seeds is an important commodity worldwide, the genes that regulate the development of these seed trichomes have not been well characterized. Cotton seed trichomes and Arabidopsis leaf trichomes are both single-celled elongated structures of epidermal origin which have similar, but distinct, regulatory mechanisms (Serna and Martin, 2006). Arabidopsis trichome development is well characterized at the molecular level and is regulated by complexes of transcription factors that either promote or inhibit trichome initiation and growth (Hülskamp, 2004; Serna and Martin, 2006). These complexes include the R2R3 MYB, GLABROUS1 (GL1), which is a positive regulator of trichome initiation, and the single-repeat MYBs, TRIPTYCHON (TRY), CAPRICE (CPC), ENHANCER OF TRIPTYCHON AND CAPRICE 1 (ETC1), ETC2, ETC3 and TRICHOMELESS 1 (TCL1) that are negative regulators acting in a partially redundant manner to suppress trichome development in different cells and tissues. The WD40 protein (TRANSPARENT TESTA GLABRA1, TTG1) and the basic helix–loop–helix proteins (GLABRA3 (GL3) and ENHANCER OF GLABRA3 (EGL3)) are involved in the trichome initiation complex with GL1 in Arabidopsis, while the homeobox (HOX) transcription factor GLABRA2 (GL2) is a downstream target and is involved in regulating trichome morphogenesis.

Potential homologues of some of these Arabidopsis trichome R2R3 MYB genes that are expressed in developing cotton fibre cells have been identified. These include six MYBs (*GhMYB1–6*) expressed in young ovules after fibre initiation (Loguerico *et al.*, 1999) and 55 MYBs identified from developing cotton ovules prior to and after fibre initiation (−3 to +3 days post-anthesis (dpa); Suo *et al.*, 2003). While these genes show sequence similarity with Arabidopsis MYBs, particularly in the DNA-binding domain (Loguerico *et al.*, 1999; Suo *et al.*, 2003), their roles in fibre development have not been detailed, with the exception of *GhMYB109* (Pu *et al.*, 2008). This highly fibre-specific MYB is similar in sequence to the trichome regulator AtGL1 and the closely related WEREWOLF (AtWER; Suo *et al.*, 2003) that is involved in root hair development, although it appears that *GhMYB109* is not required for cotton fibre initiation but is important for fibre elongation (Pu *et al.*, 2008). Ectopic expression of *GaMYB2* (a *Gossypium arboreum* fibre-expressed MYB also similar to GL1), under the control of

the *GL1* promoter, complements an Arabidopsis *gl1* mutant (Wang *et al.*, 2004). While these results suggest that these MYBs are all involved in cotton fibre development, their exact roles remain unclear and the key transcriptional regulators of fibre initiation remain to be identified. Transcriptome analyses of fibreless mutants relative to wild-type cotton (Wilkins and Arpat, 2005; Wu *et al.*, 2006; Lee *et al.*, 2007) and gene expression profiling of expanding fibre initials relative to adjacent non-fibre epidermal cells (Wu *et al.*, 2007) have identified other transcriptional regulators and metabolic pathways potentially involved in early fibre initiation events, including two specialised MYBs expressed very early in fibre development.

GhMYB25 is one of those cotton MYBs that is related in sequence to the AmMIXTA/AmMYBML1 transcription factors that direct the formation of conical epidermal cell in petals of *Antirrhinum majus* (Martin *et al.*, 2002). *GhMYB25* is expressed in fibre initials during the time of fibre initiation and elongation (Wu *et al.*, 2006), as well as in trichomes of leaves, stems and petals (Machado *et al.*, 2009). A role in the timing of fibre elongation has been suggested as RNA interference (RNAi)-mediated suppression of *GhMYB25* leads to only a slight reduction in the number of fibre initials, but much shorter cotton fibres in addition to significant reductions in trichomes on other parts of the plant (Machado *et al.*, 2009). The second MYB factor identified in the wild-type and fibreless mutant microarray comparisons (Wu *et al.*, 2006) encoded a protein with 69% sequence identity to GhMYB25 and about 36% to the *A. majus* MIXTA factor. It had an expression profile during fibre development similar to *GhMYB25* (Wu *et al.*, 2006) and was designated *GhMYB25-like*.

In this study we examine the role of *GhMYB25-like* in fibre development using RNAi-mediated gene silencing and transgenic cotton lines containing an extra copy of *GhMYB25-like* under the control of its own promoter. Reduction of *GhMYB25-like* transcripts resulted in fibreless seed, but had no effect on trichomes on other parts of the plants. There was a reduction in the levels of other known MYBs involved in early fibre development, indicating that *GhMYB25-like* is upstream in the pathway of transcription factors controlling fibre differentiation.

RESULTS

GhMYB25-like shows homology to MIXTA factors and is fibre specific

As expected for an allotetraploid (A^TD^T) like cotton, two different bacterial artificial chromosomes (BACs; represented by BAC241J3 and BAC207D10) hybridizing to the *GhMYB25-like* cDNA were identified from a *Gossypium hirsutum* cv. Acala Maxxa BAC library and these encoded very similar genes. The *GhMYB25-like* gene on BAC241J3, like that on BAC207D10, had two introns in identical positions in the coding region, but the second intron was considerably larger (1.49 kb compared with 324 bp) in 241J3 than in 207D10 (GenBank accessions HM134081 and HM134080, respectively). This larger intron appears to be specific to Acala Maxxa, as Xu142 and Coker 315 both have second introns similar in size to the 207D10 BAC. Southern blotting of genomic DNA from the allotetraploid cotton cultivar *G. hirsutum* cv. Coker 315 and two diploid ancestors, *Gossypium herbaceum* (A_1 genome) and *Gossypium raimondii* (D_5 genome), showed that *GhMYB25-like* is present as a single copy in each diploid and as two copies in the tetraploid species (Figure S1 in Supporting Information). The encoded proteins were 97.8% identical to one another and 69% identical to GhMYB25 (GenBank EU826465). Sequencing of *GhMYB25-like* from three possible diploid progenitors of tetraploid cotton, *G. raimondii* (D_5 genome; GenBank accession HM134083) *G. arboreum* (A_2 genome; GenBank accession HM134082) and *G. herbaceum* (A_1 genome; GenBank accession HM598082) suggests that the 207D10 gene is encoded by the D^T genome and the 241J3 gene by the A^T genome of the tetraploid, and they are hereafter referred to as *GhMYB25-like A* and *GhMYB25-like D*, respectively (Figure S1). The A-genome diploid species had a very similar gene structure and similar sized second intron to the D-genome species examined. Outside the highly conserved R2R3 MYB repeat region, the GhMYB25-like proteins had highest similarity to the cotton GhMYB25 protein and then to the *A. majus* MIXTA (AmMIXTA) and MIXTA-related transcription factor MIXTA-like 1 (AmMYBML1). They are more distantly related to AmMYBML2 and -3, Arabidopsis AtMYB16 and AtMYB106 and *Petunia hybrida* PhMYB1 (Figure S1). Phylogenetic analysis showed that GhMYB25 and GhMYB25-like are, together with the other known MIXTA-like factors, members of a unique MIXTA clade of MYB transcription factors, distinct from the MYBs known to be involved in Arabidopsis leaf trichome development such as AtGL1 and AtMYB66 (WER) and the cotton GL1-like MYBs, GhMYB109 and GaMYB2.

