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\(_e\)) 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; Kantartzis 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
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; Kantartzı et al., 2009; Sharma et al., 2009a). Other parameters like Expected heterozygosity (Hₑ), 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.

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

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
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 MgCl₂ (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 d²H₂O (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).
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 two cotton 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.

<table>
<thead>
<tr>
<th>EST-SSRs primer name</th>
<th>Annealing temperature (°C)</th>
<th>Number of loci</th>
<th>Range of allele size (bp)</th>
<th>Polymorphic Information Content (PIC)</th>
<th>Confusion probability (G1)</th>
<th>Discriminating power (Dj)</th>
<th>ISSRs primer name</th>
<th>Annealing Temperature (°C)</th>
<th>Number of Loci</th>
<th>Range of allele size (bp)</th>
<th>Polymorphic information content (PIC)</th>
<th>Confusion probability (G1)</th>
<th>Discriminating power (Dj)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGHES-2</td>
<td>55</td>
<td>2</td>
<td>300-340</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-801</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>MGHES-6*</td>
<td>55</td>
<td>2</td>
<td>175-190</td>
<td>0.749</td>
<td>0.485</td>
<td>0.757</td>
<td>UBC-802</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>MGHES-8</td>
<td>55</td>
<td>1</td>
<td>170</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-803</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>MGHES-10</td>
<td>55</td>
<td>2</td>
<td>130-160</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-804</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>MGHES-11a</td>
<td>55</td>
<td>1</td>
<td>200</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-805</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>MGHES-11b</td>
<td>55</td>
<td>1</td>
<td>220</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-806</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>MGHES-15</td>
<td>55</td>
<td>1</td>
<td>170</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-807*</td>
<td>52</td>
<td>5</td>
<td>540-2050</td>
<td>0.491</td>
<td>0.492</td>
<td>0.754</td>
</tr>
<tr>
<td>MGHES-16</td>
<td>55</td>
<td>1</td>
<td>210</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-808</td>
<td>52</td>
<td>6</td>
<td>510-1380</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>MGHES-21</td>
<td>55</td>
<td>1</td>
<td>210</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-809*</td>
<td>52</td>
<td>9</td>
<td>410-1520</td>
<td>0.315</td>
<td>0.674</td>
<td>0.663</td>
</tr>
<tr>
<td>MGHES-27</td>
<td>55</td>
<td>1</td>
<td>270</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-810*</td>
<td>52</td>
<td>12</td>
<td>375-1600</td>
<td>0.293</td>
<td>0.697</td>
<td>0.652</td>
</tr>
<tr>
<td>MGHES-31*</td>
<td>55</td>
<td>2</td>
<td>200-220</td>
<td>0.750</td>
<td>0.241</td>
<td>0.759</td>
<td>UBC-811*</td>
<td>52</td>
<td>6</td>
<td>280-1240</td>
<td>0.320</td>
<td>0.669</td>
<td>0.666</td>
</tr>
<tr>
<td>MGHES-38</td>
<td>55</td>
<td>1</td>
<td>220</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-812*</td>
<td>52</td>
<td>3</td>
<td>480-2200</td>
<td>0.272</td>
<td>0.718</td>
<td>0.641</td>
</tr>
<tr>
<td>MGHES-40*</td>
<td>55</td>
<td>3</td>
<td>195-230</td>
<td>0.716</td>
<td>0.276</td>
<td>0.724</td>
<td>UBC-813*</td>
<td>52</td>
<td>3</td>
<td>450-1800</td>
<td>0.228</td>
<td>0.764</td>
<td>0.638</td>
</tr>
<tr>
<td>MGHES-44</td>
<td>55</td>
<td>1</td>
<td>220</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-814</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>MGHES-49</td>
<td>55</td>
<td>1</td>
<td>210</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-815*</td>
<td>52</td>
<td>4</td>
<td>400-1900</td>
<td>0.491</td>
<td>0.492</td>
<td>0.754</td>
</tr>
<tr>
<td>MGHES-62*</td>
<td>55</td>
<td>3</td>
<td>140-160</td>
<td>0.749</td>
<td>0.485</td>
<td>0.757</td>
<td>UBC-816*</td>
<td>52</td>
<td>9</td>
<td>600-1800</td>
<td>0.147</td>
<td>0.848</td>
<td>0.576</td>
</tr>
<tr>
<td>MGHES-63</td>
<td>55</td>
<td>1</td>
<td>185</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-817</td>
<td>52</td>
<td>4</td>
<td>1450-</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>MGHES-70*</td>
<td>55</td>
<td>2</td>
<td>200-220</td>
<td>0.586</td>
<td>0.821</td>
<td>0.590</td>
<td>UBC-818*</td>
<td>52</td>
<td>9</td>
<td>500-1400</td>
<td>0.276</td>
<td>0.714</td>
<td>0.643</td>
</tr>
<tr>
<td>MGHES-73</td>
<td>55</td>
<td>2</td>
<td>210-230</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-819*</td>
<td>54</td>
<td>5</td>
<td>740-1750</td>
<td>0.064</td>
<td>0.933</td>
<td>0.067</td>
</tr>
<tr>
<td>MGHES-75</td>
<td>55</td>
<td>1</td>
<td>190</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>UBC-820</td>
<td>54</td>
<td>4</td>
<td>850-1550</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

