International Journal of Agriculture, Environment & Biotechnology
Citation: IJAEB: 7 Special Issue : 439-450 July 2014

Molecular Biology

Genetic diversity and molecular analysis among cotton genotypes by EST-SSR markers.

Kundan K Mishra[1]* and R S Fougat[2]

[1]Department of Biotechnology, Genetics and Bioinformatics, N.V. Patel College of Pure and Applied Sciences, Vallabh Vidyanagar 388 120, India.
[2]Centre of Excellence in Biotechnology, B. A. College of Agriculture, Anand Agricultural University, Anand –388 110, Gujarat, India.

**Corresponding author: Kundan K Mishra; kkmishra71075@gmail.com*

Abstract

Objective of this research work is to utilize EST- SSR markers for fibre quality traits to generate genetic diversity between tetraploid (*Gossypium hirsutum*) and diploid (*Gossypium herbaceum* and *Gossypium arboreum)* cotton species at molecular level. Twenty four (24) genotypes of cotton and thirty five (35) EST SSR primers for different fibre quality traits (fibre length, Lint percentage, Boll weight and fibre strength) were taken for this study. Almost all primers reveals amplification in both diploid and tetraploid cotton species which indicates that flanking primer sequences are conserved in both genomes of cotton. Thirty one (31) EST SSR primers generate good and enormous amplicon and produced a total of four hundred seventy eight (478) sharp, similar and variable bands in all genotypes. Average number of bands amplified by each primer was 15.419. Statistical analysis for EST SSR data was conducted using software programme NTSYS pc version 2.02e and GeneAlEx. The genetic distance (GD) among the all genotypes of cotton were also analyzed and it is ranged from 0.05 to 0.71 which indicates significant diversity between all the genotypes of both tetraploid and diploid cotton species. The average observed mean heterozygosity was 0.60 and observed mean percentage of polymorphic loci was 60%.

Highlights

- To study genetic diversity between tetraploid and diploid cotton species
- Molecular analysis for phylogenetic relationship and fiber traits in both diploid and tetraploid cotton species
- Identification and validation for EST-SSR markers in both diploid and tetraploid cotton species
- Application of EST-SSR in Marker Assisted Selection(MAS) and cotton breeding

Keywords: EST- SSR, Diploid, Tetraploid, Cotton, Fibre quality, Genetic Distance.

The genus *Gossypium* commonly called cotton occurs naturally throughout tropical and subtropical regions and belongs to family *Malvaceae* and popularly called white gold, includes 50 species split across two ploidy levels, 45 species are diploid (2n=2x=26) and five are allotetraploid (2n=2x=52) (Wendel 1989; Percival *et al*; Wendal and Cronn, 2003). There are three major lineages of diploid species, corresponding to three continents: Australia (C, G, K genomes), the Americas (D genome), and Africa/ Arabia (A, B, E, and F genomes) (Stewart *et al.,* 1995; Percival *et al.,* 1999).The genome sizes are estimated to be 880-Mb for *Gossypium raimondii* Ulbrich,-1.75-Gb for *Gossypium arboreum* L, and ~2.5-Gb for *Gossypium hirsutum* L.(Hendrix and Stewart,2005). Among these 50 species four are mainly cultivated viz. *Gossypium arboreum* L.(n=x=13, A2) and *Gossypium herbaceum* L. (n=x=13 A1) are diploid while *Gossypium barbadense* L.[n=2x=26, (AD)2] and *Gossypium hirsutum* L.[n=2x=26, (AD)2] are teraploid (Kumria *et al.*, 2003). *Gossypium herbaceum* L is limited to Gujrat

Table 1. List of genotypes

Sr. No.	Gossypium herbaceum	Sr. No.	Gossypium arboreum	Sr. No.	Gossypium hirsutum
1.	Digvijay	12.	824	19.	G. cot 12
2.	Gvhv133	13.	CINA 329	20.	G. cot H4
3.	Dhumad	14.	CINA318	21.	G. cot 16
4.	Gvhv235	15.	Jawahar Tapti	22.	G. cot H10
5.	Gvhv544	16.	G-27	23.	G cot 8
6.	V-797	17.	CINA333	24.	G. cot 20
7.	G. cot 13	18.	DLSA17	----	----
8.	G.cot 21	---	---	---	---
9.	G. cot 23	---	---	---	---

and Karnataka. *Gossypium hirsutum* L and *Gossypium arboreum* L are grown in all the major cotton growing states of India. Gujrat has been a major cotton producing state for last many years. In term of area, Gujrat ranked 2nd next to Maharastra and contributes about ~37.5% of the national cotton production from~ 26.5 area of the country.

Cotton is the world's leading textile fiber crop and it is also source of secondary products such as oil, live-stock feed (cotton seed cake) and cellulose and play an important role in global economy. In recent years, improvement in the quality of cotton fiber has been extremely important because of changes in spinning technology (Shen *et al*.,2005). Enhanced cotton resources are needed to facilitate the improvement of this important crop. Therefore, an important area of cotton genomics is germplasm characterization and utilization. Analysis of genetic diversity and relatedness between species and among genotypes is useful in plant breeding programs because it provides a tool for accurate organization of germplasm and efficient parental selection. With advancement in molecular marker technology, marker-assisted selection (MAS) combined with conventional breeding has been one way in which fiber quality can improved

Simple sequence repeat (SSR) markers (microsatellite) are tandemly repeated DNA motifs (1-6bp long) which may vary in the number of repeats at given locus, have been sucessfully employed in many genetic diversity studies (Liu *et al*., 2000b; Grutierrez *et al.,*2002) and are useful for a variety of applications in plant genetics and breeding because of their reproducibility, multiallelic nature, codominant, relative abundance and good genomic coverage (Powel *et al*., 1996). SSR are easy to use and analyze (Morgante and Olivieri 1993). The variation among microsatellite is thought to be due to the slippage of DNA polymerase during replication or unequal crossing over, resulting in differences in the copy number of the actual nucleotide sequences (Yu et al.,1999). Polymorphism among individuals arises from changes in the number of the repeats. In the other words, these markers meet most of the requirements for ideal markers for assessing gene flow. Tracking of microsatellite markers requires specifically designed primers for conserved flanking region of repeats and PCR amplification of this region. The availability and abundance of microsatellite markers throughout the cotton genome coupled with the fact that they are polymorphic, co-dominant and are based on polymerase chain reaction (PCR) make them particularly useful in genetic diversity studies of cotton (Reddy *et al* ,2001), with in excess of 1000 microsatellite primers having already been isolated from cotton DNA genomic libraries (Nguyen *et al*.,), for molecular studies of the genetic diversity (Brubaker and Wendel. 1994; Tatineni *et al*., 1996; Iqbql *et al*., 1997). However, more work need to be carried out and the purpose of work is to investigate and compare the genetic diversity of cotton plants cultivated in several country of Asia with the specific objectives of estimating the

