INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
17–0776/2018/20–2–288–296
DOI: 10.17957/IJAB/15.0492
http://www.fspublishers.org

Full Length Article

Comparative Analysis of Fiber Morphogenesis Genes of *Calotropis procera* and *Gossypium hirsutum*

Asia Khatoon[1], Nadia Iqbal[1], Muhammad Asif[1], Hafiza Masooma Naseer Cheema[2*], Muhammad Saeed[1] and Aftab Bashir[3*]

[1]*National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan*
[2]*Plant Genetic Resources Lab, Department of Plant Breeding and Genetics, University of Agricultural, Faisalabad*
[3]*Department of Biological Sciences, F.C. College University, Lahore, Pakistan*
[*]For correspondence: masooma@uaf.edu.pk; aftabb.pk@gmail.com

Abstract

Cotton (*Gossypium* spp.) is the leading fiber crop in the world. The earnings from fiber lint are associated with its quality traits. For studying the major fiber development specific genes, cDNA libraries were used to establish 1000 EST's (expressed sequence tags), each from the fast-growing fibers of *Calotropis procera* and *Gossypium hirsutum* L. Most abundantly expressed transcripts of *C. procera* coded for expansins, aquaporins and sugar/salt transporters. Four homologs of *expansin* gene (*CpEXPA*1, *CpEXPA*2, *CpEXPA*3 and *CpEXPA*4) and four of *aquaporins* (*CpTiP1, CpTiP2, CpPiP1* and *CpPiP2)* were found to be the most abundantly expressed transcripts in fast growing *C. procera* fiber cells. Cotton fiber EST's revealed two homologs of expansin, one for *G. hirsutum plasma membrane intrinsic protein* (*GhPiP1*) and one transcript for *sucrose synthase*. The *GhPiP1* had 70% amino acid identity to *CpPiP1*, while no homolog was detected for *CpPiP2* in cotton fibers. Real-time PCR data showed that CpPiP1 had about 40% higher transcript level of the gene in *C. procera* as compared to cotton. *CpPiP2* showed high level of transcripts in *C. procera* fiber only. On the other hand, *CpEXPA*3 was the closest homologue of *GhEXP15* (76% amino acid identity) and its transcript level was 80% higher in *C. procera* fibers. Expression of *sucrose synthase* was at its high level in 5–15 DPA cotton fibers. These studies demonstrated that *CpPiP2*, *CpEXPA*3 and *sucrose synthase* are the best candidate genes for fiber modification in *G. hirsutum*. © 2018 Friends Science Publishers

Keywords: Transcript profiling; Cotton fiber; Fiber genes; EST's; Transcript analysis

Introduction

The economy of Pakistan relies on cotton production that is the supreme source of raw material for textile industry and lint for fabrics. The earnings from fiber lint are associated with its quality traits. Cotton fibers are seed trichomes that originate from the ovular epidermis and can elongate from 16 to 50 mm depending upon the cotton plant species and its geographical location (Kim and Triplett, 2001). During development, the cotton fiber passes through overlapping stages of development including initiation, elongation, secondary cell wall thickening and maturation (Arpat *et al*., 2004), which is controlled by different sets of genes at each developmental stage.

Calotropis procera is a wild and perennial shrub with excellent potential to grow under drought and salt effected lands. Its seed fibers are single celled hollow trichomes, which can grow more than 45 mm in length and have good strength, uniformity and fineness (Table 1). *C. procera* fibers represent a model system to study fiber development specific genes as they do not have the secondary cell wall deposition phase. Therefore, most of the coded genes represent those which are involved in fiber length, strength and fineness. Although, Pakistan is one of the top cotton producing countries, yet there is a dire need to increase the productivity and improve fiber quality. The transgenic approaches for crop improvement mostly focus on the utilization of genes from diverse resources. *C. procera* fiber development genes are hypothesized to be the best resources for cotton fiber improvement.

Multiple sets of genes are expressed during fiber cell development phases (Wilkins and Arpat, 2005; Earl and Deborah, 2007). There are number of genes that have role in enhancement of yield as well as fiber quality. Molecular genetics technologies are direct approaches to modify specific traits by transforming these genes. *Expansin*, *aquaporins* and *sucrose synthase* (*Sus*) are the important gene families that play a very important role in fiber strength and length (Gou *et al*., 2007; Aslam *et al*., 2013). The Sus has been reported to play a central role in strengthening the plant cell wall. This enzyme also has a key role in fiber initiation. The suppression of this protein has resulted in fiber less seeds in cotton (Ruan *et al*., 2003).

To cite this paper: Khatoon, A., N. Iqbal, M. Asif, H.M.N. Cheema, M. Saeed, A. Bashir, 2018. Comparative analysis of fiber morphogenesis genes of *Calotropis procera* and *Gossypium hirsutum*. *Int. J. Agric. Biol*., 20: 288–296

Sucrose is the primary translocatable carbohydrate in the majority of plants as its metabolism plays a vital role to the regulation of photoassimilates. The developing cotton fibers require UDP-glucose as a precursor of cellulose synthesis, which is obtained in the cell from sucrose through Sus (Ruan *et al*., 2001). Acotton line (FiberMax®) with elevated Sus has been reported to produce stronger and longer fibers (Arioli, 2005). A stronger expression of *Sus* gene in the developing fibers may help to increase the fiber length and strength in elite cotton cultivars.

