Gene Expression Studies Reveal the Role of Major Genes Regulate the High Quality of Egyptian Cotton Fibers

**Abstract:** The aim of the present study was the development of a gene candidate platform through bioinformatics work for filtering candidate genes are highly correlated to cotton fiber-quality. Thereafter, gene expression analyses were performed to validate the function of genes participate in fiber development. The expression analyses of each gene were performed at elongation stage (5, 10, 15 and 20 DPA; days post-anthesis) in different genotypes; the Egyptian cotton cv. Giza 70 (*Gossypium barbadense*), the Upland cotton cv. Tamcot (*G. hirsutum*) in addition to the developed Hybrids; Hybrid A (Giza 70, female x Tamcot) and Hybrid B (Tamcot, female x Giza 70). The results of bioinformatics analyses were used to determine the candidate genes which are highly correlated to cotton fiber quality. Twenty four candidate genes belonging to 21 families were determined in this study. Thus, semi-quantitative RT-PCR and Real-time quantitative RT-PCR (qRT-PCR) analyses were performed for each gene. The results of gene expression analyses revealed that these candidate genes were highly expressed in fiber cells of Giza 70 followed by a relatively high expression in the two hybrids A&B compared to the lowest level of expression which was detected in Tamcot. These results confirmed the functional role of these genes in the development of fiber cells during the elongation stage. Thus, these genes are participating in superior fiber quality of Egyptian cotton and represent potential targets for further improvement of Egyptian cotton through an integrated metabolic engineering strategy to develop high-yield Egyptian cotton with more proper industrial added-value.

**Key words:** Cotton · *Gossypium barbadense* · Gene expression · Fiber-quality · qRT-PCR

**INTRODUCTION**

Cotton plants particularly of *Gossypium hirsutum* and *G. barbadense*, are the main cultivated species which produce a fiber yield used in the textile industry which about 26 million metric tons from about 34 million hectares in 80 countries including USA, India, China and Egypt [1-3]. Also, cotton seeds contain oil and important secondary compounds, in addition to seed proteins used for the feeding of livestock [4].

The fibers of cotton are trichomes (each is a single-cell) that emerge from the epidermal tissue of the ovules in the bolls [5]. The linear growth of cotton fiber cell resulted from intercalary and tip growth [6]. The
elongation of cotton fiber cell to reach about 30-35 mm is dependent on dramatic changes at the level of gene expression followed by major metabolic changes to sustain the growth of fiber cell [5, 7-11].

Increasing the fiber yield of cotton plants was for a long time the major aim for cotton breeders, but recently due to technological needs of textile industry the quality improvement of cotton fibers became an essential aim in combination with the yield [12]. The genetic potential exists to enhance the strength and durability of cotton fiber to successfully compete with petroleum-based synthetic fibers such as polyester. The challenge, however, is to identify desirable genes for superior fiber quality and to introgress such traits for improved fiber quality into elite breeding lines. For the past two decades, quantitative trait loci (QTL) have been used in conjunction with linkage mapping to identify loci that are primarily responsible for variation in the phenotype of complex, quantitative traits. The main goal of QTL-based studies is to ascertain the location of QTL, the number and interaction of QTL and to ultimately determine how each QTL influences the phenotype. For example, breeding programs attempt to identify loci that govern crop yield and quality and then to introgress favorable alleles into elite breeding lines. A major limitation of QTL data is that different parental combinations and/or experiments conducted in different environments often result in the identification of partly or wholly non-overlapping sets of QTL [13].

Cotton fiber improvement strategies in breeding programs around the world face similar problems. A narrow germplasm base and limited genetic diversity in breeding populations is a major obstacle that hinders the development of high density molecular genetic linkage with saturated QTL intervals that are a prerequisite to marker-assisted selection (MAS) [14]. Cotton research across the globe circles around the goal to combine Pima fiber quality traits for length, strength, fineness and maturity in a high-yielding Upland background. However, introgression of Egyptian/Pima alleles via classical breeding from interspecific crosses has not been successful. Previous efforts have focused on gene discovery and developing DNA-based markers and a molecular genetic map to facilitate marker-assisted selection (MAS) strategies for the genetic enhancement of cotton [15-18]. Functional genomic studies to define the fiber transcriptome currently estimate that ~30,000 genes, or ~50% of the cotton genome, are required to make a fiber [11]. With the availability of microarray-based gene expression platforms, the expression QTL (eQTL) is considered as a key to opening the ‘closed box’ that exists between genotype and phenotype [19, 20].

Bioinformatics studies were used in this study as an alternative and more cost-effective approach to developing eQTL for cotton fiber, thus we proposed to selectively target expression-profiling of genetically-mapped QTL using the available published QTL-based microarrays studies to build a platform of candidate genes of cotton fiber quality [6, 13, 15, 21-32]. Afterward, expression analyses of these candidate genes were carried out in two dimensions genotype and developmental stage to determine and filtering the functional genes strongly involving in fiber development. Confirming the function of genes participate in fiber development is essential to exploit these genes for improving cotton fiber quality using biotechnology [33].

Therefore the aim of the present study was the development of a gene candidate platform through bioinformatics work for filtering candidate genes are highly correlated to cotton fiber quality. Thereafter, gene expression analyses were performed to validate the function of genes participate in fiber development. The expression analyses of each gene were performed at elongation stage (5, 10, 15 and 20 DPA; days post-anthesis) in different genotypes; the Egyptian cotton cv. Giza 70 (Gossypium barbadense), the Upland cotton cv. Tamcot (G. hirsutum) in addition to the developed Hybrids; Hybrid A (Giza 70, female x Tamcot) and Hybrid B (Tamcot, female x Giza 70).

