IDENTIFICATION OF GENES DIFFERENTIALLY EXPRESSED IN ELONGATING FIBER IN COTTON CHROMOSOME SUBSTITUTION LINE CS-B25

By

Samuel Sunil Kumar Bandi

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Masters of Science
in Biochemistry
in the Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology

Mississippi State, Mississippi

August 2011

IDENTIFICATION OF GENES DIFFERENTIALLY EXPRESSED IN ELONGATING FIBER IN COTTON CHROMOSOME SUBSTITUTION LINE CS-B25

By

Samuel Sunil Kumar Bandi

Approved:

Din-Pow Ma Professor of Biochemistry and Molecular Biology Major Professor And Graduate Coordinator	Johnie N Jenkins Adjunct Professor of Plant and Soil Sciences, and Biochemistry and Molecular Biology, Research Geneticist, USDA, ARS (Committee Member)
Sukumar Saha Adjunct Professor of plant Soil Sciences, Research Geneticist, USDA, ARS (Committee Member)	Jiaxu Li Associate Professor of Biochemistry and Molecular Biology (Committee Member)
George Hopper Interim Dean of the College of Agriculture and Life Sciences	

Name: Samuel Sunil Kumar Bandi

Date of Degree: August 6, 2011

Institution: Mississippi State University

Major Field: Biochemistry

Major Professor: Dr.Din-Pow Ma

Title of Study: IDENTIFICATION OF GENES DIFFERENTIALLY EXPRESSED IN ELONGATING FIBER IN COTTON CHROMOSOME SUBSTITUTION LINE CS-B25

Pages in Study: 80

Candidate for Degree of Masters of Science

Identification of genes specific to fiber development would improve the efforts in developing cotton plants with superior fiber quality. Through genetic introgression, 17 interspecific chromosome substitution lines (CS-B lines) of upland cotton in *G. hirsutum* (TM-1) have been developed and released recently. These substitution lines have TM1 as background and contain either whole chromosome or chromosome arms of *G. barbadense* (line 3-79) chromosomes. CS-B25 has chromosome 25 from *G. barbadense* substituted into TM-1 *G. hirsutum* was reported to show superior fiber properties. In this study, suppression subtraction hybridization (SSH) combined with Affymetrix cotton genome microarray arrays were used to identify differentially expressed genes in CS-B25. An SSH fiber cDNA library was constructed with differentially expressed genes identified in CS-B25. Microarray analysis showed that 23 genes were up-regulated and 9 down-regulated. Majority of these genes were involved in Ethylene signal pathway, Ubiquitin-proteasome pathway and cell wall synthesis.

DEDICATION

I would like to dedicate this thesis to my parents, Bandi Yohan and Mukamala Roseline.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my major professor, Dr. Din-Pow Ma for his continuous guidance, kind support, patience and teaching. I would also like to thank all my committee members Dr. Johnie Jenkins, Dr. Sukumar Saha and Dr. Jiaxu Li for their guidance and support. I am especially grateful to Dr. Johnie Jenkins and Dr. Scott Willard and the Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology for financial support to help me complete my study.

I would like to give my sincere appreciation to Dr. Chuan-Yu Hsu for her valuable advices and guidance throughout my research. My sincere thanks are extended to all students and faculty in the Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology for their assistance and friendship. Finally, I would like to deeply thank my parents for their unending love and support during my studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

aa	amino acid(s)
ACC	1-Aminocyclopropane-1-carboxylic acid
ACS	1-Aminocyclopropane-1-carboxylic acid synthase
AMP	Adenosine-5□-monophosphate
ATP	Adenosine-5□- triphosphate
cDNA	complementary DNA
CS-B25	chromosome substitution line 25
CT	Threshold Cycle
ddH_2O	double distilled water
DEPC	Diethyl pyrocarbonate
DNA	2'-deoxyribonucleic acid
DPA	day post anthesis
E1	ubiquitin-activating enzyme
E2	ubiquitin-conjugating enzyme
E3	ubiquitn ligase
EDTA	ethylenediaminetetraacetic acid
EIN	Ethylene insensitive
ERF	Ethylene responsive factor

ESTs	Expressed Sequence Tags
Gh	*Gossypium hirsutum*
GO	Gene Ontology
IPTG	isopropyl-β-D-thiogalactopyranoside
Kb	kilobase(s)
kDa	kilodalton
LB	Luria-Bertani
M	molar concentration
mRNA	messenger RNA
NP-40	Nonidet P-40
OD	optical density
PCR	polymerase chain reaction
PVP	polyvinylpyrrolidine
RING	Really Interesting New Gene
RNA	ribonucleic acid
RT	reverse transcription
SAM	S-adenosyl methionine
SCF	Skp-Cullin-F-box
SSH	Suppression Subtraction Hybridization
TAE	Tris-acetate/EDTA
TM1	Texas Marker 1
Tris	tris (hydoxymethyl)-aminomethane

CHAPTER I

INTRODUCTION

Gossypium barbadence and *Gossypium hisutum* are the only two tetraploid cultivated cotton species in the world. These species have size range from large shrubs to small plants. They have an allotetraploid genome (AADD) containing a pair of A and D genomes. Upland cotton (*G. hisutum*) is the major source of fiber in the textile industry. Upland cotton is known for its better adaptation to environments and superior agronomic traits including high lint percentage and yield. Unlike Upland cotton, Pima cotton (*G. barbadense*) has a low lint percentage and low yield but has better fiber qualities in terms of length, strength, and fineness. For the past 15 years, fiber quality has been steadily decreasing due to environmental and other factors (Feng et al., 2011). Hence, considerable efforts have been made in developing new varieties of cotton with improved fiber qualities which can benefit the farmer. Insertion of new genes in cotton genome requires an in-depth understanding of cotton fiber development. Conventional plant breeding techniques such as interspecific crosses, genetic recombination, mutagenesis, introgression and transformation, have been employed in cotton. Among them, interspecific introgression chromosome substitution, an important molecular and cytogenetic tool, has been recently used to develop 17 chromosome substitution lines of upland cotton in *G. hirsutum* TM-1 background containing *G. barbadense* (3-79) chromosome or segments (Stelly et al., 2005). In each chromosome substitution line (CS-B), a pair of chromosome from TM1 was replaced by respective chromosome pair in 3-

79 via hypoaneuploid based backcross substitution. Some important agronomic and fiber traits of chromosome substitution lines have been reported in recent studies (Saha et al., 2006; 2010; 2011). Among them, CS-B14sh and CS-B22Lo had higher lint yields and seeds. The CS-B25 line was reported to have increased fiber strength and length, reduced micronaire property, and nearly the same (slightly reduced) yield relative to TM-1 (Saha et al., 2006). The paucity of information about the genes controlling quantitative traits has been the biggest challenge in cotton improvement. Since CS-B lines are nearly isogenic for 25 chormosomes, a comparative analysis of gene expression in fiber between the two lines would help identify genes specific to the substituted chromosome. Genes associated with fiber development in chromosome 22 short arm has been studied using a chromosomal substitution line CS-B22sh via subtraction hybridization and microarray analysis (Wu et al., 2008). A similarly comparative analysis of gene expression between CS-B25 and TM-1 would help identify genes associated with fiber development on the chromosome 25 of *G. barbadense* 3-79. In this research, an integrated approach of Affymetrix cotton genome array analysis and suppression subtraction hybridization (SSH) was used to identify differentially expressed genes in 10 DPA CS-B25 fibers. The candidate genes identified would help in understanding the complex nature of Quantitative Trait Loci (QTLs) associated with fiber development in CS-B25 (Saha et al., 2010). In addition, these genes will be helpful for developing Single Nucleotide Polymorphism (SNP) markers associated with fiber quality traits and enhance Upland Cotton germplasm. The developed SNP markers along with available SSR markers can be then utilized to map fiber quality QTLs using chromosome 25 specific recombinant inbred lines (CS25RIL) developed from the cross of CS-B25 and TM-1 using a single seed decent method. The identification of fiber candidate genes will reveal important

molecular markers related to marker assisted selection of fiber quality traits and facilitate map-based cloning of fiber quality genes and other genes of interests located on chromosome 25 of *G. barbadense*.

CHAPTER II

LITERATURE REVIEW

Cotton fiber development

Cotton fibers are developed and differentiated from epidermal cells of a developing seed Cotton fiber development is controlled by genotype, various phytohormones and environmental factors. The development of fiber can be divided into four overlapping stages; initiation, elongation (primary cell wall synthesis), secondary cell wall deposition and maturation (Seagull, 1992; Shimizu et al, 1997; Raun and Chourey, 1998).

Initiation

Fiber initiation begins on or near the day of anthesis, and roughly 30 % of outer epidermal cells on ovules differentiate into fiber initials. The outer epidermal cells of cotton ovules have cuboidal shape, and they are either in the state of division or interphase during the last days before anthesis. The cells to be in the interphase stop divide and develop slight protrusions above the epidermal surface. These differentiating cells are referred as fiber initials. These initial cells grow rapidly into single unicellular fiber cells without further cell division. The development of primordial fiber is characterized by the enlargement of nucleus, reduction in phenolic compounds in vacuole and an increase in cytoplasmic density, as revealed from the ultrastructure of epidermis. The cytological studies on cotton ovules have shown that an increase in endoplasmic reticulum and release of phenolic compounds from the vacuoles into cytoplasm were

predominant in cells destined to differentiate into fiber cells (Ramsey and Berlin, 1976). On the day of anthesis, majorities of lint fibers are clearly visible as fiber initials and assume a 'ballooning' shape.

Elongation

Fiber initials, once formed, begin to undergo longitudinal expansion in the elongation stage which lasts up to 20 DPA. At about 4 DPA, the longitudinal expansion of fiber initial dominates the lateral expansion, giving the approximate diameter to the mature fiber. The blunt end of fiber tips tend to taper, orienting towards the micropylar end of the ovule (Stewart, 1975). The tapering tips of the fibers tend to adhere to each other, forming a small cluster with a certain contortion. This contortion generates a spiral configuration of the fiber tube, and plays an important role with respect to protoplasmic streaming and cell wall building (Ramsey and Berlin, 1976). The fiber cells enclose a gigantic vacuole with nucleus positioned approximately between midway and the tip. The cytoplasm contains plastids, mitochondria, single ribosomes, lipid bodies and dictyosomes with electron dense mature-face cisternae. The size of spherical nucleolus in elongating fibers increases exponentially and reaches a maximum diameter of 8-10 μm. The nucleolus is thought to be the major site of active protein synthesis, which provides the increased metabolic activity for rapid growth of expanding fiber cells. The primary cell wall is composed of microfibril. Individual mircofibril is about 10-30 nm wide, 5-10 nm thick and several microns long (Albersheim et al., 1977). Many studies have shown that the microfibrils are interwoven to form microfibril bundles. These bundles are transversely oriented on the inner side of the primary wall, next to the plasmalemma. During fiber elongation, enzymes synthesized in endoplasmic reticulum are transferred to

Golgi complex via membrane flow (Westafer and Brown, 1976). The substrates (specific to cell wall synthesis) in the cisternae of golgi apparatus are converted to products which are eventually incorporated into cell membrane.

Secondary cell wall synthesis

Secondary cell wall synthesis begins at about 15 DPA and lasts until 25 DPA which involves increasing accumulation of cellulose. Electron microscopy of cotton fiber cells revealed that endoplasmic reticulum and plasma membrane were actively involved in synthesizing the secondary cell wall components (Westafer and Brown, 1976). During the transition phase from elongation to secondary cell wall synthesis (15 to 20 DPA), the rate of cellulose biosynthesis increases nearly 100 fold. The non-carbohydrate metabolism tends to transit towards cellulose biosynthesis. Cellulose deposition occurs in an ordered manner by membrane embedded terminal complexes, which can be visible under the electron microscope (Roberts and Roberts, 2007). In addition to cellulose deposition, a change in orientation of microfibrils was also observed during secondary cell wall synthesis in cotton fiber (Seagull, 1992). Several studies have shown that microfibril orientation undergoes a shift from roughly 70-90˚ to 45-55˚ in relative to longitudinal axis, and this orientation is species dependent (Seagull, 1992). The secondary cell wall of fiber cells starts to thicken at 25-32 DPA when cellulose microfibrils are deposited in each successive layer.

Fiber maturation

Fiber maturation occurs at 40-50 DPA and is the final stage in cotton fiber development. Fiber cells at this stage lose their water contents, and the cytoplasm dries against the inner surface of the cell wall and the enlarge lumen is left in the middle where

the central vacuole was previously located (Kim and Triplett, 2004). The mature fiber contains nearly 95 % of cellulose as dry weight.

Gene expression in development of cotton fiber

During the past 10 to 15 years, many molecular techniques including suppression subtraction hybridization, microarray analysis, Expressed Sequence Tags (ESTs), and Real-Time PCR have been used to identify genes associated with cotton fiber development. Cotton fiber is a single cell and has become a perfect model for studying cell differentiation and elongation. An integrated approach to understand the mechanism of fiber development at the molecular level has been carried out by Ji et al. (2003) via a comparative gene expression analysis of 10 DPA ovules in the wild type Upland cotton Xuzhou-142 and its fuzzless –lintless-seed (*fl*) mutant. Using PCR select cDNA subtraction and microarray analysis, they have identified 172 genes that may play some important roles during fiber development. Among them, genes for a putative vacuolar (H+)-ATPase catalytic subunit, a kinesin-like calmoudulin binding protein, an arabinogalactan protein, enzymes involved in long chain fatty acid biosynthesis, proteins in auxin signal transduction, mitogen-activated protein kinases (MAPKs), and expansins were found to be up-regulated in the wild type cotton. Arpat et al. (2004) has conducted an in silico expression analysis of approximately 14,000 unigenes assembled from 46,603 expression sequence tags (ESTs) and revealed that the bulk of gene transcripts in rapidly elongating fiber cells are involved in cell wall structure and biogenesis, cytoskeleton, and carbohydrate metabolism (Arpat et al., 2004). These analysis results were further supported by oligonucleotide microarrays which confirmed that the dynamic changes in gene expression associated with fiber development were highly stage specific (Arpat et

al., 2004; Al-Ghazi et al., 2009). Research efforts have henceforth progressed in identifying genes specifically expressed in fiber cells at different developmental stages. To identify genes that regulate fiber initiation, the expression profile of 0 DPA whole ovules from fiberless mutants was compared with wild type lint cotton using cDNA microarrays (Wu et al., 2006). Numerous genes differentially expressed in cotton ovules during the initiation stage have been identified. Among them, genes for a Myb transcription factor (similar to the *Antirrhinum* Myb AmMIXTA), a putative homeodomain protein, cyclin D, lipid transfer proteins, α-expansin and sucrose synthase were found to be up-regulated. Further studies have shown that *GhMyb25* and a homeodomian gene, predominantly expressed in ovules, were up-regulated on the day of anthesis in fiber initials. In cotton, high level expression of the SuSy gene encoding sucrose synthase has been reported in 0-DPA fibers (Ruan and Chourey, 1998). The *GhMYB*25 gene has been also shown to be expressed predominantly in the epidermis of ovules in developing fiber initials by Machado et al. (2009). Two lipid transfer protein encoding genes, *GhLtp3* and *GhLtp6*, highly expressed in elongating fiber, have been isolated and characterized (Ma et al., 1995; Ma et al., 1997). Two MYB proteins, GhMYB7 and GhMYB9, were reported to be involved in transcriptional regulation of the *GhLtp3* gene during fiber development (Hsu et al., 2005). Lipid transfer proteins are basic, small proteins thought to participate in the lipid transferring during biomembrane synthesis (Zhang and Liang 2000). Taliercio et al. (2007) have used microarray to analyze gene expression with RNA from 1-DPA fiber initials. Several genes associated with fiber initiation including brassinosterols regulation, GTP mediated signal transduction, cell cycle control, components of Ca^{+2} mediated signaling pathway, and a CAPRICE/TRIPTYCHON (CPC) transcription factor that lacks the activating domain

have been identified. A custom designed cDNA microarray, based on 12,233 unique ESTs from fast elongating upland cotton fiber cells, was used for expression profiling of genes related to fiber development (Shi et al., 2006) with special emphasis in identifying genes associated with regulatory pathways. Through microarray analysis and subsequent sequencing of fiber cDNA libraries, it was found that ethylene and brassinosteriod biosynthetic pathways were significantly up-regulated during fiber elongation (Shi et al., 2006). The 1-Aminocyclopropane-1-Carboxylic Acid Oxidase 1-3 (*ACO1-3*) genes, responsible for ethylene production, were expressed at higher levels during this growth stage. Furthermore, an exogenous application of ethylene to cultured ovules resulted in robust fiber cell expansion, supporting its role in promoting cell elongation. These studies strongly indicate that the brassinosteriods and ethylene regulatory pathways play an important role in fiber initiation and elongation. Furthermore, an in vitro application of Lignoceric acid, a saturated very long chain fatty acid, promoted fiber cell elongation in the fuzzless-lintless mutant by activating ethylene signaling pathway. This confirms that the ethylene signaling pathway help in maximizing the extensibility of cotton fiber (Qin et al., 2007). In another expression study, the *GhKCR3* gene encoding 3-ketoacyl-CoA reductase that catalyzes the second step of fatty acid elongation and the gene for ATP synthase δ1 subunit required to maintain a higher ATP/ADP ratio were reported to up-regulate in elongating cotton fibers (Pang et al., 2010). Cotton fibers elongate rapidly after initiation and elongation, and deposit a large amount of cellulose at the secondary wall synthesis stage. Wu et al. (2006) identified genes that were up-regulated in expanding cotton fiber initials relative to non-expanding epidermal cells. Many of the up-regulated genes encode putative cell wall and cell membrane proteins needed for rapid cell elongation. A comparative analysis of transcriptomes and metabolic profiles of fibers

at different developmental stages (3 to 18 DPA) was performed, and 633 genes were identified to be differentially expressed during fiber development (Gou et al., 2007). The expressed genes were divided into four groups: auxin signaling, wall loosening, lipid metabolism, and cellulose biosynthesis. These data demonstrated that the fiber cell at a certain stage has its own unique features, and each development stage is distinguished by different transcript and metabolite profiles. It was found that genes induced by hormones auxin and gibberellins were also up-regulated in fast elongating fibers. Five putative auxin response genes and eight aquaporins-like genes were reported to be highly expressed in developing fibers cells during 6- to 12- DPA and 3- to 15-DPA, respectively. In addition, genes for cell wall loosening proteins α-expansins, lipid transfer proteins, and lipid biosynthesis were also highly expressed at 6 DPA, and their expression levels were maintained throughout the elongation stage. Mature cotton fiber is composed of more than 90% cellulose along with minor non-cellulosic matters. Genes associated with cellulose synthesis such as cellulose synthase, sucrose synthase, invertase, pectin esterase, β-galactosidase and other poly- or oligo-saccharide hydrolyzing enzymes had an increased expression level in 18 DPA fiber. Cotton oligonucleotide microarrays have been successfully used to identify fiber specific genes which have strong homologies with genes in Arabidopsis (Udall et al., 2007). Using nucleotide sequence homologies, six *EXPANSIN* genes were determined to be differentially expressed at the early stage of fiber development, supporting their importance in fiber elongation and lateral expansion (An et al., 2007). Hovav et al. (2008) compared gene expression in developing fibers cells with different cotton tissues and found that many genes involved in vesicle coating and trafficking were over-expressed throughout all stages of fiber development. Using suppression subtraction hybridization

