



Cloning and Expression Analysis of a Brassinosteroid Biosynthetic Enzyme Gene, *GhDWF1*, from Cotton (*Gossypium hirsutum* L.)

LUO Ming, XIAO Zhong-yi, XIAO Yue-hua, LI Xian-bi, HOU Lei, ZHOU Jian-ping, HU Ming-yu and PEI Yan

Key Laboratory of Biotechnology and Crop Quality Improvement, Ministry of Agriculture/Biotechnology Research Center, Southwest University, Chongqing 400716, P.R.China

Abstract

Brassinosteroids (BRs) are an important class of plant steroidal hormones that are essential in a wide variety of physiological processes. To determine the effects of BRs on the development of cotton fibers, through screening cotton fiber EST database and contigging the candidate ESTs, a key gene (*GhDWF1*) involved in the upstream biosynthetic pathway of BRs was cloned from developing fibers of upland cotton (*Gossypium hirsutum* L.) cv. Xuzhou 142. The full length of the cloned cDNA is 1 849 bp, including a 37 bp 5'-untranslated region, an ORF of 1 692 bp, and a 120 bp 3'-untranslated region. The cDNA encodes a polypeptide of 563 amino acid residues with a predicted molecular mass of 65 kD. The deduced amino acid sequence has high homology with the BR biosynthetic enzyme, DWARF1/DIMINUTO, from rice, maize, pea, tomato, and *Arabidopsis*. Furthermore, the typical conserved structures, such as the transmembrane domain, the FAD-dependent oxidase domain, and the FAD-binding site, are present in the GhDWF1 protein. The Southern blot indicated that the *GhDWF1* gene is a single copy in upland cotton genome. RT-PCR analysis revealed that the highest level of *GhDWF1* expression was detected in 0 DPA (day post anthesis) ovule (with fibers) while the lowest level was observed in cotyledon. The *GhDWF1* gene presents high expression levels in root, young stem, and fiber, especially, at the fiber developmental stage of secondary cell wall accumulation. Moreover, the expression level was higher in ovules (with fibers) of wildtype (Xuzhou 142) than in ovules of fuzzless-lintless mutant at the same developmental stages (0 and 4 DPA). The results suggest that the *GhDWF1* gene plays a crucial role in fiber development.

Key words: cotton, *DWARF1* gene, fiber, brassinosteroids, phytosterol

INTRODUCTION

Cotton is the world's leading fiber crop. Cotton fiber, the main production of cotton plant and the most important natural fiber, is a single cell resulting from an elongated cell of the ovule epidermis, which undergoes rapid and synchronous elongation (Basra and Saha 2001). The highly elongated structure and exceptional chemical composition establish the cotton fiber as an

ideal experimental model for studies of plant cell elongation, synthesis of cellulose, and cell wall accumulation (Kim and Triplett 2001). Furthermore, because the fiber quantity and quality are established during fiber development, the yield and quality of cotton fibers are closely related to the fiber growth. Therefore, studies on fiber development will contain not only theoretical but practical values.

To investigate the growth and development mechanism of cotton fibers, until now, a few genes

This paper is translated from its Chinese version in *Scientia Agricultura Sinica*.

LUO Ming, E-mail: Luo0424@126.com, luomingyuan@swu.edu.cn, Mobile: 13996007163; Correspondence PEI Yan, E-mail: Peiyan@swau.edu.cn, Tel: +86-23-68251883

that preferentially or specifically expressed in the fiber developmental stages have been cloned and identified. Through screening the cDNA library of cotton fibers, John and Crow (1992) first identified and characterized the *E6* gene that specially expressed and mainly activated in fiber. Subsequently, the *H6* gene has been cloned and proven to express specifically in fiber cells. Further studies revealed that the *H6* mRNA was predominantly present during the formation of the primary cell wall (John and Keller 1995). *pGhEX1*, an expansion protein gene regulated developmentally during the fiber elongation stage, was likely to play an important role in fiber cell elongation with turgor-driven growth in plant cells (Orford and Timmis 1998). Several *myb*-like genes involved in cotton fiber development were isolated first by Loguercio and his colleague (Loguercio *et al.* 1999). Recently, one sucrose synthase gene (*Sus*) was identified and proven to be crucial for fiber cell initiation and elongation (Ruan *et al.* 2003). Using SSH and microarray methods, 172 genes that preferentially or specifically expressed in fiber developmental stages were identified, many of which encoded components involved in cell expansion, lipid biosynthesis, and phytohormone biosynthesis (Shi *et al.* 2006).

However, fiber growth and development were not affected in transgenic cotton plants either up- or down-regulating some cloned gene expression. Therefore, the mechanism of fiber development is hitherto largely unknown. Recently, a few studies have demonstrated that brassinosteroids (BRs) play pivotal roles in the growth and development of cotton fibers. The exogenous application of brassinolide (BL), the most bioactive BRs found to date, results in a 6.1 and 9.5% increase of the fiber length and strength in a field test, respectively. In the system of cotton ovule culture, the treatment of cultured ovules of Coker312 with BL leads to a 30–38% increase on the mean fiber length. In the same way, treatment of ovules of Ligon mutant, a lintless mutant, results in a 16.5% increase of the fiber length and a 26.1% increase of the fiber production relative to the control samples (Kasukabe *et al.* 2001). More recently, exogenous application of BL was found to promote fiber development, while addition of a brassinosteroid biosynthesis inhibitor brassinazole (Brz) inhibited fiber elongation (Sun *et al.* 2004; Sun and Allen 2005; Shi *et al.* 2006). In addition, increasing reports

indicated that the genes responsible for various biosynthetic enzymes in the BR biosynthetic pathway and the genes involved in BR perception/signaling such as *GhDET2*, *GhSMT1*, *GhBR11*, and *GhBIN2*, show the highest expression profile during the fiber initiation and elongation (Luo *et al.* 2007; Sun *et al.* 2004; Sun and Allen 2005; Shi *et al.* 2006). These results suggested that BR is an important growth-promoting hormone required for fiber cell development. However, the final production or intermediates in BR biosynthetic pathways perform different bioactivities and physiological functions in the complicated network of plant development, for example, sitosterol is a primer in cellulose biosynthesis (Peng *et al.* 2002). The effect of one gene involved in BR biosynthesis on the fiber growth and development needs to be further investigated.