information of cotton microsatellite loci and selecting a set of microsatellite primers able to differentiate between the various cultivars studied and to estimate the genetic distance among various cotton cultivars in Asia. Recent studies have revealed that gene transcripts can also contain repeat motif, and the abundance of expressed sequence tags (ESTs) makes this an attractive potential source of microsatellite markers (Kantety *et al* 2002). EST-SSRs have also been developed for *Gossypium* (Saha *et al*. 2003; Han *et al.*2004; Quereshi *et al*. 2004). More than 4lakhs *Gossypium* sequences were in genebank. A total of ~70000 ESTs derived from the *G.raimondii* D-genome, ~450000 ESTs from the A- genome, ~260000 from the AD tetraploid of *G. hirsutm* ,about ~48000 ESTs from *G. arboreum*, ~350 from *G.herbaceum*, ~300 from *G. barbadense*, are available in the dbEST. These EST-SSRs provide structural and functional genomic information that will be useful for understanding cotton fiber development.

Several EST SSR markers have been identified in tetraploid cotton species and successfully used to develop anchor SSR for cotton chromosomes to provide a basis for genetic mapping (Liu *et al*.,2000; Reddy *et al*.,2001). There are two reports that SSR primers specific to *G. hirsutm* L. have been used in mapping of *G. barbadense* L.(Liu *et al*., 2000a) and *G. nelson* L. and *G. australae* L.(Quereshi *et al* 2001). EST SSR markers seems to be a good tool for genomic studies and can minimize the laborious cloning and screening steps of SSR development (Liu *et al*.,2000; Dayanandan *et al*., 1998, 1999; Echt *et al.,* 1999). Use of EST SSR markers from tetraploid cotton in the mapping of diploid cotton species could be successfully utilized for marker-assisted selection (MAS) and genetic diversity analysis in diploid cotton.

Materials and Methods

Experimental Plant Material

Gossypium species, total 24 genotypes were selected for this comparative studies (11 *G. herbaceum,* 7 *G. arboreum* and, 6 *G. hirsutum* genotypes). The genotypes of cotton were procured from Central Institute of cotton Research Nagpur, Regional Research Station, Anand Agricultural University (AAU) Anand, Main Cotton Research Station, Surat, Regional Cotton Research Station Viramgam and Regional Cotton Research Station, Bharuch.

DNA Extraction

Plants genotypes were grown in polyhouse and DNA samples from the tetraploid and diploid Gossypium species were extracted. Genomic DNA of all genotypes was isolated at the microlevel from young leaves using a CTAB-based extraction method of Altaf *et al.*, (1997) with slight modifications. Approximately 0.3 gm of fresh young leaf tissue was homogenized in liquid nitrogen and 0.5 ml of extraction buffer [100 mM tris-HCl (pH 8.0), 20 mM EDTA (pH 8.0), 1 M NaCl, 2% CTAB, 2% PVP-40, 1 mM 1-10, phenanthroline] and 0.2% β-mercaptoethanol in a 1.5 ml centrifuge tube with the aid of a microtube pellet pestle. After incubation for 1 hr at 60ºC, the suspension was purified twice in chloroform:isoamyl alcohol (24:1) by centrifugation at 10,000 rpm on a desktop micro-centrifuge for 12 min, at room temperature and precipitated with an equal volume of ice cold isopropanol. The recovered DNA was spooled out, or pelleted by centrifuging at 10,000 rpm for 8 min, washed twice with 80% EtOH, 15 mM ammonium acetate and once with 95% EtOH, air dried, and dissolved in 50 to 100 µl of 10mMTris buffer (pH 7.5). To this sample 2 µl of Rnase A (10 mg/ml) per 100 µl of dissolved DNA was added.

Purity and Quantification test of DNA

Spectrophotometry was performed to determine DNA concentration by using Nanodrop N.D.1000 (Software V.3.3.0, Thermo Scientific, USA) at absorbance ratio 260/280 nm and the quality of obtained DNA was checked on 0.8% agarose gel. Dilution of 20 ng/µl working solution was prepared from the stock solution of the isolated DNAs. The 1.5 µl of DNA sample was loaded into the well of Nanodrop Spectrophotometer (Thermo Scientific, U.S.A.) and the concentration of DNA and absorbance at 260 nm and 280 nm were measured and the A_{260}/A_{280} ratio was automatically calculated by the software. The quality of DNA was evaluated by spectrophotometry using the 260/280 nm absorbance ratio method and by electrophoreses on a 1% (w/v) agarose gel, and the DNA concentration estimated at 260 nm (Sambrook et al., 1989). The stock DNA samples were stored at -20°C and working DNA samples (containing 50 ng/µL) at 4°C.