Using primers common to the A- and D-genome *GhMYB25-like* genes, quantitative real-time RT-PCR (qRT-PCR) detected transcripts in −2 to +10 dpa ovules. Sequence analysis of the cloned PCR products indicated that the A- and the D-genome homoeologs are present at equal levels during this period. Maximal expression occurred in ovules between −1 and +3 dpa and declined rapidly as the fibres began to elongate (Figure 1). No expression of *GhMYB25-like* was observed in vegetative tissues with the exception of the boll coat, where very low levels of transcript were present. A 1.3 kb promoter region of the *GhMYB25-like D* was cloned upstream of the β-glucuronidase (GUS) gene and transformed into Coker 315 cotton. The GUS expression patterns were examined in 10 progeny each from four independent T_1 lines using segregating wild types as

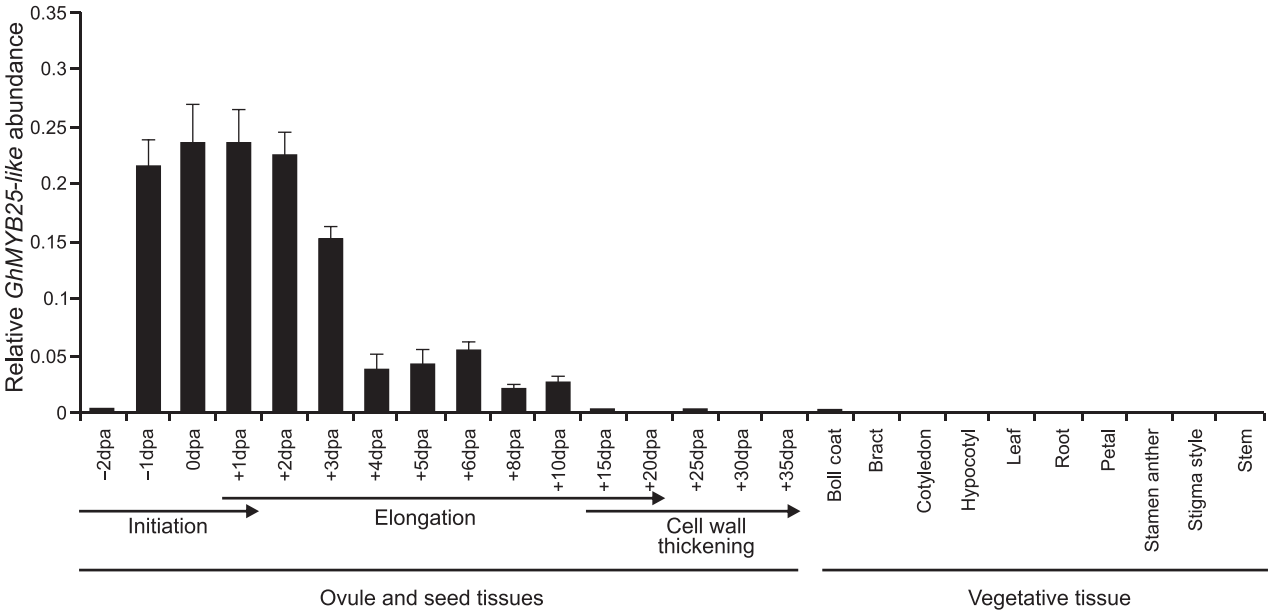

Figure 1. Quantitative real-time RT-PCR measurements of *GhMYB25-like* in cotton ovules, seeds and vegetative tissues.
Expression was normalized against the expression of the cotton ubiquitin gene. Error bars are standard deviation of six technical repeats. The various stages of fibre development are indicated with arrows. dpa, days post-anthesis.

controls. A representative example is shown in Figure 2. In each line, strong GUS activity was detected in the ovules, transmitting tissue and anthers of the flowers before anthesis (–3 dpa; Figure 2a) and in the ovules (primarily in the epidermis) and transmitting tissues of 0 dpa ovaries (Figure 2c,i,k,q), becoming restricted to fibre cells by +3 dpa ovules (Figure 2e,u). At higher magnification GUS expression was localized in the epidermal layer of 0 dpa ovules (Figure 2s) and in the epidermal and fibre cells of +3 dpa ovules (Figure 2w). The GUS activity declined between +3 and +6 dpa (Figure 2e,g) and was not detected in the elongating fibres, seed coat or embryos of older ovules (+10 dpa; Figure 2l). Although transcript abundance of the endogenous *GhMYB25-like* genes was not high in those tissues (Figure 1), GUS activity was detected in mature pollen grains (Figure 2o) and young roots (Figure 2m). The difference in expression between the reporter gene construct and direct transcript abundance could result from the absence of regulatory regions within parts of the gene not included in the promoter construct. GUS activity was not detected in the non-transgenic controls (Figure 2b,d,f,h,j,n,p,r,t,v,x).

Reduced *GhMYB25-like* expression affects only seed trichome development in transgenic cotton

A constitutive 35S promoter-driven RNAi construct targeted against both genomic copies of *GhMYB25-like* was introduced into cotton cultivar Coker 315 by *Agrobacterium tumefaciens*-mediated transformation. Seven independent transformed lines (25Li plants) were produced; most were fibreless, but otherwise indistinguishable from the wild type. Four independent single-copy 25Li lines were selected for detailed analysis and monitored through two subsequent generations to produce T_3 seed. For each line six transgene-positive and two transgene-negative T_1 plants were selfed. All 24 transgene-positive plants produced fibreless seed and the eight transgene-negative plants all had normally fibred seeds, so the fibreless phenotype co-segregated with the RNAi construct as a dominant trait, as would be expected. Detailed phenotyping and molecular analysis was performed on the T_3 plants homozygous for the presence of the transgene or wild-type segregants of the same generation for each line. The transcript levels for *GhMYB25-like* were significantly reduced in the ovules and seeds of all the silenced plants compared with their corresponding segregating wild-type plants (Figure 3). Mature cotton bolls (Figures 4a–d and S2) on T_3 homozygous 25Li plants contain essentially fibreless seeds when compared with the fibre-covered seeds in the wild-type bolls. Scanning electron microscopy of ovules of wild-type segregant plants (Figures 4 and S2) show normal differentiation and rapid emergence of fibre cells from the surface of the ovule at 0 dpa and subsequent elongation of fibres at +2 dpa (Figure 4e,g). By contrast, the surfaces of the ovules from the 25Li plants are smooth with no appearance of fibre initials (Figure 4f,h) even up to 10 dpa, and mature seeds in different transgenic lines only had low but varying numbers of a few fibres at the chalazal end (Figures 4b,d and S2) compared with wild-type seeds (Figure 4a,c). Fuzz fibres, the short hairs that initiate around 5 dpa in wild-type cotton and

