Acta Biochim Biophys Sin 2011, 43: 607–617 | © The Author 2011. Published by ABBS Editorial Office in association with Oxford University Press on behalf of the Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. DOI: 10.1093/abbs/gmr055.
Advance Access Publication 5 July 2011

Original Article

A cotton gene encoding novel MADS-box protein is preferentially expressed in fibers and functions in cell elongation

Yang Li[1†], Hua Ning[2†], Zeting Zhang[1], Yue Wu[1], Jia Jiang[1], Siyun Su[1], Fangyun Tian[1], and Xuebao Li[1*]

[1]Hubei Key Laboratory of Genetic Regulation and Integrative Biology, College of Life Sciences, HuaZhong Normal University, Wuhan 430079, China
[2]College of Chemistry and Life Science, Hubei University of Education, Wuhan 430205, China
[†]These authors contributed equally to this work.
*Correspondence address. Tel: +86-27-67862443; Fax: +86-27-67862443; E-mail: xbli@mail.ccnu.edu.cn

Cotton fibers, as natural fibers, are widely used in the textile industry in the world. In order to find genes involved in fiber development, a cDNA (designated as *GhMADS11*) encoding a novel MADS protein with 151 amino acid residues was isolated from cotton fiber cDNA library. The deduced protein shares high similarity with Arabidopsis AP1 and AGL8 in MADS domain. However, the GhMADS11 protein (being absent of the partial K-domain and normal C-terminus) is shorter than AP1 and AGL8 by the reason of gene frameshift mutation during evolution. The experimental results revealed that GhMADS11 was not a transcriptional activator, and it did not form homodimer. *GhMADS11* transcripts were specifically accumulated in elongating fibers, but no or very low signals of its expression were detected in other tissues of cotton. Overexpression of *GhMADS11* in fission yeast promotes atypical cell elongation by 1.4–2.0-fold. Furthermore, morphological analysis indicated that the transformed cells expressing *GhMADS11m*, a MIKC-type derivative of *GhMADS11* by the site-directed mutation, displayed the same phenotype as that of the transformed cells with *GhMADS11*. The concurrence of these data sets suggested that GhMADS11 protein may function in fiber cell elongation, and its MADS domain and partial K-domain are sufficient for this function.

Keywords cotton (*Gossypium hirsutum*); MADS-box protein; gene expression; fiber development; cell elongation

Received: February 10, 2011 Accepted: April 19, 2011

Introduction

Cotton (*Gossypium hirsutum*) fibers are natural fibers that are widely used in the textile industry. As a single-cell trichome, cotton fiber provides an excellent system for studying cell elongation and cell wall biosynthesis. Fiber development is a highly regulated process which consists of four distinctive but overlapping stages: fiber initiation, cell elongation, secondary cell wall biosynthesis, and maturation. Fiber cell initiation usually starts at the day of anthesis. During this stage, ~30% of epidermal cells on the ovule surface start to enlarge and elongate rapidly. The primary wall formation (cell elongation) starts at anthesis as well, overlapping with fiber cell initiation, and lasted to 19th to 20th day after anthesis (days post anthesis, DPA). This stage can also be divided into two periods: fiber expansion stage and polar elongation stage. At about 16 DPA, the secondary cell wall formation starts, and continues up to 40 DPA. During this period, abundance of cellulose deposits in the secondary cell wall because of rapid cellulose biosynthesis. After 40 DPA, cotton fiber development enters maturation stage. At this stage, cotton fiber is associated with changes in mineral content and enzyme levels/activities. Mature fiber is made up of cellulose, water, small quantities of proteins, pectins, hemicellulose, mineral substances, wax, and small amounts of organic acids, sugars, and pigments that provide excellent wearability and esthetics [1–3].

It is believed that a number of genes are required for fiber differentiation and development. A previous study has reported that the expression of 13 genes is down-regulated in ovules of six reduced fiber or fibreless mutants at anthesis, compared with those of wild-type cotton [4]. It also showed that *GhMyb25* gene regulates outgrowths of epidermal cells including fibers, and overexpression of *GhMyb25* in tobacco leads to more branched long-stalked leaf trichomes. Further studies revealed that ectopic overexpression of *GhMyb25* results in an increase in the number of leaf trichomes and the cotton fiber initiation. In contrast, *GhMyb25*-silenced cotton plants produced short fibers, less trichomes on other parts of transgenic plants, and reduction in seed production was observed [5]. *GaMYB2* is predominantly expressed early in developing cotton fibers.

Overexpression of *GaMYB2* in Arabidopsis gives rise to more seed-trichome production. After transferring into Arabidopsis *gl1* mutant, *GL1::GaMYB2* rescues trichome formation [6]. Additional data suggested that *GhMYB2* can be regulated by a bHLH protein GhDEL65 [7]. *GhGluc1* displays its high expression level in the short fibers and weak level in the intermediate and long fibers [8]. Furthermore, the genes involved in cytoskeleton formation are required for fiber development [9,10].

MADS-box proteins family is a large family in both plant and animal [11]. Many plant MADS-box proteins are MIKC-type, including two conserved domains, MADS-box and K-box, and two variable regions, I region and C terminus [12]. MADS-box is a DNA-binding domain that binds conserved DNA sequence, $CC(A/T)_6GG$, called CArG boxes [13]. I region contributes to the specification of dimerization. K-box is characterized by three α-helices that are involved in dimerization of MADS proteins [14]. C-terminus, in some case, has been identified as a transactivation domain or contributes to the formation of multimeric MADS-box protein complexes [15,16]. Plant MADS-box genes have been mostly characterized as regulators of the transition to flowering and controlling of floral organ identities [17–19]. In the past decade, their indispensable roles in vegetative tissues have been discovered. In Arabidopsis, *AGL11* and *AGL13* are preferentially expressed in ovules, and *AGL15* is preferentially expressed in embryos, while *AGL12*, *AGL14*, and *AGL17* are all preferentially expressed in roots [20]. With the loss of *AGL12*, plant will show short roots, lower rate of cell production, and abnormal root apical meristem organization [21]. A previous study has revealed that pea *MTF1* mRNA is specifically expressed in seed coats during seed development [22]. The expression of *STMADS11*, a potato MADS-box gene, is found in all vegetative organs except for floral tissues [23]. In Arabidopsis, *AGL16* is expressed in root, mature guard cells and trichomes [24], and its expression in root is affected by N starvation [25]. *FBP20*, a Petunia *MADS* gene, is expressed in vegetable tissues. Constitutive expression of *FBP20* leads to ectopic trichome formation on adaxial sides of petals [26]. In cotton, several MADS-box genes have been identified. *GhMADS1*, *GhMADS4*, *GhMADS5*, *GhMADS6*, and *GhMADS7* are expressed at flowers, ovules, and fibers, while *GhMADS3* is only expressed in flowers [27–29]. However, little is known about the roles of MADS-box proteins in fiber development of cotton so far. Here, we reported a novel cotton MADS-box protein, GhMADS11, without entire K-box and normal C-terminus, and compared it with those known MADS proteins. *GhMADS11* gene was specifically expressed in fibers, and overexpression of *GhMADS11* in yeast promoted cell elongation.

