Comparative Analysis of Genetic Diversity among *Bt* Cotton Genotypes Using EST-SSR, ISSR and Morphological Markers

J. Ashraf[1], W. Malik[1*], M. Z. Iqbal[2], A. Ali Khan[3], A. Qayyum[1], E. Noor[1], M. A. Abid[1], H. M. Naseer Cheema[3], and M. Q. Ahmad[1]

ABSTRACT

Robust information of genetic diversity among *Bt* cotton genotypes is still lacking and availability of several marker systems has prompted us to compare their utility for the detection of genetic diversity. Here, we report the comparative efficiency of morphological and molecular markers (EST-SSRs and ISSRs) in determining the genetic diversity among 30 Pakistani *Bt* cotton genotypes. Three different dendrograms based on 20 EST-SSRs, 13 ISSRs and 20 morphological markers divided the 30 *Bt* cotton genotypes into five, six, and three clusters, respectively. EST-SSRs and ISSRs revealed 0.73-1.00 and 0.77-0.97 genetics similarity among *Bt* cotton genotypes which indicated low level of genetic diversity. Further population structure analysis showed extensive allelic admixture among *Bt* cotton cultivars and identified three (EST-SSRs) and six (ISSRs) subgroups. The MGHES-31 (EST-SSRs) and UBC-807 and UBC-815 (ISSRs) showed maximum values of Polymorphic Information Contents (*PIC*) and *Dj* and had low value of *Cj*. The comparison of two marker systems showed that EST-SSRs had high value of Expected heterozygosity (H_{ep}) and low value of Marker Index (MI) as compared to ISSRs. In conclusion, high level of genetic similarity among *Bt* cotton genotypes compel the plant breeders for the introduction of divers/exotic genotypes in their breeding program and marker discriminating indices could be a potential tool for selection of particular marker system to unveil the genetic difference.

Keywords: Confusion probability, Expected heterozygosity, Marker Index, Population structure analysis, Polymorphic information contents.

INTRODUCTION

The assessment of genetic diversity was traditionally based on the morphological and biochemical markers, which are greatly influenced by the environmental factors. The molecular marker systems have been proved to be powerful tools for the analysis of genetic diversity (Zhang *et al*., 2011). Among these, the PCR-based marker systems including Random Amplified Polymorphic DNA (RAPD) (Lu and Myers, 2002), Amplified Fragment Length Polymorphism (AFLP) (Alvarez and Wendel, 2006), Simple Sequence Repeats (SSR) (Zhu *et al*., 2003; Kantartzi *et al*., 2009), inter Simple Sequence Repeats (ISSR) (Reddy *et al*., 2002; Noormohammadi *et al*., 2013), and Single Nucleotide Polymorphism (SNP) (Deynze *et al*., 2009) have been used for analysis of genetic diversity among different crop species. Among different marker systems, SSR markers obtained from expressed sequence tags (EST-SSRs) are part of the transcribed regions of DNA, which are more conserved across genera and species (Park *et*

---

[1]Department of Plant Breeding and Genetics, Faculty of Agricultural Sciences and Technology, Bahauddin Zakariya University, Multan.
[2]Agricultural Biotechnology Research Institute, AARI, Faisalabad.
[3]Department of Plant Breeding and Genetics, University of Agriculture, Faisalabad.
*Corresponding author; e-mail: waqasmalik@bzu.edu.pk

al., 2005). The EST-SSRs are helpful for comparative mapping, comparative genomics, and evolutionary studies and have greater potential for transferability between the species than genomic SSRs (Zhu *et al.*, 2009). Inter Simple Sequence Repeat (ISSR) markers use SSR repeat-anchored primers, which mainly amplify the DNA segment present at the amplifiable distance in between two identical SSR repeat regions directed in the contrary directions (Reddy *et al.*, 2002). ISSRs are simple to use as prior knowledge of the target sequences flanking the repeat regions is not required (Nagaraju *et al.*, 2002).

In the presence of various molecular markers, comparison among different markers system is imperative to decide which marker system is most suitable for the issue being researched (Scariot *et al.*, 2007; Murty *et al.*, 2013). Reliability of the different markers for diversity analysis can be improved by calculating the various parameters like Polymorphic Information Contents (PIC), Confusion probability (Cj) and Discriminating power (Dj) (Belaj *et al.*, 2003; Kantartzi *et al.*, 2009; Sharma *et al.*, 2009a). Other parameters like Expected heterozygosity (H_{ep}), Effective multiplex ratio (E) and Marker Index (MI) can also be used to test the overall efficiency of any marker system (Maras *et al.*, 2008).

Cotton (*Gossypium spp*) is the most widely cultivated fiber crop around the globe. The economic impact of the cotton industry throughout the world is about $500 billion/year with an annual consumption of about 115 million bales of cotton fiber (Zhang *et al.*, 2013). However, insect pests, specially the bollworm, constantly threaten the productivity of cotton worldwide, which are not only increasing the cost of inputs *i.e.* pesticides, but also effecting the yield and quality. In this scenario, biotechnology has opened new avenue by developing a genetically engineered cotton variety having insect resistance due to presence of *Cry1Ac* gene isolated from the soil bacterium *Bacillus thuringiensis* (Bt) var *Kurstaki* (Tabashnik *et al.*, 2008).

In Pakistan, exotic *Bt* cotton cultivars were introduced through informal channels. Afterward, breeders developed *Bt* cotton cultivars through backcrossing using exotic *Bt* cotton having *Cry1Ac* gene (MON53 event) in local cotton genotypes (Ullah *et al.*, 2012). The *Bt* cotton is popular among the farmers because it seems to increases the yield and provide effective control of important pests, primarily bollworms (Zhang *et al.*, 2004). But, in *Bt* cotton genotypes susceptibility to abiotic and biotic stresses increased due to narrow genetic bases as only few selective cotton genotypes were used as recurrent parents in the development of *Bt* cotton varieties (Ullah *et al.*, 2012).

The study of genetic divergence among *Bt* cotton genotypes will not only offer the theoretical basis for cotton germplasm conservation but will also help in improving the tolerance against different biotic and abiotic stresses. Therefore, current study was conducted to determine the genetic similarity among 30 *Bt* cotton genotypes using different marker systems and to compare the efficiency between EST-SSRs and ISSRs for diversity analysis.

MATERIALS AND METHODS

Field Experiment

The experiment was conducted at the Department of Plant Breeding and Genetics, Bahauddin Zakariya University, Multan. A set of 30 upland *Bt* cotton genotypes (Table 1) was sown in a Randomized Complete Block Design (RCBD) with three replications. The row to row and plant to plant distances were 0.75 and 0.3 m, respectively. All necessary cultural practices were applied for good crop stand. Plants were fertilized with nitrogen (143 kg ha⁻¹), phosphorus (115 kg ha⁻¹) and potassium (125 kg ha⁻¹). Nitrogen was applied in three doses *i.e.* at sowing, first irrigation, and maximum flowering stage.

Table 1. Passport information of 30 *Bt* cotton genotypes.