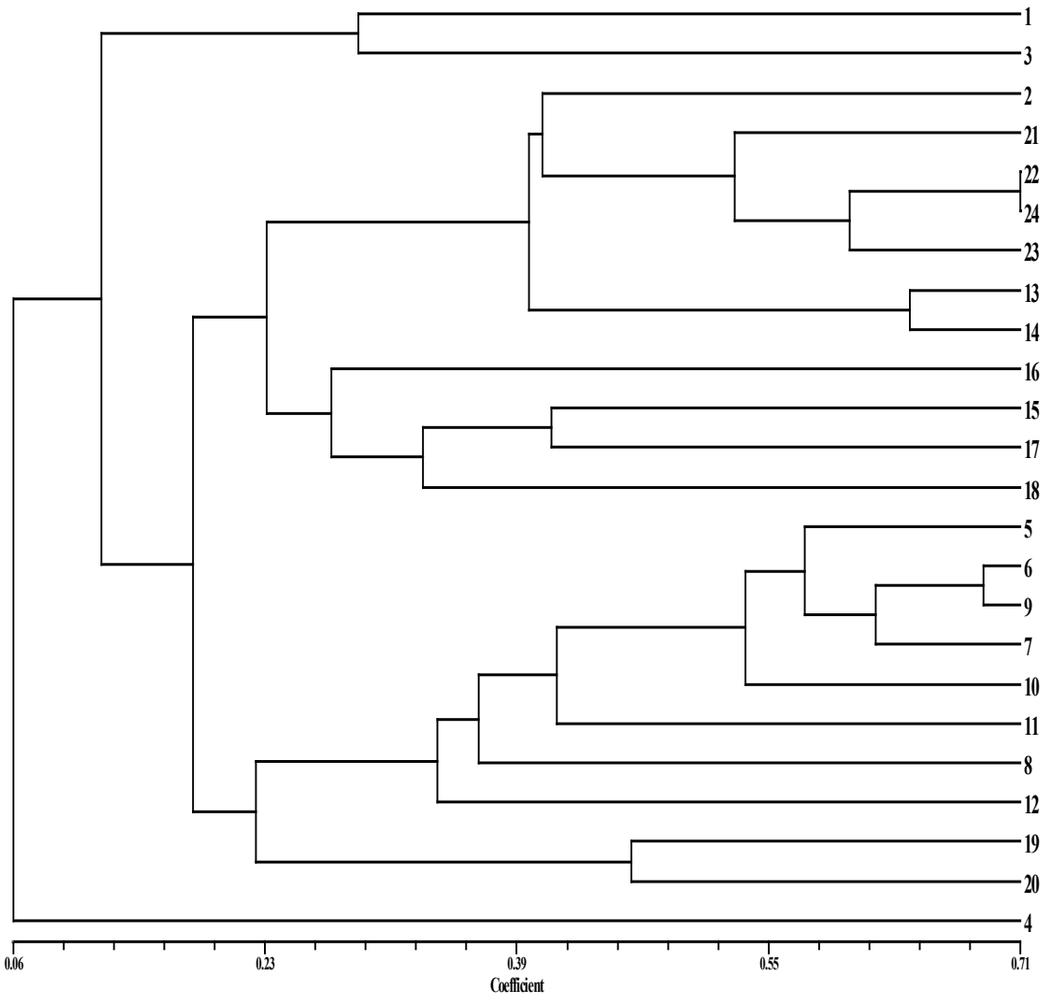
Table 2. Results of EST SSR analysis.

Sr. No.	Locus Name	Molecular band size (bp)	Total No. of bands	Total No. Alleles	PIC value
1.	BNL 2920	360-174	29	4	0.685121
2.	BNL 1317	370-176	18	3	0.549383
3.	BNL341	125-80	16	2	0.492188
4.	BNL1231	197	3	1	0
5.	BNL1414	227-127	13	2	0.473373
6.	BNL3874	328-195	7	2	0.244898
7.	BNL4030	679-156	31	5	0.553469
8.	BNL3408	148-110	11	3	0.495868
9.	BNL3482	370-230	11	3	0.396694
10.	BNL 1030	190-164	11	3	0.396694
11.	BNL3650	110-80	10	2	0.42
12.	BNL 58O	210-165	9	3	0.493827
13.	CIR30	356-117	16	3	0.554688
14.	CIR45	167-114	21	3	0.444444
15.	CIR70	296-156	7	2	0.244898
16.	CIR78	336-295	12	2	0.375
17.	CIR246	230	7	1	0
18.	JESPR 127	386-192	17	3	0.484429
19.	JESPR 307	388-94	23	4	0.729600
20.	JESPR 65	316	6	1	0
21.	JESPR 29	176-110	11	2	0.165289
22.	JESPR208	526-116	34	5	0.705247
23.	NAU326O	189-132	10	2	0.197531
24.	NAU4024	569-202	29	4	0.661407
25.	NAU1200	276-108	23	3	0.635161
26.	NAU1369	369	7	1	0
27.	NAU1531	296-96	24	4	0.672154
28.	NAU2035	282	7	1	0
29.	NAU3393	294-259	8	3	0.40625
30.	NAU3654	276-96	17	3	0.290657
31.	NAU 923	574-290	30	5	0.730489
Total		-	478	83	-
Average		-	15.419	2.67	0.465411

Table 3. List of EST-SSR Primer their Quantitative Trait Loci (QTL) /Gene name and EST-SSR linked repeat sequence.

Sl.No	Primer sequence	Primer	QTL	QTL/ Gene Name	EST-SSR Linked Repeat length
1.	TTTGGAGCCATTTACATGCA	BNL1030 (F)	QTL	qFL-D9-1	(GT)16 (CA)13
2.	AAACCACTTCTGCATCTGGA	BNL1030(R)	QTL	qFL-D9-1	(GT)16 (CA)13
3.	AAAAATCAGCCAAATTGGGA	BNL1317 (F)	QTL	qFL-D9-1	(AG)14
4.	CGTCAACAATTGTCCCAAGA	BNL1317(R)	QTL	qFL-D9-1	(AG)14
5.	AAAAACCCCTTTCCATCCAT	BNL1414(F)	QTL	qFL-D9-1	(AG)16
6.	GGGTGTCCTTCCCAAAAATT	BNL1414(R)	QTL	qFL-D9-1	(AG)16
7.	TAATAAAAGGGAAAGGAAAGAGTT	BNL1231(F)	QTL	qFS-A1-1	(AG)15
8.	TATGGCTCTAGAATATTCCCTCG	BNL1231(R)	QTL	qFS-A1-1	(AG)15
9.	CCTCCCTCACTTAAGGTGCA	BNL4030(F)	QTL	qFS-D-6-1	(GT)10
10.	ATGTTGTAAGGGTGCAAGGC	BNL4030(R)	QTL	qFS-D-6-1	(GT)10
11.	CACCATTGTGGCAACTGAGT	BNL3140(F)	QTL	qFS-D9-1	(GA)11
12.	GGAAAAGGGAAAGCCATTGT	BNL3140(R)	QTL	qFS-D9-1	(GA)11
13.	TCGATTTCCTTATTTGATTTCTG	BNL3650(F)	QTL	05 Lint	(TC)15+(TA)6
14.	AATTTGTCCAGATTCATTCTTCA	BNL3650(R)	QTL	05 Lint	TC)15+(TA)6
15.	ATCCAAACCATTGCACCACT	BNL3408(F)	QTL	LP	GT2AT(GT)12
16.	GTGTACGTTGAGAAGTCATCTGC	BNL3408(F)	QTL	LP	GT2AT(GT)12
17.	CATGTTCTAATCATATATATATGTATATATATGTGT	BNL3874(F)	QTL	LP	(AT)5+G+(TA)4(GT)10
18.	AAAATAACAAAAGCCATGGAATAA	BNL3874(R)	QTL	LP	(AT)5+G+(TA)4(GT)10
19.	TTGAGGGCATCCAAATCCAT	BNL3994(F)	QTL	LP	(CT)25+(GA)25
20.	CCTCCACCATACACGTGCTA	BNL3994(R)	QTL	LP	(CT)25+(GA)25
21.	ATTTGCCCCAGGTTTTTTTT	BNL3482(F)	QTL	LP	(AC)12
22.	GCAACACCTTTTCCTCCCTA	BNL3482(R)	QTL	LP	(AC)12
23.	TTCTTGCATTGAATAATACTGGC	BNL2920(F)	QTL	BW	(AG)12
24.	CTTAATTCTAAAAATCAATAA	BNL2920(R)	QTL	BW	(AG)12
25.	CTATGTTTGGCCTTGGCATT	BNL580(F)	QTL	BW	(GT)+(GT)+(CT)14
26.	TAGTGACAGATATCCCCGGC	BNL580(R)	QTL	BW	(GT)+(GT)+(CT)14
27.	ACCTGGGGTACTTGTCCACA	BNL341(F)	QTL	BW	(GA)14
28.	CCATCCCATTTGTGATACCC	BNL341(R)	QTL	BW	(GA)14
29.	TTAGGGTTTAGTTGAATGG	CIR246(F)	QTL	qFL-D2-1	(TG)6
30.	ATGAACACACGCACG	CIR246(R)	QTL	qFL-D2-1	(TG)6
31.	ACTAGCAGTGCGAATACA	CIR45(F)	QTL	qFL-D2-1	(TG)7
32.	TGGTTAAGGGTTGGG	CIR45(R)	QTL	qFL-D2-1	(TG)7
33.	CAATATCTCACTTGGACCT	CIR30(F)	QTL	LP	(C)8(TC)6(CA)8
34.	TGCTACACATCATAGTTGG	CIR30(R)	QTL	LP	(C)8(TC)6(CA)8
35.	TGCATGATGAAGTTAGA	CIR78(F)	QTL	LP	(GT)7
36.	ACATAAATCCCAAGAAC	CIR78(R)	QTL	LP	(GT)7
37.	CTTCATCATAGTAGCGAGTT	CIR182(F)	QTL	LP	(AC)10