Materials and Methods

Plant materials

Seeds of cotton (*G. hirsutum* cv. Coker312 and Xuzhou142) were surface-sterilized with 70% (v/v) ethanol for 1 min and 10% (v/v) H_2O_2 for 2 h, followed by washing with sterile water. The sterilized seeds were germinated on half-strength MS (16 h light/8 h dark cycle, 28°C) for 5–6 days. Roots, cotyledons, and hypocotyls were cut from these sterile seedlings. Other tissues (such as leaves, anthers, petals, ovules, and fibers) were derived from the cotton plants grown in fields.

Isolation of *GhMADS11* cDNA and genomic DNA

More than 4000 cDNA clones were randomly selected from cotton fiber cDNA library [30] for sequencing. One clone containing complete *GhMADS11* sequence was identified from the cDNAs.

For characterizing the *GhMADS11* gene, the partial sequence in 3'-terminus of *GhMADS11* genomic DNA was first isolated from cotton genome by polymerase chain reaction (PCR) using *GhMADS11-L1* primers: forward 5'-TGCTTCTGCCCCTTAACATC-3' and reverse 5'-AGGTTATTCAAGGTGGTGGC-3'. Then, the other fragments of *GhMADS11* genomic DNA was isolated by genome walking PCR. Cotton Genome Walker libraries were constructed using Genome Walker kit (Clontech, Mountain View, USA) as described previously [30]. Genome walking PCR was performed according to the manufacturer's instructions. Primers used for PCR-based DNA walking in Genome Walker Libraries were *GhMADS11-L2*-P1: 5'-TCAACTAGCCCAATTGATCAGT TGCCG-3', *GhMADS11-L2*-P2: 5'-ACCGATTCTCGGTA AACAGTCAGATTG-3', *GhMADS11-L3*-P1: 5'-TCTGGA TATCTTCTCCTTCGTAATGCC-3'.*GhMADS11-L3*-P2: 5'-TTGAAATTTGAAGCCATTGGGGGTACC-3', and the adapter sequence AP1 (5'- GTAATACGACTCACTATAG GGC-3') and AP2 (5'- ACTATAGGGCACGCGTGGT-3'). The PCR fragments of *GhMADS11* gene were cloned in pGEM-T vector for sequencing.

Quantitative reverse transcriptase-PCR

Total RNA was isolated from cotton tissues such as roots, hypocotyls, cotyledons, leaves, petals, anthers, ovules, and developing fibers (3–21 days post anthesis, DPA). The RNA was then purified by Qiagen RNeasy mini Kit (Qiagen, Hilden, Germany), and its concentration were determined by Nanodrop spectrophotometer and agarose gel electrophoresis.

Expression of the *GhMADS11* gene in cotton tissues was analyzed by real-time quantitative RT-PCR using the fluorescent intercalating dye SYBR-Green (TOYOBO,

Tokyo, Japan) in the detection system (Option 2, MJ Research). A cotton polyubiquitin gene (*GhUBI1*, access number in GenBank: EU604080) was used as a standard control in RT-PCR reactions. A two-step RT-PCR procedure was performed in all experiments using the method as described previously [9]. In brief, first-strand cDNAs were synthesized from cotton total RNAs using M-MLV reverse transcriptase (Promega, Madison, USA) according to the manufacturer's instruction. Then, the cDNAs were used as templates in real-time PCR with gene-specific primers (*MADS11-RT*-forward: 5'-GGACAAAGAACTGCAAGAACAG-3', *MADS11-RT*-reverse: 5'-GTTGTACTACTCTACACACATC-3'). The amplification of the target gene was monitored every cycle by SYBR-Green fluorescence. The Ct (cycle threshold), defined as the PCR cycle at which a statistically significant increase of reporter fluorescence was first detected, was used as a measure for the starting copy numbers of the target gene. Relative quantization of the target *GhMADS11* expression level was performed using the comparative Ct method. PCR conditions for primer pairs were optimized for annealing temperature and Mg^{2+} concentration, and PCR products were confirmed on an agarose gel. The efficiency of each primer pair was detected by using *GhMADS11* cDNA as standard template, and the RT-PCR data were normalized with the relative efficiency of each primer pair.

DNA and protein analysis

The *GhMADS11* DNA sequence and its deduced protein were analyzed using DNA analysis program (DNAStar software), and protein sequence homology analysis was performed with Clustal W (http://www.ebi.ac.uk/clustalw/). Sequences selected for phylogenetic analysis are cotton MADS-box proteins (GhMADS1, GhMADS3−7, and GhMADS11) and Arabidopsis MADS-box proteins (AP1, AGL8, SEP1, SEP3, AP3, PI, STK1, SHP1, and SHP2). A minimum evolution tree was generated in MEGA 3.1 program. A bootstrap analysis with 1000 replicates was performed to assess the statistical reliability of the tree topology.

Overexpression of *GhMADS11* gene in fission yeast

A 456 bp open reading frame (ORF) sequence of the *GhMADS11* gene was amplified by PCR using pfu DNA polymerase and the primer pair (*MADS11L* 5'-CTTCTCGAGATGGGAAGGGGTAGGGT-3' and *MADS11R* 5'-CTTGGATCCTCAGCTGCCAATGTGCC-3'), digested with *Xho*I/*Bam*HI, and cloned into yeast vector pREP-5N, producing pREP-5N-GhMADS11 construct.

To prepare its frameshift mutation, primer-based site-directed mutation of *GhMADS11* ORF sequence (*GhMADS11m*) was generated by PCR, using the primer pair: 5'-CTTTACAAAGAAACACTGAGAATCAACTTATG-3' and 5'-CATAAGTTGATTCTCAGTGTTTCTTTGTAAAG-3'. Then, the *GhMADS11m* (frameshift ORF sequence of *GhMADS11*) was cloned into yeast vector pREP-5N, generating pREP-5N-*GhMADS11m* construct.

The constructs were transferred into yeast cells by electroporation (Bio-Rad, Hercules, USA) according to the manufacturer's instructions. The transformants were selected on plates, containing minimal medium (MM) with 2 μM thiamine at 30°C. Ten colonies for each construct were picked out to incubate in liquid MM with 2 μM thiamine, which represses the nmt-1 promoter activity, until mid-log phase in a shaker (220 rev/min, 30°C). Subsequently, the yeast cells were harvested and washed three times with MM without thiamine to derepress the promoter, and then incubated in the same thiamine-free MM for 20 h (220 rev/ min, 30°C). The yeast cells were observed and fixed in 70% ethanol for 1 h. After washing three times in PBS, yeast cells were collected and stained with 1 μg/ml nucleus-specific fluorescent dye, DAPI. Then the specimens were photographed under a Nikon microscope (Nikon Co. Ltd., Yokohama, Japan). Cell length (of 50 cells per transformant) was measured for statistical evaluation of the cell elongation, using empty pREP-5N transformants as controls.

Trans-activation activity assay

Trans-activation activity assay was conducted as previously described [31]. The ORF sequences of *GhMADS11* and *GhMADS11m* amplified by PCR using proofreading *pfu* DNA polymerase were inserted into pBluescript sk- vector at *Sma*I and *BamH*I sites, respectively. Then, the fragments of *GhMADS11* and *GhMADS11m* digested from the pBluescript sk-vector were cloned into pGBKT7 by *Eco*RI and *Bam*HI sites to create the GAL4 DNA binding domain fusion constructs (pGBKT7-GhMADS11 and pGBKT7-GhMADS11m). The constructs were transformed into yeast stains AH109 and Y187 by LiAc-mediated transformations. The AH109 transformants were screened on SD/-Trp/-Ade medium (SD MM lacking Trp and Ade) and Y187 transformants were tested the β-galactosidase activity by using flash-freezing filter assay. The yeast cells containing empty pGBKT7 were used as negative control, and yeast cells containing pGBKT7-53 and pGADT7-RecT were used as positive controls.

