Plant Tissue Cult. & Biotech. **15**(2): 145-156, 2005 (December)

PTC&B

Preliminary Progress in Jute (*Corchorus* species) Genome Analysis

Ahmad S. Islam, Matthew Taliaferro, Christopher T. Lee, Craig Ingram, Rebecca J. Montalvo, Gerrit van der Ende, Shahabudin Alam[1], Javed Siddiqui[1] and Kanagasabapathi Sathasivan*

Molecular Cell and Developmental Biology, School of Biological Sciences, The University of Texas at Austin, USA

Key words: Genomic DNA, *Corchorus olitorius, C. capsularis*, WU-BLAST, TAIR, 18S rRNA, tRNA-Leu

Abstract

The paper summarizes the progress made in cloning and sequencing a limited number of genes from jute (*Corchorus olitorius* and *C. capsularis*) and discusses future applications of jute genome analysis. As of December 2005, slightly over 200 DNA sequences have been deposited in GenBank. Although many of these sequences are partial and uncharacterized, this marks the beginning of a major step in unraveling of the hitherto unknown jute genome. We have constructed both cDNA and genomic DNA libraries for *C. olitorius* var. O4 and *C. capsularis* var. CVL-1 in the plasmid vectors pSMART and pBluescript, respectively. Random clones were isolated and sequenced. These DNA sequences have been deposited in GenBank and analyzed using TAIR (www.arabidopsis.org) for similar sequences in *Arabidopsis thaliana* and other related plant species. The complete sequence for tRNA-Leu and partial sequence of DNA fragments encoding several proteins, such as RNA polymerase β subunit-1, 18S rRNA, mitochondrial DNA directed RNA polymerase and carboxytransferase β-subunit are reported in this paper. The analysis of DNA sequences of related taxa deposited in GenBank is also presented delineating the scope and applications of cloning genes of agronomic importance.

Introduction

Jute, the world's second most cultivated fiber crop next to cotton, is extensively grown in India, Bangladesh, China, Thailand, Russia, Myanmar, Nepal, Uzbekistan, Chile and Brazil (http://faostat.fao.org/). India and Bangladesh rank as the top two countries in terms of jute production. In Bangladesh, jute is the principal cash crop and jute products contribute a significant portion to

*Corresponding author (sata@mail.utexas.edu); [1]DNA Technologies, Gaithersburg, Maryland.

country's foreign exchange (http://www.bangladeshgov.org/jute/). Jute cultivation helps sustain millions of farmers in these two countries alone. The cultivated varieties of jute have been evolved from *Corchorus olitorius* L. and *C. capsularis* L. through conventional breeding and pure line selection (Ghosh 1983) based on their yield and agronomic performance. Synthetic fibers have been a major competition in the international market to the natural jute fibers. In recent years the situation has improved because natural fibers do not pollute the atmosphere as do synthetics and may help reduce the cutting of trees for making paper. In addition, jute is now being utilized to manufacture more value-added industrial products such as in the making of Geo-textiles (http://www.jute.com/geojute.html) for protecting embankments against river erosion, fiber reinforced building materials, packaging materials and in the production of paper. However, there are problems in increasing the productivity and profitability of jute. Some of the major challenges include susceptibility of the jute crop to insect pests and fungal diseases, photoperiod sensitivity, poor fiber quality, and low yield under unfavorable growth conditions such as salinity, drought, flood or cold. Addressing these challenges through traditional plant breeding program has limitations due to lack of genetic diversity among cultivated jute varieties and sexual incompatibilities between the cultivars of the two jute species and between each of the two species and wild *Corchorus* species. In spite of developing successful hybrids between two species of jute namely, *C. olitorius* and *C. capsularis* (Islam and Rashid 1960), it was not possible to release any variety from the advanced progeny of the above interspecific crosses. Hence, in recent years molecular approaches to improve the agronomic traits of jute are being considered as alternatives.

