Control of Plant Trichome Development by a Cotton Fiber MYB Gene

Shui Wang, Jia-Wei Wang, Nan Yu, Chun-Hong Li, Bin Luo, Jin-Ying Gou, Ling-Jian Wang, and Xiao-Ya Chen
National Key Laboratory of Plant Molecular Genetics, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China

Cotton (Gossypium spp) plants produce seed trichomes (cotton fibers) that are an important commodity worldwide; however, genes controlling cotton fiber development have not been characterized. In Arabidopsis thaliana the MYB gene GLABRA1 (GL1) is a central regulator of trichome development. Here, we show that promoter of a cotton fiber gene, RD22-like1 (RD1), contains a homeodomain binding L1 box and a MYB binding motif that confer trichome-specific expression in Arabidopsis. A cotton MYB protein GaMYB2/Fiber Factor 1 transactivated the RD1 promoter both in yeast and in planta. Real-time PCR and in situ analysis showed that GaMYB2 is predominantly expressed early in developing cotton fibers. After transferring into Arabidopsis, GL1::GaMYB2 rescued trichome formation of a gl1 mutant, and interestingly, 35S::GaMYB2 induced seed-trichome production. We further demonstrate that the first intron of both GL1 and GaMYB2 plays a role in patterning trichomes: it acts as an enhancer in trichome and a repressor in nontrichome cells, generating a trichome-specific pattern of MYB gene expression. Disruption of a MYB motif conserved in intron 1 of GL1, WEREWOLF, and GaMYB2 genes affected trichome production. These results suggest that cotton and Arabidopsis use similar transcription factors for regulating trichomes and that GaMYB2 may be a key regulator of cotton fiber development.

INTRODUCTION

Aerial surfaces of most land plants have epidermal hairs (trichomes), and some plants of systematically unrelated groups, such as cotton (Gossypium spp) and Salix, produce seed trichomes, which are among the longest plant cells. Trichomes are believed to protect plants against insects, microbes, herbivores, and abiotic damages and to assist seed dispersal. Furthermore, seed trichomes of Gossypium plants, namely cotton fibers, are the most widely used natural fibers in the textile industry (Wilkins et al., 2000; Larkin et al., 2003).

Trichomes provide an excellent system for investigation of cell fate determination. In the epidermis, only some cells are committed to becoming a trichome, and the location and timing of trichome differentiation present an example of cell patterning. Once initiated, trichomes often undergo endoduplication and rapid elongation. In addition, trichome production can be genetically manipulated; alterations and even depletion of trichomes often do not affect plant viability and are readily observable (Larkin et al., 2003).

For the model plant Arabidopsis thaliana, trichome development and root epidermal patterning have been studied in depth, and both processes use a common mechanism involving closely related transcription factors and a similar lateral inhibition signaling pathway (Schneider et al., 1999; Schnitter et al., 1999; Larkin et al., 2003). Two transcription factors, GLABRA1 (GL1) and GL3, function as positive regulators of trichome development, and they encode an R2-R3 MYB protein and a basic helix-loop-helix-related protein, respectively (Oppenheimer et al., 1991; Payne et al., 2000). TRANSPARENT TESTA GLABRA1 (TTG1), a WD40 domain–containing protein that is thought to mediate protein–protein interactions, also acts as a positive regulator (Walker et al., 1999). TRIPTYCHON, CAPRICE, and ENHANCER OF TRY AND CPC have been shown to negatively regulate trichome formation. All three genes encode single-repeat MYB proteins lacking a transcriptional activation domain, and they act in a partially redundant manner to mediate lateral inhibition (Wada et al., 1997, 2002; Schnitter et al., 1999; Schellmann et al., 2002; Kerk et al., 2004). It is proposed that the combinatorial action of these patterning genes regulates trichome-specific expression of GL2, which encodes a homeobox (HOX) transcription factor that promotes trichome cell differentiation and growth (Rerie et al., 1994). Thus, GL2 may be considered a regulator of downstream genes of the trichome pathway. However, it has been unclear up to now how these upstream patterning genes, such as GL1, are regulated and whether this regulation also involves lateral inhibition.

Although both are unicellular epidermal hairs, cotton fibers are distinct from Arabidopsis trichomes in that they are produced in a specific organ (seed), their initiation and development are highly synchronous, and they are unbranched and extremely elongated (Kim and Triplett, 2001). Analyses of gene expression patterns have identified numerous genes whose transcripts are enriched in fiber cells (John and Crow, 1992; Li et al., 2002; Ji et al., 2003), including those encoding metabolism enzymes that are essential to cell growth (Pear et al., 1996; Ruan et al., 2003). For transcription factors, a noteworthy work is the isolation of six MYB genes from G. hirsutum (Loguerico et al., 2000; Larkin et al., 2003).
that changing of a region by three replacements (SCAN-7, -8, and -9) strongly reduced RDL1-P3::GUS expression (Figure 2B). Sequence analysis by PLACE (http://www.dna.affrc.go.jp/htdocs/PLACE/) showed that this region contains an L1 box (TGCAATTTA) and a MYB motif (CAGTTG), which are separated by 13 nucleotides only (Figure 2B). The L1 box is the binding site of HOX proteins, including the L1 layer-specific ATML1 (Abe et al., 2001) and the trichome regulator GL2 (Ohashi et al., 2003). The MYB motif is recognized by MYB transcription factors; Arabidopsis MYB2, for example, binds to the MYB motif of RDL2 promoter (Urao et al., 1993). The Arabidopsis GL1, which controls trichome formation, and WEREWOLF (WER), which controls development of non-hair cells of root, are both MYB transcription factors (Oppenheimer et al., 1991; Lee and Schiefelbein, 1999).

