TITLE: **Functional Genomic Analysis of Early Events in Cotton Fiber Development**

DISCIPLINE: **Genomics and Biotechnology**

AUTHORS: Z.J. Chen (Corresponding author)
Institute for Cellular and Molecular Biology
The University of Texas at Austin
One University Station A-4800
Austin, TX 78712-0159
Phone: 512-475-9327
Fax: 512-471-2149
Email: zjchen@mail.utexas.edu

J.J. Lee
A.W. Woodward
Z. Han
M. Ha
E. Lackey
Section of Molecular Cell and Developmental Biology
Institute for Cellular and Molecular Biology, and Center for Computational Biology and Bioinformatics
The University of Texas at Austin
One University Station A4800
Austin, TX 78712

ACKNOWLEDGEMENT: We thank other members of the NSF-funded cotton fiber genomics project, Barbara Triplett, David Stelly, Peggy Thaxton, and Sing-Hoi Sze, for valuable comments on the research. This work was supported by grants from the National Science Foundation Plant Genome Research Program (DBI0624077), the Cotton Incorporated (04-5555), and the Texas Higher Education Coordinating Board Advanced Technology Program (000517-0218-2001) to Z. Jeffrey Chen.

ABBREVIATIONS: Cotton, Genomics, Microarrays, Gene expression, Polyploidy, Fiber

Functional Genomic Analysis of Early Events in Cotton Fiber Development

ABSTRACT

Fiber cell initiation is a complex process involving many pathways, including various signal transduction and transcriptional regulation components. Here we report gene expression patterns during fiber development using expressed sequence tags (ESTs) and microarrays. We found many parallels of gene expression between *Arabidopsis* trichome development from the leaf epidermis and cotton fiber development from the ovule epidermis. Transcript profiling and ovule culture experiments both indicate that several phytohormones mediate cotton fiber development. Auxin and gibberellins appear to promote early stages of fiber initiation; ethylene- and brassinosteroid-related genes are upregulated during the fiber elongation phase. Interestingly, the phytohormone-related genes were induced prior to the activation of *MYB*-like genes, suggesting an important role of phytohormones in cell fate determination. Moreover, AA subgenome ESTs of all functional classifications including cell cycle control and transcription factor activity were selectively enriched in *G. hirsutum* L., an allotetraploid derived from polyploidization between AA and DD genome species, a result consistent with the production of long lint fibers in AA genome species. These results suggest a mode of temporal and genomic-specific regulation of the genes involved in phytohormonal pathways and transcriptional networks during early stages of fiber development.

Cotton fibers form through a fascinating process of single-cell development involving numerous modifications to fundamental cell metabolism. As Kim and Triplett (2001) aptly stated, cotton fiber is among most exaggerated plant cell types and mature cotton fiber can reach a "length of 6 cm or one-third the height of an *Arabidopsis* plant". Thus, cotton fiber is considered to be a model system for studying cell fate, differentiation, elongation, and cell-wall biosynthesis. The importance of cotton fiber development to fundamental biological research is magnified by the fact that cotton is the leading natural fiber crop in the world and a mainstay of the US economy. The economic value of cotton fiber grown in the US is typically ~$6 billion/year plus $0.5 billion/year for cottonseed oil and meal (Smith, 1999; NCCA, 2001). Cotton fiber exports account for ~$2 billion/year of the US trade surplus. Business revenue stimulated by the crop is estimated at ~$120 billion. Texas and California each produces ~30% of the US crop. Fiber quality (length, strength, and micronaire) is a major economic factor especially given growing competition with synthetic fibers.

The genus *Gossypium* occurs naturally throughout tropical and subtropical regions, and includes about 45 species split across two ploidy levels, diploid ($2n = 2x = 26$) and tetraploid ($2n = 4x = 52$) (Wendel, 1989; Percival et al., 1999; Wendel and Cronn, 2003). An important event in cotton genome evolution was the spontaneously formation of allopolyploid cotton that has been subsequently selected and domesticated as modern cultivated cotton. The progenitors of allotetraploid cotton are most closely related to 'AA' and 'DD' extant diploid species. This polyploidization event occurred ~1.5 million years ago (Mya), and the AADD allotetraploids diverged into five species that are distributed throughout the New World and the rest of the globe (Wendel, 1989; Percival et al., 1999; Wendel and Cronn, 2003; Desai et al., 2006). Among the extant diploids resembling the presumed ancestors of tetraploid cotton, the AA progenitor

species produce both lint (long) fibers that are spinnable into yarn and shorter fibers called fuzz. In contrast, the DD genome progenitor species produce very few lint fibers that are initiated pre-anthesis, but are much shorter in length than the lint fibers of the AA genome progenitor (Percival et al., 1999; Applequist et al., 2001).