and cDNA microarrays, Wu et al. (2008) identified 36 fiber specific genes that were differentially expressed in CS-B22sh 15-DPA fiber compared to TM-1. Among them, beta-tubulins, an actin, a putative kinesin light chain, a cellulose synthase, glycosyl hydrolase family 3 proteins, pyruvate decarboxylase, glycoside hydrolase family 5, GDP-mannose pyrophosphorylase, dynamin-like protein, and annexin were up-regulated in CS-B22sh. Two alpha tubulin genes, *GhTau*2/3 and *GhTau*4, have been found to up-regulate in elongating fiber cells (Whittaker and Tripllet, 1999). In addition to tubulin proteins, cotton *GhRac*1 GTPase, similar to the group IV member of Arabidopsis *Rac/Rop* GTPase, was identified to play an important role in cytoskeleton assembly (Kim and Triplett, 2004). A *GhACT*1 gene encoding actin is identified and its role in cytoskeleton has been investigated. RNAi interference analysis has shown a significant decrease in the GhACT1 protein level which resulted in the disruption of actin cytoskeleton, suggesting a predominant expression of *GhACT*1 in elongating fiber (Li et al*.,* 2005). Using a SNP-specific microarray technology, Hovav et al. (2008) studied the putative advantage of allopolyploidy in cotton genomes. The technology was applied to fibers harvested from three developmental time points in wild and modern domesticated cotton species, *G. hirsutum* and *G. barbadense*. In both species it was found that D-genome expression was preferentially enhanced under human selection pressure, leading to convergent genetic alterations. *G. barbadense* is a widely cultivated cotton with superior luster, silkiness and relatively high yield. A comparative expression profiling via microarray was performed on fiber cells isolated at three different elongation stages in wild (K101) and domesticated (Pima S-7) cotton plants (Chaudhary et al. 2008). Studies have shown that nearly 4,200 genes were differentially expressed between wild and domesticated cotton, suggesting that replicated domestication in two different species (*G.*

hirsutum and *G. barbadense*) has resulted in overlapping, parallel, and metabolic transformations (Hovav et al., 2008). Mei et al. (2009) studied the correlation between increasing concentration of reactive oxygen species (ROS) with fiber elongation. They found that the application of NADPH oxidase inhibitor diphenyleneiodonium and peroxidase inhibitor salicyl hydroxamic acid to the wild-type cotton ovule culture significantly suppressed fiber growth. Ten *GhPOX* (1-10) genes encoding cotton class III peroxidases were identified and microarray studies revealed that *GhPOX*1 was predominantly expressed in elongating fibers. Betancur et al. (2010) have studied potential analogies between phylogenetically distinct cellulose synthase genes in Arabidopsis shoot trichromes and cotton fiber. The expression analysis of cellulase sythase genes (*CesA*) via microarray has revealed the orthologus nature of cellulose synthases between the two plants.

Molecular tools for large scale gene expression study

Microarray analysis

The state of a cell at any given stage is governed by the expression of a subset of genes at that time. According to the central dogma, the first step in gene expression is transcription, in which DNA sequences are transcribed into mRNAs. The mRNAs are then translated to synthesize proteins. The genes associated with a biological process can be studied by isolating the proteins associated with the process and followed by determining the function of the proteins. However, isolation, identification and structure determination of proteins are time consuming, expensive, and involving numerous experimental manipulations. This would make mRNAs a useful and reliable source to study the genes associated with the biological process. Based on the knowledge of what

mRNAs and in what levels, an investigator can draw conclusions pertaining to the state of a particular cell. The collection of all mRNAs of a cell is referred as transcriptome. In a comparative microarray analysis, array slides containing tens of thousands genes are exposed to different mRNA samples and the fluorescent signal intensities are then compared among the samples. From the fluorescence, an expression profile of a set of genes can be created for each sample.

Microarrays are currently used in biological research to address a wide variety of questions. It is known that cells from different tissues perform different function. Although cells can be distinguished by their phenotypes, the underlying biochemical reactions are governed by proteins which in turn depend on specific gene expression. Microarrays have been successfully used in studying overall developmental process of an organism. A comparative analysis of gene expression profiles at different growth stages can be used to track the development of the organism. Selective gene expression patterns in different organs or tissues in response to environmental and developmental cues have been studied in Arabidopsis via microarray (Ma et al., 2005). This analysis has shown that plants, depending on the presence or absence of light, undergo dramatic changes in developmental patterns. A distinct set of genes associated with each organ (leaf, flower, and cotyledons) has been identified and only 16% of them are common in all the organs examined. Using microarrays, the evolutionary relationship among species and the impact of environmental changes on the developmental process of an organism have been studied. Preuss et al. (2004) have studied the evolution of human brain by comparing gene expression in brain tissues obtained from humans, chimpanzees, and other non-human primates (squirrels and monkeys). The results showed that the rate of gene expression change in the human brain accelerated during evolution, and approximately 2-

4% of genes were differentially expressed in the cerebral cortex. Cherkasova et al. (2003) have developed two microarray methods to study the genetic diversity of RNA viruses at the level of genome recombination and nucleotide sequence heterogeneity. The experimental results implicated that these two methods could be used for large-scale full-genome sequencing screening of viral isolates, vaccine quality control, and drug treatment. Microarray experiments can be also used to study genetic diseases that result from mutations within a gene or a set of genes. Nonsense mutations in messenger RNA coding region (UAA, UAG or UGA) can lead to premature translational termination which results in truncated polypeptide products. However, a selective bioavailable non toxic drug that can read disease-causing premature codons might alleviate the pathologies of nonsense mediated diseases. Welch et al. (2007) have studied the application of a PTC124 drug in patients harboring nonsense mutations in cystic fibrosis transmembrane conductance regulator (*CFTR*) via microarray. Their studies have shown that PTC124 can reduce the severity of principle manifestations of cystic fibrosis by promoting UGA non sense suppression. Microarray analysis also can be used in identifying genes differentially expressed in diseased cells by comparing with normal cells. This would enable scientists to study genes associated with disease processes and further facilitate in designing better drugs that are specifically aimed to target a gene, protein, or signaling cascade. In addition, the changes in gene expression in diseased cells when treated with new drugs can be also monitored via microarray analysis. Gupta et al. (2010) have investigated the role of efflux pump genes in multidrug resistant *Mycobacterium tuberculosis* during stress induced by common anti-tuberculous drugs. This study has led to identification of eight important efflux genes which were upregulated when multidrug-

resistant *M. tuberculosis* isolates were exposed to common anti-tuberculous drugs (ofloxacin, carbonyl cyanide, m-chloro phenylhydrazone).

The microarray experiments have been used to perform large scale gene expression profiling. In principle, the experimental design is similar to Southern and Northern hybridization. A typical microarray experiment for gene expression analysis consists of five important steps: preparing the microarray, preparing labeled mRNA sample, hybridization of labeled sample to the microarray and washing the microarray slide, scanning the microarray and interpreting the scanned image. The microarray gene chips are commercially available from Affymetrix GeneChip system, which uses prefabricated oligonucleotide chips. The Genechip probe arrays are manufactured using a combined technology of photolithography and combinatorial chemistry. Approximately 1.3 million oligonucleotide probes are synthesized on each array. Each oligonulceotide is located on a specific area called probe cell. Probe arrays are manufactured in a series of cycles on glass substrate coated with linkers that contain a photolabile protecting group. A mask is applied to the probe arrays, and only selected portions are exposed to UV light. Illumination enables the removal of photolabile protecting groups and facilitates the addition of selective nucleoside phosphoramidite only at the exposed sites. In the next step, a different mask is applied and followed by illumination and chemical coupling. A specific set of oligonucleotide probes are synthesized in a given location by repeating the cycle. The completed probe arrays are finally packaged into cartridges. In order to detect the cDNA bound to microarrays, the cDNA sample is labeled with a reporter molecule. The reporters currently used in microarray experiments are fluorescent dyes, called Fluros or Fluorophores, which absorb energy of a specific wavelength and re-emit energy (fluorescence) at a different (but equally specific) wavelength. Streptavidin phycoerythrin

and cyanide dye are exclusively used in microarray analysis. The Streptavidin phycoerythrin, also known as SA-PE or R-PE streptavidin, is typically used for flow cytometry, microarrays, ELISA, and other applications that require either high sensitivity or simultaneous multicolor detection. Phycoerythrin is a member of a family of proteins called phycobiliproteins, derived from cyanobacteria and eukaryotic algae, exhibits extremely bright fluorescence and high quantum yields. A DNA microarray is hybridized with a biotinylated DNA probe along with diluted calf thymus DNA and then incubated with Streptavidin phycoerythrin, which has excitation and emission wavelengths at 488 nm and 570 nm respectively. The cyanide dye family including Cy3 and Cy5 are water-soluble fluorescent dyes of the cyanine dye family. The Cy3 dye emits green fluorescence light at 570 nm when excited by emission light at 550 nm, whereas Cy5 absorbs in the orange region at 649 nm and emits red color fluorescence (650/670 nm). They are synthesized with reactive groups on either one or both of the nitrogen side chains that can be chemically linked to either nucleic acids or protein molecules. The dye labeling is for visualization and quantification purposes. They are used in a wide variety of biological applications including comparative genomic hybridization and gene chip analysis, which are used in transcriptomics. They are also used to label proteins and nucleic acid for various studies including proteomics and RNA localization. In microarray experiments, DNA or RNA is labeled with either Cy3 or Cy5, which carries an N-hydroxysuccinimidyl ester (NHS-ester) reactive group. Since NHS-ester reacts readily only with aliphatic amine groups, which nucleic acids lack, nucleotides have to be first modified with an aminoallyl group. This was done by incorporating aminoallyl-modified nucleotides during the synthesis reactions. The quenching effects can be minimized by labeling every 60 bases within the nucleotide sequence such that the labels are not too close to each

other. Although Cy3 and Cy5 have been effectively used in microarrays, their differential incorporation into DNA and different emission responses to the excitation laser make their usage less congenial when compared with steptavidin phycoerythrin.

The sample preparation step involves the synthesis of fluorescently labeled cDNA often referred as target DNA. The poly (A) RNA isolated from total RNA is converted to more stable double stranded cDNAs via two reverse transcription reactions. An *in vitro* transcription (IVT) reaction is then performed to produce biotin-labeled cRNA from the double stranded cDNA. The biotin labeled cRNA is used to perform hybridization by allowing the target DNA to diffuse uniformly throughout the slide. The microarray slides are sealed and placed in a hybridization chamber, following by incubation at specific temperature for sufficient time to allow the hybridization reaction to complete. A single stranded DNA/RNA whose sequence is complementary with another single stranded DNA will bind with high affinity. An imperfect match will bind with low affinity. The stringency in hybridization in microarray experiments depends on the cRNA sequence and experimental conditions such as temperature and incubation time. Following hybridization, the microarray slides are washed to eliminate any excess labeled samples. The hybridized probe array is then incubated with streptadavin phycoerythrin conjugate which specifically binds to biotin molecule. In the final step, the microarray slide is washed, dried and scanned using Gene Array Scanner. The amount of light emitted is directly proportional to the bound target at each location on the probe array. Although the scanner detects the light emitted by the target cRNAs bound to their complementary probes, it is inevitable that light from other sources would also be detected. The scanned image is quantified in three separate steps: gridding, segmentation and intensity extraction (Bowtell and Sambrook, 2003). Gridding is a process of assigning coordinates

to each identified spot which eventually help in high throughput analysis. Segmentation allows the classification of pixels into foreground and background. Intensity extraction involves the calculation of spot intensities and quality measurements. The image analysis of microarray provides measures of foreground and background values. The background constitutes light emitted from residual samples adhering to the slide, chemicals used in processing the array, hybridization artifacts, and the array itself. The background light can be corrected by different methods: global background, spot background and smoothed background adjustment. Several commercially available image analysis software packages (http://imaging.brocku.ca/products/ARV.asp, Imaging Research inc., www.mediacy.com, MediaCybernetics and www.biodiscovery.com, BioDiscovery) are available in the market. Affymetrix uses a variation of smoothed background adjustment, called Zonal adjustment background (Amaratunga and Cabrera, 2004). This involves splitting up of microarray slide into different zones and deriving low percentile spot intensities within each zone. A threshold value for the background in each zone is then calculated and all spot intensities are adjusted based on the computed value. The background adjusted threshold values are further used to adjust systemic differences in relative intensity of each channel, a process known as normalization. The intensity values are converted to expression values using Robust Multichip Average (RMA), a normalization approach exclusively used for Affymetrix microarrays (Irizarry et al., 2003). The RMA analysis can be readily performed using Bioconductor (http://www.bioconductor.org/), a web based toolkit which helps in computing RMA measurements. The raw Affymetrix data contains about 20-25 probes for the same cRNA target. Half of these will be mistmatches, which do not bind to the target sequence. The RMA algorithm takes advantage of these mismatches and summarizes the perfect

matches using quantile normalization. The statistical significance of the normalized data is then tested using the Wilcoxon signed rank-based gene expression presence/absence program, which does not depend on the distributional assumption made by RMA (Amaratunga and Cabrera, 2004). The next step involves scaling, which allows a comparative analysis of the expression levels of genes present in the sample with the control. The outliers within calculated data are then identified by using either *t*- test or ANNOVA. A close examination of the results obtained from the above mentioned statistical analysis often results in the identification of false positives. This small variance is corrected using Multiple Testing Correction/False Discovery Rate (FDR), which is the expected proportion of false positives among the positive findings. The FDR corrected microarray data has been be used to calculate fold change for identical genes present within the sample and the control.

The probable function associated with individual genes can be identified using online tools such as BLAST (Basic Local Alignment Search Tool) at NCBI, UNIPORT and TAIR. The functional categories of the genes identified can be performed using gene ontology programs (www.geneontology.org) of Blast2go and Easygo. The production of microarrays and the numerical data obtained will only help in identifying up-regulated and down-regulated genes in the sample. On the other hand, the data does not provide information about the genes that fall under similar biological process. Two powerful analytical techniques, cluster and tree view, has been introduced to organize array data. The cluster organizes related gene expression data and tree view allows clustered data to be easily visualized. Both of these tools are web based programs (http://rana.lbl.gov), which allow a compact display of clustered genes in the form of a dendrogram.

Suppression subtraction hybridization

In higher eukaryotes, many biological processes such as cellular growth, organogenesis, differentiation, and hormone regulation are mediated by programs of differential gene expression. To understand the molecular regulation of these processes, the relevant subsets of differentially expressed genes must be cloned, identified, and studied in details. Numerous cDNA subtraction methods have been reported, and in general they involve the hybridization of cDNA from one population to excess cDNA of another population. The unhybridized cDNAs are separated using hydroxylapatite chromotagraphy or oligo$(dT)_{30}$-latex beads (Hedrick et al., 1984). This type of separation requires more than 20 μg of poly(A)-mRNA as the starting material and often the low abundant cDNAs are lost during purification. Recently an alternate method has been developed which avoids physical separation of unhybridized cDNAs from the hybridized ones. This method is often referred as representational difference analysis, which employs PCR to enrich the unhybridized cDNAs (Lisitsyn et al., 1993). The PCR-based amplification has been successfully employed to enrich differentially expressed genes that differ in size or representation. However, individual mRNAs with difference in abundance is not addressed well by representational difference analysis. A new PCR based cDNA subtraction method termed as suppression subtraction hybridization (SSH) has been developed by Clontech to address technical difficulties associated with the methods mentioned above. The SSH method involves hybridization of cDNA from one population (tester) to excess of cDNA from another population (driver) and followed by the separation of unhybridized cDNA from hybridized common sequences. SSH requires approximately 0.5-2 μg of poly (A) for the synthesis of double stranded cDNA samples. The tester and driver cDNAs are then digested with a type II restriction enzyme *Rsa* I that

yields blunt ends. The tester cDNA is then ligated with specific adaptors in two separate reactions. The adaptors do not contain a phosphate group; hence only the longer strand of each adaptor is attached to the 5' end of the cDNA. The adaptor sequence includes a DNA sequence identical to PCR primer which enables to amplify the unhybridized sequences. Two hybridization reactions are performed in SSH. In the first hybridization reaction, the adaptor-ligated test cDNA is hybridized with the excess of *Rsa* I digested tester cDNA. During this step, the concentration of abundant cDNAs is normalized. The normalization occurs because the generation of mono hybrid cDNAs during reannealing is relatively faster for abundant sequence due to second order kinetics of hybridization. The differential sequences are enriched in tester fraction and remain as common non-target cDNAs. In the second hybridization, the products of first hybridization are mixed in the presence of freshly denatured driver cDNA. In this step, the common sequences are preferentially subtracted and normalized, further enriching differentially expressed genes in the tester. The unhybridized cDNAs are then amplified using a pair of primers, which correspond to outer parts of the two adaptor sequences. This selective amplification is achieved by performing an extension reaction to fill the sticky ends of the molecules for primer annealing.

SSH has been effectively used by many research groups in identifying differentially expressed genes between distinct tissues or developmental stages. Harada et al. (2010) performed SSH analysis of two cDNA samples at two different stages of flowering opening in carnation and identified 235 different cDNA fragments with functions related to transcription, signaling, cell wall modification, lipid metabolism, and transport. Yamaura et al. (2010) studied the nitrogen fixing ability by *Frankia*, a gram positive actinobacterium, by comparing cDNAs isolated from cells grown in N^- (without

NH_4Cl) and N^+ media via SSH. They identified several genes associated with nitrogen fixation and these genes encode homologs of citrate synthase, peptide synthase, catalase, phosphophenolpyruvate mutase and crotonyl coenzyme A reductase. An integrated approach using SSH and cDNA microarray was employed to identify novel genes associated with sporulation and invasion of *Eimeria tenella* which causes avian coccidiosis, a major parasitic disease of poultry (Han et al., 2010). A comparative study of subtractive cDNA libraries for 3 stages (unsporulated oocysts, sporulated oocysts and sporozites) resulted in identification of 32 genes responsible for sporulation, invasion, and growth in *E. tenella*. Among them, genes encoding microneme proteins, surface antigens of *E. tenella*, heat shock proteins, and calcium dependent protein kinase were reported to be differentially expressed. Using SSH, Canales et al. (2010) identified a total of 225 unigenes that are differentially regulated in the roots of maritime pine trees in response to the changes in ammonium availability. Among them, genes for asparginase, aspargenase synthase, aldose 1-epimerase, transaldolase, and a granule- bound starch synthase were up-regulated under conditions of excess ammonium. SSH has been also successfully used to identify genes that control the development and inheritance of different root shapes in radish (Zaki et al., 2010). A comparative analysis of genes expressed between two cultivars of radish with different root shapes (long and thick vs taibyousoubutori and skinny) has led to the identification of 140 genes. Genes encoding phenyalanine ammonia lyase, C3H4-type RING finger, myrosinase, 1-aminocyclopropane-1-carboxylate oxidase, EDA29 transcription factor, root hair elongation protein, myb transcription factor, 4-coumarate-CoA ligase, calmodulin binding protein, and actin 2 were reported to be differentially regulated in the roots of the two radish cultivars.