To see if these two boxes were responsible for trichome expression, we performed site-directed mutations to change the L1 box from TGCAATTTA to TGCAAGTTA and the MYB motif from CAGTTG to CAGTGG. Either mutation reduced the promoter activity greatly, though weak GUS staining in trichomes was still observable. Mutation of both cis-elements, however, completely inactivated the promoter (Figure 2C). These results indicate that both L1 box and MYB motif are involved in activating the RDL1-P3 promoter in Arabidopsis trichomes.

### Transcription Factors Binding to RDL1-P3

Because both L1 box and MYB motif contribute to the trichome expression of RDL1-P3::GUS, it is possible that the trichome factors of GL1 and GL2 transactivated RDL1-P3 in transgenic Arabidopsis, and their cotton homologs regulate RDL1 expression in cotton fiber cells.

Searching the database revealed three cotton MYBs that might be related to trichome development. GhMYB1 (Loguerico et al., 1999), also called CotMYBA (Payne et al., 1999), is 33% identical to Antirrhinum majus MIXTA, a positive regulator of floral papillae development (Noda et al., 1994). GhMYB2 (Loguerico et al., 1999) and GHMYB109 (Suo et al., 2003) are 47% identical to each other, and they show 44 and 46% sequence identities to GL1, respectively (see Supplemental Figure 2 online). To avoid confusion with Arabidopsis MYB2, which is not related to trichome development, the cotton MYB2 was also named FIF1 during our investigation. For homeobox proteins, GhHOX1 (AF5309130) and GhHOX2 (AF530914) show 66 and 34% sequence identities to GL2, respectively. We then isolated all these genes or cDNAs from G. arboreum, in addition to a cDNA encoding a novel HOX protein, GhHOX3, which is 37% identical to GL2. These cotton transcription factors were then investigated further.

In yeast one-hybrid system, the G. arboreum MYB2 and MYB109 as well as Arabidopsis GL1 and WER exhibited a clear binding activity to a RDL1 promoter fragment (−221 ~ −33) that contained both the L1 box and MYB motif, whereas another cotton MYB protein, GaMYB1, did not, though its MYB motif showed 65% sequence identity to that of GaMYB2. As mentioned above, GaMYB1 is more similar to Antirrhinum MIXTA than to GaMYB2 and Arabidopsis GL1. When the MYB motif was mutated from CAGTTG to CAGTGG, the MYB proteins were no longer able to bind the DNA (Figure 3A). All the cotton and
Arabidopsis HOX proteins tested, however, did not bind to this fragment in yeast, probably because the HOX proteins fused with GAL4 activation domain were too large to be properly expressed or the binding would require other plant proteins to form a complex.

To provide in planta evidence of RDL1-P3 regulation, related MYB and HOX genes were expressed in wild-type plants of Arabidopsis under the control of the 35S promoter, and their effects on RDL1-P3::GUS expression were examined. We found that ectopic expression of the MYBs, including GaMYB2, GL1, and WER, induced a weak expression of RDL1-P3::GUS in nontrichome cells but repressed its expression in trichomes. In plants expressing 35S::GL2 or 35S::GhHOX3, the GUS activity was clearly present in trichomes, in addition to a relatively low activity in nontrichome cells. When the MYB and the HOX genes (GL1 and GL2 or GaMYB2 and GaHOX3, respectively) driven by the 35S promoter were cotransferred into Arabidopsis, a strong GUS staining was observed in both trichome and nontrichome cells (Figure 3B), resembling the ubiquitous pattern of the 35S promoter. For each pair of transcription factors, the effects seemed to be synergistic rather than additive (Figure 3C). These results suggest that GL1-like MYB and GL2-like HOX proteins operate together to activate the RDL1-P3 promoter in Arabidopsis plants. In addition, we also transformed the Arabidopsis plants with 35S::GaMYB2/35S::HOX1 and 35S::GaMYB2/35S::HOX2. We found that, unlike 35S::GaMYB2/35S::HOX3, these two combinations did not activate the RDL1 promoter efficiently (data not shown), suggesting that there might be a specific cooperation between GaMYB2 and HOX3 in regulating the RDL1 promoter.

Figure 1. Expression Patterns of G. arboreum RDL1, MYB2/FIF1, and HOX3 Genes.

