Preprints.org

Article

Genetic Diversity of Some Upland and Sea Island Cotton Genotypes by High Resolution Capillary Electrophoresis Gel

Adnan Aydin [*] and Emine Uygur Gocer

Article

Genetic Diversity of Some Upland and Sea Island Cotton Genotypes by High Resolution Capillary Electrophoresis Gel

Adnan Aydin *

Department of Agricultural Biotechnology, Faculty of Agriculture, Igdır University, Igdır, Turkey;
  * Correspondence: aydinadnann@gmail.com

Abstract: Cotton (*Gossypium* L.) is a major source of natural fibre for the world's textile industry and is also an important oil plant. Cotton fibre is the main source of textiles, the seeds are used for oil and the rest of the bagasse is a high-protein animal feed. In addition, more than 50 industries use cotton's so-called short fibre. Breeding studies in cotton are generally based on crosses between the best genotypes in terms of yield and fibre quality. However, the narrow genetic diversity of cultivars has a negative impact on cotton breeding programmes. For this reason, the identification of genetic resources and the disclosure of genetic diversity are important. Herein, high-resolution capillary gel electrophoresis was used to determine the genetic diversity of *G. hirsutum* and *G. barbadense* genotypes. Using 19 EST-SSR markers, a total of 47 genotypes were screened. The PIC values of the markers used ranged from 0.268 to 0.889. The mean PIC value was calculated to be 0.603. Concerning the clustering, PCoA and population structure analyses gave similar results, and the genotypes could be divided into 3 main groups. Genetic admixture with *G. hirsutum* was found in some genotypes of *G. barbadense* species. We can conclude that i) the EST-SSR markers used in this study are effective in the determination of genetic diversity, ii) The genetic diversity should be increased through the collection of genetic resources iii) genic EST-SSR markers in this study should be considered in breeding programmes by using them in QTL studies.

Keywords: EST_SSR; population structure; *Gossypium hirsutum*; *Gossypium barbadense*

1. Introduction

There are about 52 cotton species with 9 different cytogenetic genomes. Of these, 8 are diploids and 1 is tetraploid [1]. *Gossypium hirsutum* and *Gossypium barbadense* have tetraploid genomes. They are the most widely cultivated species in the world [2]. The cotton plant is used for both its fibre and its oil [3]. The cotton plant is the source of 35% of the world's fibre requirements [4]. Cottonseed oil can be used in the production of biodiesel [5] and edible oils [6]. For this reason, the genetic resources of such an important crop must be the subject of a comprehensive evaluation. For breeders, the role of genetic resources is very important [7]. Although genetic resources are important in biological research and molecular biology studies, they are also very important in understanding genetic variation in breeding studies [8]. For this reason, the collection, conservation and evaluation of genetic resources will greatly contribute to breeding studies for the development of new varieties [3,8,9]. It is also difficult to study genetic variation using classical methods because of the continuous collection and multiplication of germplasm. The solution is to record genetic diversity using genetic markers, which are not affected by space and time and can measure allele frequencies directly [10–12].

The traditional agarose gel electrophoresis method for visualising amplicons in the detection of genetic diversity using DNA markers is more widely used because it is more economical. In comparison to capillary gel electrophoresis, this method has been reported to be weaker in the identification of genetic diversity [13]. The most important feature of the capillary gel electrophoresis method is that it reliably distinguishes between molecular markers down to 1 or 2 bp and quickly reveals the difference between amplicons [14].

In this study, the genetic diversity of a total of 47 cultivars belonging to both *G. hirsutum* and *G. barbadense* species was investigated using 19 microsatellite markers (EST-SSR) and high-resolution capillary gel electrophoresis. The genetic diversity of the cotton plant, which has an important place in the world and in our country, was the main objective of this study.

2. Materials and Methods

Plant Material

A total of 47 different varieties from the Turkish gene pool were used in the present study (Table 1). Two of these are genotypes of Texas Marker-1 (TM-1/*Gossypium hirsutum*) and Pima 3-79 (*Gossypium barbadense* species with double hoploid characteristics), which are genetic standards. Ten of the genotypes used belong to the species *G. barbadense* and the other 37 genotypes belong to the species *G. hirsutum*. Some of these genotypes have been developed in our country through breeding studies, and some have been brought into our country's genetic resources from other countries and are used in adaptation and breeding studies.

Table 1. Information about the plant material used in the study.