MATERIALS AND METHODS

As a prerequisite step to achieve this objective we cultivated cotton plants of different genotypes during 2010, 2011, 2012, 2013 and 2014 seasons, in each season; the Egyptian cotton cv. Giza 70 (Gossypium barbadense) and the Upland cotton cv. Tamcot (G. hirsutum) were cultivated in four plots; two for parents (self-fertilization) and two for hybridization by crossing in two ways (Giza70; female x Tamcot) and (Tamcot; female x Giza70). In addition to another two plots for the developed Hybrids; Hybrid A (Giza 70, female x Tamcot) and Hybrid B (Tamcot, female x Giza 70) starting from 2011 season until the last season 2014. At maturity, the seeds of Hybrids and the parents were collected and kept for the next season starting from 2010 season until the last season 2014).
Collection of Plant Materials: Cotton plants of different genotypes were grown simultaneously under identical field conditions in a randomized block design. At the flowering stage, flowers were tagged on the day of anthesis and samples of cotton bolls were collected at 5, 10, 15 and 20 DPA (days post anthesis), respectively. All samples were collected before 9:00 am and immediately placed on ice. The samples were collected on the same day within each developmental stage (18-20 bolls) were tagged to represent one biological replicate (three independent biological replicates were collected for each genotype at each developmental stage). In Total 48 samples were collected during the season for four genotypes; (12 samples/genotype) three biological replicates x four developmental stage. Cotton bolls were washed three times with ice cooled water including liquid detergent followed by three times with sterile ice-cooled water and then surface sterilized with ice-cooled 70% Ethanol for 30 seconds followed by three times with sterile ice-cooled water and then immersed in sterile ice-cooled water. Ovules were excised from the bolls under aseptic conditions and immediately transferred to a sterile 50 ml falcon tube containing RNA stabilization solution (RNALater®, QIAGEN, 76106) and kept submerged in RNALater® for two hours at 25°C and then fibers were scraped from the ovules and homogenated in a porcelain mortar with total RNA extraction buffer.

Total RNA Extraction: RNA extractions were performed using the CTAB method with modifications, in brief; the dissected fibers (100 mg) were homogenated in a porcelain mortar with one ml of total RNA extraction buffer containing [CTAB 2%, polyvinylpyrrolidone (PVP40) 2%, 100 mM Tris–HCl (pH 8.0), 25 mM EDTA and 2.0 M NaCl] and then β-mercaptoethanol to a final concentration of 10% (v/v) just before use (900 µl of extraction buffer and 100 µl of β-mercaptoethanol) were added [28]. The fibers homogenate was transferred into a 2 ml centrifuge tubes and incubated at 65°C for 15 min; and the tubes were shaken by inverting the tubes 3-4 times during incubation. Then, RNA was extracted with 800 µl of chloroform: isoamyl alcohol: (24:1) 3–5 times, each time the mixture was shaken for 30 sec. then the mixture was centrifuged at 10,000 xg for 10 min at 4°C and the supernatant was transferred to a new tube. The next step was performed to purify the RNA from the proteins using 800 µl of phenol/chloroform (1:1), the tubes were shaken for 30 sec and were centrifuged at 10,000 xg for 10 min at 4°C. The supernatant was collected in a new tube containing an equal volume of (24:1) chloroform/isoamyl alcohol. The samples were shaken and then centrifuged at 10,000 xg for 10 min at 4°C. The supernatant was transferred to a new tube containing 1/3 volume of LiCl (8 M) and kept at -20°C for at least 4 hrs. Then, the samples were centrifuged at 10,000 xg for 20 min at 4°C. The pellet was washed with 500 µl 70% ethanol, air dried for 5 min at room temperature and then dissolved in 100 µl RNase-free water. The total RNA extract was treated with RNase-free DNase I (NEB, M0303S), 5 units (2.5 µl) for 10 minutes at 37°C in DNase I reaction buffer, then RNA was extracted following the standard phenol/chloroform extraction as above mentioned. Finally, total RNA was precipitated by adding 0.1 volume of sodium acetate (3 M) and 2 volumes of ethanol 100% and incubated at -20°C for 2 hrs. The RNA was recovered by centrifugation (13,000 xg for 15 min at 4°C), washed with 70% ethanol, air dried and then dissolved in 100 µl RNase-free water. The quality and quantity of total RNA extracts were determined spectrophotometrically and the contamination due to phenol/poly saccharides and proteins was determined by recording the OD ratios; A_{260/230} and A_{260/280}, respectively, the ratios (> 1.9) indicated that the extracts were free of polyphenols, polysaccharides and proteins, thus considered with good-quality. Thereafter, to verify RNA integrity, the extracts were fractioned by electrophoresis in a 1% agarose gel, stained with ethidium bromide and visualized under UV light. Total RNA was isolated independently three times from each biological pool to create technical replicates; in total 36 RNA extracts/genotype (9 independent extracts of total RNA were carried out for each genotype at each stage x 4 developmental stages).

Purification of mRNA from total RNA Extracts: For enhancing standard cDNA preparation and subsequently the expression profiling analyses of candidate genes, poly(A)⁺ RNA was purified from quality-controlled total RNA using NEBNext® Poly(A)⁺ mRNA Magnetic Isolation Module (NEB, E7490L), then the concentration of the purified mRNA of each sample was determined spectrophotometrically.