(A) Transcript profile of GaRDL1, GaMYB2, and GaHOX3 in plants of G. arboreum. Relative transcript amounts of each gene were normalized with respect to cotton Histone3 transcript levels (100%). Mean values were obtained from three independent PCR amplifications, and the error bars indicate the standard error of the mean. A break in the scale (—) has been incorporated to show the higher amount of GaRDL1 in fibers. R, roots; S, stems; L, leaves; O-0, 0-DPA ovules; F-3, 3-DPA fibers; F-6, 6-DPA fibers; F-9, 9-DPA fibers; F-12, 12-DPA fibers; NO-6, 6-DPA naked ovules (fibers stripped off).

(B) GUS staining of Arabidopsis plants expressing RDL1-P3::GUS. Young leaf (top), mature leaf (middle), and stem (bottom) are shown.
Complementation of Cotton MYB2/FIF1 to gl1 Mutant

To see if GaMYB2 is functionally similar to GL1, complementation was performed on an Arabidopsis gl1 mutant (SALK_039478), which has a T-DNA inserted in the third exon, resulting in a glabrous phenotype (Figure 4A). It is thus a loss-of-function mutation. The coding region of GaMYB2 gene was then placed between the upstream and downstream segments of the GL1 gene (GL1::GaMYB2); after transferring into gl1 plants, the transgene fully rescued trichome production in >90% of the transformants, similar to the efficiency obtained with GL1::GL1 and GL1::WER (Figure 4B, Table 1). When the three MYB genes driven by the 35S promoter were used for complementation, normal trichome production was also observed in plants of T1 generation. In the wild-type background, these MYB genes driven by either the GL1 sequences or the 35S promoter caused reduced production of leaf trichomes in 30 ~ 50% of the transgenic plants (Figure 4B, Table 1). Previous studies also showed two opposite effects of 35S::GL1 expression: suppressing trichome production in the wild-type background and promoting trichome formation in gl1 (Oppenheimer et al., 1991; Larkin et al., 1994), but such effect has not been reported before for MYB genes driven by the GL1 regulatory sequences. Reciprocal complementation using GL1 and WER found that the two genes encode functionally equivalent MYB proteins, though they specify distinct cell types because of nonoverlapping expression patterns (Lee and Schiefelbein, 2001). Our complementation data presented here indicate that the cotton protein GaMYB2 has the same function as Arabidopsis GL1 and WER in regulating trichome development.

Intron 1 Is Required for Patterning Gene Expression

The Arabidopsis GL1 and WER genes contain two introns, whereas the cotton genes of GaMYB2 and MYB109 contain only one. The cotton intron is equivalent to the first intron of Arabidopsis genes regarding their positions (131 to 146 bp downstream of the translation start codon ATG). During our complementation tests, we also used GaMYB2 cDNA (GaMYB2c). Surprisingly, neither GL1::GaMYB2c nor 35S::GaMYB2c rescued trichome formation of the gl1 mutant. In the wild-type background, the two constructs did not exhibit any effect on trichome production. We then extended this test to Arabidopsis GL1 and WER. Under the control of either GL1 sequences or the 35S promoter, the intronless cDNAs did not work (Figure 4B, Table 1). This intron requirement is therefore common to these trichome-related MYB genes, rather than specific to GaMYB2.

When the trichome-specific promoter RDL1-P3 was used to drive GaMYB2, both the intron 1–containing genomic coding region and the cDNA functioned in regulating trichome development in plants of T1 generation (Figure 4C). Similar results were obtained with GL1 (Table 1). In a previous study, ATML1::GL1c also rescued the gl1 mutant defects (Esch et al., 2003). The ATML1 promoter is expressed at a low level in all the epidermal cells of the developing leaf, and in situ analyses have shown its enhanced expression in young developing trichomes (Sessions et al., 1999). Therefore, cDNAs of the MYB genes could work normally, at least in T1, if expressed properly. These results imply that the intron is involved in regulating the gene expression pattern.

To confirm this assumption, we constructed two types of chimerical GL1::GUS reporter genes containing different introns in the coding region: one contained the GUS intron in pCAMBIA1301, and the other employed instead the GaMYB2
intron or GL1 intron 1. We found that, when the MYB intron was present, GUS activity was restricted to trichomes and the signal was clear. In plants expressing the original GUS gene, the GUS staining pattern was far from specific (Figure 5A).

Identification of an Intronic MYB Motif

PLACE analysis identified a MYB binding motif (CA/CGTTA) that is present in the GaMYB2 intron and in the first intron of GL1 and WER as well. Furthermore, position of this motif is conserved (22 or 23 bp upstream of the second exon) (Figure 5B). This intronic MYB motif is not found in MYB109 of G. arboreum. We then mutated this MYB motif into CA/CacTA. We found that the activity of promoting trichome formation in gl1 plants was considerably weakened after mutation, although the degree was different depending on the individual genes (Figure 5C, Table 1). For GL1::GL1-IN-MU, though the complementation rate was still high, >40% of transgenic plants showed abnormal trichome production: the amount of leaf trichomes was decreased and trichomes were mainly distributed on the leaf margin and stem. For GL1::GaMYB2-IN-MU, this mutation cost >50% of the transgenic plants without complementation. And for

Figure 3. Transcriptional Activation of RDL1-P3 by GaMYB2 and HOX3.

