Research Article

Exploring Genetic Diversity in Cotton Genotypes Using EST-SSR and ISSR Markers: A Comparative Study

Jaweria Iqbal[1], Muhammad Tanveer Altaf[2], Muhammad Faheem Jan[3], Waqas Raza[4*], Waqas Liaqat[5], Ikram ul Haq[6], Amna Jamil[7], Sameer Ahmed[1], Amjad Ali[2] and Arif Mehmood[8]

[1]Department of Plant Breeding and Genetics, Faculty of Agricultural Sciences and Technology, Bahauddin Zakariya University, Bosan Road, 60800 Multan Pakistan; [2]Faculty of Agricultural Sciences and Technologies, Sivas University of Science and Technology, Sivas 58140, Türkiye; [3]College of Agriculture, Northeast Agricultural University, Harbin 150030, China; [4]Department of Plant Pathology, Nanjing Agricultural University, Nanjing, China; [5]Department of Field Crops, Faculty of Agriculture, Institute of Natural and Applied Sciences, Çukurova University, Adana, 01330, Türkiye; [6]Department of Plant Breeding and Genetics College of Agriculture, University of Sargodha, Sargodha, Pakistan; [7]Department of Horticulture, MNS University of Agriculture, Multan, Pakistan; [8]Department of Entomology, College of Agriculture, University of Sargodha, Punjab, Pakistan

Abstract | Cotton (*Gossypium hirsutum* L.) is a significant global crop and a vital raw material for industries. Studying existing cotton germplasm is crucial for discovering new genetic resources for future breeding. In this study, the effectiveness of EST-SSRs and ISSRs markers were compared for assessing genetic variation among 45 Pakistani Bt and non-Bt cotton varieties. ISSR and EST-SSR primers yielded 108 and 28 loci, respectively. The polymorphism was found 32.40% for ISSR, while 89.28% was recorded for SSR primers. Cluster analysis revealed a high level of genetic similarity for ISSR (average 0.92) and EST-SSR (average 0.85) among cotton genotypes. The mean polymorphic information content (PIC) value was 0.25 for ISSR, whereas it was recorded 0.48 for EST-SSRs. Confusion probability (Cj) exhibited a negative association with discriminating power (Dj), while Dj displayed a positive association with PIC. Marker discriminating statistics showed that EST-SSRs have a high expected heterozygosity of polymorphic loci (Hep) as compared to ISSR, along with a higher marker index value (MI). The effective multiplex ratio for ISSRs (1.40) was greater than EST-SSRs (1.12). The structural analysis revealed 6 sub-clusters for EST-SSRs and 4 sub-clusters for ISSRs. This phylogenetic study is crucial for identifying promising genotypes for breeding programs, especially given the limited genetic diversity in cotton breeding. The study showed that Bt cotton genotypes share a high genetic similarity, emphasizing the need for introducing diverse or exotic genotypes into breeding programs to enrich genetic diversity. Additionally, marker-discriminating indices can aid in selecting effective markers to assess genetic variation, facilitating the development of improved cotton varieties with desired traits.

Received | September 11, 2023; **Accepted** | October 07, 2023; **Published** | November 01, 2023
***Correspondence** | Waqas Raza, Department of Plant Pathology, Nanjing Agricultural University, Nanjing, China; **Email:** waqasraza61@yahoo.com
Citation | Iqbal, J., M.T. Altaf, M.F. Jan,W. Raza, W. Liaqat, I. Haq, A. Jamil, S. Ahmed, A. Ali and A. Mehmood 2023. Exploring genetic diversity in cotton genotypes using EST-SSR and ISSR markers: A comparative study. *Sarhad Journal of Agriculture,* 39(4): 800-814.
DOI | https://dx.doi.org/10.17582/journal.sja/2023/39.4.800.814
Keywords | Cotton, Polymorphic information content, Genetic diversity, Confusion probability, Population structure

Introduction

Cotton belongs to the Malvaceae family and the Gossypium genus, which consists of approximately 52 species (Wang *et al.*, 2021), found in various regions with semi-arid, tropical and subtropical climates. Among these species, there are currently four cultivated cotton types. Two of them are diploids with a chromosome number of 26 (2n = 2x = 26), and the other two are allotetraploids with a chromosome number of 52 (2n = 4x = 52). Cotton (Gossypium spp.) exhibits eight different genome types, denoted as A, B, C, D, E, F, G, and K. The cultivated species *Gossypium herbaceum* and *Gossypium arboreum* are examples of diploid plants with the genetic composition AA. On the other hand, *Gossypium hirsutum* and *Gossypium barbadense* are allotetraploid species with the genetic composition AADD (Jabran *et al.*, 2019; Saleem *et al.*, 2020). *G. hirsutum* represents approximately 96-97% of the total cotton cultivation worldwide, while *G. barbadense* accounts for 2-3%. G. *herbaceum* and *G. arboreum* are grown on only 1% of the global cotton farmland (Jabran *et al.*, 2019; Basal *et al.*, 2019).

Cotton is a highly valuable crop with significant economic importance worldwide. It is widely used in the textile industry as a primary raw material, making it one of the most popular materials for clothing and other textile products (Majumdar *et al.*, 2019). However, cotton offers more than just textile production. Its seeds are used to extract oil and seed cake for the feed industries, while the stalks find application in the paper industry, making it a versatile plant (Munir *et al.*, 2020). In recent years, cottonseed oil has gained popularity as an alternative to petroleum-based fuels, contributing to biodiesel production (Sharma *et al.*, 2020). The increased demand for cotton and its by-products has led to an annual global production of approximately 27 million tons (FAOSTAT, 2020-21). Major cotton-producing countries include India, China, the United States, Brazil, Pakistan, and Uzbekistan (Tokel and Erkencioglu, 2021). The global population is projected to reach 11 billion by 2050 and has resulted in a higher demand for fuel, food, fiber and feed (Hayat *et al.*, 2020). Cotton fiber is a widely recognized and prominent textile fiber in the textile industry. It has a significant annual economic impact of $600 billion globally (Khan *et al.*, 2020).

It is critical to enhance agricultural production by at least two to three times under increasing population and changing climate. The negative impacts of climate change, resulting in various abiotic and biotic stresses have led to a decrease in global agricultural output. These issues highlight the vital need to improve crop productivity. Additionally, considering resource constraints, it is essential to effectively explore and utilize the existing genetic diversity (Hayat *et al.*, 2020).

Genetic variation in cotton is essential for sustainable development, as it allows for the creation of new gene combinations and helps in choosing the right parent plants for breeding programs. The initial step in creating better plant materials and crop varieties involves evaluating genetic diversity and the connections between various genetic resources. These resources are considered valuable sources for developing new crop varieties (Han *et al.*, 2022) and are essential for the success of crop enhancement efforts. The data related to genetic diversity is indispensable for enhancing crops and developing new varieties (Bakhsh *et al.*, 2019; Swarup *et al.*, 2021).

Molecular markers play a universally reputed and prominent role in plant breeding studies, serving various purposes (Nadeem *et al.*, 2018). One crucial application of molecular markers is assessing genetic diversity (Nadeem *et al.*, 2018; Ali *et al.*, 2019; Gautam *et al.*, 2022; Tahir *et al.*, 2022). In cotton, various DNA marker systems have been employed to identify genetic diversity including RFLP (Yu *et al.*, 1997), RAPD (Bukhari *et al.*, 2021), AFLP (Jian *et al.*, 2017), SSR (Yu *et al.*, 2012; Tang *et al.*, 2015), ISSR (Ashraf *et al.*, 2016) and iPBS-retrotransposons (Baran *et al.*, 2023). SSR markers are considered promising molecular markers in various applications, primarily because of their reliable reproducibility, co-dominant, ease of use and high polymorphism (Xiong *et al.*, 2021). These markers have been extensively used in cotton research for DNA fingerprinting, analyzing genetic diversity, facilitating marker-assisted selection, constructing molecular maps, and identifying quantitative trait loci (QTL) (Kumar *et al.*, 2021; Wang *et al.*, 2018). The utilization of SSR markers has significantly contributed to the conservation of cotton germplasm resources and the advancement of cotton varieties through genetic improvement. The use of ISSR markers has proven to be highly effective in estimating genetic diversity in various studies (Bilval *et al.*, 2017; Jamil *et al.*,

2022). In cotton, ISSR serves as an informative and simple genetic marker system, capable of detecting both inter- and intra specific variation (Farahani *et al*., 2018; Kahodariya *et al*., 2015). ISSR markers are known for their strong reliability, informative nature, rapidity, and efficiency compared to other systems in their ability to distinguish between genetic variations (Abdellatif *et al*., 2012).

The comparison of different marker systems is imperative in the presence of different molecular markers to choose which marker system is best suited to the issue being investigated (Murty *et al*., 2013). Assessing the various parameters such as Confusion probability (Cj), Discriminating power (Dj) and Polymorphic Information Contents (PIC) (Sharma *et al*., 2009; Kantartzi *et al*., 2009), can enhance the reliability of several markers for diversity assessment. Other factors such as effective multiple ratio (E), Marker index (MI) and Expected Heterozygosity (Hep) may also be used to examine the overall efficiency. This research aimed to explore the genetic diversity of cotton germplasm using the EST-SSR and ISSR marker system, which will contribute to determining population structure and easing the task of cotton breeders particularly, in the context of abiotic stresses. We also compared the effectiveness of ISSR and EST-SSR markers for assessing the genetic diversity of cotton germplasm.