Recently, research in some universities in the Indian subcontinent has resorted to molecular approaches through systematic cloning and transgenic work. By means of a combined study of AFLP and RAPD, Hossain et al. (2002, 2003) at Dhaka University have shown the importance of using molecular markers in distinguishing between cold-tolerant and cold sensitive jute varieties obtained from GenBank at Bangladesh Jute Research Institute (BJRI). Using simple sequence repeat (SSR) marker loci and AFLP assay, Basu et al. (2004) at the Indian Institute of Technology (IIT), Kharagpur, evaluated genetic diversity of 49 genotypes of the two jute species. More recently, P. K. Gupta at Charan Singh University, Meerat (CSUM), India and his associates as well as the Dhaka University team led by Haseena Khan developed genomic SSRs and deposited the sequences in GenBank. By developing more SSRs Gupta and his team at CSUM are planning to embark upon a program of gene tagging combined with the construction of a framework linkage map for QTL interval mapping. In addition to the research on markers, procedure for successful regeneration (Seraj

et al. 1992; Saha et al. 1999) and transformation of *C. capsularis* (Ghosh et al. 2002) have been developed.

At the University of Texas at Austin, work on the construction of cDNA and genomic DNA library of the above two jute species has been initiated recently. As a first step towards achieving this goal, a rapid method for high quality RNA isolation from jute was developed in this lab (Khan et al. 2004). A cDNA and genomic DNA libraries have been constructed in pSMART and pBluescript vectors from *C. olitorius* and *C. capsularis,* respectively. We have isolated and sequenced random fragments with the objective of testing these DNA libraries and deposited some of the sequences in GenBank. This paper reports analysis of the DNA sequences that show homology with other plant genes including that of *Arabidopsis thaliana*. It summarizes the results from the analysis of selected DNA sequences currently available in GenBank. We present the challenges encountered in cloning genes from jute and suggest possible solutions to overcome such problems. In addition, this paper attempts to update jute researchers on the current status in jute molecular biology and to establish liaison with any other team(s) who may be interested to collaborate in the jute genome project.

Material and Methods

Plant materials and growth: Seeds of *C. capsularis* var. CVL-1 and of *C. olitorius* var. O4 were supplied by Bangladesh Jute Research Institute through the courtesy of Haseena Khan, Dhaka University. Seeds from *C. olitorius* were surface sterilized with 10 % bleach and incubated on a moist filter paper in Petri dishes inside an incubator at room temperature (23°C). Seven-day-old seedlings grown inside Petri dishes were collected and shipped to DNA technologies, Maryland for RNA isolation and cDNA library construction. The sample weight was roughly 1g. Seeds of *C. capsularis* were surface sterilized in the same manner and planted in the greenhouse under normal lighting and allowed to grow to mature plants. The young leaves of mature plants were harvested, frozen in liquid nitrogen and shipped to DNA technologies, Maryland for genomic DNA isolation and library construction.

Isolation of the total RNA: The frozen jute tissues were directly grinded in the RNA isolation solution (Trizol-Invitrogen) in liquid nitrogen. The lysates were centrifuged at 10,000 × g for 10 min to remove unlysed cells and contaminants. The supernatant was mixed with 0.2 volume of chloroform and incubated on ice for 15 min. The mixture was centrifuged at 10,000 × g for 10 min and the upper layer was collected. The upper layer was mixed with half volume of isopropanol and centrifuged for 30 min at 10,000 × g. The pellet was washed with 70% ethanol

and air dried. The RNA was dissolved in DEPC treated autoclaved water. An aliquot of RNA was checked on denaturing agarose gel.

Isolation of the mRNA: Jute mRNA was isolated from total RNA using oligo dT cellulose. Briefly, the total RNA was mixed with 10 ml of binding buffer (10 mM Tris pH 7.5, 500 mM NaCl) and heated at 70°C for 5 min. The samples were immediately chilled on ice for 5 min and mixed with oligo dT latex beads. The mixture was incubated at room temperature for 2 h to allow the complete binding of the mRNA with oligo dT. After incubation, the mixture was centrifuged at 5,000 ×g. The pellet was washed twice with the binding buffer. Now, the pellet was washed with excess of a low salt buffer (10 mM Tris pH 7.5, 250 mM NaCl). This step was repeated several times to remove all unbound RNA molecules. The mRNA was finally eluted with 10 mM Tris pH 8.0. An aliquot of mRNA sample was checked on denaturing agarose gel.