38.	GAATCAAGCAGAGGATTT	CIR182(R)	QTL	LP	(AC)10
39.	AACCACCAACCATTCA	CIR70(F)	QTL	qFS-D8-1	(AC)8
40.	TGGGACTCGGTCATC	CIR70(R)	QTL	qFS-D8-1	(AC)8
41.	CCACCCAATTTAAGAAGAAATTG	JESPR65(F)	QTL	qFL-A5-1	(GAA)25
42.	GGTTAGTTGTATTAGGGTCGTTG	JESPR65(R)	QTL	qFL-A5-	(GAA)25
43.	CTTGGCCATGTATTCCTTCA	JESPR307(F)	QTL	qFL-D6-1	(TGA)11
44.	GAAAGACACTAAGCTGAGGC	JESPR307(R)	QTL	qFL-D6-1	(TGA)11
45.	CGCAACCAAACATATACTTCACAC	JESPR208(F)	QTL	qFL-D9-1	(CT)15
46.	CCCTTTCCATCCATAGAACG	JESPR208(R)	QTL	qFL-D9-1	(CT)15
47.	GATTTGGGTAACATTGGCTC	JESPR127(F)	QTL	qFL-D8-1	(GA)9AA(GA)5
48.	CTGCAGTGTTGTGTTGGGTAGA	JESPR127(R)	QTL	qFL-D8-1	(GA)9AA(GA)5
49.	CACCGTTTCCAAGTAAGATT	JESPR29(F)	QTL	LP	(CTT)18
50.	GGTTAATCTTAGTTGAGGTC	JESPR29(R)	QTL	LP	(CTT)18
51.	TTTTGCAGATGTTTGTAGGG	NAU3260(F)	QTL	qFL-A10-1	(CT)4 4CA(CT)4
52.	TTTCTTCAACAGGGGCTAAG	NAU3260(R)	QTL	qFL-A10-1	(CT)4 4CA(CT)4
53.	TTCGGGAAAGTTAGAGGAGA	NAU1233(F)	QTL	qFL-A10-2	(AAT)6
54.	TCCTCAGAGCTCGGAATAGT	NAU1233(R)	QTL	qFL-A10-2	(AAT)6
55.	CAACAGCAACAACCACAA	NAU1200(F)	QTL	qFL-A5-1	(CAG)11
56.	CTGCCTCGAGGACAAATAGT	NAU1200(R)	QTL	qFL-A5-1	(CAG)11
57.	CAGCCATCCCTCCTCTAATA	NAU3393(F)	QTL	qFLLG021	(CTG)7
58.	GTCAGCAGCCATTCTAACCT	NAU3393(R)	QTL	qFLLG021	(CTG)7
59.	ACAAGCATCTTCATGGACCT	NAU4024(F)	QTL	qFLLG021	(GTC)6
60	AGAAGGATGATGCAAAGAGG	NAU4024(R)	QTL	qFLLG021	(GTC)6
61	TTACCAGCAGCCAACACTAA	NAU3654(F)	QTL	qFSLG05-1	(TGA)5
62	TCCCCTTCAACATCTTCTTC	NAU3654(R)	QTL	qFSLG051	(TGA)5
63	TGGCAGAGATGAATGTAAGC	NAU1369(F)	QTL	qFS-D6-1	(AGGCGG)3
64.	GGTAACGGATGGAAAATCAC	NAU1369(R)	QTL	qFS-D6-1	(AGGCGG)3
65.	CGAGAAACTTCACTGGACCT	NAU2035(F)	QTL	qFS-D6-1	(GATA)4
66.	GAAAAGGTAGGCTTGTTGGA	NAU2035(R)	QTL	qFS-D6-1	(GATA)4
67.	GGAATTCAAGGTTGAAGGAG	NAU923(F)	QTL	qFS-D9-1	(TCTTTT)4
68.	CCTCTTCTTTGGCTCTGAAA	NAU923(R)	QTL	qFS-D9-1	(TCTTTT)4
69.	GTATCCGCCCACAAATAAAG	NAU1043(F)	QTL	qFS-LG051	(TTC)14
70.	GCATCGTGAGAGAAAGTGAA	NAU1043(R)	QTL	qFSLG051	(TTC)14



1.Digvijay; 2.Gvhv133; 3.Dhumad; 4.Gvhv235; 5.Gvhv544; 6.V-797; 7.G. cot 13; 8.G.cot21; 9.G. cot 23; 10.Gvhv505; 11.Gvhv104; 12.824; 13.CINA 329; 14.CINA318; 15.Jawahar Tapti; 16.G-27; 17.CINA333; 18.DLSA17; 19.G. cot 12; 20.G. cot H4; 21.G. cot 16; 22.G. cot H10; 23.G cot 8; 24.G. cot 20.