Expansins constitute one of the important gene families involved in plant cell expansion and other cell wall modification processes like seed germination, control of organ size, morphology, fruit softening and pollen tube invasion in the grass stigma (Sabirzhanova *et al*., 2005). *Expansin* gene family comprises of four sub-families like α-*expansin* (*EXPA*), β-*expansin* (*EXPB*), *expansin-like A* (*EXLA*) and *expansin-like B* (*EXLB*) (Kende *et al*., 2004). The *EXPA* and *EXPB* sub-families are found in all groups of land plants (Schipper *et al*., 2002; Li *et al*., 2003) and these genes are involved in cell wall loosening. On the other hand, function of EXLA and EXLB proteins have not been fully understood (Cosgrove *et al*., 1997). Another mechanism reported for plant cell elongation is the turgor pressure driven cell elongation by water channels constituted by aquaporins. The aquaporins belong to the major intrinsic protein family (MIPs), which comprise of a super family of integral membrane proteins in mammals, plants, insects, yeast, bacteria, protozoa and archaea (Kammerloher *et al*., 1994; Deen and Van, 1998; Carbrey *et al*., 2001; Kozono *et al*., 2003). The aquaporins are water channel proteins with an average size of 28–30 kDa that form channels/pores in the biological membranes and specifically regulates osmotic pressure based movement of water molecules and other small solutes in and out of living cells (Sakurai *et al*., 2005). They play a vital role in transporting a large volume of water by consuming small amount of energy (Tyerman *et al*., 2002).

The present study was planned to screen expressed sequence tags (EST) from fast growing fiber cDNA library of *C. procera and G. hirsutum.* Most abundantly expressed transcripts were selected from ESTs and their sequences were analyzed using bioinformatics tools. Expression levels of selected genes were determined in cotton (*G. hirsutum*) and *C. procera* tissues. This study would help to select the genes that may be over expressed in cotton to improve fiber quality traits.

Materials and Methods

Analysis of Fiber EST's

EST's were developed from the fast-growing fibers of *C. procera* and *G. hirsutum* L. to study the major fiber development specific genes. Fiber cDNA libraries of *G. hirsutum* and *C. procera* were available in Plant Genome Resource Lab, NIBGE and were used to establish EST's (Cheema *et al*., 2010). One thousand clones were picked from each library and high quality sequences were obtained from the commercial sequencing services provided by Eurofins, Luxembourg. The DNA sequences were initially analyzed through BLAST at NCBI and the ontology groups were obtained using KOBAS (http://kobas.cbi.pku.edu.cn/).

Analysis of Fiber Specific Genes

A total of eleven genes involved in fiber development were picked from EST's on the basis of BLAST and ontology group searches. The variant of *expansin* obtained from cotton fiber cDNA library was named as *GhEXPA15* and the four variants of *expansin* from *C. procera* fiber cDNA library were named as *CpEXPA1, CpEXPA2, CpEXPA3* and *CpEXPA4* based on their similarity with the known groups of expansins. A clone of *expansin* variant (*GhEXPA8*) from *G. hirsutum* cv. CIM707 was already available in lab (accession no. EU593897). Other genes explored from EST's included *aquaporins*, which were categorized as tonoplast and plasma membrane intrinsic proteins based on their similarities to the known *aquaporins*. One *aquaporin* gene (*GhPiP1*) was isolated from cotton EST clones, while four *aquaporins* (*CpPiP1, CpPiP2, CpTiP1* and *CpTiP2*) were isolated from *C. procera* fiber cDNA library. Only one clone of *Sus* was identified from the cotton fiber EST clones. The nucleotide and the coded amino acid sequences of all the *expansin* and *aquaporin* variants were submitted in GeneBank at NCBI (http://www.ncbi.nlm.nih.gov) and accession numbers were obtained.

Comparison of *G. hirsutum* and *C. procera* Proteins

Amino acid sequences for each of the selected genes were derived using nucleotide translate tool available at www.expasy.ch. Both the nucleic acid and amino acid sequences of *G. hirsutum* and *C. procera* fiber specific genes were compared using multiple sequence alignment. The amino acid sequences showed higher similarity due to codon degeneracy, therefore, the amino acid comparisons are included in this study. The multiple sequence alignment tool of the CLC, Denmark Main Workbench Ver 6.7.1 was used to create amino acid sequence alignments between the *G. hirsutum* and *C. procera* protein families. The alignments were then fed into the pairwise comparison tool and percent identities between the sequences were obtained. The alignment data of expansin and aquaporin proteins was also used to construct UPGMA (unweighted pair group method with arithmetic mean) based phylogenetic tree for each of the protein family using the phylogenetic tree construction tool available in the CLC Main Workbench.

The signal peptide and protease cleavage site probability in the amino acid sequences of all the variants of *expansin genes* in *G. hirsutum* and *C. procera* were determined using SignalP 3.0 at http://www.cbs.dtu.dk/services/SignalP/.

Table 1: Fiber quality traits of *G. hirsutum* cv. CIM 707 compared to *C. procera*

Trait	Scales				
Satple length (mm)	Short < 20	Medium 20-26	Long 27-32	Extra Long > 32	
CIM-707			32		
C. procera				>35	
Fiber Strength (tppsi)	Weak < 71	Fair 72-80	Average 81-88	Strong 89-97	Very Strong > 97
CIM-707				97	
C. procera				95	
Fiber fineness (µg/inch)	Very Fine < 3	Fine 3-3.9	Average 4-4.9	Coarse 5-5.9	Very Coarse > 5.9
CIM-707			4.2		
C. procera	2.09				

Table 2: Primer sequences of *18S rRNA, CpEXPA3, CpPiP2* and *sucrose synthase* for real-time PCR studies