Yeast two-hybrid analysis

The coding sequences of *GhMADS11* and *GhMADS11m* were cloned into pGADT7 by the method as described above. Both pGADT7-GhMADS11 and pGADT7-GhMADS11m were transformed into AH109, respectively. AH109 containing pGADT7-GhMADS11 or pGADT7-GhMADS11m construct and Y187 with pGBKT7-GhMADS11 or

pGBKT7-GhMADS11m vector were used in the mating [32]. The yeast zygotes were selected by quadruple dropout medium (QDO medium, SD/-Trp/-Leu/-His/-Ade), using transformants containing pGBKT7 and pGADT7 vectors as negative controls and transformants containing pGBKT7-53 and pGADT7-RecT vectors as positive controls.

Results

Isolation and characterization of *GhMADS11* gene

To investigate the roles of the genes in fiber development, we randomly selected 4000 cDNA clones, of which one gene (cDNA) encoding a novel protein with putative MADS domain was identified, from a cotton fiber cDNA library. This cDNA (designated as *GhMADS11*, accession number in GenBank: HM989877) is 905 bp in length, including a short 5'-untranslated region (UTR), a 456 bp coding region and a 3'-UTR. It encodes a MADS homolog with 151 amino acids (18.29 kDa, PI 10.87), including 34 basic amino acids, 13 acidic amino acids, 46 hydrophobic amino acids, and 44 polar amino acids. Sequence analysis revealed that the deduced GhMADS11 protein shares relatively high homology (56 and 55% identities, respectively) with Arabidopsis AGL8 and AP1 proteins. Using bioinformatics program blast, we found that GhMADS11 may belong to D homologue. GhMADS11 contains a MADS-box which forms DNA binding motif, and an I-region, like the AGL8 and AP1 proteins. Particularly, its MADS domain is highly conserved, displaying 90 and 93% identities with those of AP1 and AGL8, respectively. However, the deduced protein is absent of the partial K-domain and normal C-terminus, unlike AGL8 and AP1 proteins (**Fig. 1**). Thus, it seems to be a novel protein evolved from the common MADS protein, owing to losing a nucleotide acid at 110th codon of the ORF of the original gene, resulting in frameshift mutation of the gene, compared with those homologous MADS genes (cDNAs) (**Fig. 2**). To further investigate whether the frameshift mutation is due to mis-splicing of the single transcript, we isolated the genomic DNA sequence of *GhMADS11* gene. As shown in **Fig. 3**, the exon sequences of the gene are consistent with its cDNA sequence, demonstrating that the frameshift mutation site has been existed in *GhMADS11* gene during evolution, instead of mis-splicing its transcript.

Phylogenetic relationships of the MADS proteins

To determine divergence of the isolated GhMADS11 protein with other known MADS-box proteins in cotton and Arabidopsis during evolution, the phylogenetic relationship of 16 MADS-box proteins were generated by MEGA3.1. As shown in **Fig. 4**, the tree is divided into five distinct branches: AP1 subgroup, SEP subgroup, AP3/PI subgroup, AGL subgroup, and AG subgroup. GhMADS11 together with AGL8 and AP1 forms AP1 subgroup, suggesting that the GhMADS11 has a close evolutional relationship with AGL8 and AP1. Furthermore, GhMADS11 occupies a distinct branch that is basal to the AGL8 and AP1 clade, implying that GhMADS11 diverges earlier from the other cotton MADS proteins. On the other hand, GhMADS1 belongs to the SEP subgroup, GhMADS5 and GhMADS6 are located in the AGL11 subgroup, while GhMADS3, GhMADS4, and GhMADS7 are positioned in the AG subgroup. These results suggested that the divergence in these MADS proteins could occur before the differentiation of the two species, cotton (*G. hirsutum*) and *Arabidopsis thaliana*.

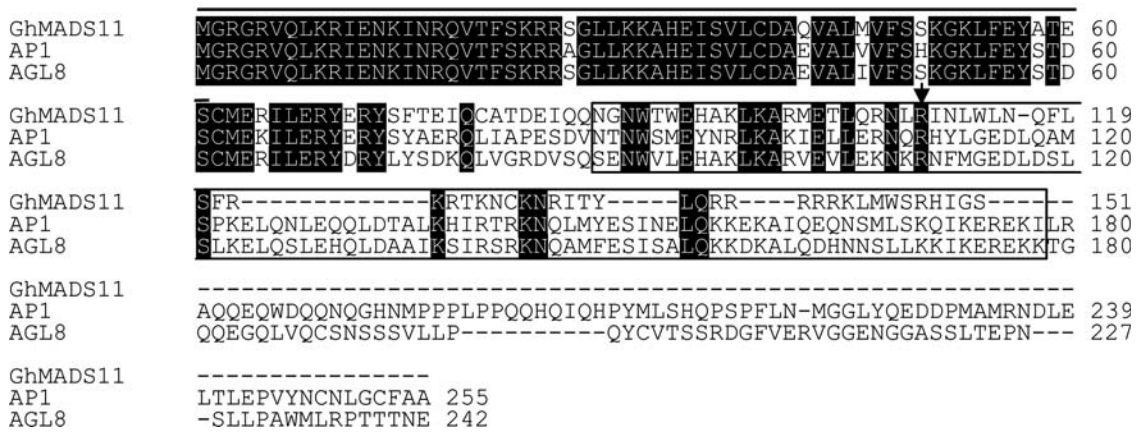

Figure 1 Alignment and sequence comparison among cotton MADS11 and Arabidopsis AP1 and AGL8 proteins Numbers on the right indicate the protein length in amimo acids. MADS-box domain is indicated by the line above the alignments. The K-box domain is indicated by white frame. The amino acid residues identical among the sequences are indicated in black. GhMADS11 in this work, and Arabidopsis AP1 (CAA78909) and AGL8 (AAA97403) from GenBank.