No	Genotype	Species	No	Genotype	Species
1	Adana-98	*G. hirsutum*	25	Nieves	*G. hirsutum*
2	Aydın-110	*G. hirsutum*	26	Özbek-100	*G. hirsutum*
3	BA-320	*G. hirsutum*	27	PG -2018	*G. hirsutum*
4	Babylon	*G. hirsutum*	28	Samon	*G. hirsutum*
5	Carisma	*G. hirsutum*	29	Sezener-76	*G. hirsutum*
6	Carmen	*G. hirsutum*	30	SG-404	*G. hirsutum*
7	Caroline Queen	*G. hirsutum*	31	Sindos-80	*G. hirsutum*
8	Claodia	*G. hirsutum*	32	ST-468	*G. hirsutum*
9	Comos	*G. hirsutum*	33	Veret	*G. hirsutum*
10	Diva	*G. hirsutum*	34	Acala Royale	*G. hirsutum*
11	DP-50	*G. hirsutum*	35	Fibermax-819	*G. hirsutum*
12	Elsa	*G. hirsutum*	36	Tamcot Spinx	*G. hirsutum*
13	Erşan-92	*G. hirsutum*	37	TM-1	*G. hirsutum*
14	Fantom	*G. hirsutum*	38	Pima 3-79	*G. barbadense*
15	Flash	*G. hirsutum*	39	G.B-58	*G. barbadense*
16	Frego	*G. hirsutum*	40	Askabat-71	*G. barbadense*
17	Gaia	*G. hirsutum*	41	Askabat-91	*G. barbadense*
18	Garant	*G. hirsutum*	42	Askabat-100	*G. barbadense*
19	Gloria	*G. hirsutum*	43	Bahar-14	*G. barbadense*
20	Gossypolsüz 86	*G. hirsutum*	44	Bahar-82	*G. barbadense*
21	Gumbo	*G. hirsutum*	45	Giza-45	*G. barbadense*
22	Ligur	*G. hirsutum*	46	Giza-70	*G. barbadense*
23	Lockette	*G. hirsutum*	47	Giza-75	*G. barbadense*
24	Mc Nair 220	*G. hirsutum*			

DNA isolation

DNA was isolated using the method developed for cottonseed by Aydin et al. [15]. To isolate the DNA, 5 seeds of each genotype were crushed under sterile conditions and the bulk method was used. The quality and quantity of the isolated genomic DNA was confirmed by NanoDrop spectrophotometer (Maestrogen, MN-013) and 1% agarose gel electrophoresis. Calculated and normalised amounts of gDNA were used for PCR studies.

PCR Amplification and SSR Markers

A total of 19 SSR markers were used in PCR studies, including 12 EST-SSR primers developed by Wang et al. [16] and 7 EST-SSR markers developed by Karaca and Ince [3] (Table 2). PCR amplifications were performed in a volume of 25 µL. 85 ng of gDNAs, 0.5 µM of each primer pair, 2.5 mM MgCI2, 0.28 mM dNTP, 1 unit of Taq DNA polymerase (Thermo, Cat:EP0402), 2.5 µL of 10X buffer were added. PCR protocol Touch-Down PCR method was used and the temperature was decreased at 0.5 ^{o}C at each cycle step for the first 10 cycles and continued for 30 cycles with the temperature at the 10th cycle. Pre-denaturation was performed at 94 ^{o}C for 3 min, 94 ^{o}C for 30 s, binding temperature at 60 and 66 ^{o}C (primer binding temperature, Table 2) for 45 s, elongation temperature at 72 ^{o}C for 1 min and final elongation at 72 ^{o}C for 10 min [17]. Amplifications were performed using a Thermo Fisher Scientific (Ref: A24812) thermal cycling device in PCR procedures.

Table 2. EST-SSR primers used in 47 cotton genotypes.

Primer ID	Primer Sequence	Motif	Alellel Number	Tm (^{o}C)	H_e	H_0	*PIC*
GA01-2651	F:AAATCCTACCTCTCCGGCCA R: CCCAGGGCAAAACAATGTCG	$(GCCGGC)_3$	6	60	0,485	0,479	0,475
GA02-54	F:GGGAAAGCGCGTCATTGATC R: GCCGAGCCCAGACCTAATAG	$(GCCGG)_4$	4	60	0,450	0,096	0,437
GA04-1418	F:GCAGGCAGAATACAAAAGATCGA R: AAGAAAAGGGGGAGGGGAGA	$(GCCGGC)_3$	33	60	0,737	0,596	0,730
GA07-410	F:GAAATTACCTTTCCGGCCACC R: GACGTCGTTTTGGAGGGCTA	$(GCCGGC)_3$	24	60	0,720	0,528	0,708
GA08-323	F:GAACCGACCTAAGGTGACTGT R: AGAGAGAAGGGAGGGGGAAG	$(CCGGCG)_3$	6	60	0,476	0,266	0,473
GD01-295	F:TTTACCTTTCCGACCACCGC R: GGTGCGTTTTGGTCCCCTAT	$(GCCGGC)_3$	7	60	0,704	0,415	0,693

GD02-301	F:AAACCCGTTGTGCAACCATG R: GGATGAGGCTGAGAAGGAGC	$(CCGGTG)_3$	15	60	0,898	0,883	0,889
GD03-2002	F:GGCCCGGCCCGAATATAATA R: GACTAGACCTGTCCATGGGC	$(CCGGGC)_3$	19	60	0,798	0,539	0,788
GD06-2808	F:GGCCCGGCCCGAATATAATAA R: CGGCCCGAAATATGGGCTTA	$(CCGGGC)_3$	14	60	0,625	0,560	0,616
GD08-420	F:AGAGAAGGGAGGGGGAAAGG R: GGGCTCTAACACCAAATCGGA	$(CGCCGG)_3$	42	60	0,741	0,571	0,738
GD09-1296	F:GGCGCACAAAACACCAAGAT R: AGGGAGGAAGGAAAGGGGG	$(GCCGGC)_3$	14	60	0,745	0,546	0,737
GD10-1664	F:GGGCTCTAACACCAAATCGGA R: AGAGAAGGGAGGGGGAAAGG	$(GCCGGC)_3$	29	60	0,679	0,493	0,675
MK086	F:CCACCAGTTTGGTAGGTATGAAC R: TCAACAGTGCAAGGACTTCATC	$(CAT)_8$	2	66	0,293	0,121	0,268
MK105	F:CAAAGATGCCGAAAGAGAGG R: GTAAGATCGGCGGGTCATC	$(CCG)_{12}$	30	66	0,666	0,319	0,659
MK126	F:ACCGTACCCGTGGCTCTTAT R: TGTTGTTGTGGGAGGCTTCT	$(CAT)_8$	4	66	0,414	0,043	0,378
MK129	F:GCTGATGCTGATTCCTCCAT R: TGCCCTTCATCTCGTTTCTT	$(CAA)_8$	7	66	0,373	0,479	0,345
MK132	F:AGCAAGGCATGAGCGATACT R: GGTGGTACCTTCCCATGTTG	$(TCAGCC)_6$	8	66	0,541	0,391	0,511
MK146	F:ATGGAGGCTGCAAAGACTGT	$(GTAGTGAGA)_3$	3	66	0,577	0,787	0,487