Name	Origin	Name	Origin
VH-282	Cotton Research Station, Vehari	Bt-121	Neelum Seeds, Multan
FH-172	Cotton Research Institute, Faisalabad	Bt-886	Cotton Research Station, Multan
NIAB-Bt-1	NIAB, Faisalabad	Bt-3701	NIBGE, Faisalabad
AA-904	Ali Akbar Seeds, Pakistan	Bt-456	Cotton Research Station, Multan
MM-58	Islamia University, Bahawalpur	FH-113	Cotton Research Institute, Faisalabad
FH-118	Cotton Research Institute, Faisalabad	AGC-777	Allahdin Group of Companies, Pakistan
IR-NIBGE-4	NIBGE, Faisalabad	Sayban-202	Auriga group of Companies, Pakistan
AA-905	Ali Akbar Seeds, Pakistan	Tarzan-402	M/s Four Brothers, Lahore
CIM-602	Central Cotton Research Institute, Multan	Sitara-11M	Aziz Group, Pakistan
FH-142	Cotton Research Institute, Faisalabad	Sitara-12	Aziz Group, Pakistan
CIM-599	Central Cotton Research Institute, Multan	Auriga-213	Auriga Group of Companies, Pakistan
IUB-222	Islamia University, Bahawalpur	Sitara-10M	Aziz Group, Pakistan
IR-NIBGE-5	NIBGE, Faisalabad	BH-178	Cotton Research Station, Bahawalpur
CIM-595	Central Cotton Research Institute, Multan	BZU-75	Bahauddin Zakariya University, Multan
CIM-598	Central Cotton Research Institute, Multan	AA-919	Ali Akbar Seeds, Pakistan

Data Recording and Analysis of Genetic Diversity Based on Morphological Markers

Each genotype of *Bt* cotton was characterized by five plants in every replication. Genotypes were evaluated for different morpho-architectural traits *i.e.* plant height (cm), length of fruiting branches (cm), number of monopods plant^{-1}, number of sympods plant^{-1}, node of first fruiting branch, number of nodes plant^{-1}, angle of leaves, angle of sympodia, leaf area (cm^2), number of bolls plant^{-1}, plant shape, plant height to node ratio, fruit branch inter-nodal distance (cm), and ratio of fruit branch length to plant height. A sample of 20 mature bolls from upper, lower and middle parts of plant were picked. Single roller ginning machine was utilized to gin seed cotton samples and different traits *i.e.* ginning out-turn (GOT%), seed index (g), lint index (g), staple length (mm), micronaire value and staple strength (g tex^{-1}) were recorded. The data of all morpho-architectural and fiber related traits was subjected to unweighted pair group method to generate a dendrogram using XLSTAT 2014 statistical package.

DNA Extraction

The genomic DNA was extracted in the Genomics Lab at Department of Plant Breeding and Genetics, Bahauddin Zakariya University, Multan, from 2-3 young leaves of randomly selected plant of each genotype frozen in liquid nitrogen following CTAB (Cetyl Trimethyl Ammonium Bromide) method (Khan *et al.*, 2004). The isolated DNA was quantified using spectrophotometer (Implen Nanophotometer, Germany) for downstream application.

PCR for EST-SSR and ISSR Markers

Twenty EST-SSR and 20 ISSR primers were used to analyze the genetic diversity among *Bt* cotton genotypes. A 20 µL per sample reaction was carried out using 2µL DNA (30 ng µL^{-1}) as template, 2 µL of 10X PCR buffer (50 mM Tris, pH 8.3, 500 mM KCl), 2 µL of MgCl$_2$ (25 mM), 1 µL of dNTPs (10 mM), 1 µL of each forward and reverse primers (30 ng µL^{-1}), 0.2 µL (1 U) of *Taq* DNA polymerase (Fermentas, USA) and 10.8 µL of d$_3$H$_2$O (double-distilled deionized

water). While for ISSR markers reaction of volume 20 µL per sample was carried by adding 1 µL of DNA (30 ng µL^{-1}), 2 µL of 10X PCR buffer (50 mM Tris, pH 8.3, 500 mM KCl), 2 µL of MgCl$_2$ (25 mM), 0.5 µL of dNTPs (10 mM), 1 µL of primer (30 ng µL^{-1}), 0.2 µL (1 U) of *Taq* DNA polymerase (Fermentas, USA) and 13.3 µL of d$_3$H$_2$O (double-distilled deionized water). PCR for EST-SSRs and ISSRs were performed using thermal cycler (MyCycler, Bio Rad, USA) with the following temperature profile, initial denaturation step of 94°C for 5 minutes followed by 35 cycles of 30 seconds at 94°C, 30 seconds at 55°C and 1 minute at 72°C for EST-SSRs. However, for ISSRs, the 2nd step included 40 cycles of 1 minute at 94°C, 1 minute at 52 or 54°C and 2 minutes at 72°C. The final extension for both (EST-SSR and ISSRs) was done at 72°C for 10 minutes.

Gel Electrophoresis

EST-SSRs and ISSRs products were separated by the electrophoresis on 2.5 and 1.5% agarose gels, respectively. A volume of 4 µL of gel loading dye (6X) was added in 20 µL of PCR product. A sample of 10 µL of the reaction mixture was loaded in the gel submerged in 1X TBE (Tris/Borate/EDTA) buffer and electrophoresis was done at 80 Volt constant current. To calculate the amplicons size, 3 µL of the 50 bp DNA ladder (Fermentas, USA) with bands of known molecular weight were loaded on both sides of the gel for the EST-SSRs and 3 µL of the 1kb DNA ladder (Fermentas, USA) for the ISSR markers. Post staining was done by submerging the gels in 1 L solution of ethidium bromide (0.5 µg mL^{-1}) for 20 minutes. Gels were documented using Gel Documentation system (Photonyx, USA).

Gel Scoring and Data Analysis

The bands generated by the EST-SSRs and ISSRs were visually scored. Each band was considered as single allele and scored present as (1) or absent (0) for each of EST-SSR and ISSR loci to generate the binary data. Similarity matrix based on Nei's coefficients (Nei, 1972) were developed using binary data of EST-SSR and ISSR markers. The NTSyspc 2.0 software was used to construct dendrograms of each marker system using Unweighted Pair Group Method of Arithmetic means (UPGMA). The EST-SSR and ISSR marker datasets were separately used for inferring population structure with the model-based (Bayesian) cluster software STRUCTURE 2.33 version (Pritchard *et al.*, 2000).

Indices for the Discrimination and Comparison of EST-SSRs and ISSRs

PIC (Polymorphic Information Contents), *Cj* (Confusion probability) and *Dj* (Discriminating power) of each primer pair was calculated (Anderson *et al.*, 1993). Besides, the comparison of two marker systems (EST-SSRs and ISSRs) was made on the basis of indices following Maras *et al.* (2008).

RESULTS

Out of the 20 EST-SSR primers pairs, five primers revealed polymorphism (Figure 1) among the 30 *Bt* cotton genotypes, while the remaining 15 were monomorphic. Different EST-SSR primers produced the amplicons of different sizes ranging from 130-340 bp. Similarly, out of 20 ISSRs, 13 primers produced scoreable PCR products ranging from 280-2,200 bp. Among these 13 primers, 10 were found polymorphic (Figure 2) for collected *Bt* cotton genotypes (Table 2).