(A) DNA–protein interactions in yeast one-hybrid system. Yeast strain YM4271 integrating RDL1-OH (−221 to −33 bp) or RDL-OH-MYB-MU was transformed with GAL4 transcriptional activation domain (AD) (1), AD-GL1 (2), AD-WER (3), AD-GaMYB1 (4), and AD-GaMYB2 (5).

(B) Effects of ectopic expression of MYB (GaMYB2 or GL1) and HOX (HOX3 or GL2) genes on the RDL1-P3::GUS expression pattern in Arabidopsis. In GL1/GL2, the box indicates the magnified region shown at right. 35S-NOS, the 35S promoter and nopaline synthase (NOS) terminator cassette; GaMYB2, 35S::GaMYB2/RDL1::GUS; HOX3, 35S::HOX3/RDL1::GUS; GaMYB2/HOX3, 35S::GaMYB2/35S::HOX3/RDL1::GUS; GL1/GL2, 35S::GL1/35S::GL2/RDL1::GUS.

(C) GUS activities in Arabidopsis plants expressing both RDL1-P3::GUS and the transcription factors as indicated. MYB2, GaMYB2.
GL1::WER-IN-MU, most of the trichome-regulating activity was lost. When GL1::GaMYB109 was introduced into gl1 plants, only a partial complementation was achieved (Figure 5C, Table 1), consistent with the fact that the intronic MYB motif is not present. These results indicate that this intronic MYB motif of GL1-related MYB genes plays a role in regulating trichome production.

Expression Patterns of GaMYB2

In Arabidopsis, the cotton GaMYB2 gene showed a high activity of controlling trichome formation. If this MYB factor, which is encoded by a single gene in G. arboreum (see Supplemental Figure 1 online), controls cotton fiber development, it should have a higher level of expression in fiber cells at an early stage of ovule development. Real-time PCR showed that GaMYB2 transcripts were indeed highly enriched in developing fiber cells, and the expression level was either rather low or undetectable in all other tissues examined (Figure 1A). A similar expression pattern was observed for the HOX3 and RDL1 genes; however, at the very early stage of ovule development (0 DAP), only GaMYB2 exhibited a substantial level of expression (Figure 1A). In situ mRNA hybridization of 0-DPA ovules localized the GaMYB2 transcripts in fiber initials (Figure 6A), confirming its fiber-specific expression in ovules.

Promoting Arabidopsis Seed-Trichome Production by GaMYB2

Seeds of Arabidopsis are completely glabrous. Our extensive observation of a large number of samples under binoculars did not find any trichomes on seed coat of wild-type plants. We then examined seeds of transgenic 35S::GaMYB2 plants. Interestingly, the transgene promoted seed-trichome formation in both wild-type and gl1 backgrounds. In the wild-type background, we found that ~20% (11/54) of T1 plants showed seed-trichome phenotype, among these plants ~20% (82/416) of the seeds produced epidermal trichomes. The trichomes were often unbranched or occasionally branched, 0.1 to 0.2 mm long (Figures 6B and 6C), and the number of trichomes on each seed was small: frequently one and at most four. Other aspects of seed development appeared normal. In plants transformed with 35S::GL1, 35S::WER, and 35S::GaMYB109, seed trichomes were not observed.

DISCUSSION

Dissection of the cotton RDL1 promoter indicated that both L1 box and MYB motif contribute to its activity of trichome-specific expression in Arabidopsis, and both elements confer a positive
regulation. It is interesting that the 302-bp DNA fragment of *RDL1*-P3 is already sufficient to drive trichome-specific expression efficiently. Previously, a 334-bp promoter of a tobacco (*Nicotiana tabacum*) *P450* gene (CYP71D16) was shown to drive gene expression in distal cells of multicellular trichomes (Wang et al., 2002). These short fragments, particularly the two cis-elements identified here, can be used to design artificial promoters for engineering trichome-based molecular farming, as well as fiber trait improvement.

GL1 is the first transcription factor cloned that controls Arabidopsis trichome development (Herman and Marks, 1989; Marks and Feldmann, 1989). The prevalent model is that GL1, GL3, and probably TTG1 form a transcription complex that regulates expression of GL2, and this homeobox protein regulates downstream target genes to control epidermal cell differentiation (Szymanski et al., 1998; Ohashi et al., 2003). Besides GL2, no other target genes have been identified for GL1, although they were suspected to exist (Larkin et al., 2003). We demonstrate in this paper that GL1-like MYB proteins (GL1, WER, and GaMYB2) are able to bind to the MYB motif of the cotton *RDL1* promoter, and both the MYB and HOX proteins are involved in transcriptional regulation of this promoter. Considering that the L1 box and MYB motif are located close to each other in *RDL1*-P3, the MYB and HOX proteins may form a complex that transactivates this promoter.