Semi-quantitative RT-PCR analyses: Semi-quantitative RT-PCR analyses were performed to validate the new gene-specific primers (Table 1) designed for the selected candidate genes and the ubiquitin 7 gene (GhUBQ7) as an internal control. The gene-specific primers were designed using PRIMER3 (http:// frodo.wi.mit.edu/
Table 1: Primers of candidate genes for fiber initiation and elongation in cotton

<table>
<thead>
<tr>
<th>No</th>
<th>Families of candidate Genes</th>
<th>Gene name</th>
<th>GenBank accession number</th>
<th>Primer sequence</th>
<th>Size of Amplified fragment (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cellulose synthase</td>
<td>CESAl</td>
<td>U58283.1</td>
<td>Fwd 5-AAG GCA GGT AGG GAG GCA AT-3</td>
<td>186</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CESAl</td>
<td>AF431201.0</td>
<td>Rev 5-TCC CAC ATA AAC TGG CCC TT-3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CESAl</td>
<td>JQ345699.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>UDP-glucose 6-dehydrogenase family protein</td>
<td>UGD1</td>
<td>FJ415166.1</td>
<td>Fwd 5-TAC CGG TGA CAC TAG GGA GA-3</td>
<td>217</td>
</tr>
<tr>
<td>3</td>
<td>UDP-glucose pyrophosphorylase</td>
<td>UGP</td>
<td>FJ415165.1</td>
<td>Fwd 5-ATT TTC ATC ACT AC-3</td>
<td>197</td>
</tr>
<tr>
<td>4</td>
<td>UDP-D-glucuronic acid 4-epimerase</td>
<td>GAE1</td>
<td>FJ415208.1</td>
<td>Fwd 5-GAT GAG CTG TTT CCC TGG AC-3</td>
<td>240</td>
</tr>
<tr>
<td>5</td>
<td>Arabinogalactan</td>
<td>AGP1</td>
<td>AY218846.1</td>
<td>Fwd 5-GCA CTC ACC TTA ACT AC-3</td>
<td>175</td>
</tr>
<tr>
<td>6</td>
<td>Xyloglucan endotrans glucosylase/ hydrolase</td>
<td>XTH1</td>
<td>HM749062.1</td>
<td>Fwd 5-CCT TGG ATT GCC TGC TG-3</td>
<td>191</td>
</tr>
<tr>
<td>7</td>
<td>Sucrose synthase</td>
<td>SUS1</td>
<td>JN376127.1</td>
<td>Fwd 5-TCA ATT GCC TGA ACC ACC TG-3</td>
<td>151</td>
</tr>
<tr>
<td>8</td>
<td>Vacuolar invertase</td>
<td>VIN1</td>
<td>FJ915120.1</td>
<td>Fwd 5-CAG ACC ATT GCC TGA ACC ACC TG-3</td>
<td>210</td>
</tr>
<tr>
<td>9</td>
<td>Myo-Inositol-1-phosphate synthase</td>
<td>MIPS1</td>
<td>FJ415168.1</td>
<td>Fwd 5-GGA GGA GCA ACC ACC TG-3</td>
<td>219</td>
</tr>
<tr>
<td>10</td>
<td>Expansin</td>
<td>EXP</td>
<td>FJ415208.1</td>
<td>Fwd 5-GTA ACC ACC ACC TG-3</td>
<td>198</td>
</tr>
<tr>
<td>11</td>
<td>Actin</td>
<td>ACT3</td>
<td>AY305725.1</td>
<td>Fwd 5-GTA ACC ACC ACC TG-3</td>
<td>198</td>
</tr>
<tr>
<td>12</td>
<td>Tubulin-beta</td>
<td>TUB-B9</td>
<td>AF521248.1</td>
<td>Fwd 5-GAG ATG TCC CGG GGT TCC TG-3</td>
<td>178</td>
</tr>
<tr>
<td>13</td>
<td>Acetyl CoA carboxylase</td>
<td>ACC</td>
<td>EF464667.1</td>
<td>Fwd 5-CAA ACT GCC TGG GGT GGA AGC AG-3</td>
<td>211</td>
</tr>
<tr>
<td>14</td>
<td>Acyl-CoA synthetase</td>
<td>ACS1</td>
<td>DQ174259.1</td>
<td>Fwd 5-TCC TCC TCC TCC ACC ACC ACC TG-3</td>
<td>207</td>
</tr>
<tr>
<td>15</td>
<td>3-ketoacyl-CoA synthase</td>
<td>KCS4</td>
<td>KT625614.1</td>
<td>Fwd 5-GCT TCC TCC TCC TCC ACC ACC ACC TG-3</td>
<td>153</td>
</tr>
<tr>
<td>16</td>
<td>Auxin efflux carrier family protein</td>
<td>SAUR1</td>
<td>KM014658.1</td>
<td>Fwd 5-CAG ACC ACC ACC ACC ACC ACC TG-3</td>
<td>150</td>
</tr>
<tr>
<td>17</td>
<td>Auxin-responsive factor protein</td>
<td>ARBF1</td>
<td>KJ627772.1</td>
<td>Fwd 5-TCC TCC TCC TCC ACC ACC ACC ACC TG-3</td>
<td>178</td>
</tr>
<tr>
<td>18</td>
<td>Protodermal Factor 1 gene</td>
<td>PDF1</td>
<td>JF421225.1</td>
<td>Fwd 5-AAG ACT TGG AGG GGA AGC AGC AG-3</td>
<td>169</td>
</tr>
<tr>
<td>19</td>
<td>WRKY TF protein</td>
<td>WRKY1</td>
<td>KF669776.1</td>
<td>Fwd 5-TCC TGG TGG GGA AGC AGC AG-3</td>
<td>161</td>
</tr>
<tr>
<td>20</td>
<td>MYB TF protein</td>
<td>MYB4</td>
<td>AF034132.1</td>
<td>Fwd 5-TGG TGG TGG GGA AGC AGC AG-3</td>
<td>205</td>
</tr>
<tr>
<td>21</td>
<td>Zinc finger TF protein</td>
<td>ZFP2</td>
<td>DT455798</td>
<td>Fwd 5-TCC TCC TCC ACC ACC ACC ACC ACC TG-3</td>
<td>155</td>
</tr>
<tr>
<td>22</td>
<td>Ubiquitin</td>
<td>UBQ7</td>
<td>DQ116441</td>
<td>Fwd 5-TCC TCC TCC ACC ACC ACC ACC ACC TG-3</td>
<td>192</td>
</tr>
</tbody>
</table>
primer3/input.htm). The samples of purified mRNA were reverse transcribed (RT) into cDNA using an oligo (dT)20 primer and the ProtoScript® First Strand cDNA Synthesis Kit (NEB™ E6500S) in accordance with the manufacturer’s instructions. Reverse transcription-polymerase chain reaction (RT-PCR) for the expression analyses was conducted in 20 µl reactions containing 5 µl of the template (diluted cDNA 1:5 H2O), 1 pmol of each primer of the target gene or the primers of UBQ7 as an internal control, 200 µM of each dNTP, 2.5 mM MgCl2, 2.5 units of Hot Start Taq DNA Polymerase (NEB™ M0495S). To validate the results of these semi-quantitative RT-PCR analyses, the number of RT-PCR cycles was reduced to only 20 cycles depending on the results of RT-PCR optimization in independent experiments including serial concentrations of the internal control gene UBQ7 and different cycle’s number. Therefore, the RT-PCR amplification was performed by the initial denaturation step at 94°C for 5 min. followed by 20 cycles; at 94°C (30 s), 60°C (30 s) and 72°C (30 s) and finally holding 1 min at 72°C for extension, employing the Applied Biosystems® Veriti™ 96-Well Thermal Cycler with 6 x 16-well VeriFlex blocks and the Applied Biosystems® ProFlex™ 3 x 32-well PCR System. Samples were run in triplicate to validate the results of semi-quantitative RT-PCR. The amplified products of the three RT-PCR replicates of each sample were mixed together and 10 µl was separated on a 1.2% agarose gel and visualized after staining with ethidium bromide.