No.	Name of gene	Primer pair	Primer sequence (5'-3')	Primer length	Amplicon size (bp)
1	*18S rRNA*	RT18SF2	AAACGGCTACCACATCCAAG	20 mer	153
		RT18SR2	CCTCCAATGGATCCTCGTTA	20 mer	
	expansins	CLCExpA3 F2	GAAGAAGGGGGAATAAGG	20 mer	236
		CLCExpA3 R2	GTTGTAGCTGGTGATTGT	18 mer	
3	*aquaporins*	CLCPIP2 F1	TTTCGGTGGCATGATCTT	18 mer	333
		CLCPIP2 R1	TCTGGCATTTCTCTTGGG	18 mer	
4	*sucrose synthase*	CLCSuSyF2	GAGAAAGTATACGGAACAGA	20 mer	390
		CLCSuSyR2	GAAAGTACTGGTGATGATGA	20 mer	

The amino acid sequences of expansins and aquaporins were used to obtain ontology groups from *Arabidopsis thaliana* database using KOBAS ver. 3 available at http://kobas.cbi.pku.edu.cn. The aquaporins being the water transport proteins were analyzed through TMPred (http://www.ch.embnet.org/software/TMPRED_form.html).

Transcript Profiling of Expressed Genes

The transcript profiling of *expansins, and Sus* gene families was carried out in various cotton and *C. procera* tissues including fibers, while the transcript profiling of *aquaporin* was performed only in the fiber tissues of *G. hirsutum* and *C. procera* using a BIO RAD iQ5 cycler, USA.

Plant Material

Cotton seeds of CIM 707 were obtained from the Central Cotton Research Institute Multan, Pakistan. The seeds were delinted using 10% H_2SO_4, washed with tap water for 3–4 times and air dried for 48–72 h at room temperature. The seeds were planted in plastic pots containing Metro Mix® and were grown in the green house set at a day and night temperature between 30–35ºC. The day flower bloomed, was marked as 0 DPA (days post anthesis) and the developing bolls were collected at 10 DPA for total RNA isolation and subsequent cDNA synthesis. *C. procera* plants growing in the NIBGE campus area were selected. Flower opening was taken as 0 DPA and 10 DPA fruiting body was used to collect fibers for isolation of total RNA and cDNA synthesis. Other plant tissues, used to obtain cDNA included leaves, shoots, stem and roots. All tissues were washed with DEPC treated water, dried with blotting paper and frozen in liquid nitrogen on sight.

RNA Isolation and First Strand cDNA Synthesis

The fibers isolated at 10 DPA and other tissues were ground separately in liquid nitrogen to fine powder using sterile pestle and mortar. Total RNA was extracted by guanidine thiocyanate method using the "Plant RNA Purification Reagent" (Invitrogen, USA). Quality of extracted total RNA was observed by electrophoresis. The RNA was quantified by its band intensity on agarose gel and spectrophotometrically by Ultrospec 3100 (GE, USA). Total RNA from each tissue was used as a template to synthesize first strand cDNA using Revert Aid H^- cDNA synthesis kit (Thermo Fisher Scientific, USA). The cDNA was used as template in qPCR assay to amplify fiber related genes. The cDNA samples were equalized by PCR amplification of 18S rRNA gene from each of the cDNA preparations. The normalized template was subsequently used in all qRT-PCR reactions.

Primer Validation

The CLC Main Workbench was used to design the primers for qRT-PCR. Multiple sets of primers were designed on each gene and used for validation on the respective gene. All primer pairs were tested on 10 DPA fiber cDNA. Single primer pairs were selected for real-time PCR studies on the basis of sharp, specific and abundant amplification of the target gene (Table 2). Recommended laboratory practices and precautionary measures were followed to conduct real-time assay.

Real-time PCR Analysis

Real-time quantitative PCR (qRT-PCR) was conducted using 12.5 µL SYBR Green Super Mix (Bio-Rad, USA) and 25 ng/µL of each gene specific primer in a total reaction

volume of 25 µL. The PCR reactions were performed in triplicate on an iQ5 cycler (BioRad, USA) using 96 well reaction plates sealed with optically clear film. Initial denaturation temperature to preheat the samples was set at 94ºC for 4 min. The subsequent 40 cycles were repeated for denaturation at 94ºC for 30 sec, annealing at 50ºC for 30 sec and extension at 72ºC for 30 sec. The melt curves for each amplification product were monitored for any nonspecific amplification. Relative fold expression was calculated by ΔCt method ($E^{Ct C.procera-Ct G. hirsutum}$), which is a built in feature of the IQ5 cycler. The machine software statistically analyzes the data by taking mean values of the three replicates for each gene, while standard error and mean gene expression data was calculated using the least significance difference (LSD) at $p \leq 0.05\%$ of mean values.

Results

Analysis of Fiber Specific Genes

Nucleotide sequences of 2000 EST clones were BLAST searched at NCBI and also annotated using KOBAS. Genes involved in different developmental processes were categorized. The genes related to cell morphogenesis were picked from these categories and used for transcript profiling in different *G. hirsutum* and *C. procera* tissues including fibers. The analysis of the sequencing data indicated that four homologs of *expansin* gene families are expressed in the developing *C. procera* fibers. The BLAST search, indicated that all the four *expansins* belonged to *expansin A* gene family. Four homologs of *expansin A* gene *i.e. CpEXPA*1, *CpEXPA*2, *CpEXPA*3 and *CpEXPA*4 were isolated from the *C. procera*, while one variant of *expansin* (*GhExpA15*) was isolated from cotton cDNA library, in addition to the previously isolated *expansin* variant (*GhEXP8*). These genes were named according to the suggested nomenclature of *expansin* gene family (Kende *et al*., 2004). GenBank accession no. of the four *C. procera expansin* homologs are EF434781, EF434782, EF434783, EF434784 respectively. The GenBank accession no. of cotton *expansin* (*GhExpA15*) is KY613800.

