Global Journal of Biotechnology & Biochemistry 4 (1): 10-18, 2009
ISSN 2078-466X

Characterization of Cytoskeletal and Cell Growth ESTs in Preferentially Expressed Extra-Long Stable Fiber of Giza88 Variety

[1]Osama A. Momtaz and [2]Ahmed El-Fatih A. El-Doliefy

[1]Deputy Director for Research, Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Center (ARC), Ministry of Agriculture (MOA), Egypt
[2]Laboratory of Gene Expression and Regulation Technologies (GERT), AGERI, ARC, MOA, Egypt

Abstract: International as well as national economics are largely influenced by cotton textile industries and its associated technologies which basically depend on fiber quality. Therefore, Egyptian cotton biotechnology approaches adapted novel advances to develop new varieties for Delta regions and reclaimed areas to enhance the superior extra-long stable fiber characteristics as well as production per area. Fiber quality is highly affected by cell structure (cytoskeleton) motives and growth regulating cascades; in addition to overlapping environmental factors. As a result, fiber pre- and post-anthesis genes (e.g., expansin) parallel with fiber cytoskeleton and especially those of preferential expression (e.g., actin and tubulin) undertook intensive research. As a step toward elucidating the genomic EST-motive network responsible for cotton fiber cell growth and quality; isolated and purified genomic Giza88 PAC clones were characterized and screened. Furthermore, PCR-based screening and restriction map positional cloning for the identified motifs were investigated. Bioinformatics diagnostic approaches for the identified motifs (GbACT, FATUB and FATEXP3F) illustrated the evolutional phylogenic relationship between *G. barbadense* and other *Gossypium spp*. This is a pioneer step to study fiber QTL tags for preservation and improvement of fiber quality as marker assistance selection procedures for Egyptian cotton breeding programs.

Key words: *G. barbadense* · QTL-PAC library · PCR-screening · Comparative evolution-pre-post-anthesis-expansin- actin- tubulin-bioinformatics-structural genomics

INTRODUCTION

Gossypium barbadense is the most important fiber crop species for textile industry in Egypt and worldwide. Although classical breeding has improved cotton quality and yield in the 20th century, further improvements of fiber quality are needed [1]. Cotton fibers are botanically extreme long single epidermal cells (trichomes) that develop on the outer surface of ovule, reaching about 5 cm in some species. Probably, both fiber types (long lint and short) are common in early differentiation of developmental pathways; however, biochemical and molecular studies were focused on lint fiber for its economic significance [2].

Recent studies of comparative gene expression profile in elongating fibers using cDNA macroarray have identified new preferentially expressed genes such as auxin-binding protein, a mitogen-activated protein kinas, RD22-like protein and actin1 [3]. Besides, some subtractive hybridization studies had identified number of enzymes and cell wall proteins, yet, the physiological or biochemical roles are to be verified [4]. Moreover, cytoskeleton genes (α- and β-tubulin) play a dual role in cell pre-anthesis elongation and post-anthesis secondary cell wall biosynthesis [5]. As well, cell wall-modifying genes, such as endo-1, 4-β-glucanase and expansin, perform in concert with cell turgor allowing cell expansion [6]. Genetic studies showed that actin cytoskeleton by interacting with ARP2/ARP3 protein complex, plays a pivotal role in controlling fiber cell shape and several other cell types [7]. Cotton fiber index is structurally depend on the cytoskeleton (actin and tubulin) that builds its shape as well as its main characteristics; however, it is affected by many cytosolic proteins such as expansin during developmental stages.

Corresponding Author: Dr. Osama A. Momtaz, Deputy Director for Research, Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Center (ARC), Ministry of Agriculture (MOA), Egypt

Fiber cell wall contains, not only cellulose, but also several other polysaccharides, which are modified by enzymes and other proteins (e.g., expansins) that able to weaken the linkage between cellulose microfibrils allowing cell wall extension through its polymer creeping which is mechanically forced by cell turgor pressure [8]. Expansins are implicated in the control of plant cell modification, growth, separation, wall dissolution, pollen-tube invasion and other processes [9] and [10]. On the other hand, microtubule proteins were integrated in the intracellular processes including morphogenesis, cell growth and division [11]. In *Arabidopsis*, 6 α-tubulin, 9 β-tubulin genes; while in maize, 7 α-tubulin, 6 β-tubulin genes; however in cotton, 9 α-tubulin and 7 β-tubulin isotypes were identified on immune-blots of 2D gels with preferential expression of α-tubulin in fiber cells [12]. Cotton fiber β-tubulin was first isolated, molecular characterized and preferentially expressed as reported by Li *et al.*, [5]. Nonetheless, α-tubulin in *G. barbadense spp*. had not been investigated. Examining the growing number of cotton expressed sequence tags (ESTs) as a genomic approach associated with elongation of cotton fibers, aimed to define different physiological and biochemical processes occurring in fiber cells at different developmental time intervals [13].