Materials and Methods

The study was performed at the Plant Breeding and Genetics Department of Bahauddin Zakariya University in Multan, Pakistan. Forty-five accessions of Bt and non-Bt cotton germplasm were used in this research and were collected from different research stations. The names and sources of these 45 cotton genotypes are listed in Table 1.

Table 1: *Passport information of studied genotypes of Bt and non-Bt cotton.*

Name	Origin	Name	Origin
BT-A-1	Central Cotton Research Institute, Multan	Crystal-1	Four brother
AGC-555	Allahdin Group of Companies, Pakistan	Eagle-1	Four brother
NS-131	Neelum Seed Company, Multan	CIM-602	Central Cotton Research Institute, Multan
NS-141	Neelum Seed Company, Multan	6/13	Cotton research station, Multan
FH-118	Cotton Research Institute, Faisalabad	MNH-1026	Cotton research station, Multan
CIM-116	Central Cotton Research Institute, Multan	CIM-600	Cotton research station, Multan
CIM-632	Central Cotton Research Institute, Multan	MNH-1020	Cotton research station, Multan
Cyto-178	Central Cotton Research Institute, Multan	GH-Baghdadi	Central cotton research institute, ghotki
IR-3701	National Institute of Biotechnology and Genetic Engineering, Faisalabad	BS-2015	Bahawalpur Research Station
Sitara-008	Aziz Group, Pakistan	AGC-999	Allahdin Group of Companies, Pakistan
CEMB-33	Center of Excellence for Molecular Biology Punjab University, Lahore	NIYAB-878	NIBBGE, Faisalabad
IUB-222	Islamia University, Bahawalpur	CYTO-124	Central Cotton Research Institute, Multan
AA-802	Ali Akbar Seeds, Pakistan	Shahkar	Warbel
FH-113	Cotton Research Station, Multan	BS-80	Bahawalpur Research Station
TARZAN-2	M/s Four Brothers, Lahore	SH-Buraq	Petron
VH-305	Cotton Research Station, Vehari	CEMB-66	CEMB Lahore
VH-363	Cotton Research Station, Vehari	Tarzan-1	M/s Four Brothers, Lahore
BH-172	Cotton Research Station, Bahawalpur	GH-Mubarak	Central cotton research institute, ghotki
BH-1999	Cotton Research Station, Bahawalpur	FH-Kahkashan	Cotton Research Station, Faisalabad
FH-Lalazar	Ayub Agriculture Research Institute, Faisalabad	MNH-886	Cotton Research Station, Multan
NBBGE-8	National Institute of Biotechnology and Genetic Engeering, Faisalabad	CIM-622	Central Cotton Research Institute, Multan
IUB-213	Islamia University, Bahawalpur	CEMB-55	Center of Excellence for Molecular Biology Punjab University, Lahore
Nibge-7	NIBGE, Faisalabad		

Extraction of genomic DNA

Seeds were planted in small plastic pots in a greenhouse. After three weeks of sprouting, we collected fresh and new leaf samples from each genotype. The leaves were carefully taken from the plants, washed with distilled water, and placed in tubes, at -80 °C until DNA isolation started. The genomic DNA extraction was carried out using cetyl trimethyl ammonium bromide (CTAB) method (Khan *et al*., 2004), with slight modifications. DNA quantification was done using a spectrophotometer (Implen Nano photometer, Germany). To assess the quality of the DNA obtained was also confirmed by using 1% agarose gel. The concentration of the DNA was adjusted to 30 ng µL-1 and stored in the freezer (-20°C) for further use in Polymerase Chain Reaction (PCR) amplification.

PCR to amplify molecular markers

A set of 25 ISSR and 25 EST-SSR primers was taken to examine the genetic diversity within the studied cotton genotypes. A 20µL reaction was used for PCR amplification for EST-SSR primers. This mixture contained 2 µL (30ng/µL) of DNA as the template, 0.5 µL of dNTPs (10 mM), 2 µL of 10X PCR buffer (composed of 50 mM Tris, pH 8.3, and 500 mM KCl), 1 µL of each forward and reverse primers (30 ng µL-1), 2 µL of MgCl2 (25 mM), 0.2 µL (1 U) of Taq DNA polymerase from Fermentas (USA), and 11.3 µL of d3H2O (double-distilled deionized water). Similarly, for the ISSR primers, a 20 µL PCR reaction volume was prepared by including 1 µL of DNA (30 ng µL-1), 2 µL of 10X PCR buffer (50 mM Tris, pH 8.3, and 500 mM KCl), 0.5 µL of dNTPs (10 mM), 1 µL of primer (30 ng µL-1), 2 µL of MgCl2 (25 mM), 13.3 µL of d3H2O (double-distilled deionized water) and 0.2 µL (1 U) of Taq DNA polymerase from Fermentas (USA). The PCR was performed using the following temperature profile. Initially, the DNA denaturation step was carried out at 94°C for 5 minutes. This was followed by 35 cycles of amplification for EST-SSRs, consisting of 30 seconds at 94°C, 30 seconds at 55°C, and 1 minute at 72°C. However, for ISSRs, the second step involved 40 cycles of amplification, with each cycle comprising 1 minute at 94°C, 1 minute at 52 or 54°C, and 2 minutes at 72°C. A final extension was kept at 72°C for 10 minutes for both ISSRs and EST-SSRs.

The amplified DNA fragments obtained from ISSR were separated by electrophoresis. Electrophoresis was performed using a 1.5% agarose gel in 1× TBE buffer, with a voltage of 80V applied for around 2 hours. Due to the smaller amplicon size of EST-SSRs in comparison to ISSRs, the results obtained through gel electrophoresis were not sufficient to take satisfactory findings. Therefore, polyacrylamide gel electrophoresis (PAGE) was utilized for EST-SSRs. To prepare the PAGE gel, we used 12% agarose with a gel solution volume of 22.5 milliliters, 50 microliters of Temed, 700 microliters of 10% APS solution, and 52.5 milliliters of 1X TBE buffer. After the gel polymerized, the samples were loaded into wells at a low voltage. The gel was washed twice with distilled water. To determine the size of the DNA fragments produced by PCR, a known 50bp DNA ladder for EST-SSRs and a 1kb ladder for ISSRs were loaded onto the gel. The gel was stained with ethidium bromide to enhance the visibility of the DNA bands. For polyacrylamide gels, silver staining was performed using a 0.2% silver nitrate solution. The gels were lightly shaken for 30 minutes and then visualized under a UV transilluminator to detect the bands. Finally, a Gel Documentation system (Photonyx, USA) was used for further documentation (Figures 1, 2A, B).

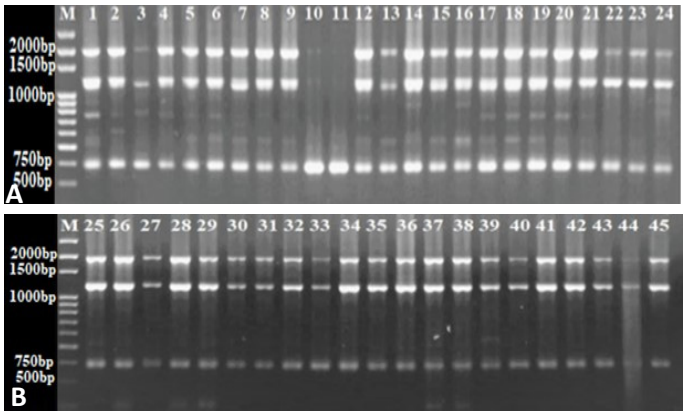

Figure 1: *(A and B). Gel Documentation of the ISSR Marker Systems.*

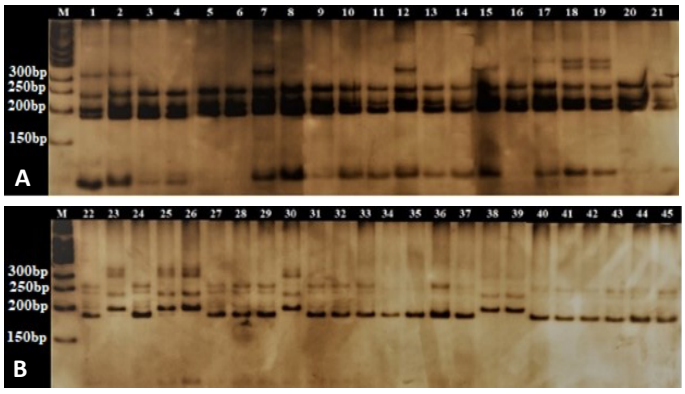

Figure 2: *(A and B) Gel Documentation of the EST-SSR Marker Systems.*

Analyzing data and evaluating gel results

The ISSR and EST-SSR amplicons were noted manually in a binary system. Each band was considered as an allele, with a score of 0 indicating its absence and a score of 1 indicating its presence. To generate a dendrogram for the EST-SSR and ISSR marker systems, the software NTSyspc 2.10e was utilized. The unweighted-pair group method of arithmetic means (UPGMA) was employed for this purpose. The binary data of the EST-SSR and ISSR markers were used to generate a Similarity matrix using Nei's coefficient (Nei, 1972).