This type of trichome–gene regulation is not an isolated case, for the cotton *LTP3* gene, which encodes a putative lipid transfer protein, also has L1 box and MYB motifs in its upstream sequence; in transgenic tobacco plants, the *LTP3* promoter directed trichome-specific expression of *GUS* as well (Liu et al., 2000). In cotton ovules, transcripts of both *RDL1* and *LTP3* genes are distributed mainly in developing fiber cells. It is thus highly possible that the two elements are recruited by cotton transcription factors to confer fiber cell expression of a group of cotton genes. Because both *RDL1* and *LTP3* are cotton genes, we searched the Arabidopsis genome database to find endogenous genes that have a similar promoter. We then identified several candidates, including a putative membrane bound O-acyl transferase gene (At5g55330). Promoter of this gene indeed directs trichome-specific expression, and its expression requires GL1 (our unpublished results). Based on these findings, we propose that similar molecular machinery is shared by cotton fibers and Arabidopsis trichomes in regulating a subset of downstream target genes, which involves both the GL1-like MYB and GL2-like HOX transcription factors.

Because function of a cell fate–determining gene is closely related to its expression pattern, regulatory elements are of central importance. Previous investigations of GL1 expression provided a vague picture. A promoter from the *GL1* 5’ noncoding region directed *GUS* expression in stipules only (Oppenheimer et al., 1991), whereas in situ hybridization showed a predominant distribution of the transcripts in developing trichomes and protoderm cells; the downstream sequence was then suggested to be required for normal expression (Larkin et al., 1993). In this

---

**Table 1. Regulation of Arabidopsis Trichome Development by GL1-Like Genes**

<table>
<thead>
<tr>
<th>Construct</th>
<th>Wild-Type Background</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>gl1 Background</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1 Plants Examined</td>
<td>+ (%)</td>
<td>− (%)</td>
<td>T1 Plants Examined</td>
<td>+ (%)</td>
<td>− (%)</td>
<td>+/− (%)</td>
<td>− (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL1::GaMYB2</td>
<td>47</td>
<td>57.5</td>
<td>42.5</td>
<td>54</td>
<td>92.6</td>
<td>0.0</td>
<td>7.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL1::GaMYB2c</td>
<td>46</td>
<td>100.0</td>
<td>0.0</td>
<td>52</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL1::GaMYB109</td>
<td>48</td>
<td>100.0</td>
<td>0.0</td>
<td>51</td>
<td>11.8</td>
<td>64.7</td>
<td>23.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL1::GaMYB109c</td>
<td>48</td>
<td>100.0</td>
<td>0.0</td>
<td>50</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL1::GL1</td>
<td>48</td>
<td>56.2</td>
<td>43.8</td>
<td>54</td>
<td>96.3</td>
<td>0.0</td>
<td>3.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL1::GL1c</td>
<td>46</td>
<td>100.0</td>
<td>0.0</td>
<td>54</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL1::WER</td>
<td>48</td>
<td>52.1</td>
<td>47.9</td>
<td>53</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL1::WERc</td>
<td>48</td>
<td>100.0</td>
<td>0.0</td>
<td>54</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SS::GaMYB2</td>
<td>54</td>
<td>68.5</td>
<td>31.5</td>
<td>54</td>
<td>92.6</td>
<td>0.0</td>
<td>7.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SS::GaMYB2c</td>
<td>52</td>
<td>100.0</td>
<td>0.0</td>
<td>54</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SS::GaMYB109</td>
<td>54</td>
<td>100.0</td>
<td>0.0</td>
<td>48</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SS::GaMYB109c</td>
<td>50</td>
<td>100.0</td>
<td>0.0</td>
<td>48</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SS::GL1</td>
<td>54</td>
<td>66.7</td>
<td>33.3</td>
<td>53</td>
<td>98.1</td>
<td>0.0</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SS::GL1c</td>
<td>54</td>
<td>100.0</td>
<td>0.0</td>
<td>50</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SS::WER</td>
<td>53</td>
<td>66.0</td>
<td>34.0</td>
<td>54</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SS::WERc</td>
<td>53</td>
<td>100.0</td>
<td>0.0</td>
<td>52</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDL1::GaMYB2</td>
<td>32</td>
<td>72.9</td>
<td>28.1</td>
<td>30</td>
<td>96.7</td>
<td>0.0</td>
<td>3.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDL1::GaMYB2c</td>
<td>31</td>
<td>74.2</td>
<td>25.8</td>
<td>32</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDL1::GL1</td>
<td>31</td>
<td>77.4</td>
<td>22.6</td>
<td>32</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDL1::GL1c</td>
<td>32</td>
<td>78.1</td>
<td>21.9</td>
<td>31</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL1::GaMYB2-IN-MU</td>
<td>36</td>
<td>69.4</td>
<td>30.6</td>
<td>36</td>
<td>11.1</td>
<td>33.3</td>
<td>55.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL1::GL1-IN-MU</td>
<td>36</td>
<td>63.9</td>
<td>36.1</td>
<td>36</td>
<td>58.3</td>
<td>41.7</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL1::WER-IN-MU</td>
<td>36</td>
<td>61.1</td>
<td>38.9</td>
<td>36</td>
<td>11.1</td>
<td>0.0</td>
<td>88.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

+, Trichomes normal; −, trichome amount reduced (less than one-third of wild-type trichomes); +/−, partial complementation to gl1 (less than half of wild-type trichomes); −, glabrous.
study, we show that the trichome-specific expression can be achieved as long as intron 1 is present. This intron can be considered a repressor in nontrichome cells because it prevented ectopic expression of the MYB genes in nontrichome cells. On the other hand, despite the restricted expression domain because of the presence of this intron, the overall transcript level still increased (data not shown), and the GUS signal in trichomes was comparatively strong (Figure 5A); these data suggest that the intron also acts as an enhancer in trichome cells. This cell-specific bidirectional regulation mediated by the intron identifies with the lateral inhibition model proposed for trichome formation.