Real Time RT-PCR

Real-Time polymerase chain reaction (RT-PCR) is one of the most powerful and sensitive techniques available for gene analysis. Its broad range of application includes quantitative gene expression analysis, genotyping, SNP analysis, pathogen detection, drug target validation, and RNA interference measurement. RT-PCR is often combined with reverse transcription to quantify messenger RNA (mRNA) and MicroRNA (miRNA). This technique is also called quantitative polymerase chain reaction (qPCR) since it measures PCR amplification as the reaction proceeds. In traditional PCR methods, the concentration of nucleic acid is determined after the amplification reaction is complete, making it impossible to determine the starting target DNA concentration. Real-Time PCR focuses on the exponential phase amplification, which allows more accurate quantification of nucleic acids. Within the exponential phase, the real-time PCR instrument calculates the threshold value, which is the level of detection at which the PCR product reaches a fluorescent intensity above the background. The PCR cycle at which the amplified DNA reaches this level is called the Threshold Cycle (Ct), which is often used in downstream quantitation or detection of the presence/absence of target DNA. The amount (or copy number) of target DNA sample can be accurately determined by comparing the Ct values of the target sample with a series of standards with known concentrations. In real-time PCR reactions, a fluorescent reporter molecule is generally used to detect target molecules. Two types of chemistries have been developed for target detection: TaqMan Probe and SYBR Green dye. The TaqMan Probe is a fluorogenic-labeled oligonucleotide that contains a reporter fluorescent dye at the 5' end and a quencher dye at the 3' end. While the probe is intact, the **quencher (Q)** fluorophore (long wavelength colored dye, such as red) reduces the fluorescence from the **reporter**

(R) fluorophore (short-wavelength colored dye, such as green). The quenching is based on Fluorescence (or Forster) Resonance Energy Transfer (FRET), which is the inhibition of fluorescence emission from a reporter caused by a quencher without the emission of a photon. During the PCR amplification, the probe binds to target DNA downstream from one of the primer binding sites and the reporter dye is removed by the 5' to 3' exonuclease activity of Taq DNA polymerase. This results in separation of reporter dye from the quencher dye, thereby increasing the reporter dye signal. The cleavage removes the probe from the template, allowing the primer extension to continue. Thus, the inclusion of TaqMan probes does not inhibit overall PCR amplification. However, the main disadvantage using TaqMan Chemistry is that the synthesis of different probes is required for different sequences.

The SYBR Green I dye is used to detect PCR products by binding to amplified double-stranded DNA. The binding and detection of the SYBR Green I dye is comparable to ethidium bromide, and the dye binds only to double-stranded DNA, not to single-stranded DNA, RNA, or small deoxyoligonucleotides. During PCR amplification, the SYBR Green I dye binds to every new copy of double stranded DNA synthesized. As the quantity of target amplicon increases, so does the amount of fluorescence emitted from the dye. The result is an increase in fluorescence intensity directly proportionate to the amount of PCR product produced.

Cotton chromosome substitution lines

Upland cotton (*Gossypium hirsutum*, 2n=52) is the most widely cultivated cotton species in the world. Upland cotton is known for its better adaptation to environments and production of high percentage lint fiber and high yields. In contrast, *Gossypium*

barbadence (pima cotton) produces low lint percentage and low yields but with better fiber qualities. There is an increasing demand for superior fiber qualities due to recent changes in the textile technology. An understanding of genetic diversity in the cotton genome is essential for improving fiber qualities. Three major approaches including mutagenesis, germplasm introgression and transformation have been utilized to increase genetic diversity. Conventional plant breeding methods such as interspecific crosses and genetic recombination have resulted in producing cytological abnormalities, distorted chromosome segregation, and non random products. An alternate method to traditional introgression is to develop chromosomal substitution lines. Recently 17 interspecific chromosome substitution lines of upland cotton in *G. hirustum* TM1 background have been developed and released to the public (Stelly et al., 2006). In each chromosome substitution line (CS-B), a pair of chromosome from TM-1 was replaced with respective chromosome pair in *G. barbadence* 3-79 via hypoaneuploid based backcross substitution. Some important agronomic and fiber traits associated with these CS-B lines have been reported in recent studies (Saha et al., 2006; 2010; 2011). The substitution lines, CS-B16, CS-B22sh and CS-B22Lo have higher lint percentage than their parents TM1 and 3-79. The substitution lines CS-B14sh and CS-B22Lo have higher lint yield and higher seeds, and CS-B25 has increased fiber length and strength, reduced micronaire property, and nearly the same yield as TM-1. The genes associated with cotton fiber development in the chromosomal substitution line CS-B22sh has been recently studied using suppression subtraction hybridization and microarray analysis (Wu et al., 2008). Thirty-six fiber specific genes have been identified to be differentially expressed in CS-B22sh when compared with TM1. Since CS-B25 has better fiber qualities and it is nearly isogenic to the parent TM-1 for 25 pairs of chromosomes, a comparative analysis of gene expression

in fiber between TM-1 and CS-B25 will provide a useful tool to identify genes on the chromosome 25 of *G. barbadence* 3-79 associated with good fiber traits.

CHAPTER III

MATERIALS AND METHODS

Plant materials

Cotton plants (TM-1 and CS-B25) were grown at the plant science research center at Mississippi State University. Cotton flowers were tagged on the day of anthesis (0 DPA) and cotton bolls were collected at 10 DPA. All Chemicals and reagents used in this study were purchased from Sigma (St. Louis, MO), Clontech (Mountain View, CA), GE Healthcare Life Sciences, Promega (Madison,WI), and Applied Biosystems (Carlsbad, CA). Deoxy-oligonucleotide primers were obtained from Invitrogen (Carlsbad, CA).

Isolation of total RNA from TM-1 and CS-B25

Cotton fibers were manually removed from seeds using sterile forceps. Approximately 1 g of fiber was ground to fine powder in a pre-cooled mortar and pestle in the presence of liquid nitrogen. The ground powder was immediately transferred to a 50 ml tube and mixed with 5ml of preheated extraction buffer (0.2 M Borax, 30 mM EGTA, 1% (w/v) SDS, 1% (w/v) sodium deoxycholate, 2% (v/v) PVP-40, 0.5% (v/v) NP-40), 50 µl of DTT (1 M), and 100 µl of Proteinase K (25 mg/ml) by vigorously vortexing. The sample was incubated at 42˚C for 1.5 hr with mild agitation. The sample solution was then adjusted to 160 mM KCl with 1 M stock solution, chilled on ice for 1 hr, and centrifuged at 12,000 xg for 20 min at 4˚C. The supernatant was transferred to a sterile 50 ml tube and one third volume of 8 M LiCl was added to a final concentration of 2 M. The sample was incubated on ice overnight and centrifuged at 12,000 xg for 20 min

at 4˚C. The supernatant was discarded and the RNA pellet was washed twice with cold (4˚C) 2 M LiCl by dispensing the pellet via gentle vortexing and followed by centrifugation at 12,000 xg. The pellet was dissolved in cold 10mM Tris-HCl (pH 7.5), and the insoluble materials were removed by centrifugation at 12,000 xg for 10 min at 4˚C. The supernatant was transferred into a fresh 50 ml tube and 100 µl of potassium acetate (2.5 M, pH 5.5) were then added to precipitate residual proteins and polysaccharides. After centrifugation, the supernatant was transferred to a new sterile 50 ml tube and RNA was precipitated with 2.5 volumes of 96% ethanol at -20˚C overnight. The RNA pellet is collected by centrifugation at 12,000 xg at 4˚C for 20 min, washed with 5 ml of 70% ethanol, and vacuum dried. The isolated RNA was resuspended in 200 µl of DEPC-treated water and stored at -80˚C.

Isolation of fiber poly (A) RNA

An mRNA purification kit from GE Healthcare was employed to isolate poly(A) RNA from total fiber RNA. An oligo(dT)-cellulose column from the kit was resuspended by inverting the column several times. The column is placed in a 15 ml sterile tube and the storage buffer in the column was then drained out. One ml of high salt buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA, and 0.5 M NaCl) was applied to the column and allowed the buffer to drain by gravity. This high salt buffer wash was then repeated one more time. The total RNA sample mixed with 200 µl of sample buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA, and 3 M NaCl) was preheated at 65˚C for 5 min, immediately cooled on ice, and applied to the column. The contents were allowed to flow into cellulose bed by gravity and then centrifuged at 350 xg for 2 min. After further washing with high salt buffer, 250 µl of low salt buffer (10 mM Tris-HCl, pH 7.4, 1mM EDTA, and 0.1 M

NaCl) was applied to the column and followed by centrifugation at 350 xg for 2 min. The wash and centrifugation were repeated twice. The flow through from the column was discarded. The column was then placed in a new sterile 1.5 ml microcentrifuge tube, and 250 µl of preheated elution buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA) at 65˚C were applied to the column. The column was centrifuged at 350 xg for 2 min, and followed by three times elution each with 250 µl aliquots of the elution buffer. A total of 1 ml eluent was collected and the eluent was equally divided into two tubes. To each tube 50 µl of sample buffer, 5 µl of glycogen solution (10 mg/ml) and 1.25 ml of cold ethanol (96%) were added and kept at -20˚C for 2 hr. The samples were centrifuged at 12,000 xg for 10 min at 4˚C and the RNA pellet was then vacuum dried. The pellet was finally dissolved in 6 µl of DEPC-treated H_2O and stored at -80˚C.

Identification of differentially expressed genes via suppression subtraction hybridization

Synthesis of double stranded cDNA

Poly(A) RNA isolated from 10 DPA fibers was used for the synthesis of double stranded cDNA with a SMARTer PCR cDNA synthesis Kit (Clontech). First strand cDNA synthesis was initiated by mixing 1 µl of Poly (A) RNA (0.5 ng/µl), 1 µl of 3' SMART CDS primer II A (12 µM, 5'-AAGCAGTGGTATCAACGCAGAGTACT$_{(30)}$N-$_1$N-3'), and 2.5 µl of sterile H_2O in a 0.6 ml microcentrifuge tube. The sample was incubated at 72˚C for 3 min and then 42˚C for 2 min. A master reaction mixture containing 2 µl of 5X first strand buffer (250 mM Tris-HCl, pH8.3, 375 mM KCl, and 30 mM $MgCl_2$), 0.25 µl of DTT (100 mM), 1µl of dNTP mix (10 mM each), 1 µl of SMARTer II A Oligonucleotide (12 µM, 5'-AAGCAGTGGTATCAACGCAGAGTAC-

XXXXX-3', X represents undisclosed base), 0.25 µl of RNase inhibitor (40 U/µl), and 1 µl of SMARTScribe Reverse Transcriptase (100 U/µl) was then added to the 0.6 ml microcentrifuge tube. The reverse transcription reaction was carried out at 42˚C for 1 hr, and the reaction was terminated by heating the tube at 70 ˚C for 10 min. The cDNA sample was diluted by adding 140 µl of TE buffer and stored at -80˚C. The diluted first strand cDNA was used to synthesize second strand cDNA via PCR amplification. The PCR master mixture was prepared by mixing 30 µl of 10X Advantage 2 PCR buffer, 6 µl of 50X dNTP mix (10 mM), 6 µl of 5' PCR Primer IIA (12 µM), 6 µl of 50X Advantage 2 Polymerase mix, 30 µl of first strand cDNA, and 222 µl of ddH_2O into a sterile 0.6 ml microcentrifuge tube. Aliquots (100 µl) from the PCR mixture were individually transferred to three 0.2 ml PCR tubes labeled as A, B and C and then subjected to PCR amplification. The PCR reaction program was set at 95˚C for 1 min for pre-denaturation and followed by 27 cycles of 95˚C for 15 sec, 65˚C for 30 sec and 68˚C for 3 min. At the 15^{th} cycle of amplification, the program was paused and 70 µl from each tube (A, B and C) were transferred to a new PCR tube. The PCR reaction cycle was resumed and 5 µl of aliquot was transferred to a 0.6 ml tube after every third cycle (18, 21, 24 and 27 cycles). The PCR products (5 µl of aliquot) was electrophoresed on a 1.2% (w/v) agarose gel (20 mM Tris-HCl, pH 7.5, 7.5 mM sodium acetate, 0.5 mM EDTA, 1 µg/µl Ethidium bromide) with 1X TAE buffer (20 mM Tris-HCl, pH 7.5, 7.5 mM sodium acetate, 0.5 mM EDTA) at 60 V for 1.5 hr. The gel was examined and photographed under UV illumination. The optimal number cycles for the given sample was determined from the gel picture. The aliquot of 70 µl obtained at the 15^{th} cycle was pooled into one tube and additional PCR cycles were run until the optimal cycle number was reached.

Purification of PCR product

The PCR product was purified using a PCR purification kit obtained from Qiagen. The PCR product (210 µl) was mixed with 5 volumes (1,050 µl) of PB buffer in a fresh 1.5 ml microcentrifuge tube. A column containing silica gel membrane was placed in a 2 ml tube provided from the kit and the mixed sample was applied to the column. After centrifugation at 12,000 xg for 30 sec, 250 µl of PE buffer was applied to the column and centrifuged again at 12,000 xg for 30 sec. An additional spin for 1 min at 12,000 xg was then performed to remove traces of PE buffer. The column was last placed in a fresh 1.5 ml microcentrifuge tube and cDNA bound to the column was eluted with 50 µl of sterile H_2O.

Restriction digestion of cDNA with *Rsa* I

The double strand TM1 and CS-B25 cDNA samples synthesized from Poly(A) RNA were digested with the restriction enzyme *Rsa* I according to the manufacturer's instructions. The purified ds cDNA (43.4 ml) was mixed with 5 µl of 10X *Rsa* I restriction buffer (100 mM Bis Tris propane-HCl, pH 7.0, 100 $MgCl_2$ and 1 mM DTT) and 1.5 µl of *Rsa* I (10 units/µl) in a 0.6 ml tube. The DNA sample was digested at 37˚C for 1.5 hr. To examine the *Rsa* I digestion efficiency, 5 µl of digested DNA sample were analyzed by electrophoresis on a 1% Agarose-Ethidium bromide gel. The digestion reaction was terminated by adding 2.5 µl of 20X EDTA/Glycogen mixtures (0.2 M EDTA and 1mg/ml glycogen), and the digested DNA samples were purified using the PCR purification kit as previously described. The purified DNA fragments were vacuum dried and the DNA pellet was dissolved in 6 µl of sterile distilled H_2O.

Synthesis of adaptor-ligated CS-B25

The *Rsa* I digested CS-B25 (tester) cDNA was used for the synthesis of adaptor-ligated double stranded cDNA using two specific adaptors; adaptor 1 (5' CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT 3') and adaptor 2R (5' CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGT 3') provided by Clonetech. One µl of *Rsa* I digested CS-B25 cDNA was diluted with 5 µl of sterile H_2O and the diluted cDNA was used for ligation reaction. A ligation master mix was first prepared by mixing 3 µl of sterile H_2O, 2 µl of 5X ligation buffer (250 mM Tris-HCl pH 7.8, 50 mM $MgCl_2$, 10 mM DTT and 0.25 mg/ml BSA), and 1 µl of T4 DNA ligase (400 units/µl) in a sterile 0.6 ml tube. Two separate adaptor ligation reactions were set up by mixing 2 µl of diluted CS-B25 cDNA, 2 µl adaptor 1 (or adaptor 2R) (10 µM), and 6 µl of master mix in two 0.6 microcentrifuge tubes. In a fresh 0.6 ml microcentrifuge tube 2 µl of ligation reaction 1 and reaction 2 were mixed, which acted as an unsubtracted CS-B25 control. The adaptor-ligation reactions were performed by incubating the samples at 16˚C overnight and terminated by adding 1 µl of EDTA/Glycogen and subsequent heating at 72˚C for 5 min. The samples were briefly centrifuged and stored at -20˚C.

First hybridization reaction

The CS-B25 tester cDNA ligated with adaptor 1 (or adaptor 2R) was individually hybridized with the *Rsa* I digested TM-1 driver cDNA. The 4X hybridization buffer was warmed up to room temperature for 20 min. The hybridization reaction was performed in two separate 0.6 ml microcentrifuge tube. In the first tube, 1.5 µl of *Rsa* I digested TM-1 driver cDNA was mixed with 1.5 µl of adaptor 1-ligated CS-B25 tester cDNA and 1µl of 4X hybridization buffer. The second tube had identical reagents as tube 1 except that the

2R-ligated tester cDNA (1.5 µl) was included for hybridization reaction. The samples were overlaid with a drop of mineral oil, incubated at 98˚C for 1.5 min, and followed by hybridization at 68˚C for 8 hr. A second hybridization reaction was performed immediately after the completion of first hybridization.

Second hybridization reaction

Second hybridization was performed by mixing the two samples obtained from the first hybridization with fresh denatured TM-1 driver cDNA to further enrich differentially expressed CS-B25 cDNA. Fresh denatured driver cDNA was prepared by mixing 1 µl of driver cDNA, 1 µl of 4X hybridization buffer, and 2 µl of sterile H_2O in a 0.2 ml sterile PCR tube. The sample was overlaid with a drop of mineral oil followed by incubation at 98˚C for 1.5 min in a thermal cycler. The contents from sample 2 of the first hybridization product were carefully drawn by touching the oil/sample interface into a pipette tip. A small amount of air bubble was then created in the pipette below the sample 2 and freshly denatured driver cDNA was then drawn into the pipette. The entire contents were transferred into the tube containing sample 1 of the first hybridization product. The sample was mixed by pipetting up and down and followed by brief centrifugation. The second hybridization reaction was then performed by incubating the sample at 68˚C overnight. A 100 µl of dilution buffer (20 mM HEPES, pH 6.6, 20 mM NaCl, and 0.2 mM EDTA, pH 8.0) was added to second hybridization reaction and mixed. The sample was subsequently heated at 68˚C for 7 min and then stored at -20˚C.

PCR amplification of hybridization products

The differentially expressed genes in CS-B25 were selectively amplified via two PCR reactions. The PCR master mixture was prepared by mixing 2.5 µl of 10X PCR reaction buffer, 0.5 µl of dNTP mix (10 mM), 1 µl PCR primer 1 (5'-CTAATACGACTC-ACTATAGGGC-3', 10 µM), 0.5 µl of 50X Advantage cDNA polymerase mix, and 19.5 µl of sterile water in a 0.6 ml microfuge tube. Two separate PCR reactions were set up by mixing 24 µl of master mix with 1µl of dilute subtracted cDNA from second hybridization or 1µl of diluted unsubtracted CS-B25 tester control as template. The primary PCR reaction was conducted using MultiGene II thermal cycler (Labnet International) with a preheating step at 75˚C for 5 min to extend the adaptors, following by 27 cycles at 94˚C for 30 sec, 66˚C for 30 sec, and 72˚C for 1.5 min. The subtracted cDNA obtained from primary PCR amplification was further enriched in the second PCR amplification. Three µl of each primary PCR mixture (subtracted and unsubtracted) were diluted in 27 µl of ddH_2O and 1 µl of the diluents was used as template for second PCR amplification. The second PCR reaction mixture consisted of 2.5 µl of 10X PCR reaction buffer, 0.5 µl of dNTP mix (10 mM), 1 µl of Nested primer 1 (5'-TCGAGCGGCCGCCCGGGCAGGT-3', 10 µM), 1 µl Nested PCR primer 2R (5'-AGCGTGGTCGCGGCCGAGGT-3', 10 µM), 0.5 µl of 50X Advantage cDNA polymerase mix, and 18.5 µl of sterile water. The PCR reaction was conducted in a MultiGene II thermal cycler for 12 cycles at 94˚C for 30 sec, 66˚C for 30 sec, and 72˚C for 1.5 min. The samples were then stored at -20˚C.