The STRUCTURE software was used to analyze the genetic composition of cotton germplasm using the Bayesian clustering technique. The burn-in period was modified to 50,000 and the Markov chain Monte Carlo (MCMC) iterations were extended to 100,000. To estimate the population structure, 10 independent runs were set as parameters for each favorable population and each run. In the analysis of STRUCTURE, the criteria suggested by Evanno *et al.*(2005) were used to identify the suitable number of clusters or subpopulations (referred to as K). The most optimal value of K was determined using STRUCTURE Harvester, an online tool available at http://taylor0.biology.ucala.edu/structureHarvester/. The selection was based on the principle of choosing the highest K value.

Statistics for distinguishing markers

The Cj (confusion probability), Dj (discriminating power) and PIC (polymorphic information content) values were computed for each primer pair as described by Anderson *et al.*(1993) and Tessier *et al.*(1999).

Results and Discussion

Among the set of 25 SSR primer pairs, it was observed that 13 primers exhibited polymorphism while 12 primer pairs were found to be monomorphic. The sizes of the amplified fragments produced by the SSR primers varied from 140 to 650 base pairs. Likewise, out of the 25 ISSR primers, 14 primers exhibited polymorphism, and the sizes of the PCR products ranged from 300 to 2,000 base pairs. The ISSR UBC-815 primer ranged from 750-2000bp while EST-SSR NAU-1014 ranged from 170-300bp (Figure 1, 2). The amplification profiles of 45 genotypes by 25 EST-SSR showed a total of 28 polymorphic bands out of 53 reproducible products (Table 2), relating to 52.83 percent polymorphism. While ISSRs revealed 35 polymorphic bands out of 108 reproducible products (Table 2), corresponding to 32.40% polymorphism. The number of amplicons/ SSR primers were from 1 to 4 having an average of 1.6 alleles per locus while ISSRs have one to seven bands per locus with an average of 4.323.

Table 2: *Marker discrimination indices for EST-SSR.*

Primers	An. Temp (°C)	No. of loci	Allele size (bp)	PIC	Cj	Dj
NAU915	55	1	210	--	--	--
NAU1014	55	4	170-300	0.471	1.454	0.515
NAU1023	55	2	230-320	--	--	--
NAU1070	55	2	160-170	0.542	0.913	0.543
NAU1350	55	1	250	--	--	--
NAU1362	55	1	240	--	--	--
NAU2651	55	1	250	--	--	--
NAU3009	55	1	330	--	--	--
NAU3120	55	1	230	--	--	--
NAU3201	55	1	170	--	--	--
NAU3203	55	3	150-650	0.28	2.151	0.282
NAU3558	55	3	210-340	0.705	0.579	0.710
NAU3735	55	3	400-650	--	--	--
NAU3773	55	4	230-250	0.403	1.778	0.47
NAU3920	55	1	230	--	--	--
NAU4047	55	2	330-350	0.086	0.912	0.087
NAU4086	55	2	180-220	0.802	0.191	0.808
NAU5024	55	2	270-380	--	---	---
NAU5061	55	2	240-290	--	---	---
NAU5109	55	3	150-370	0.375	1.862	0.379
MGHES-6	55	2	180-190	0.746	0.494	0.752
MGHES-31	55	2	200-250	0.086	0.912	0.087
MGHES-40	55	4	190-400	0.886	0.216	0.891
MGHES-62	55	3	140-235	0.335	1.322	0.338
MGHES-70	55	2	190-200	0.542	0.913	0.543

An. Temp. Annealing temperature; PIC, polymorphic content; CJ, confusion probability; DJ, Discriminating power

The PIC value for EST-SSR primers varied between 0.086 and 0.886, with an average value of 0.486. The maximum PIC value (0.886) was found for MGHES-40 followed by MGHES-6 (0.746), MGHES-70(0.542) and MGHES-62 (0.335). The maximum value (0.891) of discriminating power (Dj) and the smallest level (0.21) of confusion probability (Cj) were obtained for the MGHES-40 primer. MGHES-70 primer shows the uppermost Cj value (0.913) (Table 2). The analysis of EST-SSR markers

derived from the NAU series revealed that the number of alleles per locus ranged from 1 to 4, with an average of 1.6 alleles per locus. The observed polymorphic information content (PIC) values exhibited a range of 0.086 to 0.802, with a mean value of 0.444. The lowest PIC value was observed for primer NAU-4047 along with Dj (0.087) and Cj (0.912) (Table 2). Among the 25 ISSR primers, 14 primers showed polymorphism. The PIC value (for ISSR) ranged 0.518-0.043 with an average of 0.25. UBC-819 showed the highest PIC value (0.518) along with the highest Dj (0.580) and lowest Cj (0.438). The primer UBC-840 has the lowest PIC value (0.043). The Dj value ranged from 0.522 to 0.710 and the Cj value extended from 0.913-0.438 (Table 3).

Table 3: *Indices of marker discrimination for ISSRs.*

Primers	An. Temp (°C)	No.of Loci	Allele size (bp)	PIC	Cj	Dj
UBC807	52	7	400-1500	0.195	0.8	0.6
UBC810	52	5	500-1100	---	---	---
UBC813	52	1	550	---	--	--
UBC814	52	3	550-1400	0.057	0.941	0.529
UBC815	52	4	750-2000	0.261	0.732	0.633
UBC817	52	2	1700-1750	---	----	---
UBC818	52	3	830-1000	0.221	0.773	0.613
UBC819	54	3	850-1700	0.518	0.438	0.580
UBC820	54	6	300-1500	0.395	0.595	0.702
UBC821	52	5	420-1050	0.195	0.8	0.6
UBC822	52	3	350-1000	---	---	---
UBC823	50	3	320-1300	---	---	---
UBC824	52	4	500-1500	0.388	0.603	0.698
UBC825	52	5	470-1150	---	---	---
UBC826	52	5	330-1500	0.084	0.913	0.543
UBC828	52	4	500-1800	---	---	---
UBC840	52	5	470-1250	0.043	0.955	0.522
UBC841	52	5	330-1500	0.410	0.579	0.710
UBC842	48	6	350-1500	---	---	---
UBC845	50	6	380-1400	0.35	0.646	0.676
UBC846	50	5	550-2000	---	---	---
UBC848	52	6	300-950	0.334	0.657	0.671
UBC849	52	3	330-1400	---	---	---
UBC850	52	5	400-2000	---	---	---
UBC867	52	4	500-1500	0.345	0.646	0.676

An. Temp. Annealing temperature; PIC, polymorphic content; CJ, confusion probability; DJ, Discriminating power.

Conduct cluster analysis and generate a similarity matrix for ISSRs

A dendrogram was generated through the utilization of UPGMA-based cluster analysis, employing 25 ISSR primers, resulting in the production of 108 loci. The similarity index obtained from pairwise comparison indicated values ranging from 0.83 to 1.00 having a mean value of 0.92. The dendrogram was truncated at 0.97 similarity values and it divided 45 cotton accessions into 4 major clusters (Figure 3) and comprising nine independent genotypes i.e., BT-A1, AGC-555, VH-636, FH-118, IUB-222, Cyto-178, BH-1999, NS-131 and SH-Buraq. Cluster A was furthered distributed into four sub clusters, 1A (NS-141 and AA-802), 2A (CIM-116 and TARZAN-2), 3A (CEMB-33 and FH-Lalazar) and 4A (VH-305 and BH-172) along with five independent genotypes i.e., CIM-632, Sitara-008, IR-3701, NIBGE-8, FH-113. Cluster B comprised of two sub-clusters i.e., 1B (Crystal-1 and MNH-1026) and 2B (MNH-1020 and BS-2015) along with two independent genotypes i.e., NIBGE7 and GH-Baghdadi. Cluster C comprised three sub-clusters i.e., 1C (CYTO-124 and BS-80), 2C (FH-Kahkashan and MNH-886) and 3C (CEMB-66 and CIM-622) along with two independent genotypes i.e NIAB-78 and Shahkar. Custer D contained three sub clusters including 1D (IUB-213 and CIM-600), 2D (Eagle-1, CIM-602 and 6/13), and 3D (GH-Mubarak and CEMB-55) along with two independent genotypes i.e AGC-999 and Tarzan-1.