Construction of the cDNA library: Two μg (micrograms) of mRNA were mixed with oligo dT primer (18 mer) and heated at 65°C for 10 min. The mixture was cooled on ice and the following cDNA synthesis reagents were added: dNTP, RNase inhibitor, first strand cDNA synthesis buffer and superscript reverse transcriptase. The reaction mixture was incubated at 42°C for 2.5 h. The second strands were synthesized using dNTP mix, *E. coli* DNA polymerase and RNase H in second strand buffer at 16°C for 2h. After the second strand synthesis, the ends were polished using Pfu polymerase at 72°C for 30 min. The cDNA was extracted with phenol: chloroform: isoamyl alcohol.

EcoRI adapter ligation, kinasing of cDNA ends and size fractionation of the cDNA: The end-polished cDNA was mixed with EcoRI adapter, T4 DNA ligase buffer and a high concentration of T4 DNA ligase. The mixture was incubated at 8°C for 48h. The mixture was heated at 50° C for 5 min. The ends of cDNA were phosphorylated (addition of 5'-phosphate) by T4 kinase. The reaction was continued at 37° C for 1 h. The kinased cDNA was precipitated with ethanol and 3M sodium acetate. The adapter ligated cDNA was re-suspended in 10 mM TE buffer and electrophoresed on 1% low melting point agarose gel for 3 h. The cDNA in the size range of 0.5 kb and higher was recovered from low melting point agarose by extraction with phenol; phenol chloroform and chloroform. The cDNA was precipitated with ethanol.

Vector digestion, purification and the ligation of cDNA: For making the library, pSMART vector from Lucigen was used. The sequences and the restriction map of the vector are available in www.lucigen.com. The vector was digested with EcoRI and treated with calf intestinal alkaline phosphatase (CIAP). The digested vector was purified on the agarose gel. An aliquot of digested vector was ligated and used to electroporate *E. coli* DH10B cells to test the efficiency of the digestion. After getting successful results, the cDNA was separately ligated into

the vector in the presence of T4 DNA ligase at various concentrations. The samples were incubated at 8°C for 48 h. Ligated cDNA was precipitated with ethanol and re-suspended in water. An aliquot of the ligated samples were electroporated in *E. coli* DH10B cells. Immediately after electroporation, the cells were mixed with SOC medium and allowed to recover for 1.5 h at 37°C. Libraries were centrifuged at 5000 rpm for 10 min. The medium was discarded and the pellet was re-suspended in SOC medium containing the appropriate antibiotic and 15% glycerol. The cells were plated on LB ampicillin plates and grown at 37°C overnight. DNA from single colonies was isolated and screened for the presence of inserts.

Genomic DNA library construction: Pre-chilled (– 80°C) sterile grinder and pestle were used to grind young jute (*C. olitorius* var. O-4) leaves until the leaf became powdery. Five ml of lysis buffer-I containing proteinase K were added to the powdered leaves and the mixture was incubated at 37°C in a shaker overnight with a low-speed agitation. The incubated material was spun at 10,000 rpm for 15 min at 4°C and the supernatant was collected. After adding 5 ml phenol : chloroform : isoamyl alcohol (PCI) to the supernatant, it was vortexed and spun at 10,000 rpm for 5 min. The aqueous phase was saved and 3M sodium acetate (0.3M final) was added to it. Nearly 2.5 volumes of 100% ethanol (cold) were added to the aqueous phase; the material was well mixed and stored at –20°C for 30 min. The re-suspended pellet was spun down after 1 ml of TE was added to it. Genomic DNA was then isolated by Easy-DNA Isolation Kit (Invitrogen, Carlsbad, CA; (cf. manufacturer's protocol). In short, 350 µl of solution A were added to the DNA and incubated at 65°C for 10 min. Thereafter, 150 µl of solution B were added and the mixture was vortexed vigorously until the precipitation was found to move freely. After adding 500 µl chloroform to the previous mixture, it was vortexed until the viscosity of the material was found to decrease. The next step was centrifugation at 14,000 rpm for 20 min. To the upper transferred aqueous phase, 1 ml of cold 100% ethanol was added, mixed well and left at – 20°C for 30 min. The centrifugation at 14,000 rpm for 20 min was repeated. To the transferred upper aqueous phase, 500 µl of cold 80% ethanol were added and spun at 14,000 rpm for 5 min. The pellet was re-suspended in 100 µl TE containing RNase. The last step was incubation at 37°C for 30 min and the incubated product was stored at 4°C. The quality of the isolated genomic DNA (gDNA) was checked by running the samples on a 0.7% agarose gel and taking OD readings at 260 nm.

