Front. Agric. China 2008, 2(1): 1–9
DOI 10.1007/s11703-008-0009-3

REVIEW ARTICLE

Diqiu LIU, Xianlong ZHANG

Gene cloning: exploring cotton functional genomics and genetic improvement

Abstract Cotton is the most important natural fiber plant in the world. The genetic improvement of the quality of the cotton fiber and agricultural productivity is imperative under the situation of increasing consumption and rapid development of textile technology. Recently, the study of cotton molecular biology has progressed greatly. A lot of specifically or preferentially expressed cotton fiber genes were cloned and analyzed. On the other hand, identification of stress response genes expressed in cotton was performed by other research groups. The major stress factors were studied including the wilt pathogens *Verticillium dahliae*, *Fusarium oxysporum* f. sp. *vasinfectum*, bacterial blight, root-knot nematode, drought, and salt stress. What is more, a few genes related to the biosynthesis of gossypol, other sesquiterpene phytoalexins and the major seed oil fatty acids were isolated from cotton. In the present review, we focused on the major advances in cotton gene cloning and expression profiling in the recent years.

Keywords cotton, expression profile, gene cloning, fiber development, stress

1 Introduction

Cotton fiber is the most important natural fiber materials for textile industry. Obviously cotton has an important strategic status in the international economy. Under the efforts of breeders, cotton fiber yield and quality have been greatly improved. At the same time, other traits such as disease resistance and insect resistance have been improved, as well. However, the development of cotton cultivars is lagging behind the rapid development of textile industry. The fiber quality does not meet the demands of modern textile technology. Cotton fiber yield traits are negatively related with quality traits, and therefore, it is difficult to improve the yield and quality at the same time using traditional breeding methods. It is hoped that cotton cultivars with high yields and good quality could be bred by combining the methods of genetic engineering with traditional breeding. The development of cotton molecular biology is the base of genetic improvement in cotton breeding. Fortunately, cotton gene cloning, functional gene research and expression profile analysis have given us a better understanding of cotton.

2 Gene expression profiles and gene cloning of cotton fiber development

Mature fiber cells can grow up to 2.5 to 4.0 cm in length. Each fiber is a single cell. During the elongation stage, millions of fiber cells undergo rapid and simultaneous elongation without concurrent cell division and multicellular development. Moreover, mature fibers are almost purely composed of cellulose which dominates approximately 90 percent of the dry weight of cotton fibers. Consequently, the cotton fiber is a perfect experimental model for the study of the mechanism of plant cell elongation, wall development, cellulose biosynthesis and many other basic biological processes in plants. It is well known that fiber development is composed of four distinct and overlapping stages: fiber cell enlargement and initiation from -3 to 1 day post-anthesis (DPA), fiber elongation and primary cell wall synthesis after anthesis until 25 DPA, secondary cell wall cellulose deposition from 15 DPA to 50 DPA, and fiber cell dehydration and maturation after 45 DPA (Basra and Malik, 1984).

In order to gain an integrated understanding of the developmental mechanisms of fiber at the molecular level, *G. hirsutum* ‘Xuzhou-142’ and a fuzzless-lintless-seed (*fl*) mutant of ‘Xuzhou-142’ were used to study fiber development. About 60 clones preferentially expressed in

Received June 7, 2007; accepted June 24, 2007

Diqiu LIU, Xianlong ZHANG (✉)
National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China
E-mail: xlzhang@mail.hzau.edu.cn

cotton fibers were screened out from 'Xuzhou-142' cDNA library and genes encoding E6, expansin, annexin, actin, α- and β-tubulins were included (Li et al., 2002a). Ji et al. (2003) isolated 172 cDNA preferentially expressed during the cotton fibers' rapid elongation. A lot of genes which accumulated in the rapidly elongating fibers were identified. Some of them were genes encoding the putative vacuolar (H^+)-ATPase catalytic subunit, kinesin-like calmodulin binding protein and arabinogalactan proteins. A huge expressed sequence tags (ESTs) project was completed by Arpat et al. (2004) and 14000 unique genes were assembled from 46603 ESTs which were generated from 7–10 DPA fiber cDNA of *G. arboreum*. Oligonucleotide microarrays analysis shows that more than 2500 "expansion-associated" genes were down regulated when the developmental stage switched from elongation to the secondary cell wall synthesis and a group of 81 novel genes were found preferentially expressed at the beginning of secondary cell wall synthesis. Lee et al. (2006) used *G. hirsutum* genetic standard line 'TM-1' and its naked seed mutant *N1N1* to identify the fiber-associated genes. Many fiber-associated genes were down-regulated in the *N1N1* mutant and may lead to a defective process in the early stages of *N1N1* fiber development. Gene regulation associated with fiber development was influenced in both negative and positive ways in the *N1N1* mutant. Cotton ovules expression profiling was changed by treatment with cycloheximide (Wu et al., 2005). Expression of protein synthesis genes was intensively suppressed, whereas protein phosphorylation genes, transport genes and transcription factors were induced with higher expression than that expressed in the normal ovules. They also found that cycloheximide treatment clearly enriched the number of novel genes. In order to identify genes that might regulate lint fiber initiation, a microarray analysis of multiple mutant lines was done by Wu et al. (2006). Only 13 different genes were down-regulated in some or all of the mutants with an increased expression in wild-type cotton fiber, such as transcription factors MYB and homeodomain genes, a cyclin D gene. Samuel et al. (2006) studied the gene expression during the early stages of fiber cell development by computational and expressional analyses of 32789 high-quality ESTs derived from 'TM-1' immature ovules. Genes encoding putative transcription factors and involved in auxin, brassinosteroid (BR), gibberellic acid (GA), abscisic acid (ABA) and ethylene signaling pathways were obviously assembled in the early stages of fiber development. Shi et al. (2006) found that ethylene played a major role in cotton fiber cell elongation and was more efficacious than BR in regulating fiber cell elongation.