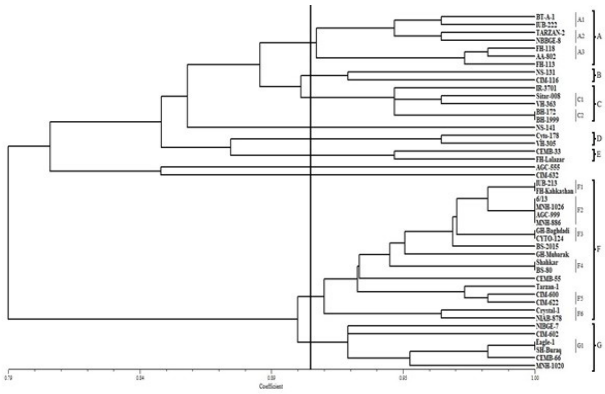

Figure 3: *Dendrogram of forty-five Bt and non-Bt genotypes based on EST-SSR.*

Cluster analysis and similarity matrix for EST-SSRs

A total of 53 loci were produced by 25 EST-SSR primers and UPGMA was utilized to generate the dendrogram. The similarity index obtained from pairwise comparison indicated values ranging from 0.64 to 1.00 having a mean value of 0.85. The dendrogram was truncated at 0.91 similarity value and it divided 45 cotton genotypes into seven main clusters (Figure 4) along with three independent

genotypes i.e. NS-141, AGC-555, and CIM-632. Cluster A comprised of three sub-clusters including 1A (BT-A1 and IUB-222), 2A (Tarzan-2 and NIBGE-8) and A3 (FH-118 and AA-802) along with one independent genotype i.e. FH-113. Cluster B contained two genotypes i.e., NS-131 and CIM-116. Cluster C contained two sub-clusters including C1 (Sitara-008 and VH-363) and C2 (BH-172 and BH-1999) along with one independent genotype i.e., IR-3701. Cluster D comprised of two genotypes i.e., Cyto-178 and VH-305 while Cluster E also included two genotypes i.e., CEMB-33 and FH-Lalazar. Cluster F contained 6 sub clusters via F1 (FH-Kahkashan and IUB-213), F2 (MNH-1026, 6/13, MNH-886 and AGC-999), F3 (CYTO-124 and GH-Baghdadi), F4 (Shahkar and BS-80), F5 (CIM-600 and CIM-622) and F6 (Crystal-1 and NIAB-878) along with 6 independent genotypes i.e BS-2015, GH-Mubarak, CEMB-55, Tarzan-1, NIBGE-7 and CIM-602. Cluster G contained one sub-cluster i.e., G1 (Eagle-1 and SH-Buraq) along with two independent genotypes i.e., CEMB-66 and MNH-1020.

Indices for discriminating ISSRs and EST-SSR markers

Phylogenetic studies of EST-SSRs showed one to four alleles per loci. Table 2 displays a range of PIC values from 0.086 to 0.886, with a mean value of 0.48. The MGHES-40 primer showed the highest value of PIC (0.886) along with Dj (0.891) while having the lowest Cj (0.21). The lowest PIC (0.086) was observed for primer NAU-4047 along with Dj (0.087) and Cj (0.912). Similarly, ISSRs revealed one to seven numbers of loci having an average of 4.32 (Table 3). The PIC values varied from 0.043 to 0.518 with a mean value of 0.25. The primer UBC-840 showed the lowest PIC value (0.043) along with the lowest Dj (0.522) and highest Cj (0.955). The primer UBC-819 depicted the highest PIC value (0.518) along with the highest Dj (0.580) and lowest Cj (0.438).

Analyzing the distinctions between EST-SSR and ISSR marker systems

A number of different parameters were recorded to find differences between EST-SSRs and ISSRs marker systems. The total assay unit for EST-SSRs and ISSRs was 25, while the polymorphic bands per assay count for EST-SSRs and ISSRs were 1.12 and 1.4, respectively. ISSRs showed a high number of loci per assay (4.32) while EST-SSRs showed a high value of marker index (0.54) (Table 4).

Table 4: *Comparison criteria for the EST-SSR and ISSR marker systems.*

Statistics	Marker system	
	EST-SSR	ISSR
Total assay unit (U)	25	25
Polymorphic bands (Np)	28	35
Monomorphic bands (Mnp)	25	73
Polymorphic bands per assay (Np/U)	1.12	1.4
Total loci generated (L)	53	108
Locus per assay unit (nu)	2.12	4.32
Expected heterozygosity (Hep)	0.48	0.25
Fraction of polymorphic bands (B)	0.53	0.32
PIC (mean)	0.48	0.25
Effective multiplex ratio (E)	1.12	1.40
Marker Index value (MI	0.54	0.34

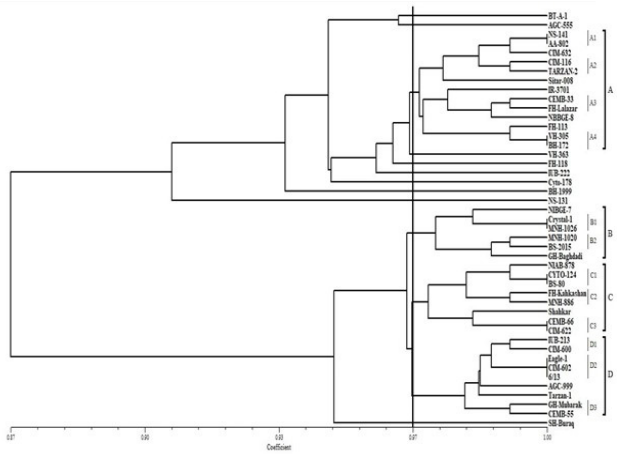

Figure 4: *Dendrogram of forty-five Bt and non-Bt genotypes based on ISSR.*

Structure analysis using ISSR and EST-SSR markers

A set of 45 genotypes was clustered through "Structure Software" by using an admixture model to determine a mixed grouping by means of correlated allelic frequency between various populations. By using K value ranging from 2 to 10 with 20 repetitions, the data was run through software with burn in period length 3,000 and MCMC reps of 30,000. For dominant markers, the logarithm of data likelihood Ln P (D) declined to put satisfactory results however ad hoc quantity (ΔK) based system was applied to estimate the best value of K. The results of the analysis, particularly with EST-SSR markers (as shown in Figures 5 and 7), indicated the presence of four distinct clusters (K) that optimized the DK parameter.

In EST-SSR, K1 comprised of 13.6%, K2 (16.5%), K3 (19.4%), K4 (18.7%), K5 (21.7%), and K6

(10%) proportion of genotypes. The average genetic divergence between subpopulations was high for K2-K3 (0.1485) while it was lowest for K2-K4 (0.0253). The genetic divergence was highest between individuals of K6 (0.1504) and lowest between individuals of K2 (0.0261). On the other hand, the highest value of Delta K (ΔK) for ISSR marker was obtained at ΔK = 4 (Figures 6 and 8). K1 comprised16.0%, K2 (5.6%), K3 (18.7%), K4 (37.1%), K5 (19.3%) and K6 contained 3.3% proportion of genotypes. The mean genetic difference among subpopulations was maximum for K2-K3 (0.1105) while it was lowest for K2-K4 (0.0319). The genetic divergence was highest between individuals of K6 (0.0630) and shortest between individuals of K3 (0.0133).

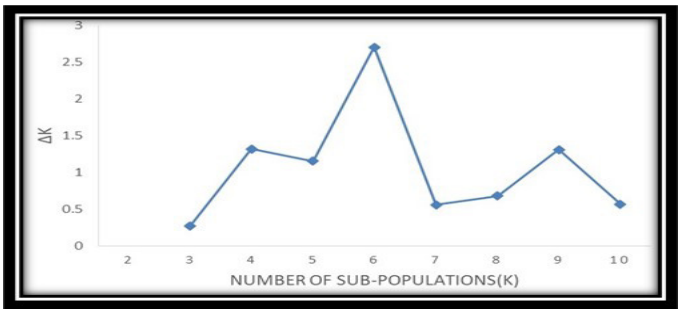

Figure 5: *Delta K for EST-SSR markers.*

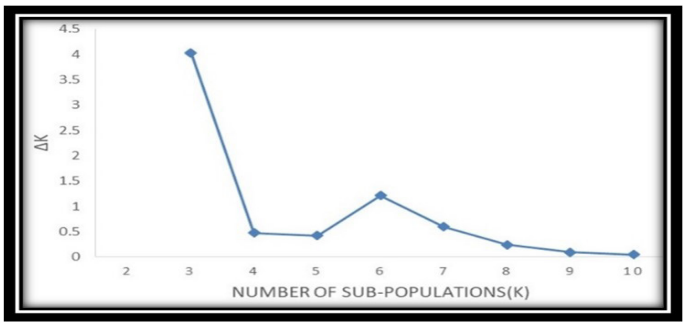

Figure 6: *Delta K for ISSR markers.*

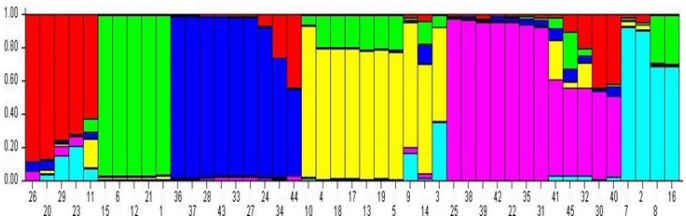

Figure 7: *Population structure of 45 cotton accessions revealed by EST-SSR marker system.*

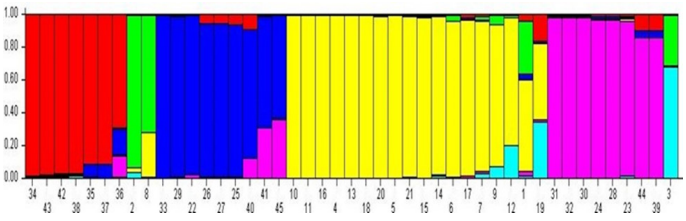

Figure 8: *Population structure of 45 cotton accessions revealed by ISSR marker system.*