This work represents a pioneer study for characterization of cytoskeletal and cell growth ESTs in preferentially expressed Egyptian cotton extra-long stable fiber of Giza88 variety. Besides, it is to introduce extra-long stable high quality superior regulating factors markers at the inter-varietal level in Egyptian cotton breeding program. PCR-based screening for some pre and post-anthesis gene regulatory sequences took place and further confirmatory investigation for the PAC clone characterization and analyses was reported. Fragment gene identification of each cloned gene family was subcloned, sequenced and analyzed using the bioinformatics software's in a comparative study. Finally, it is used as an intermediate marker for identification and characterization of particular related-gene families from a wide- genomic screening pool of *G. barbadense* Giza 88 PAC library.

MATERIALS AND METHODS

Plant Material and Nucleic Acid Isolation: The seeds of the Egyptian extra long stable variety (Giza88) *Gossypium barbadense* L. were de-linted and planted in greenhouse using 8 /16 hours night/day at 28°C with humidity set up on 50%. Leaves at 4-6 weeks age were collected, frozen using liquid nitrogen and stored at -80°C. Genomic DNA was isolated from young leaves of Giza88 using Qiagen DNeasy™ Plant Mini kit (Cat. No. 69104) following the manufacturer manual. This genomic DNA was used as positive control templates for all PCR reactions. Nucleic acids NO (s) samples were stored at -80°C.

DNA Vector Preparation: Two plasmids were used in this study, the Promega pGEMTeasy® vector system I (Cat. No.A1360). that was used for sub-cloning maintenance of the identified fragments (GbACT, FATEXP3F and FATUB). The second was pPACe4 (19.5 Kb) vector that developed by [14] and purchased from The Children's Hospital Oakland Research Institute (CHORI) (http://bacpac.chori.org/ppac4.htm) and used for the construction of the genomic library. pPACe4's bacteria were reactivated, prepared in large scale and plasmid isolated using Plasmid Qiagen Max kit (Cat. No. 12163) following the manufacturer manual. The pPACe4 vector was restriction digested using *BamH*I (Amersham Pharmacia, USA) at 30°C overnight. Restriction enzyme digested pPACe4 vector was de-phosphorylated using 0.1 unit of calf intestinal alkaline phosphatase (CIAP) (Biolabs, USA) per p.mol DNA ends at 37°C for 30 min. Reaction was stopped using stopping solution (5mM EDTA, pH:8; 0.5% SDS and 0.1 mg/ml proteinase K). Dephosphorylation was then heat inactivated, followed by double phenol/chloroform/iso-amyl (25/24/1 v/v) purifications. Afterwards, ethanol precipitation took place and samples were dried and dissolved in 50 µl ddH_2O. The large dephosphorelated fragment vector (16.7 Kb) was fractionated and extracted from the gel using QIAEX® II Gel Extraction kit (Cat. No.20051). By then, PAC vector DNA was ready for ligation reaction and stored in 5 µl aliquots at -80°C.

Isolation and Purification of pPACe4 Clones: Some modifications were applied for Giza88 nuclei isolation and purification from the young leaves yet basically depended on the protocol presented by Zahng *et al.* [15] and Paterson *et al.* [16]. Modifications included pretreatment of leaves using diethyl ether for few seconds followed by 1X TE (pH: 8) washing. Washed leaves were ground in liquid nitrogen then transferred into 100 ml of 1X homogenization buffer (HB), however; β-mercabtoethanol and TritonX100 were immediately added to the homogenized mix. Homogenized tissue mix was squeezed and filtered into 250 ml centrifuge bottle. Bottles were Beckman centrifuged at 2500 rpm for 21 minutes at 8°C. Nuclei pellet was re-suspended,

washed, re-filtered, re-centrifuged and transferred to 1.5 ml eppendorf tube using cut end pipette tips. Afterwards, the pellet was washed at bench-top centrifuge (1100 rpm) for 2 minutes at 4°C as 3 times wash using 1 ml of nuclei wash buffer without disturbing the nuclei pellet. Finally, nuclei pellet was suspended and embedded in low melting point (LMP) agarose plugs as described by Paterson *et al.* [16].

Nuclei plugs were lysed at 50°C water bath for 48 hours, washed several times using TE buffer, restriction digested using 1.3 units of *BamH*I (Pharmacia) and applied to the Bio-Rad clamped-contour homogenous electrophoresis field (CHEF) DR® III System for 18 hours; (4 and 40 sec) as initial and final ramping time, respectively; 120 angle and 6 V/cm with pump speed set at 80 and 1X TAE run buffer cooled at 4°C. Cotton high molecular weight (HMW) DNA was excised from agarose, purified using β-agarase enzyme (Biolabs) as specified in the manufacturer manual; and then spot-dialyzed on VSWP filter membrane 0.025 μM (Millipore, USA) as described by Peterson *et al.* [16]. Two μl (20ng) dephosphorylated pPACe4 fragment (DPF) (16.7 Kb) and 2 μl (80-100ng) size selected insert DNA (SSID) were gently mixed and incubated at 50°C for 2 min followed by ligation using 200 units of T4 DNA ligase enzyme in 15 μl total volume. Ligation ratio was applied as 1:5 (DPF: SSID) for 3 h at room temperature, then overnight at 16°C. Recombinant DNA was transformed into electro-competent *DH101β* strain cells using Bio-Rad Gene Pulser™. Ratio of PAC-Ligation mix to competent cells was set as 1:15-20 (v/v), as described by Momtaz *et al.* [17]. Constructed PAC colonies were grown on agar LB/sucrose/Kanamycin plates then picked and stored at -80°C as glycerol stocks.