To identify genes preferentially expressed during early and late cotton fiber development, we constructed two fiber suppression subtracted libraries on the basis of PCR-selected subtraction using a pool of non-fiber tissues as the same driver with 10 DPA and 20 DPA fiber cells as testers, respectively (Liu et al., 2006). Finally, 292 clones were identified in both libraries as being preferentially expressed during fiber development through differential screening. Many previously reported cotton fiber related genes are included in the two libraries. Northern blot analysis revealed the genes from 10 DPA fiber subtracted library expressed at high level during early fiber development, such as the genes for putative gibberellin-regulated protein, putative tonoplast intrinsic protein and putative plasma membrane intrinsic protein. On the other hand, those genes screened out of the 20 DPA fiber subtracted library, like genes for arabinogalactan protein and fiber glycosyl hydrolase family 19 proteins, were found to be highly expressed after 15 DPA. What is more, we constructed a normalized fiber cDNA library (from −2 to 25 DPA) of *G. barbadense* 'Pima 3-79' (Tu et al., 2007). More than 9000 plasmids randomly selected from the library were arrayed on the nylon membranes and 'Pima 3-79' fiber RNA from five developmental stages was used to screen the cDNA microarray. The result was that 887 high-quality ESTs were identified as fiber related genes and most of them were assigned to functional categories using gene ontology. The two largest functional categories were energy/carbohydrate metabolism (18%) and protein metabolism (18%). Additionally, subarray analysis of the sequenced cDNAs indicated that most cDNAs were highly expressed from 5 to 15 DPA. For the *G. barbadense* fiber development, we found that expression profiles of genes related to phytohormone synthesis and signal transduction pathway usually varied. The gibberellins (GAs) and their pathways are important in fiber initial and early elongation. What is more, auxin signal transduction pathways and ethylene were revealed to be involved in cotton fiber elongation. The members of *GbCTLs* and *GbTLPs* may participate in the secondary cell wall synthesis.

Isolation and characterization of fiber-specific and fiber-enriched genes represents the first step toward understanding the mechanism of fiber development. As of today, many genes preferentially expressed in cotton fibers have been identified. Some of them are involved in cell elongation. Some play important roles in cell wall biogenesis, while some take part in cellulose synthesis. There are approximately 50 genes cloned from cotton fiber (Table 1). Unfortunately, this is just a vague understanding about the molecular mechanism of fiber development and the integration of these fiber-related genes is still insufficient to elucidate the mechanism of cotton fiber development. Although it is very hard to perform functional verification in cotton plants, tobacco and *Arabidopsis thaliana* could be used as excellent models to verify the putative functions of preferentially expressed cotton fiber genes. The trichome development of tobacco and *Arabidopsis* has some common features

Table 1 A survey of cotton fiber specifically or preferentially expressed genes

gene	expression pattern (DPA)	putative function during fiber development	references
E6	5–28	Probable cell wall structural proteins	John and Crow, 1992
H6	10–30	Expressed predominantly in fiber cells; possibly functions in the secondary cell wall assembly	John and Keller, 1995
B6	15–35	Probable cell wall structural protein	John, 1995
Rac13	5–35	Encoding small GTP-binding proteins analogous to mammalian rac; *Rac13* highly expressed in cotton fibers with the maximal expression occurring at the time of transition between primary and secondary wall synthesis, which may be involved in the signal transduction pathway that controls cytoskeletal organization.	Delmer et al., 1995
Rac9			
GhRac1	8–20	A member of the Rac/Rop GTPase gene family; highly expressed during the elongation stage; may be a potential regulator of fiber elongation by controlling cytoskeletal assembly.	Kim and Triplett, 2004a
GhRacA *GhRacB*	5–16	Two novel *Rac* genes from cotton were expressed in root, hypocotyls, stem, leaf, and fibers; the highest level of transcripts was to accumulate during early cotton fiber development.	Li et al., 2005b
Fbl2A	15–35	Highly expressed during late primary and early secondary wall synthesis stages; *Fb12A* may have structural roles in the secondary wall of fibers.	Rinehart et al., 1996
FS5	6–14	96% nucleotide identity with *E6*	Orford and Timmis, 1997
FS6	6–14	Encoding phospholipid transfer proteins (LTPs)	
FS17	6–14	Encoding a putative polypeptide enriching proline; *FS17* may be a cell-wall protein with a possible N-terminal signal peptide of 23 amino acids.	
FS18	6–20	Encoding a protein of 71 amino acids	
GhEX1	6–24	Encoding an expansin with an N-terminal signal peptide; *GhEX1* is involved in cell wall extension during fiber elongation.	Orford and Timmis, 1998
GhExp1	6–28	Encoding α-expansins; specifically accumulated in developing fibers and involved in cell wall extension during fiber elongation.	Harmer et al., 2002
GhExp6	6–26		
GH3	5–20	Encoding an LTP; possible outer cellular location in fiber cells and specifically expressed in fiber cells; *GH3* may participate in cutin synthesis during fiber development.	Ma et al., 1995
LTP6	5–20	Encoding an LTP, and 64% homologous to *GH3*; *LTP6* may participate in cutin synthesis.	Ma et al., 1997
FSltp1	6–12	Encoding full-length LTPs; expression patterns of them were similar to each other and were expressed in fiber and other non-fiber tissues.	Orford and Timmis, 2000
FSltp2	6–12		
FSltp3	6–18		
GhCelA1	17–35	Encoding the catalytic subunit of cellulose synthase; *GhCelA1* was expressed at high levels during active secondary wall cellulose synthesis in the development of cotton fibers; involved in cellulose biosynthesis.	Pear et al., 1996
GhCAP	10–17	Encoding adenylyl cyclase-associated protein (CAP); *GhCAP* was expressed mainly in young fibers.	Kawai et al., 1998
SS3	5–35	Encoding sucrose synthase; *SS3* controls sucrose mobilization to fiber and cotyledonary cells for biosynthesis of cellulose and seed protein/oil.	Ruan et al., 1997
GhACP	2–20	Encoding acyl carrier protein (ACP), and predominantly expressed during fiber elongation; *GhACP* may play a role in rapidly elongating cotton fibers by involved in the synthesis of membrane lipids.	Song and Allen, 1997
GhMYB1-6	−9–35	Myb-domain genes of R2R3-MYB family; *GhMYB-1*, *-2*, and *-3* were expressed in all examined tissues, but *GhMYB-4*, *-5*, and *-6* showed tissue-specific expression patterns.	Loguercio et al., 1999
GhMYB109	1–8	Encoding MYB transcription factors, and specifically expressed in cotton fiber initial cells and elongating fibers; *GhMYB109* may be involved in cotton fiber initiation and elongation.	Suo et al., 2003
GhMyb7 *GhMyb9*	5–20	Encoding R2R3-MYB proteins, and specifically expressed in elongating fiber and flowers; *GhMyb7* and *GhMyb9* transcriptionally regulated the lipid transfer protein gene *LTP3*.	Hsu et al., 2005
GaMYB2	0–12	Predominantly expressed early in the development of cotton fibers; *GaMYB2* rescued the trichome formation of *Arabidopsis gl1* mutant.	Wang et al., 2004