In recent decades, molecular markers have gained widespread use for assessing genetic diversity, crucial for enhancing species genetics. Marker selection depends on specific objectives, expected polymorphism levels, resource availability, time, and budget constraints (Kumar *et al.*, 2009). Combining multiple markers can yield superior results compared to individual markers (Serra *et al.*, 2007). Past studies have employed various molecular markers like ISSR, IPBS-retrotransposons, SSR, and RAPD to investigate genetic variability among cotton genotypes (Bukhar *et al.*, 2021; Ashraf *et al.*, 2016; Baran *et al.*, 2023). EST-SSR markers, with co-dominant inheritance, are ideal for fingerprinting, and are valuable due to their origin in functional gene sequences and high transferability. On the other hand, ISSR markers are also multi-locus markers and exhibit dominant inheritance, which makes them highly effective for analyzing genetic diversity (Nadeem *et al.*, 2018; Sethi *et al.*, 2016; Malik *et al.*, 2014). This study evaluates genetic similarity in 45 non-Bt and Bt cotton genotypes using these markers and also compares ISSRs and EST-SSRs for diversity analysis. The EST-SSRs showed 52.83% polymorphism whileISSRs found 32.40%. The number of amplicon SSR primers varied from 1 to 4 having an average of 1.6 alleles per locus while ISSRs have one to seven bands per locus with an average of 4.32 in 25 ISSR primers. The mean SSR polymorphism band per primer in this investigation was lower than ISSR. The level of polymorphism (SSR = 89.28%) found in our research was greater than in earlier cotton studies using different markers (Bilval *et al.*, 2017; Dahab *et al.*, 2013; Tyagi *et al.*, 2015). The results of the current study (no of loci 1 to 4 per SSR) were in line with the previous study by McCarty *et al.* (2022) who studied genetic variation among cotton germplasm applying SSR markers. Bilval *et al.* (2017) investigated genetic polymorphism (the number of alleles/loci ranged from 1 to 4) using sixteen SSR primers. Zhu *et al.* (2019), using 557 accessions of G. hirsutism, reported 6.02 alleles per locus. The genetic diversity of 22 cotton collections utilizing 30 SSR markers was studied by Javaid *et al.* (2017), who found 3.72 alleles per locus. Similar to this, Gurmessa (2019) found 3.8 alleles per locus in cotton genotypes, but McCarty *et al.* (2022), revealed a significant number of alleles (7.9) per locus. Ali *et al.* (2019) reported a 6.3 number of alleles in cotton germplasm. Lacape *et al.* (2007) and Zhang *et al.* (2011) found an average of 5.5 alleles per locus, ranging from 2 to 26 per locus. Moreover, a low level of polymorphism for ISSR primer was shown

in our investigation against earlier studies (Dongre *et al*., 2004). Bardak and Bolek (2012) revealed a total of 173 alleles, including 3.93 alleles per locus using 5 ISSRs and 39 SSRs applied for 25 genotypes of cotton.

PIC values vary among genotypes, and higher values indicate greater genetic diversity and allelic differentiation. Menezes *et al*. (2015) reported that markers with higher PIC values are more effective in identifying polymorphism within a specific population. In a study conducted by Cai *et al*. (2014), who analyzed two, *G. barbadense* and 99 *G. hirsutum* genotypes, and the average PIC value for 20 SSRs was determined to be 0.46. The average PIC value obtained in our investigation aligns with the findings of De Magalhães Bertini *et al*. (2006), who reported a value of 0.48 while examining the genetic relationship between multiple Brazilian cotton genotypes employing SSR markers. However, our PIC value findings were greater than the value of 0.46 reported by Tu *et al*. (2014), evaluating the genetic relationship of multiple upland cotton varieties exploiting SSR markers. Correspondingly, the genetic diversity assessment conducted by Guang and Xiong-Ming, 2006 using SSR markers on various upland cotton genotypes from diverse ecological areas in China yielded a lower value of 0.62. Zhang *et al*. (2011) reported a value of 0.80 when they examined the genetic diversity between different cultivars of cotton from China by EST-SSR markers. In related studies, different research teams found varying PIC values. Abdurakhmonov *et al*. (2008) measured an average PIC value of 0.122 using 287 accessions and 95 SSRs. Tyagi *et al*. (2014) achieved a value of 0.17 with 378 accessions and 120 SSRs. Moiana *et al*. (2015) obtained a value of 0.361 from 20 accessions and 27 SSRs, while Qin *et al*. (2015) reported a mean PIC of 0.3 from their study involving 241 accessions and 333 SSRs. Kuang *et al*. (2022) obtained PIC of the SSR markers fluctuated from 0.18 to 0.90, with a mean of 0.64 in 79 cotton genotypes. Çelik (2022) found PIC of SSR markers varying from 0.49 to 0.10 with an average PIC value of 0.312. Seyoum *et al*. (2018) attained PIC values ranging from 0.371 to 0.019 (mean 0.225) through SSR.

In context to ISSR, PIC values varied from 0.518 to 0.043, with an average value of 0.25 (Table 3). Zaki and Hussein (2023) found average PIC value 0.239 in cotton genotypes using ISSR markers. The PIC values obtained in the study conducted by Abdellatif and Soliman (2013) were higher than in our study. Tyagi *et al*. (2014) observed PIC values, ranging from 0.86 to 0.90, when employing ISSR primers in fifteen cotton genotypes.

Effective primers play a vital role in genetic diversity studies. MGHES-31 exhibited the highest PIC value of 0.750 for EST-SSRs, whereas UBC-807 and UBC-815 had a PIC value of 0.491 for ISSRs in the current study. These primers also showed a high level of Dj value and a low level of Cj value, indicating their strong capability to detect differences in alleles. Both of these primers showed a greater tendency to differentiate among genotypes.

Our research found that there is a substantial amount of genetic similarity among 45 cotton genotypes. This similarity ranged from 73% to 100% for EST-SSRs and from 77% to 97% for ISSRs. In a study by Bilval *et al*. (2017), they found genetic similarities between 54% and 96% using SSR markers. Similarly, in a study by Ashraf *et al*. (2016), they observed comparable levels of genetic similarity among different Bt cotton types, with genetic similarity ranging from 73% to 100% using EST-SSR markers and from 77% to 97% using ISSR markers. Ullah *et al*. (2012) also noted high genetic similarity, ranging from 0.90 to 0.98, among 19 Bt cotton varieties. Previous research by Iqbal *et al*. (1997), Lukonge *et al*. (2007) and Rahman *et al*. (2008) also reported significant genetic similarities among various cotton types. The examination of genetic similarity among different cotton genotypes revealed a significant level of resemblance (Ullah *et al*., 2012; Kalivas *et al*., 2011)

To check the relationship among cotton germplasm, Population structure and dendrogram were used as clustering algorithms (Figures 3,4,7,8) and here we will briefly explain the dendogram. The clustering analysis of the dendrograms, constructed using both EST-SSR and ISSR markers, revealed that a significant proportion of clusters consisting of genotypes were from both (public and private) sectors. This similarity is likely because the same gene pool has been used repeatedly, leading to limited genetic diversity in the available germplasm (Zhang *et al*., 2011). Breeders also frequently used closely related elite parental lines, which resulted in Bt cotton genotypes with close genetic ties. In our study, we classified the germplasm into seven distinct groups based on EST-SSR markers

(Figure 3). Tyagi *et al*. (2014) categorized 381 cotton genotypes into five groups using SSR markers. Khan *et al*. (2009) used SSR markers to divide 40 genotypes into three groups, with an average similarity ranging from 36% to 89%. Anderica *et al*. (2018) examined 48 cotton genotypes with 62 SSR markers, resulting in three main clusters. Kuang *et al*. (2022) employed SSR markers, grouping all samples into five classes at a similarity coefficient of 0.57.

Cluster analysis based on ISSR categorized 45 cotton germplasm into four major groups (Figure 4). Tyagi *et al*. (2014) examined the cluster analysis of fifteen cotton genotypes divided into four groups similar to the current study.

The effectiveness of primers can be assessed using important parameters such as marker index (MI) (Powell *et al*., 1996) and polymorphic information content (PIC) (Anderson *et al*., 1993) which can be determined using various statistical tools. Based on our findings, the evaluation of two marker systems in terms of their discriminating efficiency exposed that the EST-SSR marker exhibited the superior ability to demonstrate allelic variations between particular genotypes of cotton. This was evident from the higher expected heterozygosity (Hep) observed in the EST-SSR marker compared to the ISSR markers (Rohlf, 2000; Belaj *et al*., 2003). The comparison of two marker systems in a study involving forty-five cotton genotypes revealed that the expected heterozygosity (Hep) value was higher for the EST-SSR marker (0.48) compared to the ISSR marker (0.25). An earlier study by Ashraf *et al*. (2016) found that EST-SSR had higher expected heterozygosity (0.71) than ISSR (0.29) for distinguishing variations.

The study found that EST-SSRs had a higher marker index value compared to ISSRs. ISSRs had a lower marker index value due to a higher multiplex ratio (E = 1.40), a unique characteristic of these markers. This uniqueness is likely due to the higher number of alleles produced by each ISSR primer rather than allelic heterozygosity between genotypes (Maras *et al*., 2008). Using both types of molecular markers provided valuable insights into cotton's genetic diversity, emphasizing its importance in characterizing cotton germplasm. The high genetic similarity among cotton genotypes highlights the need for diversifying parental lines in breeding programs. To address limited genetic variation, new approaches like transgenic development and wide hybridization are necessary in existing cotton varieties.