By combining all effects of BRs on fiber development, it is proposed that BRs are an important clue to elucidate the molecular mechanism of cotton fiber growth and development and to improve the quantity and quality of cotton fibers through gene engineering. Therefore, it is necessary to clone the genes involved in BR biosynthesis and signal transduction from cotton fibers. In this article, we have cloned the *GhDWF1* gene involved in BRs upstream biosynthetic pathway by screening the cotton EST database and contigging the candidate EST sequences. The expression patterns of *GhDWF1* gene in various organs and various stages of fiber development were obtained. The results revealed the relationship of the *GhDWF1* gene with the growth and development of cotton fibers.

MATERIALS AND METHODS

Plant material and growing condition

Gossypium hirsutum cv. Xuzhou 142 and its fuzzless-lintless mutant (*fl*) were provided by the Cotton Research Institute, Chinese Academy of Agricultural Sciences, and were grown in field with normal administration.

Isolation of cotton genomic DNA and total RNA

Genomic DNA was isolated from young, expanding cotton leaves using the CTAB method as described

previously (Xiao *et al.* 2002). For isolation of total RNA from various organs and tissues in cotton plant, the roots, cotyledons, and hypocotyls were obtained from 7-day seedlings grown in sand, and the leaves, flowers, ovules, and fibers of different developmental stages were obtained from mature plants. Isolation of total RNA followed the protocol described previously (Luo *et al.* 2003).

Synthesis of first-strand cDNA

20 µg total RNA from various organs and tissues was used for each sample to synthesize first-strand cDNA by the instructions of first strand cDNA synthesis kit (TaKaRa).

Screening cotton EST sequences with high homology to plant DWF1/DIM genes

The amino acid sequence of *Arabidopsis* DWARF1/DIMINUTO (Klahre *et al.* 1998) was found out in the GenBank database and was used as a probe to screen the cotton EST database in GenBank. The cotton ESTs with high sequence homology to *Arabidopsis* DWARF1/DIMINUTO were identified using the BLAST program (<http://www.ncbi.nlm.nih.gov/blast>).

Cloning the full length of cDNA of the *GhDWF1* gene

Candidate EST sequences identified were subjected to contig analysis with the SeqMan program (DNASar, Madison, WI). BLASTX analysis (<http://www.ncbi.nlm.nih.gov/blast>) was performed to determine the putative full length open reading frame (ORF). A pair of specific primers P1 (5'-GGAGCTTAAGAGGAAGA GGCATTC-3') and P2 (5'-ATCCCATACATCAGTTC CGACCC-3') designed based on the putative ORF was used to amplify the cDNA sequence of *GhDWF1* from 4-DPA (day post anthesis) ovules (with fibers). The PCR thermocycling condition was as follows: 94°C for 5 min, followed by 35 cycles of 94°C for 1 min, 56°C for 45 s, and 72°C for 1.5 min, and a final extension at 72°C for 10 min. The PCR product was purified and ligated with pUCm-T vector and sequenced.

Sequence analysis of the deduced GhDWF1 protein

The protein similarities of the deduced GhDWF1 protein with other DWARF1/DIMINUTO homologues were determined by the program of Basic Local Alignment Search Tool (BLAST) at NCBI. The alignment analysis of multiple amino acid sequences was performed with the Megalign program (DNASar, Madison, WI).

RT-PCR analysis

Reverse transcription PCR (RT-PCR) assays were carried out according to the instruction of the TaKaRa RT-PCR kit. The amplified production of cotton *Ubiquitin* gene was used as a RNA loading control and cDNA quality control. A pair of primers named Ubi-up and Ubi-dn (Ubi-up: 5'-CAGATCTTCG TCAAAACC-3', Ubi-dn: 5'-GACTCCTTCTGGATG TTGTA-3') was used for *Ubiquitin* amplification.

Southern hybridization

For Southern analysis, about 20 µg cotton genomic DNA were digested by four groups of restriction enzymes, such as *Xba* I, *Bam*H I and *Eco*R V, *Eco*R I and *Spe* I, and *Xho* I. The cotton DNA digested completely were subjected to fractionate on 0.7% (m/v) agarose, and then transferred onto positively charged nylon membrane (Bio-Rad) as described previously (Sambrook and Russell 2001). The fragment contained a partial 3'-untranslate region, and a partial coding sequence of *GhDWF1* was used as the probe labeled by α -³²P using Ready-To-Go Labeling Beads (Pharmacia). After washed, Kodak X-ray films (Eastman-Kodak, Rochester, NY, USA) were exposed to the nylon membranes for two days with the aid of intensifier screens.

RESULTS

Screening and contigging EST sequence with the homology to *Arabidopsis* DWF1/DIMNUTO

To obtain a *DWF1* gene from upland cotton (*Gossypium hirsutum* L.), a few cotton ESTs were identified ac-

cording to the amino acid sequence similarity to DWF1, an *Arabidopsis* cell elongation protein (Klahre *et al.* 1998), using the BLASTN program (<http://www.ncbi.nlm.nih.gov/blast>). Seven of them (GenBank accession No. AI725977, AI332040, AW587455, BE053495, BF271039, BG439971 and BG445535) located on N-terminus, middle region, and C-terminus of DWF1 protein were selected for contig analysis through the SeqMan program of DNASTar package. A consensus of 1 980 bp was obtained and subjected to sequence analysis in the GenBank database. The results show that the consensus sequence contains a full ORF of 1 692 bp (from 45 to 1 763 bp)(Fig.1). The deduced protein from the putative ORF matches well with the other DWF1/DIMINUTO homologues.