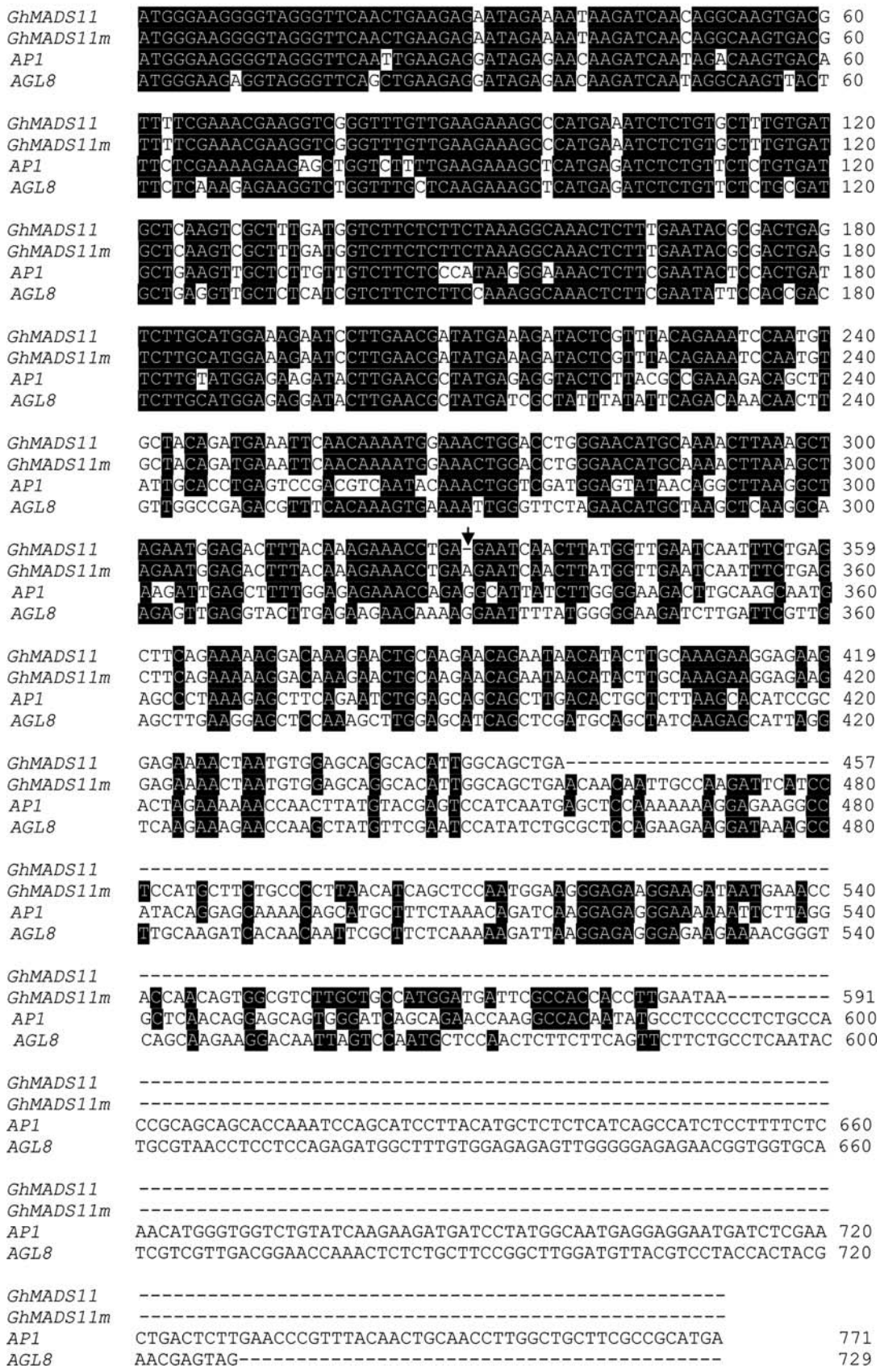

Figure 2 Sequence comparison between *GhMADS11*, *GhMADS11m*, *AP1* and *AGL8* cDNAs Numbers on the right indicate the nucleic acid length. The nucleotides identical among the sequences are indicated in black. The arrows indicate the frameshift mutation site of *GhMADS11* sequence. *GhMADS11* in this work, Arabidopsis *AP1* (AT1G69120) and *AGL8* (AT5G60910).

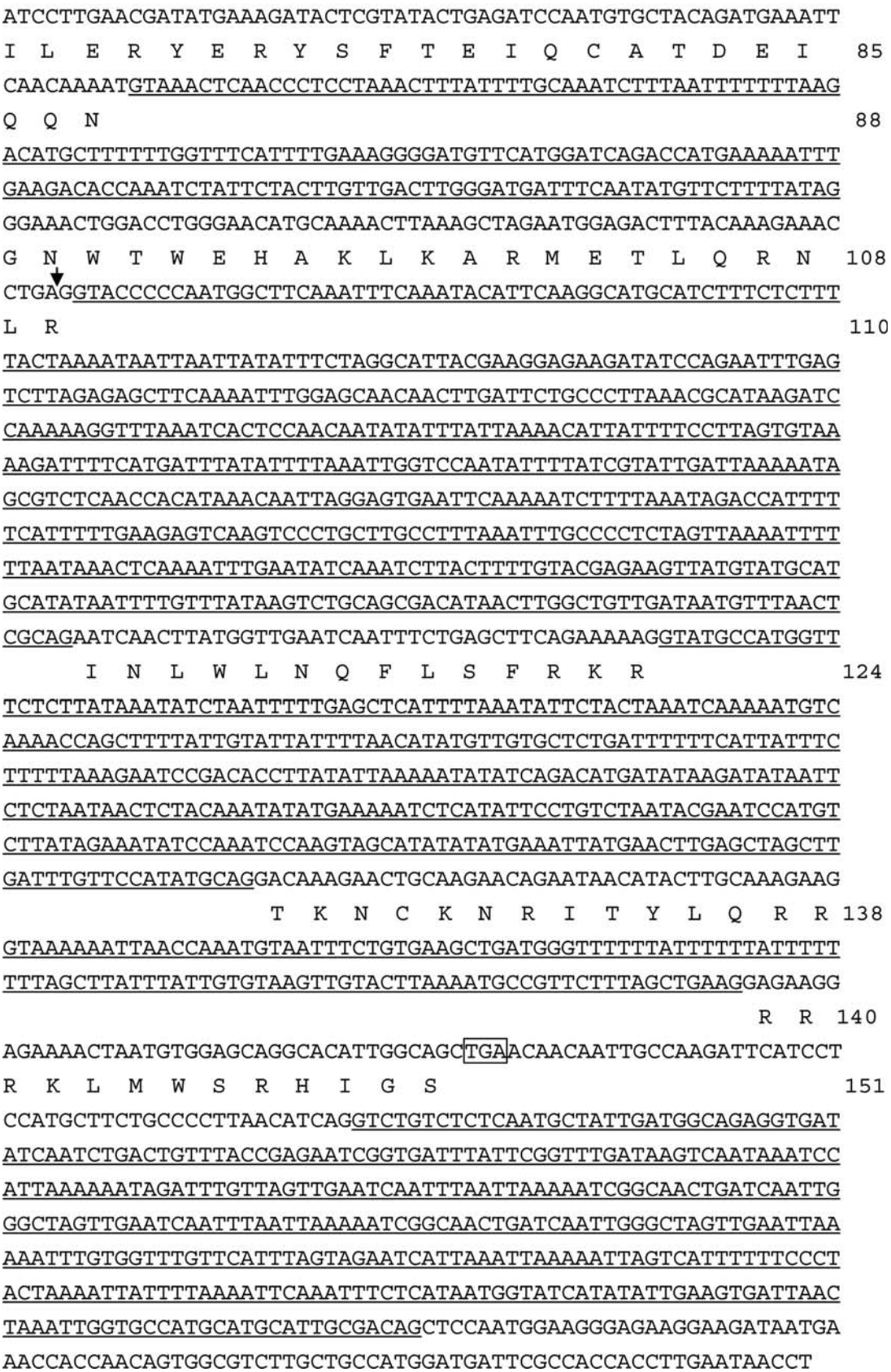

Figure 3 Analysis of genomic DNA sequence of *GhMADS11* gene and its deduced amino acid sequence Partial sequence of *GhMADS11* gene is shown. Numbers on the right indicate the polypeptide length. The underlines show the introns, and the black box indicates the termination codon of the *GhMADS11*gene. The arrows indicate the frameshift mutation site of *GhMADS11* sequence.