Conclusions and Recommendations

Cotton holds immense economic significance in both Pakistan and worldwide agriculture, contributing substantially to foreign exchange earnings for different countries. This study primarily concentrated on the molecular characterization of 45 cotton genotypes native to Pakistan. The selection of a suitable molecular marker technique for assessing genetic diversity entails careful consideration of factors such as statistical power, reliability, and the extent of polymorphisms. In this context, EST-SSR and ISSR markers are suitable choices for genotype screening and molecular characterization. Furthermore, Dj (Discriminating Power) and PIC (Polymorphic Information Contents) represent more dependable marker-discriminating indices for the purpose of germplasm characterization. Additionally, ISSRs may be particularly advantageous for assessing genetic variability among Bt and non-Bt cotton genotypes due to their capacity to generate a greater number of bands per reaction. Conversely, EST-SSRs, characterized by their high Expected (Hep) value and co-dominant nature, are also best suited for genome mapping applications.

Acknowledgements

The authors are highly thankful to the Department of Plant Breeding and Genetics, Bahauddin Zakariya University, Multan for providing necessary materials for successful completion of the research.

Novelty Statement

Exploring the genetic differences in cotton types using molecular markers is like going on an exciting adventure. This study helps us better understand cotton's genetic diversity and can be valuable for cotton breeders working on developing new cotton varieties.

Author's Contribution

Jaweria Iqbal and Sameer Ahmed: Designed and conducted the study.
Muhammad Tanveer Altaf and Amna Jamil: performed analysis.
Muhammad Tanveer Altaf, Jaweria Iqbal and Waqas Liaqat: Manuscript writing.

Muhammad Faheem Jan, Amjad Ali and Arif Mehmood: Helped in writing and formatting.
Waqas Raza, Muhammad Tanveer Altaf and Ikram ul Haq: Helped in writing and proof reading of the manuscript.

Conflict of interest
The authors have declared no conflict of interest.

References

Abdellatif, K.F. and Y.A. Soliman. 2013. Genetic relationships of cotton (*Gossypium barbadense* L.) genotypes as studied by morphological and molecular markers. Afric. J. Biotech., 12(30): 4736-4746. https://doi.org/10.5897/AJB2013.12361

Abdellatif, K.F., Y.A. Khidr, Y.M. El-Mansy, M.M. El-Lawendey and Y.A. Soliman. 2012. Molecular diversity of Egyptian cotton (*Gossypium barbadense* L.) and its relation to varietal development. J. Crop Sci. Biotechnol., 15: 93–99. https://doi.org/10.1007/s12892-011-0120-5

Abdurakhmonov, I.Y., R.J. Kohel, J.Z. Yu, A.E. Pepper, A.A. Abdullaev, F.N. Kushanov, I.B. Salakhutdinov, Z.T. Buriev, S. Saha and B.E. Scheffler. 2008. Molecular diversity and association mapping of fiber quality traits in exotic *G. hirsutum* L. germplasm. Genomics, 92: 478-487. https://doi.org/10.1016/j.ygeno.2008.07.013

Ali, I., N.U. Khan, S. Gul, Khan, Z. Bibi, K. Aslam, G. Shabir, H. Haq, S. Aslam and I. Hussain. 2019. Genetic diversity and population structure analysis in upland cotton germplasm. Int. J. Agric. Biol., 22: 669–676.

Andérica, E.P., F.R. Montero, M.L. García and J.G. Ligero. 2018. Genetic diversity and phylogenetic relationships of a potential cotton collection for European breeding research. Turk. J. Bot., 42(2): 172-182. https://doi.org/10.3906/bot-1706-18

Anderson, J.A., G.A. Churchill, J.E. Autrique, S.D. Tanksley and M.E. Sorrells. 1993. Optimizing parental selection for selection for genetic linkage maps. Genome, 36: 181-186. https://doi.org/10.1139/g93-024

Ashraf, J., W. Malik, M. Iqbal, A. Khan, A. Qayyum, E. Noor and M.Q. Ahmad. 2016. Comparative analysis of genetic diversity among bt cotton genotypes using est-ssr, issr and morphological markers. J. Agric. Sci. Technol., 18(2): 517-531.

Bakhsh, A., M. Rehman S. Salman and R. Ullah. 2019. Evaluation of cotton genotypes for seed cotton yield and fiber quality traits under water stress and non-stress conditions. Sarhad J. Agric., 35: 161–170. https://doi.org/10.17582/journal.sja/2019/35.1.161.170

Baran, N., F. Shimira, M.A. Nadeem, M.T. Altaf, M. Andirman, F.S. Baloch and M.G. Temiz. 2023. Exploring the genetic diversity and population structure of upland cotton germplasm by iPBS-retrotransposons markers. Mol. Biol. Rep., 50(9): 1-13. https://doi.org/10.1007/s11033-023-08399-0

Bardak, A. and Y. Bolek. 2012. Genetic diversity of diploid and tetraploid cottons determined by SSR and ISSR markers. Turk. J. Field Crops, 17(2): 139-144.

Basal, H., E. Karademir, H.K. Goren, V. Sezener, M.N. Dogan, I. Gencsoylu and O. Erdogan. 2019. Cotton production in turkey and Europe. In: Jabran K, Chauhan BS (eds) cotton production. John Wiley and Son Inc., New York, NY, USA. pp. 297–321. https://doi.org/10.1002/9781119385523.ch14

Belaj, A., Z. Satovic, G. Cipriani, L. Baldoni, R. Testolin, L. Rallo and I. Trujillo. 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. https://doi.org/10.1007/s00122-003-1301-5

Bilval, B.B., K.V. Vadodariya, B.K. Rajkumar and G.R. Lahane. 2017. Genetic diversity of parents using RAPD, ISSR and SSR molecular markers in upland cotton (*Gossypium hirsutum* L.). Bull. Environ. Pharmacol. Life Sci., 6: 51-57.

Bukhari, S.A., M.A. Iqbal and S. Naz. 2021. Studies on genetic diversity of cotton using RAPD markers. Pure. Appl. Biol., 3(3): 95-100. https://doi.org/10.19045/bspab.2014.33002

Cai, C., W. Ye, T. Zhang and W. Guo. 2014. Association analysis of fiber quality traits and exploration of elite alleles in Upland cotton cultivars/accessions (*Gossypium hirsutum* L.). J. Integr. Plant Biol., 56: 51-62. https://doi.org/10.1111/jipb.12124

Çelik, S., 2022. Genetic diversity analysis of some upland cotton (*Gossypium hirsutum* L.) genotypes using SSR markers. Tr. Doğa ve Fen Derg., 11(1): 80-89. https://doi.org/10.46810/tdfd.995786

Dahab, A.A., M. Saeed, B.B. Mohamed, M.A. Ashraf, A.N. Puspito, K.S.B.A.A. Shahid and T. Husnain. 2013. Genetic diversity assessment of cotton (*Gossypium hirsutum* L.) genotypes from Pakistan using simple sequence repeat markers. Aust. J. Crop Sci., 7(2): 261-267.

De Magalhães Bertini, C.H.C., I. Schuster, T. Sediyama, E.G. de Barros and M.A. Moreira. 2006. Characterization and genetic diversity analysis of cotton cultivars using microsatellites. Genet Mol. Biol., 29(2): 321–329. https://doi.org/10.1590/S1415-47572006000200021

De Menezes, I.P.P., L.V. Hoffmann and A.P.V. Barroso. 2015. Genetic characterization of cotton landraces found in the Paraíba and Rio Grande do norte states. Crop. Breed. Appl. Biotechnol., 15: 26–32. https://doi.org/10.1590/1984-70332015v15n1a4

Dongre, A., V. Parkhi and S. Gahukar. 2004. Characterization of cotton (*G. hirsutum*) germplasm by ISSR, RAPD markers and agronomic value. Ind. J. Biotechnol., 3: 388-393.

Evanno, G., S. Regnaut and J. Goudet. 2005. Detecting the number of clusters of individuals using the software structure: A simulation study. Mol. Ecol., 14: 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

FAOSTAT, 2020-21. Food outlook biannual report on global food markets. Food and Agriculture Organization Statistics (FAO), Rome, Italy.

Farahani, F., M. Sheidai and F. Koohdar. 2018. Genetic finger printing of cotton cultivars by ISSR molecular markers. Genetika-Belgrade, 50: 627–634. https://doi.org/10.2298/GENSR1802627F

Gautam, A.K., R.K. Verma, S. Avasthi, Y. Bohra, B. Devadatha, M. Niranjan and N. Suwannarach. 2022. Current insight into traditional and modern methods in fungal diversity estimates. J. Fung., 8: 226. https://doi.org/10.3390/jof8030226

Guang, C.H.E.N. and D.U. Xiong-Ming. 2006. Genetic diversity of source germplasm of upland cotton in China as determined by SSR marker analysis. Acta Genet. Sin., 33(8): 733-745. https://doi.org/10.1016/S0379-4172(06)60106-6

Gurmessa, D., 2019. Genetic diversity study of improved cotton (*G. hirsutum* L.) varieties in Ethiopia using simple sequence repeats markers. J. Biotechnol., 7: 6–14.