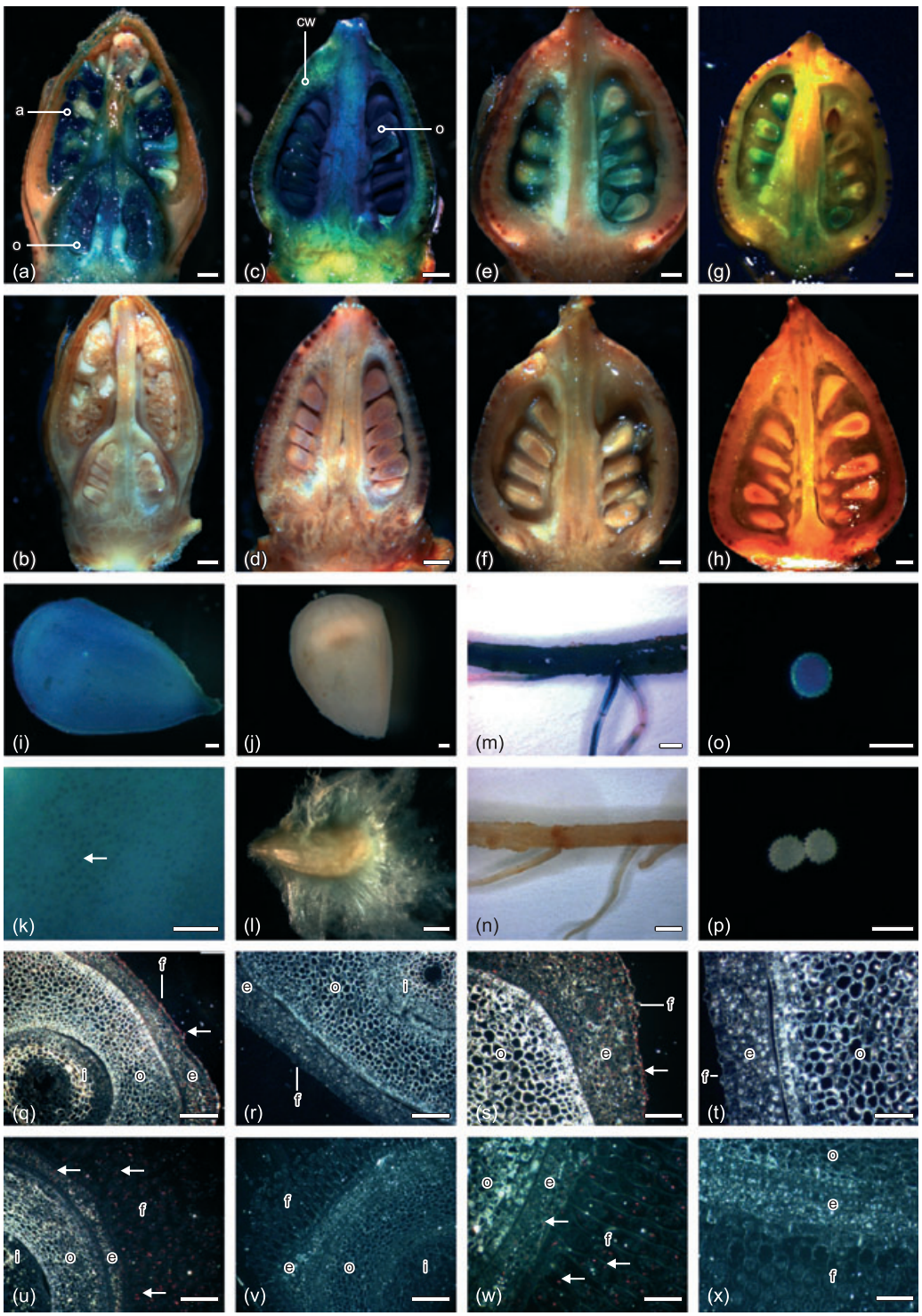

Figure 2. Histochemical localization of *GhMyb25-like*–GUS expression in mature transgenic and control cotton plants.
(a)–(h) Longitudinal sections of cotton flower bolls. cw, carpel wall; o, ovule; a, anther. GUS expression is detected at −3 days post-anthesis (dpa). (a), 0 dpa (c), +3 dpa (e) and +6 dpa (g), but not in the corresponding wild-type controls (b, d, f, h). Staining is observed on the surface of the transgenic ovule (i) compared with the non-transgenic ovule (j) at 0 dpa and higher expression is observed in fibre initials (k) as indicated by the arrow. Scale bars = 1 mm. Darkfield micrographs of 10 μm thick cross sections of 0 dpa (q, s) and non-transgenic 0 dpa ovules (r, t) and +3 dpa transgenic ovules (u, w) and non-transgenic + 3 dpa ovules (v, x). f, fibre; e, epidermis; o, outer integument of ovule; i, inner integument of ovule. Scale bars = 100 μm. Expression is localized to the epidermal layer and fibre cells in transgenic ovules (as indicated by arrows; q, s, u, w), but not in non-transgenic ovules (r, t, v, x) at +3 dpa. No GUS activity is detected in elongating fibres at 10 dpa (l). Staining in the root vasculature, root tip and root cap tissues of 10-day-old T_1 seedlings (m), but not in T_1 seedlings of non-transgenic plants (n). Expression is detected in pollen grains (o) of the flower at 0 dpa, but not in the non-transgenic plants (p).

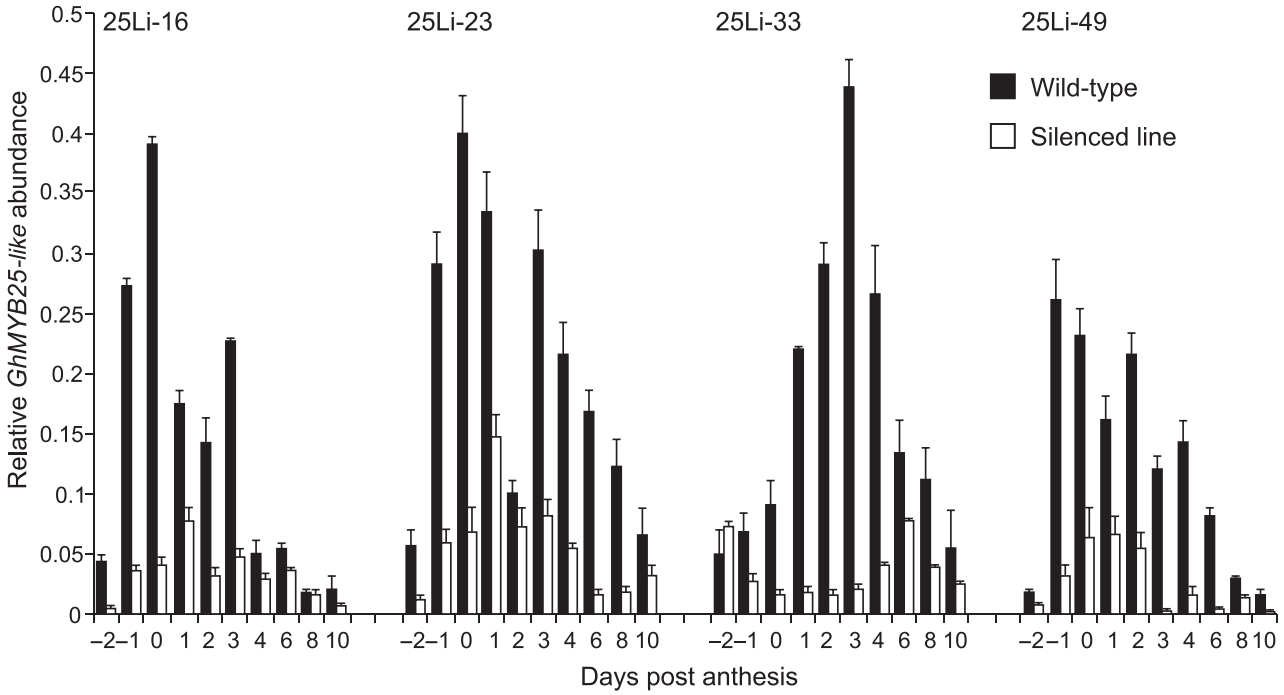

Figure 3. Quantitative real-time RT-PCR measurements of *GhMYB25-like* transcripts in ovules and developing seeds of *GhMYB25-like* RNA interference (RNAi) silenced (25Li) transgenic cotton lines and their corresponding segregating wild-type plants.
The four different transgenic lines are indicated at the top of the graph. Expression was normalized against the expression of the cotton ubiquitin gene. Error bars are standard deviation of six technical repeats.

that remain attached to the seed after removal of the long fibres during ginning, are also completely absent from all 25Li plants (Figure 4d). The trichomes on leaves, petals and petioles of the 25Li lines are unaffected (Figures 4i–n and S2) and the RNAi plants were indistinguishable from their wild-type segregant controls in respect of leaf shape, height, flowering time, flower structure and other observable morphological features (not shown).

An extra copy of the *GhMYB25-like* gene does not enhance fibre initiation

The full-length *GhMYB25-like D* driven by its own promoter was introduced into cotton and 13 T_0 lines were produced (OX25L plants). Segregating T_1 plants from independent transformants OX25L-11, OX25L-18, OX25L-59 and OX25L-78, each with a single additional copy of the *GhMyb25-like D* gene, were repeatedly selfed to produce T_3 seeds of homozygotes of both transgenic and wild types. These four lines were grown side by side and in each line a total of six transgene-positive and two negative plants were analysed for *GhMYB25-like* expression in 0 dpa ovules. All lines showed an increase in expression of *GhMYB25-like* transcripts relative to the wild type (Figure 5a). The number of fibre initials was counted from scanning electron microscope (SEM) images in the middle of 0 dpa ovules, but there was no significant difference in the number of fibre initials in the transgenic lines relative to their corresponding segregating wild-type controls (Figure 5b,c).

GhMyb25-like gene expression is reduced in Xu142 *fl*

The fibreless phenotype of Xu142 *fl* is similar to that observed in the *GhMYB25-like* silenced lines (Figure 4o,p). The transcripts encoding *GhMyb25-like* are reduced in Xu142 *fl* compared with its parental Xu142 at 0 dpa as originally shown by microarray hybridization and semi-quantitative RT-PCR (Wu *et al.*, 2006). This was confirmed by qRT-PCR, where the levels of *GhMYB25-like* are reduced, but not abolished, in the fibreless mutant compared with the parental line during early fibre development from –1 to +4 dpa (Figure 6a). Sequence analysis of the cloned qRT-PCR products showed that the A- and the D-genome homoeologs are present in comparable amounts in the parental Xu142 line at 0 dpa, but in the fibreless Xu142 *fl* line only the A genome sequence is expressed. The genomic sequences of both the A and D homoeologs of *GhMYB25-like* were amplified from Xu142 and Xu142 *fl* and no differences were detected in the promoter and 3′ untranslated regions (UTRs) of these genes. However, the A-genome copy in Xu142 *fl* is mutated at a single nucleotide in the coding region that causes a non-conservative (lysine → valine) amino acid substitution in the R2R3 MYB DNA-binding domain (Figure 6b).