Real-time Quantitative RT-PCR analyses: Real-time quantitative RT-PCR assays were performed using the same gene-specific primers which were successfully examined by semi-quantitative RT-PCR analyses. For each gene the samples were run in triplicate on each plate to validate the results of quantitative RT-PCR; 3 samples/genotype (representing 3 biological replicates) at each developmental stage were analyzed and repeated three times in three independent experiments (on three plates) as three technical replicates, therefore, the obtained results represent means (± SE) of 9 replicates; (n = 9). The real-time quantitative RT-PCR analyses were performed using an ABI 7500 Real-Time PCR system (Applied Biosystems®). The reactions were performed in a final volume of 20 µl containing 5 µl of (diluted cDNA 1: 5 H2O), 10 µL of 2x SYBR Green Master Mix (Applied Biosystems®) and 1 pmol of forward and reverse primers of the target gene or the primers of the UBQ7 gene as an internal control. The thermal cycling conditions were; initial denaturation step at 94°C for 5 min. followed by 40 cycles; at 95°C (30 s), 60°C (30 s) and 72°C (30 s). The relative gene expression levels were presented as 2 ΔΔCT [34].

Statistical analysis: Data were statistically analyzed on the basis of a completely randomized design, with a two-way factorial arrangement [35]. Data were analyzed by two-way ANOVA for main effects and interaction between (G; genotypes and DS; developmental stages). All values are means and standard deviation (± SE) of 9 replicates; three biological replicates of plant materials and three technical replicates in each biological sample (n = 9). Significant differences were calculated using the Least-Significant-Difference (L.S.D.) test at (*p < 0.05 level) to demonstrate the significant differences and at (**P<0.01 level) for highly significant differences in gene expression between genotypes during elongation stage at different samples date (5, 10, 15 and 20 DAP).

RESULTS AND DISCUSSION

Bioinformatics Work for Developing the Fiber QTL-specific Gene Candidate Platform: In this study, we developed a fiber QTL-specific gene candidate platform to study gene expression of fiber QTL. This gene candidate platform was developed depending on the continuous Bioinformatics work which was carried out since the beginning of this study (2010) to date as an ongoing process till the end of this work to enrich the developed gene candidate platform through regular cycles of Bioinformatics analyses in our laboratory using different resources; the available recent literature and public databases, especially TropGene (tropgened.cirad.fr/en/cotton.html), CMAP (cottondb.org/cdbpages/cmap.html), (http://www.cottongen.org/node/49448) and (http://www.ncbi.nlm.nih.gov) [6, 13, 15, 21-32]. Bioinformatics analyses were performed using Vector NTI Advance v. 10 program and (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Thus, the gene candidate platform was included enzymes of cell wall biosynthesis, carbohydrate metabolism, phytohormones, fatty acids, secondary metabolism and transcription factors related genes associated with fiber elongation as shown in (Table 1). Design of gene-specific primers of cotton fiber gene candidate platform was carried out as shown in (Table 1) including the gene of cotton ubiquitin extension protein GhUBQ7; GenBank accession number DQ116441 as an internal control to normalize the amount of template cDNA added in each reaction for standard normalization of expression levels and to evaluate the reproducibility of the new gene
candidate platform by real-time quantitative RT-PCR analyses [36, 37]. The present analyses of gene expression profiling of fiber QTL were carried out using the developed cotton fiber QTL gene candidate platform in two dimensions of study (genotypes and fiber development). The expression profiling of these genes were carried out to determine the specific changes in gene expression patterns between the Egyptian cotton Giza 70 and the Upland Tamcot in comparison to the developed Hybrids (Giza 70, female x Tamcot) and (Tamcot, female x Giza 70) at different samples date (5, 10, 15 and 20 DPA) during elongation stage.