Analysis of *C. procera* fiber EST's also showed four variants of aquaporins (*CpPiP1, CpPiP2, CpTiP1* and *CpTip2*). The BLAST search indicated that two aquaporins of *C. procera* are plasma membrane intrinsic proteins, while other two are tonoplast intrinsic proteins. The accession numbers of these variants are; KY569400, KY5699399, KY569398 and KY569397 respectively. The aquaporin isolated from *G. hirsutum* (*GhPiP1*) has the accession no. KY613801.

On the basis of ontology group, the variants of expansins and aquaporins were categorized according to their role in developmental processes. Different functions of expansin and aquaporin were inferred by searching the gene ontology database developed in KOBAS for *A. thaliana* (Table 3A and B).

Comparison of *Gossypium* and *C. procera* Expansions

Predotar analysis of all the expansins indicated the association of these proteins with endoplasmic reticulum (ER). The SignalP 3.0. Analysis using neural network (NN) indicated the length and cleavage site of the signal peptides associated with each expansin. Signal peptide cleavage site of both cotton expansins (*GhExpA8* and *GhExpA15*) was at amino acid position 24. The *C. procera* expansins (*CpEXPA1, CpEXPA2, CpEXPA3* and *CpEXPA4*) had signal peptide cleavage site at 23, 27, 25 and 18 amino acid positions, respectively. The maximum cleavage site probabilities were 0.762, 0.751, 0.971, 0.682 and 0.976 for *C. procera* expansins, respectively, while the cleavage site probability for both *G. hirsutum* expansins was 0.88.

Identity of *G. hirsutum* and *C. procera* Expansions

The identity among *C. procera* expansins (CpEXPA1, CpEXPA2, CpEXPA3 and CpEXPA4) and cotton expansins (GhEXPA8 and GhEXPA15) at amino acid level showed maximum percent identity between GhEXPA8 *and* GhEXPA15 (98.3%), while the percent identity of CpEXPA3 with GhEXPA8 and GhEXPA15 remained 76%. The overall identity between the *G. hirsutum* and *C. procera* ranged from 64 to 98%, which indicated that the expensins of the two genera are much similar to each other as compared to aquaporins (Table 4). The evolutionary relatedness of cotton and *C. procera* expansin indicated that GhExP8 and GhExP15 were paralogues and similarly CpEXPA3 and EXPA4 fall in the same clade with the least distance among them. On the other side CpEXPA1 and CpEXPA2 look like outgroups, which indicate that these are highly divergent and might be orthologous to the *C. procera* (Fig. 1).

Comparison of *G. hirsutum* and *C. procera* Aquaporins

The TMPred analysis of *GhPiP1* indicated that it has five transmembrane regions, while all the aquaporins from *C. procera* had six transmembrane regions (http://www.ch.embnet.org/software/TMPRED_form.html). The pairwise comparison between the amino acid sequences of aquaporins indicated that CpTiP1 and CpTiP2 84.52% identity, while *GhPiP1* has 70.24% and 54.67% with *CpPiP1* and *CpPiP2* respectively. Overall, the tonoplast intrinsic proteins (TiP's) had the lowest identity with plasma membrane intrinsic proteins (PiP's) that range between 21–29% (Table 5).The phylogenetic relation of cotton and *C. procera* aquaporins indicate that CpTiP1 and CpTiP2 are paralogues in *C. procera* and have evolved recently. GhPIP1 and CpPIP1 fall in the same clade but the distance between them indicates that those two have evolved a long time ago and are arthologues. The association of CpPiP2 with the same clade indicates that this protein might have evolved a long time ago from the same parent that contributes to CpPiP1 and GhPiP1 (Fig. 2).

Table 3A: Ontology groups of the *expansin* EST's (*GhExp8, GhExp15, CpEXPA1, CpEXPA2, CpEXPA3* and *CpEXPA4*) identified using KOBAS

Category	Gene ontology (GO)	GhEXP8	GhEXP15	CpEXPA1	CpEXPA2	CpEXPA3	CpEXPA4
Plant-type cell wall organization	0009664	+	+	+	-	+	-
Multidimensional cell growth	0009825	+	+	+	-	+	-
Unidimensional cell growth	0009826	+	+	+	-	+	-
Plant-type cell wall modification	0009827	+	+	+	-	+	-
Plant-type cell wall loosening	0009828	+	+	+	-	+	-
Plant-type cell wall modification involved in multidimensional cell growth	0009831	+	+	+	-	+	-

Table 3B: Ontology groups of the *aquaporin* EST's (*GhPiP1*, *CpPiP1*, *CpPiP2*, *CpTiP1* and *CpTiP2*) identified using KOBAS

Category	Gene ontology (GO)	GhPIP1	CpPIP1	CpPIP2	CpTIP1	CpTIP2
Transporter activity	0005215	-	+	+	+	+
Water transmembrane transporter activity	0005372	-	+	+	+	+
integral component of plasma membrane	0005887	-	+	+	+	+
water transport	0006833	-	+	+	+	+
cell volume homeostasis	0006884	-	+	+	+	+
regulation of cell size	0008361	-	+	+	+	+
cellular water homeostasis	0009992	-	+	+	+	+

Table 4: The % similarity between amino acid sequences of expansins from *C. procera* and *G. hirsutum*

Expansins	GhEXPA8 (EU593897)	GhEXPA15 (KY613800)	CpEXPA3 (EF434783)	CpEXPA4 (EF434784)	CpEXPA2 (EF434782)	CpEXPA1 (EF434781)
GhEXPA8(EU593897)	100	98.3	76.37	70.04	61.51	68.03
GhEXPA15(KY613800)	98.3	100	76.79	70.89	61.9	68.85
CpEXPA3 (EF434783)	76.37	76.79	100	80	66.67	66.8
CpEXPA4 (EF434784)	70.04	70.89	80	100	70.2	64.61
CpEXPA2 (EF434782)	61.51	61.9	66.67	70.2	100	57.71
CpEXPA1 (EF434781)	68.03	68.85	66.8	64.61	57.71	100