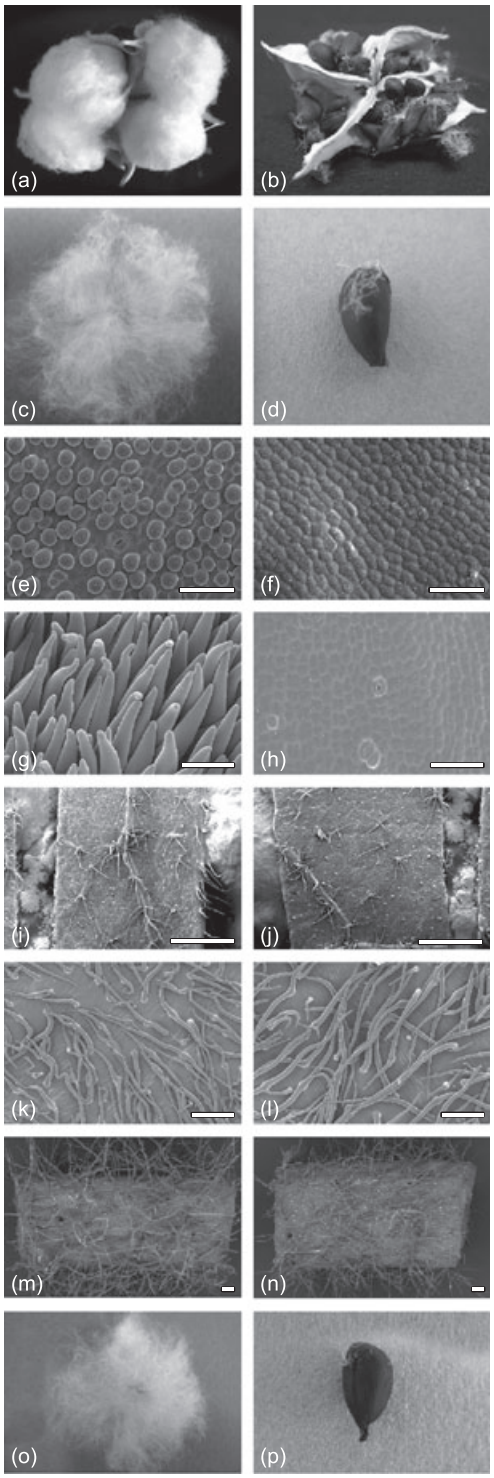

Figure 4. *GhMYB25-like* silenced lines have a fibreless phenotype but no change in other trichomes. Photographs of a mature boll and seed from wild-type control (a, c) and *GhMYB25-like* silenced line 25Li-23 (b, d). Scanning electron microscope (SEM) images of (e, f) 0 dpa ovules; (g, h) +2 dpa ovules; abaxial surfaces of leaves (i, j), petal bases (k, l) and petioles (m, n) in wild-type control (left panel) and *GhMYB25-like* silenced line 25Li-23 (right panel). The SEM images were taken at a similar position in the middle of ovules or leaves of the same age. Scale bar = 50 μm (e–l) and scale bar = 200 μm (m, n). Mature seed of (o) the cultivar Xu142 and (p) its fibreless mutant derivative Xu142 *fl.*

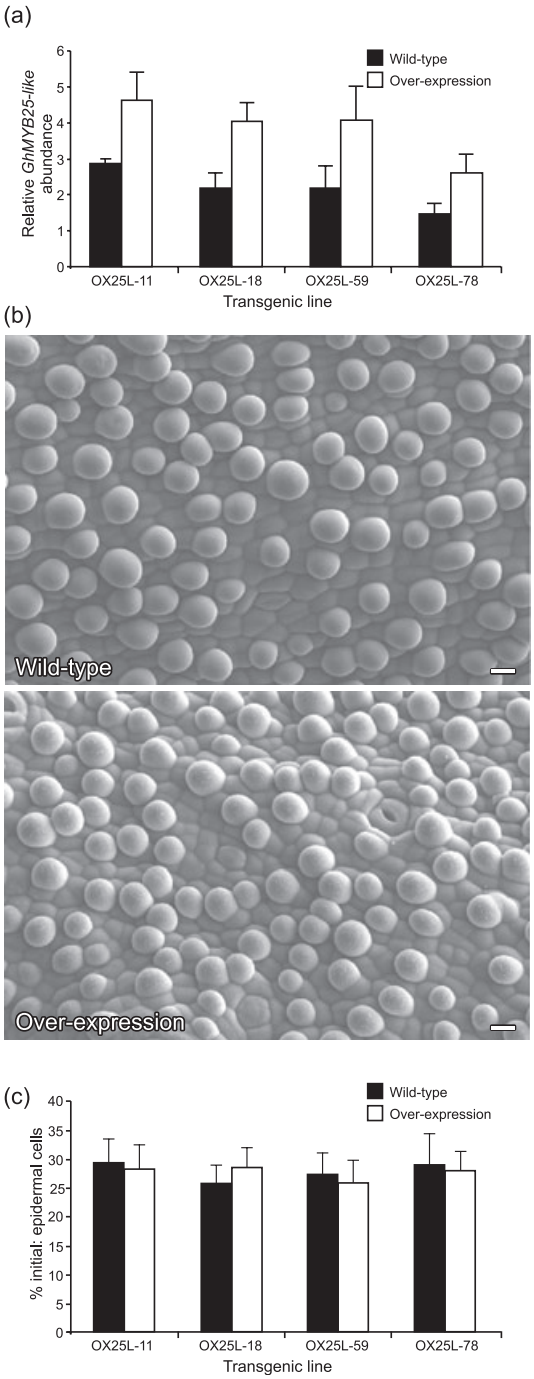

Figure 5. Over-expression of *GhMYB25-like* under its own promoter does not affect fibre initial number. (a) Quantitative real-time RT-PCR measurements of *GhMYB25-like* transcripts in 0 dpa (days post-anthesis) ovules of transgenic cotton containing an additional copy of *GhMYB25-like* under the control of its own promoter. Expression was normalized against the expression of the cotton ubiquitin gene. Error bars are standard deviation of six technical repeats.
(b) Scanning electron microscope (SEM) images of 0 dpa ovules of the wild-type control (upper panel) and over-expression line OX25L-11 (lower panel).
(c) Fibre initials were counted from SEM images of 0 dpa ovules and were expressed as the percentage of initials per thousand non-fibre epidermal cells. The values were averaged over eight selected ovules covering a total area of about 240 mm^2. Scale bar = 50 μm.