Similarity Matrix and Cluster Analysis for EST-SSRs

Twenty EST-SSRs produced 30 alleles and UPGMA (Unweighted Pair-Group Method using Arithmetic average) based
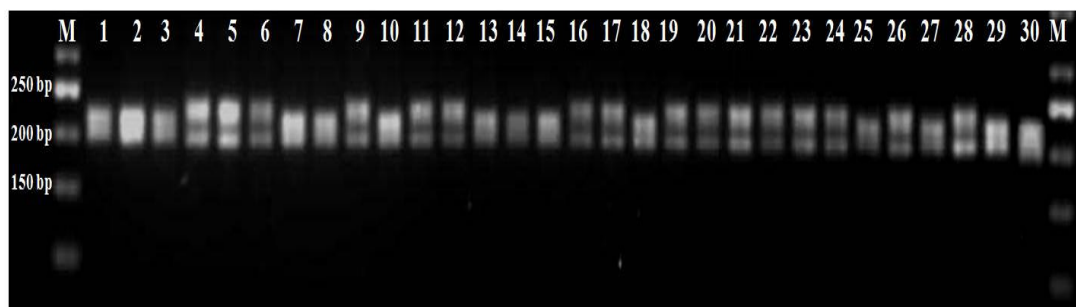
Figure 1. EST-SSRs generated from 30 *Bt* cotton genotypes using primer MGHES-40.
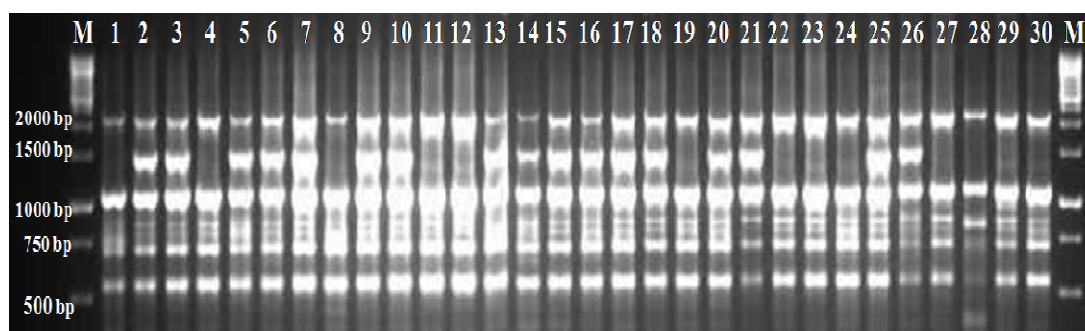

Figure 2. ISSRs generated from 30 *Bt* cotton genotypes using primer UBC-807.

dendrogram was generated. EST-SSRs based pair wise comparison depicted that the similarity among *Bt* cotton genotypes ranged from 1.00 to 0.73 with a mean of 0.88 (Table 3). Dendrogram was truncated at level of 0.91 genetic similarity and divided the 30 *Bt* cotton genotypes into five main clusters (1, 2, 3, 4 and 5) and two genotypes *i.e.* IUB-222 and BH-178 were found independent (Figure 3). Cluster 1 had 10 genotypes and was further divided into the 3 sub-clusters (1a, 1b and 1c). Sub-clusters 1a and 1c were comprised of four genotypes *i.e.* 1a (VH-282, FH172, IR-NIBGE-4, Bt-3701) and 1c (MM-58, FH-113, AGC-777, Sitara-11M), while, sub-cluster 1b had two genotypes (AA-904 and Sayban-202). Cluster 2 was comprised of two cotton genotypes *i.e.* NIAB-Bt-1 and FH-118. Cluster 3 had sub-cluster 3a (FH-142 and Sitara-12) and an independent genotype IR-NIBGE-5. Cluster 4 was divided into 3 sub-clusters 4a (AA-905, Sitara-10M, CIM-595), 4b (CIM-599 and Bt-886) and 4c (CIM-598, AA-919, BZU-75) and an independent genotype Auriga-213. Cluster 5 had a sub-

cluster 5a (Bt-121, Tarzan-402 and Bt-456) and an independent genotype CIM-602.

Similarity Matrix and Cluster Analysis for ISSRs

A total of 79 loci were identified by the 13 ISSR primers and UPGMA based cluster analysis was performed to generate the dendrogram. ISSRs based pair wise comparison indicated a mean of 0.88 genetic similarity among *Bt* cotton genotypes which ranged from 0.77 to 0.97 (Table 3). The dendrogram was truncated at the 0.91 similarity value which divided the 30 *Bt* cotton genotypes into six main clusters and three independent genotypes *i.e.* VH-282, CIM-595 and Sitara-11M (Figure 4). Cluster 1 was further sub divided into a sub-cluster 1a (AA-904 and FH-118) and two cotton genotypes *i.e.* FH-142 and MM-58 also existed in this cluster. Cluster 2 had two sub-clusters, 2a having three *Bt* cotton genotypes *i.e.* AA-905, CIM-602 and CIM-599, and 2b comprised of two cotton genotypes (IR-

Table 2. Marker discriminating indices of EST-SSRs and ISSRs.

EST-SSRs primer name	Annealing temperature (C)	Number of loci	Range of allele size (bp)	Polymorphic Information Content (PIC)	Confusion probability (Cj)	Discriminating power (Dj)	ISSRs primer name	Annealing Temperature (C)	Number of Loci	Range of allele size (bp)	Polymorphic information content (PIC)	Confusion probability (Cj)	Discriminating power (Dj)
MGHES-2	55	2	300-340	----	----	----	UBC-801	----	----	----	----	----	----
MGHES-6 ^{a}	55	2	175-190	0.749	0.485	0.757	UBC-802	----	----	----	----	----	----
MGHES-8	55	1	170	----	----	----	UBC-803	----	----	----	----	----	----
MGHES-10	55	2	130-160	----	----	----	UBC-804	----	----	----	----	----	----
MGHES-11a	55	1	200	----	----	----	UBC-805	----	----	----	----	----	----
MGHES-11b	55	1	220	----	----	----	UBC-806	----	----	----	----	----	----
MGHES-15	55	1	170	----	----	----	UBC-807 ^{a}	52	5	540-2050	0.491	0.492	0.754
MGHES-16	55	1	210	----	----	----	UBC-808	52	6	510-1380	----	----	----
MGHES-21	55	1	210	----	----	----	UBC-809 ^{a}	52	9	410-1520	0.315	0.674	0.663
MGHES-27	55	1	270	----	----	----	UBC-810 ^{a}	52	12	375-1600	0.293	0.697	0.652
MGHES-31 ^{a}	55	2	200-220	0.750	0.241	0.759	UBC-811 ^{a}	52	6	280-1240	0.320	0.669	0.666
MGHES-38	55	1	220	----	----	----	UBC-812 ^{a}	52	3	480-2200	0.272	0.718	0.641
MGHES-40 ^{a}	55	3	195-230	0.716	0.276	0.724	UBC-813 ^{a}	52	3	450-1800	0.228	0.764	0.618
MGHES-44	55	1	220	----	----	----	UBC-814	----	----	----	----	----	----
MGHES-49	55	1	210	----	----	----	UBC-815 ^{a}	52	4	400-1900	0.491	0.492	0.754
MGHES-62 ^{a}	55	3	140-160	0.749	0.485	0.757	UBC-816 ^{a}	52	9	600-1800	0.147	0.848	0.576
MGHES-63	55	1	185	----	----	----	UBC-817	52	4	1450-1850	----	----	----
MGHES-70 ^{a}	55	2	200-220	0.586	0.821	0.590	UBC-818 ^{a}	52	9	500-1400	0.276	0.714	0.643
MGHES-73	55	2	210-230	----	----	----	UBC-819 ^{a}	54	5	740-1750	0.064	0.933	0.067
MGHES-75	55	1	190	----	----	----	UBC-820	54	4	850-1550	----	----	----

^{a} Polymorphic primers.