Analysis of PCR product

The PCR products were analyzed by electrophoresis on a 2% (w/v) agarose gel containing 20 mM Tris-HCl, pH 7.5, 7.5 mM sodium acetate, 0.5 mM EDTA, 1 µg/µl

ethidium bromide with 1X TAE buffer (20 mM Tris-HCl, pH 7.5, 7.5 mM sodium acetate, 0.5 mM EDTA) at 60 V for 1.5 hr. The gel was examined and photographed under UV illumination.

Cloning of subtracted cDNA into T/A cloning vector

Preparation of *E.coli* XL1-Blue competent cells

E.coli XL1-Blue cells were inoculated into 2 ml of LB medium (Bacto-tryptone 10 g/l, Bacto-yeast extracts 5 g/l, NaCl 10 g/l, pH 7.0) and grown at 37 ˚C overnight with vigorous shaking. One ml of overnight culture was inoculated in 50 ml of LB medium and grown at 37 ˚C with vigorous shaking until the $O.D_{600}$ reached 0.2 -0.3. The culture was then kept on ice, and 20 ml of cell culture were transferred to a 50 ml centrifuge tube and centrifuged at 3,000 xg at 4˚C for 10 min. The supernatant was discarded, and the cells were resuspended in 20 ml of ice cold 0.1 M $CaCl_2$ and kept on ice for 15 min. The sample was subsequently centrifuged at 3,000 xg at 4˚C for 10 min and the cells were resuspended in 1 ml of 0.1 M $CaCl_2$. The prepared competent cells were immediately used for transformation.

Cloning of subtracted cDNA into pGEM-T Easy vector

The subtracted cDNA product from the second PCR amplification contained an Adenosine residue at 3' ends which can be cloned directly into a T/A cloning vector pGEM-T Easy. Briefly, 1 µl of subtracted cDNA, 5 µl of 2X rapid ligation buffer (60 mM Tris-HCl, pH 7.8, 20 mM $MgCl_2$, 20 mM DTT, 2 mM ATP, and 10% polyethylene glycol), 1 µl of pGEM-T Easy cloning vector (50 ng/µl), 1 µl of T4 DNA ligase (3 units/µl), and 2 µl of sterile H_2O were mixed in 0.6 ml microcentrifuge tube and then incubated at 4˚C overnight. To the ligation mixture, 200 µl of freshly prepared competent

cells were added and mixed by gently vortexing. The sample was incubated on ice for 30 min, heat shocked at 42˚C for 2 min, and then chilled on ice for 30 min. The transformed cells were mixed with 500 µl of LB medium and incubated at 37˚C for 1 hr. Aliquots of transformed cells along with 15 µl of Isopropyl-β-D-thiogalactopyanoside (IPTG) (200 mg/ml) and 10 µl of 5-bromo-4-chloro-3-indolyl-β-D-galacto-pyranoside (X-gal) (50 mg/ml) were spread on LB/ampicillin (50 µg/ml) agar plates. The plates were incubated at 37˚C overnight.

Isolation of recombinant pGEM-T Easy plasmids

Twenty five white colonies were picked from LB/amp plates and transferred into a 15 ml culture tube containing 2 ml of LB medium with 50 µg/ml ampicillin. The culture tubes were incubated at 37˚C overnight, and 1 ml of culture was transferred to 1.5 ml microcentrifuge tube and centrifuged at 14,000 xg for 10 min. The supernatant was discarded and the pellet was resuspended in 100 µl of P1 buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA) by vortexing. To the cell suspension, 200 µl of P2 lysis buffer (200 mM NaOH, 1% (w/v) SDS) were added and gently mixed. The cell lysate was then gently mixed with 150 µl of P3 Neutralization Buffer (3M potassium acetate, pH 5.5), incubated at room temperature for 10 min, and then centrifuged. The supernatant was transferred into a fresh 1.5 ml microcentrifuge tube, mixed with 1 ml of 96% ethanol, and kept at -20˚C for 30 min. The sample was centrifuged at 14,000 xg for 10 min, and the supernatant was discarded. The DNA pellet was then washed with 1 ml of 70% ethanol, vacuum dried, and resuspended in 50 µl of sterile ddH_2O.

Sequencing recombinant pGEM-T Easy Plasmids

Recombinant plasmid DNA was sequenced using an ABI PRISM Big dye Terminator Cycle sequencing Ready Kit (Applied Biosystems) with ABI PRISM 310 DNA Genetic analyzer. The sequencing reaction mix contained 2 µl of Plasmid DNA (1.2 µg/µl), 1 µl of T7 primer (5' TAATACGACTCACTATAGGGCGA 3', 10 µM), 4 µl of Terminator Ready Reaction Mix, and 17 µl of sterile water. The cycling sequencing reaction was conducted by PCR via Pre-denaturation at 96˚C for 2 min and followed by 25 cycles at 96˚C for 10 sec, 50˚C for 5 sec and 60˚C for 4 min in a MultiGene II thermal cycler. The reaction mixture was then transferred into a new 1.5 ml microcentrifuge tube containing 50 µl of 96% of ethanol and 2 µl of 3M sodium acetate, pH 4.8. The sample was kept at room temperature for 15 min and centrifuged at 14,000 xg for 15 min. The DNA pellet was washed with 250 µl of 70% ethanol, dried under vaccum, and resuspended in 20 µl of Hi-Di Formamide. The DNA samples were denatured at 95˚C for 2 min, immediately cooled on ice, and loaded onto the ABI PRISM 310 DNA Genetic Analyzer (Applied Biosystems). The DNA sequence data was collected and analyzed using online softwares BLAST at NBCI, TAIR and UNIPORT.

Identification of differentially expressed genes in CS-B25 via Microarray analysis

Affymetrix cotton genome arrays that contain 23,977 probe sets representing 21,854 cotton transcripts were used to analyze gene expression in 10 DPA fiber in CS-B25 and TM-1 lines. Three replications of RNA samples of CS-B25 and TM-1 were used for array analyses. The double-strand cDNAs prepared using total RNA from the two lines were reverse transcribed using Affymetrix One Cycle cDNA synthesis reactions. The cDNA was converted to cRNA and labeled with Biotin by an in vitro transcription (IVT) method using T7 RNA polymerase and biotin-labeled ribonucleotides. The cRNA

was then used as probe to hybridize to the cotton genome microarray. The DNA array was then washed and fluorescent hybridization signals were scanned using laser confocal devices. The microarray hybridization and post-hybridization scanning of arrays were conducted in the Genomics Core Facilities, Howell and Elizabeth Heflin Center for Human Genetics, The University of Alabama at Birmingham (UAB). The signal intensity obtained for individual spots correlating with the concentration of target mRNA. Data mining was also conducted using microarray data analyzing software, statistical programs and algorithms to determine if the gene of interest is up-regulated, down-regulated, or unchanged. Data was then filtered and processed at the University of South Dakota, Department of Mathematics.

Quantitative Real Time PCR analysis

Synthesis of first strand cDNA

The total RNA samples isolated from TM1 and CSB25 initially used for microarray analysis were used for Real-time RT-PCR analysis. One μg of DNase I-treated total RNA was used to synthesize first strand cDNA. A reverse transcription reaction mixture (total 20 μl) containing 1 μl of total RNA (1 μg/μl), 2 μl of 10X M-MLV reverse transcriptase buffer, 2 μl of dNTPs (5 mM each), 1 μl of random primer (10 μM), 1 μl (40 U/μl) of RNA OUT inhibitor (Invitrogen), and 1 μl of M-MLV reverse transcriptase (Promega), and 12 μl of ddH_2O was prepared in a 0.6 ml tube. The tube was incubated at 42°C for 1 hr, and the synthesized cDNA sample was stored at -20°C.

Real Time RT-PCR analysis of 10 DPA fiber transcripts in CS-B25 and TM-1 lines

Real time RT-PCR was used to validate the expression levels of genes identified via SSH and microarray analysis. The real time RT-PCR analysis was conducted using

the Power SYBR Green PCR Master Mix kit and the 7500 Fast Real-Time PCR system (Applied Biosystems, Foster City, CA) in a manner of three technical replications per RNA (cDNA) sample. For each technical replication, a master mixture was prepared as the counts of total number of examined samples plus at least one additional reaction, and contained all required reaction reagents excluding cDNA template. After mixing and short centrifugation, a 9.0-µl aliquot was transferred to individual 0.2-ml PCR tubes followed by adding cDNA template into each tube. Thus, each real time RT-PCR reaction mixture contained 1 µl of cDNA template, 5 µl of SYBR Green Mix, 0.25 µl of forward primer (10 µM), 0.25 µl of reverse primer (10 µM), and 4 µl of ddH_2O. Each reaction mixture was finally transferred to individual well of the Fast Optical 96-well plate (ABI, Foster City, CA), which was then sealed with a piece of optical adhesive film (ABI, Foster City, CA) before loading to the 7500 Fast Real-Time PCR system (ABI, Foster City, CA). The real time PCR analysis was performed with a program of initial denaturation at 95°C for 10 min, following by 40 cycles of 95°C for 15 sec, and 60°C for 1 min. To validate the specificity of each amplicon and the primer-dimer formation, a dissociation curve analysis was included after each real time PCR analysis. The data of each real time RT-PCR assay was analyzed based on the comparative Ct ($2^{-\Delta\Delta Ct}$) method (Livak and Schmittgen, 2001). The *18S rRNA* gene was used as a reference gene in the study for the data analysis.

Genes and primer sequences for validating transcript levels include, 18S RNA (control) (Forward primer- 5' GGAATTGACGGAAGGGCACCACCAGGC3' and Reverse primer- 5' GGACATCTAAGGGCATCACAGACCTG3'), LTP 3 (Forward primer- 5'ATTAATTACTTGAGAGGCAG3' and Reverse primer- 5' GAGTCCGCTTGCAAGGCCAA 3'). LTP 4 (Forward primer- 5'

ATTGGTTACTTGACAGGGAA 3' and Reverse primer- 5'

GAGTCCGCTTGCAAGGCCAT 3'), E2 (Forward primer- 5'

TCTGCTCTCAATCTGCTGGT 3' and Reverse primer- 5'

CACACCAACCACTGTCATCC 3'), RING (Forward primer- 5'

GCCGTCTCCAACCTCTCAC 3' and Reverse primer- 5'

CGGACCTTAACCTTCTCACC 3'), AGU (Forward primer- 5'

CGGAAGTGGAGCCAAGATAG 3' and Reverse primer- 5'

TCTGTGGAGCTGGTGTGGT 3'), CYP (Forward primer- 5'

TACCCTCAAAGCTGGTGGAC 3' and Reverse primer- 5'

TGATAGCCACTGCAAGAGGA 3'), PROK (Forward primer- 5'

ATGCTGGATCTTTACCCTCA 3' and Reverse primer- 5'

CAAGTGTTGCTTCCCACTCA 3'), ERF (Forward primer- 5'

GAAGCCGCCAAGCGTATC 3' and Reverse primer- 5' CTCCAACTCACTCTCCACCA

3'), RING1 (Forward primer- 5' CAAACAGGACAGACCGAGTTC 3' and Reverse

primer- 5' AGGCTGGTCAGAAAGGGTTG 3'), RIB (Forward primer- 5'

TCATTGTGATGGCTGCTGAT 3' and Reverse primer- 5'

TGGCTTCCTTCGTTTGATGT 3') and DHF (Forward primer- 5'

TTGAAGACCACATTTCCCAAG 3' and Reverse primer- 5'

TCATTAGGCAAGGGCAGTTT 3').

CHAPTER III

RESULTS

Isolation of poly (A) RNA from 10 DPA cotton fiber tissues

Total RNA was isolated from 10-DPA fiber of CS-B25 and TM-1 lines using a modified hot Borate method. The concentration of the RNA samples was determined using a NanoDrop 1000 spectrophotometer and found to be 1.2 µg/µl and 0.98 µg/µl for CS-B 25 and TM-1, respectively. The integrity of the RNA was examined on a 1% ethidium bromide- Agarose gel (Figure 1). The appearance of two rRNA bands and their relative intensity indicated that the RNA samples were intact and not degraded. A total of five RNA samples were isolated from each cotton line.

Poly(A) RNA samples were subsequently isolated from total RNAs using a mRNA purification kit. The concentration poly(A) RNA was also determined using the NanoDrop 1000 spectrophotometer, which was found to be 20.1 ng/µl and 32.7 ng/µl for TM1 and CS-B25, respectively. The two poly(A) RNA samples from CS-B25 and TM-1 were also examined on a 1% Agarose gel as shown in Figure 2.

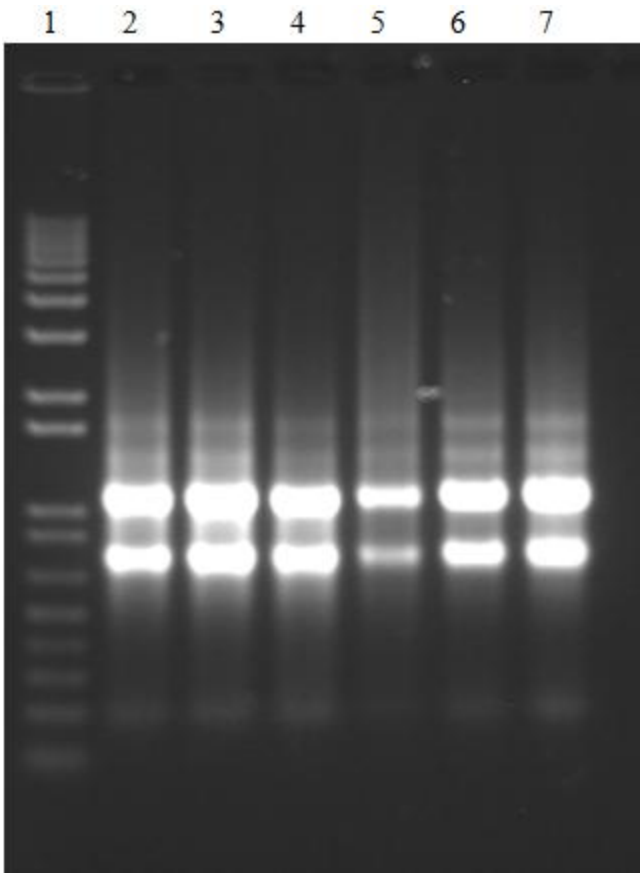

Figure 1 Analysis of total RNA isolated from 10-DPA fiber of TM-1 and CS-B25 on a 1% Agarose gel. Lanes 2, 3 and 4 represent total RNAs from TM-1 and Lanes 5, 6 and 7 are total RNAs from CS-B25. Lane 1 is 1 Kb plus DNA marker.

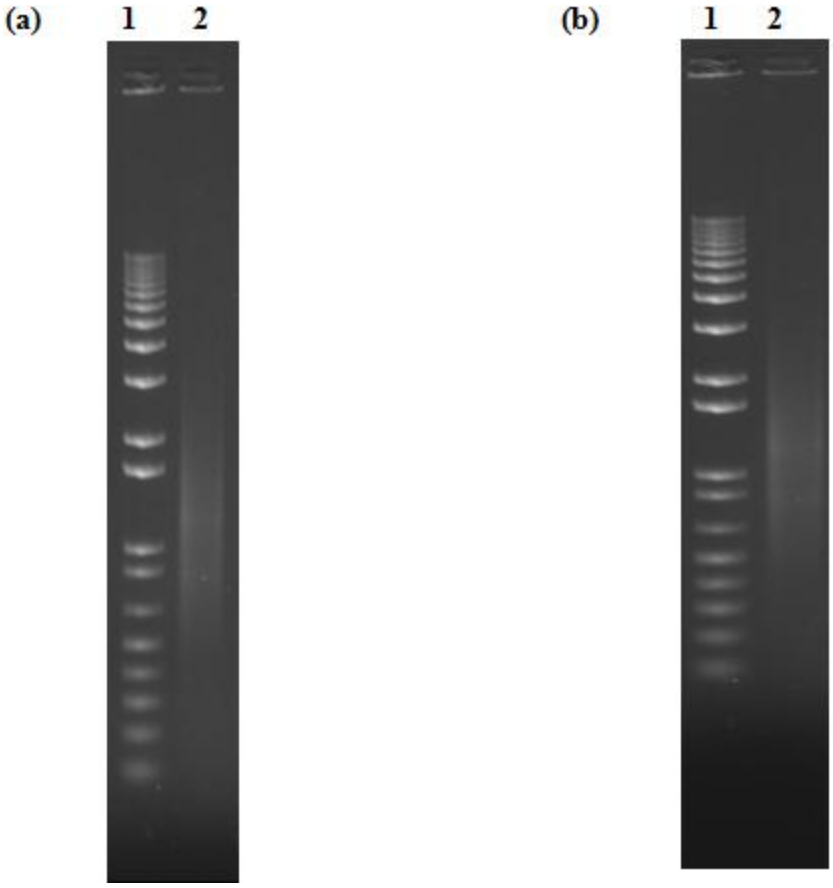

Figure 2 Analysis of 10-DPA Poly(A) RNA isolated from TM-1 (a) and CS-B25 (b) on a 1% agarose gel Lane 1 is kilobase plus DNA ladder and lane 2 contains fiber poly (A) RNA.

Identification of differentially expressed genes in CS-B25 via Subtraction hybridization

Subtraction hybridization was performed using a PCR select subtraction hybridization kit with TM-1 as driver and CS-B25 as tester. The double stranded cDNA from TM-1 and CS-B25 was digested with *Rsa* I restriction enzyme and the digested products were analyzed on a 1% agarose gel (Figure 3). The two cDNA samples appeared to be completely digested by *Rsa* I from the formation of small-size DNA fragments in lanes 3 and 5. The digested products were subsequently purified and their concentration was determined to be 9.68 and 9.55 ng/µl for CS-B25 and TM-1,

respectively. The digested tester cDNA was then ligated with specific adaptors (adaptor 1 and adaptor 2R). First hybridization and second hybridization reactions were performed by adding excess *Rsa* I digested TM-I (driver) cDNA. The unhybridized cDNA of CS-B25 was amplified using PCR, and 8 µl of the PCR product were analyzed on a 2% agarose gel (Figure 4). The subtracted cDNA was then cloned to the pGEM-T Easy vector. A total of 907 colonies grew on LB/ampicillin plates after transformation, which includes 402 white and 505 blue colonies. Four hundred white colonies were picked and plasmids were isolated using the standard alkaline detergent method. Twelve clones were selected and sequenced using the ABI310 genetic analyzer. The sequenced genes were subsequently identified using the NCBI- BLAST program. The putative function of these sequenced genes was annotated using a gene ontology program and shown in Table 1. The seven identified genes in CS-B25 via SSH analysis encode RNA binding protein, protein kinase family protein, papain-like cysteine proteinase, lipid transfer protein, drought-stressed protein, RING-H2 finger protein, and an ubiquitin-conjugating E2 enzyme.