Note: Low identity=0-50% Medium identity=51-75% High identity=76-100%

Table 5: The % identity between amino acid sequences of aquaporins from *C. procera* and *G. hirsutum*

Aquaporins	CpPIP1 (KY569400)	GhPIP1 (KY613801)	CpPIP2 (KY569399)	CpTIP1 (KY569398)	CpTIP2 (KY569397)
CpPIP1 (KY569400	100	70.24	65.67	27.48	28.48
GhPIP1 (KY613801)	70.24	100	54.67	21.62	21.89
CpPIP2 (KY569399)	65.67	54.67	100	29.66	28.97
CpTIP1 (KY569398)	27.48	21.62	29.66	100	84.52
CpTIP2 (KY569397)	28.48	21.89	28.97	84.52	100

Note: Low identity=0-50% Medium identity=51-75% High identity=76-100%

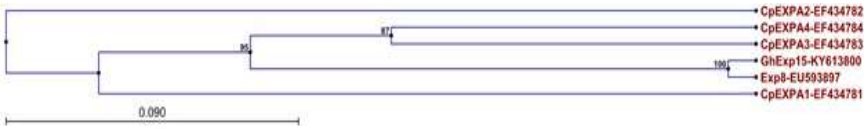

Fig. 1: Phylogenetic tree was constructed using amino acid sequences of *expansin* gene family isolated from *G. hirsutum* (GhEXPA8 and GhEXPA15) and *C. procera* fibers (CpEXPA1, CpEXPA2 and CpEXPA3)

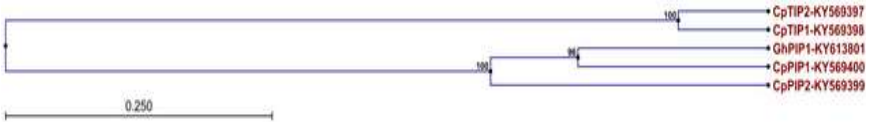

Fig. 2: Phylogenetic tree was constructed using amino acid sequences of *aquaporins* gene family isolated from *G. hirsutum* (GhPiP1) and *C. procera* (CpPiP1, CpPiP2, CpTiP1 and CpTiP2) fibers

Expression Analysis of Cotton and *C. procera* Fiber Specific Genes

The total RNA isolated from different *C. procera* and *G. hirsutum* tissues was of high quality (Fig. 3). Equal amount of templated cDNA was used for qRT-PCR, as represented by the amplification of 18srRNA (Fig. 4a). RT-PCR result showed the differential transcription of *sucrose synthase*, *expansin* and *aquaporin* genes (Fig. 4b, 4c and 4d). The transcriptome profiling of the three fiber morphogenesis gene families (*expansins*, *aquaporins* and *Sus*) isolated from *G. hirsutum* and *C. procera* indicated that the expansins and aquaprins are highly transcribed in developing *C. procera* fibers as compared to *G. hirsutum*. The *C. procera* has more variants of the expansins however, none of these expansins is specific to the *C. procera* fibers. The *C. procera* fiber transcripts indicated highest transcript level of *CpEXPA3*, while the other three expansins were found to be at markedly low concentrations. On the other hand, the GhExp8 and GhExP15 are the fiber specific proteins in cotton (Fig. 5A).

Expression profiling revealed that expression level of *C. procera equaporin (CpPiP2)* was highest in *C. procera* fibers, while no expression was seen in cotton fibers. The expression level of *CpPIP1* was detected in *C. procera* as well as in cotton fibers (Fig. 5B). The expression level of *C. procera expansins (CpExpA1, CpExpA2, CpExpA3 and CpExpA4)* was detected in root, shoot, fiber and leaves in *C. procera*. Highest expression level was detected in cotton stem and leaves, while minimum expression was seen in cotton roots. In fiber, *CpExpA2 and CpExpA4* showed minimum expression, while *CpExpA1* showed normal expression level. High level of expression was expressed by *CpExpA3* (Fig. 5A). Expression analysis of Cotton *Sus* gene revealed that high transcripts were present in 5, 10 and 15 DPA cotton fibers. Minimum expression level was noticed in stem, shoots and leaves at 0 and 20 DPA (Fig. 5C).

Discussion

Fiber development is a complex process in cotton due to primary and secondary cell wall deposition, which requires a large number of gene families in the fiber development process (Arpat *et al*., 2004). Modification of cotton fiber traits utilizing the classical breeding and transgenic technology for improving traits like length, strength and finesse is a crucial task for researchers due to negative association between fiber length and other quality traits (Azhar *et al*., 2004). The work on the development of fiber EST's of *G. hirsutum* and *C. procera* revealed several gene families including lipid transport proteins, aquaporins, actins, tubulins, E6 family proteins, proline rich proteins, expansins and sucrose synthases represent the most frequent transcripts in developing fibers. The *C. procera* produces long, fine, strong and shiny single celled seed fibers as compared to other fiber producing resources.