Table 3. EST-SSRs and ISSRs based similarity matrix among 30 *Bt* cotton genotypes.[a]

Genotypes	VH-282	FH-172	NIAB-Bt-1	AA-904	MM-58	FH-118	IR-NIBGE-4	AA-905	CIM-602	FH-142	CIM-599	IUB-222	IR-NIBGE-5	CIM-595	CIM-598	Bt-121	Bt-886	Bt-3701	Bt-456	FH-113	AGC-777	Sayban-202	Tarzan-402	Sitara-11M	Sitara-12	Auriga-213	Sitara-10M	BH-178	BZU-75	AA-919
VH-282																														
FH-172	0.91	1.00																												
NIAB-Bt-1	0.89	0.92																												
AA-904	0.92	0.91	0.89																											
MM-58	0.90	0.91	0.90	0.90																										
FH-118	0.91	0.95	0.91	0.96	0.94																									
IR-NIBGE-4	0.87	0.94	0.84	0.87	0.87																									
AA-905	0.90	0.89	0.86	0.92	0.87																									
CIM-602	0.90	0.89	0.86	0.90	0.90																									
FH-142	0.90	0.91	0.86	0.92	0.92																									
CIM-599	0.90	0.86	0.87	0.90	0.92																									
IUB-222	0.86	0.90	0.85	0.91	0.87																									
IR-NIBGE-5	0.84	0.91	0.89	0.89	0.89																									
CIM-595	0.87	0.92	0.87	0.87	0.87																									
CIM-598	0.89	0.91	0.86	0.89	0.86																									
Bt-121	0.90	0.89	0.89	0.87	0.86																									
Bt-886	0.87	0.86	0.88	0.87	0.87																									
Bt-3701	0.85	0.86	0.86	0.90	0.86																									
Bt-456	0.87	0.90	0.89	0.92	0.85																									
FH-113	0.90	0.92	0.85	0.89	0.86																									
AGC-777	0.86	0.91	0.85	0.87	0.90																									
Sayban-202	0.90	0.92	0.85	0.91	0.86																									
Tarzan-402	0.91	0.85	0.86	0.86	0.85																									
Sitara-11M	0.84	0.91	0.80	0.84	0.86																									
Sitara-12	0.85	0.87	0.80	0.84	0.84																									
Auriga-213	0.84	0.85	0.80	0.87	0.84																									
Sitara-10M	0.81	0.85	0.84	0.86	0.84																									
BH-178	0.89	0.89	0.82		0.87																									
BZU-75	0.87	0.87			0.81																									
AA-919	0.86																													

[a] Values in bold and italic represents the similarity among *Bt* cotton genotypes revealed by ISSRs.
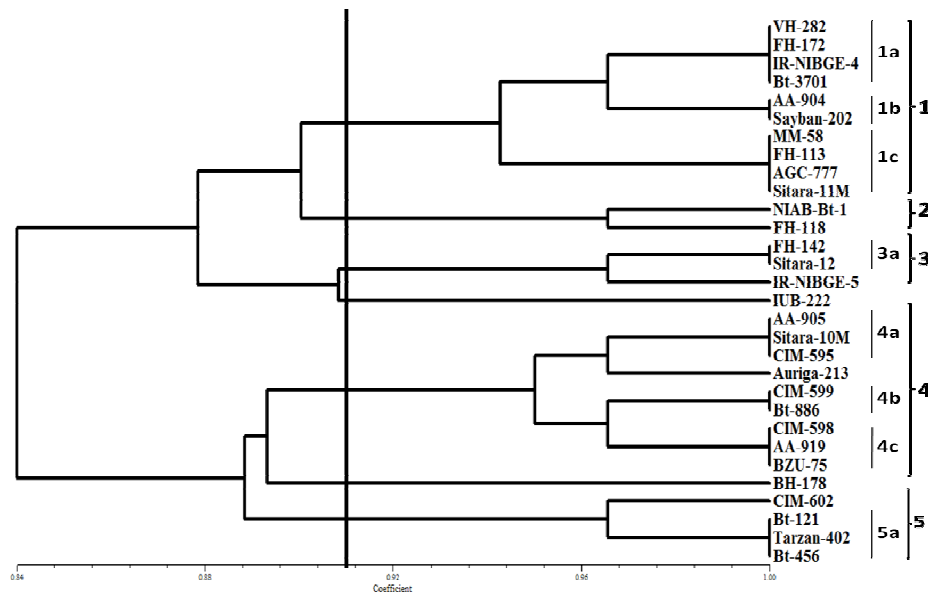
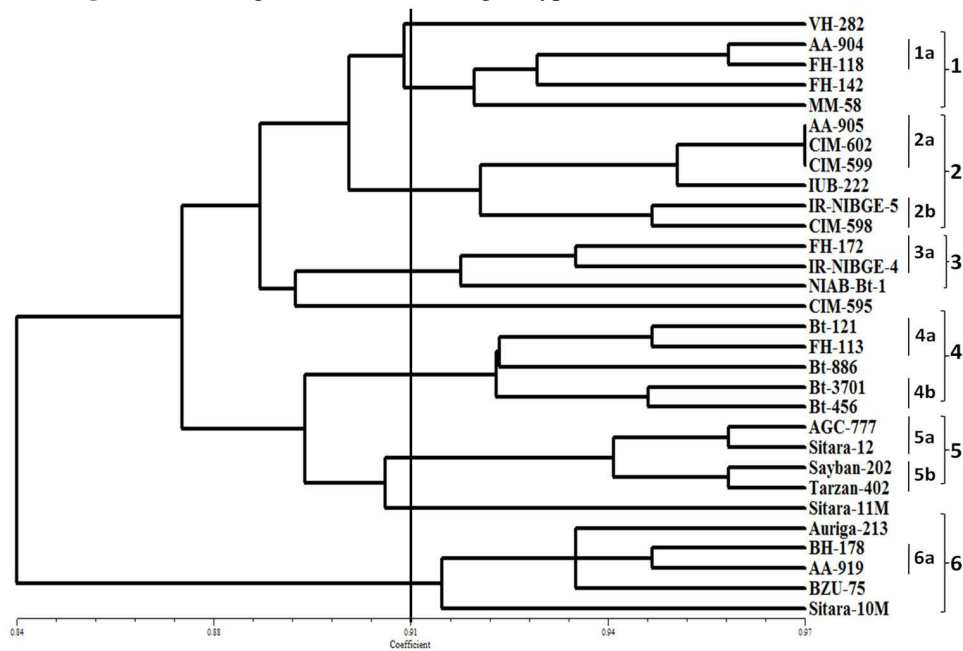**Figure 3.** Dendrogram of 30 *Bt* cotton genotypes based on EST-SSR markers.