Plasmid DNA from PAC clones was randomly selected, isolated and identified using a modified alkaline lyses procedure from 5 ml overnight cultures in (LB) medium containing kanamycin 50 mg/ml [17]. PAC clones were size characterized using pulsed-field gel electrophoresis (PFGE) Momtaz *et al.* [20]. Purified recombinant PAC DNA was eluted in ddH_2O, rather than TE buffer, used by most BAC/PAC library manuals, to overcome troubleshooting in further restriction enzymatic reactions affected by EDTA presence.

PCR-Screening for Pre- and Post-Anthesis Gene Families: Random DNA clones from the constructed library were selected for PCR-based screening using three different related-gene family's specific primers covering the targeted pre- and post-anthesis expressed sequences, as shown in Table 1.

PCR screening reaction was performed in 50 μl total reaction volume as described by Momtaz *et al.* [17]. PCR conditions included 3 min. for first DNA denaturation step at 94°C followed by 35 cycles at 94°C for 30 sec., annealing step at 58°C (actin), 60°C (expansin) and 50°C (tubulin) for 30 sec. and extension step at 72°C for 90 sec. while the final extension cycle at 72°C for 10 min. Exclusively, tubulin PCR was performed for 40 cycles. All PCR-based screening reactions were fractionated on 3% TAE agarose gel.

Sequencing and Data Analysis: Fragments generated from PCR-based screening were sub-cloned in pGem vector as the manufacturing manual. Vector was then subjected to the automated DNA sequencing method. Sequencing reactions were performed using ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (PE applied Biosystems, USA) in conjunction with ABI PRISM (310 Genetic Analyzer). Cycle sequencing was performed using the Gene Amp 2400 Thermal Cycler, the reaction was conducted in a total volume of 20 μl containing 8 μl of terminator ready reaction mix, 1 μg of plasmid DNA and 3.2 pmol of M13 universal forward primer. The cycle sequencing program was 96°C for 10 sec., 50°C for 5 sec. and 60°C for 4 min. repeated for 25 cycles with rapid thermal ramping, the nucleotide sequence was determined automatically by the electrophoresis of the cycle sequencing reaction product on 3100 Genetic Analyzer. The data were provided as fluorimetric scans from which the sequence assembled using the sequence analysis software. Data analysis for each sequence was determined using the BLASTn alignment technology at the websites of NCBI (http://www.ncbi.nlm.nih.gov), in addition to the DNASTAR Lasergene software program (ver. 7.1.0 (44) copy right 1989-2008).

RESULTS AND DISCUSSION

Screening of the Giza88 PAC Library: The cloned PAC vector was prepared by using a high quality and purified Qiagen plasmid isolation protocol as described by manufacturer manual in contrary to the cesium gradient protocol with ultra speed centrifugation generally used for that purpose. This modification was processed to overcome, mega-base DNA tricky situations such as nuclei low yielding, shearing and oxidation of polyphenolic compounds, which makes DNA useless. The protocol of high pH supplemented with sucrose gradient buffer developed by Zhang *et al*. 20 was

Table 1: Pre- and Post- Anthesis degenerate primer sequences

Fragment name	Degenerate primer sequence	Gene family
GbACT	F5'-TGC CCG GAA GTC CTC TTC CAG-3' R5'-ATT TTC CCA GA AGT TTG ACC GCG C-3'	Actin
FATEXP3	F5'-GCT GAA GTG TGT GAG TGA TCC-3' R5'-CAA CTG GAT GGA GCA ACG TCG-3'	Expansin
FATUB	F5'-CAT GGC TTG YTG TTT GAT GTA YCG-3' R5'-CCT CAC GAG CCT CAG AGA AYT CTC C-3'	Tubulin

performed [18]. This modification was processed to overcome, mega-base DNA tricky situations such as nuclei low yielding, shearing and oxidation of polyphenolic compounds, which makes DNA useless. The protocol of high pH supplemented with sucrose gradient buffer developed by Zhang *et al.*, [19] was performed. In addition, antioxidants such as Triton X-100 and β-mercabtoethanol that destroy polyphenoles, chloroplast and mitochondria membranes were added. Dual effect of low speed centrifugation within the protocol not only, saves intact nuclei from negative effects, but also, it discards starch grains from nuclei pellets [16]. Although the manual does not work well, when directly applied on the Giza88, however, some modifications were applied. The major modification was centrifuge speed by which the nuclei were collected. Both Zhang's and Peterson's manuals used 2700 xg for cereal's nuclei precipitation, while the successful speed for Giza88 nuclei's was 2500 rpm for the 250 ml bottle using Beckman centrifuges and 1100 rpm for the bench top centrifuges using eppendorf tubes. Forward wise, nuclei lyses step, double concentration of proteinase K was used instead of the 0.1-0.5 mg/ml used by Zhang *et al*. [15].