Cloning and characterization of *GhDWF1* gene

Based on the result of the consensus sequence analysis, a pair of specific primers (P1 and P2) located on the flank sequence of ORF was designed and synthesized. The cDNA from 4 DPA ovules (with fibers) was used as the template to amplify the sequence expected. The resulting sequence indicated that the obtained fragment was 1 849 bp, 5 bp shorter in the 3'-untranslated region relative to the corresponding consensus (Fig.1). The full length of the ORF was 1 692 bp, starting at nucleotide 38 and stopping at nucleotide 1 729, corresponding to 563 amino acids (Fig.1), and designed as *GhDWF1* (GenBank accession No. AF538359).

The homology analysis of the deduced *GhDWF1*

To further characterize the homology of *GhDWF1* with other species DWF1 and investigate their evolution relationship, a comparison of these amino acids was done. The results revealed that *GhDWF1* was 84.0% identity and 91.0% similarity with the *Arabidopsis thaliana* DWARF1/DIMINUTO (GenBank accession No. NP850616), and was 86.0% identity and 91.0% similarity with the cell elongation protein DWARF1/DIMINUTO derived from rice (GenBank accession No. NP921328) (Fig.2). These indicate a significant similarity to the members of various plants, such as *A. thaliana*, *Pisum sativum* Linn., *Oryza sativa*, and *Lycopersicon esculentum*. The phylogenetic analysis

suggests that the *GhDWF1* protein is more closely related to *Zea mays* and *Oryza sativa* than any other dicotyledonous plant, such as *P. sativum* Linn., *L. esculentum*, and *A. thaliana* (Fig.3). In addition, the deduced protein consisting of one N-terminal peptide formed a transmembrane domain (24-47), which was likely to be involved in transmembrane and protein location. Furthermore, a homology of FAD-oxidase (109-217) and a structure of the FAD-binding domain (159-170) were located in the middle region (Fig.2). These results reveal that *GhDWF1* is a DIM/DWF1 homologue from cotton.

Southern hybridization

To determine the structure of *GhDWF1* in the genomic DNA, Southern hybridization analysis was used to detect the copy number. Cotton genomic DNA was isolated from the young leaves of cotton plants and digested with four groups of restriction enzymes namely *Xba* I, *Bam*H I and *Eco*R V, *Spe* I and *Eco*R I, and *Xho* I, respectively. A fragment of *GhDWF1* cDNA labeled by α -[^{32} P]dCTP was used as the probe for hybridization. The hybridization result indicates that only one band was detected in each lane, which suggests that the *GhDWF1* gene is a single copy in cotton genome (Fig.4).

Expression pattern of *GhDWF1* in cotton plant and at various stages of fiber development

The expression pattern is a significant data for indicating the function of the target gene. To determine the role of the *GhDWF1* gene in the growth and development of cotton plants, especially in the development and growth of fibers, the expression levels of the *GhDWF1* gene were detected by RT-PCR. The total RNA was obtained from the root and cotyledon of seedlings, and from leaves, buds, and 0 DPA ovules of mature plants of *Gossypium hirsutum* cv. Xuzhou 142. The total RNA was used to synthesis the first strand of cDNA, which serves as a template in the PCR process. As seen in Fig.5, the relative expression levels of *GhDWF1* were various in all detected organs and tissues. The peak of mRNA accumulation corresponded to the period of fiber initiation at 0 DPA. The accumulation of the transcript