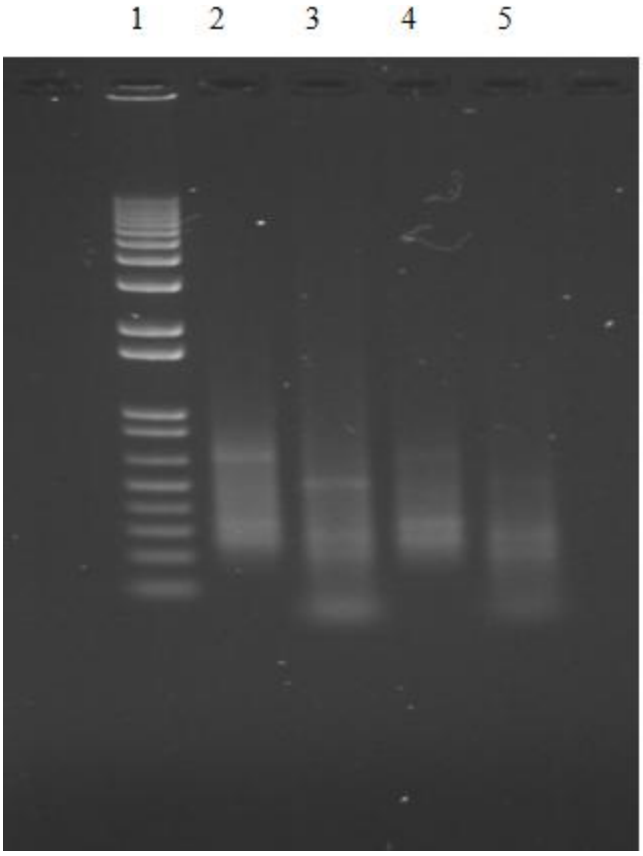

Figure 3 Analysis of *Rsa* I digested double-stranded cDNA from TM-1 and CS-B25 on a 1% agarose gel. Lanes 2 and 3 represent undigested and digested TM-1 cDNA, and lanes 4 and 5 are undigested and digested CS-B25 cDNA. Lane 1 is 1Kb plus DNA marker.

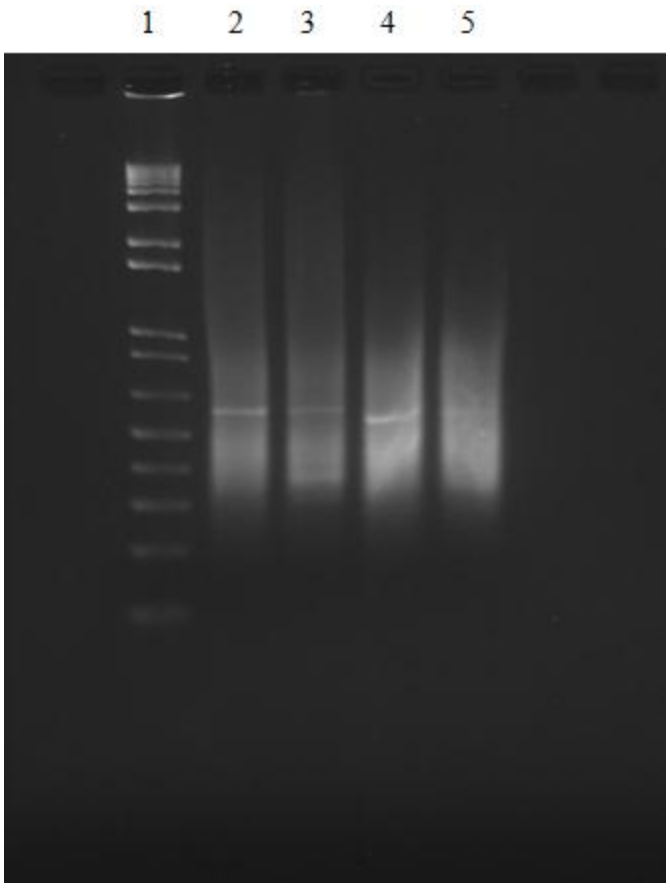

Figure 4 Analysis of PCR products amplified unhybridized CS-B25 (tester) cDNA on a 2% agarose gel.

Table 1 Identification of up-regulated genes in CS-B25 via subtraction hybridization

GENES IDENTIFIED	Frequency	ANNOTATION	e-value	Functional category
RNA binding / protein binding [Arabidopsis thaliana]	1	NP_564591.1	1.07E-16	RNA metabolic process, gene expression
protein kinase family protein / peptidoglycan-binding LysM domain-containing protein [Arabidopsis thaliana]	1	AAF99862.1	4.67E-46	kinase activity, phosphotransferase activity, alcohol group as acceptor
putative papain-like cysteine proteinase [Gossypium hirsutum]	2	CAE54306.1	7.35E-33	cysteine-type peptidase activity, endopeptidase activity
Gossypium hirsutum aquaporin PIP1-2 mRNA	1	ABR68794.1	1.35E-100	transport, cellular process
LTP 4 Gossypium hirusitum/Fs LTP 2	3	ABO42261	1.01E-56	Lipid localization and transport
Populus EST from severe drought-stressed leaves	2	No significant similarities	3.01E-38	unclassified
RING-H2 finger protein ATL2N, putative [Ricinus communis]	1	EEF34075.1	1.76E-08	protein binding, metal ion binding
ubiquitin-conjugating enzyme E2 [Gossypium hirsutum]	1	EEF29726.1	2.60E-28	ubiquitin-dependent protein catabolic process

Identification of differentially expressed genes in CS-B25 via Microarray analysis

Affymetrix cotton genome arrays that contain 23,977 probe sets representing 21,854 cotton transcripts were used to analyze gene expression in 10 DPA fibers in CS-B25 and TM-1 lines. Three replications of RNA samples of CS-B25 and TM-1 were used for array analyses. The pre-scanned microarray data was normalized using the RMA (Robust Multiple-chip Average) algorithm and the presence/absence of a gene was identified using the Wilcoxon signed rank based genes expression presence/absence

program. The mean expression values obtained for the TM-1 and CS-B25 can be represented in the form of either a scatter plot or a volcanic plot. Both of these plots are useful tools for quick visualizing the variation within (between) arrays and also helpful in comparing the transcript levels between the two samples. Figures 5 and 6 represent scatter and volcanic plots, showing nearly symmetrical distribution of signal intensities across the median representing a linear association between two samples. Further, samples that have the most similar expression profiles are clustered together in the form of a dendrogram (Figure 7). This dendrogram or the Heat Map is useful for identifying outlying samples within the replicates. In order to identify differentially expressed genes, an unpaired Bayesian t-test followed by Multiple Testing Correction/False Discovery Rate (FDR) was applied to the microarray dataset. All the genes possessing statistical q-value greater than 0.25 with fold change less than 1 were filtered from the data set. A total of 198 which include 151 up-regulated and 27 down-regulated genes for CS-B25 line were identified. A further filtration of the data set with a fold change greater than 1.5 reduced the number of up-regulated and down-regulated genes to 21 and 9, respectively. The possible function of the identified genes was determined using online tools like BLAST at NCBI, TAIR and UNIPORT. The functional annotation was conducted using blast2go and Easygo for up- and down-regulated genes (Tables 2 and 3)

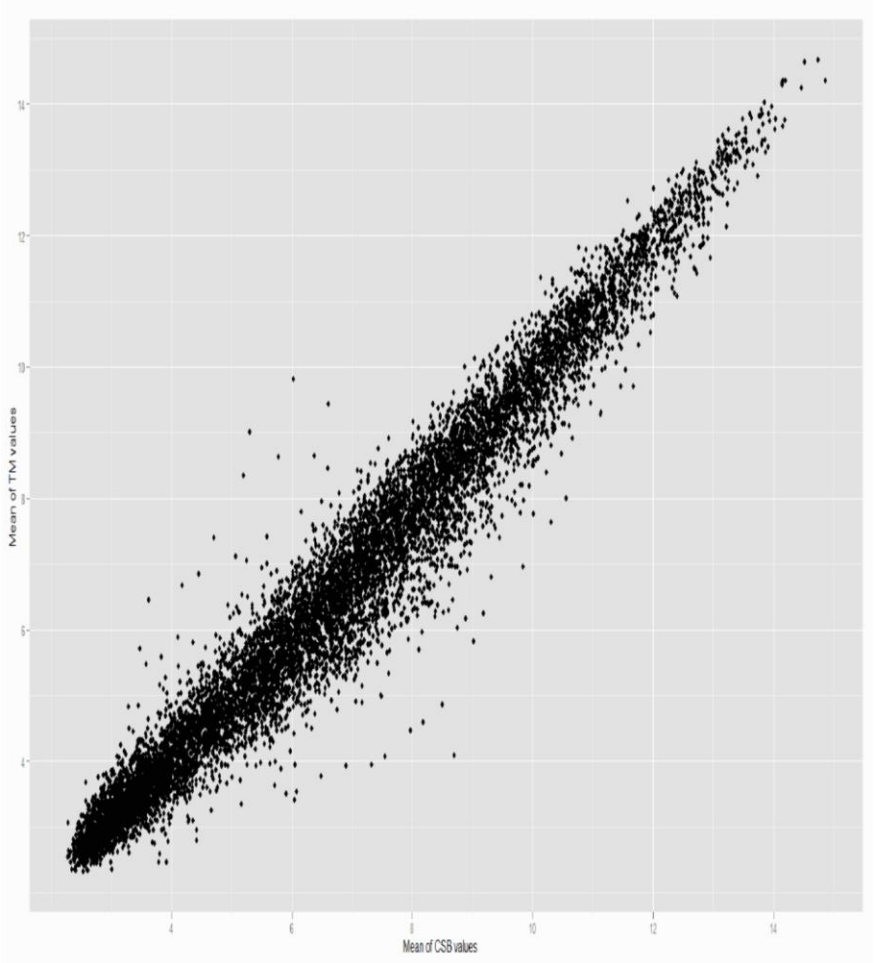

Figure 5 A Scattered Box-Plot of Normalized data (TM-1 vs CS-B25).

The microarray data was normalized using RMA algorithm. A box plot was prepared using CS-B25 on X-axis and TM-1 on Y-axis. The points present on median represent genes with relatively low fold change compared to the points presents away from median.

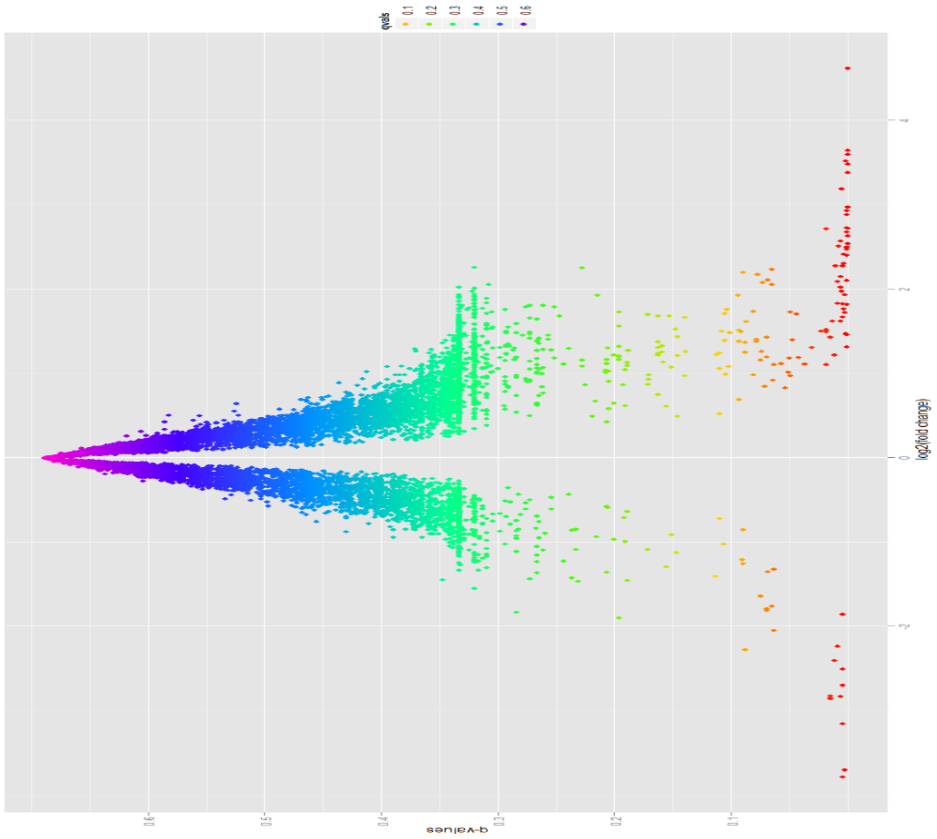

Figure 6 A volcanic plot of q-value vs fold change

A pair wise comparison between the FDR corrected TM-1 and CS-B25 probe sets were represented in a volcanic plot. The probe sets present outside the triangular region were considered to be regulated in the volcanic plot. All genes possessing statistical q-value greater than 0.25 were filtered out from the data.

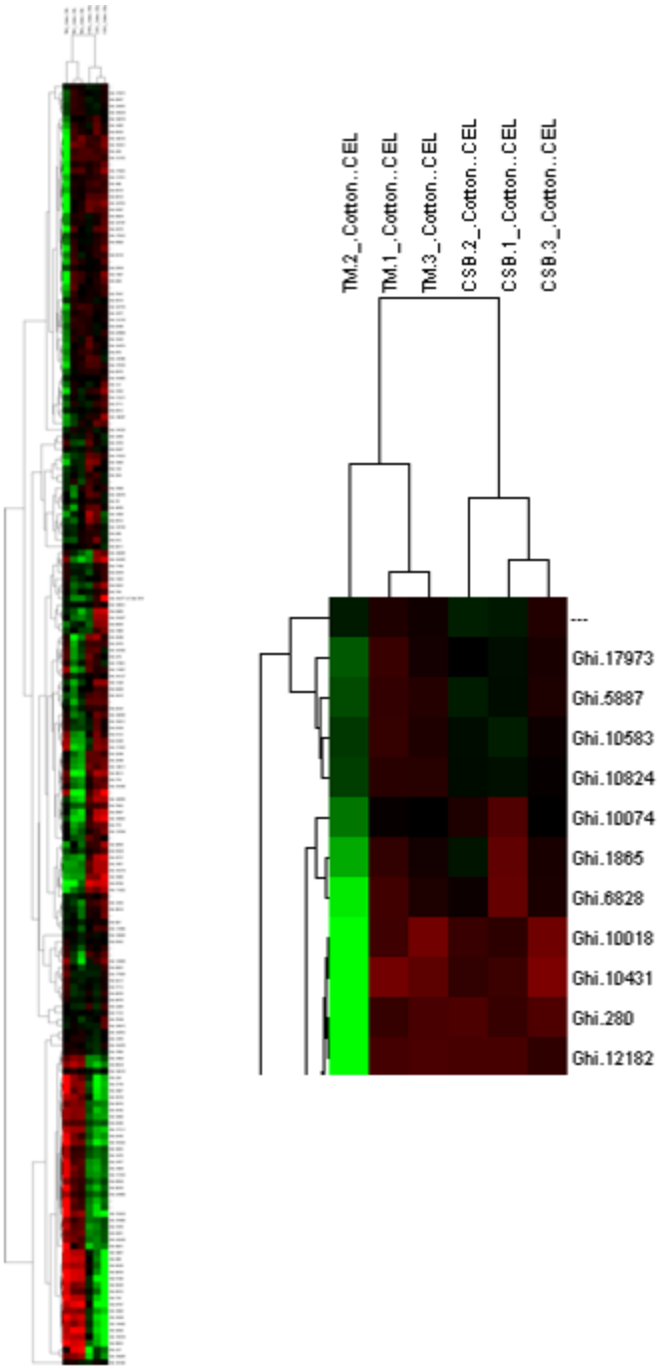

Figure 7 Heat maps of cotton genome arrays

Heat maps were generated based on image intensities by cluster analysis of specific genes in both TM-1 and CS-B25.

Table 2 Up-regulated genes identified from microarray analysis with fold change ≥1.5.

UNIGENE ID	Gene Description	E-Value	Similarity	Score	Fold change	GO ID's
Ghi.10753	putative protein stem-specific protein Nicotiana tabacum supported by full-length cDNA	2.35E-50	78	200.67	1.54	N/A
Ghi.1326	Encodes a basic helix–loop–helix transcription factor that is expressed in the hypophysis-adjacent embryo cells	8.94E-24	76	75.87	1.53	GO:0005634
Ghi.10655	Encodes a glutamine-dependent asparagine synthetase, the predicted ASN1 peptide contains a purF-type glutamine-binding domain	2.11E-11	76	71.24	1.78	GO:0009063
Ghi.3049	Haloacid dehalogenase like hydrolase family protein which is known to function as a hydrolase	1.22E-22	90	108.61	1.53	GO:0008152
Ghi.3049	Haloacid dehalogenase like hydrolase family protein which is known to function as a hydrolase	2.33E-53	68	210.69	1.59	GO:0016787
Ghi.3451	Encodes an enzyme putatively involved in trehalose biosynthesis. The protein has a trehalose synthase (TPS)-like domain	2.56E-33	85	144.05	1.77	GO:0003824
Ghi.4737	Encodes ribonuclease RNS3	2.00E-17	83	91.27	1.75	GO:0009718
Ghi.4813	unknown protein with molecular function and biological process unknown	N/A	N/A	N/A	1.68	N/A

Table 2 (continued)

Ghi.5997	Encodes a member of the *TBL (TRICHOME BIREFRINGENCE LIKE)* gene family containing a plant-specific DUF231	1.23E-22	81	108.61	1.72	GO:0012505
Ghi.6465	Encodes EXLB1 (expansin-like B1), a member of the expansin family.	1.13E-15	77	58.15	1.58	N/A
unknown	Arabidopsis NAC domain containing protein 47	2.18E-16	91	583.17	1.5	GO:0005488
Ghi.6780	DNAJ heat shock protein, putative; functions include: heat shock protein binding, protein folding	1.08E-26	76	122.09	2.1	GO:0031072
Ghi.16374	*Gossypium hirsutum* ACC oxidase 1 (ACO1) mRNA	4.62E-44	100	137.50	1.75	GO:0009693
Ghi.7066	Involved in biological process unknown. Located in cell wall, vacuole. known to express in 22 plant structures	8.38E-46	70	186.03	1.85	GO:0016023
Ghi.16277 /// Ghi.7874	Encodes a member of the ERF (ethylene response factor) subfamily B-2 of the plant specific ERF/AP2 transcription factor family	N/A	N/A	N/A	1.58	N/A
Ghi.16693	Encodes a protein that appears to have 1-amino-cyclopropane-1-carboxylic acid oxidase activity based on mutant analyses.	6.26E-19	100	96.28	1.85	GO:0009693
Ghi.14611	BURP domain-containing protein / polygalacturonase, putative; Functions: polygalacturonase activity	2.54E-34	70	139.42	1.54	N/A

Table 2 (continued)

N/A	Hypothetical protein	2.45E-32	75	142.12	1.79	N/A
Ghi.11202	Unknown protein with molecular function and biological process unknown located as cellular component	N/A	N/A	N/A	1.72	N/A
Ghi.4550	zinc finger (C3HC4-type RING finger) family protein; functions include ubiquitin-protein ligase activity, protein binding	1.13E-28	85	128.64	1.5	GO:0008270
Ghi.778	zinc finger (C3HC4-type RING finger) family protein; functions include protein binding, zinc ion binding.	1.37E-34	72	148.28	1.5	GO:0008270
Ghi.12182	zinc finger (C3HC4-type RING finger) family protein; functions include protein binding, zinc ion binding	1.89E-27	86	115.54	1.34	GO:0008270
Ghi.944	zinc finger (C3HC4-type RING finger) family protein; functions include protein binding, zinc ion binding.	3.44E-46	80	185.26	1.28	GO:0008270

Table 3 Down-regulated genes identified from microarray analysis with fold change ≥1.5.