Separations of the mega-base DNA, in addition to the first and second size selections of the HMW DNA that will take part in PAC cloning, were applied on the CHEF apparatus. The dephosphorylated pPACe4 fragment (DPF) (16.7 Kb) and size selected insert DNA (SSID) were gently mixed and incubated for ligation reaction in a ratio of 1:5 (vector: insert) [20]. Characterization of undigested recombinant PACs was performed against the dephosphorylated non-recombinant pPACe4 vector as described by Momtaz *et al*. [20]. Giza88 genomic PAC library comprised 8900 PAC clones with 70 Kb average sizes; representing 0.3 equivalents to the haploid genome (2118 Mb) of *Gossypium barbadense*. The library was alphabetically labeled with numbers and stored in two sets at -80°C as well as glycerol stocks [17].

Screening was approached using PCR, for its fast and accurate results. Conserved fiber structural and quality related primers that matched the *G. hirsutum* NCBI gene bank database for actin, expansin and tubulin mRNAs were used (Table 1). Nevertheless these degenerate primers were used to amplify and screen the specific motifs from a genomic DNA of Giza88 PAC clones. Randomized recombinant PACs were isolated and PCR-based screened. PCR screening produced 33 positive fiber structural and quality related gene families. Precisely, 14 PAC clones comprised some actin-conserved fragments [17], 10 for expansin and 9 for tubulin. After second PCR confirmation, four clones represented actin positive PAC clones containing the conserved actin fragment sizes (~200 bp, ~450 bp, 550 bp and ~859 bp) as shown (Fig. 1a). These data partially agreed with those of Li *et al*. [3] using real-time PCR in *G. hirsutum* stating that actin was preferentially expressed in cotton fiber expansion stage. On the other hand, EPAC 30, 42 and 77 clones were produced in relation to the expansin gene family represented by (FATEXP3) at molecular size 532 bp (Fig. 1b). Similarly, NPAC 40 and 41 clones representing tubulin gene family (FATUB) with a molecular size 402 bp (Fig. 1c) were produced. Neither GbACT and FATEXP3 nor FATUB bands were appeared in the pPACe4 vector PCR reaction (negative control). The data suggested that these bands refer to some gene related regions found in the inserted DNA fragment encountered by these positive PAC clones and not related to any sequence of the pPACe4 vector itself.

Identification and Positional Cloning of Some Fiber Conserved Motives: MPAC94 clone was isolated and purified then DNA inserted in this clone was restriction enzyme digested and agarose gel simulated for fingerprinting as described by Xu *et al.* [21]. The resulted fragments present in the digestion pattern of MPAC94 and were absent in the digested pPACe4, negative control, were selected, eluted, purified, fractionated and sequenced as individual PCR-based positional cloning. Positional cloning contributed the smallest restricted fragment (16.26 Kb) comprising (GbACT) the actin conserved region (859 bp) as determined by the Gel Doc 2000 System (Bio-Rad) software [17].

Comparative Analysis of Giza88 Identified Motifs: PCR screening of the identified PAC clones have produced fiber pre-anthesis preferentially expressed sequences (GbACT and FATUB) in addition to a post-anthesis sequence (FATEXP3F). Comparative analysis of FATUB showed bias to α-tubulin sequences (Fig. 2a). The FATUB DNA sequence alignment, using NCBI database, emphasized similarity bias to α tubulin sequences against the rest of the tubulin family

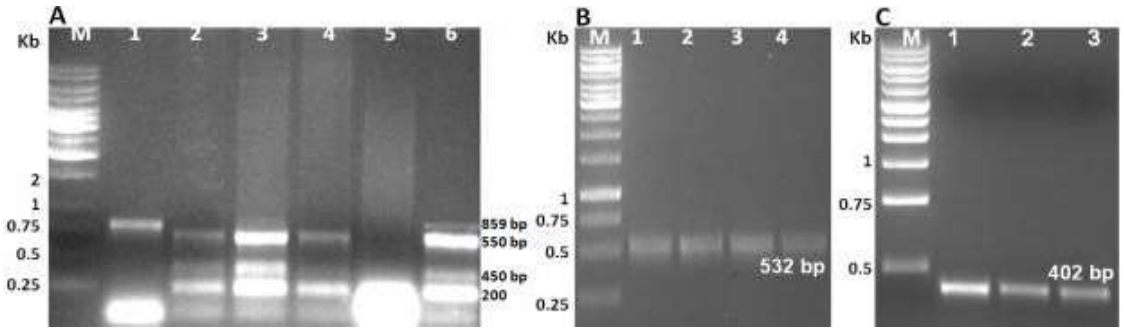

Fig. 1: PCR based confirmation screening for positive PAC clones. **A**). Lane M: 1Kb DNA ladder (Gibco). Lane 1: PCR reaction of the actin (859 bp) from genomic DNA positive control. Lanes (2-6): MPAC clones 9, 11, 12, 14 and 94, respectively. The five lanes shared actin conserved regions 200, 450, 550 and 750 bp. **B**). lane (M): 1Kb DNA ladder (Fermintus). Lane 1: The genomic positive expansin. Lanes (2-4): expansin positive EPAC clones 30, 42 and 77. All expansin fragments were at the same molecular size (532 bp). **C**). Lane (M): 1Kb DNA ladder (Biolabs). Lane 1: the genomic positive tubulin. Lanes (2-3): tubulin positive NPAC clones 40 and 41. All tubulin fragments were in the size of 402 bp.