CS:	GTGCTTTGGAGC TTAAGAGGAAGAGGC ATCTTTTCACTCCCAAATGTCAGATCTTCAAGCACCCTTCGCCCAAAGAGGAAG	13
AS:	GGAGCTTAAGAGGAAGAGGC ATCTTTTCACTCCCAAATGTCAGATCTTCAAGCACCCTTCGCCCAAAGAGGAAG	76
AA:	M S D L Q A P L R P K R K	13
CS:	AAGGGCTTGGTGGAC TTTTGGTCCAG TTTGGTGGGTTTGTATATTTTGTCTCTCC TTTTCGGCCTGTGATTACTTT	167
AS:	AAGGGCTTGGTGGAC TTTTGGTCCAG TTTGGTGGGTTTGTATATTTTGTCTCTCC TTTTCGGCCTGTGATTACTTT	160
AA:	K G L V D F L V Q F R W I F V I F F V L P F S T L Y Y F	41
CS:	CTCATATATCTTGGAGATGTCAGATCCGAGATGAAG TCC TACAAGCAGCGTCAGAGGAAC ATGATGAAAATGTTAAGAAAGTA	251
AS:	CTCATATATCTTGGAGATGTCAGATCCGAGATGAAG TCC TACAAGCAGCGTCAGAGGAAC ATGATGAAAATGTTAAGAAAGTA	244
AA:	L I Y L G D V R S E M K S Y K Q R Q K E H D E N V L K V	69
CS:	GTGAAGCGCTCTCAACAG AGGAATCCAAAAAGGATGGTCTTGTATGACAGC CCGCAAC CATGGATTG CKGTGGGA TGGCG	335
AS:	GTGAAGCGCTCTCAACAG AGGAATCCAAAAAGGATGGTCTTGTATGACAGC CCGTAAAC CATGGATTG CCGTGGGA ATGGCG	328
AA:	V K R L K Q R N P K D G L V C T A R K P W I A V G M R	9
CS:	AATGTARAC TATAAGAGAGCTCGCCAT TATGAAGTTGATTGTG TGTCTTCG TAACATTC TTGAAATTG ATARAGAGAGAAATG	419
AS:	AATGTAGAC TATAAGAGAGCTCGCCAT TATGAAGTTGATTGTG TGTCTTCG TAACATTC TTGAAATTG ATAAACAG AGAATG	412
AA:	N V D Y T K R A R H Y E V D L S A C F R N I L E I D K Q R M	125
CS:	ATTGCAAGGGTTGAGCCACTTGTAAAC ATGGGGCAG ATTACAG TGTCACAGT TCCAAATGAATCTTCCCTTGCTGTG TTGCA	503
AS:	ATTGCAAGGGTTGAGCCACTTGTAAAC ATGGGGCAG ATAACAGTGTCACAGT TCCAAATGAATCTTCCCTTGCTGTG TTGCA	496
AA:	I A R V E P L V N M G Q I T R V T V P M N L S L A V V A	153
CS:	RAGCTTGACGATCTTACAGTAGGTGGTCTCATCAATGGCTACGGGATTGAAGG AARCTCAC ACATCTATG GCGTGTCTTCTGAT	587
AS:	GAGCTCGACGATCTTACAGTAGGTGGTCTCATCAATGGCTACGGGATTGAAGG AAGCTCAC ACATCTATG GCGTGTCTTCTGAT	580
AA:	E L D D L T V G G L I N G Y G I E G S S H I Y G L F S D	181
CS:	ACTGTTGTAGCTTATGAG ATAGTTTGTGCTGATGGC CGTGTGT TAGAGCTAC MAAGGACAATGAATATCTGATCTTT TCTAT	671
AS:	ACTGTTGTAGCTTATGAG ATAGTTTGTGCTGATGGC CGTGTGT TAGAGCTAC CAAGGACAATGAATATCTGATCTTT TCTAT	664
AA:	T V V A Y E I V L A D G R V V R A T K D N E Y S D L F Y	209
CS:	GCTATCCCA TGGTCTCAAGGA CTCTTGGATTCTTGTGTGCTGCG AAATCAAGCTTATAC CTGTTAAAG AATACATG AGACTG	755
AS:	GCTATCCCA TGGTCTCAAGGA CTCTTGGATTCTTGTGTGCTGCG AAATCAAGCTTATAC CTGTTAAAG AATACATG AGACTG	748
AA:	A I P W S Q G T L G F L V A A E I K L I P V K E Y M R L	237
CS:	ACATACAGC CGTGTAGTGGGAATTTG CAGGAGCTTGTCAAGG TTATATGGACTCTTTTG CACCCAGAG ATGGTGATCAGGAT	839
AS:	ACATACAGC CGTGTAGTGGGAATTTG CAGGAGCTTGTCAAGG TTATATGGACTCTTTTG CACCCAGAG ATGGTGATCAGGAT	832
AA:	T Y T P V V G N L Q D L A Q G Y M D S F A P R D G D Q D	265
CS:	AATCCAGAG AAAGTTCCCGATTCTGTAGAAAGCATG GTCTACTC ACCCAGTGAAGGTGTGT TCATGAYTG GGAGATATG CCTCT	923
AS:	AATCCAGAG AAAGTTCCCGATTCTGTAGAAAGCATG GTCTACTC ACCCAGTGAAGGTGTGT TCATGAC TG GGAGATATG CCTCT	916
AA:	N P E K V P D F V E G M V Y S P T E G V F M T G R Y A S	293
CS:	AAAGAAGAG GCCAAGAG AAGGGGAAT AAAATTAAC AATGTAGG TTGGTGGT TAAACCTGGTCTAC AACATGCGC AAACG	1007