Figure 4. Dendrogram of 30 *Bt* cotton genotypes based on ISSR markers.

NIBGE-5 and CIM-598). Cluster 3 had only one sub-cluster 3a (FH-172 and IR-NIBGE-4) and an independent cultivar NIAB-Bt-1. Dendrogram revealed that the clusters 4 and 5 were also divided into 2 sub-clusters *i.e.* 4a, 4b and 5a, 5b, respectively. Cluster 6 was grouped into one sub-cluster 6a having BH-178 and AA-919, while the three independent entries *i.e.* Sitara-10M, BZU-75 and Auriga-213 were also present in the 6th cluster.

Structure Analysis Using EST-SSR and ISSR Markers

Analysis of population structure using EST-SSRs depicted that the Logarithm of

the Data likelihood [Ln P(D)] on average continued to increase with increasing numbers of assumed subpopulations (K) from 2 to 10, however, results were not sufficient to delineate sub-populations. The ad hoc quantity based on the second order rate of change in the log probability (ΔK) showed a clear peak at *K*= 3 (Figures 5-a and -b), which confirmed that a K value of three was the most probable prediction for the number of subpopulations. The numbers of individual cotton varieties were 26.3% in K1, 41.1% in K2 and 32.6% in K3. The average genetic divergence (Fst) between subpopulations was high between K1-K2 (0.107) and K2-K3 (0.102), while genetic divergence was very low between K1-K3 (0.033). Similarly, genetic divergence based on *Fst* were higher among the individuals of K1 (0.08), followed by K3 (0.05).

In case of ISSR markers, similar findings were observed where Logarithm of the Data likelihood [Ln P(D)] failed to infer population structure. Contrastingly, relatively higher subpopulations were identified by ISSR markers using ad hoc quantity based on the second order rate of change in the log probability (ΔK) which showed a clear peak at *K*= 6 (Figure 5 c, d).

The numbers of individual cotton varieties ranged between 12.3 (K2) and 19.8% (K1) with an average of 16.6%. The average genetic divergence (Fst) between subpopulations ranged between 0.03 (K2-K4) and 0.14 (K3-K6), while genetic divergence was very low between K1-K3 (0.033). Similarly, genetic divergence based on *Fst* were highest among the individuals of K3 (0.023) and lowest among individuals of K2 (0.01).

Marker Discriminating Indices for EST-SSRs and ISSRs

The analysis of EST-SSRs revealed that the number of alleles per locus ranged from one to three with an average of 1.5 (Table 2). The Polymorphic Information Content (PIC) value varied from 0.586 to 0.750. The MGHES-31 depicted the highest *PIC* (0.750) and *Dj* (0.759) value and had lowest (0.241) value of *Cj*. The highest value of *Cj* (0.821) and lowest *Dj* (0.590) and *PIC* (0.586) value was observed for primer MGHES-70 (Table 2). Similarly, ISSR markers showed that the number of loci varied from 3 to 12, with an average of 6.08.
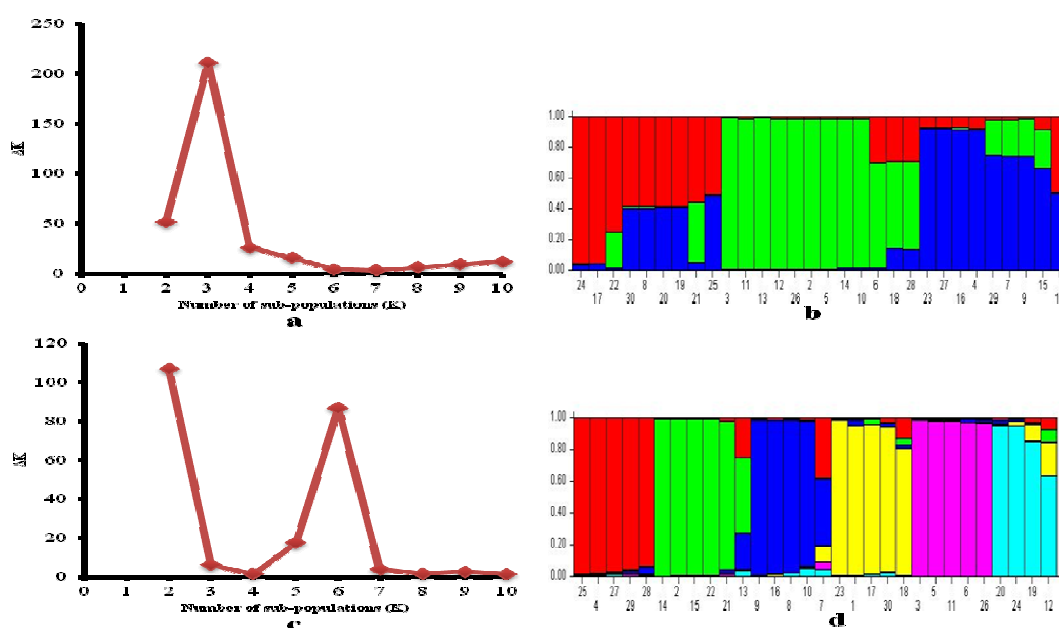

Figure 5. Population structure in *Bt* cotton varieties by EST-SSRs (a, b) and ISSRs (c, d).

Among 13 ISSR primers, 10 were polymorphic and their *PIC* value ranged from 0.064 to 0.491. The maximum *PIC* value (0.491) was estimated for the primers UBC-807 and UBC-815 followed by the UBC-811 (0.320) and UBC-809 (0.315), respectively. The *Cj* value varied from 0.492 to 0.933, and *Dj* value ranged from 0.067 to 0.754. The primer UBC-819 had the highest *Cj* (0.933) and lowest *Dj* (0.067) and PIC (0.064) estimates. The highest *Dj* value (0.754) and lowest *Cj* value (0.492) were calculated for primers UBC-807 and UBC-815 (Table 2).

Cluster Analysis Based on the Morphological Markers

Cluster analysis based on 20 morphological markers divided the 30 *Bt* cotton genotypes into three main clusters (1, 2, and 3) and two independent entries *i.e.* NIAB-Bt-1 and AA-919 when the

constructed dendrogram was truncated at the similarity level 0.97 (Figure 6). Cluster 1 contained the maximum number of genotypes, and it was subdivided into seven sub-clusters *i.e.*1a, 1b, 1c, 1d, 1e, 1f and 1g each contained two genotypes. Cluster 2 was grouped into two sub-clusters (2a, 2b) and also contained three independent genotypes (IUB-222, Sitara-10M and Bt-121). Sub cluster 2a contained Auriga-213 and Sitara-12 genotypes, and 2b having Sitara-11M and IR-NIBGE-4. Cluster 3 was classified into one sub-cluster (3a) having the FH-113 and Bt-456 genotypes and an independent entry FH-142 also existed in the 3rd cluster.