1	CACCGTGTCA	ACCAATGGCG	GCCGCGGGAA	TTCAGATTCA	TGGCTTGAAT	50
51	TTTCGATGTA	CCGAGGAGAG	AGATGTGCCC	AAAAATGTGA	ATGCAGCTGT	100
101	GGCTACCATC	AAGACCAAGC	GCACAATCCA	ATTTGTCGAT	TGGTGCCCTA	150
151	CCGGATTTAA	GAGCGGAATC	AACTACCAGT	CCACCAACTG	TTGTTACGGG	200
201	TGGAGACCTT	GCCTAGGTTC	AGAGGGCCAG	TATGAATGAT	CTCTAACTCA	250
251	ACCAGTGTCG	AGGAACTGTT	CTCTCGCATT	GACCACAAAA	TTCTATCTCA	300
301	TGTATGCCAA	ACGTGCCTTC	GTGCACTGGT	ATGTTGTTGA	CGGAATGGCG	350
351	GAAAGATCAA	TTCTCCGACG	CTCGTTGAGG	CCTCACAAGT	GCATTTCGCG	400
401	GG					402

Fig. 2A: The DNA sequence of the identified tubulin fragment FATUB. PCR-based screening of the Giza88 PAC library produced a fragment of tubulin related-gene family which was in the molecular size (402 bp).

1	GTGAAGTGTG	TGAGTGATCC	ACAATGGTGC	CTGCCTGGTT	CCATTGTGGT	50
51	CACTGCCACA	AACTTCTGCC	CTCCAGGAGG	CTGGTGTGAC	CCTCCCAATC	100
101	ACCCACTTTG	ATCTCTCTCA	GCCTATTTTT	CAACACATTG	CGCAATACGG	150
151	AGCTGGTATT	GTTCCTGTAA	TGTACAGAAG	GTAAACAAAC	AATGAAAACA	200
201	GCAACATAGT	TTTTCAAAAT	TCATTTTCCA	CCGACACACT	AACTAATAAT	250
251	TATTTTTTAT	CAGGGTAAAG	TGCAGGAGAA	GTGGGGGTAT	CAGGTTCACA	300
301	ATCAATGGAC	ATTCCTACTT	CAATTTAGTA	CTGATCACCA	ATGTTGGAGG	350
351	TGCTGGGGAT	GTACATTCAG	TATCCATTAA	AGGGTCAAAG	ACCAGATGGC	400
401	AAGCAATGTC	AAGGAACTGG	GGCCAGAACT	GGCAAAGCAA	TTCTTACCTT	450
451	AATGGCCAAA	GCCTCTCTTT	CATCGGCACG	ACCAGTAACG	GGACAGAGTG	500
501	TTGTTTCGTT	CGACGTTGCT	CCATCCAGTT	GA		532

Fig. 2B: The DNA sequence of the identified expansin fragment FATEXP3F. PCR-based screening of the Giza88 PAC library produced a fragment of expansin related-gene family which was in the molecular size (532 bp).

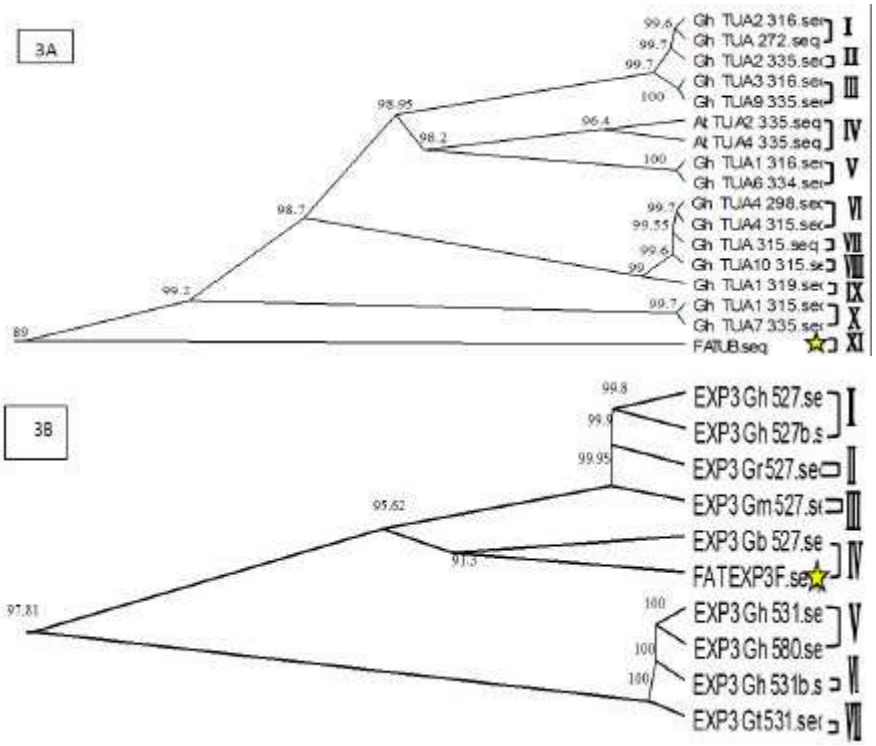

Fig. 3A,B: **A:** Phylogeny tree showing the percenatge of identity of the 16 tubulin sequences identified from *G. hirsutum* and *A. thaliana* against the FATUB sequence that identified from Giza88 (*G. barbadense*).
B: Phylogeny tree summarizing the relationship between the FATEXP3F and the different nine expansin sequences(Table 2a) at the level of genomic DNA sequence. Sequences degenerated from *G. barbadense* were highlighted with a side star.