Restriction enzyme digestion of gDNA and size fractionation: Genomic DNA was partially digested with restriction enzyme Apo I and Sma I (Roche Diagnostics, Indianapolis, IN) in separate tubes for ligation with their specific vectors. Approximately, an amount of 10 µg of gDNA was taken for restriction enzyme

digestion for 90 min; about one-third of the partially digested gDNA was removed from the digestion reaction and mixed with 10 µl of 0.5 M EDTA every 30 min, followed by their storage at 4°C. The three restriction enzyme-digested DNA fractions were combined. The next step was to purify the digested material using Qiagen MinElute Gel extraction Kit (cf. manufacturer's protocol).

Vector preparation: Seven µg of pBluescript SK (+/-) vector (www.Stratagene.com) was digested twice with EcoRI and SmaI (Roche Diagnostics, Indianapolis, IN) in separate tubes for the ligation with their specific inserts. Briefly, the DNA was digested for 2.5 h at the specific temperature; thereafter more enzymes were added and the digestion period was prolonged for another 2 h. The digested DNA was inactivated by heating at 70°C for 10 min and treated with Shrimp alkaline phosphatase, (SAP; Roche Diagnostics, Indianapolis, IN) after it had cooled down to room temperature, (cf. the manufacturer's protocol). Restriction-enzyme digested and SAP-treated vector DNA was then purified with PCI followed by treatment with chloroform: isoamyl alcohol (CIA), and finally precipitated with ethanol. The DNA was then re-digested with the same set of enzymes and the same time frame as above followed by SAP treatment, purified with PCI followed by CIA and precipitated with ethanol. The vector DNA was re-suspended with an appropriate amount of sterile H_2O/TE. The quality check (QC) of the isolated vectors was performed by running the samples on a 0.7% agarose gel.

Ligation and electroporation: The ApoI digested gDNA was ligated with the pBluescript SK (+/-) vector which was earlier digested with EcoRI and the Sma I digested gDNA was ligated with the pBluescript SK (+/-) vector which was earlier digested with Sma I, respectively. The ligation reaction was made for 10 µl and incubated for two days at 4°C. Two µl of the ligated reaction mixture were added to 70 µl of Electro-10 Blue competent cells (thawed at 4°C) and electroporated in a 0.1 cm gap pre-chilled cuvette at 1700 volts. One ml of SOC was added immediately. Cells were transferred to a new 15 ml culture tube, incubated at 37°C for 1 h and 100 µl of cells were spread on LB-agar plates containing ampicillin. The plates were incubated at 37°C overnight. The number of colonies was counted.

Quality control (QC) of the library: Ten colonies were randomly picked up and grown in LB (+ antibiotic) overnight. QC was performed by restriction enzyme digestion (Fig. 1) or PCR methods. Once the pilot ligations were found effective (> 90% recombinants), large-scale ligations were set up to get at least 10^6 primary clones. The whole library was recovered in 10 ml of SOC at 37°C. One hundred µl of the library were plated on LB-plates (with ampicillin). The plates were incubated overnight. On the following day, the number of colonies was counted. The library was centrifuged and re-suspended in 2 ml of LB-antibiotic. Approxi-

mately, 40 - 50,000 colonies were spread on each 150 mm plate. The plates were incubated overnight. The last step consisted of aliquoting the libraries in 1 ml aliquot, labeling each container and storing them at –70°C.

