TÜBİTAK

Turkish Journal of Botany

http://journals.tubitak.gov.tr/botany/

Research Article

Turk J Bot
(2014) 38: 1274-1286

doi:10.3906/bot-1405-78

Genetic diversity of Turkish commercial cotton varieties revealed by molecular markers and fiber quality traits

Eminur ELÇİ[1,*], Yaşar AKIŞCAN[2], Batuhan AKGÖL[1]

[1]Department of Research and Development, ProGen Seed Company, Antakya, Hatay, Turkey
[2]Department of Field Crops, Faculty of Agriculture, Mustafa Kemal University, Antakya, Hatay, Turkey

Received: 27.05.2014 • **Accepted:** 30.10.2014 • **Published Online:** 17.11.2014 • **Printed:** 28.11.2014

Abstract: To assess the genetic diversity and relationships among commercial *Gossypium* species released in Turkey between 1964 and 2014, 96 cotton varieties were analyzed using morphological and molecular markers. Morphological analysis was performed based on 4 fiber quality traits including fiber length, strength, fineness, and uniformity, and the mean values of each trait for each genotype were calculated using 2-year data. The results showed that most of the genotypes have long fiber length, very high fiber strength, coarse (45 genotypes) or average (50 genotypes) fiber fineness, and high uniformity. Twenty-six simple sequence repeat (SSR) markers and 14 markers linked to quantitative trait loci (QTLs) for fiber quality traits produced a total of 103 alleles, with an average of 2.57 alleles per locus ranging from 80 bp to 300 bp products, with an average polymorphism information content (PIC) value of 0.233. Markers DPL513 and DPL431 (among 26 SSR markers) and markers CIR246 and BNL4108 (among 14 molecular markers) were found to be very informative, with 0.724, 0.663, 0.749, and 0.583 PIC values, respectively. The combined morphological and molecular data analysis resulted in more than 8 clades using the unweighted pair group method with arithmetic average (UPGMA). The upland cotton varieties were distinctly separated from the lowland cotton variety Maydos Yerlisi (*Gossypium herbaceum* L.). Within the upland cotton varieties, the Egyptian cotton variety Giza 70 (*G. barbadense* L.) was distinctly separated from commercial cotton varieties of Turkey (*G. hirsutum* L.), as revealed by both morphological and molecular dendrograms. Principal component analysis (PCA) derived from combined data was in agreement with UPGMA analysis. It is concluded that commercial Turkish cotton varieties have a good genetic diversity with high fiber quality, considering the upland cotton's narrow genetic structure. These results can provide a useful guide for selecting specific germplasm with distinct genetic backgrounds in cotton breeding programs.

Key words: *Gossypium* spp., HVI, PIC, UPGMA, PCA, Turkey

1. Introduction

Cotton, as an annual crop, is the world's leading natural fiber crop and an important crop for bioenergy production (Lusas and Jividen, 1987; Chen et al., 2007). It belongs to the genus *Gossypium* of the family Malvaceae and consists of approximately 50 species, including 45 diploids (2n = 26) and 5 allotetraploids (2n = 52) (Fryxell et al., 1992). The species are grouped into A through G and K genomes, based on chromosome structures (Endrizzi et al., 1984). The allotetraploid species arise through the hybridization of an A-genome taxon related to *Gossypium herbaceum* L. (2n = 2x = 26), with a D-genome taxon related to *G. raimondii* Ulbrich and *G. gossypioides* L. (2n = 2x = 26) (Beasley, 1942; Wendel et al., 1992). They consist of *G. barbadense* L., *G. darwinii* Watt, *G. hirsutum* L., *G. tomentosum* Nuttall, and *G. mustelinum* Miers ex Watt (Percival et al., 1999; Wendel and Crohn, 2003).

Two economically important cultivated tetraploid species of cotton, *G. hirsutum* (also known as "Upland" cotton) and *G. barbadense* (Caribbean "Sea-Island", "Extra Long Staple", and modern "Pima" and "Egyptian" cultivars), dominate world cotton production. *G. hirsutum* L. is the principal cultivated cotton and accounts for about 90% of the world's cotton production (Chen et al., 2007). Conventional breeding methods generally aim to improve agronomically important traits by combining characters present in different parental lines of cultivated species or their wild relatives. The prediction of genetic similarities among genotypes is very important for crop improvement and accurate selection of parental combinations and the maintenance of sufficient diversity in breeding programs is necessary (Lacape et al., 2010).