Comparison of EST-SSR and ISSR Marker Systems

The parameters used for the comparison of EST-SSR and ISSR markers showed the differences between these marker techniques. The results showed that the
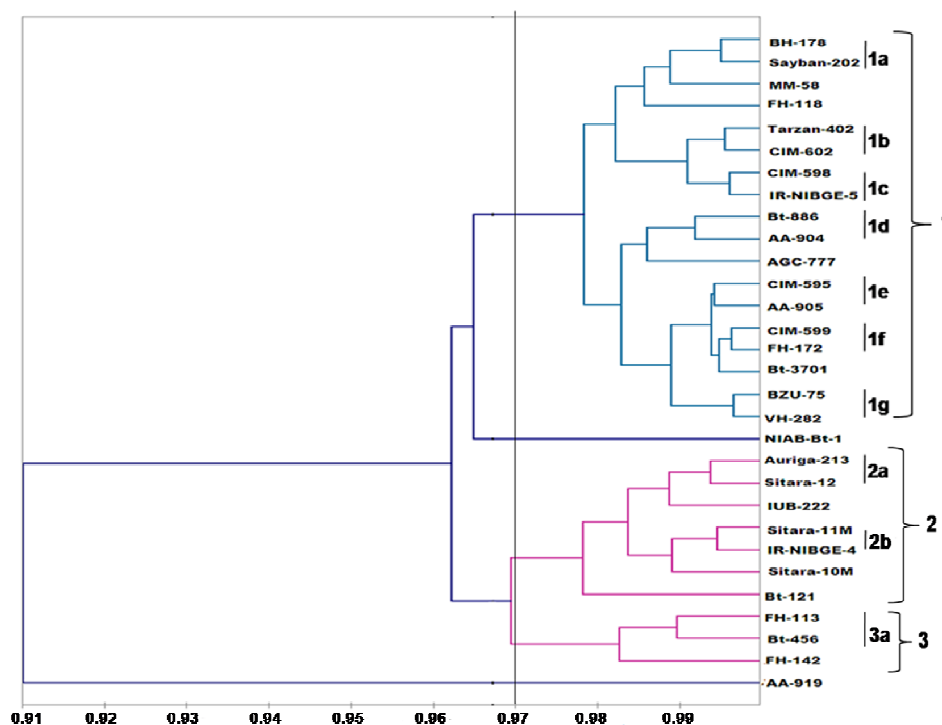

Figure 6. Dendrogram of 30 *Bt* cotton genotypes based on morphological markers.

number of polymorphic bands per assay unit were 2.69 and 0.40 for the ISSR and EST-SSR markers, respectively. The average number of the loci per assay unit was also high for the ISSR markers (6.08) as compared to the EST-SSR markers (1.5). The Expected heterozygosity (H_{ep}) for the ISSR and EST-SSR markers was 0.29 and 0.71, respectively. The estimates of the effective multiplex ratio (2.69) and marker index (0.78) were high for the ISSR markers, whereas the EST-SSR had the lower values of the effective multiplex ratio (0.40) and marker index (0.28), respectively (Table 4).

DISCUSSION

The success of any cotton breeding program depends on the level of genetic diversity available for desired alleles and the precise characterization of the variability prevailing in the existing germplasm (Sharma *et al*., 2009b). The morphological evaluation of germplasm is a traditional method for determination of evolutionary and pedigree relationship. With the development of DNA markers, it is essential that these techniques should be used for assessment of genetic diversity in germplasm collection to supplement and improve the classification based on morphological markers (Russell *et al*.,

1997). Among different DNA makers, inter simple sequence repeats (ISSR) is a reproducible and efficient marker system (Reddy *et al*., 2002). These (ISSRs) markers are useful for identifying the genetic polymorphism among genotypes by producing large number of the markers that target multiple SSR loci dispersed across the genome (Dalamu *et al*., 2012). In addition, EST-SSRs are valuable marker systems as they are rooted in sequences of functional gene, directly linked with transcribed genes, and have high transferability (Park *et al*., 2005). Keeping in view all these advantages of ISSRs and EST-SSRs, in the current study, they were used to reveal the genetic differences among *Bt* cotton genotypes.

Our study revealed high genetic similarity (73 to 100% by EST-SSRs and 77 to 97% by ISSRs) among 30 *Bt* cotton genotypes. Ullah *et al*. (2012) also reported high level of genetic similarity (0.90 to 0.98) among 19 *Bt* cotton genotypes. High genetic similarity among cotton genotypes was also reported in many earlier studies (Lukonge *et al*., 2007; Rahman *et al.,* 2008). The prevalence of high level of genetic similarity is likely due to monoculture of a small number of successful varieties and their recurrent use in *Bt* cotton breeding program (Van Esbroeck *et al*., 1998). In 1998, the introduction of *Bt* cotton (Bollgard) in Pakistan through informal sources and recurrent use of three *CLCuD* resistant lines *i.e* CP15/2,

Table 4. Indices for the comparison of ISSR and EST-SSR markers.

Indices with their abbreviations		Markers system	
		ISSR	EST-SSR
Number of assay unit	U	13	20
Number of Polymorphic bands	n_p	35	8
Number of monomorphic bands	n_{np}	44	22
Number of polymorphic bands per assay	n_p/U	2.69	0.40
Number of loci	L	79	30
Number of loci per assay unit	n_u	6.08	1.50
Expected heterozygosity of polymorphic loci	H_{ep}	0.29	0.71
Fraction of polymorphic bands	β	0.44	0.27
Effective multiplex ratio	E	2.69	0.40
Marker Index	MI	0.78	0.28

LRA5166, and CEDIX for development of new cotton cultivars also contributed in narrowing the genetic base of new *Bt* cotton genotypes (Rahman *et al*., 2002; Ullah *et al*., 2012).

In our study, dendrograms based on EST-SSR and ISSR markers showed the differences in total number of clusters and position of the genotypes within the clusters. This difference may be due to the fact that different markers identified distinctive regions of DNA variation within the genome (Dongre *et al*., 2004). Further, clustering based on the morphological markers was also different from the clustering based on molecular markers. The possible reason for these differences might be that most of the quantitative traits are polygenic and greatly affected by environmental factors. Besides, molecular markers are distributed throughout the genome (coding and non-coding region) and nearly 90% of the regions of the genome are not phenotypically expressed. Further, non-coding region is not accessible for phenotypic expression (Dalamu *et al*., 2012). The clustering of the dendrograms based on EST-SSR, ISSR and morphological markers exhibited that most of the clusters contained genotypes of both public and private sectors. This similarity might be due to repeated utilization of some outstanding gene pool, leading to narrow genetic base of available germplasm (Zhang *et al*., 2011). Moreover, breeders usually used the similar elite parental lines in their breeding program which leads to the close affinity of the *Bt* cotton cultivars.