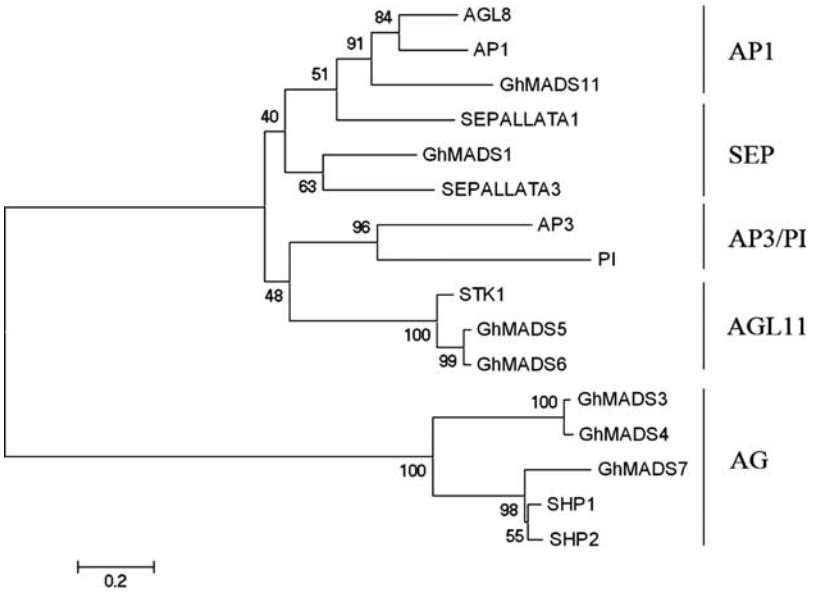

Figure 4 Phylogenetic relationships of cotton MADS proteins and Arabidopsis MADS proteins MADS-box subfamilies are indicated on the right. GhMADS11 is identified from this work, and the others are selected from GenBank. The accession numbers of the protein sequences in GenBank are as follows: AP1 (CAA78909), AGL8 (AAA97403), PI (AAD51998), AP3 (AAD51903), SEP1 (AAU81996), SEP3 (O22456), SHP1 (AAA32730), SHP2 (AAA32735), STK1 (AAC49080), GhMADS1 (AF538965), GhMADS3 (AAL92522), GhMADS4 (ABM69042), GhMADS5 (ABM69043), GhMADS6 (ABM69044), and GhMADS7 (ABM69045).

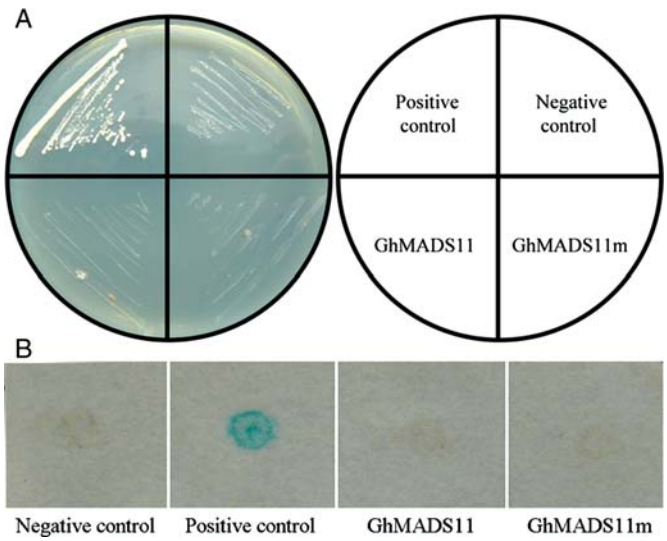

Figure 5 Trans-activation activity assay of GhMADS11 and GhMADS11m in yeast cells (A) Yeast transformants were streaked on SD/-Trp/-Ade medium (SD minimal medium lacking Trp and Ade). (B) Flash-freezing filter assay of the β-galactosidase activity. The yeast cells harboring empty pGBKT7 vector were used as negative control, and harboring pGBKT7-53 and pGADT7-RecT were used as positive control.

GhMADS11 does not act as a transcriptional activator

To determine whether GhMADS11 protein acts as a transcriptional activator, we fused GhMADS11 with the GAL4 DNA-binding domain and transformed it into *Saccharomyces cerevisiae* AH109 and Y187, and then tested the report gene. The transformed yeast cells only harboring GAL4 DNA-binding were used as the negative control, and the cells containing pGBKT7-53 and pGADT7-RecT were used as the positive control. The results revealed that, like negative controls, the cells containing *GhMADS11* could not grow on SD/-Trp/-Ade medium [**Fig. 5(A)**] or activate the *LacZ* reporter gene expression [**Fig. 5(B)**]. By contrast, positive control cells could grow in SD/Trp/-Ade medium and showed β-Galactosidase activity (**Fig. 5**). These results suggested that single GhMADS11 protein may not be a transcriptional activator.

To examine whether the GhMADS11 protein without K-box and normal C-terminus lead to loss transcriptional activity, we created *GhMADS11m* (the frameshift mutation of *GhMADS11*) encoding a derivative of GhMADS11 by primer-based site-directed mutation (**Fig. 2**). *GhMADS11m* was also cloned into pGBKT7 and transformed into AH109 and Y187, and then tested the report gene. The results showed that the cells containing *pGBKT7-GhMADS11* did not grow on SD/-Trp/-Ade medium, and did not activate the *LacZ* reporter gene expression either (**Fig. 5**).

GhMADS11 does not form homodimers

To identify whether GhMADS11 proteins form homodimers, yeast two-hybrid technology was employed to analyze the interaction between two GhMADS11 molecules (see the section Methods). As shown in **Fig. 6**, the interaction between two GhMADS11 molecules was not

detected in the yeast two-hybrid assays, suggesting that GhMADS11 molecules may not form homodimers.

To determine whether GhMADS11 protein without K-box and normal C-terminus leads to loss of the capability of forming homodimers, GhMADS11m was also cloned into pGADT7. Just like GhMADS11, the interaction between two GhMADS11m molecules or between GhMADS11 and GhMADS11m was not detected in the yeast two-hybrid assays (**Fig. 6**).

GhMADS11 is preferentially expressed in fibers

To investigate the expression pattern of *GhMADS11* gene, quantitative RT-PCR was carried out using gene-specific primers and cDNAs from different cotton tissues as templates. Our results revealed that *GhMADS11* gene was strongly expressed in fibers, but negligible amounts of its transcripts were detected in other tissues of cotton [**Fig. 7(A)**].

To confirm whether *GhMADS11* gene expression was developmentally regulated in fibers, quantitative RT-PCR was also performed to analyze the gene expression pattern during fiber development in the same cultivar of cotton. The results demonstrated that *GhMADS11* expression levels varied with fiber development [**Fig. 7(B)**]. At early fiber development stage, very weak signals of the gene expression were detected in 3 and 6 DPA fibers. As fiber further developed, *GhMADS11* transcripts were increased in 9 DPA fibers, and reached its peak value in 15 DPA fibers. Thereafter, its expression activity was quickly declined to very low level in 18 and 21 DPA fibers. The above data indicated that *GhMADS11* expression was fiber-preferential and developmental-regulated, implicating that the gene may participate in fiber development.