UNIGENE ID	Gene Description	E-Value	Similarity	Score	Fold change	GO ID's
Ghi.227	unknown protein	1.09E-07	64	58.9214	1.529729	N/A
Ghi.3944	unknown protein	N/A	N/A	N/A	1.539709	N/A
Ghi.4280	unknown protein	N/A	N/A	N/A	1.598694	N/A
Ghi.5198	unknown protein	N/A	N/A	N/A	1.780521	N/A
Ghi.6797	unknown protein	N/A	N/A	N/A	1.605947	N/A
Ghi.728	unknown protein	N/A	N/A	N/A	1.69766	N/A
Ghi.9030	unknown protein	N/A	N/A	N/A	1.627004	N/A
Ghi.9683	Dihydroflavonol reductase, Catalyzes the conversion of dihydroquercetin to leucocyanidin in the biosynthesis of anthocyanins	7.3E-20	91	99.3673	1.642606	GO:0009813
Ghi.18424	ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein, functions include RNA binding and ribosome biogenesis	N/A	N/A	N/A	1.573492	N/A

Enrichment analysis of differentially expressed genes was conducted using agriGO. All genes identified as differentially expressed in CS-B25 were grouped together based on common GO's against cotton genome arrays. Enrichment analysis was performed for both up- and down-regulated genes (Figures 8 and 9). The results suggest that the genes expressed in response to stimulus form a significant proportion of the genes differentially expressed in CS-B25.

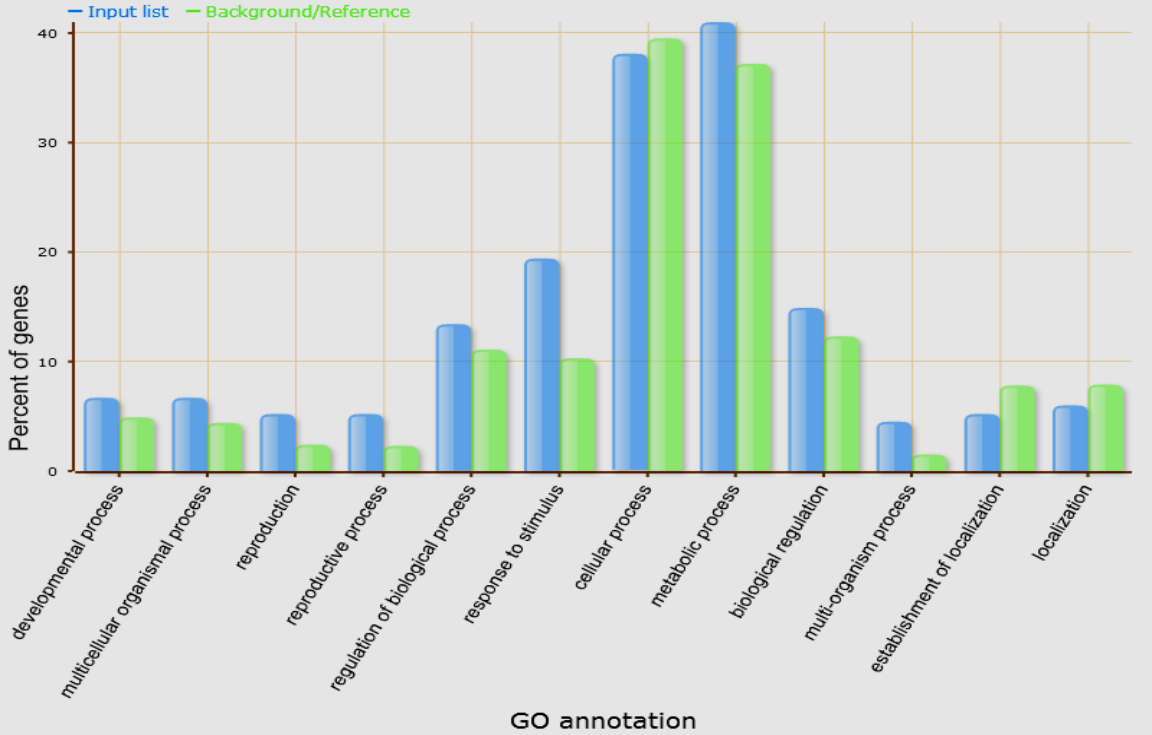

Figure 8 Enrichment analysis of differentially up-regulated genes in CS-B25 using agriGO.

All up-regulated genes in CS-B25 were grouped together in blue based on common GO's against cotton genome array represented in green.

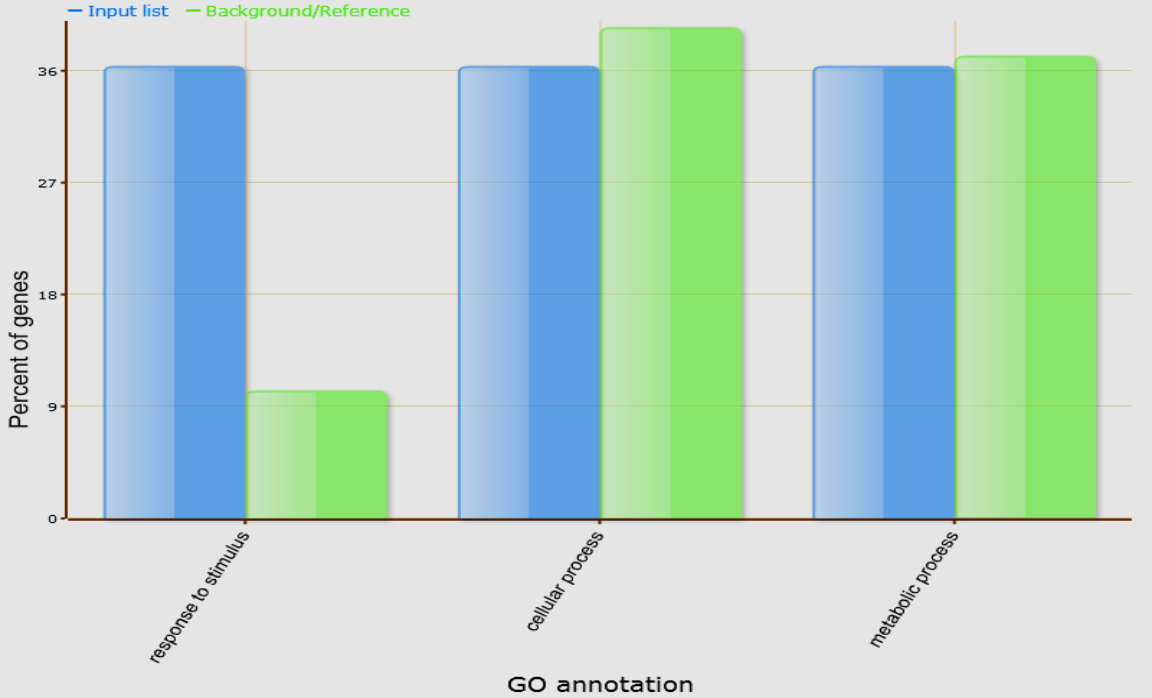

Figure 9 Enrichment analysis of differentially down-regulated genes in CS-B25 using agriGO.

All down-regulated genes in CS-B25 were grouped together in blue based on common GO's against cotton genome array represented in green.

Gene expression analysis of identified genes via Real Time RT-PCR in CS-B25

The transcript levels of differentially expressed genes have been validated using Real-time RT-PCR. As shown in Figure 10, genes for Aquaprotein (AGU), ERF and RING1 are expressed higher in CS-B25 in relative to TM-1 and Cystein proteinase (CYP), Dihydroflavonol reductase (DHS), Ribsomal binding protein (RIB) and E2 are lower in CS-B25 compared to TM-1. Expression profile for AGU, ERF, RING1, DHS and ribosomal binding protein are in agreement with subtraction hybridization and microarray data.

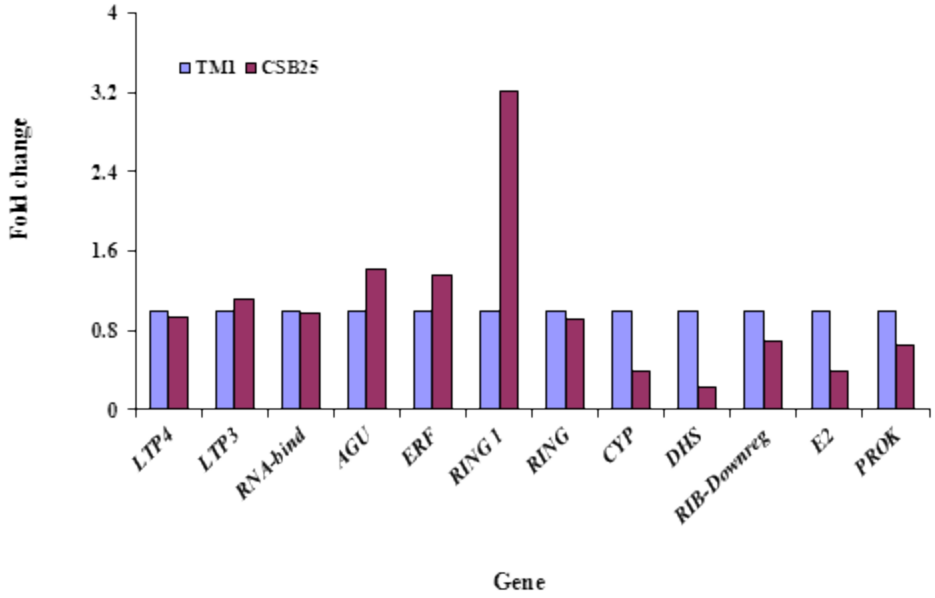

Figure 10 Validation of differentially expressed genes in CS-B25 by Real Time RT-PCR.

Genes for Aquaprotein (AGU), ERF and RING1 are upregulated and cystein proteinase (CYP), dihydroflavonol reductase (DHS), ribsomal binding protein (RIB), and E2 are down regulated.

CHAPTER IV

DISCUSSION

Suppression subtractive hybridization (SSH) and microarray analysis have been successfully used to identify genes that are differentially expressed in the 10-DPA CS-B25 fiber in this study. SSH analysis has identified 9 genes which are subsequently annotated using Gene Ontology program (Table 1). The microarray analysis has resulted in identifying 198 genes, of which 151 were up-regulated and 47 down-regulated in CS-B25. All these up-and down-regulated genes have fold change in expression less than 2.0. A reset of fold change value ≥ 1.5 resulted in reducing the number of the up-regulated genes to 21 and down-regulated genes to 9 (Tables 2 and 3). The full-length sequences of cDNAs identified via microarray analysis were obtained from Plex database and functional annotation was established using Gene Ontology program. The genes were further grouped under specific metabolic pathways. Several inferences can be made from the identified genes and their possible role during cotton fiber elongation is elucidated.

The events accompanied with cotton fiber development are highly regulated and phytohormone signaling pathways play an important role in overall development. The ethylene signaling pathway is known to play an important role in plants with response to biotic and abiotic stress, fruit ripening, and senescence. The role of ethylene in cotton fiber development is not well documented. Ethylene biosynthesis is tightly regulated by a variety of biotic and abiotic factors. S-adenosyl-methionine (S-AdoMet) and 1-aminocyclopropane-1-carboxylic acid (ACC) are the precursors in ethylene biosynthesis

(Peiser et al., 1984). Nearly 80% of cellular methionine is converted to S-adenosyl methionine (SAM) by S-AdoMet synthatase at the expense of ATP (Ravanel et al., 1998). SAM acts as a methyl group donor but also as a substrate in multiple biochemical pathways including ethylene biosynthesis. SAM is converted to ACC by the enzyme ACC synthase (ACS). The final step in ethylene biosynthesis is catalyzed by ACC oxidase, which requires oxygen and oxidizes ACC to form ethylene, CO2, and cyanide. ACS is the key enzyme that catalyzes the rate limiting step in the conversion of SAM to ACC. The regulation of *ACC* gene expression was studied in several plant species and well documented in Arabidopsis and tomatoes. In several plant species, ACC synthetase is encoded by a multigene family. ACC synthetase has a tertiary structure resembling pyridoxal 5'-phosphate dependent amino-transferases. Twelve *ACS* genes have been identified (*ACS*1 to *ACS*12) in the arabidopsis genome (Yamagami et al., 2003). Studies have shown that the *ACS*1 gene encodes a protein with the loss of function and *ACS*3 is a psuedogene (Liang et al., 1995). In contrast, *ACS4*, *ACS 5* and *ACS6* encode functional proteins and are induced by indole acetic acid (IAA) (Abel et al., 1995), cytokinin (Vogel et al., 1998) and ozone (Vahal et al., 1998), respectively. The dimerization of ACS resulting in post translation modification of C-terminal domain is found to be essential for its stability. Besides ACC synthase and ACC oxidase, three CUL3-BTB E3 (CUL 3a, 3b and 3c) unbiquitin ligases were also involved in regulating ethylene biosynthesis (Risseeuw et al., 2003). CUL3a and CUL3b are known to interact with C-terminal domain of the type 2 ACS which is targeted for degradation in proteasome (Yoshida et al., 2006). This explains the low levels of ACS and suggests that ethylene biosynthesis is tightly regulated. In cotton, the full-length *GhACS* cDNA encoding a 476-aa ACS protein has been isolated and characterized. The *GhACS* falls in the group II PLP

dependent aminotransferases and is expressed in response to wound, ABA, and $CuCl_2$ (Wang et al., 2007; Mishra et al., 2008). Shi et al. (2006) demonstrated that exogenous application of ethylene to cultured ovules resulted in robust fiber elongation and expansion. The increased fiber length was almost three fold when compared to untreated ovules. Accordingly, the presence of an ethylene biosynthetic inhibitor AVG resulted in reduced fiber length in cultured cotton ovules. The microarray analysis of 5-10 DPA cotton fibers conducted by Shi et al. (2006) has led to the identification of three functional *ACO* genes encoding 1-amino-cyclopropane-1-carboxylic acid oxidase. The gene expression profile indicated that adding galactose to cultured ovules also increased *ACO* gene expression and therefore ethylene biosynthesis. Qin et al. (2007) found that very long chain saturated fatty acids (VLCFA) (C24:0) promoted fiber elongation. It was suggested that very long chain fatty acids (VLCFA) might act upstream of the ethylene pathway by promoting biosynthesis of ACO. The microarray data from this research clearly showed an upregulation of ACO in CS-B25 with fold change of 1.75. The role of ACO in fiber elongation needs to be further studied.

Ethylene is perceived by a family of membrane proteins similar to histidine protein kinase receptors (Kendrick and Chang, 2008). The effect of ethylene on dark grown seedling often referred as 'TRIPLE RESPONSE' led to the identification of components involved in ethylene signal transduction pathway. In the absence of ethylene, the receptors are negative regulators of the ethylene signaling pathway. The downstream component to the receptor is CTR1, a member of the Raf family of mitogen activated kinase kinase kinase (MAPKKK). CTR1 regulates ethylene pathway by repressing the positive regulator ethylene insensitive (EIN2) protein. EIN2 is a member of NRamp family of metal ion transporters, and it regulates ethylene response by altering ion

concentrations. EIN2 relays signal to a family of transcription factors by an unknown mechanism that includes EIN3 and EIN3-like (EIL) proteins, which in turn activate ethylene responsive factor (ERF1). Binding of ethylene blocks the activation of CTR1, allowing EIN2 to positively regulate ethylene signal by an unknown mechanism. EIN2 levels are regulated by two F-box proteins, namely EBF1 and EBF2 that promote EIN2 degradation in the absence of ethylene (Qiao et al., 2009). In this research, an EIN3 like gene was identified in the microarray data set with a fold change of 1.22. In addition, two Skip-cullin-F-box E3 ligases, EBF1 and EBF2, that promote degradation of EIN3 and EIL1 via the 26S proteasome (Gou and Ecker, 2003; Potuschak et al., 2003; Lee et al., 2006) were also identified. The stability of EIN3 was thought to be controlled by two MAPKs; one required for stabilization and the other involved in degradation (Yoo et al., 2008). Because EIN3 is considered as a common point of convergence of ethylene, glucose, and light signaling pathways (Chao et al., 1997; Yanagisawa et al, 2003; Lee et al., 2006), it is possible that EIN2/ EIN3 levels are tightly regulated. This explains the observed low fold change. However, the complete nucleotide of EIN2/ EIN3 gene needs to be determined and studied for dissecting its role in ethylene signaling transduction. Downstream to EIN2, the ERF1 activates ethylene responsive genes PDF1-2 (Chen et al., 2005). Two ethylene response element binding proteins (GhEREB2 and GhEREB3) were isolated in cotton using a yeast one-hybrid system (Daun et al., 2006). Both of the proteins contain an acidic activation domain and one conversed DNA binding domain, and might act as transcription factors in biotic stress signal transduction pathways (Daun et al., 2006). Two cotton specific ERFs (GhERF4 and GHERF1) were isolated and characterized from *G. hirsutum*. Both of these proteins contain 58 amino acid long AP2/ERF domain and the RKRP nuclear localization signal. GhERF4 is a plant specific

transcription factor and plays an important role in response to ABA, ethylene and environmental stress (Qiao et al., 2008). Based on an in silico study and phylogenic reconstruction of conserved domains in ERFs, some ERFs were predicted to play a key role in the intergration of jasmonate, ethylene, and pathogen responses (Champion et al., 2009). Jin et al (2010) identified three novel ERFs, GhERF2, GhERF3 and GhERF6, which were thought to function as trans-acting factors in response to ethylene and stress. Protein degradation plays an important role in enhancing the sensitivity of ethylene signal response. Although the direct involvement of ERF proteins in cotton fiber development is not established yet, it can be implicated that ERF might play a role in activating downstream targets essential to cotton fiber development. ERF was identified via microarray analysis in this study, which is upregulated in CS-B25 with a relative fold change of 1.58 compared with TM1. Its expression level was also validated through Real-Time RT-PCR analysis. Previous studies showed that Arabidopsis ETR2 receptor is degraded rapidly in the presence of 26S proteasome, which resulted in dramatic increase in receptor gene synthesis. However, this increase did not affect fruit ripening (Gou and Ecker, 2003; Potuschak et al., 2003; Gagne et al., 2004). Hence, it is implicated that ERFs in cotton could be regulated via E3 F-box ligases. Two important ectoapyrases (GhAPγ1 and GhAPγ2) that function as dephosphorylases were identified in cotton (Clark et al., 2010). Ectoapyrases act on extracellular nucleotide tri-or di-phosphates and may play an important role in cotton fiber development. It has been demonstrated that ATP accumulation in extra cellular matrix causes an inhibition in fiber elongation. This condition was reserved upon the application of ACC precursor. The increased concentration of ATP resulted in increased fiber growth.