* Polymorphic primers.
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 $C_j$ value varied from 0.492 to 0.933, and $D_j$ value ranged from 0.067 to 0.754. The primer UBC-819 had the highest $C_j$ (0.933) and lowest $D_j$ (0.067) and $PIC$ (0.064) estimates. The highest $D_j$ value (0.754) and lowest $C_j$ 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

![Dendrogram](image)

**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,
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 (Hep) 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 ($H_e$) will make them more suitable for genome mapping.

REFERENCES


تجزیه تطبیقی (مقاومت یافته) تنوع زننیکی بين Znوتیپ های پنهان Bt با استفاده از EST-SSR و نشانگرهای ریخت‌شناسی.

چکیده

بدون وجود اطلاعات موثر در باره تنوع زننیکی Znوتیپ های پنهان Bt و در دسترس بودن جدید سامانه نشانگر ما را بر آن داشت که کار بردن آن ها را برای بررسی و نشان‌آوری تنوع زننیکی مقاومه کنیم. در این مقاله، کارآیی نسبی (تطبیقی) نشانگرهای ملکولی و ریخت‌شناسی (ISSRs) EST-SSR (I) تنوع زننیکی میان 30 Znوتیپ پنهان Bt با 30 نشانگر ریخت‌شناسی، 30 Znوتیپ پنهان Bt با 30 عدد EST-SSR و 13 عدد ISSRs به ترتیب در برنامه‌های دام و سه زیر-گروه در EST-SSR و در نهایت نشانگر Hep و شناسایی شد. همچنین، Dj و PIC و UBC-815 و UBC-807 و MGHES-31 (EST-SSRs) در مقایسه EST-SSR و با دادسته‌ای (ML) یاین کردن که نشان از کم بودن مطلب Bt تنوع زننیکی بود. تحلیل بیشتر ساختار جمعیت، اختلاف آلفای گسترده ای بین کوتین‌های پنهان Bt این ترتیب نشان داد و سه زیر-گروه در EST-SSR و در نهایت نشانگر Hep و شناسایی شد. همچنین، Dj و PIC و UBC-815 و UBC-807 و MGHES-31 (EST-SSRs) در مقایسه EST-SSR و با دادسته‌ای (ML) یاین کردن که نشان از کم بودن مطلب Bt تنوع زننیکی بود. تحلیل بیشتر ساختار جمعیت، اختلاف آلفای گسترده ای بین کوتین‌های پنهان Bt این ترتیب نشان داد و سه زیر-گروه در EST-SSR و در نهایت نشانگر Hep و شناسایی شد. همچنین، Dj و PIC و UBC-815 و UBC-807 و MGHES-31 (EST-SSRs) در مقایسه EST-SSR و با دادسته‌ای (ML) یاین کردن که نشان از کم بودن مطلب Bt تنوع زننیکی بود. تحلیل بیشتر ساختار جمعیت، اختلاف آلفای گسترده ای بین کوتین‌های پنهان Bt این ترتیب نشان داد و سه زیر-گروه در EST-SSR و در نهایت نشانگر Hep و شناسایی شد. همچنین، Dj و PIC و UBC-815 و UBC-807 و MGHES-31 (EST-SSRs) در مقایسه EST-SSR و با دادسته‌ای (ML) یاین کردن که نشان از کم بودن مطلب Bt تنوع زننیکی بود. تحلیل بیشتر ساختار جمعیت، اختلاف آلفای گسترده ای بین کوتین‌های پنهان Bt این ترتیب نشان داد و سه زیر-گروه در EST-SSR و در نهایت نشانگر Hep و شناسایی شد. همچنین، Dj و PIC و UBC-815 و UBC-807 و MGHES-31 (EST-SSRs) در مقایسه EST-SSR و با دادسته‌ای (ML) یاین کردن که نشان از کم بودن M

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

چما. و. م. احمد