**Semi-quantitative RT-PCR Analyses of Expressed Genes in Developing Fibers:** Results of semi-quantitative RT-PCR analyses of expressed genes in developing cotton fibers of different genotypes; Tamcot (G. hirsutum), Giza 70 (G. barbadense), Hybrid A (Giza 70 x Tamcot) and Hybrid B (Tamcot x Giza 70) at different samples date (5, 10, 15 and 20 DPA) revealed that these genes were differentially expressed and exhibited a distinct pattern for expression of each gene in different genotypes during the fiber development as shown in (Fig. 1).

The expression of cellulose synthase A9 gene (CESA9) showed that the level of expression was remarkably increased at mid and late elongation stage (10, 15 and 20 DPA) in all genotypes with pronounced elevated expression in cv. Giza 70 and the two Hybrids A&B in comparison to cv. Tamcot. These results confirmed the overexpression of this gene in the G. barbadense genotype and this can explain the obvious alteration in the expression level in the Hybrid genotypes due to the introgressed chromosome segments of G. barbadense genotype into the genome of these Hybrids. Furthermore, the results of the level of expression of all genes in G. barbadense genotype and in the Hybrids A&B exhibited a similar pattern to the CESA9 gene with very little alterations in comparison to the G. hirsutum genotype.

The results of semi-quantitative RT-PCR analyses of the expressed genes in developing cotton fibers of the two Hybrids A&B indicated that the presence of the G. barbadense copy of these genes in the genome of these two Hybrids lead to increasing the expression of these genes in developing cotton fibers at different samples date (5, 10, 15 and 20 DPA) during elongation stage. Hence, these findings strongly emphasize that these genes are involving in the regulation of fiber initiation and elongation in G. barbadense and further analyses by real-time quantitative RT-PCR are of great importance to validate these results which may explain the detected higher expression of these genes in this genotype which possesses the particular high fiber-quality traits.

The results of semi-quantitative RT-PCR analyses of the expressed genes in developing cotton fibers of Giza 70 and the two Hybrids A&B genotypes compared to Tamcot proved that the high expression of these genes is correlated with the presence of G. barbadense chromosomes in these genotypes and these results are in accordance with several reports indicated that most of the fiber quality genes are expressed in genotypes with a high A-genome background such as (G. barbadense) more than in genotypes with a high D-genome background such as (G. hirsutum) [27, 38-41].

**Real-time Quantitative RT-PCR Analyses of Expressed Genes in Developing Fibers:** Real-time quantitative RT-PCR analyses of cellulose synthase genes; CESA1, CESA4 and CESA9 revealed that these genes were differentially expressed during the stages of fiber development and the expression level showed progress increase and reached the maximum level at late elongation stage (20 DPA) in all genotypes with particular elevated expression in cv. Giza 70 and followed by Hybrid A (Giza 70 x Tamcot) and with very little alterations in Hybrid B (Tamcot x Giza 70) in comparison to G. hirsutum genotype cv. Tamcot at different samples date (5, 10, 15 and 20 DPA) during elongation stage as shown in (Fig. 2). The expression of cellulose synthase genes; CESA1, CESA4 and CESA9 exhibited highly significant differences were recorded in Giza 70 (G. barbadense) and significant differences in Hybrid A (Giza 70 x Tamcot) in comparison to Hybrid B (Tamcot x Giza 70) and G. hirsutum genotype. These results are in accordance with the obtained results of semi-quantitative RT-PCR analyses of CESA9 gene and strongly confirmed the predominantly overexpression of these genes in developing fibers of Egyptian cotton at different samples date (5, 10, 15 and 20 DPA) during elongation stage. The results of the expression of these genes (CESA1, CESA4 and CESA9) proved that highly significant differences were recorded in Giza70 compared to the control (Tamcot) at different samples date (5, 10, 15 and 20 DPA) during elongation stage. Thus, these results remarkably indicated that these genes (CESA1, CESA4 and CESA9) functionally participate in the elevation of fiber quality in Giza 70. Therefore, these results probably contribute to enhancing the resolution of the genetic profile of cotton fiber quality in Egyptian cotton. These results are in concordance with those obtained by Fang.
Fig. 2: Real-time quantitative RT-PCR analyses of expression of cellulose synthase genes. Changes in the relative expression level of cellulose synthase genes; CESA1, CESA4 and CESA9 were detected with significant differences (*P<0.05) and with highly significant differences (**P<0.01) according to L.S.D. test in two cotton genotypes; Giza 70 (G. barbadense) and Hybrid A (Giza 70 x Tamcot) in comparison to Tamcot (Gossypium hirsutum) and Hybrid B (Tamcot x Giza 70) at different samples date (5, 10, 15 and 20 DPA) during elongation stage. Changes in the relative expression level of UGD1; UDP-glucose 6-dehydrogenase 1 gene, UGP; UDP-glucose pyrophosphorylase gene, UGP1; UDP-glucose pyrophosphorylase 1 gene were detected with significant differences (*P<0.05) and with highly significant differences (**P<0.01) according to L.S.D. test for GAE1 and XTH1 genes in cotton genotypes; Giza 70 and Hybrid A&B in comparison to Tamcot at late elongation stage (15 and 20 DPA). The UGD1 gene showed a highly significant difference in Tamcot and the two Hybrids at mid and late elongation stage (10, 15 and 20 DPA). The relative expression level of each gene was determined using the ubiquitin 7 gene (UBQ7) as an internal control. Results represent means (± SE) of 9 replicates; three biological replicates of plant materials and three technical replicates for each biological sample (n = 9).

et al. [27] who reported that different genes of cellulose synthase (CESA3, CESA4 and CESA7) exhibited a significant upregulation in the hybrid line (CSIL-35431) containing G. barbadense chromosomes during fiber elongation stage (from 5 DPA to 20 DPA).