members of *Gossypium spp*. Even though, *A. thaliana* had contributed, in the alignment, with two sequences; the least query was 68% to 85%, while the maximum identity was 78% to 90%. According to Fig. 3a, the highest similarity 100% in clades III and V was related to *G. hirsutum*. However, tubulins from *A. thaliana* in clade IV were resmblence to each other by 96.4%, while raised to 98.2% against clades I, II, III and V from *G. hirsutum*. The percentage of similarity between *G. hirsutum* tubulins in clades I, II and III was 99.7% decreased to 98.95% in clad V. This indicates that tubulin sequences within the same organism *A. thilian* or *G. hirsutum* was high in similarity while less similar when compared against each other. Precisely, the disperate calde XI of FATUB which is, *G. barbadense*, scoring percentage of identity 89% away from other tubulin's clades of *G. hirsutum* as well as *A. thaliana*. Although the FATUB and GhTUAs, all were descended from the same genus *Gossypium*, the percentage of divergence between all *G. hirsutum* tubulins was ragned 0-0.55. Only when the comparsion include different genus like *A. thaliana*, the divergence raised to range 1.1-1.8. However, when a related species from the same genus was included, *G. barbadense*, divergence range is least to the value (2) with FATUB sequence (Fig. 3a). It means that the inter-species divergence (within the genus) is higher than the intra-species divergence (between geni).

To detect the similarty at the level of DNA sequence between GbACT and the other actins of *G. hirsutum* from the NCBI gene bank, a comaparison study was achieved by Momtaz *et al.* [22] at the level of coding sequences (CDs). Additionally, DNA STAR programe for the mega alignment was used to develop the pedigree. The percentage of similarity at (CDs) level was 93.8%, 97.7% and 97.4% with ACT 1, 2 and 5, respectively. However, the pedigree value was 98.52% with the clades of ACT2 and ACT5, while was 97.77% with ACT1. These results implicated that the sequence of ACT2 is similar and related to those of ACT5 as well as GbACT, meanwhile, all (ACT 2, 5 and GbACT) were far from ACT1 [17].

Table 2A: Nine DNA sequences with highest percentage of identity against FATEXP3F. Each accession number is representing one of the *Gossypium* species. Moreover, it shows the genome type that the sequence is identified from; and the given code that was used in the mega-alignment

Accession #	Plant	Genome	Used code for alignment	
AF512541	*G. hirsutum*	Precursor	EXP3 Gh 580	bp
EF644243	*G. barbadense*	D	EXP3 Gb 527	bp
EF644247	*G. raimondii*	D	EXP3 Gr 527	bp
EF644246	*G. mustelinum*	D	EXP Gm 527	bp
EF644245	*G. hirsutum*	D	EXP3 Gh 527	bp
EF644244	*G. hirsutum*	D	EXP3 Gh 527b	bp
EF644242	*G. tomentosum*	A	EXP3 Gt 531	bp
EF644241	*G. hirsutum*	A	EXP3 Gh 531	bp
EF644239	*G. hirsutum*	A	EXP3 Gh 531b	bp

Table 2B: Different expansin genes. The identified expansin genes with its most related *Gossypium* spp. Sources explanied with genome type and percentage of identity (% ID) detected against the identified FATEXP3F of Giza88 (*G. barbadense*).

G. 88% of ID	Plant	Gene	Genome
99%	*G. barbodense*	EXP3	D
98%	*G. hirsutum*	EXP3	D
	G. mustelinum	EXP3	D
	G. raimondii	EXP3	D
96%	*G. hirsutum*	EXP3	A
	G. barbodense	EXP3	A
	G. tomentosum	EXP3	A
	G. abboreum	EXP3	A
94%	*G. hirsutum*	EXP3	A
	G. barbodense	EXP3	A
	G. tomentosum	EXP3	A
	G. mustelinum	EXP3	A
93%	*G. arboreum*	EXP3	
82%	*G. hirsutum*	EXP2	
80%	*G. hirsutum*	EXP1, EXP2 and EXP8	D
	G. raimondii	EXP1 and EXP2	
79%	*G. hirsutum*	EXP2 and EXP6	D and A
	G. barbodense	EXP1, EXP2 and EXP6	D and A
	G. raimondii	EXP6	
	G. tomentosum	EXP6	A and D
	G. mustelinum	EXP6	A and D
	G. arboreum	EXP6	
78%	*G. hirsutum*	EXP1, EXP2 and EXP15	A and D
	G. tomentosum	EXP1 and EXP2	D
	G. mustelinum	EXP1 and EXP2	A and D
	G. arboreum	EXP1 and EXP2	
77%	*G. hirsutum*	EXP6 and EXP1	A and D
	G. barbodense	EXP6	A and D
	G. raimondii		
	G. tomentosum		
	G. mustelinum		A and D
	G. arboreum		
73%	*G. arboreum*	EXP6	
72%	*G. hirsutum*	EXP6	A
	G. barbodense		A
	G. raimondii		
	G. tomentosum		A and D
	G. mustelinum		D
71%	*G. hirsutum*	EXP6	D
69%	*G. hirsutum*	EXP4	A and D
	G. barbodense		D
	G. raimondii		
	G. tomentosum		D