Ubiquitin mediated protein degradation is the most promeint mechaninsm utilized by eukaryotes to modulate protein levels in the cell. In this system, Ubiquitin (Ub) is covalently attached to the target protein which is marked for destruction by the multisubunit 26S proteasome (Herhko and Ciechanovre, 1998). Ubiquitin is attached to target protein by a sequential action of three enzymes: ubiquitin activating enzyme (E1), ubiquitn conjugating enzyme (E2) and ubiquitin protein ligase (E3) (Pickart, 2001). The Ub is activated to form E1-Ub intermediate in a two-step reaction mechanism which requires ATP. The first step involves activation of Ub, which results in the formation an acyl phosphoanhydride bond between AMP moiety of ATP and C-terminal glycine residue on the Ub. E1 then binds to activated AMP-Ub to form E1-Ub intermediate with simultaneous release of AMP. The E1-Ub complex is then transferred to a Cysteinyl residue on E3. The E2-Ub complex then either directly conjugate ubiquitin to the target protein in the presence of E3 ligase or forms an intermediate E3-Ub, which then transfers Ub to the substrate protein (Herhko and Ciechanovre, 1998). Binding of a single Ub result in local modification of the protein, this enhances binding of second Ub protein (Mukhopadhyay and Riezman, 2007). The process of reiteration results in the formation of Polyubiquitin chain, a substrate more often degraded by 26S proteasome (Voges et al., 1999). The Ubiquitin-proteasome system has a profound effect on many cellular processes like cell cycle, hormone signaling, defence, embryogenesis, circardian rythms and floral development. The E3 ubiquitin ligases are boardly classified into different protein families based on the type of interaction domain used by E2 conjugating enzymes. The majority of E3 ligases are broadly classified into two superfamilies containing either RING domain or U-Box domain/HECT domain (Gagne et al., 2002; Moon et al., 2004). In Arabidopsis the majority of E3 ligases contain a 70 amino acid

motif known as RING finger under the family Really Interesting New Gene (RING) proteins (Freemont, 2000). The RING finger is a zinc-binding motif that binds to the E2 conjugating enzyme. U-box E3 ligases are characterized by a 70 amino acid U-box domain which is structurally similar to RING motif but do not use zinc ions to stabilize secondary structure (Yee and Goring, 2009). The smallest E3 subfamily is HECT (Homology to E6-AP C Terminus) domain proteins that contain of approximately 350 amino acids with separate ubiquitin-binding site and E2-binding site (Pickart, 2001). In plants, there are more than 3 types of multisubunit RING E3s that are comprised of RING protein RBX1, a cullin scaffold-like protein, and additional substrate recognition protein that either bind to cullin directly or bind indirectly via adaptor proteins (Moon et al., 2004; Chen et al., 2006). Based on the type of the cullin domain, the RING proteins complexes are grouped into three broad categories, Skp-Cullin-F-box (SCF), Broad-complex, Tramtrack, Bric-a-Brac (CUL3-BTB), and DNA-DAMAGE BINDING 1 (CUL4-DDB1) (Santner and Estelle, 2010). Among the mutlisubunit E3 ligases, the SCFs are the best characterized and the most abundant proteins. The SCF group proteins contain four subunits namely SKP1 (ASK in plants), Cullin, F-box, and RING-Box1 (RBX1) (Herhko and Ciechanovre, 1998). Cullin servers as a scaffold protein with RBX1protein bound to C-terminus and F-box protein bound to N-terminus region with an adaptor SKP1 (Zheng et al., 2002).

Several proteins, including expansins, EXLB1, and TBL, involved in cellulose biosynthesis and present in primary cell wall have been identified in the microarray analysis. Expansins are known to be present in primary cell wall and play an important role in microtubulin rearrangement, thereby promoting cell elongation (Choi et al, 2006; An et al., 2007). *EXLB*1, a member of the expansin family, was up-regulated in CS-B25

with a 1.58 fold change relative to TM1. A gene similar to TBL (TRICHOME BIREFRINGENCE LIKE) gene family containing a plant-specific *DUF*231 domain has been also identified from the microarray data. This gene showed an up-regulation in CS-B25 with considerable 1.72 fold change when in comparison with TM1. Reports have shown that Arabidopsis trichome birefringence (*tbr*) mutant has severely reduced crystalline cellulose in trichomes (Bischoff et al., 2010). TBR and TBR-like3 are also known to be trancriptionally coordinated with primary and secondary cellulose synthase (CESA) genes.

REFERENCES

Abel, S., Nguyen, M. D., Chow, W. and Theologis, A. 1995. *ACS4*, a primary indoleacetic acid-responsive gene encoding 1-aminocyclopropane- 1-carboxylate synthase in *Arabidopsis thaliana*. J. Biol. Chem. 270: 19093–19099.

Aleman, L., Kitamura, J., Abdel-mageed, H., Lee, J., Sun, Y., Nakajima, M., Ueguchi-Tanaka, M., Matsuoka, M. and Allen, R. D. 2008. Functional analysis of cotton orthologs of GA signal transduction factors GID1 and SLR1. Plant Mol. Biol. 68: 1-16.

Albersheim, P., McNeil, M. and Labavitch, J.M. 1977. The molecular structure of the primary cell wall and elongation growth. Plant Growth Regulation (ed.), Pilet PP. 1-12.

Al-Ghazi, Y., Bourot, S., Arioli, T., Dennis, E. S., & Llewellyn, D. J. 2009. Transcript profiling during fiber development identifies pathways in secondary metabolism and cell wall structure that may contribute to cotton fiber quality. Plant Cell Physiol. 50: 1364-1381.

Amaratunga, D and Cabrera, J. 2004 Exploration and analysis of DNA microarray and protein array data. John Wiley and Sons Ch 4: 39-58.

An, C., Saha, S., Jenkins, J. N., Scheffler, B. E., Wilkins, T. A. and Stelly, D. M. 2007. Transcriptome profiling, sequence characterization, and SNP-based chromosomal assignment of the *EXPANSIN* genes in cotton. Mol. Genet. Genomics 278: 539–553.

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

Betancur, L., Singh, B., Rapp, R. A., Wendel, J. F., Marks, M. D., Roberts, A. W. and Haigler, C.H. 2010. Phylogenetically distinct cellulose synthase genes support secondary wall thickening in Arabidopsis shoot trichomes and cotton fiber. J. Integr. Plant Biol. 52: 205-220.

Bischoff, V., Nita, S., Neumetzler, L., Schindelasch, D., Urbain, A., Eshed, R., Persson, S., Delmer, D. and Scheible, W. R. 2010. TRICHOME BIREFRINGENCE and its homolog AT5G01360 encode plant-specific DUF231 proteins required for cellulose biosynthesis in Arabidopsis. Plant Physiol. 153: 590–602.

Botwell, D. and Sambrook, J. 2003. DNA microarray: a molecular cloning manual. CSHL press Ch 7: 509-594.

Canales, J., Flores-Monterrosso, A., Rueda-Lopez, M., Avila, C. and Canovas, F. M. 2010. Identification of genes regulated by ammonium availability in the roots of maritime pine trees. Amino Acids 39: 991-1001.

Cedroni, M. L., Cronn, R. C., Adams, K. L., Wilkins, T. A. and Wendel, J. F. 2003. Evolution and expression of MYB genes in diploid and polyploid cotton. Plant Mol. Biol. 51: 313-325.

Chae, H. S., Faure, F. and Kieber, J. J. 2003. The *eto1*, *eto2*, and *eto3* Mutations and Cytokinin Treatment Increase Ethylene Biosynthesis in Arabidopsis by Increasing the Stability of ACS Protein. Plant Cell 15: 545-559.

Champion, A., Hebrard, E., Parra, B., Bournaud, C., Marmey, P., Tranchant, C. and Nicole, M. 2009. Molecular diversity and gene expression of cotton ERF transcription factors reveal that group IXa members are responsive to jasmonate, ethylene and Xanthomonas. Mol. Plant Pathol. 10: 471-485.

Chao, Q., Rothenberg, M., Solano, R., Roman, G., Terzaghi, W. and Ecker, J. R. 1997. Activation of the ethylene gas response pathway in Arabidopsis by nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. Cell 89: 1133-1144.

Chaudhary, B., Hovav, R., Rapp, R., Verma, N., Udall, J. A., and Wendel, J. F. 2008. Global analysis of gene expression in cotton fibers from wild and domesticated *Gossypium barbadense*. Evol. Dev. 10: 567-582.

Chen, H., Shen, Y., Tang, X., Yu, L., Wang, J., Guo, L., Zhang, Y., Zhang, H., Feng, S., Strickland, E., Zheng, N. and Deng, X. W. 2006. Arabidopsis CULLIN4 forms an E3 ubiquitin ligase with RBX1 and the CDD complex in mediating light control of development. Plant Cell 18: 1991–2004.

Chen, Y. F., Etheridge, N. and Schaller, G. E. 2005. Ethylene signal transduction. Ann. Bot. 95: 901–915.

Cherkasova, E., Laassri, M., Chizhikov, V., Korotkova, E., Dragunsky, E., Agol, V. I. and Chumakov, K. 2003. Microarray analysis of evolution of RNA viruses: evidence of circulation of virulent highly divergent vaccine-derived polioviruses. Proc. Natl. Acad. Sci. USA 100: 9398-9403.

Choi, D., Cho, H.-T. and Lee, Y. 2006 Expansins: expanding importance in plant growth and development. Physiol. Plant 126: 511–518.

Christians, M. J., Gingerich, D. J., Hansen, M., Binder, B. M., Kieber, J. J., Vierstra, R. D. 2009. The BTB ubiquitin ligases ETO1, EOL1 and EOL2 act collectively to regulate ethylene biosynthesis in Arabidopsis by controlling type-2 ACC synthase levels. Plant J. 57: 332–345.

Clark, G., Torres, J., Finlayson, S., Guan, X., Handley, C., Lee, J., Kays, J. E., Chen, Z. J. and Roux, S. J. 2010. Apyrase (nucleoside triphosphate-diphosphohydrolase) and extracellular nucleotides regulate cotton fiber elongation in cultured ovules. Plant Physiol. 152: 1073-1083.

Debouck, C. and Goodfellow, P. N. 1999. DNA microarrays in drug discovery and development. Nat. Genet. 21: 48-50.

Delanghe, E.S. 1986. Lint development. In Cotton Physiology. Edited by Mauney, J. R. and Stewart, J. McD. Cotton Foundation, Memphis, TN, USA. 325–349.

Dharmasiri, S. and Estelle, M. 2002. The role of regulated protein degradation in auxin response. Plant Mol. Biol. 49: 401–409.

Diatchenko, L., Lau, Y. F., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E. D. and Siebert, P. D. 1996. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. USA 93: 6025-6030.

Duan, H., Li, F., Wu, X., Ma, D., Wang, M. and Hou, Y. 2006. Cloning and characterization of two EREBP transcription factors from cotton (*Gossypium hirsutum* L.). Biochemistry (Mosc.) 71: 285-293.

Feng, L., Vinicius, B. B., Cory, I. M., Hequet, E., Bordovsky, J. P., Keeling, W., Boman, R. and Bednarz, C. W. 2011. Effects of Irrigation, Cultivar, and Plant Density on Cotton Within-Boll Fiber Quality. Agron. J. 103: 297-303.

Freemont, P.S. 2000. RING for destruction? Curr. Biol. 10: 84–87.

Gagne, J. M., Downes, B. P., Shiu, S. H., Durski, A. M. and Vierstra, R. D. 2002. The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in *Arabidopsis.* Proc. Natl. Acad. Sci. USA 99: 11519–11524.

Gagne, J. M., Smalle, J., Gingerich, D. J., Walker, J. M., Yoo, S. D., Yanagisawa, S. and Vierstra, R. D. 2004. Arabidopsis EIN3-binding F-box 1 and 2 form ubiquitin-protein ligases that repress ethylene action and promote growth by directing EIN3 degradation. Proc. Natl. Acad. Sci. USA 101: 6803-6808.

Gerrish, R. S., Gill, A. L., Fowler, V. G. and Gill, S. R. 2010. Development of pooled suppression subtractive hybridization to analyze the pangenome of *Staphylococcus aureus*. J. Microbiol. Methods 81: 56-60.

Gou, J. Y., Wang, L. J., Chen, S. P., Hu, W. L. and Chen, X. Y. 2007. Gene expression and metabolite profiles of cotton fiber during cell elongation and secondary cell wall synthesis. Cell Res.17: 422-434.

Graves, D.A. and Stewart, J. M. 1988. Analysis of the protein constituency of developing cotton fibers. J. Exp. Bot. 39: 59-69.

Gray, W. M., del Pozo, J. C., Walker, L., Hobbie, L., Risseeuw, E., Banks, T., Crosby, W. L., Yang, M., Ma, H. and Estelle, M. 1999. Identification of an SCF ubiquitin–ligase complex required for auxin response in *Arabidopsis thaliana*. Genes Dev. 13: 1678–1691.

Guilfoyle, T.J. and Hagen, G. 2007. Auxin response factors. Curr. Opin. Plant Biol. 10: 453–460.

Guo, H. and Ecker, J. 2003. Plant responses to ethylene gas are mediated by SCFEBF1/EBF2-dependent proteolysis of EIN3 transcription factor. Cell 115: 667-677.

Gupta, A. K., Katoch, V. M., Chauhan, D. S., Sharma, R., Singh, M., Venkatesan, K. and Sharma, V. D. 2010. Microarray analysis of efflux pump genes in multidrug-resistant *Mycobacterium tuberculosis* during stress induced by common anti-tuberculous drugs. Microb. Drug Resist. 16: 21-28.

Han, H. Y., Lin, J. J., Zhao, Q. P., Dong, H., Jiang, L. L., Xu, M. Q., Zhu, S. H. and Huang, B. 2010. Identification of differentially expressed genes in early stages of *Eimeria tenella* by suppression subtractive hybridization and cDNA microarray. J. Parasitol. 96: 95-102.

Harada, T., Torii, Y., Morita, S., Masumura, T. and Satoh, S. 2010. Differential expression of genes identified by suppression subtractive hybridization in petals of opening carnation flowers. J. Exp. Bot. 61: 2345-2354.

Hedrick, S. M., Cohen, D. I., Nielsen, E. A. and Davis, M. M. 1984. Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature 308: 149-153.

Hershko, A. and Ciechanover, A. 1998. The ubiquitin system. Annu. Rev. Biochem. 67: 425–479.

Hovav, R., Udall, J. A., Chaudhary, B., Hovav, E., Flagel, L., Hu, G. and Wendal, J. F. 2008. The evolution of spinnable cotton fiber entailed prolonged development and a novel metabolism. PLoS Genet. 4: e25.

Hovav, R., Udall, J. A., Hovav, E., Rapp, R., Flagel, L. and Wendel, J. F. 2008. A majority of cotton genes are expressed in single-celled fiber. Planta 227: 319-329.

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

Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B. and Speed, T. P. 2003. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res. 31: e15.

Ji, S. J., Lu, Y. C., Feng, J. X., Wei, G., Li, J., Shi, Y. H., Fu, Q., Liu, D., Luo, J. C. and 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.

Jiang, M., Ryu, J., Kiraly, M., Duke, K., Reinke, V. and Kim, S. K. 2001. Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 98: 218-223.

Jin, L. G. and Liu, J. Y. 2008. Molecular cloning, expression profile and promoter analysis of a novel ethylene responsive transcription factor gene *GhERF4* from cotton (*Gossypium hirstum*). Plant Physiol. Biochem. 46: 46-53.

Jin, L. G., Li, H. and Liu, J. Y. 2010. Molecular characterization of three ethylene responsive element binding factor genes from cotton. J. Integr. Plant Biol. 52: 485-95.

Kendrick, M. D. and Chang, C. 2008. Ethylene signaling: New levels of complexity and regulation. Curr. Opin. Plant Biol. 11: 479–485.

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

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

Kimura, S., Laosinchai, W., Itoh, T., Cui, X., Linder, C. R. and Brown, R. M., Jr. 1999. Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant *Vigna angularis*. Plant Cell 11: 2075-2086.

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. and Chen, Z. J. 2006. Developmental and gene expression analyses of a cotton naked seed mutant. Planta 223: 418–432.

Lee, J. J., Woodward, A. W. and Chen, Z. J. 2007. Gene expression changes and early events in cotton fibre development. Ann. Bot. 100: 1391-1401.

Lee, J. H., Deng, X. W. and Kim, W. T. 2006. Possible role of light in the maintenance of EIN3/EIL1 stability in Arabidopsis seedlings. Biochem. Biophys. Res. Comm. 350: 484-491.

Lee, S. W., Tomasetto, C. and Sager, R. 1991. Positive selection of candidate tumor-suppressor genes by subtractive hybridization. Proc. Natl. Acad. Sci. USA 88: 2825-2829.

Leida, C., Terol, J., Marti, G., Agusti, M., Llacer, G., Badenes, M. L. and Rios, G. 2010. Identification of genes associated with bud dormancy release in *Prunus persica* by suppression subtractive hybridization. Tree Physiol. 30: 655-66.

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

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

Li, X. R., Wang, L. and Ruan, Y. L. 2010. Developmental and molecular physiological evidence for the role of phosphoenolpyruvate carboxylase in rapid cotton fiber elongation. J. Exp. Bot. 61: 287-295.

Li, Y., Liu, D., Tu, L., Zhang, X., Wang, L., Zhu, L., Tan, J. and Deng, F. 2010. Suppression of *GhAGP4* gene expression repressed the initiation and elongation of cotton fiber. Plant Cell Rep. 29: 193-202.

Liang, X., Oono, Y., Shen, N. F., Kohler, C., Li, K., Scolnik, P. A. and Theologis, A. 1995. Characterization of two members (*ACS1* and *ACS3*) of the 1-aminocyclopropane-1-carboxylate synthase gene family of *Arabidopsis thaliana*. Gene 167: 17–24.

Lisitsyn, N., Lisitsyn, N. and Wigler, M. 1993. Cloning the differences between two complex genomes. Science 259: 946-951.

Liu, D., Tu, L., Wang, L., Li, Y., Zhu, L. and Zhang, X. 2008. Characterization and expression of plasma and tonoplast membrane aquaporins in elongating cotton fibers. Plant Cell Rep. 27: 1385-1394.

Livak, K. J. and Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. Methods 25: 402-428.

Loguerico, L. L., Zhang, J. Q. and 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, M., Tan, K., Xiao, Z., Hu, M., Liao, P. and Chen, K. 2008. Cloning and expression of two sterol C-24 methyltransferase genes from upland cotton (*Gossypium hirsuturm* L.). J. Genet. Genomics 35:357-363.