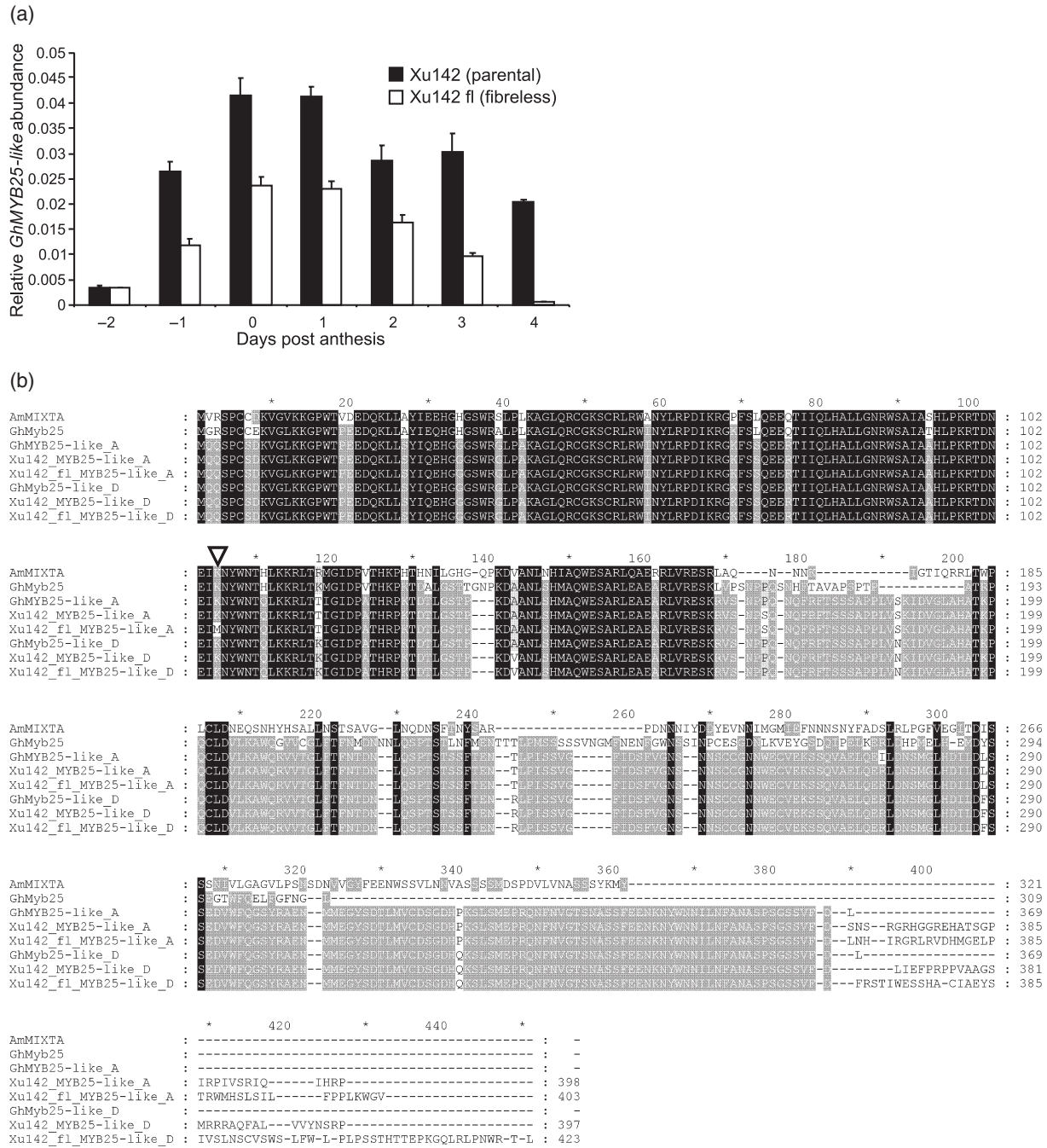

Figure 6. *GhMYB25-like* transcripts are reduced in the Xu142 *fl* mutant. (a) Quantitative real-time RT-PCR measurements of *GhMYB25-like* transcripts in ovules and developing seeds of in Xu142 (parental) and Xu142 *fl* (fibreless mutant) cotton lines. Expression was normalized against the expression of the cotton ubiquitin gene. Error bars are standard deviation of six technical repeats.
(b) Alignment of GhMYB25-like A and D protein sequences from Xu142 (Xu142 GhMYB25-like A and D) and Xu142 *fl* (Xu142 fl MYB25-like A and D) with GhMYB25 and AmMIXTA. Conserved amino acids are shaded and the mutation in the DNA-binding domain of the Xu142 *fl* MYB25-like is indicated with an arrowhead.

A regulatory pathway for cotton fibre initiation

To begin to elucidate the pathway of factors that regulate cotton fibre development the expression of the two fibre transcription factors, *GhMYB25* and *GhMYB109*, implicated in fibre expansion and elongation were quantified in the *GhMYB25-like* silenced lines (Figures 7 and S4). In wild-type cotton *GhMYB25* begins expressing very early during fibre development, peaking around 0 dpa (Machado *et al.*, 2009; Figure S3), while *GhMYB109* is expressed at high levels a little later during early fibre elongation, peaking around +3 to +4 dpa (Suo *et al.*, 2003; Figure S3). The RNAi-mediated

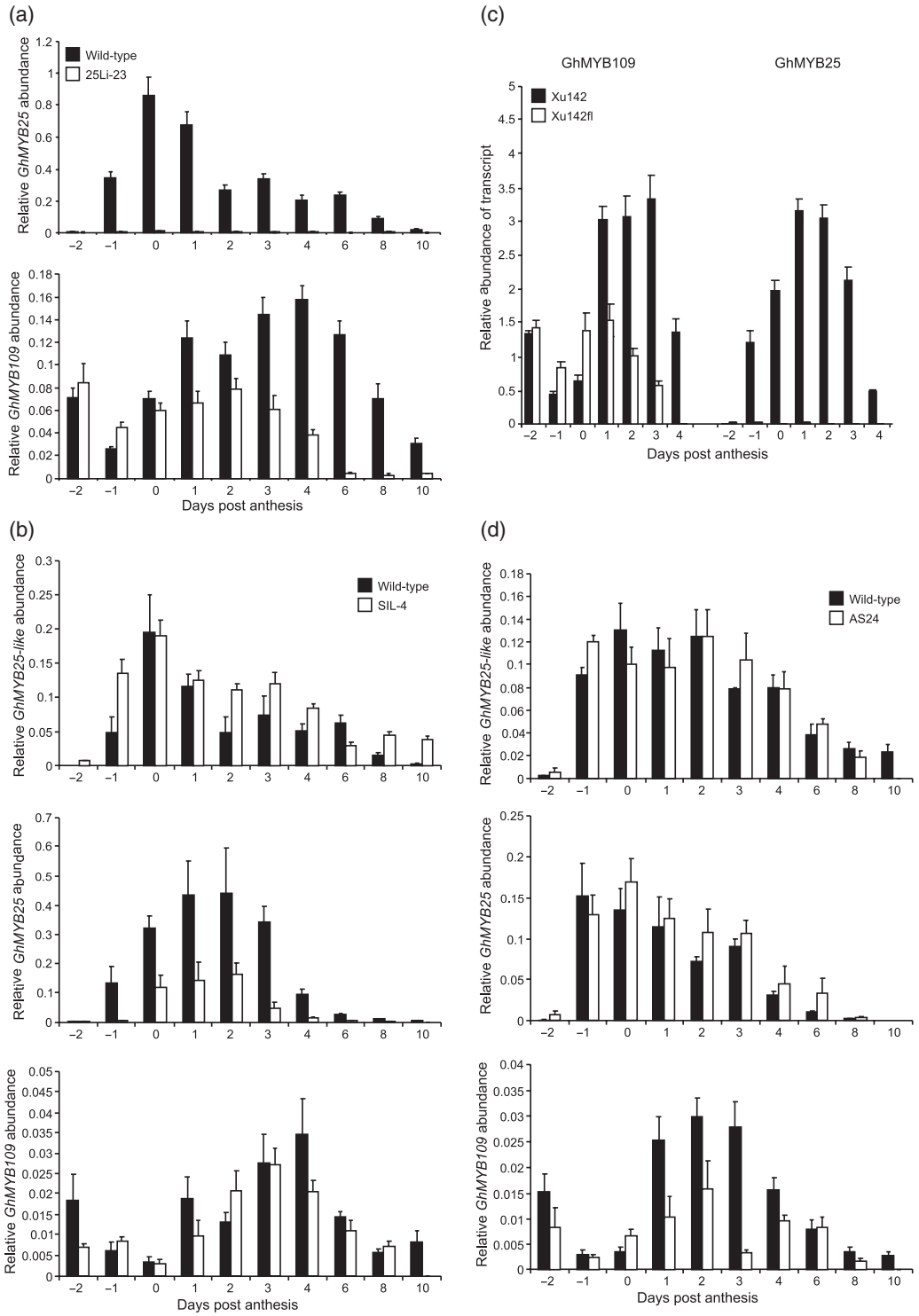

Figure 7. Quantitative real-time PCR defines a regulatory pathway for cotton fibre initiation. (a) Quantitative real-time RT-PCR measurements of *GhMYB25* and *GhMYB109* transcripts in ovules and developing seeds of the *GhMYB25-like* silenced transgenic cotton line (25Li-23).
(b) Quantitative real-time RT-PCR measurements of *GhMYB25-like*, *GhMYB25* and *GhMYB109* transcripts in ovules and seeds of the *GhMYB25* silenced cotton line (SIL-4).
(c) Quantitative real-time RT-PCR measurements of *GhMYB109* and *GhMYB25* transcripts in Xu142 and Xu142 *fl* ovules and seeds.
(d) Quantitative real-time RT-PCR measurements of *GhMYB25-like*, *GhMYB25* and *GhMYB109* in ovules and developing seeds of *GhMYB109* silenced transgenic lines (AS24 plants). Expression was normalized against the expression of the cotton ubiquitin gene.
Error bars are the standard deviation of six technical repeats.

reduction in the levels of *GhMYB25-like* transcripts almost abolished transcription of *GhMYB25* and caused a significant reduction in *GhMYB109* transcripts, particularly during fibre elongation from +3 dpa (Figures 7a and S4). There is insufficient nucleotide similarity between any of these genes for this to be the result of cross-silencing. The pattern of expression for these transcription factors in the Xu142 *fl* mutant line (Figure 7c) was the same as in the *GhMYB25-like* silenced lines. Expression of all three MYBs was also examined in the *GhMYB25*-silenced plants reported by Machado *et al.* (2009) and the *GhMYB109* silenced plants reported by Pu *et al.* (2008). *GhMYB25-like* and *GhMYB109* transcripts were essentially unchanged in the *GhMYB25*-silenced plants, although *GhMYB25* transcripts were clearly reduced by three-fold or more (Figures 7b and S5). *GhMYB25-like* and *GhMYB25* levels were also unaffected in *GhMYB109* silenced lines, although *GhMYB109* transcripts were reduced by two-fold or more (Figures 7d and S6). The increase in abundance of *GhMYB25-like* transcripts in transgenic cotton containing an additional copy of the *GhMYB25-like* gene under the control of its own promoter also resulted in a significant increase in *GhMYB25* during fibre initiation, but had no effect on *GhMYB109* expression (Figure S7).