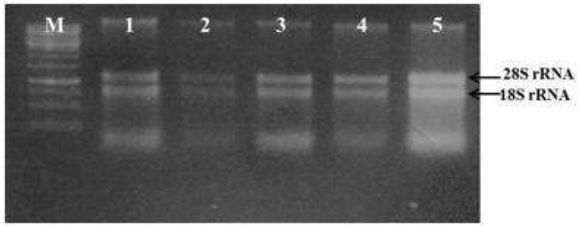

Fig. 3: Total RNA isolated from different cotton tissues. M: 1 kb ladder; Lane 1-5: total RNA isolated from root, shoot, stem, leaves and fibers, respectively

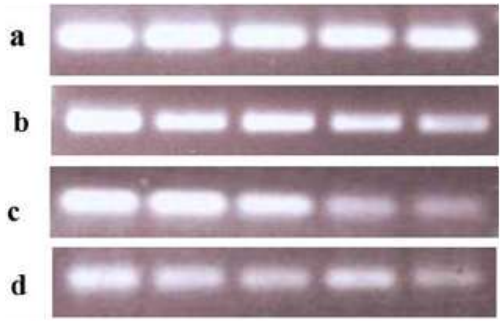

Fig. 4: PCR with different primer pairs at 10 fold dilution series of cotton fiber cDNA. a) PCR with *18S rRNA* primers, b) *Sucrose synthase*, c) *expansin*, d) *aquaporin*

BLAST search indicated that *C. procera* fiber transcripts are rich in *aquaporins*, *expansins* and ion transporters. The isolation of four *expansin* homologs (*CpEXPA1*, *CpEXPA2*, *CpEXPA3* and *CpEXPA4*) and four *aquaporins* homologs (*CpPiP1, CpPiP2, CpTiP1 and CpTip2)* in elongating fibers suggested a strong role of these gene families in *C. procera* fiber development. SignalP 3.0 is a computational method used to predict N-terminal signal peptides and is the most prevalent and advanced version (Bendtsen *et al*., 2004). Results of SignalP 3.0 showed that expansin protein is ER oriented where the cleavage of signal peptides occurred.

Pairwise alignment between the *G. hirsutum* and *C. procera* expansins indicated that CpEXPA3 has 76% amino acid identity with GhEXPA15 and GhEXPA8, and was the closest homologue of *G. hirsutum* expansins (Table 4). The real-time transcriptome profiling of *CpEXPA3* revealed that this expansin homologue is highly observed in *C. procera* fibers and might have a major involvement in fiber elongation. Phylogenetic analysis of the expansin genes from two genera indicated that the two homologues from *G. hirsutum* are paralogous and have recently been evolved. The CpEXPA3 and CpEXPA4 had 80% amino acid identity and were observed in the same clade indicating that those are paralogues too but evolved a long time ago. The CpEXPA1 and CpEXPA2 were found to be highly divergent from all the identified expansins in *G. hirsutum* and *C. procera*. On the other hand, the transcript level of *CpEXPA3* was 40% higher than *CpEXPA1*.

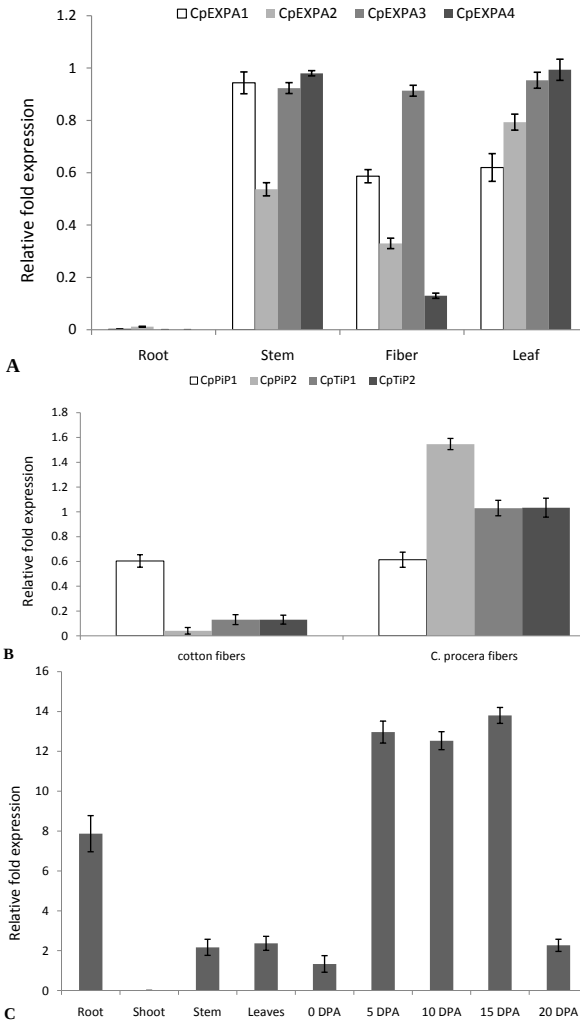

Fig. 5: Transcript profiling of fiber morphogenesis genes in *C. procera* and various tissues of cotton plant by Q-PCR (A) Transcript profiling of *C. procera expansins* in various tissues including roots, stem, fibers and leaves. *CpEXPA3* is highly expressed in *C. procera* fibers (B) Expression of *aquaporins* variants in *C. procera* and *G. hirsutum* fibers. The expression of *PiP2* is high in *C. procera* fibers (C) Expression profiling of *Susy.* in cotton tissues and fibers (0-20 DPA). Expression level was high at 5-15DPA

This indicated that *C. procera* fiber elongation requires two highly divergent expansins (CpEXPA3 and CpEXPA1), while CpEXPA2 and CpEXPA4 might play a negligible role in elongating *C. procera* fibers.