(*Continued*)

gene	expression pattern (DPA)	putative function during fiber development	references
GhRab11a *GhRab11b*	6	Encoding small GTP-binding proteins; *GhRab11a* and *GhRab11b* are involved in cotton fiber elongation.	John et al., 1999
CFL1	7–24	Homologous to yeast *FKS1*, a putative *β*-1,3-glucan synthase gene; highly expressed during primary wall development; *CFL1* could play a role in callose synthesis.	Cui et al., 2001
GhRGP1	2–40	Encoding putative reversibly glycosylated polypeptide (RGP); *GhRGP1* may be involved in plant cell wall non-cellulosic polysaccharide biosynthesis.	Zhao and Liu, 2001
Ghprp1	5–31	Encoding cell wall proline-rich proteins (PRPs); *GhRGP1* was predominantly expressed in the fiber during the elongation stage.	Tan et al., 2001
GhTUB1	4–21	Encoding *β*-tubulin, and preferentially accumulated at high levels in fiber; *GhTUB1* may play a distinct and required role in fiber development.	Li et al., 2002b
Gh-BTubL	10–20	Encoding β-tubulin; *Gh-BTubL* possibly plays a role in cotton fiber elongation.	Ji et al., 2002
GhACT1	5–28	Encoding an actin; *GhACT1* was involved in fiber elongation, but not in fiber initiation.	Li et al., 2005a
GhWBC1	5–15	Encoding an ATP-binding cassette transporter of the WBC subfamily; the over-expressed *GhWBC1* interferes with Arabidopsis seed and silique development.	Zhu et al., 2003
GhKCBP	10–17	Encoding kinesin-like calmodulin-binding protein (KCBP), and expressed in elongating fibers; *GhKCBP* plays a role in interphase cell growth likely by interacting with cortical microtubules.	Preuss et al., 2003
GhlecRK	6	Encoding a lectin-like protein kinase, and enriched in the boll and shoot development; *GhlecRK* is probably involved in signal transduction during fiber development.	Zuo et al., 2004
GhGlcAT1	5–30	Encoding glucuronosyltranferase, and highly expressed in fiber cells; *GhGlcAT1* may be involved in non-cellulose polysacchrides biosynthesis of the cotton cell wall.	Wu and Liu, 2005
AnnGh1 *AnnGh2*	7–24	Encoding annexin, and highly expressed in the elongation stages of cotton fiber development; *AnnGh1* and *AnnGh2* might play roles in cell elongation.	Potikha and Delmer, 1997; Shin and Brown, 1999
GhPFN1	3–21	Encoding profiling, and primarily expressed during the elongating stages; *GhPFN1* may be involved in the rapid elongation of cotton fibers by promoting actin polymerization.	Wang et al., 2005
GhGLP1	4–20	Encoding a germin-like protein (GLP), and accumulating to their highest levels during the period of fiber expansion; *GhGLP1* may be important for fiber elongation.	Kim and Triplett, 2004b
GhKCR1 *GhKCR2*	3–20	Encoding 3-ketoacyl-CoA reductases and preferentially expressed during the cotton fiber elongation period; *GhKCR1* and *GhKCR2* play an important role in very long chain fatty acids biosynthesis.	Qin et al., 2005
GhRLK1	6–27	Encoding LRR receptor-like protein kinase and expressed during fiber secondary wall synthesis; *GhRLK1* is likely involved in signal transduction pathway of cotton fiber development.	Li et al., 2005c
GhBRI1-1 *GhBRI1-2*	0–30	Similar to Arabidopsis brassinosteroid insensitive 1 genes and strongly expressed in rapidly expanding regions including fibers and hypocotyls; *GhBRI1-1* and *GhBRI1-2* might be functional components of the BR signaling pathway.	Sun et al., 2004
GhCTL1 *GhCTL2*	15–31	Encoding chitinase-like proteins and preferentially expressed during secondary wall deposition in cotton fiber; *GhCTL1* and *GhCTL2* are essential for cellulose synthesis in primary and secondary cell walls.	Zhang et al., 2004
GhZIP	5–15	Encoding the basic leucine zipper (bZIP) protein and expressed in ovule and fiber cells; *GhZIP* might be involved in the transcription regulation of genes during cotton fiber elongation.	Jiang et al., 2004
GhTTG1 *GhTTG3*	6–24	Encoding WD-repeat proteins; expressed in developing fibers, ovules and other tissues; *GhTTG1* and *GhTTG3* can complement *ttg1* mutant phenotypes and restores trichome formation; may be pivotal in fiber development.	Humphries et al., 2005