AS:	AAAGAAGAG GCCAAGAG AAGGGGAAT AAAATTAAC AATGTAGG TTGGTGGT TAAACCTGGTCTAC AACATGCGC AAACG	1000
AA:	K E E A K K G N K I N N V G W W F K P W F Y Q H A Q T	321
CS:	GCCTTAAAG AAGGGAGAG TTTGTAGAG TACATTCCT ACAAGAGAATATTAC CAGGCACAC AAGATGTT TGTATTGGG AGGGG	1084
AS:	GCCTTAAAG AAGGGAGAG TTTGTAGAG TACATTCCT ACAAGAGAATATTAC CAGGCACAC AAGATGTT TGTATTGGG AGGGG	1082
AA:	A L K K G E F V E Y I P T R E Y Y H R H T R C L Y W E G	349
CS:	AAGCTCATC CTTCATCTCGAGATCAATGGTGGT TGGTTCTCTTGGGCTG GTTGATGC CACCAAGG TTTCTTGC TCAAG	1175
AS:	AAGCTCATC CTTCATCTCGAGATCAATGGTGGT TGGTTCTCTTGGGCTG GTTGATGC CACCAAGG TTTCTTGC TCAAG	1168
AA:	K L I L P F G D Q W W F R F L L G W L M P P K V S L L K	377
CS:	GCTACTCAAGGTGAATCT ATAAGAAAC TATTACCATGAGATGCA TGTGATTCAAGACATGC TTGTTCTCT TTTACAAG TTGGG	1259
AS:	GCTACTCAAGGTGAATCT ATAAGAAAC TATTACCATGAGATGCA TGTGATTCAAGACATGC TTGTTCTCT TTTACAAG TTGGG	1252
AA:	A T C G E S I R N Y Y H E M H V I Q D M L V P L Y K V G	405
CS:	GATGCGCTTGA TTGGGTC CACCATGAG ATGGAGATC TATCCCAT TTGGCTG TCCCGCACC GACTGTTC AAGCTTCTG TCAAG	1343
AS:	GATGCGCTTGA TTGGGTC CACCATGAG ATGGAGATC TATCCCAT TTGGGCTG TCCCGCACC GACTGTTC AAGCTTCTG TCAAG	1336
AA:	D A L E W V H E M E I Y P I W L C P H R L F K L P V K	433
CS:	ACAATGGTG TATCTGAAC CAGGCTTTGAGCAGCAT CACAGACAAGGCGACAC ACCATATG CTCAGATGTT CACCGATG TTGGG	1427
AS:	ACAATGGTG TATCTGAAC CAGGCTTTGAGCAGCAT CACAGACAAGGCGACAC ACCATATG CTCAGATGTT CACCGATG TTGGG	1420
AA:	T M V Y P E P G F E Q H R R Q G D T P Y A Q M F D V G	481
CS:	GTGTATTATGCTCCAKGC CCGTATTG AGGGGTGAAGTATTGTA TGGTGCAGAGGAGTTC GTAAATGGAGCAATGGC TGATC	1511
AS:	GTGTATTATGCTCCAKGC CCGTATTG AGGGGTGAAGTATTGTA TGGTGCAGAGGAGTTC GTAAATGGAGCAATGGC TGATC	1504
AA:	V Y Y A A P G P V L R G E V F D G A E A V R K L E Q W L I	489
CS:	AAAAACCAT AGTTTCCAG CCACAGTATGCAGTGTGTCR GAGCTGAAC GAGAAGGA TTTCTGGAG GATGTTG ATGCTGAGC TGAT	1595
AS:	AAAAACCAT AGTTTCCAG CCACAGTATGCAGTGTGTCR GAGCTGAAC GAGAAGGA TTTCTGGAG GATGTTG ATGCTGAGC TGAT	1588
AA:	K N H S F Q P Q Y A V S E L N E K D F W R M F D A D L Y	517
CS:	GAGCATGTG CGTARGAAG TACGAGCTGTGGGAACA TTCATGAG TGTGTACTACAATCCAAGAAAGGAAGGAGACC GAAAAA	1679
AS:	GAGCATGTG CGTARGAAG TACGAGCTGTGGGAACG TTCATGAG TGTGTACTACAATCCAAGAAAGGAAGGAGACC GAAAAA	1672
AA:	E H V R R K Y G A V G T F M S V Y Y K S K K G R K T E K	545
CS:	GAGGTCCAAGAGCGGAACAAGCCAC CTGAAACTGCGTATGC AGAGGCTGA TTAGTGAG CATGGGGAG TTTGGTAG ATTTT	1763
AS:	GAGGTCCAAGAGCGGAACAAGCCAC CTGAAACTGCGTATGC AGAGGCTGA TTAGTAAG CATGGGGAG TTTGGTAG AAGTT	1756
AA:	E V Q E A E Q A H L E T A Y A E A D *	563
CS:	ATGTAGAGAGGAATCTTGTACCTTGTGTGCTC TTTAACTAGATTGGC TTTACGCG ATTGTATTA TGGGGGCTGGAAC	1847
AS:	ATGTGA-----ATCTTTG CATTGCTT TGTGTGCTC TTTAATTTAGTATTGCT TTTACGCGA TGGTATTA TGTGGGCTGGAAC	1835
CS:	TGATGTTATGGGATAGGTTAAAGCAGTGAATGTTGTGG ATGAWTTATGGCTA TAGTTTAG C-TAT	1931
AS:	TGATGTTATGGGAT	1849
CS:	AATGAGAGGCTTTTGGTTT TWAASSTCTC AAAAAAAAAA AAAAAAAAAA AAAAA	1980