	R: CCACTCCGACTAAAAGATCAGC						
MK173	F:GGGGTCCACAGATACAGG R: GTCCAAAACTTGTCCCATTAG	(TATG)$_9$	13	66	0,867	0,106	0,858

(H$_e$: expected heterozygosity, H$_o$: observed heterozygosity, *PIC*: Polymorphism Information Content).

Detection of PCR Amplicons

The PCR amplicons were determined by capillary gel electrophoresis. With this method, PCR products were performed on the QIAxcel Advanced instrument (Qiagen, Cat.No./ID:90001941) using the QIAxcel DNA High Resolution Kit (Cat.No. 929002). Amplicons were analysed using QIAxcel ScreenGel version 1.6.

Data Analysis

Different statistics were performed on the data obtained by analyzing the amplicons. Polymorphism information content (PIC) for each marker [8], Jaccard's coefficient with the Multi-Variate Statistical Package (MVSP, version 3.13O, Kovach Computing Services, Pentraeth, UK) for PCoA analysis, Bayesian statistics with MrBayes software v3.2.1 x64 for phylogenetic dendogram [18] and Structure software version 2.3.4 for population structure, and the results were extracted using the STRUCTURE HARVESTER online tool (http://taylor0.biology.ucla.edu/structureHarvester/) to calculate the optimal K value [19,20]. PowerMarker v3.25 software was used for expected heterozygosity (H$_e$) and observed heterozygosity (H$_o$) analyses [21].

3. Results

3.1. Polymorphism Analysis of EST-SSR Markers

In this study, 19 EST-SSR markers were used to screen 47 cotton genotypes. Accordingly, primers used were observed to generate between 2 and 42 amplicons (Table 2). While the lowest number of alleles was observed in primer MK086, the highest number of alleles was observed in primer GD08-420. A total of 280 alleles were analyzed from 19 EST-SSR primers. The average number of amplicons per primer was calculated to be 14.7. In 47 genotypes, the primers used in markers MK086, MK126, MK129 had the lowest PIC values. PIC values were very high in the remaining 16 primers. The PIC values ranged from 0.268 (MK086) to 0.889 (GD02-301). While the average PIC value was 0.603, 11 EST-SSR markers (GA04-1418, GA07-410, GD01-295, GD02-301, GD03-2002, GD06-2808, GD08-420, GD09-1296, GD10-1664, MK105, MK173) were found to be above this average.

The expected average ratio of heterozygosity for all the primers was 0.620 and the observed average ratio of heterozygosity was 0.433. The lowest heterozygosity ratio was 0.293 for the MK086 marker and the highest was 0.898 for the GD02-301 marker. In addition, a total of 11 EST-SSR markers were above average. Similarly, marker GD02-301 had the highest Ho value of 0.883, whereas the lowest Ho value was 0.043 for marker MK126.