with cotton fiber development. There is a long way to go before we can comprehensively understand the mechanism of cotton fiber development.

3 Isolation of stress related genes and expression profiles of cotton in response to the wilt pathogens

There are many stress factors that threaten cotton plant development, such as fungal and bacterial pathogens, a variety of insects, drought, heat, salt, water, and so on. Each of these stress factors can cause a severe decrease in fiber production and damages the fiber and cottonseed oil quality. Understanding the molecular mechanisms of cotton defensive responses during stress will facilitate the development of desirable cotton cultivars by genetic engineering. Some genes with full-length sequences related to stress responses have been isolated from cotton. Their expression patterns and putative functions are described in Table 2. There are a few genetically modified stress-tolerant cultivars that have been released for field production. We believe that, as the work progresses, the stress-tolerance improvement of cotton using biotechnology would be a key area both in research and in application.

The NBS-LRR gene family is composed of a large multigene family, whose members encode the nucleotide-binding site (NBS) and leucine-rich repeat (LRR) motif. About 75% of the plant disease-resistance genes (R-genes) belong to this gene family. NBS and LRR are short and conservative motifs of NBS-LRR genes. In cotton, the first 33 resistance gene analogues (RGAs) were cloned from root-knot nematode (RKN) resistant *G. hirsutum* line 'M-249' (Tan et al., 2003). Homologous searching in GenBank revealed that cotton RGAs were somewhat identical to RGAs and R-genes from other plant species and all RGAs had conserved NBS domain. In our laboratory, 31 RGAs were isolated from *G. Barbadense* line 'Pima 90' (Tu et al., 2003). The 31 RGAs plus those RGAs from Tan et al. (2003) were multi-aligned and they were divided into two classes. Most of upland cotton RGAs were grouped into Class I, but all sea-island cotton RGAs, as well as several upland cotton RGAs, fell into Class II. He et al. (2004) identified 61 RGAs from *G. hirsutum* RKN resistance line 'Auburn 634'. Among the 61 RGAs, 15 were mapped to 7 chromosomes or linkage groups. The distribution of RGAs between A and D subgenomes seemed biased and the majority of RGAs were mapped to the A subgenomes. Recently, Hinchliffe et al. (2005) isolated 55 novel RGAs from RKN resistance genetic source *G. hirsutum* 'Auburn 634 RNR'. In total, 61 RGA-sequence-tagged site (STS) primers were designed according to the consensus sequences or individual RGA sequences. Nine RGA markers were mapped to a couple of homologous chromosomes 12 and 26.

F. oxysporum f. sp. *vasinfectum* and *V. dahliae* are the two major fungal pathogens of cotton. Both of them are root-invading vascular pathogens with similar life cycles in the plant. These two diseases cause the wilting of cotton in many cotton-growing areas and greatly influence the fiber yield and quality. Although conventional breeding has done much to improve disease tolerances, these two fungal diseases are still the important factors causing yield losses in wide areas.

Table 2 The cotton genes related to stress response

gene	homologous sequences, expression patterns, and putative functions	references
cMsrA	Similar to *Brassica napus* methionine sulfoxide reductase; enriched later and more weakly in resistant variety than the sensitive under the salt and pathogens stress; *cMsrA* might protect the cells against oxidative damage caused by stress.	Zhao et al., 2003
peroxidase gene family	Encoding peroxidase; induced by *Xanthomonas campestris* pv. *Malvacearum* and somewhat down-regulated during pathogen infection; members of the gene family may have various functions in the defensive response to *Xanthomonas* infections.	Assigbetsé et al., 1999; Delannoy et al., 2003
dHG-6-OMT	Encoding desoxyhemigossypol-6-O-methyltransferase; methylating the phenolic group of dHG at the 6-position; *dHG-6-OMT* may alter the resistance of terpenoids.	Liu et al., 2002
GbERF	Encoding ethylene responsive element binding factors; *GbERF* can be induced by ethylene, *V. dahliae* infection, salt, cold, and drought stresses; response to ethylene, ABA, environmental stresses, and pathogen attack.	Qin et al., 2004
GhHb1	Encoding nonsymbiotic hemoglobin; inducing the expression in root by *V. dahliae* and H_2O_2; *GhHb1* may be involved in the defensive response of cotton to *V. dahliae* invasion.	Qu et al., 2005
GaPR-10	Encoding pathogenesis-related class 10 protein and inducing the expression by *V. dahliae* and jasmonate; *GaPR-10* might be involved in activated responses of plants to pathogens.	Zhou et al., 2002
MIC-3	No significant homology to the known sequences; inducing the expression in root by nematode; *MIC-3* may play a critical role in the resistance response to root-knot nematode.	Zhang et al., 2002
GhNHX1	Encoding plant vacuolar-type Na^+/H^+ antiporters; strongly induced by salt stress and abscisic acid in cotton seedlings; *GhNHX1* plays an important role in salt tolerance of cotton.	Wu et al., 2004
TPS	Encoding trehalose 6-phosphate-synthase; increasing the expression in drought stressed leaves and roots; *TPS* participates in trehalose biosynthesis and stress signal transduction.	Kosmas et al., 2006