That a MYB binding motif is found in the intron (intron 1) of GL1, WER, and GaMYB2 is unlikely a coincidence because this motif is required for normal trichome production. However, unlike deletion of the intron, disruption of this motif reduced, rather than eliminated, the trichome-regulating potential (Table 1). This suggests that the intron may contain other cis-elements. Alternatively, intron removal by a spliceosome may play an additional role.

Figure 5. Effects of Intron 1 and the Intronic MYB Motif on Expression of Trichome-Related Genes.
(A) GUS staining of Arabidopsis seedlings expressing GL1::GUS containing the GUS intron (top) and the first intron of GL1 (middle and bottom).
(B) Intron positions in GL1, WER, and GaMYB2 genes and the MYB motif in the first intron of each gene. MYB2, GaMYB2.
(C) Reduced production of trichomes in gl1 plants expressing GL1::GaMYB109, whose intron contains no MYB motif, or the MYB genes with disrupted intronic MYB motif. See also Table 1.

35S::GL1 but not 35S::GL1c rescued the trichome formation of the gl1 mutant. Elaboration of trichomes by 35S::GL1 or related transgenes is likely owing to the intron (intron 1) because the 35S promoter is active in practically all tissues. However, unlike GL1::GL1, 35S::GL1 could not maintain a normal trichome production in plants of T2 and subsequent generations. This indicates that the 5’ and 3’ noncoding sequences of GL1 are also involved in patterning trichome formation. Further dissection is required to clarify the regulatory role for each of these DNA segments.

Most eukaryotic genes are interrupted by one or more introns, which can influence many aspects of gene expression. In plants, a well-characterized example is Arabidopsis AGAMOUS (AG), a MADS box gene required for the development of the internal two whorls (stamens and carpels) of flower. Proper expression of AG involves sequences located in the second intron (Sieburth and Meyerowitz, 1997; Deyholos and Sieburth, 2000; Lohmann et al., 2001; Hong et al., 2003). Besides intronic cis-acting elements, presence of the first intron may enhance plant gene expression.
expression (Callis et al., 1987; Fu et al., 1995; Cai et al., 1998). It is now clear that proper expression of GL1 also requires the first intron. This intron-involved regulation of trichome formation is likely conserved in flowering plants because Arabidopsis and cotton represent two distantly related families.

Arabidopsis trichomes and cotton fibers are both unicellular hairs of epidermis, and they may share similar molecular machinery of regulation. Since the cloning of the Arabidopsis GL1 gene, scientists have been searching for MYB transcription factors that control cotton fiber initiation and development. GhMYB1 shows certain homology to MIXTA; when expressed in tobacco under the 35S promoter, it caused phenotypic changes that were global rather than specific, though supernumerary trichomes were observed. In Arabidopsis, however, its expression did not alter trichome production (Payne et al., 1999); and in our yeast one-hybrid assay, GaMYB1 failed to bind to the RDL1 promoter (Figure 3A). The cotton MYB109 shows a relatively high homology to GL1. However, GL1::MYB109 complemented gl1 mutation of Arabidopsis only partially (Table 1), and its transcripts were barely detectable in ovules at 0 DPA (Suo et al., 2003), when fiber initiation already started (Kim and Triplett, 2001). It is unlikely involved in early initiation of cotton fibers.

Three lines of evidence presented here suggest the cotton MYB2 to be a key regulator of cotton fiber development. First, it showed a clear binding activity to the cotton fiber-specific promoter RDL1 in yeast and, together with cotton HOX3, it regulated RDL1::GUS expression in Arabidopsis plants. Second, it is highly expressed in fiber cells at an early developmental stage, preceding the expression of other fiber genes, such as RDL1, HOX3 (Figure 1A), and MYB109 (Suo et al., 2003). Clearly, the spatial and temporal expression pattern of GaMYB2 fits its proposed role of controlling fiber cell formation. Third, GaMYB2 protein not only rescued trichome formation of an Arabidopsis gl1 mutant but also promoted the model plant to produce seed trichomes. Besides influencing trichome formation, GaMYB2 did not exhibit other observable effects on Arabidopsis development. Currently, we are transferring sense and antisense constructs of GaMYB2 into cotton plants to obtain more direct evidence of its role in regulating cotton fiber development.

It has been indicated that similar MYBs may play distinct roles depending on cell types in which they are expressed (Glover et al., 1998; Lee and Schiefelbein, 2001). It is therefore not totally unexpected that 35S::GaMYB2 triggered seed-trichome production in Arabidopsis because seed coat should be included in

Figure 6. In Situ Hybridization of Cotton Ovule and the Arabidopsis Seed Trichome.