Fig. 1 Co-alignment of the sequences of consensus, amplified fragment, and deduced amino acid. In consensus sequences: S = G or C, K = G or T, M = A or C, Y = C or T, R = G or A, W = A or T (DNASar package, SeqMan program). Lowercase indicates the nucleotide deficit in at least one sequence in consensus sequences. CS, consensus sequences; AS, amplified fragment sequences; AA, amino acid sequences. The sequences of primers (P1 and P2) are indicated by the underline. The box shows the translation start codon and the asterisk shows the terminal codon.

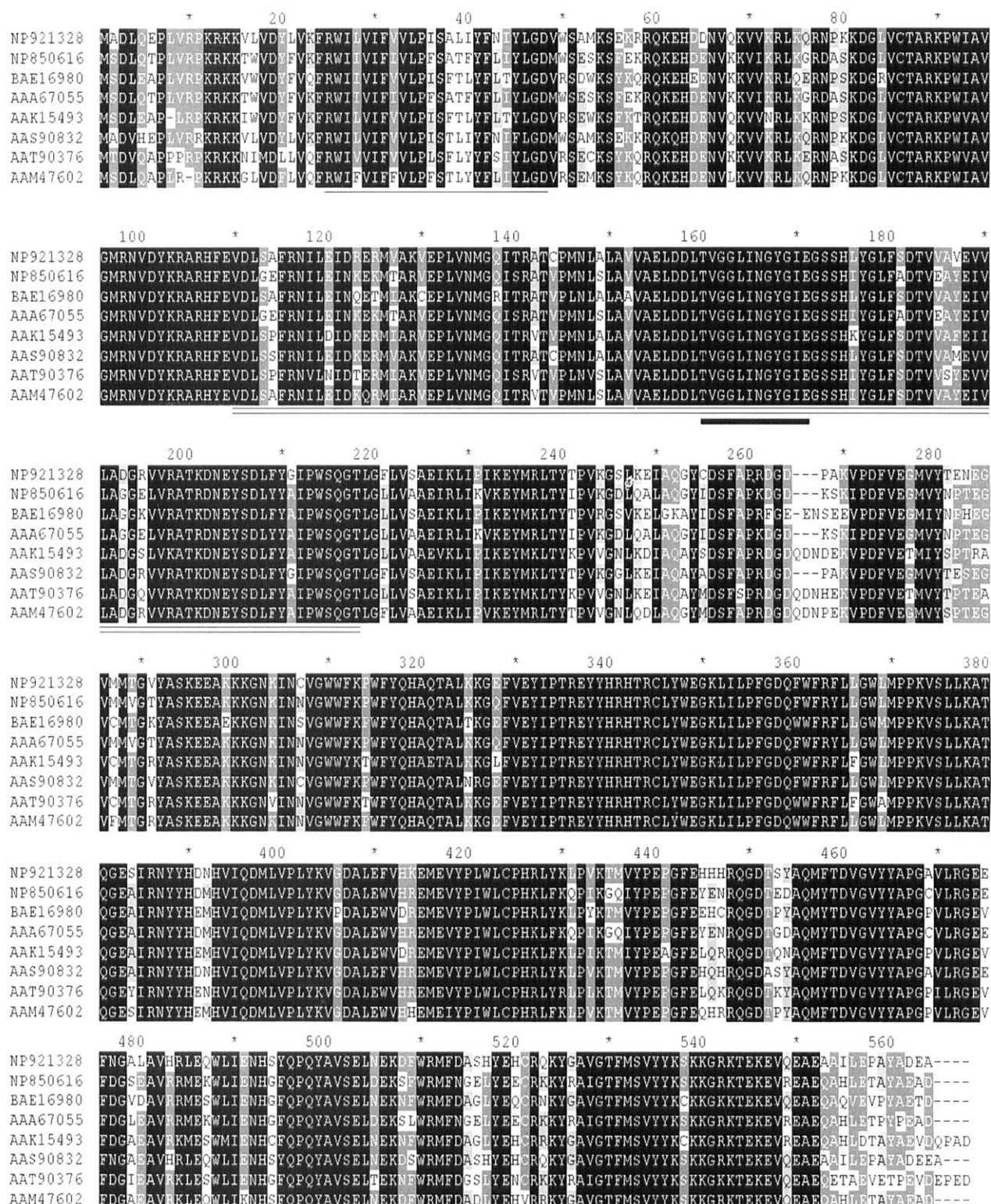


Fig. 2 Alignment of proteins between the deduced amino acid sequence of GhDWF1 and other DWARF1/DIMINUTO proteins. AAM47602, brassinosteroid biosynthetic enzyme in cotton (GhDWF1); AAK15493, brassinosteroid biosynthetic enzyme in pea (LKB); AAT90376, DWARF1/DIMINUTO of tomato; NP921328, cell elongation protein DIMINUTO/DWARF1 of *Oryza sativa*; BAE16980, sterol C-24 reductase in *Zinnia elegans*; AAS90832, brassinosteroid biosynthetic enzyme in *Zea mays* (ZmDWF1); NP850616, *Arabidopsis* DWF1/DIMINUTO1; AAA67055, *Arabidopsis* DIMINUTO. The putative transmembrane domain is underlined by the single line (amino acid residues from 24-47). The sequence homology to FAD-dependent oxidase is underlined by the double lines (109-217). The bold black line below the amino acid sequence (amino acid residues from 159-170) indicates PP-loop, which is the deduced FAD-binding site.

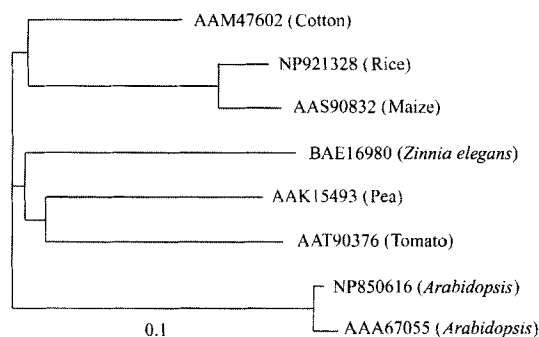


Fig. 3 The phylogenetic tree for DWARF1/DIMINUTO proteins.

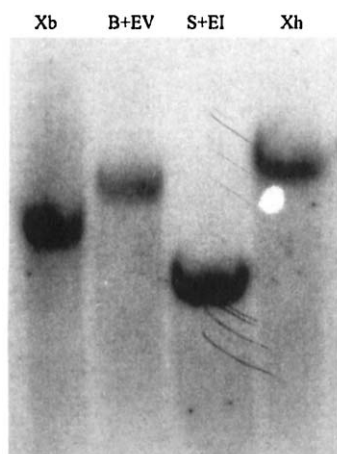


Fig. 4 The Southern-blot pattern of the *GhDWF1* gene. 15 µg genomic DNA of upland cotton (cv. Xuzhou142) was digested by restriction endonucleases and fractionized in agar gel, and then, membrane bolting was performed. An 879 bp fragment, including partial ORF and 3'-untranslated region, was labeled by ³²P, and was used as a probe.

of *GhDWF1* was high in the root and decreased in the order of young stems, buds, leaves, and cotyledons. To determine the expression feature of the *GhDWF1* gene in various developmental stages of ovules and fibers, the cDNA from 0 DPA ovules (with fibers), 4 DPA ovules (with fibers), 12 DPA fibers, and 23 DPA fibers served as the template in PCR. The results showed that the accumulation of mRNA was the highest in the 0 DPA ovules (with fibers), and then moderate expression levels were present in 4 DPA ovules (with fibers) and 12 DPA fibers. The expression level was higher in 23 DPA fibers than in 12 DPA fibers (Fig.5). These results indicated that the *GhDWF1* gene plays an important role in the development of cotton fibers and ovules, especially in fiber cell initiation and secondary cell wall biosynthesis. To further test the role of the