Jian, W.A.N.G., H.O.U. Lu, R.Y. Wang, M.M. He and Q.C. Liu. 2017. Genetic diversity and population structure of 288 potato (*Solanum tuberosum* L.) germplasms revealed by SSR and AFLP markers. J. Integr. Agric, 16(11): 2434-2443. https://doi.org/10.1016/s2095-3119(16)61619-2

Han, P., X. Tian, Y. Wang, C. Huang, Y. Ma, X. Zhou, Y. Yu, D. Zhang, H. Xu, Y. Cao, B. Zhu, Z. Feng, S. He, X. Du, Z. Lin, L. Zhu, C. You, Z. Pan and X. Nie. 2022. Construction of a core germplasm bank of upland Cotton (*Gossypium hirsutum* L.) based on phenotype, genotype and favorable alleles. Genet. Resour. Crop Evol., 69: 2399–2411. https://doi.org/10.1007/s10722-022-01379-6

Hayat, K., A. Bardak, D. Parlak, F. Ashraf, H.M. Imran, H.A. Haq, M.A. Mian, Z. Mehmood and M.N. Akhtar. 2020. Biotechnology for cotton improvement. In: (eds. Ahmad, S. and Hasanuzzaman, M.) Cotton production and uses. Springer, Singapore. https://doi.org/10.1007/978-981-15-1472-2_25

Iqbal, M.J., N. Aziz, N.A. Saeed, Y. Zafar and K.A. Malik. 1997. Genetic diversity evaluation of some elite cotton varieties by RAPD analysis. Theor. Appl. Genet., 94: 139-144. https://doi.org/10.1007/s001220050392

Jabran, K, S. ul-Allah, B.S. Chauhan and A. Bakhsh. 2019. An Introduction to global production trends and uses, history and evolution, and genetic and biotechnological improvements in cotton. Chapter 1, pp. 1-22. Hoboken, NJ: Wiley Online Library. https://doi.org/10.1002/9781119385523.ch1

Jamil, A., K. Razzaq, I.A. Rajwana, A. Naz, G. Akhtar, S. Ullah, H.N. Faried, S.B. Hussain, Y. Li, M. Amin, M.A. Sher, M.T. Altaf,H. Hamad, M.A.A. Ahmed, A. Saleh.and M.J. Ansari. 2022. Characterization of indigenous phalsa (*Grewia subinequalis*) genotypes using morphological traits and ISSR markers. J. King Saud Univ. Sci., 34(7): 102237. https://doi.org/10.1016/j.jksus.2022.102237

Javaid, A., F.S. Awan, F.M. Azhar and I.A. Khan. 2017. Assessment of allelic diversity among drought-resistant cotton genotypes using microsatellite markers. Genet. Mol. Res., 16: 28549206. https://doi.org/10.4238/gmr16029664

Jiang, W., H.B. Zhu and J.M. He. 2008. Genetic diversity in germplasm resources of cotton from

different area based on ISSR markers. Cotton Sci., 20(5): 348-353.

Kahodariya, J., P. Sabara and D. Vakharia 2015. Assessment of genetic diversity in old world and new world cotton cultivars using RAPD and ISSR markers. Indian J. Biotechnol., 14: 511–517.

Kalivas, A., F. Xanthopoulos, O. Kehagia and A.S. Tsaftaris. 2011. Agronomic characterization, genetic diversity and association analysis of cotton cultivars using simple sequence repeat molecular markers. Genet. Mol. Res., 10: 208-217. https://doi.org/10.4238/vol10-1gmr998

Kantartzi, S.K., M. Ulloa, E. Sacks and J.Mc.D. Stewart. 2009. Assessing genetic diversity in *Gossypium arboreum* L. cultivars using genomic and EST-derived microsatellites. Genetica, 136: 141-147. https://doi.org/10.1007/s10709-008-9327-x

Khan, A.I., Y.B. Fu and I.A. Khan. 2009. Genetic diversity of Pakistani cotton cultivars as revealed by simple sequence repeat markers. Commun. Biometry Crop Sci., 4(1): 21-30.

Khan, I.A., F.S. Awan, A. Ahmad and A.A. Khan. 2004. A modified mini-prep method for economical and rapid extraction of genomic DNA in plants. Plant Mol. Biol. Rep., 22(1): 89-89. https://doi.org/10.1007/BF02773355

Khan, M.A., A. Wahid, M. Ahmad, M.T. Tahir, M. Ahmed, S. Ahmad and M. Hasanuzzaman. 2020. World cotton production and consumption: An overview. In: (eds. Ahmad, S. and Hasanuzzaman, M.,) cotton production and uses. Springer, Singapore. https://doi.org/10.1007/978-981-15-1472-2_1

Kuang, Z., C. Xiao, M.K. Ilyas, D. Ibrar, S. Khan, L. Guo and H. Chen. 2022. Use of SSR markers for the exploration of genetic diversity and DNA finger-printing in early-maturing upland cotton (*Gossypium hirsutum* L.) for future breeding program. Agronomy, 12(7): 1513. https://doi.org/10.3390/agronomy12071513

Kumar, P., S. Nimbal, R.S. Sangwan, N. Budhlakoti, V. Singh, D.C. Mishra and R.R.S. Choudhary. 2021. Identification of novel marker–trait associations for lint yield contributing traits in upland cotton (*Gossypium hirsutum* L.) using SSRs. Front. Plant Sci., 12: 855. https://doi.org/10.3389/fpls.2021.653270

Kumar, P., V. Gupta, A.K. Misra, D.R. Modi and B.K. Pandey. 2009. Potential of molecular markers in plant biotechnology. Plant Omics, 2(4): 141-162.

Lacape, J.M., D.M. Dessauw, J.L. Rajab, B. Noyer and B. Hau. 2007. Microsatellite diversity in tetraploid Gossypium germplasm: assembling a highly informative genotyping set of cotton SSRs. Mol. Breed, 19: 45–58. https://doi.org/10.1007/s11032-006-9042-1

Lukonge, K., L. Hersalman and M.T. Labuschagne. 2007. Genetic diversity of tanzanian cotton revealed by AFLP analysis. Afr. Crop Sci. Conf., 8: 773-776.

Ma, X., X.M. Du and J.L. Sun. 2003. SSR fingerprinting analysis on 18 coloured cotton lines. J. Huazhong Agric. Univ., 23: 606-609.

Majumdar, G., S.B. Singh and S.K. Shukla. 2019. Seed production, harvesting, and ginning of cotton. Cotton Production, pp. 145-174. https://doi.org/10.1002/9781119385523.ch8

Malik, W., J. Ashraf, M.Z. Iqbal, A.A. Khan, A. Qayyum, M.A. Abid, E. Noor, M.Q. Ahmad and G.H. Abbasi. 2014. Molecular markers and cotton genetic improvement: Current status and future prospects. Sci. World J., pp. 1-15. https://doi.org/10.1155/2014/607091

Maras, M., J. Sustar -Vozlic, B. avornik and V. Meglic. 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. https://doi.org/10.2478/v10014-008-0009-2

McCarty, J.C., D.D. Deng, J.N. Jenkins and L. Geng. 2022. Genetic diversity of day-neutral converted landrace *Gossypium hirsutum* L. accessions using SSR markers. Euphytica, 214: 173. https://doi.org/10.1007/s10681-018-2264-6

Menezes, I.P.P.D., L.V. Hoffmann. and P.A.V. Barroso. 2015. Genetic characterization of cotton landraces found in the Paraíba and Rio Grande do Norte states. Crop Breed. Appl. Biotechnol, 15: 26-32. https://doi.org/10.1590/1984-70332015v15n1a4

Moiana L.D., P.S.V. Filho, M.C. Gonçalves-Vidigal and L.P. Carvalho. 2015. Genetic diversity and population structure of upland cotton Brazilian cultivars (*Gossypium hirsutum* L. race latifolium H.) using SSR markers. Aust. J. Crop Sci., 9: 143-152.

Munir, H., F. Rasul, A. Ahmad, M. Sajid, S. Ayub, M. Arif, P. Iqbal, A. Khan, Z. Fatima, S. Ahmad and M.A. Khan. 2020. Diverse

uses of cotton: From products to byproducts. Cotton production and uses: Agronomy, crop protection, and postharvest technologies, Springer Singapore. pp. 629-641. https://doi.org/10.1007/978-981-15-1472-2_30

Murty, S.G., F. Patel, B.S. Punwar, M. Patel, A.S. Singh and R.S. Fougat. 2013. Comparison of RAPD, ISSR, and DAMD markers for genetic diversity assessment between accessions of *Jatropha curcas* L. and its related species. J. Agric. Sci. Tech., 15: 1007-1022.