Cotton fibers are highly elongated single cells of the epidermal surface of the seed and they are widely used as the raw material in the textile industry. Fiber quality

* Correspondence: eminur@gmail.com

is defined by characteristics of the cotton fiber that allows its processing into yarn and end products (Shen et al., 2011). Different physical characteristics of cotton fibers are measured ranging from fiber length and length uniformity, strength, elongation (degree of extensibility), maturity (extent of cell wall thickening), micronaire (resistance to air flow across a plug of fibers) and fineness (linear density, a function of diameter and thickness), to color indices (reflectance and yellowness) (Lacape et al., 2010). These mentioned characteristics are associated with the proficient spinning and weaving processes that alter the fiber into fabrics. Therefore, it is very important to improve fiber quality in locally dominating cotton genotypes to accomplish the requirements of the growing textile industry, processing and end uses (Ali et al., 2008).

More than a hundred quantitative trait loci (QTLs) influencing fiber quality properties related to fiber length, length uniformity, micronaire, and strength were identified (Chee and Campbell, 2009). These QTLs allow the estimation of the number of genes, their locations, and the phenotypic and genetic effects of individual QTLs on fiber traits. Identified DNA markers, tightly linked to fiber quality QTLs, promise to assist breeders in selecting genotypes from a large heterogeneous population for desired and valuable allele combinations during cotton cultivar development (Shen et al., 2011; Zhang et al., 2011).

Molecular markers originate from DNA mutations such as point mutations, translocations, duplications, insertions, or deletions that generally occur in noncoding regions (Mondini et al., 2009) and provide a useful tool for detection, efficient evaluation, and selection of plant materials (Kumar et al., 2009). Microsatellites, also known as SSRs, are highly polymorphic and tandemly repeated sequences of DNA, comprising basic short motifs generally varying between 2 and 6 base pairs in genome. They have advantages because they possess co-dominant inheritance, high allelic diversity, high abundance, and high reproducibility (Mondini et al., 2009).

The aim of the present study was to determine the genetic diversity of commercial cotton varieties released in Turkey since 1964, through morphological fiber traits and molecular SSR markers. We evaluate some important fiber quality characteristics and analyze the genetic profile of cotton varieties, using SSR markers and molecular markers linked to QTLs for fiber quality traits.

2. Materials and methods

2.1. Plant material

Ninety-six cotton (94 *G. hirsutum* L., 1 *G. herbaceum* L., and 1 *G. barbadense* L.) cultivars released in Turkey (except *G. barbadense* L.) between 1964 and 2014 (Table 1) were collected from preserved gene banks and breeding companies. Seeds of Giza 70 (*G. barbadense* L., Egyptian cotton) were used as an out-group control in the genetic analysis.

2.2. Morphological analysis

2.2.1. Field experiments

The seeds of 96 varieties were sown in the field under randomized complete block design with 3 replications during the 2012 and 2013 growing seasons in the Karaali region of Hatay, Turkey. All of the recommended agronomic and plant protection practices were followed from sowing to harvesting of the cotton crop. The matured bolls were collected from the first position of the middle fruiting branches of each individual plant and used for quality analysis.

2.2.2 Fiber quality analysis

The collected boll samples were ginned with a single roller electrical gin on an individual plant basis to obtain lint for fiber analysis. Before fiber quality analysis, lints were conditioned at 21 ± 1 °C and 65 ± 2% relative humidity for 48 h in a controlled room. An HVI 1000 (Uster, Switzerland) was used to analyze fiber quality traits and the most important cotton fiber properties, i.e. fiber length (mm), fiber strength (g tex^{-1}), fiber fineness (micronaire), and fiber uniformity (%), were examined.

2.3. Molecular analysis

2.3.1. Genomic DNA extraction

The DNA was extracted from leaves by cetyl trimethylammonium bromide (CTAB) extraction method (Doyle and Doyle, 1987) with a few modifications. For each 100 mg of tissue, 300 µL of CTAB isolation buffer (2% hexadecyltrimethylammonium bromide, 1.4 M NaCl, 0.2% β-ME, 20 mM EDTA, 100 mM Tris-HCl, pH 8) was added to each tube, and homogenized by TissueLyser (Qiagen, Germany). More CTAB extraction buffer (450 µL) was added to each tube and the samples were incubated at 65 °C for 60 min with occasional mixing. Due to the high content of polyphenolic compounds in cotton tissues, 750 µL of phenol/chloroform/isoamyl alcohol (25:24:1 v/v) was added to each sample and the samples were vortexed and then centrifuged. The supernatants were transferred to a new tube and 500 µL of chloroform/isoamyl alcohol (24:1 v/v) solution was added. Next, 500 µL of ice-cold isopropanol was added to each tube and the tubes were incubated for 30 min at room temperature. The samples were centrifuged and the supernatants were discarded. The pellets were air-dried and then resuspended in 50 µL of 10 mM Tris, pH 8.0, 1 mM EDTA buffer. Nucleic acids were measured quantitatively and qualitatively by spectrophotometer NanoDrop 2000c (Thermo Fisher Scientific, Waltham, MA, USA). The extracted DNA was stored at –20 °C.