With the development of the biotechnology, the defense mechanisms employed in cotton plants had been partially known at the molecular level. Gene expression profiles were different between *F. oxysporum* f. sp. *vasinfectum* infected cotton root and hypocotyl tissues (Dowd et al., 2004). Defense-related gene expression is induced in hypocotyls while they appeared to be constitutively expressed in the infected root tissues. It is interesting that the number of plant genes induced in roots was less than that of the repressed. It might be the strategy for pathogen to suppress the host responses during the infection. The pathogenesis-related genes, gossypol and lignans biosynthesis genes were the major infection-induced genes. In order to discover the molecular character of the defensive response of the cotton plants to *V. dahliae*, some progress has been made through constructing cDNA libraries enriching defense-response related genes. Hill et al. (1999) identified 86 cDNA clones as *V. dahliae* up-regulated genes. These cDNAs includes phenylalanine ammonia-lyase, putative oxidoreductase, translationally controlled tumour protein, glucose-6-phosphate/phosphate-translocator, 14-3-3-like protein, pathogenesis-related (PR) proteins, and so on. Meanwhile, differentially expressed ESTs involved in the defense reaction of sea-island cotton to *V. dahliae* were isolated by suppression subtractive hybridization (SSH) (Zuo et al., 2005). Among the 147 differentially expressed ESTs, 131 ESTs were considered as over-expressed and 16 ESTs were down-regulated. As for our research group, Zhu et al. (2005) also constructed an SSH library of sea-island cotton upon the infection of *V. dahliae*. There are 78 clones which were up-regulated and were putatively involved in the defense response. Sequence and functional analysis demonstrated that three kinds of genes were highly expressed during *V. dahliae* infection. They were the root-knot nematode-induced gene *MIC-3*, PR protein genes, and *Arabidopsis* disease resistance-responsive family protein genes.

4 Isolation genes involved in the biosynthesis of cotton terpenoids and fatty acids

Cotton plants accumulate gossypol and other terpenoids in the pigment glands of foliage and seed. Gossypol is the major terpenoid in the seed and root. However, there is a unique group of terpenes contained in foliage glands except for gossypol. These compounds are elicited in response to bacterial or fungal infection and are thought to play a major role in defending cotton against pathogens and insects. Apparently, understanding the biosynthesis of the terpenoids will result in utilization of these natural defense compounds to increase cotton resistance toward insects and fungal pathogens. Martina et al. (2003) proposed a putative biosynthetic pathway for cotton terpenoids. As for cloning the key genes that control the biosynthesis of cotton terpenoids, several reports have indicated progress being made. A (+)-δ-cadinene synthase gene was isolated from *G. arboreum* suspension cells infected by *V. dahliae* and the translated product of the gene functioned as farnesyldiphosphate (FPP) cyclase that produced the (+)-δ-cadinene (Chen et al., 1995). A year later, more cDNAs encoding (+)-δ-cadinene synthase were reported by Chen et al. (1996). Obviously, the cotton (+)-δ-cadinene synthase genes are composed of a gene family. Xu et al. (2004) isolated a transcription factor gene *GaWRKY1* from *G. arboreum* which can regulate the (+)-δ-cadinene synthase-A. Expression of *GaWRKY1* was down-regulated in glandless or reduced-gossypol cotton cultivars. As a transcription factor of (+)-δ-cadinene synthase, *GaWRKY1* plays an important role in regulation of sesquiterpene biosynthesis in cotton. A cytochrome P450 mono-oxygenase was found to catalyse the hydroxylation of (+)-δ-cadinene and form 8-hydroxy-(+)-δ-cadinene (Luo et al., 2001).

The fatty acid is the major structural components of plant cells and the biosynthetic pathway of fatty acid is a primary metabolic pathway found in every cell. As of present, several key genes involved in fatty acid biosynthetic pathway have been isolated and characterized from cotton. Two genes, *ghSAD-1* and *ghFAD2-1*, were cloned from the upland cotton and found to be the key enzyme determining the fatty acid composition of cottonseed oil (Liu et al., 1996; Liu et al., 1999). They encoded cotton stearoyl-acyl-carrier protein $\Delta 9$-desaturase and microsomal ω-6 fatty acid desaturase, respectively. The *ghFAD2-1* was highly and specifically expressed in the cottonseed. What is more, a palmitoyl-acyl carrier protein thioesterase (PATE) gene was screened out from a cotton genomic library (Yoder et al., 1999). Another cDNA encoding PATE was obtained by screening a cDNA library with a heterologous *Arabidopsis FatB1* probe (Pirtle et al., 1999). The latter PATE gene was expressed in *E. coli* and a series of assays demonstrated that this gene encodes a functional thioesterase with a preference for saturated acyl-ACPs (FatB) over unsaturated acyl-ACPs (FatA) (Huynh et al., 2002). Pirtle et al. (2001) characterized a cotton gene, *FAD2-3*, which encodes an enzyme that catalyzes the desaturation of oleate into linoleate.