Isolation of plasmid DNA: The plasmids containing either cDNA or gDNA cloned materials were isolated via alkaline lysis procedure through Qiagen spin columns as per manufacturer's instructions. Once these plasmids were isolated, the quantity and quality of the samples were determined based on spectrophotometric (NanoDrop) analysis. In order to identify which plasmids possibly contained cDNA inserts, we digested each plasmid sample with EcoRV and EcoRI restriction enzymes and then separated the results of the digest on a 0.8% agarose gel stained with ethidium bromide, SyBr Green or SyBr Gold. The recombinant plasmids were identified and sent for sequencing at the core DNA sequencing facility at The University of Texas at Austin. Upon obtaining positive sequencing results, a BLAST nucleotide search was performed on each sample in order to determine to which genes the fragments were most closely related. In addition, we have confirmed the sequence alignment results with the WU-BLAST service in TAIR (www.arabidopsis.org) and the results have been recorded.

Results and Discussion

The cDNA library constructed in pSMART cloning vector (www.lucigen.com) into the NotI/Blunt restriction site and stably maintained in the *E. coli* host cell DH10B-T1 contained 10^6 primary clones comprising about 90% recombinants. The average insert size as determined by PCR was 100 to 500 bp. The inserts were relatively small ranging from 100 to 500 bp. The short strands of cDNA could possibly be due to some inhibitory compound that might have co-purified

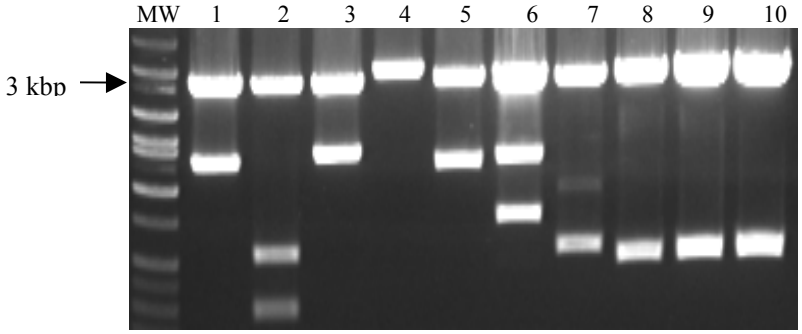

Fig. 1. Results from gDNA library; ten randomly picked clones digested with Pvu II indicating the presence of DNA inserts in nine out of ten samples.

with the jute RNA during the process of isolation. Special attention to prevent such contamination may help in the future cDNA synthesis and library

construction. Random cDNA library clones were cultured on LB kanamycin and the DNA isolated was sequenced.

The total genomic DNA of *C. capsularis* was used in the genomic DNA library constructed in pBluescript KS (www.stratagene.com). The genomic library contained 10^6 to 10^7 primary clones and the average insert size was about 100 to 1000 bp. Random clones were cultured in LB ampicillin and the plasmid DNA isolated was sent for sequencing. The cDNA and gDNA sequences were used to search for similar sequences through the TAIR (www.arabidopsis.org) web site through WU-BLAST (Wisconsin University - BLAST). The sequence analysis from randomly selected clones is summarized in the Table 1.

Table 1. Examples of the DNA sequences obtained from the jute genomic DNA and cDNA sequences. Repetitive gene names are mentioned here to illustrate the frequency of such clones in the library.

Seq. No.	TAIR result	Score	E value
7	AT1G68990. DNA-directed RNA polymerase, mitochondrial	96	2.00E-18
13	AT3G41768. 18SrRNA	287	2.00E-76
14	AT3G41768. 18SrRNA	344	4.00E-93
5	ATCG00180. RNA polymerase β subunit-1	1068	0
4	ATCG00180. RNA polymerase β subunit-1	337	2.00E-91
9	ATCG00180. RNA polymerase β' subunit-1	1207	0
11	ATCG00180. RNA polymerase β' subunit-1	371	1.00E-101
15	ATCG00490. large subunit of RUBISCO.	492	1.00E-138
3	ATCG00500. carboxytranserase β subunit	151	3.00E-35
2	ATCG00860. hypothetical protein	101	3.00E-20
8	ATCG00860. hypothetical protein	898	0
10	ATCG00905. chloroplast gene encoding ribosomal protein s12	295	8.00E-79
12	ATCG01180. chloroplast-encoded 23S ribosomal RNA	315	9.00E-85
6	ATCG01280. hypothetical protein	870	0
1	tRNA-Leu	161	3.00E-38