Fig-1 Dendrogram showing the genetic relationship between different cotton genotypes based on Nei's (1978) similarity coefficients using UPGMA as the clustering method for the polymorphism data obtained at 31 EST SSR loci.

EST SSR (Microsatellite) PCR Amplification

We used 35 primer pairs to investigate genetic variation and phylogeny in our 24 cotton cultivars. Thirty one of 35 primer pairs easily produced detectable products. Amplifications were carried out in 200-µL microtubes containing a 25 µL reaction mix consisting of 20 ng template, 0.2 µM of each primer, 1 U Taq DNA polymerase (Finzymes), 0.2 mM of each dNTP, 0.2 to 0.325 mM MgCl2, and 1X reaction buffer (10 mM Tris-HCl and 50 mM KCl, pH 8.3). The amplification was carried out in a Thermal Cycler (Whatman Biometra T-Gradient, Germany) using a program consisting of a denaturation step of 5 min at 94°C followed by 40 cycles of 60 s at 94°C, 60 s at 55°C, and 1.5 min at 72°C. The program ended with an extension step at 72°C for 7 min. Agarose gel of 2.5% concentration was prepared in 1X TBE (2.5 g agarose in 100 ml 1X TBE and 2.5 µl Ethidium bromide from 10 mg/ml stock). PCR amplified products (9 µl and 1 µl 6X loading dye) were loaded into the wells. The 1000 bp standard DNA ladder (1 µl) (marker) was also run along with the samples. The electrophoresis was conducted at 100 V current (constant) for 2.5 hrs. to separate the amplified bands. The separated bands were visualized under UV transilluminator (Biometra,Germany) and photographed using gel documentation system (Bio-rad, California).

Scoring and analysis of data

Clear and distinct bands amplified by 35 SSR primers were scored for the presence and absence of the corresponding band among the all 24 cotton genotypes. The scores 1 and 0 indicates the presence or absence of bands respectively. The softwares used for the analysis of the scored data were NTSTSpc version 2.02 (Rholf 1998), GeneAlEx (Peakall *et al.,* 2010) and SPSS. The major part of the analysis was carried out using NTSYSpc version 2.02 (Rholf, 1998) except for the calculation of the Shannon index, observed and effective number of alleles which were calculated using GeneAlEx6.4.

Results and Discussion

The 31 pairs of SSR primers amplified a total of 83 alleles with an average of 2.48 per SSR locus and PIC values varying from 0.165289 to 0.730489 with an average of 0.465411. Molecular marker JESPR 29 was proved to be the least informative marker with a PIC 0.165289, while, on the contrary, NAU923 was proved to be the most informative with a PIC of 0.730489. Markers with higher PIC values have the potential to reveal the variation between alleles and could be used more effectively for molecular mapping and analysis of genetic variation in a population .

The 31 polymorphic molecular markers amplified from 1-4 alleles for each locus with BNL4030 found to amplify 5 alleles. This marker is an informative marker with a PIC of 0.553469 while a similar PIC has been reported by Liu et al. (2000b). An SSR profile was constructed for each cultivar using the 31 primer pairs, which were able to discriminate between 24 cultivars studied. Phylogenetic analysis revealed that the 24 cotton cultivars that were examined belong to 2 main groups. The finding of genetic distance between the cotton cultivars using SSR molecular markers can be useful in marker-assisted selection (MAS). This acquires greater importance for cotton where the genetic base of improved cultivars is limited enough, because EST SSR molecular markers have higher analytical force, and hence, they reveal higher variation between genotypes compared to other molecular markers.

In the present SSR analysis with 35 microsatellite markers only 31 markers gave the results. All thirty one markers produced 83 alleles. The average number of alleles per locus was found to be 2.47. A maximum of 5 alleles were recorded for marker BNL 4030, JESPR 208 and NAU 923 was followed by BNL 2920, JESPR 307, NAU 4024 AND NAU 1531 which produced 4 alleles. BNL 1317, BNL 580, CIR30, NAU 1200 and NAU3393 produced three alleles each and BNL1531, CIR 246, CIR 307, JESPR65, NAU 1369, NAU2035 which was lowest in the present investigation.

The highest PIC value was recorded for NAU 923 (0.730489) and lowest for JESPR 29 (0.165289); whereas mean PIC value from all tested microsatellite was 0.465411 The average heterozygosity for 31 EST-SSR markers was 0.60. Dendrogram Based on Nei's (1978) unbiased measures of genetic distance by UPGMA method formed two major clusters which grouped all 24 cotton genotypes (Fig.). Cluster 1(B) which was smallest

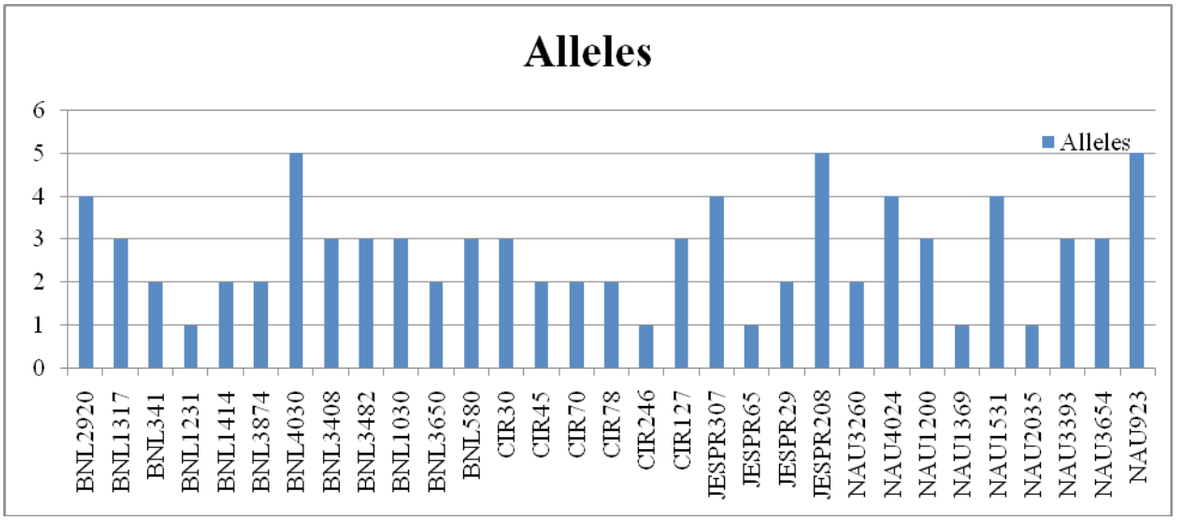


Fig. 2. Number and frequency distribution of alleles for all primer locus.