3.2. Clustering and PCoA Analysis

Corresponding to the amplicons generated with 19 EST-SSR markers, data from 47 cotton genotypes used in the study were analysed. In this regard, The phylogenetic tree was constructed using Bayesian statistics with the programme MrBayes. (Figure 1). The tree presented by the Bayesian statistics has been coloured with the programme FigTree v1.4.4.
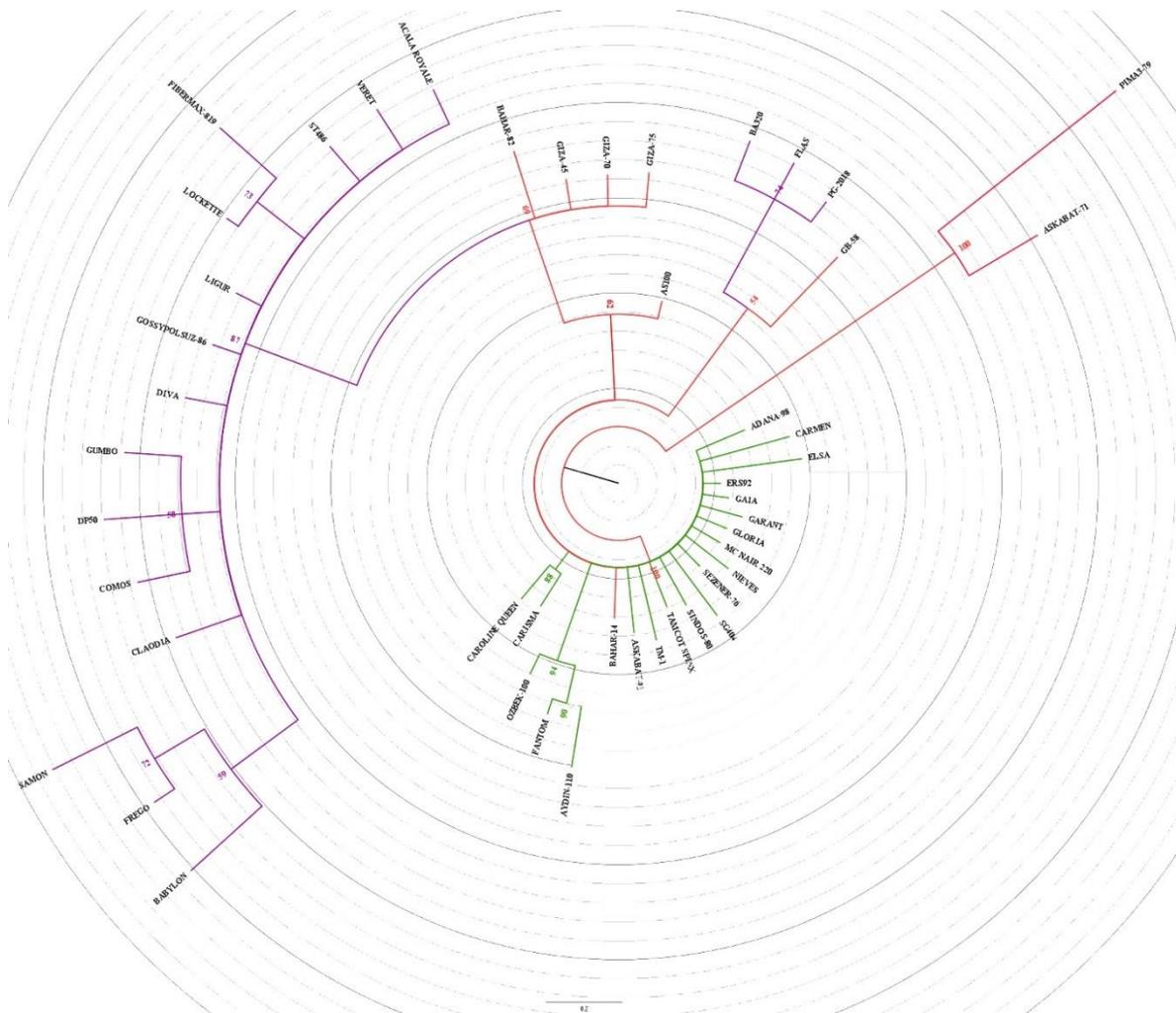
Figure 1. Constructed with Bayesian Statistics Consensus Tree. (The polar tree is presented in unrooted format, values shown at the nodes Post-Probability values, bar below the tree indicates the base change scale).

In the clustering analysis, the genotypes Pima3-79 *(G. barbadense)* and Askabat-71 (*G. barbadense*) differed from the other genotypes with a post-probability value of 100%.

The remaining 45 genotypes were divided into different groups. They were coloured green, red and purple. The green and purple coloured ones are the genotypes that belong to the *G. hirsutum* species and the red coloured ones are the genotypes that belong to the *G. barbadense* species. Of the green coloured genotypes, *G. barbadense* species and cultivars are known to be Bahar-14 and Askabat-91. On the other hand, although there were genotypes belonging to the red coloured *G. barbadense* species among the purple stained genotypes, their clustering was observed as a different group. Only genotype GB-58 was grouped in a cluster with BA320, Flas and PG-2018 with a post-probability value of 54%.

Principal Coordinates Analysis (PCoA) was performed on the EST-SSR data using Jaccard's similarity matrix to provide a different view of the genetic relationships between genotypes (Figure 2).
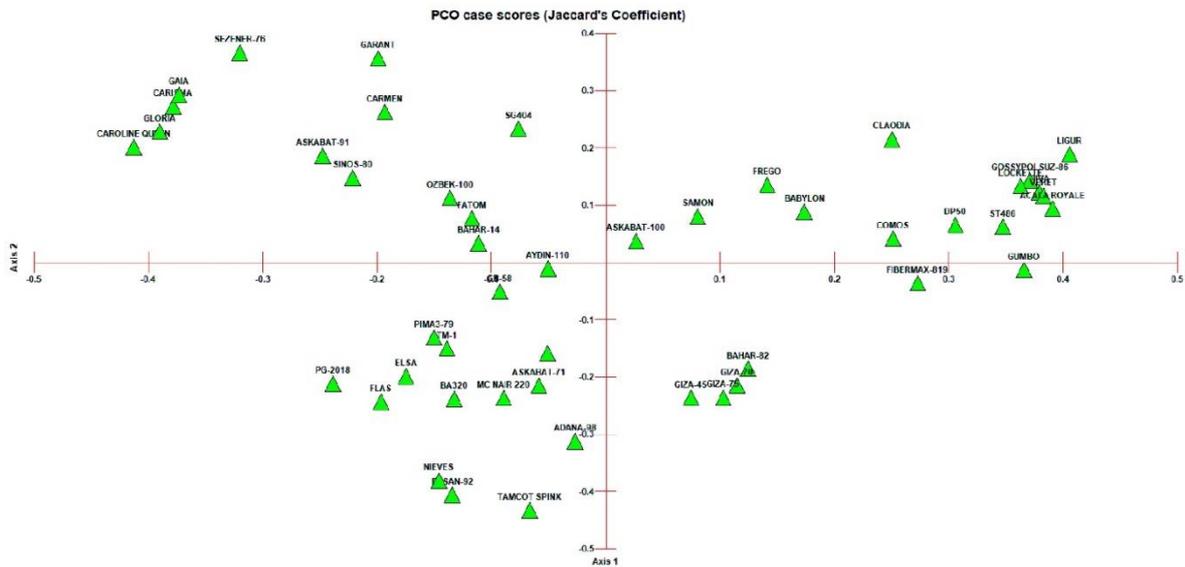
Figure 2. PCoA analysis based on Jaccard's similarity index values (Axis1 and Axis2 show 6.769% and 11.914% variation, respectively).