Luo, M., Xiao, Y., Li, X., Lu, X., Deng, W., Li, D., Hou, L., Hu, M., Li, Y. and Pei, Y. 2007. GhDET2, a steroid 5alpha-reductase, plays an important role in cotton fiber cell initiation and elongation. Plant J. 51: 419-430.

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

Ma, L., Sun, N., Liu, X., Jiao, Y., Zhao, H. and Xing, W. D. 2005. Organ-specific expression of arabidopsis genome during development. Plant Physiol. 138: 80-91.

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

Machado, A., Wu, Y., Yang, Y., Llewellyn, D. J. and Dennis E. S. 2009. The MYB transcription factor GhMYB25 regulates early fibre and trichome development. Plant J. 59: 52-62.

Mei, W., Qin, Y., Song, W., Li, J. and Zhu, Y. 2009. Cotton *GhPOX1* encoding plant class III peroxidase may be responsible for the high level of reactive oxygen species production that is related to cotton fiber elongation. J. Genet. Genomics 36: 141-150.

Michiels, S., Koscielny, S. and Hill, C. 2005. Prediction of cancer outcome with microarrays: a multiple random validation strategy. Lancet 365: 488-492.

Mishra, A., Khare, S., Trivedi, P. K. and Nath, P. 2008. Ethylene induced cotton leaf abscission is associated with higher expression of cellulase (*GhCel1*) and increased activities of ethylene biosynthesis enzymes in abscission zone. Plant Physiol. Biochem. 46: 54-63.

Moon, J., Parry, G. and Estelle, M. 2004. The ubiquitin-proteasome pathway and plant development. Plant Cell 16: 3181–3195.

Mukhopadhyay, D. and Riezman, H. 2007. Proteasome-independent functions of ubiquitin in endocytosis and signaling. Science 315: 201–205.

Orford, S. J. and Timmis, J. N. 1998. Specific expression of an expansin gene during elongation of cotton fibres. Biochem. Biophys. Acta 1398: 342-346.

Pang, Y., Song, W., Chen, F. and Qin, Y. 2010. A new cotton SDR family gene encodes a polypeptide possessing aldehyde reductase and 3-ketoacyl-CoA reductase activities. Biochem. (Moscow) 75: 320-326.

Pang, Y., Wang, H., Song, W. Q. and Zhu, Y. X. 2010. The cotton ATP synthase δ1 subunit is required to maintain a higher ATP/ADP ratio that facilitates rapid fiber cell elongation. Plant Biol. (Germany) 12: 903-909.

Pear, J. R., Kawagoe, Y., Schreckengost, W. E., Delmer, D. P. and 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.

Peiser, G. D., Wang, T. T., Hoffman, N. E., Yang, S. F., Liu, H. W. and Walsh, C. T. 1984. Formation of cyanide from carbon 1 of 1-aminocyclopropane-1-carboxylic acid during its conversion to ethylene. Proc. Natl. Acad. Sci. USA 81: 3059-3063.

Pickart, C.M. 2001. Mechanisms underlying ubiquitination. Annu. Rev. Biochem.70: 503–533.

Potuschak, T., Lechner, E., Parmentier, Y., Yanagisawa, S., Grava, S., Koncz, C., Genschik, P. 2003. EIN3-dependent regulation of plant ethylene hormone signaling by two Arabidopsis F-Box proteins: EBF1 and EBF2. Cell 115: 679-689.

Preuss, T. M., Caceres, M., Oldham, M. C. and Geschwind, D. H. 2004. Human brain evolution: insights from microarrays. Nat. Rev. Genet. 5: 850-860.

Qiao, H., Chang, K. N., Yazaki, J. and Ecker, J. R. 2009. Interplay between ethylene, ETP1/ETP2 F-box proteins, and degradation of EIN2 triggers ethylene responses in Arabidopsis. Genes Dev. 23: 512–521.

Qiao, Z. X., Huang, B. and Liu, J. Y. 2008. Molecular cloning and functional analysis of an ERF gene from cotton (*Gossypium hirsutum*). Biochem. Biophys. Acta 1779: 122-127.

Qin, Y. M., Hu, C. Y., Pang, Y., Kastaniotis, A. J., Hiltunen, J. K. and Zhu, Y. X. 2007. Saturated very-long-chain fatty acids promote cotton fiber and arabidopsis cell elongation by activating ethylene biosynthesis. Plant Cell 19: 3692-3704

Qin, Y. M., Hu, C. Y. and Zhu, Y. X. 2008. The ascorbate peroxidase regulated by H_2O_2 and ethylene is involved in cotton fiber cell elongation by modulating ROS homeostasis. Plant Signal Behav. 3: 194-196.

Qin, Y. M., Pujol, F. M., Shi, Y. H., Feng, J. X., Liu, Y. M., Kastaniotis, A. J., Hiltunen, J. K. and 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.

Ramsey, J. C. and Berlin, J. D. 1976. Ultrastructure of early stages of cotton fiber differentiation. Bot. Gaz. 137: 11-19.

Ramsey, J. C. and Berlin, J. D. 1976. Ultrastructural aspects of early stages of cotton fiber elongation. Amer. J. Bot. 63: 868-876.

Ravanel, S., Gakière, B., Job, D. and Douce R. 1998. The specific features of methionine biosynthesis and metabolism in plants. Proc. Natl. Acad. Sci. USA 95: 7805-7812.

Rebrikov, D. V., Britanova, O. V., Gurskaya, N. G., Lukyanov, K. A., Tarabykin, V. S. and Lukyanov, S. A. 2000. Mirror orientation selection (MOS): a method for eliminating false positive clones from libraries generated by suppression subtractive hybridization. Nucleic Acids Res. 28: e90.

Reed, J. W. 2001. Roles and activities of Aux/IAA proteins in Arabidopsis. Trends Plant Sci. 6: 420–425.

Risseeuw, E. P., Daskalchuk, T. E., Banks, T. W., Liu, E., Cotelesage, J., Hellmann, H., Estelle, M., Somers, D. E. and Crosby, W. L. 2003. Protein interaction analysis of SCF ubiquitin E3 ligase subunits from Arabidopsis. Plant J. 34: 753–767.

Roberts A. W. and Roberts. E. 2007. Evolution of the cellulose synthase (CesA) gene family: insights from green algae and seedless plants. cellulose: Molecular and Structural Biology. 17-34.

Rojas-Mayorquin, A. E., Torres-Ruiz, N. M., Gudino-Cabrera, G. and Ortuno-Sahagun, D. Subtractive hybridization identifies genes differentially expressed by olfactory ensheathing cells and neural stem cells. Int. J. Dev. Neurosci. 28: 75-82.

Ruan, Y. L. and Chourey, P. S. 1998. A fibreless seed mutation in cotton is associated with lack of fiber cell initiation in ovule epidermis and alterations in sucrose synthase expression and carbon partitioning in developing seeds. Plant Physiol. 118**:** 399-406.

Ruan, Y. L., Llewellyn, D. J. and Furbank, R. T. 2001. The control of single-celled cotton fiber elongation by developmentally reversible gating of plasmodesmata and coordinated expression of sucrose and K^+ transporters and expansin**.** Plant Cell 13: 47-60.

Ruan, Y. L., Llewellyn, D. J. and Furbank, R. T. 2003. Suppression of sucrose synthase gene expression represses cotton fiber cell initiation, elongation, and seed development**.** Plant Cell 15: 952–964.

Ruan, Y. L. and Chourey, P. S. 1998. A fiberless seed mutation in cotton is associated with lack of fiber cell initiation in ovule epidermis and alterations in sucrose synthase expression and carbon partitioning in developing seeds. Plant Physiol. 118: 399-406.

Ruegger, M., Dewey, E., Gray, W. M., Hobbie, L., Turner, J. and Estelle, M. 1998. The TIR1 protein of Arabidopsis functions in auxin response and is related to human SKP2 and yeast grr1p. Genes Dev. 12: 198–207.

Saha, S., Wu, J., Jenkins, J. N., McCarty, J. C., Hayes, R. and Stelly, D. M. 2010. Genetic dissection of chromosome substitution lines of cotton to discover novel *Gossypium barbadense* L. alleles for improvement of agronomic traits. Theor. Appl. Genet. 120: 1193–1205.

Saha, S., Jenkins, J. N., Wu, J., McCarty, J. C., Gutierrez, O. A., Percy, R. G., Cantrell, R. G. and Stelly, D. M. 2006. Effects of chromosome-specific introgression in upland cotton on fiber and agronomic traits. Genetics 172: 1927–1938.

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. and Jeffrey, C. 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.

Saha, S., Wu, J., Jenkins, J. N., McCarty, J. C., Hayes, R. and Stelly, D. M. 2011. Delineation of interspecific epistasis on fiber quality traits in *Gossypium hirsutum* by ADAA analysis of intermated *G. barbadense* chromosome substitution lines. Theor. Appl. Genet. 122: 1351-136.

Santner, A. and Estelle, M. 2010. The ubiquitin-proteasome system regulates plant hormone signaling. Plant J. 61: 1029-40.

Seagull, R. W. 1992. A quantitative electron microscopic study of changes in microtubule arrays and wall microfibril orientation during in-vitro cotton fiber development. J. Cell Sci. 101: 561–577.

Seagull, R. W. 1993. Cytoskeletal involvement in cotton fiber growth and development. Micron**.** 24: 643-660.

Shi, H., Wang, X., Li, D., Tang, W., Wang, H., Xu, W. and Li X. 2007. Molecular characterization of cotton 14-3-3L gene preferentially expressed during fiber elongation. J. Genet. Genomics 34: 151-9.

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. and 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.

Shimizu, Y., Aotsuka, S., Hasegawa, O., Kawada, T., Sakuno, T., Sakai, F. and Hayashi, T. 1997. Changes in levels of mRNAs for cell wall-related enzymes in growing cotton fiber cells. Plant Cell Physiol. 38: 375–378.

Shirley, M. W. 1995. Eimeria species and strains of chickens. In Biotechnology-Guidelines on techniques in coccidiosis research, J. Eckert, R. Braun, M. W. Shirley, and P. Coudert (eds.). The European Commission DGXII, Luxembourg City, Luxembourg, 24 p.

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

Stelly, D.M., Saha, S., Raska, D.A., Jenkins, J.N., McCarty, J.C. and Gutierres, O.A. 2005. Registration of 17 Upland (*Gossypium hirsutum*) cotton germplasm lines disomic for different *G. barbadense* chromosome or arm substitutions. Crop Sci. 45: 2663-2665.

Stewart, J. M. 1975. Fiber initiation on the cotton ovule (*Gossypium hirsutum*). Am. J. Bot. 62: 723-730.

Sun, Y. and Allen, R. D. 2005. Functional analysis of the BIN 2 genes of cotton. Mol. Genet. Genomics 274: 51-59.

Suo, J., Liang, X., Pu, L., Zhang, Y. and 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.). Biochem. Biophys. Acta 1630: 25-34.

Taliercio, E. W. and Boykin, D. 2007. Analysis of gene expression in cotton fiber initials. BMC Plant Biol. 7: 22.

Tan, X. and Zheng, N. 2009. Hormone signaling through protein destruction: A lesson from plants. Am. J. Physiol. Endocrinol. Metab. 296: E223–E227.

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

Vahal, J., Schlagnhaufer, C. D. and Pell, E. J. 1998. Induction of an ACC synthase cDNA by ozone in light-grown *Arabidopsis thaliana* leaves. Physiol. Plant. 103: 45–50.

Vogel, J. P., Woeste, K. W., Theologis, A. and Kieber, J. J. 1998. Recessive and dominant mutations in the ethylene biosynthetic gene *ACS5* of Arabidopsis confer cytokinin insensitivity and ethylene overproduction, respectively. Proc. Natl. Acad. Sci. USA 95: 4766–4771.

Voges, D., Zwickl, P. and Baumeister, W. 1999. The 26S proteasome: a molecular machine designed for controlled proteolysis. Annu. Rev. Biochem. 68: 1015–1068.

Wang Q. Q., Liu, F., Chen, X. S., Ma X. J., Zeng, H. Q. and Yang, Z. M. 2010. Transcriptome profiling of early developing cotton fiber by deep-sequencing reveals significantly differential expression of genes in a fuzzless/lintless mutant. Genomics 96: 369-376.

Wang, S., Wang, J. W., Yu, N., Li, C. H., Luo, B., Gou, J. Y., Wang, L. J. and Chen, X. Y. 2004. Control of plant trichome development by a cotton fiber MYB gene. Plant Cell 16: 2323-34.

Wang, X., Zhang, Y., Zhang, J., Cheng, C. and Guo, X. 2007. Molecular characterization of a transient expression gene encoding for 1-aminocyclopropane-1-carboxylate synthase in cotton (*Gossypium hirsutum* L.). J. Biochem. Mol. Biol. 40: 791-800.

Welch, E. M., Barton, E. R., Zhuo, J., Tomizawa, Y., Friesen, W. J., Trifillis, P., Paushkin, S., Patel, M., Trotta, C. R., Hwang, S., Wilde, R. G., Karp, G., Takasugi, J., Chen, G., Jones, S., Ren, H., Moon, Y. C., Corson, D., Turpoff, A. A., Campbell, J. A., Conn, M. M., Khan, A., Almstead, N. G., Hedrick, J., Mollin, A., Risher, N., Weetall, M., Yeh, S., Branstrom, A. A., Colacino, J. M., Babiak, J., Ju, W. D., Hirawat, S., Northcutt,, V. J., Miller, L. L., Spatrick, P., He, F., Kawana, M., Feng, H., Jacobson, A., Peltz, S. W. and Sweeney, H. L. 2007. PTC124 targets genetic disorders caused by nonsense mutations. Nature 447: 87-91.

Westafer, J. M. and Brown, R. M., Jr**.** 1976. Electron microscopy of the cotton fiber: new observations on cell wall formation. Cytobios 15: 111-138.

Whittaker, D. J., Triplett, B. A. 1999. Gene specific changes in alpha-tubulin transcript accumulation in developing cotton fibers. Plant Physiol. 121: 181-188.

Wu, Y., Machado, A. C., White, R. G., Llewellyn, D. J. and Dennis, E. S. 2006. Expression profiling identifies genes expressed early during lint fiber 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. and 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.

Wu, Z., Soliman, K. M., Bolton, J. J., Saha, S. and Jenkins, J. N. 2008. Identification of differentially expressed genes associated with cotton fiber development in a chromosomal substitution line (CS-B22sh). Funct. Integr. Genomics 8: 165-174.

Xiao, Y. H., Li, D. M., Yin, M. H., Li, X. B., Zhang, M., Wang, Y. J., Dong, J., Zhao, J., Luo, M., Luo, X. Y., Hou, L., Hu, L. and Pei, Y. 2010. Gibberellin 20-oxidase promotes initiation and elongation of cotton fibers by regulating gibberellin synthesis. J. Plant Physiol. 167: 829-837.

Xiao, Z., Tan, K., Hu, M., Liao, P., Chen, K. and Luo, M. 2008. Cloning and expression analysis of *GhDET3*, a vacuolar H^+-ATPase subunit C gene, from cotton. J. Genet. Genomics 35: 307-312.

Xu, D., Wei, J., Cui, H., Gong, J., Yan, Y., Lai, R. and Qin, Q. 2010. Differential profiles of gene expression in grouper *Epinephelus coioides*, infected with Singapore grouper iridovirus, revealed by suppression subtractive hybridization and DNA microarray. J. Fish Biol. 77: 341-360.

Xu, W. L., Wang, X. L., Wang, H. and Li, X. B. 2007. Molecular characterization and expression analysis of nine cotton *GhEF1A* genes encoding translation elongation factor 1A. Gene 389: 27-35.

Yajun, H., Wangzhen, G., Xinlian, S. and Tianzhen, Z. 2008. Molecular cloning and characterization of a cytosolic glutamine synthetase gene, a fiber strength-associated gene in cotton. Planta 228: 473-483.

Yamagami, T., Tsuchisaka, A., Yamada, K., Haddon, W. F., Harden, L. A. and Theologis, A. 2003. Biochemical diversity among the 1-Amino-cyclopropane-1-carboxylate synthase isozymes encoded by the Arabidopsis gene family. J. Biol. Chem. 278: 49102–49112

Yamaura, M., Uchiumi, T., Higashi, S., Abe, M. and Kucho, K. 2010. Identification by suppression subtractive hybridization of Frankia genes induced under nitrogen-fixing conditions. Appl. Environ. Microbiol. 76: 1692-1694.

Yanagisawa, S., Yoo, S. D. and Sheen, J. 2003. Differential regulation of EIN3 stability by glucose and ethylene signalling in plants. Nature 425: 521-525.

Yang, S. F. and Hoffman, N. E. 1984. Ethylene biosynthesis and its regulation in higher plants. Ann. Rev. Plant physiol. 35: 155-189.

Yee, D. and Goring, D. R. 2009. The diversity of plant U-box E3 ubiquitin ligases: from upstream activators to downstream target substrates. J. Exp. Bot. 60: 1109–1121.

Yoo, S., Cho, Y., Tena, G., Xiong, Y. and Sheen, J. 2008. Dual control of nuclear EIN3 by bifurcate MAPK cascades in C2H4 signaling. Nature 451: 789-795.

Yoshida, H., Wang, K. L., Chang, C. M., Mori, K., Uchida, E. and Ecker, J. R. 2006. The ACC synthase TOE sequence is required for interaction with ETO1 family proteins and destabilization of target proteins. Plant Mol. Biol. 62: 427–437.

Yu, X., Zhu, Y., Lu, S., Zhang, T., Chen, X., and Xu, Z. 2000. A comparative analysis of a fuzzless-lintless mutant of *Gossypium hirsutum* L. cv. Xu-142. Sci. China C. Life Sci. 43: 623-30.

Zaki, H. E. M., Yokoi, S. and Takahata, Y. 2010. Identification of genes related to root shape in radish (*Raphanus sativus*) using suppression subtractive hybridization. Breed. Sci. 60: 130-138.

Zhang, M. and Liang, C. 2000. Progress in the studies of lipid-transfer protein. Progr. Biochem. Biophys. 27: 246-247.

Zeng, F., Baldwin, D. A. and Schultz, R. M. 2004. Transcript profiling during preimplantation mouse development. Dev. Biol. 272: 483-496.

Zhao, P. M., Wang, L. L., Han, L. B., Wang, J., Yao, Y., Wang, H. Y., Du, X. M., Luo, Y. M. and Xia, G. X. 2010. Proteomic identification of differentially expressed proteins in the *Ligon lintless* mutant of upland cotton (*Gossypium hirsutum* L.). J. Proteome Res. 9:1076-1087.

Zheng, N., Schulman, B. A., Song, L., Miller, J. J., Jeffrey, P. D., Wang, P., Chu, C., Koepp, D. M., Elledge, S. J., Pagano, M., Conaway, R. C., Conaway, J. W., Harper, J. W. and Pavletich, N. P. 2002. Structure of the Cul1–Rbx1–Skp1–F box–Skp2 SCF ubiquitin ligase complex. Nature 416: 703–709.