The information in the above Table was compiled from all sequences collected from June to November 2005. All sequences were compared against known sequences in the TAIR database (www.arabidopsis.org/Blast) with introns and untranslated regions included (AGI Genes (+introns. +UTRs) (DNA)). Any entered sequence that gave homologous sequences with a maximum score of 50 bits or less was discarded. Thus, of 52 sequences, 37 were discarded. Those with maximum scores over 50 bits were analyzed further. The scores and error values for the homologous sequences were recorded. The highest score for each input sequence was noted. The probable location of the genes of interest, their placement in the chromosome, and the specific protein encoded by each individual gene were also taken note of. Analysis of the

sequences revealed the partial DNA sequence of three hypothetical proteins, RNA polymerase β subunit-1, 18S rRNA, mitochondrial DNA directed RNA polymerase, carboxytransferase β subunit, and tRNA-Leu. As mentioned above a known partial base sequence of putative phosphate transport ATP-binding protein gene, partial cds was determined in *C. capsularis* var. CVL-1.

The early submission to GenBank (http://www.ncbi.nih.gov/index.html) of the DNA base sequences of *Corchorus* species was of *C. olitorius* from Harvard University Herbaria, MA, USA (Alverson et al. 1999). In the course of the next four years, this group (Whitlock et al. 2003) working at University of Massachusetts, Amherst published partial cDNA base sequences of NADH dehydrogenase *(ndhF)* gene of a number of wild *Corchorus* species namely, *C. sidoides, C. bricchettii, C. argutus, C. siliquosus* as well as fiber-yielding, species, *C. olitorius* and *C. capsularis.* On the basis of similarity of base sequence of *ndhF* gene, Whitlock et al. (2003) suggested that *Oceanopapaver neocaledonicum* be transferred to *Corchorus.*

Another two important submissions to GenBank were by Basu et al. (2003) at the Indian Institute of Technology, Kharagpur. Working with C. *capsularis*, they determined complete cDNA sequence of caffeoyl-CoA-O-methyltransferase and cinnamyl alcohol dehydrogenase which are two of the three genes involved in lignin biosynthesis. Recently, Liu et al. (2005) at the Chinese Medical University, Taiwan deposited at the GenBank, base sequence of 18S ribosomal RNA gene of *C. olitorius* and *C. capsularis as* well as that of *C. aestuans* var. *brevicaulis*, a wild *Corchorus* species from Taiwan. A phylogenetic tree was constructed comprising several species of *Corchorus* and *Gossypium robinsonii*. Nucleotide sequences of 18s rRNA from the four plants were compared using a clustalw multiple alignment (http://align.genome.jp) by uploading a Fasta file of the nucleotide sequences into the clustalw site. After the alignment a phylogenetic tree was constructed using the clustalw site from the alignment score page. The NJ tree format was used.

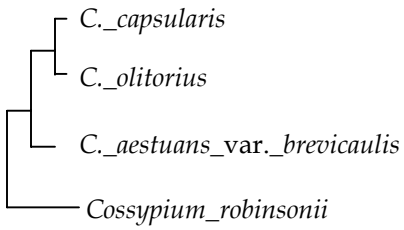

Fig. 2. Phylogenetic tree of selected *Corchorus* species and *Gossypium robinsonii* based on 18S rRNA gene sequences.

The phylogenetic tree shows that *Gossypium* and *Corchorus* diverged from a common ancestor and that all species of the *Corchorus* are more closely related to

each other than to *Gossypium*. Furthermore, the tree shows that *C. capsularis* and *C. olitorius* are more closely related than each of these two species to *C. aestuans*. Thus the molecular data presented in this paper are in agreement with the current taxonomic classification and the crossing relationships between different *Corchorus* species as reported by the first author (Islam 1958). This phylogenetic tree based on 18S rRNA sequences needs to be compared taking into consideration of similar molecular characters before a consensus tree is generated.