Table 1. Fiber quality traits and registration information of Turkish cotton varieties analyzed in this study.

No.	Variety name	Fiber quality traits				Registration year
		FL mm	FS g/tex	FF mic	FU %	
1	BA-525	Long	Very high	Coarse	High	2006 [a]
2	Nazilli 84	Long	Very high	Average	High	1984 [b]
3	DP 419	Long	Very high	Average	High	2007 [c]
4	DP 5409	Long	Very high	Average	High	1999 [c]
5	Claudio	Long	Very high	Average	Very high	2010 [d]
6	Sandra	Long	Very high	Coarse	Very high	2010 [e]
7	Menderes 2005	Long	Very high	Average	Very high	2005 [b]
8	NATA	Long	Very high	Average	High	1999 [f]
9	Nazilli 663	Long	Very high	Coarse	High	2003 [b]
10	DP 5690	Long	Very high	Coarse	High	1999 [c]
11	Ege-69	Long	Very high	Average	High	1977 [b]
12	ST 488	Long	Very high	Average	Very high	2007 [f]
13	DP 50	Long	Very high	Coarse	Very high	1999 [c]
14	Furkan	Long	Very high	Average	High	2011 [g]
15	Nazilli 84-S	Long	Very high	Average	High	1998 [b]
16	Julia	Long	Very high	Average	High	2008 [d]
17	Beyaz Altın 151	Long	Very high	Coarse	Very high	2005 [a]
18	Gossypolsüz Nazilli	Long	Very high	Average	Very high	2002 [b]
19	Özbek 142	Medium long	Very high	Coarse	Average	2002 [b]
20	ADN P01	Long	Very high	Coarse	High	2008 [h]
21	Aydın-110	Long	Very high	Average	Very high	2001 [b]
22	DP 388	Long	Very high	Coarse	High	2001 [c]
23	ST 474	Long	Very high	Coarse	Very high	2008 [c]
24	Diva	Long	Very high	Average	Very high	2002 [i]
25	SG 404	Long	Very high	Coarse	High	1999 [c]
26	DP332	Long	Very high	Coarse	High	2011 [c]
27	Deltapine 15121	Long	Very high	Coarse	High	1964 [c]
28	Coker 100-A	Long	Very high	Average	High	1964 [c]
29	Beren	Long	Very high	Average	Very high	2010 [h]
30	Gürelbey	Long	Very high	Average	Very high	2002 [b]
31	Maydos Yerlisi	Medium	Average	Coarse	Average	1964 [b]
32	Delcerro	Long	Very high	Average	Very high	1977 [b]
33	DP 396	Long	Very high	Coarse	High	2009 [c]
34	Primera	Long	Very high	Average	Very high	2011 [j]
35	Aksel	Long	Very high	Coarse	High	2008 [a]
36	Fantom	Long	Very high	Average	High	2008 [j]
37	Nazilli 66-100	Medium long	Very high	Coarse	High	1975 [b]
38	GAPEYAM-1	Medium long	Very high	Average	High	2006 [k]
39	PG 2018	Medium long	Very high	Average	Very high	2011 [a]
40	SG 1001	Long	Very high	Coarse	High	1999 [c]
41	DP493	Long	Very high	Coarse	High	2004 [c]
42	Nazilli 87	Medium long	Very high	Coarse	High	1987 [b]
43	Adana 98	Long	Very high	Coarse	High	1998 [h]
44	ST 373	Long	Very high	Average	Very high	2006 [f]
45	DP 5614	Long	Very high	Average	High	1999 [c]
46	Assos	Long	Very high	Average	High	2008 [l]
47	Nazilli M-503	Long	Very high	Average	High	1992 [b]
48	Ayhan 107	Medium long	Very high	Coarse	High	2007 [b]
49	Coşkun-1	Long	Very high	Average	High	2006 [b]
50	DP 499	Medium long	Very high	Coarse	High	2008 [c]
51	SG-501	Long	Very high	Coarse	Very high	1999 [c]

Table 1. Continued.