Population structure is considered as a powerful tool for genetic divergence studies. The detection of different population structures using two different marker systems in the same population indicated that complex genetic structures exist, largely drifted by strong relatedness within some germplasm. The results are largely in agreement with the cluster analysis which also clustered the varieties in similar manners with slight differentiation. Recently, Noormohammadi *et al*. (2013)

revealed allelic admixture among cotton genotypes by STRUCTURE analysis and Q matrix showed close affinity between diploid genotypes. Thus, results of this study suggest that there is an important need to introduce more diversity into *Bt* cotton cultivars. Further, a regular genetic evaluation of the genetic resources will facilitate in reducing redundancy and developing a core germplasm collection.

Primer efficiency is an important parameter, especially in cases where primers are employed for genetic diversity studies. The highest *PIC* value for the MGHES-31 (0.750) in EST-SSRs and UBC-807 and UBC-815 (0.491) in ISSRs coupled with high *Dj* and low *Cj* value showed that these primers have strong ability to reveal allelic variation and both of these primers had more tendency to discriminate/distinguish between two genotypes (Sharma *et al*., 2009a). Moreover, comparison of the two marker systems based on the discriminating efficiency exhibited that the Expected heterozygosity (H_{ep}) of the EST-SSR markers was greater than the ISSR markers, suggesting that EST-SSRs had greater ability to show the allelic variation among *Bt* cotton genotypes (Belaj *et al*., 2003). It was also observed that the value of the Marker Index (MI) was greater for the ISSRs as compared to EST-SSRs. The high value of the marker index for ISSRs is the result of a high multiplex ratio (E= 2.69), which showed the distinctive nature of ISSR markers.

CONCLUSIONS

Conclusively, induction of diverse parental lines, wide hybridization, development of transgenics and *TILLING* (Targeted Induced Local Lesions In Genomes) populations could be a useful tool to cater the prevalence of narrow genetic base of Pakistani *Bt* cotton genotypes. Further, among marker discriminating indices, *PIC* (Polymorphic Information Contents) and *Dj* (Discriminating power) are believed to be

more reliable indices while selecting a marker or combination of markers for characterization of germplasm. Also, between marker systems, ISSRs could be suitable for studying genetic diversity among *Bt* cotton genotypes because of their ability to produce more number of bands per reaction, while the co-dominant nature of EST-SSRs with high value of Expected heterozygosity (H_{ep}) will make them more suitable for genome mapping.

REFERENCES

1. Alvarez, I. and Wendel, J. F. 2006. Cryptic Interspecific Introgression and Genetic Differentiation within *Gossypium aridum* (*Malvaceae*) and Its Relatives. *Evolution*, **60:** 505-517.
2. Anderson, J. A., Churchill, G. A., Autrique, J. E., Tanksley, S. D. and Sorrells, M. E. 1993. Optimizing Parental Selection for Genetic Linkage Maps. *Genome,* **36:** 181-186.
3. Belaj, A., Satovic, Z., Cipriani, G., Baldoni, L., Testolin, R., Rallo, L. and Trujillo, I. 2003. Comparative Study of the Discriminating Capacity of RAPD, AFLP and SSR Markers and of their Effectiveness in Establishing Genetic Relationships in Olive. *Theor. Appl. Genet.,* **107:** 736-744.
4. Dalamu Behera, T. K., Gaikwad, A. B., Saxena, S., Bharadwaj, C. and Munshi, A. D. 2012. Morphological and Molecular Analyses Define the Genetic Diversity of Asian Bitter Gourd (*Momordica charantia* L.). *Aus. J. Crop Sci.,* **6:** 261-267.
5. Deynze, A. V., Stoffel, K., Lee, M., Wilkins, T. A., Kozik, A., Cantrell, R. G., Yu, Z., Kohel, R. J. and Stelly, D. M. 2009. Sampling Nucleotide Diversity in Cotton. *BMC Plant Biol.,* **9:** 125.
6. Dongre, A., Parkhi, V. and Gahukar, S. 2004. Characterization of Cotton (*G. hirsutum*) Germplasm by ISSR, RAPD Markers and Agronomic Value. *Ind. J. Biotechnol.*, **3:** 388-393.
7. Kantartzi, S. K., Ulloa, M., Sacks, E. and Stewart, J. Mc. D. 2009. Assessing Genetic Diversity in *Gossypium arboreum L.* Cultivars Using Genomic and EST-derived Microsatellites. *Genetica,* **136:** 141-147.
8. Khan, I. A., Awan, F. S., Ahmad, A. and Khan, A. A. 2004. A Modified Mini Prep. Method for Economical and Rapid Extraction of Genomic DNA in Plants. *Plant Mol. Biol. Repor.,* **22:** 89a-89e.
9. Lu, H. J. and Myers, G. O. 2002. Genetic Relationships and Discrimination of Ten Influential Upland Cotton Varieties Using RAPD Markers. *Theor. Appl. Genet.,* **105:** 325-331.
10. Lukonge, K., Hersalman, L. and Labuschagne, M. T. 2007. Genetic Diversity of Tanzanian Cotton Revealed by AFLP Analysis. *Afr. Crop Sci. Conf. Proceed.,* **8:** 773-776.
11. Maras, M., Sustar-Vozlic, J., Javornik, B. and Meglic, V. 2008. The Efficiency of the AFLP and SSR Markers in Genetic Diversity Estimation and Gene Pool Classification of Common Bean (*Phaseolus vulgaris L.*). *Acta Agric. Slov.,* **91:** 87-96.
12. Murty, S. G., Patel, F., Punwar, B. S., Patel, M., Singh, A. S. and Fougat, R. S. 2013. Comparison of RAPD, ISSR, and DAMD Markers for Genetic Diversity Assessment between Accessions of *Jatropha curcas* L. and Its Related Species. *J. Agr. Sci. Tech.,* **15**: 1007-1022
13. Nagaraju, J., Kathirvel, M., Kumar, R. R., Siddiq, E. A. and Hasnain, S. E. 2002. Genetic Analysis of Traditional and Evolved Basmati and Non-Basmati Rice Varieties by Using Fluorescence-based ISSR-PCR and SSR Markers. *Pro. Nat. Acad. Sci.,* **99:** 5836-41.
14. Nei, M. 1972. Genetic Distance between Populations. *Am. Nat.,* **106:** 283-292.
15. Noormohammadi, Z., Farahani, Y. H. A., Sheidai, M., Baraki, S. G. and Alishah, Q. 2013. Genetic Diversity Analysis in Opal Cotton Hybrids Based on SSR, ISSR, and RAPD Markers. *Genet. Mol. Res.,* **12:** 256-269.
16. Noormohammadi, Z., Taghavi, E., Foroutanb, M., Sheidai, M. and Alishah, O. 2013. Structure Analysis of Genetic Diversity in Tetraploid and Diploid Cotton Genotypes. Int. *J. Plant Anim. Sci.,* **3:** 79-90.
17. Park, Y. H., Alabady, M. S., Ulloa, M., Sickler, B., Wilkins, T. A., Yu, J., Stelly, D. M., Kohel, R. J., El-Shihy, O. M. and Cantrell, R. G. 2005. Genetic Mapping of New Cotton Fiber Loci Using EST-derived Microsatellites in an Interspecific

Recombinant Inbred Line Cotton Population. *Mol. Genet. Geno.,* **274:** 428-441.