5 Isolation and characterization of genes associated to cotton somatic embryogenesis

In order to accelerate the production of transgenic cotton varieties, it is imperative to improve the embryogenic competence and regenerability of a wider range of cotton cultivars. The identification and isolation of the vital genes of somatic embryogenesis (SE) are of great importance. Our research group has made advances in the molecular dissection cotton SE. A systematic and comprehensive analysis of the genes involved in the process of cotton SE had been preformed by Zeng et al. (2006). Totally 242 SE related genes were isolated by suppressive subtraction hybridization and macroarray. More than half had not been identified previously as SE-related genes such as genes involved in transcription, post-transcription, and transportation. Additionally, many novel genes were found involved in cotton somatic embryogenesis. Our study suggests that a complicated and concerted mechanism involved with multiple cellular pathways is responsible for cotton SE.

Acknowledgements The financial support of this research from the National High-Tech Research and Program of China (No. 2006AA10A109-4) and the National Natural Science Foundation of China (Grant No. 30671323) were appreciated.

References

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

Assigbetsé K, Cuny G, Valette V, Delannoy E, Bresson E, Jalloul A, Daniel J F, Geiger J P, Nicole M (1999). Cloning and characterization of a bacterial-induced peroxidase encoding cDNA from cotton (GenBank accession No. AF155124). Plant Physiol, 121: 312

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

Chen X Y, Chen Y, Heinstein P, Davisson V J (1995). Cloning expression and characterization of (+)-δ-cadinene synthase: a catalyst for cotton phytoalexin biosynthesis. Arch Biochem Biophys, 324: 255–266

Chen X Y, Wang M, Chen Y, Davisson V J, Heinstein P (1996). Cloning and heterologous expression of a second (+)-δ-cadinene synthase from *Gossypium arboreum*. J Nat Prod, 59: 944–951

Cui X, Shin H, Song C, Laosinchai W, Amano Y, Brown R M Jr (2001). A putative plant homolog of the yeast β-13-glucan synthase subunit *FKS1* from cotton (*Gossypium hirsutum* L.) fibers. Planta, 213: 223–230

Delannoy E, Jalloul A, Assigbetse K, Marmey P, Geiger J P, Lherminier J, Daniel J F, Martinez C, Nicole M (2003). Activity of Class III peroxidases in the defense of cotton to bacterial blight. Mol Plant Microbe Interactions, 16: 1030–1038

Delmer D P, Pear J R, Andrawis A, Stalker D M (1995). Genes encoding small GTP-binding proteins analogous to mammalian Rac are preferentially expressed in developing cotton fibers. Mol General Genet, 248: 43–51

Dowd C, Wilson I W, McFadden H (2004). Gene expression profile changes in cotton root and hypocotyl tissues in response to infection with *Fusarium oxysporum* f. sp. *Vasinfectum*. Mol Plant Microbe Interactions, 17: 654–667

Harmer S E, Orford S J, Timmis J N (2002). Characterization of six alpha-expansin genes in *Gossypium hirsutum* (upland cotton). Mol Genet Genomics, 268: 1–9

He L, Du C, Covaleda L, Xu Z, Robinson A F, Yu J Z, Kohel R J, Zhang H B (2004). Cloning characterization and evolution of the NBS-LRR-encoding resistance gene analogue family in polyploid cotton (*Gossypium hirsutum* L.). Mol Plant Microbe Interactions, 17: 1234–1241

Hill M K, Lyon K J, Lyon B R (1999). Identification of disease response genes expressed in *Gossypium hirsutum* upon infection with the wilt pathogen *Verticillium dahliae*. Plant Mol Biol, 40: 289–296

Hinchliffe D J, Lu Y Z, Potenza C, Segupta-Gopalan C, Cantrell R G, Zhang J F (2005). Resistance gene analogue markers are mapped to homologous chromosomes in cultivated tetraploid cotton. Theor Appl Genet, 110: 1074–1085

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

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

Huynh T T, Pirtle R M, Chapman K D (2002). Expression of a *Gossypium hirsutum* cDNA encoding a FatB palmitoylacyl carrier protein thioesterase in *Escherichia coli*. Plant Physiol Biochem, 40: 1–9

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

Ji S J, Lu Y C, Li J, Wei G, Liang X J, Zhu Y X (2002). A β-tubulin-like cDNA expressed specifically in elongating cotton fibers induces longitudinal growth of fission yeast. Biochem Biophys Res Commun, 296: 1245–1250

Jiang J X, Guo W Z, Zhang T Z (2004). Cloning and expression analysis of a bZIP cDNA in *Gossypium hirsutum* L. Acta Genetica Sinica, 31: 616–621

John C, Hee J K, Triplett B A (1999). Characterization of two cDNAs encoding small GTP-binding proteins (Accession Nos. AF165095 and AF165096) from immature cotton (*Gossypium hirsutum* L.) locules. Plant Physiol, 121: 685

John M E (1995). Characterization of a cotton (*Gossypium hirsutum* L.) fiber-mRNA (*Fb-b6*). Plant Physiol, 107: 1477–1478

John M E, Crow L J (1992). Gene expression in cotton (*Gossypium hirsutum* L.) fiber: cloning of the mRNAs. Proc Natl Acad Sci USA, 89: 5769–5773

John M E, Keller G (1995). Characterization of mRNA for a proline-rich protein of cotton fibers. Plant Physiol, 108: 669–676

Kawai M, Aotsuka S, Uchimiya H (1998). Isolation of a cotton CAP gene: a homologue of adenylyl cyclase-associated protein highly expressed during fiber elongation. Plant Cell Physiol, 39: 1380–1383

Kim H J, Triplett B A (2004a). Characterization of GhRac1 GTPase expressed in developing cotton (*Gossypium hirsutum* L.) fibers. Biochim Biophys Acta, 1679: 214–221

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

Kosmas S A, Argyrokastritis A, Loukas M G, Eliopoulos E, Tsakas S, Kaltsikes P J (2006). Isolation and characterization of drought-related trehalose 6-phosphate-synthase gene from cultivated cotton (*Gossypium hirsutum* L.). Planta, 223: 329–339