No.	Variety name	Fiber quality traits				Registration year
		FL mm	FS g/tex	FF mic	FU %	
52	Sure-Grow 96	Long	Very high	Average	High	2003 [c]
53	Sayar 314	Long	Very high	Average	High	1980 [h]
54	Erşan 92	Long	Very high	Average	High	1992 [m]
55	Nazilli 342	Long	Very high	Average	High	2003 [b]
56	Nazilli 954	Medium long	Very high	Coarse	High	2003 [b]
57	Sure-Grow 747	Long	Very high	Coarse	Very high	2002 [c]
58	Dicle 2002	Long	Very high	Coarse	High	2002 [n]
59	Flora	Long	Very high	Coarse	High	2007 [d]
60	Nazilli M-342	Long	Very high	Average	High	1998 [b]
61	Carolina Queen	Medium long	Very high	Coarse	High	1968 [o]
62	Maraş 92	Long	Very high	Coarse	High	1992 [m]
63	DP 565	Long	Very high	Average	Very high	2002 [c]
64	NP EGE 2009	Long	Very high	Coarse	High	2009 [b]
65	GSN-12	Long	Very high	Average	High	2007 [b]
66	BA Gold	Long	Very high	Average	High	2006 [i]
67	Gossipolsuz-86	Long	Very high	Average	Very high	1986 [p]
68	Celia	Long	Very high	Average	Very high	2007 [d]
69	NP ÖZBEK 100	Long	Very high	Average	Very high	2009 [b]
70	DP-5111	Long	Very high	Coarse	Very high	2001 [c]
71	Napa 122	Long	Very high	Coarse	High	2007 [b]
72	Çukurova 1518	Long	Very high	Coarse	High	1982 [o]
73	Deltaopal	Long	Very high	Average	High	1999 [c]
74	Barut 2005	Long	Very high	Coarse	High	2005 [b]
75	Cosmos	Long	Very high	Coarse	High	2011 [l]
76	Carisma	Medium long	Very high	Coarse	High	2013 [a]
77	SG 125	Long	Very high	Average	Very high	1999 [c]
78	BA-320	Medium long	Very high	Average	High	2005 [a]
79	Şahin-2000	Long	Very high	Coarse	High	2001 [b]
80	Lider	Long	Very high	Average	High	2004 [i]
81	Ekşi-911	Long	Very high	Coarse	High	2002 [b]
82	Nazilli M39	Medium long	Very high	Coarse	High	1992 [b]
83	DP-20	Long	Very high	Average	High	1999 [c]
84	Flash	Long	Very high	Average	Very high	2008 [a]
85	Nazilli 303	Medium long	Very high	Coarse	High	2003 [b]
86	Stoneville-453	Long	Very high	Average	High	1995 [d]
87	Famosa	Long	Very high	Coarse	High	2011 [j]
88	Lydia	Long	Very high	Coarse	Very high	2012 [a]
89	ST 468	Medium long	Very high	Coarse	High	2006 [f]
90	Lachata	Medium long	Very high	Coarse	High	1999 [f]
91	Campo	Long	Very high	Average	Very high	2004 [i]
92	Nazilli 143	Long	Very high	Average	High	1998 [b]
93	Elsa	Long	Very high	Average	Very high	2011 [d]
94	Carmen	Long	Very high	Average	High	2001 [d]
95	PG-910	Medium long	High	Average	High	2014 [a]
96	Giza 70	Extra long	Very high	Fine	High	Control

Scale: Fiber length: medium (22–25 mm), medium-long (26–28 mm), long (29–34 mm), extra-long (>34 mm); fiber strength: average (25–27), high (28–30), very high (>30); fiber fineness: fine (3.1–3.9), average (4.0–4.9), coarse (5.0–5.9); fiber uniformity: average (80–82), high (83–85), very high (>86).

Maintainers: [a]Progen Tohum A.Ş., [b]Nazilli Pamuk Araştırma İstasyonu, [c]Monsanto Gıda ve Tarım Tic. Ltd. Şti., [d]Bayer Türk Kimya San. Ltd. Şti., [e]Demet Tarım Tic. Ltd. Şti., [f]May-Agro Tohum San. Tic. A.Ş., [g]Doğu Akdeniz Geçit Kuşağı Tarımsal Araştırma İstasyonu, [h]Doğu Akdeniz Tarımsal Araştırma Enstitüsü, [i]Özbuğday Tarım İşletmeleri ve Tohumculuk A.Ş., [j]Golden West Tohumculuk Tic. Ltd. Şti., [k]GAP Tarımsal Araştırma Enstitüsü, [l]Birlik Tohum San. Tic. Ltd. Şti., [m]Kahramanmaraş Tarımsal Araştırma Enstitüsü, [n]GAP Uluslararası Tarımsal Araştırma ve Eğitim Merkezi, [o]Çukurova Tarımsal Araştırma Enstitüsü, [p]Ege Üniversitesi Ziraat Fakültesi.