The molecular basis of the fiber initiation stage remains largely mysterious. About 15 to 25% of the protodermal cells differentiate and develop lint fibers (Basra and Malik, 1984; Tiwari and Wilkins, 1995; Kim and Triplett, 2001). Fiber cell initiation process is quasi-synchronous and rapid. Cell fate determination undoubtedly precedes the formation of morphologically visible fiber cell initials. Phytohormone treatments from 2-3 days preanthesis to the day of anthesis induce fiber production on cultured ovules (Beasley and Ting, 1974), but few fibers are produced if phytohormones are applied after the day of anthesis (Graves and Stewart, 1988). Recent studies indicate cotton MYB transcription factors, sucrose synthetase, brassinosteroid, and ethylene are involved in fiber cell initiation and elongation (Loguercio et al., 1999; Suo et al., 2003; Sun et al., 2004; Wang et al., 2004; Shi et al., 2006; Wu et al., 2006). The fiber elongation phase is perhaps the best-studied period of fiber development (Kim and Triplett, 2001; Arpat et al., 2004; Haigler et al., 2005; Wilkins and Arpat, 2005). The temporal boundaries separating fiber elongation from the prior initiation phase and the subsequent maturation phase are not discrete, but elongation is generally defined to be from 5 to 20 DPA. Fiber cell initiation and elongation are orchestrated in each fiber cell through regulation of gene expression and intercellular signaling pathways and many metabolic activities (Kim and Triplett, 2001).

RESULTS AND DISCUSSION

We examined molecular and cellular events of fiber cell development in the naked seed mutant (*N1N1*) and its isogenic line of cotton (*Gossypium hirsutum* L. cv. Texas Maker-1, TM-1). The dominant mutation not only delayed the process of fiber cell formation and elongation but also reduced the total number of fiber cells, resulting in sparsely distributed short fibers. Gene expression changes in TM-1 and *N1N1* mutant lines among four tissues were analyzed using spotted cotton oligo-gene microarrays. Using the *Arabidopsis* genes, we selected and designed ~1,334 70-mer oligos from a subset of cotton fiber ESTs. Statistical analysis of the microarray data indicates that the numbers of significantly differentially expressed genes were 856 in the leaves compared to the ovules (3 days post-anthesis, DPA), 632 in the petals relative to the ovules (3 DPA), and 91 in the ovules at 0 compared to 3 DPA, all in TM-1. Moreover, 117 and 30 genes were expressed significantly different in the ovules at 3 and 0 DPA, respectively, between TM-1 and *N1N1*. Quantitative RT-PCR analysis of 23 fiber-associated genes in seven tissues including ovules, fiber-bearing ovules, fibers, and non-fiber tissues in TM-1 and *N1N1* indicates a mode of temporal regulation of the genes involved in transcriptional and translational regulation, signal transduction, and cell differentiation during early stages of fiber development. Suppression of the fiber-associated genes in the mutant may suggest that the *N1N1* mutation disrupts temporal regulation of gene expression, leading to a defective process of fiber cell elongation and development.

We also compared and analyzed expression analyses of 32,789 high-quality ESTs derived from *Gossypium hirsutum* L. Texas Marker-1 (TM1) immature ovules (GH_TMO) (Yang et al., 2006). The ESTs were assembled into 8,540 unique sequences including 4,036 tentative consensus sequences (TCs) and 4,504 singletons, representing ~15% unique sequences in the

cotton EST collection. Compared to ~178,000 existing ESTs derived from elongating fibers and non-fiber tissues, GH_TMO ESTs showed a significant increase in the percentage of the genes encoding putative transcription factors such as MYB and WRKY and the genes encoding predicted proteins involved in auxin, brassinosteroid (BR), gibberellic acid (GA), abscisic acid (ABA) and ethylene signaling pathways. Cotton homologues related to *MIXTA*, *MYB5*, *GL2* and eight genes in auxin, BR, GA and ethylene pathways were induced during fiber cell initiation but repressed in the naked seed mutant (*N1N1*) that is impaired in fiber formation. The data agree with the known roles of MYB and WRKY transcription factors in *Arabidopsis* leaf trichome development and the well-documented phytohormonal effects on fiber cell development in immature cotton ovules cultured *in vitro*. Moreover, the phytohormone-related genes were induced prior to the activation of *MYB*-like genes, suggesting an important role of phytohormones in cell fate determination. Significantly, AA subgenome ESTs of all functional classifications including cell cycle control and transcription factor activity were selectively enriched in *G. hirsutum* L., an allotetraploid derived from polyploidization between AA and DD genome species. The data are consistent with the production of long lint fibers in AA genome species.