Previous reports indicated that at least fifteen genes of cellulose synthase (CESA) participate in the biosynthesis of cellulose in G. raimondii [42]. In addition, Li et al. [43] reported that these genes (CESA3 - CESA5 - CESA6 - CESA9 - CESA10) are involving in the
formation of the primary cell wall during the elongation stage. Gene expression analyses of (UGP gene; UDP-glucose pyrophosphorylase, UGP1 gene; UDP-glucose pyrophosphorylase 1 and AGP1 gene; Arabinogalactan 1) genes exhibited that highly significant changes were occurred in developing cotton fibers at late elongation stage (15 and 20 DPA) in Giza 70 and the same trend was detected with less extent in Hybrid A and these changes reached the significant level in Hybrid B only at 20 DPA for UGP and UGP1 genes in comparison to Tamcot, as shown in (Fig. 2).

Results of expression analyses of (GAE1 gene; UDP-D-glucuronic acid 4-epimerase 1 and XTH1 gene; xyloglucan endotransglycosylase/hydrolase 1) genes showed highly significant differences in Giza 70 and Hybrid A&B at late elongation stage (15 and 20 DPA) in comparison to Tamcot as shown in (Fig. 3). On the contrary, the UGD1 gene; UDP-glucose 6-dehydrogenase 1 gene, showed a highly significant difference in Tamcot at late elongation stage (20 DPA) and significant differences in Hybrid B at mid and late elongation stage (10, 15 and 20 DPA) in addition to Hybrid A at 10 DPA. The alteration in relative expression level of SUS1 gene; sucrose synthase 1, VIN1 gene; vacuolar invertase 1 and MIPS gene; myo-inositol-1-phosphate synthase, genes were detected at a highly significant level in developing cotton fibers at late elongation stage (15 and 20 DPA) in Giza 70 and Hybrid A and the same trend was detected with less extent in Hybrid B in comparison to Tamcot, as shown in (Fig. 3). In addition SUS1 gene showed a pronounced increase in expression levels in Giza 70, which recorded a significant difference at initial stage (5 DPA) and highly significant differences at mid and late developing stages of cotton fibers (10, 15 and 20 DPA) in comparison to Tamcot and even to the Hybrids which only recorded a highly significant difference in Hybrid A and a significant difference in Hybrid B at 20 DPA.

Different genes involved in carbohydrate metabolism such as; UGP, UGP1, GAE1, XTH1, UGD1, SUS1 and VIN1 are of great importance for cellulose biosynthesis in addition to MIPS gene for plasma membrane of fiber cell. It was observed that the expression of these genes was markedly upregulated in Giza 70 and at less extent in the Hybrids A&B compared to Tamcot during the elongation stage (5-20 DAP). Thus collectively these results suggest vital roles of these candidate genes in elongation of cotton fiber cell. These results are in accordance with those obtained by Fang et al [27] who reported that, many genes are involved in the biosynthesis of the primary cell wall of cotton fiber including xyloglucan endotransglycosylases (XET), tubulins, expansins and arabinogalactans, thus these genes were upregulated during fiber elongation in the chromosome-introgressed-lines (CSILs) containing segments of G. barbadense chromosomes. Several reports indicated that xyloglucan endotransglycosylases has a role in cotton fiber development [43-45]. The genes XET22, XET28 and other XETs were markedly upregulated during fiber elongation in the CSILs [27].

Confirming the function of genes participate in fiber development is essential to exploit these genes for improving cotton fiber quality using biotechnology [33]. For instance, sucrose is hydrolyzed in fiber cells into hexoses (glucose and fructose) by two enzymes sucrose synthase (EC 2.4.1.13) and invertase (EC 3.2.1.26) for cell wall biosynthesis, metabolism and signaling [46-48]. Thus, the role of SUS gene in fiber cell elongation was reported by [49, 50]. In addition, the role of invertase in the growth of cotton fiber cell was reported by Wang et al. [46]. The silencing of SUS gene in cotton resulted in fiberless seeds [51]. Furthermore, the introduction of the sucrose synthase gene of potato in G. hirsutum resulted in increasing fiber length, seed number and consequently the yield of cotton was increased [52].

Results of gene expression analyses of (ACT3 gene; Actin 3 and TUB-B9 gene; tubulin-beta 9) genes showed that these genes were upregulated and recorded highly significant differences in Giza 70 in comparison to Tamcot particularly at late elongation stage (15 and 20 DPA). In addition, the expansin gene (EXP) was highly upregulated in developing cotton fibers of Giza 70 at all samples and only with less extent at 20 DPA, as shown in (Fig. 4). The alteration in expression levels of (ACT3, TUB-B9 and EXP) genes exhibited a similar trend for Hybrid A compared with Giza 70, in particular for the EXP gene and with less extent for ACT3 and TUB-B9 genes. The polymers of actin proteins participate in the cytoskeletons and direct the polar growth of eukaryotic cells [53, 54]. The linear growth of cotton fiber cell resulted from intercalary and tip growth [6].