The PCoA plot shows that the genotypes are distributed almost evenly. Only in the lower right corner of the Axis2 axis (the positive part of the X-axis and the negative part of the Y-axis) was the number of genotypes collected low, and more than half of the genotypes were cultivars belonging to the *G. barbadense* species. Some genotypes grouped together in certain regions. Among these, Caroline Queen, Gloria, Carisma, Gaia and Sezener-76 genotypes were grouped in the upper left shelf of Axis2; Ligur, Gosyypol-free-86, Lockette, Diva, Veret and Acala Rpyale genotypes were closely grouped in the upper right shelf of Axis1 (Figure 2).
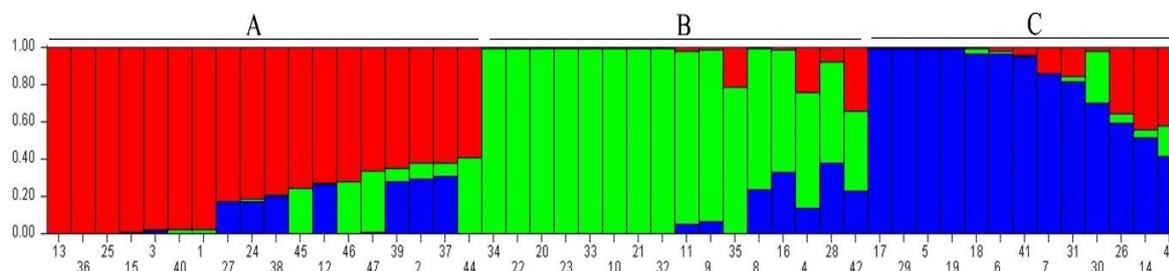
3.3. Population Structure Analysis

A population analysis was carried out on 47 cotton genotypes with 280 alleles belonging to 19 markers obtained by means of EST-SSR markers. SUTRUCTURE v2.3.4 software was used to perform Bayesian clustering for population analysis. Using the STRUCTURE software, the analysis was carried out with 10000 burn-in periods and 100000 replicates, and the Delta K value was 3 (Figure 3). This result showed that the 47 cotton genotypes used were divided into 3 groups. Genotypes with a membership coefficient of 0.8 and greater than 0.8 were considered pure [21]. Of the 47 cotton genotypes used, only 30 were identified as pure in this case. According to the number in Table 1; genotypes 1, 3, 5, 6, 7, 9, 10, 11, 13, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 29, 30, 31, 32, 33, 34, 36, 38, 40, and 41 were considered pure.
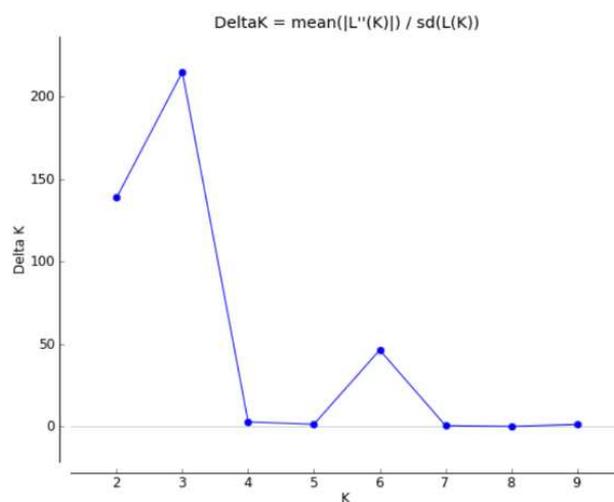
Figure 2. Line graph of ∆K values changing with K value.

The groupings by blotting are also shown in Figure 4. According to this figure, 18 genotypes were calculated in group A, 16 in group B and 13 in group C. The group with the lowest number of genotypes is group C and in this group there are 2 genotypes of *G. barbadense* (Ashgabat-91, Bahar-14) in contrast to *G. hirsutum*. In group B, only 1 genotype out of 16 belongs to *G. barbadense* (Askabat-100). Group A is the most populated group and contains 18 genotypes, of which 7 belong to *G. barbadense*.

Figure 2. Q-plot of the population structure for 47 cotton genotype.

Pima -79 and G.B-58 genotypes of *G. barbadense* species in group A were confused with group C, while genotypes Askabat-71, Giza-45, Giza-70 and Giza-75 were confused with group B. Bahar-82 (*G. barbadense*) genotype in group B showed confusion with group A. However, Askabat-100 (*G. barbadense*) genotype showed confusion with both groups A and C. Apart from this, Askabat-91 (*G. barbadense*) genotype in group C showed some confusion with group A.