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

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

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

Li X B, Fan X P, Wang X L, Cai L, Yang W C (2005a). The cotton *ACTIN1* gene is functionally expressed in fibers and participates in fiber elongation. Plant Cell, 17: 859–875

Li X B, Xiao Y H, Luo M, Hou L, Li D M, Luo X Y, Pei Y (2005b). Cloning and expression analysis of two Rac genes from cotton (*Gossypium hirsutum* L.). Acta Genetica Sinica, 32: 72–78 (in Chinese)

Li Y L, Sun J, Xia G X (2005c). Cloning and characterization of a gene for an LRR receptor-like protein kinase associated with cotton fiber development. Mol Genet Genomics, 273: 217–224

Liu D, Zhang X, Tu L, Zhu L, Guo X (2006). Isolation by suppression-subtractive hybridization of genes preferentially expressed during early and late fiber development stages in cotton. Mol Biol, 40: 825–834

Liu J, Benedict C R, Stipanovic R D, Magill C W, Bell A A (2002). Cloning and expression of desoxyhemigossypol-6-O- methyltransferase from cotton (*Gossypium barbadense*). J Agric Food Chem, 50: 3165–3172

Liu Q, Singh S P, Brubaker C L, Sharp P J, Green A G, Marshall D R (1999). Molecular cloning and expression of a cDNA encoding a microsomal ω-6 fatty acid desaturase in cotton (*Gossypium hirsutum* L.). Australian J Plant Physiol, 26: 101–106

Liu Q, Singh S P, Sharp P J, Green A G, Marshall D R (1996). Nucleotide sequence of a cDNA from *Gossypium hirsutum* encoding a stearoyl-acyl carrier protein desaturase (accession No. X95988) (OGR 96-012). Plant Physiol, 110: 1436

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.). Mol Gen Genet, 261: 660–671

Luo P, Wang Y H, Wang G D, Essenberg M, Chen X Y (2001). Molecular cloning and functional identification of (+)-δ-cadinene-8-hydroxylase, a cytochrome P450 mono-oxygenase (CYP706B1) of cotton sesquiterpene biosynthesis. Plant J, 28: 95–104

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

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

Martina G S, Liua J, Benedicta C R, Stipanovicb R D, Magillc C W (2003). Reduced levels of cadinane sesquiterpenoids in cotton plants expressing antisense (+)-δ-cadinene synthase. Phytochemistry, 62: 31–38

Orford S J, Timmis J N (1997). Abundant mRNAs specific to the developing cotton fiber. Theor Appl Genet, 94: 909–918

Orford S J, Timmis J N (1998). Specific expression of an expansin gene during elongation of cotton fibers. Biochim Biophys Acta, 1398: 342–346

Orford S J, Timmis J N (2000). Expression of a lipid transfer protein gene family during cotton fiber development. Biochim Biophys Acta, 1483: 275–284

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

Pirtle I L, Kongcharoensuntorn W, Nampaisansuk M, Knesek J E, Chapman K D, Pirtle R M (2001). Molecular cloning and functional expression of the gene for a cotton Δ-12 fatty acid desaturase (FAD2). Biochim Biophys Acta, 1522: 122–129

Pirtle R M, Yoder D W, Huynh T T, Nampaisansuk M, Pirtle I L, Chapman K D (1999). Characterization of a palmitoyl-acyl carrier protein thioesterase (FatB1) in cotton. Plant Cell Physiol, 40: 155–163

Potikha T S, Delmer D P (1997). cDNA clones for annexin *AnnGh1* (accession No. U73746) and *AnnGh2* (accession No. U73747) from *Gossypium hirsutum* (cotton) (PGR 97-003). Plant Physiol, 113: 305

Preuss M L, Delmer D P, Liu B (2003). The cotton kinesin-like calmodulin-binding protein associates with cortical microtubules in cotton fibers. Plant Physiol, 132: 154–160

Qin J, Zhao J Y, Zuo K J, Cao Y F, Ling H, Sun X F, Tang K X (2004). Isolation and characterization of an ERF-like gene from *Gossypium barbadense*. Plant Sci, 167: 1383–1389

Qin Y M, Pujol F M, Shi Y H, Feng J X, Liu Y M, Kastaniotis A J, Hiltunen J K, Zhu Y X (2005). Cloning and functional characterization of two cDNAs encoding NADPH dependent 3-ketoacyl-CoA reductased from developing cotton fibers. Cell Res, 15: 465–473

Qu Z L, Wang H Y, Xia G X (2005). *GhHb1*: A nonsymbiotic hemoglobin gene of cotton responsive to infection by *Verticillium dahliae*. Biochim Biophys Acta, 1730: 103–113

Rinehart J A, Petersen M W, John M E (1996). Tissue-specific and developmental regulation of cotton gene *FbL2A*: Demonstration of promoter activity in transgenic plants. Plant Physiol, 112: 1331–1341

Ruan Y L, Chourey P S, Delmer D P, Perez-Grau L (1997). The differential expression of sucrose synthase in relation to diverse patterns of carbon partitioning in developing cotton seed. Plant Physiol, 115: 375–385

Samuel Y S, Cheung F, Lee J J, Ha M, Wei N E, Sze S H, Stelly D M, Thaxton P, Triplett B, Town C D, Jeffrey Chen Z (2006). Accumulation of genome-specific transcripts transcription factors and phytohormonal regulators during early stages of fiber cell development in allotetraploid cotton. Plant J, 47: 761–775