(A) In situ hybridization of GaMYB2 transcripts in the G. arboreum ovule (0 DPA) with antisense (top) and sense (middle) probes. Arrowheads indicate fiber cells; the mRNA was localized in fiber cells.

(B) Trichome(s) on seed of wild-type Arabidopsis expressing 35S::GaMYB2 observed under binoculars. Arrowheads indicate the seed-trichome.

(C) Scanning electron micrographs depicting the trichome on seed of transgenic Arabidopsis expressing 35S::GaMYB2. The box indicates the magnified region shown in the bottom panels. Arrowhead indicates the seed-trichome.
the expression domain of the 3SS promoter. Previous investigations reported that Arabidopsis plants expressing 3SS::GL1 produced one to several ectopic trichomes on the cotyledon, but trichomes on seed were not mentioned (Larkin et al., 1994). In this investigation, we did not find any seed trichomes in plants expressing either 3SS::GL1 or 3SS::WER. Thus, the possibility exists that a specific structure of cotton MYB2 is related to its activity of promoting seed-trichome production. Although leaf trichomes are commonly found in flowering plants, seed-trichome formation is a sporadic event of evolution that happened in several unrelated plant families, including Malvaceae, Salicaceae, and Bombacaceae. Dissection of functional domains of cotton MYB2 and related MYB proteins holds a potential of probing into the mystery of seed-trichome evolution.

METHODS

Plant Material

Cotton plants (Gossypium arboreum cv Qingyangxiaozai and G. hirsutum cv Xuzhou-142) were grown in a greenhouse. Ovules were harvested at various developing stages; fibers were collected by scraping the ovule in liquid nitrogen. Plants of Arabidopsis thaliana (Colombia-0) were grown indoors at 22°C under continuous illumination. A T-DNA insertion mutant of gl1 (SALK_039478) was obtained from ABRC (Ohio State University, Columbus, OH). Arabidopsis tissues above the hypocotyl of the 2-week-old seedlings (four pairs of visible leaves) were collected for analysis, unless otherwise indicated.

DNA and RNA Analysis

Genomic DNA and total RNA were extracted as described (Zhu et al., 2003). For DNA gel blot analysis, hybridization and washing were performed as described (Sambrook et al., 2001). Probes were generated using the Prime-a-Gene Labeling System (Promega, Madison, WI).

Genomic walking was performed on G. arboreum DNA to isolate RDL1 5′ upstream segment according to the Genome Walker kits (Clontech, Palo Alto, CA). To obtain the 5′-terminal sequence of G. arboreum RDL1 transcript, 5′-RACE was performed using the SMART RACE cDNA amplification kit (Clontech).

The cDNAs used in this investigation were amplified by RT-PCR from the total RNAs, which were isolated from G. arboreum ovules and fiber cells, and from A. thaliana seedlings, respectively. GHBOX3 was isolated from a G. hirsutum cDNA library (Zhu et al., 2003). The genes or gene fragments were PCR amplified from the genomic DNA. Pyrobest DNA polymerase (TakaRa, Dalian, China) was used for the PCR amplification.

The mRNA was quantitatively analyzed by real-time PCR, with gene-specific TagMan primers designed using Primer Express (Applied Biosystems, Foster City, CA), and the primers were given in Supplemental Table 1 online. The 5′ and 3′ ends of these probes were labeled with fluorescent dyes FAM (6-carboxyfluorescein) and TAMRA (6-carboxytetramethyl-rhodamine), respectively. Total RNA of 1 μg from various tissues of G. arboreum was reverse transcribed in a 20-μL reaction using the RNA PCR (AMV) kit (TakaRa). TaqMan RT-PCR assays were performed on a Rotor-Gene 3000 (Corbett Research, Mortlake, Australia). In a final volume of 25 μL, the reaction contained 25 ng of reverse transcription mixture, 0.2 μM forward and reverse primers, and 0.1 μM probe. All reactions were performed in triplicate. Cotton Histone3 (AF024716) mRNA was used as an internal control, and the relative amounts of the mRNA were calculated using the comparative threshold cycle method. Amplification efficiency (82 to 90%) for the four primer-probe sets was determined by amplification of cDNA reverse transcribed from a dilution series of total RNA using 1, 0.5, 0.25, and 0.125 μg per reaction (data not shown).

DNA Mutations and Linker Scanning

Deletion derivatives were PCR constructed using sets of primers (see Supplemental Table 2 online). Linker scanning and point mutations were performed as described (Gustin and Burk, 2000). After PCR amplification, the DNA fragments were inserted into the HindIII-BamHI sites of pBl212 vector (Clontech) to replace the 3SS promoter region. Presence of the expected mutations was confirmed by sequencing.