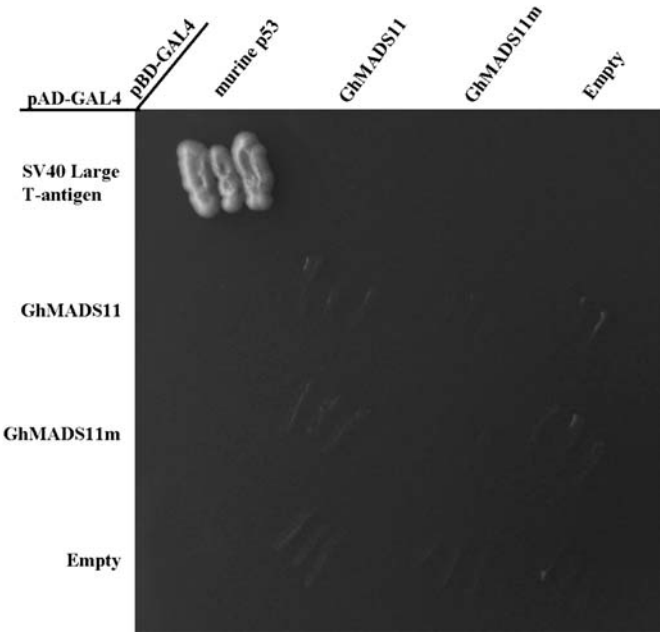

Figure 6 One to one interaction between GhMADS11 and GhMADS11m Yeast cells harboring pGADT7-GhMADS11 (or pGADT7-GhMADS11m) and pGBKT7-GhMADS11 (or pGBKT7-GhMADS11m) were assayed on QDO nutritional selection medium, using empty pGBKT7 and pGADT7 vectors as negative controls, and pGBKT7-53 and pGADT7-RecT vectors as positive controls.

Overexpression of *GhMADS11* in yeast promotes cell elongation

To investigate whether *GhMADS11* gene plays a role in cell elongation, the coding sequence of *GhMADS11* gene was cloned into a yeast vector pREP-5N, and introduced into yeast (*Schizosaccharomyces pombe*) cells. After 20 h of inductive cultivation, morphological changes of yeast cells were detected under the optical microscope. The results indicated that transformed cell lines expressing *GhMADS11* grown in induction medium were longer than those harboring the same vector but grown in uninduced conditions [**Fig. 8(B,E)**]. In contrast, the transformed cell lines with the empty pREP-5N vector displayed normal length grown in either induction medium or non-induction medium [**Fig. 8(A,D)**]. Moreover, it was observed that the *GhMADS11* overexpressed cells were monocyte when identified by DAPI staining (**Fig. 8**). This result indicated

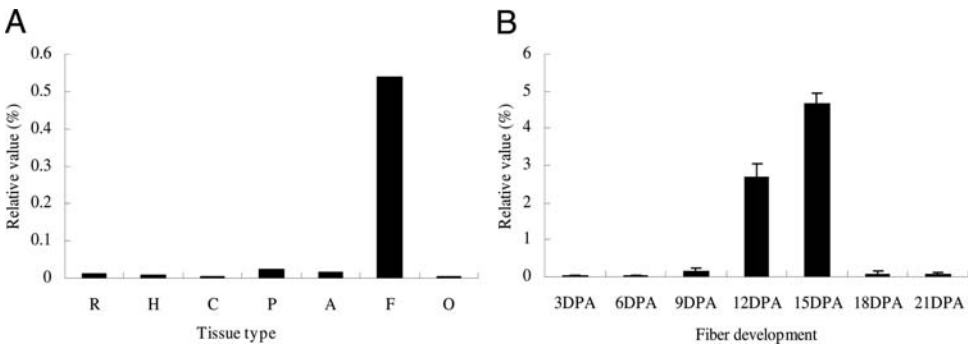

Figure 7 Quantitative RT-PCR analysis of *GhMADS11* gene expression (A) Expression of *GhMADS11* in cotton tissues. (B) Expression of *GhMADS11* in fiber development of cotton. DPA, day post-anthesis. R, roots; H, hypocotyls; C, cotyledons; P, petals; A, anthers; O, ovules; F, fibers; 3 DPA, 6 DPA, 9 DPA, 12 DPA, 15 DPA, 18 DPA and 21 DPA refer to 3, 6, 9, 12, 15, 18, 21 DPA fibers.

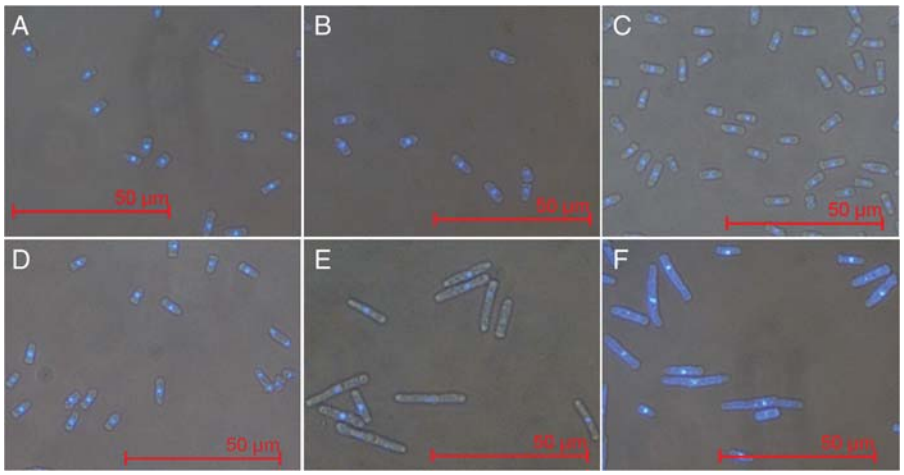

Figure 8 Overexpression of *GhMADS11* and *GhMADS11m* promoted typical cell elongation of fission yeast Micrographs were taken using microscopy. (A) Yeast cells harboring empty pREP-5N vector were cultured in non-induction medium. (B) Yeast cells harboring pREP-5N-*GhMADS11* were cultured in non-induced medium. (C) Yeast cells harboring pREP-5N-*GhMADS11m* were cultured in non-induction medium. (D) Yeast cells harboring empty pREP-5N vector in induction medium exhibited normal morphology and length as those grown under non-induction conditions. (E) Yeast cells harboring pREP-5N-*GhMADS11* were cultured in induction medium to undergo atypical longitudinal elongation. (F) Yeast cells harboring pREP-5N-*GhMADS11m* also showed the atypical longitudinal elongation in induction medium. The nucleus in yeast cells is shown by nucleus-specific fluorescent dye DAPI staining.

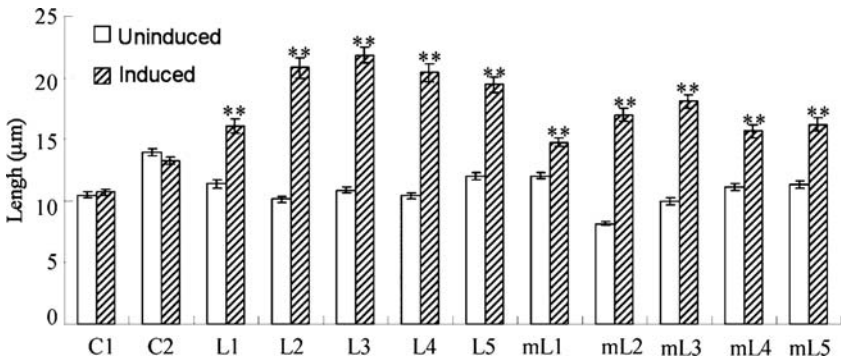

Figure 9 Statistical analysis of the length of yeast cells C1 and C2, controls (yeast cells with empty vectors); L1－L5, transformed yeast cell lines harboring *GhMADS11* gene; mL1－mL5, transformed yeast cell lines harboring *GhMADS11m* gene. $^{**}P < 0.01$. Independent *t*-tests demonstrated that there was significant difference in cell length between the transformed yeast cells and controls (*t*-test for equality of means).

that the longer cells were caused by cell elongation, rather than cell division. The five transformed cell lines and two controls were randomly selected for measuring cell length (50 cells each line). Statistical analysis indicated that the induced yeast cells were 1.4−2.0-fold longer than those uninduced yeast cells and controls (**Fig. 9**), suggesting that overexpression of the *GhMADS11* gene stimulated the longitudinal growth of the host cells.