To further dissect gene expression changes during early stages of fiber development, we developed a new set of spotted oligo-gene microarrays in cotton using ~210,000 ESTs including 33,000 derived from fiber-bearing ovules (Udall et al., 2007). Using the microarrays, we have examined gene expression changes in integuments and epidermal fibers that are separated using a laser dissection method. The expression data obtained in early stages of fiber cell development will be reported and discussed.

REFERENCES

Applequist, W.L., R. Cronn, and J.F. Wendel. 2001. Comparative development of fiber in wild and cultivated cotton. Evol. Dev. 3: 3-17.

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

Basra, A., and C.P. Malik. 1984. Development of the cotton fiber. Int. Rev. Cytol. 89: 65-113.

Beasley, C.A., and I.P. Ting. 1974. The effects of plant growth substances on *in vitro* fiber development from unfertilized cotton ovules. Amer. J. Bot. 61: 188-194.

Desai, A., P.W. Chee, J. Rong, O.L. May, and A.H. Paterson. 2006. Chromosome structural changes in diploid and tetraploid A genomes of *Gossypium*. Genome 49: 336-345.

Graves, D.A., and J.M. Stewart. 1988. Chronology of the differentiation of cotton (*Gossypium hirsutum* L.) fiber cells. Planta 175: 254-258.

Haigler, C.H., D.H. Zhang, and C.G. Wilkerson. 2005. Biotechnological improvement of cotton fibre maturity. Physiologia Plantarum 124: 285-294.

Kim, H.J., and B.A. Triplett. 2001. Cotton fiber growth in planta and in vitro: Models for plant cell elongation and cell wall biogenesis. Plant Physiol. 127: 1361-1366.

Loguercio, L.L., J.Q. Zhang, and T.A. Wilkins. 1999. Differential regulation of six novel MYB-domian genes defines two distinct expression patterns in allotetraploid cotton (*Gossypium hirsutum* L.). Mol. Gen. Genet. 261: 660-671.

NCCA. 2001. Profile of a resourceful industry [Online]. Available at http://www.cotton.org/pubs/index.cfm (National Cotton Council of America).

Percival, A.E., J.F. Wendel, and J.M. Stewart. 1999. Taxonomy and germplasm resources. p. 33-63. *In* C.W. Smith and J.T. Cothren (eds.) Cotton: Origin, history, technology, and production. John Wiley & Sons, Inc. New York, NY.

Shi, Y.H., S.W. Zhu, X.Z. Mao, J.X. Feng, Y.M. Qin, L. Zhang, J. Cheng, L.P. Wei, Z.Y. Wang, and Y.X. Zhu. 2006. Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. Plant Cell 18: 651-664.

Smith, C.W. 1999. Production statistics. p. 435-450. *In* C.W. Smith and J.T. Cothren (eds.) Cotton: Origin, history, technology, and production. John Wiley & Sons, Inc., New York, NY.

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

Suo, J., X. Liang, L. Pu, Y. Zhang, and Y. Xue. 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.

Tiwari, S.C., and T.A. Wilkins. 1995. Cotton (*Gossypium hirsutum*) seed trichomes expand via diffuse growing mechanism. Can. J. Bot. 73: 746-757.

Udall, J.A., L.E. Flagel, F. Cheung, A.W. Woodward, R. Hovav, R.A. Rapp, J.M. Swanson, J.J. Lee, A.R. Gingle, D. Nettleton, C.D. Town, Z.J. Chen, and J.F. Wendel. 2007. Spotted cotton oligonucleotide microarrays for gene expression analysis. BMC Genomics 8: 81.

Wang, S., J.W. Wang, N. Yu, C.H. Li, B. Luo, J.Y. Gou, L.J. Wang, and X.Y. Chen. 2004. Control of plant trichome development by a cotton fiber MYB gene. Plant Cell 16: 2323-2334.

Wendel, J.F. 1989. New World tetraploid cottons contain Old World cytoplasm. Proc. Natl. Acad. Sci. USA. 86: 4132-4136.

Wendel, J.F., and R.C. Cronn. 2003. Polyploidy and the evolutionary history of cotton. Adv. in Agron. 78: 139-186.

Wilkins, T.A., and A.B. Arpat. 2005. The cotton fiber transcriptome. Physiologia Plantarum 124: 295-300.

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

Yang, S.S., F. Cheung, J.J. Lee, M. Ha, N.E. Wei, S.H. Sze, D.M. Stelly, P. Thaxton, B. Triplett, C.D. Town, and Z.J. Chen. 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.