According to Momtaz *et al.* [23], intervarietal sequence divergence between *G. hirsutum* actin gene families, at DNA level, showed differential range 0.5–3%; however at ORF level, it reached 0.6–2.9%; meanwhile at amino acid sequence level it decreased to 0.0%. On the other hand, percentage of divergence increased noticabley to 9.9% at the amino acid level, when the intravarietal comparsion included *G. barbadene*; while at the ORF level it ranged 11.9–13.6% and it was 1.0–2.9% at DNA sequence level. All these data confirmed that the percenatge of molecular sequence divergence (MSD) within *G. Hirsutum* sequences were less than MSD between *G. hirsutum* and *G. barbadense*.

DNA sequence of the identified fiber structural and quality expansin gene family (FATEXP3) (532 bp) (Fig. 2b) took place using the NCBI Blastn program. The sequence was aligned used as a limited search to record matching entrez query: *Gossypium* [ORGN]. The alignment showed 102 hits. While the lower identity was 69% with the *G. hirsutum* α-expansin 4 (EXP A4); the maximum identity was 99% with *G. barbadenes* α-expanisn 3 (EXP A3). Nine expansins gene sequences were chosen for more specific analysis against FATEXP3F sequence (Table 2a). Mega-alignment using Lasergene DNASTAR program was performed using Clustal method which produced a pedigree (Fig. 3b). Since FATEXP3F was identified from the cultivated (Giza88) *G. barbadense* and based on the percent of identity as well as the pedigree in Fig. 3b; FATEXP3F was 91.3% more related to the sequence of EXP3 Gb527 (*G. barbadense*) at clad IV, while it was 95.62% at clades I, II and III (all not less than 99.8%). On the other hand, the percent of identity between the other clades V, VI and VII was 100.0% while when compared to the cluster of FATEXP3F the percent of identity lowered to reach 97.81% (Fig. 3b). This indicates the difference on the DNA sequence level, of Giza88 FATEXP3F from the related expansin genes of upland cotton.

Expansin sequences of *G. hirsutum* were high in similarity 100.0% and inter-related to *G. tomentosum* (Fig. 3b, clad VII). Although, they were descended from different ancestors represented by three lineages with different clades, hitherto it gives a pitch that they may have some common ancient ancestors. This proposal is advocated the statement of Cosgrove *et al.* [23] who deduced the high percentage of similarity in the expansin gene family and its historical evolution. Moreover, vacant data out from the phylogeny tree indicated that the evolutional aspect of expansin lineage

from *G. barbadense* (IV) is just differing from the other lineages of other expansins from the other *Gossypium* SPP. (Fig. 3b; clades I, II and III). As indicated by Cosgrove *et al.* [23], from a genomic comparative evolutionary vision, we may indicate that genome type A and D is playing a role in the new classification of expansin gene family evolution that further was emphasized by Table 2b and Figure 3b. The three expansins (EXPs) from genome type A were almost identical (V, VI and VII); in contrary to those from genome type D which was less identical (Table 2a and Fig. 3b). As if, genomes A and D were descended from common ancestral clades in the ancient ages. Nevertheless, the high percentage of identity between EXPs of the tetraploid *G. hirsutum* and the diploid *G. tomentosum* is, probably, because they have shared or decanted from the same genome type A, or, at least, from one ancestor of *Gossypium* evolution.

Furthermore, it gives a plausible example about less stringency along with less conservation that genome type (D) may endure within its sequences. Through the evolutionary stages, those sequences were hard to raise its potential for being distinctive (heterogeneous) in comparison to the high similarity in the same gene family (expansin) in genome type A. Accordingly, Cosgrove *et al.* [23], studied EXPs sequence evolution in *A. thaliana* and rice while our proposition states that the new-world genome type D was the later in the expansin evolution, in contrary to that of the old-world type A. This may be due to its high divergence and less homogeneity at the expansin sequence level as it shown in Figure 3a. Many fiber quantitative trait loci (QTL) have been identified in the D sub-genome of tetraploid cotton, suggesting that the D-genome contains important genes or regulators of fiber morphogenesis and fiber properties [24].

Identifying, the Egyptian cotton Giza88 variety, fiber quality related gene families for its high superior fiber traits in intra-varietal manner is a laborious task for Egyptian cotton. As a first step, defining pre-, post-anthesis regulating genes as well as investigating its structural and developmental role in comparative analysis relationships with the upland cotton illustrates evolutional phylogenic relationship between *G. barbadense* and other *Gossypium spp*. This is a pioneer step to study fiber QTL tags for preservation and improvement of fiber quality as marker assistance selection procedures for Egyptian cotton breeding programs.