Nadeem, M.A., M.A. Nawaz, M.Q. Shahid, Y. Doğan, G. Comertpay, M. Yıldız and F.S. Baloch. 2018. DNA molecular markers in plant breeding: Current status and recent advancements in genomic selection and genome editing. Biotechnol. Biotechnol. Equip., 32(2): 261-285. https://doi.org/10.1080/13102818.2017.1400401

Nei, M., 1972. Genetic distance between populations. Am. Nat., 106: 283-292. https://doi.org/10.1086/282771

Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey and A. Rafalski. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed, 2: 225-238. https://doi.org/10.1007/BF00564200

Qin, H., M. Chen, X. Yi, S. Bie, C. Zhang, Y. Zhang, J. Lan, Y. Meng, Y.Yuan and C. Jiao. 2015. Identification of associated SSR markers for yield component and fiber quality traits based on frame map and upland cotton collections. PLoS One, 10: e0118073. https://doi.org/10.1371/journal.pone.0118073

Rahman, M.T., N. Yasmin, I. Tabassum, I. Ullah, M. Asif and Y. Zafar. 2008. Studying the extent of genetic diversity among *Gossypium arboreum* L., genotypes/cultivars using DNA fingerprinting. Genet. Resour. Crop Evol., 55: 331-339. https://doi.org/10.1007/s10722-007-9238-1

Rohlf, F.J., 2000. NTSYS-pc: Numerical taxonomy and multivariate analysis system, Version 2.1. Exeter Software, New York.

Saleem, M.A., A. Qayyum, W. Malik and M.W. Amjid. 2020. Molecular breeding of cotton for Drought stress tolerance. In: (eds. Ahmad, S. and Hasanuzzaman, M.) cotton production and uses. Springer, Singapore. pp. 495-508. https://doi.org/10.1007/978-981-15-1472-2_24

Salunkhe, S.N. and Deshmukh, Y.A., 2009. Molecular characterization of elite cotton cultivars using ISSR markers. Asian J. Bio Sci., 4(2): 249-253.

Serra, I.A., G. Procaccini, M.C. Intrieri, M. Migliaccio, S. Mazzuca and A.M. Innocenti. 2007. Comparison of ISSR and SSR markers for analysis of genetic diversity in the seagrass *Posidonia oceanica*. Mar. Ecol. Prog. Ser., 338: 71-79. https://doi.org/10.3354/meps338071

Sethi, K., P. Siwach and S.K. Verma. 2016. Simple sequence repeats (SSR) and interspersed sequence repeats (ISSR) markers for genetic diversity analysis among selected genotypes of *Gossypium arboreum* race bengalense. Afr. J. Biotechnol., 15(1): 7-19. https://doi.org/10.5897/AJB2015.14878

Seyoum, M., X.M. Du, S.P. He, Y.H. Jia, Z. Pan and J.L. Sun. 2018. Analysis of genetic diversity and population structure in upland cotton (*Gossypium hirsutum* L.) germplasm using simple sequence repeats. J. Genet., 97: 513-522. https://doi.org/10.1007/s12041-018-0943-7

Sharma, A., P. Kodgire and S.S. Kachhwaha. 2020. Investigation of ultrasound-assisted KOH and CaO catalyzed transesterification for biodiesel production from waste cotton-seed cooking oil: Process optimization and conversion rate evaluation. J. Clean. Prod., 259: 120982. https://doi.org/10.1016/j.jclepro.2020.120982

Sharma, S.N., V. Kumar and S. Mathur. 2009. Comparative analysis of RAPD and ISSR markers for characterization of sesame (*Sesamum indicum* L.) genotypes. J. Plant Biochem. Biotechnol., 18: 37-43. https://doi.org/10.1007/BF03263293

Sun, D.L., J.L. Sun, Y.H. Jia, Z.Y. Ma and X.M. Du. 2009. Genetic diversity of colored cotton analyzed by simple sequence repeat markers. Int. J. Plant Sci., 170: 76-82. https://doi.org/10.1086/593037

Swarup, S., C.J. Cargill, K. Crosby, L. Flagel, J. Kniskern and K.C. Glenn. 2021. Genetic diversity is indispensable for plant breeding to improve crops. Crop Sci., 61: 839–852. https://doi.org/10.1002/csc2.20377

Tahir, A., I. Iqbal, K.M. Talib, J. Luhuai, X. Chen, A. Akbar, A. Asghar and I. Ali. 2022. Modern tools for the identification of fungi, including yeasts. In: Extremophilic fungi: Ecology, Physiology and Applications. Singapore: Springer Nature Singapore. pp. 33-51. https://doi.org/10.1007/978-981-16-4907-3_2

Tang, S., T. Zhonghua, Z. Tengfei, X. Fang, F. Liu, D. Liu, J. Zhang, D. Liu, S. Wang, K. Zhang, Q. Shao, Z. Tan, H.P. Andrew and Z. Zhang. 2015. Construction of genetic map and QTL analysis of fiber quality traits for upland cotton (*Gossypium hirsutum* L.). Euphytica, 201: 195–213. https://doi.org/10.1007/s10681-014-1189-y

Tessier, C., David, J., P.This, J.M. Boursiquot and A. Charrier. 1999. Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. TheoR. Appl. Genet., 98: 171-177. https://doi.org/10.1007/s001220051054

Tessier, C., J. David, P. This, J.M. Boursiquot and A. Charrier. 1999. Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. Theor. Appl. Genet., 98: 171-177. https://doi.org/10.1007/s001220051054

Tokel, D. and B.N. Erkencioglu. 2021. Production and trade of oil crops, and their contribution to the world economy. Oil crop genomics, Springer Cham., pp. 415-427. https://doi.org/10.1007/978-3-030-70420-9_20

Tu, J.L., M.J. Zhang, X.Q. Wang, X.L. Zhang and Z.X. Lin. 2014. Genetic dissection of upland cotton (*Gossypium hirsutum*) cultivars developed in Hubei Province by mapped SSRs. Genet. Mol. Res. 13(1): 782–790. https://doi.org/10.4238/2014.January.31.4

Tyagi, P., M.A. Gore, D.T. Bowman, B.T. Campbell, J.A. Udall and V. Kuraparthy. 2014. Genetic diversity and population structure in the US Upland cotton (*Gossypium hirsutum* L.). Theor. Appl. Genet., 127: 283-295. https://doi.org/10.1007/s00122-013-2217-3

Tyagi, S.A., S.K. Verma. K.I. Nehra and D.R. Bajya. 2015. Molecular marker based genetic diversity analysis in cotton using RAPD and SSR markers. Int. J. Adv. Res. Biol. Sci., 29: 232-236.

Ullah, I., A. Iram, M.Z. Iqbal, M. Nawaz, S.M. Hasni and S. Jamil. 2012. Genetic diversity analysis of Bt cotton genotypes using simple sequence repeats markers. Genet. Mol. Res., 11: 597-605. https://doi.org/10.4238/2012.March.14.3

Van Esbroeck, G.A., D.T. Bowman, D.S. Calhoun and O.L. May. 1998. Changes in the genetic diversity of cotton in the USA from 1970 to 1995. Crop Sci., 38: 33-37. https://doi.org/10.2135/cropsci1998.0011183X003800010006x

Wang, M., J. Li, P. Wang, F. Liu, Z. Liu and G. Zhao. 2021. Comparative genome analyses highlight transposon-mediated genome expansion and the evolutionary architecture of 3D genomic folding in cotton. Mol. Biol. Evol., 38(9): 3621–3636. https://doi.org/10.1093/molbev/msab128

Wang, X.Y., X.Y. Li, Z.L. Gong, J. Wang, L. Fan, J. Zheng, Y. Liang, J. Guo, M. Mamat and X. Ai. 2018. DNA fingerprinting construction and genetic diversity analysis based on SSR markers for upland cotton in Xinjiang. Cotton Sci., 30: 308–315.

Xiong, Y., X. Lei, S. Bai, Y. Xiong, W. Liu, W. Wu, Q. Yu, D. Dong, J. Yang and X. Ma. 2021. Genomic survey sequencing, development and characterization of single-and multi-locus genomic SSR markers of *Elymus sibiricus* L. BMC Plant Biol., 21: 1–12. https://doi.org/10.1186/s12870-020-02770-0

Yu, J.R., J.S. Chung, S. Huh, S.H. Lee and J.Y. Chai. 1997. PCR-RFLP patterns of three kinds of Metagonimus in Korea Korean J. Parasitol., 35: 271-276. https://doi.org/10.3347/kjp.1997.35.4.271

Yu, J.Z., R.J. Kohel, D.D. Fang, J. Cho, A. Van Deynze, M. Ulloa and S.M. Hoffman, A.E. Pepper, D.M. Stelly, J.N. Jenkins, S. Saha, S.P. Kumpatla, M.R. Shah, W.V. Hugie and R.G. Percy. 2012. A high density simple sequence repeat and single nucleotide polymorphism genetic map of the tetraploid cotton genome. G3, 2(1): 43-58. https://doi.org/10.1534/g3.111.001552

Zaki, H. and N.R. Hussein 2023. Inter simple sequence repeats and morphological traits to identify cultivated cotton varieties (*Gossypium barbadense* L.) in Egypt. Genet. Resour. Crop. Evol., 70(3): 993-1006. https://doi.org/10.1007/s10722-022-01483-7

Zhang, Y., X.F. Wang, Z.K. Li, G.Y. Zhang and Z.Y. Ma. 2011. Assessing genetic diversity of cotton cultivars using genomic and newly developed expressed sequence tag-derived microsatellite markers. Genet Mol. Res., 10: 1462-1470. https://doi.org/10.4238/vol10-3gmr1277

Zhu, L., P. Tyagi, B. Kaur and V. Kuraparthy. 2019. Genetic diversity and population structure in elite US and race stock accessions of upland cotton (*Gossypium hirsutum*). J. Cotton Sci., 23: 38–47. https://doi.org/10.56454/GLUV4792