The elongation of cotton fiber cell to reach about 30-35 mm is dependent on dramatic changes at the level of gene expression followed by major metabolic changes to sustain the growth of fiber cell [5, 7-11]. Fifteen genes of actin were characterized in G. hirsutum [55]. The suppression of ACT1 gene by RNAi in G. hirsutum resulted in restriction of fiber cell elongation [55]. In addition, different actin-binding proteins are involved in elongation of cotton fiber cells [56-61].
Fig. 3: Real-time quantitative RT-PCR analyses of expressed genes in developing cotton fibers. Changes in the relative expression level of \( GAE1 \); UDP-D-glucuronic acid 4-epimerase 1 gene, \( AGP1 \); arabinoxylan 1 gene and \( XTH1 \); xyloglucan endotransglucosylase/hydrolase 1 gene were detected with significant differences (*) \( P < 0.05 \) and with highly significant differences (**) \( P < 0.01 \) according to L.S.D. test for these genes in Giza 70 and Hybrid A&B in comparison to Tamcot at different samples date (5, 10, 15 and 20 DPA) during elongation stage. Changes in the relative expression level of \( SUS1 \); sucrose synthase 1 gene, \( VIN1 \); vacuolar invertase 1 gene and \( MIPS \); myo-Inositol-1-phosphate synthase gene were detected with significant differences (*) \( P < 0.05 \) and with highly significant differences (**) \( P < 0.01 \) according to L.S.D. test for these genes in cotton genotypes; Giza 70 and Hybrid A&B in comparison to Tamcot at late elongation stage (15 and 20 DPA). The \( SUS1 \) gene showed highly significant differences in Giza 70 at mid and late elongation stage (10, 15 and 20 DPA). The relative expression level of each gene was determined using the ubiquitin 7 gene (UBQ7) as an internal control. Results represent means (± SE) of 9 replicates; three biological replicates of plant materials and three technical replicates for each biological sample \((n = 9)\).

Thyssen et al. [32] reported that the Ligon lintless Li1 mutant with short fibers resulted from the mutation in the actin gene by single amino acid Gly65Val substitution. Expansin proteins play an important role in fiber length and quality [62]. The Expansins and tubulins proteins are essential for cell wall expansion and elongation of cotton fiber cell [7, 62-64]. α-Expansins
Fig. 4: Real-time quantitative RT-PCR analyses of expressed genes in developing cotton fibers. Changes in the relative expression level of \textit{ACT3}; Actin 3 gene, \textit{EXP}; expansin gene and \textit{TUB-B9}; tubulin- beta 9 gene were detected with significant differences (**\(P<0.05\)) and with highly significant differences (**\(P<0.01\)) according to L.S.D. test for these genes in Giza 70 and Hybrid A and with little alteration in Hybrid B in comparison to Tamcot particularly at late elongation stage (15 and 20 DPA). The \textit{EXP} gene showed highly significant differences in Giza 70 and Hybrid A and with less extent in Hybrid B at initial, mid and late elongation stage (5, 10, 15 and 20 DPA). The relative expression level of \textit{ACCC}; acetyl CoA carboxylase gene, \textit{ACS1}; acyl-CoA synthetase 1 gene and \textit{KCS4}; 3-ketoacyl-CoA synthase 4 gene were detected with significant differences (**\(P<0.05\)) and with highly significant differences (**\(P<0.01\)) according to L.S.D. test for these genes in Giza 70 and Hybrid A&B and with little alteration in Tamcot particularly for \textit{ACS1} and \textit{KCS4} genes at late elongation stage (20 DPA). The \textit{ACCC} gene showed highly significant differences in Giza 70 and Hybrid A&B at late elongation stage (15 and 20 DPA). The relative expression level of each gene was determined using the ubiquitin 7 gene (\textit{UBQ7}) as an internal control. Results represent means (± SE) of 9 replicates; three biological replicates of plant materials and three technical replicates for each biological sample (\(n = 9\)).
Fig. 5: Real-time quantitative RT-PCR analyses of expressed genes in developing cotton fibers. Changes in the relative expression level of ARF5; Auxin-responsive factor 5 gene, SAUR1; auxin-responsive small suxin 1 gene and PDF1; Protodermal Factor 1 gene were detected with significant differences (*P<0.05) and with highly significant differences (**P<0.01) according to L.S.D. test for these genes in Giza 70 and Hybrid A and with little alteration in Hybrid B in comparison to Tamcot particularly at late elongation stage (15 and 20 DPA). The PDF1 gene showed highly significant differences at all samples in Giza 70, for Hybrid A and for B with less extent at mid and late elongation stage (15 and 20 DPA) and significant difference at 20 DPA for Hybrid B. Changes in the relative expression level of MYB4; MYB TF 4 gene, WRKY7; WRKY TF 7 gene and ZFP2; Zinc finger TF 2 gene were detected with significant differences (*P<0.05) and with highly significant differences (**P<0.01) according to L.S.D. test for these genes in Giza 70 and Hybrid A and with little alteration in Hybrid B in comparison to Tamcot particularly at late elongation stage (15 and 20 DPA). The MYB4 gene showed highly significant differences in Giza 70 at mid and late elongation stage (10, 15 and 20 DPA) and at 20 DPA for Hybrid A. The relative expression level of each gene was determined using the ubiquitin 7 gene (UBQ7) as an internal control. Results represent means (+ SE) of 9 replicates; three biological replicates of plant materials and three technical replicates for each biological sample (n = 9).
(EXPAs) form various complexes with polysaccharide (xyloglucan and pectin) and bind to the microfibrils of cellulose during cell wall biosynthesis [65].

The results of gene expression analyses of EXP and TUB-B9 genes are in accordance with those obtained by Fang et al. [27] who reported that the EXP A1 genes were markedly upregulated in the CSILs and the TUB-B6 gene was upregulated only in CSIL-35431 at the early stage of fiber elongation.

The alteration in expression levels of (ACS1 gene; acyl-CoA synthetase 1 and KCS4 gene; 3-ketoacyl-CoA synthase 4) genes in developing cotton fibers of Giza 70 exhibited significant differences at all samples as shown in (Fig. 4), in addition, the ACCC gene; acetyl CoA carboxylase, exhibited significant differences in Giza 70 and also the Hybrids A&B at late elongation stage (15 and 20 DPA).

It is worthy to mention that, the expression levels of ACS1 and KCS4 genes were also upregulated in Tamcot particularly at late elongation stage (15 and 20 DPA). The fatty acids are essential for plasma membrane biosynthesis and involved in elongation of fiber cells. Several reports pointed to the upregulated genes of fatty acids metabolism such as: ACC synthase, long-chain acyl-CoA synthetase (LACS) and 3-ketoacyl-CoA synthase (KCS) in cotton fiber cells [66-68].