DISCUSSION

The cotton fibre has a complex and dynamic transcriptome, with 90% or more of the genome active at the various stages from initiation through to final fibre cell wall thickening and fibre death (Hovav *et al.*, 2008; Al-Ghazi *et al.*, 2009). Orchestration of these transcriptional changes is also complex and involves a large number of regulatory genes encoded by homoeologous genes in the two genomes comprising allotetraploid cotton. A number of transcription factors, related to MYBs known to be involved in the regulation of epidermal cell patterning, have been identified in cotton at the onset of cell expansion and elongation (Loguerico *et al.*, 1999; Suo *et al.*, 2003; Wilkins and Arpat, 2005; Lee *et al.*, 2006, 2007; Wu *et al.*, 2006). It is, however, difficult to assign a function to these transcription factors as cotton has a large number of MYBs, many of which are not specific to fibres (Cedroni *et al.*, 2003). The difficulties associated with the production of transgenic plants in cotton have hindered the validation of the functional importance in fibre development of all but a couple of these genes. Here we have shown that an R2R3 MYB transcription factor, *GhMYB25-like*, has quite restricted tissue expression and has an essential role in cotton fibre initiation.

GhMYB25-like belongs to a distinct class of MYBs involved in cotton seed fibre differentiation

GhMYB25-like was originally identified as a gene whose transcript abundance was decreased during fibre initiation in a fibreless mutant (Wu *et al.*, 2006). The cotton genome contains two homoeologs which are very similar, although they have a single amino acid difference in the DNA-binding domain. The GhMYB25-like proteins are 69% identical to GhMYB25, a cotton MYB with a role in regulating seed fibre elongation and trichome development (Machado *et al.*, 2009). Although they share sequence and structural similarities in the R2R3 DNA-binding domain with other MYBs, these two MYBs appear to represent a unique class of MYB in cotton that is distinct in sequence from the other 80 or more cotton MYBs so far identified (see the Cotton Transcription Factor Database: http://planttfdb.cbi.pku.edu.cn/web/index.php?sp=gh). The GhMYB25-like and GhMYB25 proteins are more closely related to the petal-expressed MYBs of Arabidopsis, *Petunia* and *Antirrhinum* (Figure S1) than they are to MYBs known to regulate trichome development in Arabidopsis. The group includes *AmMIXTA1*, *AmMYBML1*, *-2* and *-3* from *A. majus* (Baumann *et al.*, 2007), *PhMYB1* from *P. hybrida* (van Houwelingen *et al.*, 1998) and *AtMYB16*, *AtMYB106* and *AtMYB17* from *Arabidopsis thaliana* (Romero *et al.*, 1998). A number of these petal-specific transcription factors have been demonstrated to have specialized functions in the regulation of cell morphogenesis in the petal epidermis (Martin *et al.*, 2002; Perez-Rodriguez *et al.*, 2005; Baumann *et al.*, 2007; Jaffe *et al.*, 2007), although *AtMYB106*, recently identified as the *NOECK* gene, while expressed in petals appears to also regulate Arabidopsis trichome branching (Jakoby *et al.*, 2008), so they can have multiple functions as appears to be the case for *GhMYB25* that is involved in both trichome and fibre development.

Even though GhMYB25-like shares sequence similarity with the petal MIXTA MYBs (Stracke *et al.*, 2001) its expression is restricted to the cotton ovule epidermal cells and young fibre cells and it is not expressed in other tissues that develop trichomes, including petals and leaves (Figure 1). While cotton may contain direct homologues of MIXTA involved in petal development, in the absence of a genome sequence they remain to be identified. *GhMYB25-like,* therefore, has most likely evolved independent ovule-specific functions to direct the expansion of epidermal cells to produce the fibres of cotton.

GhMYB25-like is a key regulator in cotton fibre initiation

RNAi-mediated silencing of *GhMYB25-like* transcripts abolishes expansion and elongation of ovule epidermal cells, resulting in a completely fibreless seed (except for a few remnant fibres at the chalazal end of the seed) with no impact on the trichomes, on leaves or petals, or any other obvious morphological phenotype. This is consistent with the pattern of expression of *GhMYB25-like* observed in transcriptional and reporter studies that show it is most highly expressed in early fibre development only in the epidermal and young fibre cells of the ovules (Figures 1

and 2). Broader expression of the promoter–reporter GUS construct may be due to a lack of regulatory sequences in this construct. It is most highly expressed at 0 dpa, but an increase in expression is already seen before fibre initiation (radial epidermal cell expansion) and expression remains high for the first few days of fibre development (up to +3 dpa) suggesting that GhMYB25-like may act as a signal for the expansion and elongation of epidermal cells on the surface of the ovule.

Genetic studies on the Xu142 *fl* mutant indicated that there may be at least two recessive loci (*li3* and *li4*), or even more, conferring the fibreless trait in this mutant (Du *et al.*, 2001). A point mutation which causes a non-conservative amino acid substitution is present within the highly conserved R2R3 MYB domain of GhMYB25-like A. Other MYBs that differ at this point from the consensus have little or no DNA-binding activity (Serna and Martin, 2006); however, we cannot be certain if this is a loss of function mutation in this case and hence one of the recessive loci identified by Du *et al.* (2001).

GhMYB25-like appears to be the most upstream of the different fibre MYBs we have studied and is specifically involved in controlling fibre development with no apparent role in leaf or petal trichome differentiation. The fibre-expressed MYBs, *GhMYB109* and *GhMYB25,* are downstream from *GhMYB25-like.* Both *GhMYB109* and *GhMYB25* have been shown to be regulators of cotton seed fibre development (Pu *et al.*, 2008; Machado *et al.*, 2009) and both contain MYB recognition sequences within their own promoters that could be targets for GhMYB25-like binding. A reduction in the transcripts of these two genes was observed in the *GhMYB25-like* silenced plants as well as in the Xu142 *fl* mutant. *GhMYB25* was the most drastically affected, suggesting it may be a direct target for *GhMYB25-like*. An increase in *GhMyb25-like* transcript levels in transgenic cotton containing an additional copy of the *GhMYB25-like* gene under the control of its own promoter resulted in an increase in *GhMYB25* transcription during early fibre initiation, but *GhMYB109* expression was unaffected by increasing *GhMYB25-like* so any interaction must either be saturated or indirect. Transcription of *GhMYB109* peaks a few days later than *GhMYB25* which has maximal expression at −1 to 0 dpa, but reductions in *GhMBY25* have no effect on *GhMYB109*, suggesting that they lie in independent or at least partially redundant pathways. The phenotypes of *GhMYB109*- and *GhMYB25*-silenced cotton are very similar, in that they both show a delay in the timing of fibre initiation and a reduction in fibre length, although it appears that *GhMYB109*, unlike *GhMYB25*, has no role in leaf trichome development. These results indicate that *GhMYB25-like* is a key regulatory component, acting upstream of *GhMYB25* and *GhMYB109*, in the pathway specifically regulating epidermal cell differentiation to form cotton seed trichomes, although the other MYBs may also function in trichome differentiation in other parts of the plant. *GhMYB25-like* must also be an important regulator of fuzz fibre development, as the silenced lines lack both the long lint fibres normally harvested for textile production and also the short fuzz fibres that normally initiate between 5 and 10 dpa in a second wave after the lint fibres have already elongated considerably. The fuzz fibres, however, never elongate past a millimetre or so. Mutants that have normal lint fibre and that lack fuzz fibre (so-called naked seed mutants) exist in cotton (e.g. Du *et al.*, 2001), but not the converse, so fuzz fibre is epistatic to lint fibre and must share common regulators, including *GhMYB25-like*.

GhMYB25-like: a candidate to manipulate cotton fibre yield?