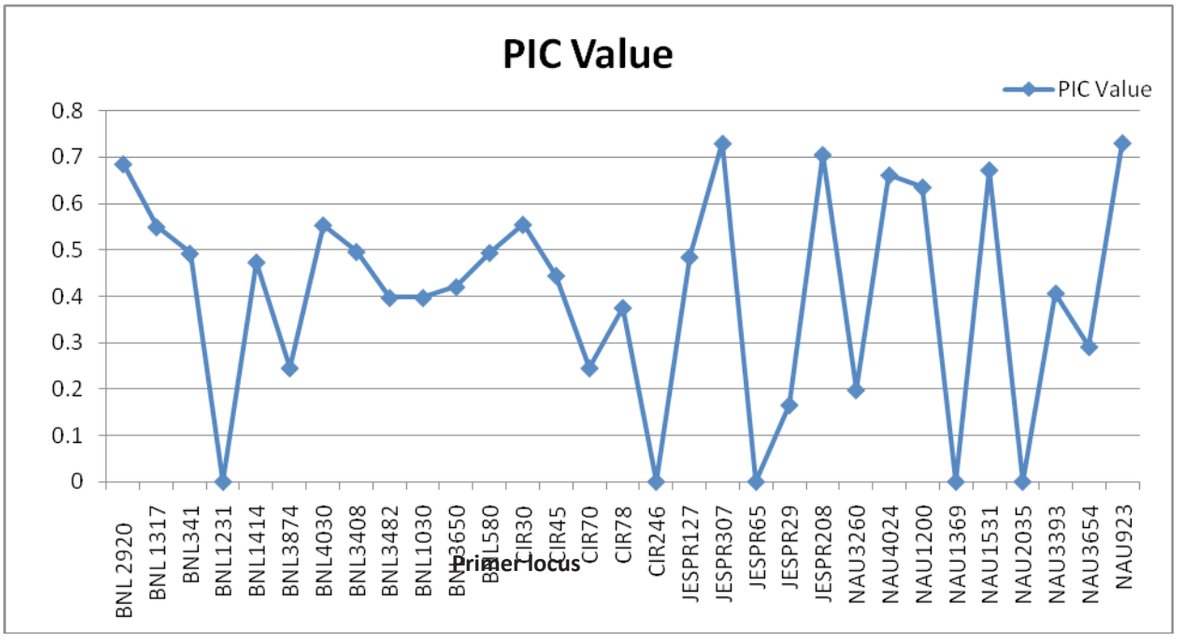


Fig. 3 Frequency of PIC value of all primers for different fibre quality traits and their diversity for EST SSR Marker.

in all clusters included Gvhv 235. Second cluster formed two subgroups viz. A1 and A2. Cluster A1 included two genotypes Digvijay and Dhumad. The A2 cluster again formed two sub group viz. A2A and A2B. A2A includes 11 genotypes viz. Gvhv133, G.cot16, G.cotH10, G.cot 20, G.cot8, CINA329, CINA318, CINA344, Jawar Tapti, CINA333 and DLSA17. A2B includes 10 genotypes viz. G.cotH4, G.cot12, 824, G. cot21, Gvhv104, Gvhv505, G.cot13, G.cot13, V-797 and Gvhv544. Gvhv133, G.cot 16, G.cotH10, G.cot20, G.cot8, CINA329, CINA318, G27, Jawar Tapti, CINA333 and DLSA17 which leads us to assume that the A2A subgroup cluster included the varieties of similar genetic origin and which can be helpful for consideration of these varieties for the hybridization programme, while the A2B subgroup cluster contained G.cotH4, G.cot12, 824, G. cot21, Gvhv104, Gvhv505, G.cot13, G.cot13, V-797 and Gvhv544. The involvement of genotypes in the same cluster again indicated that the varieties with the same genetic composition were falling in the one cluster and this can help to conduct a hybridization programme with these genotypes. The similarity coefficients ranged from 0.05 to 0.71 for all accessions. Thus, the relatedness of the cultivars studied was efficiently established through the use of SSR markers though with some differences in the positioning of some cultivars at various clusters. Both the dendrograms reflect relationships among most of the cultivars, depending upon their fiber quality traits.

Application of Tetraploid Derived EST SSR Primers in Diploid cotton Species.

The electrophoretic profiles obtained with the study of the primers, 31 primers gave amplification in both diploids and tetraploids, out of which 26 primers gave polymorphism in both diploids and tetraploids. JESPR29, NAU3260, BNL3874, CIR70, NAU3654, CIR78, BNL1030, BNL3482, NAU3393, BNL3650, CIR45, BNL1414, CIR307, BNL341, BNL580, BNL3408, BNL1317, BNL4030, CIR30, NAU1200, NAU4024, NAU1531, BNL2920, JESPR208, JESPR127 and NAU923 produced significant polymorphism among the *G. herbaceum, G.arboreum* and *G.hirsutum,*genotypes analyzed. EST-SSR markers amplified by primer CIR 30, NAU1200, NAU4024, NAU1531, BNL2920, JESPR208, JESPR127 and NAU923 gave high polymorphism and they are also abundant in both diploid (*G. herbaceum, G.arboreum*) genotypes therefore, they are best markers for marker assisted selection (MAS) in both diploid (*G. herbaceum and G.arboreum*) genotypes for fibre quality traits. Fibre quality linked SSR markers of tetraploid cotton in diploid species may be used to amplify DNA fragments in most of diploids genome.

EST-SSR markers used for this study reveals diversity and similarities among and between genotypes of diploid and tetraploid cottons genotypes based on molecular characterization through EST-SSR Primers.