Yeast One-Hybrid Assay

Yeast one-hybrid assay was performed using the MATCHMAKER one-hybrid system (Clontech). An RDL1 promoter region (−221 to −33 bp), RDL1-Oh, was ligated into the EcoRI-XbaI sites of pHiSi1 and HindIII-SaiI sites of pLacZi, respectively. These two bait constructs were linearized with XhoI and NcoI, respectively, and integrated into the yeast genome (strain YM4271). The level of background HIS was determined by culturing the yeast strain on the synthetic dropout (SD)/-His medium containing various concentrations of 3-amino-1, 2, 4-triazole. A yeast dual reporter strain was prepared by first integrating pLacZi-RDL1, followed by pHiSi1-RDL1, into the genome of yeast. Similarly, yeast strain YM4271 integrating RDL1-1OH-MYB-MU was formed and tested. The dual reporter strain was selected and maintained on SD/-His/-Ura medium. For construction of the GAD-MYB or GAD-HOX fusions, the cDNA (open reading frame) was ligated with GAL4 activation domain in pGAD424. Yeast transformants were tested on SD/-Leu/-His/-Ura medium containing 5 mM 3-amino-1, 2, 4-triazole.

Constructs

RDL1-P3, a 302-bp fragment of RDL1 promoter, was used to substitute the 3SS promoter region in the HindIII-NcoI sites of pCAMBIA1301 (CAMBIA, Canberra, Australia), forming pRDL1::GUS. The 3SS promoter and NOS terminator were isolated from pBl212 by PCR and inserted into the EcoRI-SacI and the Sphi-HindIII sites of pRDL1::GUS, respectively. Coding regions of MYB and HOX genes were isolated by PCR with gene-specific primers (see Supplemental Table 2 online) and placed between the 3SS promoter and NOS terminator, generating p3SS::MYB/RDL1::GUS and p3SS::HOX/RDL1::GUS. Yeast one-hybrid assay was performed using the MATCHMAKER one-hybrid system (Clontech). An RDL1 promoter region (−221 to −33 bp), RDL1-Oh, was ligated into the EcoRI-XbaI sites of pHiSi1 and HindIII-SaiI sites of pLacZi, respectively. These two bait constructs were linearized with XhoI and NcoI, respectively, and integrated into the yeast genome (strain YM4271). The level of background HIS was determined by culturing the yeast strain on the synthetic dropout (SD)/-His medium containing various concentrations of 3-amino-1, 2, 4-triazole. A yeast dual reporter strain was prepared by first integrating pLacZi-RDL1, followed by pHiSi1-RDL1, into the genome of yeast. Similarly, yeast strain YM4271 integrating RDL1-1OH-MYB-MU was formed and tested. The dual reporter strain was selected and maintained on SD/-His/-Ura medium. For construction of the GAD-MYB or GAD-HOX fusions, the cDNA (open reading frame) was ligated with GAL4 activation domain in pGAD424. Yeast transformants were tested on SD/-Leu/-His/-Ura medium containing 5 mM 3-amino-1, 2, 4-triazole.
This GL1 upstream fragment was then obtained by PCR and inserted back into BsREI, forming pgL1-2. The coding region of the GUS gene was isolated from pCAMBIA1301 by PCR and inserted into the Psal and Sau sites of pgL1-2. The intron of GUS gene was replaced by GL1 intron 1 or GaMYB2 intron, and modified GUS coding regions were inserted into the Psal and Sau sites of pgL1-2, generating pgL1::GUS-GL1-IN1 and pgL1::GUS-GaMYB2-IN, respectively.

Plant Transformation and GUS Assay
Binary constructs were introduced into Agrobacterium tumefaciens strain GV3101 and subsequently transferred into Arabidopsis using a floral dip method (Clough and Bent, 1998). Transgenic plants were selected on media containing 50 μg/mL of kanamycin or hygromycin. Histochemical and quantitative GUS assays were performed as described (Jefferson et al., 1987). Samples were stored in 70% ethanol before microscopic observation. For quantitative assays, fluorescence was measured in a minifluorometer (Hofer, San Francisco, CA).

In Situ Hybridization and Scanning Electron Microscopy
G. arboresum ovules were collected at 0 DPA and embedded in paraffin. GaMYB2 mRNA was detected with a digoxigenin-labeled riboprobe using the method as described (Long and Barton, 2000; http://www.wisc.edu/genetics/CATG/barton/protocols.html). PCR was conducted on GaMYB2 cDNA to generate a template for the probe, which was complementary to the 3’ region of GaMYB2 transcript downstream of the conserved MYB domain.

The seed trichome of transgenic Arabidopsis plants expressing 35S::GaMYB2 was examined under a JSM-6360LV scanning electron microscope (JEOL, Tokyo, Japan).

Sequence data from this article have been deposited with the EMBL/GenBank data libraries under accession numbers AY641990, AY626160, AY633559, and AY626159 for the GaRDL1, GaMYB2 (FIF1), GaMYB109, and GHOX3 cDNA clones, respectively. AY633558 is the accession number for the RDL1-P3 promoter.

ACKNOWLEDGMENTS
We thank Thea Wilkins, Yongbiao Xue, Jeffrey Chen, and Hong Ma for their help and discussions. This research was supported by the Ministry of Science and Technology of China (2002CB111301, 2001AA222121, and JY03-A-02) and by the National Natural Science Foundation of China (30270846).

Received June 2, 2004; accepted July 5, 2004.

REFERENCES