While the addition of an extra copy of *GhMYB25-like* under the control of its own promoter increased total expression of *GhMYB25-like* and enhanced the expression levels of *GhMYB25*, the number of initials on the surface of developing cotton ovules was unaffected. Machado *et al.* (2009) have reported that increasing the levels of *GhMYB25* by over-expressing it under the control of a strong, constitutive promoter resulted in a modest increase in the number of fibres on ovules, but the biggest effects appeared to be on the number of trichomes on leaves and petals. There is clearly some feedback regulation at either the transcriptional, translational or post-translational level or saturation of other factors yet to be identified that prevents a significant alteration in the number of epidermal cells becoming fibres, despite the increases in expression of these critical regulatory factors. Perhaps more targeted expression of *GhMYB25-like* under the control of a promoter lacking MYB-binding sites or the use a strong ovule-specific promoter, such as the FBP7 (Floral Binding Protein 7) promoter (Colombo *et al.*, 1997), would increase its expression sufficiently to increase the number of initials on the ovule surface and hence improve fibre yield, or it may be necessary to concomitantly increase expression of other interacting protein partners before any significant increase in fibre differentiation could be achieved.

Even though transgenic cotton lines are difficult to generate, they clearly provide a powerful means to validate the functional significance of genes involved in cotton fibre development. Silencing and over-expression lines are essential tools to unravel the complex regulatory networks that control fibre initiation and will allow the identification of potential new candidates for the manipulation of cotton fibre traits such as quality and yield.

EXPERIMENTAL PROCEDURES

Plant material

Wild-type and mutant cotton cultivars of *G. hirsutum* (Coker 315, Xuzhou 142, also called Xu142), and the fibreless line Xu142 *fl*, and accessions of *G. arboreum*, *G. herbaceum* and *G. raimondii* were

used in this study. The Xu142 *fl* mutant and its parental Xu142 were provided by Professor Xiao-Ya Chen (Institute of Plant Physiology, Chinese Academy of Science, Shanghai, China), while other lines were obtained from Dr Greg Constable (Australian Cotton Research Institute, Narrabri, Australia). The *GhMYB109*-silenced line was provided by Professor Y. Xue (Institute of Genetics and Developmental Biology, Chinese Academy of Sciences and National Centre for Plant Gene Research, Beijing, China). Vegetative and reproductive tissues were harvested from cotton plants grown in a glasshouse under conditions as described previously (Wu *et al.*, 2006). Developing ovules were excised from developing flower buds or bolls on selected days before or post-anthesis (dpa) relative to the day of anthesis (0 dpa).

GhMYB25-like silencing and over-expression constructs

The 560-bp C-terminal fragment of GhMYB25-like downstream of the conserved R2R3-binding domain was amplified using Herculase II Fusion polymerase (Stratagene, http://www.stratagene.com) from cotton 0-dpa ovule cDNA using the gene-specific primers (forward 5′-GGTTCACGTCTTCATCG-3′; reverse 5′-CAGAAGGGGAAGCATT C-3′) with either a 5′-attB1 extension (5′-GGGGACAAGTTTGTCAA AAAAAGCAGGCT-3′) or a 5′-attB2 extension (5′-GGGGACCACTT TGTACAAGAAAGCTGGGT-3′), respectively, for Gateway cloning (Invitrogen, http://www.invitrogen.com/) following the manufacturer's protocol. The amplified PCR fragment was recombined in both orientations into the pHellsgate8 silencing vector via an intermediate pDONR201 vector (Invitrogen) as described previously (Machado *et al.*, 2009). While in the absence of a genome sequence for cotton we cannot yet confirm that there are no other unintended targets of the silencing construct, when the region used was interrogated against the extensive *Gossypium* fibre expressed sequence tag (EST) collection in GenBank it only had hits to six GhMYB25-like sequences that matched one or other of our two homoeologues.

The BAC library screening with the *GhMYB25-like* cDNA sequence and BAC sequencing were as described in Machado *et al.* (2009). Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 4.0 (Tamura *et al.*, 2007). A 4.6-kb *Hin*d III subclone of the genomic sequence from BAC 207D10 (*G. hirsutum* cv. Acala Maxxa BAC library, Clemson Genomics Institute, Clemson, SC, USA) together with its own promoter was cloned into the pPLEX506 expression vector (Schünmann *et al.*, 2003) to produce an over-expression construct that was introduced into *A. tumefaciens* strain AGL1 by electroporation.

GhMyb25-like GUS reporter gene construct

The *GhMYB25-like D* promoter region was amplified using Herculase II Fusion polymerase (Stratagene, http://www.stratagene.com) from DNA of BAC207D10 using forward and reverse primers 5′-TGTTCTTTAGCGGCGTTTGTG-3′ and 5′-GTTCTTTGAAGTCAAG-CCAGGCCG-3′ with appropriate attB1 and attB2 extensions for Gateway cloning as recommended by the manufacturer. The 1.3 kb promoter was recombined into the GUS reporter vector pSir-ogateIV-GUS via the pDONR201 intermediate vector, as described above, as a transcriptional fusion with the GUS coding region and *Flaveria* malic enzyme terminator. The construct was transformed into AGL1 for transformation into cotton as described below.

Production of transgenic cotton plants

For the production of transgenic plants, cotyledons were cut from sterile Coker 315 seedlings and used as explants for transformations according to the protocol of Murray *et al.* (1999). Transgenic plantlets were grown to maturity for the collection of T_1 seeds and homozygous T_3 lines were obtained by repeated selfing as described previously (Machado *et al.*, 2009). A PCR analysis was used to confirm the presence of transgenes in transformants and progeny of over-expression and silencing lines as reported previously (Machado *et al.*, 2009).

Histochemical assay of GUS gene expression

Histochemical assays of GUS activity in T_1 transgenic cotton plants were conducted as described by Jefferson *et al.* (1987). Stained cotton ovules at the developmental stage of 0 and +3 dpa were embedded in LR White Resin. The samples were cut into 10 μm thick sections using a Leica microtome and were photographed under a Ziess DMR microscope (Zeiss, http://www.zeiss.com/) equipped with dark-field optics. All other samples were examined and photographed under a Zeiss MC80-DX microscope with attached digital camera.

PCR amplification of genomic clones and sequencing

The full-length *GhMYB25-like* genes were amplified from genomic DNA isolated as described by Patterson *et al.* (1993) from Coker 315, *G. arboreum, G. herbaceum* and *G. raimondii*. The forward and reverse primers used were 5′-GAAGGGTATAATGGTCATATAGG CTT-3′ and 5′-CAATTAATCTCAAAAGACAGAAGAAC-3′, respectively, generating a 1200-bp fragment. The PCR reaction contained 100 ng of genomic DNA, 200 μM of each of the dNTPs, 0.2 μM each of the primers, 1 unit of Phusion Taq DNA polymerase and 1× Phusion buffer (Phusion, http://www.neb.com) in a 50 μl reaction as recommended. An initial denaturation cycle at 98°C for 1 min was followed by 30 cycles at 98°C for 10 sec, 53°C for 10 sec and 72°C for 40 sec, with a final step of 72°C for 3 min. The PCR products were cloned into pGEM-T Easy and sequenced.

Southern blot hybridisation

A total of 20 μg of genomic DNA extracted from *G. hirsutum* cv. Coker 315, *G. herbaceum* and *G. raimondii* was digested with *Bgl*I and transferred to Hybond N^+ (Amersham Biosciences, http://www1.gelifesciences.com/) according to the manufacturer's recommendations. The Southern blot of the various A, D or AD genome cotton species was hybridised with the same 560-bp C-terminal coding region fragment of *GhMYB25-like* used in the RNAi construct.

Scanning electron microscopy (SEM)

Fibre and trichome development on ovules, leaves, petals and petioles was visualized by SEM. Cotton ovules were collected from multiple flowers of transgenic and non-transgenic plants at developmental stages of 0, +1, +2, +3, +4, +6, +8 and +10 dpa as described previously (Wu *et al.*, 2006). Ovules were frozen in an Oxford CT 1500 cryotrans system, gold coated and imaged using a JEOL 6400 scanning electron microscope (http://www.jeol.com) as described by Craig and Beaton (1996). Trichomes were visualized by SEM on the abaxial leaf surface of the first fully expanded leaf and on the petioles attaching those same leaves and on petals of recently opened flowers.