Gene expression analyses of (ARF5 gene; auxin-responsive factor 5, SAUR1 gene; auxin-responsive small suxin 1 and PDF1 gene; protodermal factor 1) genes revealed that these genes were highly upregulated in developing cotton fibers of Giza 70 at late stage (15 and 20 DPA) in comparison to Tamcot as shown in (Fig. 5). On the contrary, Fang et al. [27] reported that auxin-response genes such as auxin-response factors and SAUR-like auxin-responsive genes were obviously upregulated in particular at the early stage of fiber elongation (5-10 DAP).

In addition, the PDF1 gene exhibited a pronounced highly significant upregulation in Giza 70 at all samples as shown in (Fig. 5) and also the Hybrid A at the late stage (15 and 20 DPA) and with less extent for Hybrid B in comparison to Tamcot. The protodermal factor 1 gene (GhPDF1) was upregulated during fiber initiation and elongation particularly at the early stage [9]. The silencing of GhPDF1 gene resulted in retardation of fiber growth and the produced fibers were shorter than the wild-type. In addition, the expression of some genes related to sugar transport, pectin synthesis and ethylene were markedly reduced in fiber cells of transgenic cotton plants [69].

Results of gene expression analyses of the transcription factor genes (MYB4, WRKY7 and ZFP2) in developing cotton fibers of Giza 70 exhibited a highly significant upregulation in comparison to Tamcot, in addition, Hybrid A only at 20 DPA and with less extent for Hybrid B, as shown in (Fig. 5). These results confirmed the overexpression of these regulatory genes (MYB4, WRKY7 and ZFP2) in the Egyptian cotton cv. Giza 70; G. barbadense genotype, as well as in the Hybrids A&B due to the introgressed chromosome segments of Giza 70 into the genome of these Hybrids in comparison to the Upland cotton cv. Tamcot; G. hirsutum genotype. The role of different families of transcription factors such as MYB ERF, WRKY and NAC were studied during cotton fiber development [27]. The expression of different numbers of genes belonging to these TF families (16 genes of ERF, 36 genes of MYB, 43 genes of WRKY and 29 genes of NAC) was upregulated in the hybrid lines (CSIL-35431 and CSIL-31134) at 10 DPA [27]. Many reports indicated that a number of genes of MYB family are associated with elongation and development of cotton fiber cell [31, 70-75]. Ding et al. [29] reported that several genes of WRKY family were upregulated during elongation and development of fiber cell in different cotton species G. raimondii and G. arboretum.

Finally, the alterations in gene expression levels of 24 fiber-quality related genes belonging to 21 families were detected using semiquantitative RT-PCR analyses and were confirmed using the precise real-time quantitative RT-PCR analyses. Thus, the results of gene expression profiling of cotton fiber-quality related genes revealed that these genes were differentially expressed in developing cotton fibers of all genotypes at different samples date (5, 10, 15 and 20 DPA) during elongation stage. The pronounced highly significant increases in the level of expression of these genes have occurred in cv. Giza 70, whereas moderate or low expression has occurred in Hybrid A and Hybrid B, respectively, in comparison to the Upland cotton cv. Tamcot. Therefore, according to the emerged results using a bioinformatics approach to determine candidate genes related to cotton fiber-quality QTL and the validation of these candidate genes by real-time quantitative RT-PCR analyses, a group of genotype-specific and developmental stage-specific co-expressed genes was identified and proved to be essential for exhibiting the superior fiber quality of Egyptian cotton.
CONCLUSIONS

The planned real-time quantitative RT-PCR analyses of gene expression profiling of cotton fiber-quality related genes in this study were carried out using the developed cotton fiber-quality gene candidate platform in two dimensions of study (genotypes and fiber development). The alterations in gene expression levels were detected using semiquantitative RT-PCR analyses and were confirmed using the precise real-time quantitative RT-PCR analyses. The expression profiling of cotton fiber-quality genes were carried out to determine the specific changes in gene expression patterns between the Egyptian cotton Giza 70 and the Upland Tamcot in comparison to the developed two Hybrids (Giza 70, female x Tamcot) and (Tamcot, female x Giza 70) at elongation stage. The results in this study of gene expression profiling of cotton fiber-quality related genes revealed that these genes were differentially expressed in developing cotton fibers of all genotypes at different samples date (5, 10, 15 and 20 DPA) during elongation stage. The pronounced highly significant increases in the level of expression of these genes have occurred in cv. Giza 70, whereas moderate or low expression has occurred in Hybrid A and Hybrid B, respectively, in comparison to the Upland cotton cv. Tamcot. Therefore, according to the obtained results using a bioinformatics approach to determine candidate genes related to cotton fiber-quality QTL a group of 24 genes was determined. Thereafter, the validation of these candidate genes was performed using real-time quantitative RT-PCR analyses. Thus, a group of genotype-specific and developmental stage-specific co-expressed genes was identified and proved to sustain the superior fiber quality of Egyptian cotton. These results confirmed the functional role of these genes in the development of fiber cells during the elongation stage. Thus, these genes are participating in superior fiber quality of Egyptian cotton and represent potential targets for further improvement of Egyptian cotton through an integrated metabolic engineering strategy to develop high-yield Egyptian cotton with more proper industrial added value.

ACKNOWLEDGMENT

The authors wish to acknowledge the Science and Technology Development Fund (STDF), Ministry of Higher Education and Scientific Research, that this work was supported through a grant for a research project (ID: 1085). Heartfelt thanks for team work members at the Scientific Center of Excellence in Plant Biotechnology (CU-PBRL).

REFERENCES