To examine whether the GhMADS11 protein without K-box and normal C-terminus changed its function, we created *GhMADS11m* (the frameshift mutation of *GhMADS11*) encoding a derivative of GhMADS11 by primer-based site-directed mutation (**Fig. 2**). The *GhMADS11m* was cloned into PREP-5N vector, and then transferred into yeast cells. The experimental results indicated that the phenotype of the transformed yeast cells expressing *GhMADS11m* is similar to that of the transformed yeast cells expressing *GhMADS11* [**Fig. 8(C,F)**]. Statistical analysis revealed that there was significant difference in cell length between the transformed yeast cell lines and controls grown in induction medium (**Fig. 9**). These results indicated that The N-terminus including MADS domain, but not the additional K-box and normal C-terminus, of GhMADS11 protein functions in cell elongation.

Discussion

In this study, we identified a cotton gene, *GhMADS11* which encodes a novel MADS protein with a shortened sequence produced by frameshift mutation during evolution. The *GhMADS11* expression was fiber specific and

developmental regulated. Similarly, previous studies have revealed that the transcripts of *GhFLA1*, *GhFLA2*, and *GhFLA4* are accumulated at relatively high levels in fibers. As a result, the highest levels of *GhFLA1* and *GhFLA4* transcripts are detected in 10 DPA fibers, while *GhFLA2* shows its highest expression level in 20 DPA fibers [33]. *GaMYB2* transcripts are largely accumulated in 3–9 DPA fibers, and then decline with further fiber development [6]. These genes can be loosely described as fiber specific, and may provide important clues about their roles during fiber development.

The data presented in this study indicated that overexpression of *GhMADS11* in fission yeast (*S. pombe*) significantly promoted atypical longitudinal growth of the host cells, like some fiber-specific genes, such as cotton *TUA9* [10] and *14-3-3s* [32]. However, *GhMADS11* activity reached its peak value in the relatively later stage of fiber elongation (12–15 DPA), whereas *GhTUA9* and *Gh14-3-3* genes show their highest expression levels in relatively earlier fiber development (3–10 DPA). This suggested that *GhMADS11* gene may be only involved in fiber polar elongation, whereas *GhTUA9* and *Gh14-3-3s* may play roles in both cell expansion stage and polar elongation stage during early fiber development.

It has been demonstrated that MADS proteins are conserved in virtually every plant. GhMADS11 falls within the AP1 clade of A-type gene coding protein and shows a high degree of similarity with AGL8 and AP1 of Arabidopsis. Although GhMADS11 shares relatively high amino acid sequence similarity with AGL8 and AP1, the expression pattern of *GhMADS11* gene is very different from those of *AGL8* and *AP1* genes. Previous studies have reported that *AP1* is highly expressed in young flower primordial, sepals, and petals [34], while *AGL8* mRNA accumulates at the highest level in inflorescence apical meristem [35]. A later study has revealed that AGL8 belonging to AP1 subfamily is involved in both determination of meristem identity and of carpel development [36]. In contrast, *GhMADS11* transcripts were accumulated specifically in 12–15 DPA fibers. These data provided some hints that GhMADS11 may function mainly in developing fibers, which is different from the traditional A-type MADS proteins.

In plants, most MADS proteins are MICK-type, including two conserved domain, the MADS-box and K-box, and two variable segments, I-region, and C-terminus [12]. Sequence analysis revealed that GhMADS11 contains highly conserved MADS domain, but is devoid of partial K-domain and normal C-terminus (**Fig. 1**), compared with the known MADS proteins in plants, which implies that *GhMADS11* may originate from a traditional MADS gene in cotton, owing to its frameshift mutation during evolution. Previous study have indicated that A-type MADS proteins, such as AP1 and AG, usually form homodimers to regulate the expression of downstream target genes [16]. Similarly, in this study, single GhMADS11 could not activate the reporter genes expression in yeast cells. Previous studies further showed that AP1 can interact with SEP3, and AP1-SEP3 complex can function as AP1 homodimer [16]. Therefore, in cotton fibers, the formation of complex may provide the molecular basis for the combinatorial interaction of GhMADS11 and other MADS-box protein instead of itself. It has been demonstrated that the K domain mediates specific protein/protein interactions and is required for the formation of DNA-binding dimmers and AP3-PI heterodimer [37]. In contrast with AP3 and PI, the K domain is dispensable for the formation of DNA-binding dimmers and homodimer of AP1 and AG [38,39]. In Arabidopsis, the 77 amino acids in the N-terminus of AGL2 are sufficient for DNA-binding and dimerization, similar to AP1 and AG [40]. Moreover, some evidence also indicated that both full-length AP1 protein and AP1-MIK (AP1 protein without C-terminus) can interact with AP3-PI complex and SEP3. Otherwise, with the deletions of C-terminus of both AP1 and SEP3, the complex of those genes cannot be formed [16]. In this study, our data suggested that GhMADS11 protein lost a part of the K-box and entire C-terminus by gene frameshift mutation during evolution. It is noteworthy that overexpression of GhMADS11 in fission yeast showed the same phenotype as its derivative GhMADS11m which contains the entire K-domain and C-terminus. These results may be explained by the possibility that GhMADS11 protein could reserve the capability of combining other MADS-box proteins and binding to target DNA for regulating cell elongation in fission yeast.

In summary, the data presented here indicated that the fiber-preferential *GhMADS11* gene encodes a novel MADS-box protein that lost partial K-domain and entire C-terminus. Like other MADS proteins, single GhMADS11 was not a transcriptional activator. Overexpression of this gene in fission yeast promoted atypical longitudinal growth of the host cells. Thus, the results of this work may provide some clues to understand the role of this novel MADS-box protein in fiber development of cotton.

Funding

This work was supported by grants from the Ministry of Agriculture of China for transgenic research (2009ZX08009-117B and 2008ZX08009-003) and the National Natural Science Foundation of China (30871317).