The joint team led by P.K. Gupta at Charan Singh University, Meerat, India and Haseena Khan at Dhaka University, Bangladesh submitted to GenBank DNA base sequences of SSR markers in 195 accessions of *C. olitorius*. These markers were used to distinguish accessions that were cold-tolerant from susceptible ones. The SSR sequences submitted to Genbank (http://www.ncbi.nlm.nih.gov/BLAST/) by Sharma et al (2005) were retrieved and entered in TAIR website (www.arabidopsis.org/) to check for homology to sequences in other plant species, especially the genome of *Arabidopsis thaliana*. The WU-BLAST search was set to include introns and UTRs (AGI Genes (+introns, +UTRs) (DNA)). Sequences that gave significant homology were recorded. In some instances, a sequence showed homology to multiple sequences in other plant taxa, in which case the sequence with the greatest score was selected and recorded. In the case of multiple sequences having the same score, the one with the lowest E value was selected and recorded. Any sequence that returned matches with a maximum score of 50 bits or less was recorded as "No Hits."

Table 2. Summary of the sequence analysis of the jute SSR sequences by Sharma et al. (2005) posted in the GenBank.

Genes	Frequency	Score	E value
Nuclear	79	197 to 2074	0.79 to 2.7E-88
Mitochondrial	6	308 to 2310	6.9E-12 to 2.6E-122
Choloroplast	3	596 to 1550	3.1E-21 to 7.6E-65
Unidentified	107		

Of the 195 sequences analyzed, 79 showed homology to nuclear DNA with scores ranging from 197 to 2074, six showed homology to mitochondrial DNA with scores ranging from 308 to 2310, three showed homology to chloroplast DNA with scores ranging from 596 to 1550, and 107 were unidentified.

The DNA libraries were made using mRNA extracted from leaves of young seedlings of the two jute species. So far the results submitted to GenBank are partial cDNA sequences of putative phosphate transport ATP-binding protein gene of *C. capsularis* var. CVL-1 (Accession No. DQ151661) and partial genomic base sequences of 18S ribosomal RNA gene of *C. olitorius* var. O4 (Accession No. DQ151662). Study of the randomly selected genomic DNA sequences of inserts

indicated that they represent partial sequences of the following: RNA polymerase β subunit, chloroplast-encoded 23S ribosomal, chloroplast gene encoding ribosomal s-12, DNA-directed RNA polymerase, mitochondrial carboxytransferase β subunit and three hypothetical proteins. These sequences will soon be submitted to GenBank.

There should be a note of caution in deciphering genomic base sequences of *Corchorus.* The record of nucleotide sequences in *Corchrous* in the GenBank show that the majority of authors have not given names of the cultivar associated with the two fiber-yielding species. It is an important omission because a large amount of differences exist between cultivars of the two species of *Corchorus* in both morphological and physiological traits that are bound to be reflected in their DNA profiles.

The ultimate objective of the present study was to construct a complete cDNA and genomic DNA libraries of both *C. olitorius* and *C. capsularis* so that the genes of interest can be cloned and used to transform jute cultivars for obtaining value-added industrial products. Since this enormous task cannot be undertaken without adequate funding and collaboration with institutions willing to work on a tropical crop, we have been studying the partial cDNA and genomic DNA libraries on a limited scale. So far the genes we have isolated are either chloroplast-, ribosomal- or mitochondrial genes or the enzymes that reside in these structures. The main reason for obtaining genes from the above organelles alone is that DNA has so far been isolated from young leaves that contain numerous chloroplasts and mitochondria, and such genes are more abundant than others. No efforts were made so far to select unique DNA fragments. Future analysis will involve screening of the current library with probes of these abundantly expressed genes to avoid repeated sequencing. Also, many of the sequences are partial and full length sequences need to be isolated by cloning the cDNA ends. Further study aims at collecting suitable material from other parts of the plant such as young stem, bark tissue and roots to clone genes controlling many important traits of agronomic importance including those that confer resistance to insects, fungal and viral pathogens, and those for regulating lignin biosynthesis. Both random sequencing and selective cloning strategies will be employed. In addition, tissue specific promoters will be also be cloned for potential genetic improvement. Since the entire jute genome analysis is a major effort, an international collaboration among the scientists from India, Bangladesh, USA and other countries is essential to help accomplish this task. The clones and sequences should be made available for the genetic improvement of jute to ultimately benefit the millions of jute farmers in the long run.