Expansin gene family has multifunctional groups of proteins, which have very specific role in developmental processes like seed germination, control of organ size, morphology, fruit softening, pollen tube invasion in the grass stigma and fiber development (Lee *et al.*, 2003; Sabirzhanova *et al.*, 2005). The EXPA and EXPB group proteins are involved in cell wall loosening. It has been proposed that under the given cellular turgor pressure the expansins weaken the non-covalent binding between cell wall polysaccharides, thereby allowing the cellulose fibrils to slide over each other (Marga *et al*., 2005). Over expression of *CpExpA3* (having highest expression in *C. procera* fibers) in cotton fibers might increase fiber length by enhancing the primary cell wall loosening (Lee *et al*., 2003; Sabirzhanova *et al*., 2005; Wilkins and Arpat, 2005; Cheema *et al*., 2010; Bajwa *et al*., 2013; Iqbal *et al.*, 2016). The expression level of the four *C. procera expansin* genes was detected in four different *C. procera* tissues. All the variants showed high expression in stem and leaves. None of the variants was significantly detected in root tissues indicating that *C. procera* might have some negligible amount of root specific expansins (Fig. 5A).

C. procera is a wild plant and grows well under extreme draught, temperature and salt conditions. It is expected that the plant has acquired the highly efficient genes for water and salt transport. Aquaporins are the key water transport proteins in plants (Siefritz *et al*., 2002). In developing fibers, these are located at the plasmodesmatal connections of the fiber cells (Iqbal *et al*., 2008). Two types of aquaporins, tonoplast and plasma membrane associated, were detected in the *C. procera* fiber EST's. The CpTiP1 and CpTiP2 represented the tonoplast associated membrane proteins, while CpPiP1 and CpPiP2 represented plasma membrane associated proteins. The CpPiP1 was found to have 70% amino acid identity to the GhPiP1 indicating that it was the closest homologue of GhPiP1. Interestingly, the CpPiP1 and CpPiP2 were identified to be highly expressed in *C. procera* fibers, indicating that CpPiP2 is acting as a second water pump in building up the turgor pressure for greater expansion in *C. procera* fiber tissues. The phylogenetic analysis revealed that CpPiP1 lies in the same clade of GhPiP1 but the 70% amino acid identity indicated that the two proteins are distantly related. The CpPiP2 had 65% and 54% amino acid identity with the CpPiP1 and GhPiP1 respectively, which indicated that the two proteins are highly divergent both in cotton and *C. procera*. However, CpPiP2 might be a good candidate gene for enhancing the buildup of water turgor pressure in developing cotton fibers.

Transcript level of *Sus* gene from cotton revealed high expression in rapidly elongating cotton fibers at 5, 10 and 15 DPA (Fig. 5C), which coincides well with the previous reports (Bolton *et al*., 2009; Iqbal *et al.*, 2016). Transcript level of the gene dropped after 15 DPA that might be due to the deposition of secondary cell wall and progression towards cell death (Brill *et al.*, 2011). It is revealed that Sus has the key role in strengthening the plant cell wall as well as fiber initiation (Amor *et al*., 1995). These findings correlated with the findings of another experiment, which described that suppression of *Sus* has resulted in fiber less seeds in cotton. The over expression of *Sus* in cotton lines (Fiber Max) have been reported to produce stronger and longer fibers (Arioli, 2005). It is postulated that the natural

expression of this gene might not coordinate well with the rapidly expanding fiber cells when the extraneous *expansin* gene is used to increase fiber expansion in transgenic plants. It is, therefore, proposed that the genes for fiber expansion (*expansin*) should be used together with the *Sus* gene to compensate cellulose synthesis in the cotton transgenic plants for longer and stronger fibers.

C. procera aquaporin *CpPiP2* displayed the highest expression level in *C. procera* fibers while no expression was seen in cotton fibers (Fig. 5B). Although aquaporins are associated with water transport, they can play crucial role in cotton (*G. hirsutum*) fibre elongation. Remarkable reduction in fiber elongation was detected when the expression of *GhPIP2* in cotton was suppressed by RNA interference (Cosgrove, 1986). It is suggested that *CpPiP2* is a candidate gene for cotton transformation to enhance the fiber length because previous study (Aslam *et al*., 2013) showed that over expression of *CpPiP2* in tobacco result in elevating the length of stem and leaves trichomes.

Conclusion

The analysis of fiber ESTs and transcriptome profiling of fiber genes revealed that *aquaporins*, *expansin* and *S. synthase* are closely associated with elongating fibers. It is concluded from the present study that novel genes can be explored for fiber morphogenesis from diverse resources by applying the bioinformatics tools and transcriptome profiling. The isolated *Aquaporin* and *expansin* homologs from *C. procera* represent useful natural resources to modify the fiber characteristics. Transcriptome profile analysis revealed that over expression or genes from diverse resources for cell expansion (*expansins*), turgor pressure development (*aquaporins*) and UDP glucose pathway (*Sucrose synthase*) can be used to enhance the cotton fiber characteristics related to length and strength.

References

Amor, Y., C.H. Haigler, S. Johnson, M. Wainscott and D.P. Delmer, 1995. A membrane-associated form of Sus and its potential role in synthesis of cellulose and callose in plants. *Proc. Natl. Acad. Sci*., 92: 9353–9357

Arioli, T., 2005. Genetic engineering for cotton fiber improvement. *Pflanzenschutz-Nachrichten. Bayer*, 58: 140–150

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

Aslam, U., A. Khatoon, H.M.N. Cheema and A. Bashir, 2013. Identification and characterization of plasma membrane aquaporins isolated from fiber cells of *Calotropis procera. Biomed. Biotechnol*., 14: 586–595

Azhar, F.M., M. Naveed and A. Ali, 2004. Correlation analysis of seed cotton yield with fiber characteristics in *Gossypium hirsutum* L. *Int. J. Agric. Biol*., 6: 656–658

Bajwa, K.S., A.A.R. Shahid, A. Qayyum, A. Kiani, M. Sarfraz, M.A. Ashraf, A.A. Dahab, A. Bakhsh, A. Latif, M.A.U. Khan, A.N. Puspito, A. Aftab, A. Bashir and T. Husnain, 2013. Expression of *Calotropis procera* expansin gene *CpEXPA3* enhances cotton fibre strength. *Aus. J. Crop Sci*., 7: 206–212