4. Discussion

Effects of Molecular Markers on Population

Using DNA sequences and regions of the genome with different characteristics, many marker techniques have been developed. The choice of marker techniques is in accordance with the purpose of the study [22]. SSR markers are the preferred choice for studies such as genetic diversity and the identification of QTLs [23–25]. These markers have been developed for many crops and cotton since the discovery of EST-SSRs by Cardle et al. [3,26–28]. Some of the reasons why EST-SSR markers are the most preferred are that they are reliable and reproducible, co-dominant, cheap and easy to use, and easily transferable between species [29]. It has been extensively used for genetic diversity in cotton [8,30,31], linkage studies [32], determination of abiotic stress tolerance [33], mapping [34,35].

In this study, a total of 19 EST-SSR markers were used, which were developed for cotton by Wang et al. [16] and Karaca and Ince [3]. As the markers used are located in the expressed regions of the genome, this gives us the information that these markers have a high population discriminatory power, depending on the population used. Karaca and Ince [3] found that markers MK086, MK132, MK146 and MK173 were monomorphic in *G. hirsutum* and *G. barbadense* species by agarose gel electrophoresis (AGE). They reported that when converted to the CAPS marker (using the *Hinf* I restriction enzyme), only the MK129 marker was polymorphic. However, with the method of capillary gel electrophoresis (CGE) that we have used, we have observed that these markers form direct polymorphic amplicons. PIC values ranged from 0.268 to 0.858 and the mean PIC value was calculated as 0.501. Since the CGE method has a higher resolution, it is more advantageous than the AGE method and its information content has been reported by various researchers [13,36,37]. The EST-SSRs that were developed by Wang et al. [16] and the EST-SSRs that were used in our study have, in general, very similar repeat sequences. However, the PIC values of the markers ranged from 0.437 to 0.889. The average value was 0.663. In addition, the markers that were developed were located on different chromosomes of the A and D genomes that were selected at random. The PIC values of the markers in the D genome were also found to be higher than in the A genome in this study. Molecular data reported that all allopolyploids in *Gossypium* share a common ancestry,

supporting the hypothesis that polyploid formation occurs only once [38]. Furthermore Wendel [39] reported that each allopolyploid genome contains the chloroplast of genome A, the old world cotton. The differences in genome D may therefore be responsible for the higher PIC values in genome D.

Cluster Analysis and Population Structure

For the conservation and utilisation of genetic resources and for breeding studies, the determination of genetic diversity is very important [8,21]. Genetic variability, resulting from genetic relatedness and genetic diversity between plant groups, is critical to the success of plant breeding [40,41]. Better parental selection to produce different varieties can greatly improve breeding studies. The use of markers is important in genetic studies to identify heterotic groups, understand population structure, and distinguish between basic lineages [10]. To reveal the diversity of genotypes in our unit, both cluster analysis and population structure were used.

The genetic relationship between *G. hirsutum* and *G. barbadense* was analysed by cluster analysis using Bayesian statistics. In the analysis, except for 2 genotypes (Bahar-14 and Askabat-91) belonging to *G. barbadense* species, the remaining genotypes were grouped differently. Furthermore, Aydın [15] reported that some seeds of Askabat-91 genotype had the same characteristics as *G. hirsutum*. There is therefore a high probability that seeds of this genotype are contaminated. Except for Pima 3-79, the other 9 genotypes of *G. barbadense* (Giza-45, Giza-70, Giza-75, GB-58 originating from Egypt; Bahar-14, Bahar-82, Askabat-71, Askabat-91 and Askabat-100 originating from Turkmenistan) are cultivated for their fiber (Aydın, 2018). Clustering analysis separated Pima 3-79 and Aksabat-71 from other genotypes with 100% power. It has been reported that the Pima 3-79 genotype is a doubled haploid and is considered to be the genetic standard [3,15]. Therefore, the fact that this genotype is in another cluster can be related to Bayesian statistics. Accordingly, genotype Askabat-71 is likely to contain high levels of genetic material belonging to the *G. barbadense* species. The clusters were determined by colouring the phylogenetic tree. Here, genotypes coloured purple and green are genotypes belonging to *G. hirsutum* species, while those coloured red are genotypes belonging to *G. barbadense* species. Genotypes belonging to *G. hirsutum* species formed two distinct clusters in the clustering analysis. While the purple cluster contained 18 genotypes, the green cluster separated them with 19 genotypes. The level of variation is also negatively affected by the fact that cotton's genetic base has a narrow genetic structure compared to other crops [31,42,43]. The development of cotton varieties is generally based on the use of existing varieties. This leads to the protection of the narrow genetic structure. In terms of population structure, the 47 genotypes in use were divided into 3 main groups. Of these, 30 genotypes were accepted as pure and the remaining 17 genotypes were concluded to be mixed. When this analysis is examined, it can be interpreted that heterozygosity is low. The high number of pure individuals is an indication that the variation is low and new studies should be initiated to increase this variation. Although pure lines are used in crossbreeding studies, researchers conducting breeding programmes prefer to have a high level of variation in order to develop varieties with the desired characteristics.

5. Conclusions

In this study, different statistical methods were used to investigate the genetic diversity of *G. hirsutum* and *G. barbadense* genotypes. Although the PIC values of the EST-SSR markers used were found to be high for this population, PCoA, Clustering and population structure analyzes gave similar results and It was concluded that the genetic diversity of the species was low. Various approaches to recording genetic resources and increasing genetic variation should be used in these studies and analyses. The use of EST-SSR markers in QTL studies is also extremely important. As these markers are the parts of the gene that can be directly expressed, they will be useful in developing cotton varieties with the desired characteristics.