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. Plant Cell, 18: 651–664

Shin H, Brown R M Jr (1999). GTPase activity and biochemical characterization of a recombinant cotton fiber annexin. Plant Physiol, 119: 925–934

Song P, Allen R D (1997). Identification of a cotton fiber-specific acyl carrier protein cDNA by differential display. Biochim Biophys Acta, 1351: 305–312

Sun Y, Fokar M, Asami T, Yoshida S, Allen R D (2004). Characterization of the brassinosteroid insensitive 1 genes of cotton. Plant Mol Biol, 54: 221–232

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

Tan H, Callahan F E, Zhang X D, Karaca M, Saha S, Jenkins J N, Creech R G, Ma D P (2003). Identification of resistance gene analogs in cotton (*Gossypium hirsutum* L.). Euphytica, 134: 1–7

Tan H, Creech R G, Jenkins J N, Chang Y F, Ma D P (2001). Cloning and expression analysis of two cotton (*Gossypium hirsutum* L.) genes encoding cell wall proline-rich proteins. DNA Seq, 12: 367–380

Tu L L, Zhang X L, Liang S G, Liu D Q, Zhu L F, Zeng F C, Nie Y C, Guo X P, Deng F L, Tan J F, Xu L (2007). Genes expression analyses of sea-island cotton (*Gossypium barbadense* L.) during fiber development. Plant Cell Rep, 26: 1309–1320

Tu L L, Zhang X L, Zhu L F, Nie Y C, Guo X P (2003). Origin diversity and evolution of NBS-type disease-resistance gene analogues in sea-island cotton (*Gossypium barbadense* L.). Acta Genetic Sinica, 30: 1071–1077 (in Chinese)

Wang H Y, Yu Y, Chen Z L, Xia G X (2005). Functional characterization of *Gossypium hirsutum* profilin 1 gene (*GhPFN1*) in tobacco suspension cells. Planta, 222: 594–603

Wang S, Wang J W, Yu N, Li C H, Luo B, Gou J Y (2004). Control of plant trichomes development by a cotton fiber MYB gene. Plant Cell, 16: 2323–2334

Wu C A, Yang G D, Meng Q W, Zheng C C (2004). The cotton *GhNHX1* gene encoding a novel putative tonoplast Na^+/H^+ antiporter plays an important role in salt stress. Plant Cell Physiol, 45: 600–607

Wu Y T, Liu J Y (2005). Molecular cloning and characterization of a cotton glucuronosyltranferase gene. J Plant Physiol, 162: 573–582

Wu Y, Machado A C, White R G, Llewellyn D J, Dennis E S (2006). Expression profiling identifies genes expressed early during lint fibre initiation in cotton. Plant Cell Physiol, 47: 107–127

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

Xu Y H, Wang J W, Wang S, Wang J Y, Chen X Y (2004). Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-δ-cadinene synthase-A. Plant Physiol, 135: 507–515

Yoder D W, Nampaisansuk M, Pirtle I L, Chapman K D, Pirtle R M (1999). Molecular cloning and nucleotide sequence of a gene encoding a cotton palmitoyl-acyl carrier protein thioesterase. Biochim Biophys Acta, 1446: 403–413

Zeng F, Zhang X, Zhu L, Tu L, Guo X, Nie Y (2006). Isolation and characterization of genes associated to cotton somatic embryogenesis by suppression subtractive hybridization and macroarray. Plant Mol Biol, 60: 167–183

Zhang D, Hrmova M, Wan C H, Wu C, Balzen J, Cai W, Wang J, Densmore L D, Fincher G B, Zhang H, Haigler C H (2004). Members of a new group of chitinase-like genes are expressed preferentially in cotton cells with secondary walls. Plant Mol Biol, 54: 353–372

Zhang X D, Callahan F E, Jenkins J N, Ma D P, Karaca M, Saha S, Creech R G (2002). A novel root-specific gene *MIC-3* with increased expression in nematode-resistant cotton (*Gossypium hirsutum* L.) after root-knot nematode infection. Biochim Biophys Acta, 1576: 214–218

Zhao G R, Liu J Y (2001). Isolation of a cotton RGP gene: a homolog of reversibly glycosylated polypeptide highly expressed during fiber development. Biochim Biophys Acta, 1574: 370–374

Zhao J, Zuo K, Wang J, Cao Y, Zhang L, Tang K (2003). cDNA cloning and characterization of a cotton peptide methionine sulfoxide reductase (cMsrA). DNA Seq, 14: 303–310

Zhou X J, Lu S, Xu Y H, Wang J W, Chen X Y (2002). A cotton cDNA (*GaPR-10*) encoding a pathogenesis-related 10 protein with in vitro ribonuclease activity. Plant Sci, 162: 629–636

Zhu L F, Tu L L, Zhang X L, Nie Y C, Guo X P, Xia Q Z (2005). Construction and analysis of SSH library of *Gossypium barbadense* upon infection with *Verticillium dahliae*. Acta Genetica Sinica, 32: 528–532 (in Chinese)

Zhu Y Q, Xu K X, Luo B, Wang J W, Chen X Y (2003). An ATP-binding cassette transporter *GhWBC1* from elongating cotton fibers. Plant Physiol, 133: 580–588

Zuo K J, Wang J, Wu W, Chai Y, Sun X F, Tang K X (2005). Identification and characterization of differentially expressed ESTs of *Gossypium barbadense* infected by *Verticillium dahliae* with suppression subtractive hybridization. Mol Biol, 39: 191–199

Zuo K, Zhao J, Wang J, Sun X, Tang K (2004). Molecular cloning and characterization of *GhlecRK* a novel kinase gene with lectin-like domain from *Gossypium hirsutum*. DNA Seq, 15: 58–65