REFERENCES

1. Jones, M.E. and G. Keller, 1996. Metabolic pathway engineering in cotton: Biosynthesis of polyhydroxybutyrate in fiber cells. Proc. Natl. Acad. Sci. USA, 93/23: 12768-73.
2. Humphries, J.A., A.R. Walker, J.N. Timmis and S.J. Orford, 2005. Two WD-repeat genes from cotton are functional homologues of the *Arabidopsis thaliana* Transparent Testa Glabra1 (TTG1) gene. Plant Mol. Biol., 57: 67-81.
3. Li, X.B., F. Xiao-Ping, W. Xiu-Lan, C. Lin and Y. Wei-Cai, 2005. The cotton *ACTIN*1 gene is functionally expressed in fibers and participates in fiber elongation.The Plant Cell, 17: 859-75.
4. Ji, S.J., Y.C. Lu, J.X. Feng, G. Wei *et al*. 2003. Isolation and analyses of gene preferentially expressed during early cotton fiber development by subtractive PCR and cDNA array. Nucl. Acids Res., 31: 2534-43.
5. Li, X.B., L. Cai, N.H. Cheng and J.W. Liu, 2002. Molecular characterization of the cotton *GhTUB1* gene that is preferentially expressed in fiber. Plant Phys., 130: 666-74.
6. Shimizu, Y., S. Aotsuka, O. Hasegawa, T. Kawada, T. Sakuno *et al.*, 1997. Changes in levels of mRNAs for cell wall-related enzymes in growing cotton fiber cells. Plant Cell Phys., 38: 375-78.
7. Mathur, J., N. Mathur, V. Kirik, B. Kernebeck *et al.*, 2003. Arabidopsis CROOKED encodes for the smallest subunit of the ARP2/3 complex and controls cell shape by region specific fine F-actin formation. Development, 130: 3137-46.
8. Cosgrove, D.J., 2000. Loosening of plant cell walls by expansins. Nature, 407: 321-26.
9. Cosgrove, D.J., L.C. Li, H.T. Cho, S. Hoffmann-Benning *et al.*, 2002. The growing world of expansins. Plant Cell Phys., 43: 1436-44.
10. Li, Y., L. Jones and S. McQueen-Mason, 2003. Expansins and cell growth. Curr. Opin. Plant Biol., 6: 603-10.
11. Wolfgang, J., A. Baur, P. Nick, R. Reski and G. Gorr, 2004. A large plant β-tubulin family with minimal C-terminal variation but differences in gene expression. Gene, 340: 151-60.
12. Whittaker, D.J. and B.A. Triplett, 1999. Gene-specific changes in α-tubulin transcript accumulation in developing cotton fibers. Plant Phys., 121: 181-88.

13. Arpat, A.B., M. Waugh, J.P. Sullivan, M. Gonzales *et al.*, 2004. Functional genomics of cell elongation in developing cotton fibers. Plant Mol. Biol., 54: 911-29.
14. Frengen, E., D. Weichenhan, B. Zhao, K. Osoegawa *et al.*, 2000. Modular bacterial artificial chromosome vectors for transfer of large inserts into mammalian cells. Genomics, 58: 118-26.
15. Zhang, H.B., 2000. Construction and manipulation of large-insert bacterial clone libraries- manual. Texas A&M University, College Station. Available via http://hbz.tamu.edu.
16. Peterson, D.G., J.P. Tomkins, D.A. Frisch, R.A. Wing and A.H. Paterson, 2000. Construction of plant bacterial artificial chromosome libraries: An illustrated guide. J. Agri. Genomics, http://www.cabi-publishing.org/jag/index.html.
17. Momtaz, O.A., A.E.A. El-Doliefy, A.H. Fahmy and A.A. Diab, 2007. Comparative sequence analysis of actin related gene family isolated from *Gossypium barbadense*. World J. Agri. Sci., 3/1:130-39.
18. Luo, Z.W., C.A. Hackett, J.E. Bradshaw, J.W. Mcnicol and D. Milbourne, 2001. Construction of a genetic linkage map in tetraploid species using molecular markers. Genetics, 157: 1369-85.
19. Zhang, H.B., X.P. Zhao, X.L. Ding, A.H. Paterson and R.A. Wing, 1995. Preparation of mega-base DNA from plant nuclei. Plant J., 7: 175-84.
20. Momtaz, O.A., M.I. Shehata, S. El-Khodary and H.Z. Hassan, 2006. Idntification and isolation of *smHSPs* gene family from *Gossypium barbadense* L.c.v. Giza90 Egyptian cotton variety. Arab. J. Biotech., 9/1: 203-15.
21. Xu, Z., S. Sun, L. Covaleda, K. Ding, A. Zhang *et al*. 2004. Genome physical mapping with large-insert bacterial clones by fingerprinting analysis: Methodlogy, resource clone genome coverage and contig map quality. Genomics, 84: 941-51.
22. Momtaz, O.A., A.E.A. El Doliefy, A. Bahieldin and E.H. Hassanein, 2007a. Identification of actin related gene family from *Gossypium barbadense* Giza88 cultivar using PCR-based positional cloning. Arab. J. Biotech., 1: 27-40.
23. Cosgrove, D.J., R.E. Carey and J. Sampedro, 2006. Genome histories clarify evolution of the expansin superfamily: new insights from the poplar genome and pine ESTs. J. Plant. Res., 119: 11-21.
24. Guo, W., C. Cai, C. Wang and Z. Han *et al.* 2007. A microsatellite-based, gene-rich linkage map reveals genome structure, function and evolution in *Gossypium*. Genetics, 176: 527-41.