Acknowledgment

The authors thank Professor Haseena Khan and Samuil Haque of Dhaka University, Bangladesh for providing jute seeds with proper phytosanitary certificate. They thank Ray Neubauer, Thomas Garrad, Mohamad Wazni, Lindsay Stone and Nabila Anwar who helped them with the plasmid DNA

isolation and laboratory work. Sincere thanks to Dena Sutton and her associates for helping with the growing of jute plants in the greenhouse of The University of Texas at Austin. Finally they express their gratitude to the University CoOp and the University of Texas at Austin for supporting the undergraduate research fellowship program without which the present research could not have been undertaken.

References

Alverson WS, Whitlock BA, Nyffeler R, Bayer C and **Baum DA** (1999) Phylogeny of the core Malvales: evidence from ndhF sequence data. Am. J. Bot. **86**(10), 1459-1471 (1999)

Basu A, Maiti MK and **Sen SK** (2003) Cloning and sequencing of mRNA for caffeoyl-CoA-o-methyltransferase from *Corchorus capsularis*. Direct submission to GenBank. [Accession No. AY500159]

Basu A, Maiti MK and **Sen SK** (2003) Nucleotide sequence of cinnamyl alcohol cDNA isolated from *Corchorus capsularis.* Direct submission to GenBank [Accesssion No. AY504425]

Basu A, Ghosh M, Meyer R, Powell W, Basak SL and **Sen SK** (2004) Analysis of Genetic Diversity in Cultivated Jute Determined by Means of SSR Markers and AFLP Profiling. Crop Sci. **44**: 678-685.

Ghosh T (1983) Handbok on jute. FAO, Rome.

Ghosh M, Saha T, Nayak P and **Sen SK** (2002) Genetic transformation by particle gun bombardment of cultivated jute, *Corchorus capsularis* L. Plant cell reports. **20**: 936-942.

Hossain MB, Haque S and **Khan H** (2002). DNA Fingerprinting of Jute Germplasm by RAPD. J. Biochem. and Mol. Biol. **35**(4): 414-419. (BSRK and Springer-Verlag)

Hossain MB, Awal A, Rahman MA, Haque S and **Khan H** (2003) Distinction between cold sensitive and cold tolerant jute by a combined RAPD and AFLP study. J. Biochem. and Mol. Biol. **36**(5): 427-432 (BSRK and Springer-Verlag).

Islam AS (1958) Crossing relationship in the genus *Corchorus.* Proc. Intl. Botanic Congress held in Montreal, Canada.

Islam AS and Rashid A (1960) First successful hybridization between the two jute yielding species. Nature **185**: 258-259.

Khan F, Islam AI and **Sathasivan K** (2004) A rapid method for high quality RNA isolation from Jute: *Corchorus capsularis* and *C. olitorius* L. Plant Tissue Cult. **14**: 63-68.

Liu SL, Chang YS and **Hsieh CC** (2005) Identification of Corchorus species based on ITS sequence of nuclear ribosomal DNA in Taiwan. GenBank Submission. 08-NOV-2005. Graduate Institute of Chinese Pharmaceutical, China Medical University, Taichung, Taiwan

Saha T, Ghosh M and **Sen SK** (1999) Plant regeneration from cotyledonary explants of jute. *Corchorus capsularis* L. Plant Cell Rep. **18**: 544-548.

Seraj ZI, Sarker AB and **Islam, AI** (1992) Plant regeneration in a jute species (*C. capsularis*) and its possible relationship with glyoxylase-I. Plant Cell Rep. **12**: 29-33.

Sharma S, Singh R, Rustgi S, Goyal A, Gau A, Tyagi AK, Khan H, Sinha MK, Balyan HS and **Gupta PK** (2005) Development and use of simple sequence repeat (SSR) markers for the study of DNA polymorphism, transferability and genetic diversity in jute. (Personal communication and sequences posted in GenBank as of Dec 2005)

Whitlock BA, Karol KG and **Alverson WS** (2003) Chloroplast DNA sequences confirm the placement of the enigmatic *Oceanopapaver* within *Corchorus* (Grewioideae: Malvaceae *sensu lato*., formerly Tiliaceae). Intl. J. Plant Sci. **164**(1): 35-41.

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis

Preliminary Progress in Jute (*Corchorus* spp.) Genome Analysis