Bendtsen, J.D., H. Nielsen, G. von Heijne and S. Brunak, 2004. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol*., 4: 783–795

Bolton, J.J., K.M. Soliman, T.A. Wilkins and J.N. Jenkins, 2009. Aberrant expression of critical genes during secondary cell wall biogenesis in a cotton mutant, ligonlintless-1 (Li-1). *Comp. Funct. Genom.*, 659301: 1-8

Brill, E., M. van Thournout, R.G. White, D. Llewellyn, P.M. Campbell, S. Engelen, Y.L. Ruan, T. Arioli and R.T. Furbank, 2011. A novel isoform of sucrose synthase is targeted to the cell wall during secondary cell wall synthesis in cotton fibre. *Plant Physiol*., 157: 40–54

Carbrey, J.M., M. Bonhivers, J.D. Boeke and P. Agre, 2001. Aquaporins in *Saccharomyces*: characterization of a second functional water channel protein. *Proc. Natl. Acad. Sci*., 98: 1000–1005

Cheema, H.M.N., A. Bashir, A. Khatoon, N. Iqbal, Y. Zafar and K.A. Malik, 2010. Molecular characterization and transcriptome profiling of *expansin* genes isolated from *Calotropis procera* fibers. *Electr. J. Biotechnol*., 13: 1–14

Cosgrove, D.J., P. Bedinger and D.M. Durachko, 1997. Group I allergens of grass pollen as cell wall-loosening agents. *Proc. Natl. Acad. Sci*., 12: 6559–6564

Cosgrove, D.J., 1986. Biophysical control of plant cell growth. *Annu. Rev. Plant Physiol*., 37: 377–405

Deen, P.M.T. and C.H. van Os, 1998. Epithelial aquaporins. *Curr. Opin. Cell Biol*., 10: 435–442

Earl, W.T. and B. Deborah, 2007. Analysis of gene expression in cotton fiber initials. *BMC Plant Biol*., 7: 22–34

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

Iqbal, S., A. Bashir, H.M. Naseer, M. Ahmed and K.A. Malik, 2008. Identification of differentially expressed genes in developing cotton fibers (*Gossypium hirsutum* L) through Differential Display. *Electr. J. Biotechnol*., 11

Iqbal, N., A. Khatoon, M. Asif and A. Bashir, 2016. Expression analysis of fiber related genes in cotton (*Gossypium hirsutum* L.) through real-time PCR. *Pak. J. Bot*., 48: 1099–1106

Kammerloher, W., U. Fischer, G.P. Piechottka and A.R. Schaffner, 1994. Water channels in the plant plasma membrane cloned by immune selection from a mammalian expression system. *Plant J*., 6: 187–199

Kende, H., K.J. Bradford, D.A. Brummell, H.T. Cho, D.J. Cosgrove, A.J. Fleming, C. Gehring, Y. Lee, S. Mcqueen-Mason, J.K.C. Rose and L.A.C.J. Voesenek, 2004. Nomenclature for members of the expansin superfamily of genes and proteins. *Plant Mol. Biol*., 3: 311–314

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

Kozono, D., X.D. Ding, I. Iwasaki, X.Y. Meng, Y. Kamagata, P. Agre and Y. Kitagawa, 2003. Functional expression and characterization of an archaeal aquaporin-AqpM from *Methano thermobacter marburgensis*. *J. Biol. Chem*., 278: 10649–10656

Lee, D.K., J.H. Ahn, S.K. Song, Y.D. Chol and J.S. Lee, 2003. Expression of an expansin gene is correlated with root elongation in soybean. *Plant Physiol*., 3: 985–997

Li, L.C., P.A. Bedinger, C. Volk, A.D. Jones and D.J. Cosgrove, 2003. Purification and characterization of four *β-expansins* (Zea m 1 isoforms) from maize pollen. *Plant Physiol*., 4: 2073–2085

Marga, F., M. Grandbois, D.J. Cosgrove and T.I. Baskin, 2005. Cell wall extension results in the coordinate separation of parallel microfibrils: evidence from scanning electron microscopy and atomic force microscopy. *Plant J*., 2: 181–190

Ruan, Y.L., D.J. Liewellyn and R.T. Furbank, 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: 47–63

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

Sabirzhanova, I.B., B.E. Sabirzhanov, A.V Chemeris, D.S. Veselov and G.R. Kudoyarova, 2005. Fast changes in expression of expansin gene and leaf extensibility in osmotically stressed maize plants. *Plant Physiol. Biochem*., 4: 419–422

Sakurai, J., F. Ishikawa, T. Yamaguchi, M. Uemura and M. Maeshima, 2005. Identification of 33 rice aquaporin genes and analysis of their expression and function. *Plant Cell Physiol*., 46: 1568–1577

Schipper, O., D. Schaefer, R. Reski and A. Fleming, 2002. Expansins in the bryophyte *Physcomitrella patens. Plant Mol. Biol*., 4: 789–802

Siefritz, F., M.T. Tyree, C. Lovisolo, A. Schubert and R. Kaldenhoff, 2002. PIP1 plasma membrane aquaporins in tobacco: from cellular effects to function in plants. *Plant Cell*, 14: 869–876

Tyerman, S.D., C.M. Niemietz and H. Bramley, 2002. Plant aquaporins: multifunctional water and solute channels with expending roles. *Plant Cell Environ*., 25: 173–194

Wilkins, T.A. and A.B. Arpart, 2005. The cotton fiber transcriptome. *Physiol. Plant*., 124: 295–300

(Received 15 July 2017; Accepted 04 October 2017)