References

1. Wang, K., Wendel, J. F., & Hua, J. (2018). Designations for individual genomes and chromosomes in *Gossypium*. *J. Cotton Res.*, 1(1), 1-5.

2. Witt, T. W., Ulloa, M., Schwartz, R. C., & Ritchie, G. L. (2020). Response to deficit irrigation of morphological, yield and fiber quality traits of upland (*Gossypium hirsutum* L.) and Pima (*G. barbadense* L.) cotton in the Texas High Plains. *Field Crops Res.*, 249, 107759.

3. Karaca, M., & Ince, A. G. (2011). New non-redundant microsatellite and CAPS-microsatellite markers for cotton (*Gossypium* L.) *Turk. J. Field Crops.*, 16(2), 172-178.

4. Kumlay, A. M. , Eren, B. , Demirel, S. , Demirel, F. & Yıldırım, B. (2021). Reveal of Genetic Variation by iPBS Analysis in Some Cotton Varieties. *Eur. j., eng. sci., tech.*, (21) , 67-73 . DOI: 10.31590/ejosat.809479

5. Rashid, U., Anwar, F., & Knothe, G. (2009). Evaluation of biodiesel obtained from cottonseed oil. *Fuel Process. Technol.*, 90(9), 1157-1163.

6. Sekhar, S. C., & Rao, B. V. K. (2011). Cottonseed oil as health oil. *Pertanika J. Trop. Agric*. Sci, 34(1), 17-24.

7. Guo, Q., Liu, J., Li, J., Cao, S., Zhang, Z., Zhang, J., ... & Li, Y. (2022). Genetic diversity and core collection extraction of *Robinia pseudoacacia* L. germplasm resources based on phenotype, physiology, and genotyping markers. *Ind Crops Prod.*, 178, 114627.

8. Shi, J., Zhang, Y., Wang, N., Xu, Q., Huo, F., Liu, X., & Yan, G. (2023). Genetic diversity and population structure analysis of upland cotton (*Gossypium hirsutum* L.) germplasm in China based on SSR markers. *Genet. Resour. Crop Evol.*, 1-12.

9. Ghuge, S. B., Mehetre, S. S., Chimote, V. P., Pawar, B. D., & Naik, R. M. (2018). Molecular characterization of cotton genotypes using SSR, ISSR and RAPD markers in relation to fiber quality traits. *J. Cotton Res.*, 32(1), 1-12.

10. Tyagi, P., Gore, M. A., Bowman, D. T., Campbell, B. T., Udall, J. A., & Kuraparthy, V. (2014). Genetic diversity and population structure in the US Upland cotton (*Gossypium hirsutum* L.). *Theor. Appl. Genet.*, 127, 283-295.

11. Moiana, L. D., Filho, P. S. V., Gonçalves-Vidigal, M. C., & de Carvalho, L. P. (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(2), 143-152.

12. Morales, N., Ogbonna, A. C., Ellerbrock, B. J., Bauchet, G. J., Tantikanjana, T., Tecle, I. Y., ... & Mueller, L. A. (2022). Breedbase: a digital ecosystem for modern plant breeding. G3, 12(7), jkac078.

13. Zhao, X., Zhang, Q., Tang, X., Xu, Y., Du, Y., Zhao, B., ... & Zhao, K. (2019). Multiplex detection of transgenic maize by microdroplet PCR combined with capillary gel electrophoresis. *Acta Biochim. Biophys. Sin.*, 51(5), 535-538.

14. Cheng, B., Xia, J., Gong, J., & Yang, S. (2011). Application of capillary electrophoresis detection with fluorescent SSR markers in rice DNA fingerprint identification. *Chin. J. Rice Sci.*, 25(6), 672-676.

15. Aydın, A. (2018). A Study on the Molecular Characterization of Some Commercial Cotton Varieties Registered in Turkey. Ph.D. Dissertation. Akdeniz University (Antalya/TÜRKİYE).

16. Wang, Q., Fang, L., Chen, J., Hu, Y., Si, Z., Wang, S., ... & Zhang, T. (2015). Genome-wide mining, characterization and development of microsatellite markers in *Gossypium* species. *Sci. Rep.*, 5(1), 10638.

17. Ince, A. G., & Karaca, M. (2012). Species-specific touch-down DAMD-PCR markers for Salvia species. *J Med Plants Res*, 6, 1590-1595.

18. Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., ... & Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space., Syst. Biol., 61(3), 539-542.

19. Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genet.*, 155(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

20. Eren, B., Keskin, B., Demirel, F., Demirel, S., Türkoğlu, A., Yilmaz, A., & Haliloğlu, K. (2023). Assessment of genetic diversity and population structure in local alfalfa genotypes using iPBS molecular markers. *Genet. Resour. Crop Evol.*, 70(2), 617-628.

21. Demirel, S., & Demirel, F. (2023). Molecular identification and population structure of emmer and einkorn wheat lines with different ploidy levels using SSR markers. *Genet. Resour. Crop Evol.*, 1-10.

22. Karaca, M., Ince, A. G., Aydin, A., Elmasulu, S. Y., & Turgut, K. (2015). Microsatellites for genetic and taxonomic research on thyme (*Thymus* L.). *Turk. J. Biol.*, 39(1), 147-159.