Quantitative real-time PCR

The expression levels of *GhMYB25-like* and other fibre-related genes were analysed by quantitative real-time PCR (qRT-PCR) in an ABI7900 sequence detection system according to the manufacturer's protocol (Applied Biosystems, http://www.appliedbiosystems.com). A cotton ubiquitin gene (EU604080) was used as a reference gene for normalization. The forward and reverse primers used for ubiquitin

were 5′-CCAGAAGGAATCCACTTTGC-3′ and 5′-CCAGCTCACATCAGCATACG-3′, respectively, and the resulting amplified fragment was 132 bp long. The forward and reverse primers for *GhMYB25-like* were 5′-GAGAAATCGAGCCAAGTTGC-3′ and 5′-GATCCCCAGAATCACAAACC-3′ respectively, producing a fragment of 168 bp from both the A- and D-genome *GhMYB25-like* genes. The primers for *GhMYB25* were 5′-TCACCCACAAGCCTAAAACC-3′ and 5′-GGCAGTGAAATGGTTGCTTT-3′ producing a fragment of 155 bp and the primers for *GhMYB109*, 5′-CACGGGTATAACAGTGGATCAG-3′ and 5′-CACGGGTATAACAGTGGATCAG-3′ and 5′-ATGGTCAGGAATCCAGAAAGTG-3′ produced a 101-bp fragment. Total RNA was isolated from tissues preserved in RNAlater solution (Ambion, http://ambion.com) as described by Wu *et al.* (2002) and digested with RNase-free TURBO DNase (Ambion, http://ambion.com) according to the manufacturer's recommendations. A total of 1 μg of RNA was reverse-transcribed from an oligo$(dT)_{18}$ primer using Superscript III according to the manufacturer's recommendations (Invitrogen). The cDNA samples were diluted 10-fold and 2 μl of the dilution was used as a template in a reaction containing 5 μl SYBR Green JumpStart Taq Ready Mix (Sigma, http://www.sigmaaldrich.com), 1 μl each of the forward and reverse primers (10 μM) and 1 μl of PCR-grade water. Cycling conditions were as in Machado *et al.* (2009). Relative expression levels were determined by the ΔΔCt method.

ACKNOWLEDGEMENTS

This work was partly funded by the Australian Cotton Research and Development Corporation and the Cotton Breeding Australia Joint Venture between CSIRO and Cotton Seed Distributors. The authors would like to thank Todd Collins, Merran Hunter, Donna Frasin, Judy Radik, Jackie Oliver and Mark Talbot for their excellent technical assistance; Dr Cheng at the Australian National University Electron Microscope Unit for assistance with SEM and Professor Y. Xue for providing the *GhMYB109* silenced transgenic lines.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Southern blot, sequence alignments and phylogenetic analysis of the cotton *GhMYB25-like* genes.

Figure S2. Phenotypes of mature bolls, seeds and trichome bearing tissues of wild-type and different *GhMYB25-like* silenced transgenic cotton lines.

Figure S3. Time course of expression of cotton fibre MYBs in wild-type Coker 315 during ovule and seed development.

Figure S4. *GhMYB25* and *GhMYB109* expression in ovules and developing seeds of additional *GhMYB25-like* silenced transgenic cotton lines.

Figure S5. *GhMYB25-like*, *GhMYB25* and *GhMYB109* expression in ovules and developing seeds of additional *GhMYB25* silenced transgenic cotton lines.

Figure S6. *GhMYB25-like, GhMYB25* and *GhMYB109* expression in ovules and developing seeds of additional *GhMYB109* silenced transgenic cotton lines.

Figure S7. *GhMYB25-like*, *GhMYB25* and *GhMYB109* expression in ovules and developing seeds of GhMYB25-like over-expression transgenic cotton line OX25L-11A.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

REFERENCES

Al-Ghazi, Y., Bourot, S., Arioli, T., Dennis, E.S. and Llewellyn, D.J. (2009) Transcript profiling during fibre development identifies pathways in secondary metabolism and cell wall structure that may contribute to cotton fibre quality. *Plant Cell Physiol.* **50**, 1364–1381.

Baumann, K., Perez-Rodriguez, M., Bradley, D., Venail, J., Bailey, P., Jin, H., Koes, R., Roberts, K. and Martin, C. (2007) Control of cell and petal morphogenesis by R2R3 MYB transcription factors. *Development*, **134**, 1691–1701.

Cedroni, M.L., Cronn, R.C., Adams, K.L., Wilkins, T.A. and Wendel, J.F. (2003) Evolution and expression of MYB genes in diploid and polyploid cotton. *Plant Mol. Biol.* **51**, 313–325.

Colombo, L., Franken, J., Van der Krol, A.R., Wittich, P.E., Dons, H.J.M. and Angenent, G.C. (1997) Down-regulation of ovule-specific MADS box genes from petunia results in maternally controlled defects in seed development. *Plant Cell*, **9**, 703–715.

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

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 Breed.* **120**, 519–522.

van Houwelingen, A., Souer, E., Spelt, K., Kloos, D., Mol, J. and Koes, R. (1998) Analysis of flower pigmentation mutants generated by random transposon mutagenesis in *Petunia hybrida*. *Plant J.* **13**, 39–50.

Hovav, R., Udall, J., Hovav, E., Rapp, R., Flagel, L. and Wendel, J. (2008) A majority of cotton genes are expressed in single-celled fibre. *Planta*, **227**, 319–329.

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

Jaffe, F.W., Tattersall, A. and Glover, B.J. (2007) A truncated MYB transcription factor from *Antirrhinum majus* regulates epidermal cell outgrowth. *J. Exp. Bot.* **58**, 1515–1524.

Jakoby, M.J., Falkenhan, D., Mader, M.T., Brininstool, G., Wischnitzki, E., Platz, N., Hudson, A., Hülskamp, M., Larkin, J. and Schnittger, A. (2008) Transcriptional profiling of mature Arabidopsis trichomes reveals that *NOECK* encodes the MIXTA-like transcriptional regulator MYB106. *Plant Physiol.* **148**, 1583–1602.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) GUS fusions: ß-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901–3907.

Lee, J.J., Hassan, O.S.S., Gao, W., Wei, N.E., Kohel, R.J., Chen, X.Y., Payton, P., Sze, S.H., Stelly, D.M. and Chen, Z.J. (2006) Developmental and gene expression analysis of a cotton naked seed mutant. *Planta*, **223**, 418–432.

Lee, J.J., Woodward, A.W. and Chen, Z.J. (2007) Gene expression changes in early events in fibre development. *Ann. Bot.* **100**, 1391–1401.

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.

Machado, A., Wu, Y., Yang, Y., Llewellyn, D.J. and Dennis, E.S. (2009) The MYB transcription factor *GhMYB25* regulates early fibre and trichome development. *Plant J.* **59**, 52–62.

Martin, C., Bhatt, K., Baumann, K., Jin, H., Zachgo, S., Roberts, K., Schwarz-Sommer, Z., Glover, B. and Perez-Rodriguez, M. (2002) The mechanics of cell fate determination in petals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **357**, 809–813.

Murray, F., Llewellyn, D., McFadden, H., Last, D., Dennis, E.S. and Peacock, W.J. (1999) Expression of the *Talaromyces flavus* glucose oxidase gene in cotton and tobacco reduces fungal infection, but is also phytotoxic. *Mol. Breed.* **5**, 219–232.

Patterson, 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.

Perez-Rodriguez, M., Jaffe, F.W., Butelli, E., Glover, B.J. and Martin, C. (2005) Development of three different cell types is associated with the activity of a specific MYB transcription factor in the ventral petal of *Antirrhinum majus* flowers. *Development*, **132**, 359–370.

Pu, L., Li, Q., Fan, X., Yang, W. and Xue, Y. (2008) The R2R3 MYB transcription factor GhMYB109 is required for cotton fibre development. *Genetics*, **180**, 811–820.

Romero, I., Fuertes, A., Benito, M.J., Malpica, J.M., Leyva, A. and Paz-Ares, J. (1998) More than 80 *R2R3 MYB* regulatory genes in the genome of *Arabidopsis thaliana*. *Plant J.* **14**, 273–284.

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.

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

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.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596–1599.

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.

Wilkins, T.A. and Arpat, A.B. (2005) The cotton fibre transcriptome. *Physiol. Plant.* **124**, 295–300.

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., Machado, A., White, R.G., Llewellyn, D.J. and Dennis, E.S. (2006) Identification of early genes expressed during cotton fibre initiation using cDNA microarrays. *Plant Cell Physiol.* **47**, 107–127.

Wu, Y., Llewellyn, D.J., White, R., Ruggerio, R., Al-Ghazi, Y. and Dennis, E.S. (2007) Laser capture microdissection and cDNA microarrays used to generate gene expression profiles of the fibre initial cells on the surface of cotton ovules. *Planta*, **226**, 1475–1490.