Conclusion

Comparative study between diploid and tetraploid cotton species for various fibre quality traits indicated that tetraploid-derived EST SSR primers for fibre quality traits could be successfully used in diploid species of cotton for marker assisted breeding. Large portions of EST-SSR flanking regions appear to be conserved among diploid and tetraploid genomes of cotton. EST-SSR markers amplified by primer CIR 30, NAU1200, NAU4024, NAU1531, BNL2920, JESPR208, JESPR127 and NAU923 gave high polymorphism and they are also abundant in both diploid (*G. herbaceum,G.arboreum*) species therefore, they are best markers for marker assisted selection (MAS) in both diploid (*G. herbaceum* and *G.arboreum*) species for fibre quality traits. These studies also determined that tetraploid derived EST SSR marker primers could be used to amplify DNA fragment in diploid cotton species.

Dendrograms generated for EST- SSR markers of cotton genotypes reflect relationships among most of the diploids (*G. herbaceum, G.arboreum*) and tetraploids (*G.hirsutum*) cultivars, depending upon their fibre quality traits. The involvement of genotypes in the same cluster again indicated that the varieties with the same genetic composition were falling in the one cluster and this can help to conduct a hybridization programme with these genotypes.

References

Abdurakhmonov IY, Buriev ZT, Saha S, Pepper AE, et al. 2007. Microsatellite markers associated with lint percentage trait in cotton, *Gossypium hirsutum. Euphytica* **156:** 141-156.

Adawy, S.S., Ebtissam, H.A. Hussein and A.E. Hanaiya 2006. Molecular characterizationand genetic relationship among cotton genotypes 2-AFLP analysis. *Arab J. Biotech*. **9**(3): 478-492.

Bertini CHCM, Schuster I, Sediyama T, and Barros EG,2006. Characterization and genetic diversity analysis of cotton cultivars using microsatellites. *Genet. Mol. Biol*. **29**: 321-329.

Brubaker CL and Wendel JF 1994. Reevaluating the origin of domesticated cotton (*Gossypium hirsutum*, Malvaceae) usingnuclear restriction fragment length polymorphism (RFLPs). *American Journal of Botany* **81**:1309-1326.

Candida H.C.de, Magalhaes Bertini, Ivan Schuster, Tocio Sediyama, Everaldo Goncalves de Barros and Maurilio Alves Moreira 2006. Characterization and genetic diversity analysis of cotton cultivars using microsatellites. *Genetics and Molecular Biology* **29**(2): 321–329.

Chen ZJ, Scheffler BE, Dennis E, Triplett BA, Zhang TZ, Guo WZ, Chen XY, Stelly DM, Rabinowicz PD, Town CD, Arioli T, Brubaker C, Cantrell RG, Lacape JM, Ulloa M, Chee P, Gingle AR, Haigler CH, Percy R, Saha S, Wilkins T, Wright RJ, Deynze AV, Zhu YX, Yu SX, Abdurakhmonov I, Katageri I, Kumar PA, Rahman M, Zafar Y, Yu JZ, Kohel RJ, Wendel JF and Paterson AH 2007. Toward sequencing cotton (Gossypium) genomes. *Plant Physiol* **145:**1303–1310. doi:10.1104/pp.107.107672.

Chen L, Zhang ZS, Hu MC, Wang W, Zhang J, Liu DJ, Zheng J, Zheng FM, Ma J 2008. Genetic linkage map construction and QTL mapping for yield and fiber quality in upland cotton. *Acta Agron Sin* **34:**1199–1205. doi:10.3724/SP.J.1006.2008.01199.

Chuanfu An, Johnie N. Jenkins, Jixiang Wu, Yufang Guo, Jack C. McCarty 2010. Use of fiber and fuzz mutants to detect QTL for yield component, seed, and fiber traits of upland cotton. *Euphytica* **172:**21-34.

Connell, J. P., S. Pammi, M. J. Iqbal, T. Huizinga, and A. S. Reddy 1998. A high through-put procedure for capturing microsatellites from complex plant genomes. *Plant Mol. Biol*. **16:**341-349.

Cronn R, Cedroni M, Haselkorn T, Grover C, Wendel JF 2002. PCR-mediated recombination in amplification products derived from polyploid cotton. *Theoretical Applied Genetics* **104**:482–489.

Dao-Hua He, Zhong-Xu Lin, Xian-Long Zhang, Yi-Chun Nie, Xiao-Ping Guo, Yan-Xin Zhang, Wu Li, 2007. QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of *Gossypium hirsutum X Gossypium barbadense. Euphytica* **153**:181-197.

Decroocq V, Fave MG, Hagen L, Bordenave L, Decroocq S 2003. Development and transferability of apricot and grapeEST microsatellite markers across taxa. *Theory of Applied Genetics* **106**:912–922.

Diqui Liu, Xiaoping Guo, Zhongxu Lin, Yichun Nie and Xianlong Zhang 2006. Genetic diversity of Asian cotton (*Gossypium arboreum L)* in china avaluated by microsatellite analysis*. Genetic Resources and Crop Evolution* **53**:1145-1152.

Ebtissam H.A. Hussein, Marwa H.A. Osman, Mona H. Hussein and Sami S. Adawy 2007. Molecular characterization of Cotton Genotypes Using PCR-based Markers. *Journal of Applied Science Research* **3**(10): 1156-1169.

Felsenstein J 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**:783-791.

Feng, X., S. Saha, and K. M. Soliman, 1997. DNA fingerprinting in cotton using AFLP. *Focus* **19**:11-12.

Georgios Michailidis, Anagnostis Argiriou, Nikos Darzentas, Athanasios Tsaftaris 2009. Analysis of Xyloglucan endotransglycosylase/hydrolase (XTH) genes from allotetraploid (*Gossypium hirsutum)* cotton and its diploid progenitors expressed during fiber elongation*. Journal of Plant physiology* **166**: 403-416.

Goldstein, D. B., and D. D. Pollock. 1997. Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *J. Hered*. **88**:335-342.

Gupta P.K. and Varshney R.K. 2000. The development and use of microsatellite markers for genetic analysis and plant breeding withemphasis on bread wheat. *Euphytica.* **113**:163–185.

Guo Wangzhen, Sun Jing and Zhang Tianzhen 2003. Gene cloning and molecular breeding to improve fiber qualities in cotton. *Chinese Science Bulletin* **48**: 8709-717.

Guo WZ, Sang ZQ, Zhou BL, Zhang TZ 2007. Genetic relationships of D-genome species based on two types of EST-SSR markers derived from *G. arboreum and G. raimondii. Plant Science* **172**(4): 808-814.