18. Pritchard, J. K., Stephens, M. and Donnelly, P. 2000. Inference of Population Structure Using Multilocus Genotype Data. *Genetic.,* **155:** 945-959.

19. Rahman, M., Hussain, D. and Zafar, Y. 2002. Estimation of Genetic Divergence among Elite Cotton Cultivars-Genotypes by DNA Fingerprinting Technology. *Crop Sci.,* **42:** 2137-2144.

20. Rahman, M., Yasmin, T., Tabbasam, N., Ullah, I., Asif, M. and Zafar, Y. 2008. Studying the Extent of Genetic Diversity among *Gossypium arboreum L.* Genotypes/Cultivars Using DNA Fingerprinting. *Genet. Resour. Crop Evol.,* **55:** 331-339.

21. Reddy, M. P., Sarla, N. and Siddiq, E. A. 2002. Inter Simple Sequence Repeat (ISSR) Polymorphism and Its Application in Plant Breeding. *Euphytica,* **128:** 9-17.

22. Russell, J. R., Fuller, J. D., Macaulay, M., Hatz, B. G., Jahoor, A., Powell, W. and Waugh, R. 1997. Direct Comparison of Levels of Genetic Variation Among Barley Accessions Detected by RFLPs, AFLPs, SSRs and RAPDs. *Theor. Appl. Genet.,* **95:** 714-722.

23. Scariot, V., De-keyser, E., Handa, T. and De-Riek, J. 2007. Comparative Study of the Discriminating Capacity and Effectiveness of AFLP, STMS and EST Markers in Assessing Genetic Relationships among Evergreen Azales. *Plant Breed.,* **126:** 207-212.

24. Sharma, S. N., Kumar, V. and Mathur, S. 2009a. Comparative Analysis of RAPD and ISSR Markers for Characterization of Sesame (*Sesamum indicum* L.) Genotypes. *J. Plant Biochem. Biotechnol.,* **18:** 37-43.

25. Sharma, M. V., Kantartzi, S. K. and Stewart, J. M. 2009b. Molecular Diversity and Polymorphism Information Content of Selected *Gossypium hirsutum* Accessions. *Summ. Ark. Cotton Res.,* 124-126.

26. Tabashnik, B. E., Gassmann, A. J., Crowder, D. W. and Carriere, Y. 2008. Insect Resistance to *Bt* Crops: Evidence Versus Theory. *Nat. Biotechnol.,* **26:** 199-202.

27. Ullah, I., Iram, A., Iqbal, M. Z., Nawaz, M., Hasni, S. M. and Jamil, S. 2012. Genetic Diversity Analysis of *Bt* Cotton Genotypes in Pakistan Using Simple Sequence Repeat Markers. *Genet. Mol. Res.,* **11:** 597-605.

28. Van Esbroeck, G. A., Bowman, D. T., Calhoun, D. S. and May, O. L. 1998. Changes in the Genetic Diversity of Cotton in the USA from 1970 to 1995. *Crop Sci.,* **38:** 33-37.

29. Zhang, B. H., Wang, Q. L., Wang, K. B., Zhou, D. and Liu, F. 2004. *Bt* Cotton in India. *Curr. Sci.,* **86:** 758-760.

30. Zhang, Y., Wang, X. F., Li, Z. K., Zhang, G. Y. and Ma, Z. Y. 2011. Assessing Genetic Diversity of Cotton Cultivars Using Genomic and Newly Developed Expressed Sequence Tag-Derived Microsatellite Markers. *Genet Mol. Res.,* **10:** 1462-1470.

31. Zhang, T., Qian, N., Zhu, X., Chen, H., Wang, S., Mei, H. and Zhang, Y. 2013. Variations and Transmission of QTL Alleles for Yield and Fiber Qualities in Upland Cotton Cultivars Developed in China. *PLoS ONE,* **8:** e57220.

32. Zhu, L. F., Zhang, X. L. and Nie, Y. C. 2003. Analysis of Genetic Diversity in Upland Cotton (*Gossypium hirsutum L.*) Cultivars from China and Foreign Countries by RAPDs and SSRs. *J. Agri. Biotechnol.,* **11:** 450-455.

33. Zhu, H. Y., Zhang, T. Z., Yang, L. M. and Guo, W. Z. 2009. EST-SSR Sequences Revealed the Relationship of *D*-Genome in Diploid and Tetraploid Species in *Gossypium. Plant Sci.,* **176:** 397-405.

تجزیه تطبیقی (مقایسه ای) تنوع ژنتیکی بین ژنوتیپ های پنبه *Bt* با استفاده از EST- SSR، ISSR، و نشانگرهای ریخت شناسی

ج. اشرف، و. مالک، م. ز. اقبال، ع. علی خان، ا. قیوم، ا. نور، م. ا. عبید، ح. م. ناصر چما، و م. ق. احمد

چکیده

عدم وجود اطلاعات موثق در باره تنوع ژنتیکی ژنوتیپ های پنبه *Bt* و در دسترس بودن چندین سامانه نشانگر ما را بر آن داشت که کار برد آن ها را برای بررسی و شناسایی تنوع ژنتیکی مقایسه کنیم. در این مقاله، کارآیی نسبی (تطبیقی) نشانگرهای ملکولی و ریخت شناسی (EST-SSR و ISSRs) در تعیین تنوع ژنتیکی میان ۳۰ ژنوتیپ پنبه *Bt* پاکستانی گزارش می شود. در این بررسی، سه عدد دندروگرام مبتنی بر ۲۰ عدد EST-SSR و۱۳ عدد ISSRs و ۲۰ نشانگر ریخت شناسی، ۳۰ ژنوتیپ پنبه *Bt* را به ترتیب در پنج، شش، و سه خوشه دسته بندی کردند. EST-SSR و ISSRs به ترتیب تشابه ژنتیکی بین ژنوتیپ هی پنبه *Bt* را به میزان ۱.۰۰-۰/۷۳ و ۰/۹۷-۰/۷۷ آشکار کردند که نشان از کم بودن سطح تنوع ژنتیکی بود. تحلیل بیشتر ساختار جمعیتی، اختلاط آللی گسترده ای بین کولتیوارهای پنبه *Bt* نشان داد و سه زیر-گروه در EST-SSR و شش زیر-گروه در ISSRs شناسایی شد. همچنین، MGHES-31 (EST-SSRs)وUBC-807 و (ISSRs) UBC-815 حد بیشینه *PIC* و *Dj* را داشتند ولی *Cj* آنها پایین بود. مقایسه این دو سامانه نشانگری نشان داد که EST-SSR در مقایسه با ISSRs دارای ناخالصی (*Hep*) بالا و شاخص نشانگری (*MI*) پایینی بود. جمعبندی کلی این است که تشابه زیاد در میان ژنوتیپ های پنبه *Bt* بهنژادگران را وادار به معرفی ژنوتیپ های متنوع و وارداتی در برنامه های اصلاح ژنتیکی می کند و نمایه های تمایز نشانگر ها مستعدا ابزار مناسبی برای انتخاب سامانه نشانگری ویژه ای برای آشکار ساختن تفاوت های ژنتیکی است.