References

1 Basra AS and Malik CP. Development of the cotton fiber. Int Rev Cytol 1984, 89: 65–113.

2 Ryser U. Cell wall biosynthesis in differentiating cotton fibers. Eur J Cell Biol 1985, 39: 236–256.
3 Arthur JC. Cotton. In: Kroschwitz JI ed. Polymers: Fibers and Textile: A Compendium. New York: Wiley, 1990, 119–141.
4 Wu Y, Machado AC, White RG, Llewelly DJ and Dennis ES. Expression profiling identifies genes expressed early during lint fibre initiation in cotton. Plant Cell Physiol 2006, 47: 107–127.
5 Machado A, Wu Y, Yang Y, Llewellyn DJ and Dennis ES. The MYB transcription factor GhMYB25 regulates early fibre and trichome development. Plant J 2009, 59: 52–62.
6 Wang S, Wang JW, Yu N, Li CH, Luo B, Gou JY and Wang LJ, *et al*. Control of plant trichome development by a cotton fiber MYB gene. Plant Cell 2004, 16: 2323–2334.
7 Shangguan XX, Xu B, Yu ZX, Wang LJ and Chen XY. Promoter of a cotton fibre MYB gene functional in trichomes of Arabidopsis and glandular trichomes of tobacco. J Exp Bot 2008, 59: 3533–3542.
8 Ruan YL and Xu SM. Genotypic and developmental evidence for the role of plasmodesmatal regulation in cotton fiber elongation mediated by callose turnover. Plant Physiol 2004, 136: 4104–4113.
9 Li XB, Fan XP, Wang XL, Cai L and Yang WC. The cotton ACTIN1 gene is functionally expressed in fibers and participates in fiber elongation. Plant Cell 2005, 17: 859–875.
10 Li L, Wang XL, Huang GQ and Li XB. Molecular characterization of cotton *GhTUA9* gene specifically expressed in fibre and involved in cell elongation. J Exp Bot 2007, 58: 3227–3238.
11 Messenguy F and Dubois E. Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. Gene 2003, 316: 1–21.
12 Pařenicová L, de Folter S, Kierrer M, Horner DS, Favalli C, Busscher J and Cook HE, *et al*. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. Plant Cell 2003, 15: 1538–1551.
13 Reichmann JL, Wang M and Meyerowitz EM. DNA-binding properties of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. Nucleic Acids Res 1996, 24: 3134–3141.
14 Davies B, Egea-Cortines M, de Andrade Silva E, Saedler H and Sommer H. Multiple interactions amongst floral homeotic MADS box proteins. EMBO J 1996, 15: 4330–4343.
15 Egea-Cortines M, Saedler H and Sommer H. Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in *Antirrhinum majus*. EMBO J 1999, 18: 5370–5379.
16 Honma T and Goto K. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature 2001, 409: 525–529.
17 Samach A, Onouchi H, Gold SE, Ditta GS, Schwarz-Sommer Z, Yanofsky MF and Coupland G. Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. Science 2000, 288: 1613–1616.
18 Theissen G. Development of floral organ identity: stories from the MADS house. Curr Opin Plant Biol 2001, 4: 75–85.
19 Theissen G and Melzer R. Molecular mechanisms underlying origin and diversification of the angiosperm flower. Ann Bot 2007, 100: 603–619.
20 Rounsley SD, Ditta GS and Yanofsky MF. Diverse roles for MADS box genes in Arabidopsis development. Plant Cell 1995, 7: 1259–1269.
21 Tapia-López R, García-Ponce B, Dubrovsky JG, Garay-Arroyo A, Pérez-Ruíz RV, Kim SH and Acevedo F, *et al*. An AGAMOUS-related MADS-box gene, XAL1 (AGL12), regulates root meristem cell proliferation and flowering transition in Arabidopsis. Plant Physiol 2008, 146: 1182–1192.
22 Buchner P and Boutin JP. A MADS box transcription factor of the *AP1/AGL9* subfamily is also expressed in the seed coat of pea (*Pisum sativum*) during development. Plant Mol Biol 1998, 38: 1253–1255.
23 Carmona MJ, Ortega N and Garcia-Maroto F. Isolation and molecular characterization of a new vegetative *MADS-box* gene from *Solanum tuverosum L*. Planta 1998, 207: 181–188.
24 Alvarez-Buylla ER, Liljegren SJ, Pelaz SP, Gold SE, Burgeff C, Ditta GS and Vergara-Silva F, *et al*. MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. Plant J 2000, 24: 457–466.
25 Gan YB, Filleur S, Rahman A, Gotensparre S and Forde BG. Nutritional regulation of *ANR1* and other root-expressed MADS-box genes in *Arabidopsis thaliana*. Planta 2005, 222: 730–742.
26 Ferrario S, Busscher J, Franken J, Gerats T, Vandenbussche M, Angenent GC and Immink RG. Ectopic expression of the petunia MADS box gene UNSHAVEN accelerates flowering and confers leaf-like characteristics to floral organs in a dominant-negative manner. Plant Cell 2004, 16: 1490–1505.
27 Guo Y, Zhu Q, Zheng S and Li MJ. Cloning of a MADS box gene (*GhMADS3*) from cotton and analysis of its homeotic role in transgenic tobacco. J Genet Genomics 2007, 34: 527–535.
28 Zheng SY, Guo YL, Xiao YH, Luo M, Hou L, Luo XY and Pei Y. Cloning of a MADS box protein gene (*GhMADS1)* from cotton (*Gossypium hirsutum*). Yi Chuan Xue Bao, 2004, 31: 1136–1141. (in Chinese).
29 Lightfoot DJ, Malone KM, Timmis JN and Orford SJ. Evidence for alternative splicing of MADS-box transcripts in developing cotton fibre cells. Mol Genet Genomics 2008, 279: 75–85.
30 Li XB, Cai L, Cheng NH and Liu JW. Molecular characterization of the cotton GhTUB1 gene that preferentially expressed in fiber. Plant Physiol 2002, 130: 666–674.
31 Shao SQ, Li BY, Zhang ZT, Zhou Y, Jiang J and Li XB. Expression of a cotton MADS-box gene is regulated in anther development and in response to phytohormone signaling. J Genet Genomics 2010, 37: 805–816.
32 Zhang ZT, Zhou Y, Li Y, Shao SQ, Li BY, Shi HY and Li XB. Interactome analysis of the six cotton 14-3-3s that are preferentially expressed in fibres and involved in cell elongation. J Exp Bot 2010, 61: 3331–3344.
33 Huang GQ, Xu WL, Gong SY, Li B, Wang XL, Xu D and Li XB. Characterization of 19 novel cotton FLA genes and their expression profiling in fiber development and in response to phytohormones and salt stress. Physiol Plant 2008, 134: 348–359.
34 Mandel MA, Gustafson-Brown C, Savidge B and Yanofsky MF. Molecular characterization of the Arabidopsis floral homeotic gene APETALA1. Nature 1992, 360: 273–277.
35 Mandel MA and Yanofsky MF. The Arabidopsis AGL8 MADS box gene is expressed in inflorescence meristems and is negatively regulated by APETALA1. Plant Cell 1995, 7: 1763–1771.
36 Gu Q, Ferrandiz C, Yanogsky MF and Martienssen R. The FRUITFULL MADS-box gene mediates cell differentiation during Arabidopsis fruit development. Development 1998, 125: 1509–1517.
37 Yang Y, Fanning L and Jack T. The K domain mediates heterodimerization of theArabidopsis floral organ identity proteins, APETALA3 and PISTILLATA. Plant J 2003, 33: 47–59.
38 Mizukami Y, Huang H, Tudor M, Hu Y and Ma H. Functional domains of the floral regulator AGAMOUS: characterization of the DNA binding domain and analysis of dominantnegative mutations. Plant Cell 1996, 8: 831–845.
39 Riechmann JL, Wang M and Meyerowitz EM. DNAbinding properties of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. Nucleic Acids Res 1996, 24: 3134–3141.
40 Huang H, Tudor M, Su T, Zhang Y, Hu Y and Ma H. DNA binding properties of two Arabidopsis MADS domain proteins: binding consensus and dimer formation. Plant Cell 1996, 8: 81–94.