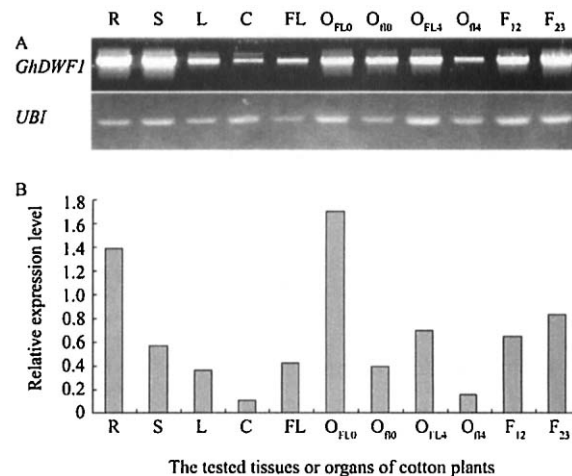


Fig. 5 The expression level of the *GhDWF1* gene in various tissues of cotton plants and at different developmental stages of fibers. A, RT-PCR analysis of *GhDWF1* gene. R, root; S, young stem; L, young leaf; C, cotyledon; FL, bud; O_{FL0}, the 0 DPA ovules (with fibers) of Xuzhou 142 wildtype; O_{FL0}, the 0 DPA ovules of Xuzhou 142 fuzzless-lintless mutant; O_{FL4}, the 4 DPA ovules (with fibers) of Xuzhou 142 wildtype; O_{FL4}, the 4 DPA ovules of Xuzhou 142 fuzzless-lintless mutant; F₁₂, the 12 DPA fibers of Xuzhou 142 wildtype; F₂₃, the 23 DPA fibers of Xuzhou 142 wildtype. *UBI* is a loading control. B, the ratio of *GhDWF1* expression levels. The values of light density were detected by FURI SmartView 2001 package. The ratio of *GhDWF1* to *UBI* in the same sample represents the relative expression level of *GhDWF1*.

GhDWF1 gene in fiber cell development, the *Gossypium hirsutum* cv. Xuzhou 142 and its fuzzless-lintless mutant (*fl*) were subject to compare the expression levels at the same stages. The results indicated that the accumulations of transcript in 0 DPA and 4 DPA ovules (with fibers) of wildtype were higher than in that of *fl* mutant (Fig.5). All the above suggest that *GhDWF1* plays a crucial role in the fiber initiation, elongation, and accumulation of secondary cell wall.

DISCUSSION

The origin and variation of *GhDWF1* gene

In this article, by screening the cotton fiber EST database and contigging the candidate ESTs, a *DWARF1/DIMINUTO* homologue, the *GhDWF1* gene was cloned from the cDNA of 4 DPA ovules (with fibers) and characterized. Although cotton is a dicot plant, the phylogenetic analysis shows that the *GhDWF1* protein is more closely related to the monocotyledon plants such as *O. sativa* and *Z. mays*, whereas it has a farther relationship with the dicot plants such as *A.*

thaliana, *P. sativum* Linn, and *L. esculentum*. The reason for this result may be the few genes with high homology counted in phylogenetic analysis. Further research must be done to analyze the genesis of DWF1 when more DIM/DWF1 homologues will be cloned from various species. Upland cotton (*Gossypium hirsutum* L.) is allotetraploid, whose genome contains two sub-genomes of A and D. Generally, a gene holds at least two copies in upland cotton genome. The result of Southern hybridization indicated only one copy of the *GhDWF1* gene in genomic DNA. This may be explained by the hypothesis proposed by Adams *et al.* (2003). In the polyploidization, the duplicated genes may be retained or lost either as an immediate consequence of polyploidization or on an evolutionary timescale.

The role of *GhDWF1* gene in the development of cotton fibers

Expression analysis indicated that the transcript of the *GhDWF1* gene accumulated preferentially in young and developing tissues, such as ovule, fiber, root, young stem, and expanding leaf. However, less accumulation of mRNA of *GhDWF1* were detected in mature organs and tissues such as cotyledon. The expression pattern is consistent largely with the previous reports about the *DWF1* expression feature in pea and maize (Schultz *et al.* 2001; Takahito *et al.* 1999; Tao *et al.* 2004). This result assumes that *DWF1* has the same function as it acts in different species and the functions of *GhDWF1* can be deduced from the result studied on the model plant, for example, Klahre *et al.* (1998) concluded that *Arabidopsis* DIM/DWF1 is a protein directly involved in the biosynthesis of sterols and BRs. In the period of fiber initiation and rapid elongation, *GhDWF1* has a relative higher expression level that is similar to the other genes involved in the biosynthesis pathway of BRs such as *GhDET2*, *GhSMT1*, and signal transduction of BRs such as *GhBR11* and *GhBIN2* (Luo *et al.* 2007; Shi *et al.* 2006; Sun *et al.* 2004; Sun and Allen 2005). It may be required for the fiber initiation and elongation to produce more BRs resulting from the expression increase of genes responsible for BR biosynthesis. However, the expression pattern of *GhDWF1* is not exactly the same as *GhDET2*, which highly expressed

at the fiber initiation stage and fiber elongation stage, whereas it slightly expressed at the stage of secondary cell wall accumulation. *GhDWF1* displays a relatively higher expression level at the stage of secondary cell wall biosynthesis (23 DPA) with a risen trend, suggesting that it plays a significant role in the cellulose deposit. Recent report demonstrates that sitosterol is a primer for cellulose biosynthesis in plant (Peng *et al.* 2002). More sitosterol is required for more cellulose biosynthesis at the stage of secondary cell wall formation. In this way, the *GhDWF1* gene highly expressed because sitosterol is a direct product of DWF1 enzyme (Choe *et al.* 1999; Fujioka and Yokota 2003).