23. Nadeem, M. A., Nawaz, M. A., Shahid, M. Q., Doğan, Y., Comertpay, G., Yıldız, M., ... & Baloch, F. S. (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.

24. Sabev, P., Valkova, N., & Todorovska, E. G. (2020). Molecular markers and their application in cotton breeding: progress and future perspectives. *Bulg. J. Agric. Sci.*, 26(4).

25. Mishra, N., Tripathi, M. K., Tiwari, S., Tripathi, N., Gupta, N., Sharma, A., ... & Tiwari, S. (2022). Characterization of soybean genotypes on the basis of yield attributing traits and SSR molecular markers. book: Innovations in Science and Technology. 2022b, 3, 87-106.

26. Cardle, L., Ramsay, L., Milbourne, D., Macaulay, M., Marshall, D., & Waugh, R. (2000). Computational and experimental characterization of physically clustered simple sequence repeats in plants. *Genet.*, 156(2), 847-854.

27. Saha, S., Karaca, M., Jenkins, J. N., Zipf, A. E., Reddy, O. U. K., & Kantety, R. V. (2003). Simple sequence repeats as useful resources to study transcribed genes of cotton. *Euphytica*, 130, 355-364.

28. Qureshi, S. N., Saha, S., Kantety, R. V., & Jenkins, J. N. (2004). EST-SSR: a new class of genetic markers in cotton. *J. Cotton Sci.*, 8, 112–123

29. Karaca, M., Ince, A. G., Aydin, A., & Ay, S. T. (2013). Cross-genera transferable e-microsatellite markers for 12 genera of the *Lamiaceae* family. *J. Sci. Food Agric.*, 93(8), 1869-1879.

30. Noormohammadi, Z., Rahnama, A., & Sheidai, M. (2013). EST-SSR and SSR analyses of genetic diversity in diploid cotton genotypes from Iran. *Nucleus*, 56, 171-178.

31. Saleem, M. A., Amjid, M. W., Ahmad, M. Q., Riaz, H., Arshad, S. F., & Zia, Z. U. (2020). EST-SSR based analysis revealed narrow genetic base of in-use cotton varieties of Pakistan. *Pak. J. Bot*, 52(5), 1667-1672.

32. 3Hinchliffe, D. J., Turley, R. B., Naoumkina, M., Kim, H. J., Tang, Y., Yeater, K. M., ... & Fang, D. D. (2011). A combined functional and structural genomics approach identified an EST-SSR marker with complete linkage to the Ligon lintless-2 genetic locus in cotton (*Gossypium hirsutum* L.). *BMC genomics*, 12, 1-18.

33. Wang, B. H., Zhu, P., Yuan, Y. L., Wang, C. B., Yu, C. M., Zhang, H. H., ... & Li, P. (2014). Development of EST-SSR markers related to salt tolerance and their application in genetic diversity and evolution analysis in *Gossypium*. *Genet. Mol. Res*, 13(2), 3732-3746.

34. Han, Z. G., Guo, W. Z., Song, X. L., & Zhang, T. Z. (2004). Genetic mapping of EST-derived microsatellites from the diploid *Gossypium arboreum* in allotetraploid cotton. *Mol. Genet. Genom.*, 272, 308-327

35. Han, Z., Wang, C., Song, X., Guo, W., Gou, J., Li, C., ... & Zhang, T. (2006). Characteristics, development and mapping of *Gossypium hirsutum* derived EST-SSRs in allotetraploid cotton. *Theoretical and Applied Genetics.*, 112, 430-439.

36. Sánchez-Pérez, R., Ballester, J., Dicenta, F., Arús, P., & Martínez-Gómez, P. (2006). Comparison of SSR polymorphisms using automated capillary sequencers, and polyacrylamide and agarose gel electrophoresis: implications for the assessment of genetic diversity and relatedness in almond. *Sci. Hortic.*, 108(3), 310-316.

37. McCudden, C. R., Mathews, S. P., Hainsworth, S. A., Chapman, J. F., Hammett-Stabler, C. A., Willis, M. S., & Grenache, D. G. (2008). Performance comparison of capillary and agarose gel electrophoresis for the identification and characterization of monoclonal immunoglobulins. Am. J. Clin. Pathol., 129(3), 451-458.

38. Wendel, J. F., Brubaker, C. L., & Seelanan, T. (2010). The origin and evolution of *Gossypium. In Physiology of cotton* (pp. 1-18). Dordrecht: Springer Netherlands.

39. Wendel, J. F. (1989). New World tetraploid cottons contain Old World cytoplasm. *Proc. Natl. Acad. Sci.*, 86(11), 4132-4136.

40. Rana, M. K., & Bhat, K. V. (2005). RAPD markers for genetic diversity study among Indian cotton cultivars. *Current Science*, 1956-1961.

41. Abdellatif, K. F., & Soliman, Y. A. (2013). Genetic relationships of cotton (*Gossypium barbadense* L.) genotypes as studied by morphological and molecular markers. Afr. J. Biotechnol., 12(30).

42. Zhang, J., Percy, R. G., & McCarty, J. C. (2014). Introgression genetics and breeding between Upland and Pima cotton: a review. *Euphytica*, 198, 1-12.

43. Li, Y., Si, Z., Wang, G., Shi, Z., Chen, J., Qi, G., ... & Zhang, T. (2023). Genomic insights into the genetic basis of cotton breeding in China. *Mol Plant*, 16(4), 662-677.