Hendrix, B., and J. McD. Stewart. 2005. Estimation of the nuclear DNA content of *Gossypium* species. *Annals of Botany* **95**: 789-797.

Hong Chen, Neng Qian, Wangzhen Guo, Qingping Song, Baocheng Li, Fujun Deng, Cunguang Dong and Tianzhen Zhang 2010. Using three selected overlapping RILs to fine-map the yield component QTL on Chro.D8 in upland cotton. *Euphytica* **176**:321-329.

Hongde Qin, Wangzhen Guo, Yuan-ming Zhang, Tianzhen Zhang 2008. QTL mapping of yield and fiber traits based on a four-way cross population in (*Gossypium hirsutum* L.) *Theor Appl Genet* **117**:883-894.

Imelfort, M. and Edwards, D 2009. De novo sequencing of plant genomes using second-generation technologies. *Brief Bioinform* **10**(6):609-18.

Iqbal, MJ, Reddy OUK, El-Zik KM, Pepper AE 2001. A genetic bottleneck in the evolution under domestication of upland cotton *Gossypium hirsutum* examined using DNA fingerprinting. *Theor. Appl. Genet.* **103**: 547-554.

Jixiang Wu, Osman Ariel Gutierrez, Johnie N. Jenkins, Jack C. McCarty, Jun Zhu 2009.Quantitative analysis and QTL mapping for agronomic and fiber traits in an RI population of upland cotton. *Euphytica* **165**:231-245.

Kaliyaperumal Ashok kumar, R. Ravikesavan, and Kirubakaran Silvas Jrba kumar Prince 2010. Combining Ability Estimates for Yield and Fiber Quality Traits in Line X Tester Crosses of Upland Cotton, (*Gossypium hirsutum L*). *International Journal of Biology* **2**(1).

Kalivas A. , Xanthopoulos F, Kehagia O. and Tsaftaris A.S. 2011. Agronomic characterization, genetic diversity and association analysis of cotton cultivars using simple sequence repeat molecular markers, *Genetics and Molecular Research*, 10(1): 208-217.

Kantartzi SK and Stewart JM 2008, Association analysis of fibre traits in *Gossypium arboreum* accessions. *Plant Breed*. **127**: 173-179.

Lascape JM, Dessauw D, Rajab M, Noyer JL, Hau B 2007. Microsatellite diversity in tetraploid Gossypium germplasm:assembling a highly informative genotyping set of cotton SSRs. *Mol. Breed.* **19**(1): 45-58.

M. Altaf-Khan, Samina N. Qureshi, Sukumar Saha, Johnie N. Jenkins, Curt L. Brubaker and Oumesh U. K. Reddy 2006. Usefulness of SSR derived from Tetraploid *Gossypium spp*. for analyses of Diploid *Gossypium spp, Journal of Crop Improvement* **16**(1/2) (#31/32).

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583-590.

Neng Qian, Xing-Wei Zhang, Wang-Zhen Guo, Tian-Zhen Zhang 2009. Fine mapping of open-bud duplicate genes in homoelogous chromosomes of tetraploid cotton. *Euphytica* **165**:325-331.

Nguyen TB, Giband M, Brottier P, Risterucci AM, Lascape JM 2004. Wide coverage of the tetraploid cotton genome using newly developed microsatellite markers. *Theoretical Applied Genetics*, **109** : 167-175.

Osman A. Gutierrrez, David M. Stelly, Sukumar Saha, Johnie N. Jenkins, Jack C. McCarty Jr, Dwaine A. Raska, Brian E. Scheffler 2006. Integrative placement and orientation of non-reductant SSR loci in cotton linkage groups by deficiency analysis. *Molecular Breeding*, DOI 10.1007/s11032-009-9266-y.

Park YH, Alabady MS, Ulloa M, Sickler B, Wilkins TA, Yu J, Stelly DM, Kohel RJ, El-Shihy OM, Cantrell RG 2005. Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred line cotton population. *Mol Gen Genomics* **274**: 428–441.

Percival A. E., Wendel J. F, and Stewart J. M 1999. Taxonomy and germplasm resources. In: Smith, C. W., and J. T. Cothren (eds.) Cotton: origin, history, technology, and production. *John Wiley & Sons, Inc., New York,* pp. 33-64.

Powell W. 1996. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci*. **1**:215–222.

PZ Wang, Preetha, S and Raveendren T. S 2008. Molecular marker technology in cotton. *Biotechnology and Molecular Review* **3**(2), 032-045.

Qureshi, S.N., S. Saha, R.V. Kantety, and J.N. Jenkins. 2004. EST-SSR: A new class of genetic markers in cotton. *Journal of Cotton Science* **8**: 112-123.

Rakshit Aruna, Rakshit S., singh J., Chopra S. K.. Balyan H. S., Gupta P.K and Bhatt Shripad R. 2010. Association of AFLP and SSR markers with agronomic and fiber quality traits in *Gossypium hirsutum* L. *Journal of Genetics* **89**(2).

Rohlf F.J. 1998. NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, **2**(1), *User Guide. Exeter Software,* New York.

Saha Sukumar, Karaca Mehmet, Johnie N. Jenkins, Allan E. Zipf, Reddy O. Umesh K. and Kantety Ramesh V. 2003. Simple sequence repeats as useful resource to study transcribed genes of cotton. *Euphytica,* **130**:355-364.

Samina N. Qureshi, Sukumar Saha, Ramesh Kantely V and Jenkins J. N. 2004. EST-SSR, A New Class of Genetic Markers in Cotton. *The Journal of Cotton Science*, **8**:112-123.

Wangzhen, Caiping Cai, Changbiao, Zhiguo Han, Xianliang Song, Kai Wang, Xiaowei Niu, Cheng Wang, Keyu Lu, Ben Shi and Tianzhen Zhang 2007. A Microsatellite-Based, Gene-Rich Linkage Map Reveals Genome Structure, Function and Evolution in *Gossypium*. *Genetics,* **176**: 527-541.

ZG Han, Tianzhen Zhang 2004. EST-SSR Based Genetic Diversity Assessment of Xinjing Cultivars in Upland Cotton. *Journal of Genetics and Molecular Biology* **15**(171-181).

Zhiguo Han, Changbio Wang, Xianliang Song, Wangzhen Guo, Jinying Gou, Chunhong Li, Xiaoya Chen, Zhang Tianzhen 2006. Characteristics, development and mapping of *Gossypium hirsutum* derived EST-SSR in allotetraploid cotton. *Theor Appl Genet*, **112**: 430-439.