The fuzzless-lintless mutant (*fl*) derived from *cv.* Xuzhou 142 (*Gossypium hirsutum* L.) was considered as a good model for the study of fiber initiation and elongation because the *fl* mutant fails to perform fiber cell differentiation and initiation from ovule epidermis at 0 DPA. On one hand, the accumulation of *GhDWF1* transcript on the ovules (with fibers) in wild type at the 0 DPA has the highest level, while that in *fl* mutant is considerably lower at the same stage. On the other hand, the expression level of *GhDWF1* is the lowest at the 4 DPA in *fl* mutant while a relative higher level has already kept in wild type at the same stage. These results further confirm that the *GhDWF1* gene plays a crucial role in the fiber initiation and elongation.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (30530490, 30370904, and 30671258), and the National High Technology Research and Development Program of China (2006AA10Z121).

References

- Adams K L, Cronn R, Percifield R, Wendel J F. 2003. Gene duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of Sciences of the USA*, **100**, 4649-4654.
- Basra A S, Saha S. 2001. Growth regulation of cotton fibers. In: Basra A S, ed, *Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing*. Food Products Press, New York. pp. 47-58.
- Choe S, Dilkes B P, Gregory B D, Ross A S, Yuan H, Noguchi T,

- Fujioka S, Takatsuto S, Tanaka A, Yoshida S, Tax F E, Feldmann K A. 1999. The *Arabidopsis* dwarf1 mutant is defective in the conversion of 24-methylenecholesterol to campesterol in brassinosteroid biosynthesis. *Plant Physiology*, **119**, 897-907.
- Fujioka S, Yokota T. 2003. Biosynthesis and metabolism of brassinosteroids. *Annual Review of Plant Biology*, **54**, 137-164.
- John M E, Keller G. 1995. Characterization of mRNA for a proline-rich protein of cotton fiber. *Plant Physiology*, **108**, 669-676.
- John M E, Crow L J. 1992. Gene expression in cotton (*Gossypium hirsutum* L.) fiber: Cloning of the mRNAs. *Proceedings of the National Academy of Sciences of the USA*, **89**, 5769-5773.
- Kasukabe Y, Fujisawa Y, Nishiguchi S, Maekawa Y, Allen R D. 2001. Production of Cotton Fiber with Improved Fiber Characteristics. United States Patent Application 20010018773.
- Kim H J, Triplett B A. 2001. Cotton fiber growth in planta and in vitro: Models for plant cell elongation and cell wall biogenesis. *Plant Physiology*, **127**, 1361-1366.
- Klahre U, Noguchi T, Fujioka S, Takatsuto S, Yokota T, Nomura T, Yoshida S, Chua N H. 1998. The *Arabidopsis* *DIMINUTO/DWARF1* gene encodes a protein involved in steroid synthesis. *The Plant Cell*, **10**, 1677-1690.
- Loguercio L L, Zhang J Q, Wilkins T A. 1999. Differential regulation of six novel MYB-domain genes defines two distinct expression patterns in allotetraploid cotton (*Gossypium hirsutum* L.). *Molecular and General Genetics*, **261**, 660-671.
- Luo M, Xiao Y H, Hou L, Luo X Y, Li D M, Pei Y. 2003. Cloning and expression analysis of a LIM-domain protein gene from cotton (*Gossypium hirsutum* L.). *Acta Genetica Sinica*, **30**, 175-182. (in Chinese)
- Luo M, Xiao Y H, Li X B, Li, Lu X F, Deng W, Li D M, Hou L, Hu M Y, Li Y, Pei Y. 2007. GhDET2, a steroid 5 α -reductase, plays an important role in cotton fiber cell initiation and elongation. *The Plant Journal*, **51**, 419-430.
- Orford S J, Timmis J N. 1998. Specific expression of an expansin gene during elongation of cotton fibers. *Biochimica et Biophysica Acta*, **1398**, 342-346.
- Peng L, Kawagoe Y, Hogan P, Delmer D. 2002. Sitosterol- β -glucoside as primer for cellulose synthesis in plants. *Science*, **295**, 147-150.
- Ruan Y L, Llewellyn D J, Furbank R T. 2003. Suppression of sucrose synthase gene expression represses cotton fiber cell initiation, elongation, and seed development. *The Plant Cell*, **15**, 952-964.
- Sambrook J, Russell D W. 2001. *Molecular Cloning: A Laboratory Manual*. 3rd ed. Cold Spring Harbor Laboratory Press, New York. pp. 474-489.
- Schultz L, Kerckhoffs L H J, Klahre U, Yokota T, Reid J B. 2001. Molecular characterization of the brassinosteroid-deficient *lkb* mutant in pea. *Plant Molecular Biology*, **47**, 491-498.
- Shi Y H, Zhu S W, Mao X Z, Feng J X, Qin Y M, Zhang L, Cheng J, Wei L P, Wang Z Y, Zhu Y X. 2006. Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. *The Plant Cell*, **18**, 651-664.
- Sun Y, Allen R D. 2005. Functional analysis of the *BIN2* genes of cotton. *Molecular Genetic Genomics*, **274**, 51-59.
- Sun Y, Fokar M, Asami T, Yoshida S, Allen R D. 2004. Characterization of the *Brassinosteroid insensitive 1* genes of cotton. *Plant Molecular Biology*, **54**, 221-232.
- Takahito N, Yukiko K, Suguru T, James B R, Motohiro F, Takao Y. 1999. Brassinosteroid/sterol synthesis and plant growth as affected by *lka* and *lkb* mutations of pea. *Plant Physiology*, **119**, 1517-1526.
- Tao Y Z, Zheng J, Xu Z M, Zhang X H, Zhang K, Wang G Y. 2004. Functional analysis of *ZmDWF1*, a maize homolog of the *Arabidopsis* brassinosteroids biosynthetic *DWF1/DIM* gene. *Plant Science*, **167**, 743-751.
- Xiao Y H, Luo M, Fang W G, Luo K M, Hou L, Luo X Y, Pei Y. 2002. PCR walking in cotton genome using YADE method. *Acta Genetica Sinica*, **29**, 62-66. (in Chinese)

(Edited by ZHANG Yi-min)