2.3.2. PCR analysis

The 26 SSR primer pairs (DPL primer sets) and 14 molecular markers (10 genomic SSRs and 4 EST-SSRs), linked to QTLs for important fiber quality traits as length, strength, fineness, and uniformity, were used for PCR analysis (Table 2). The PCR was carried out with 2 µL of pure DNA, 0.5 µL of 10 µM dNTP mix, 1 µL of 25 mM $MgCl_2$, 2.5 µL of 5X PCR buffer, and 0.5 µL of 10 µM of each primer with 0.25 µL of 5 U/µL Taq DNA polymerase (Promega, Madison, WI, USA). Reactions incubated at 94 °C for 2 min and following 35 amplification cycles (30 s at 95 °C, 30 s at 50–60 °C, and 30 s at 72 °C) were performed. The final PCR products were visualized under UV light after electrophoresis on ethidium bromide-stained 3% metaphor-agarose gels.

2.4. Data analysis

Morphological data analysis was performed based on the 4 fiber quality traits' HVI measurements, i.e. length (mm), uniformity (%), fineness (mic), and strength (g tex^{-1}), and the mean values of each trait for each genotype were calculated using the 2-year data. To standardize the assessments of numerical mean values of genotypes, they were scaled based on Bradow and Davidonis' (2000) index and USTER Technologies indices (Switzerland, 2010) (Table 3). The mean value of each trait for each genotype was then subjected to both principal component analysis (PCA) using the principal components procedure and hierarchical 2-way clustering analysis using the unweighted pair group method with arithmetic averages (UPGMA). These methods were recommended for the classification of the fiber quality properties to define the patterns of variation when different sizes of groups and numbers of characters were used (Franco et al., 1997); the analyses were performed using the JMP statistical program (version 10; SAS Institute, Cary, NC, USA). The Gower general similarity coefficient (Gower, 1971) was used in the cluster analysis of agronomical traits.

For the molecular data analysis, polymorphism information content (PIC) values of molecular markers were calculated according to the following formula: $PIC = 1 - \Sigma P^2i$, where P*i* is the frequency of the *i*th allele (Anderson et al., 1993). For genetic analysis based on molecular data, each amplified band was scored based on the presence (1) and absence (0) of bands. The binary qualitative data matrix was used to construct similarity matrices based on Jaccard similarity coefficients (Jaccard, 1908) and to construct dendrograms using UPGMA on JMP software (version 10; SAS Institute).

Combined data analysis was performed using both morphological and molecular marker data in a join data set. PCA and UPGMA analysis were performed using this combined data set containing 107 characters, including 4 morphological and 103 molecular SSR characters. The Gower general similarity coefficient (Gower, 1971) was used in the combined data analysis. The cubic clustering criterion (CCC), which is implemented in the JMP statistical program, was used to estimate the number of clusters for morphological, molecular, and combined data analysis.

3. Results

3.1. Fiber quality characterization

The fiber length analysis, based on Bradow and Davidonis' (2000) index, showed that Maydos Yerlisi (*G. herbaceum* L.) has medium (23.7 mm) fiber length, and among the studied *G. hirsutum* L. varieties Nazilli 87, Carisma, PG 2018, GAPEYAM-1, BA 320, Nazilli 66-100, DP 499, ST 468, Lachata, Nazilli M39, Nazilli 954, Nazilli 303, Carolina Queen, Ayhan 107, PG 910, and Özbek 142 have medium-long fiber length. Moreover, Giza 70 (*G. barbadence* L.) has extra-long fiber length (35.8 mm), while the remaining *G. hirsutum* L. genotypes have long fiber length. According to the USTER Technologies (Switzerland, 2010) indices all the genotypes have very high fiber strength except PG 910 (high fiber strength with 28.9 g tex^{-1}) and Maydos Yerlisi (average fiber strength with 27.7 g tex^{-1}). The analysis of fiber fineness showed that Giza 70 has fine fiber fineness (3.7 mic), while the remaining genotypes have coarse (45 genotypes) or average (50 genotypes) fiber fineness. Fiber uniformity measurements showed that most of the genotypes (66) have high uniformity, whereas some (28) have very high fiber uniformity. Only 2 genotypes, Maydos Yerlisi and Özbek 142, have average fiber uniformity (80.7% and 81.8%, respectively) (Table 1).

The UPGMA analysis based on morphological markers clustered the genotypes into 4 main clades. The first clade included Delcerro and Giza 70; the second clade Özbek 142 and Maydos Yerlisi; the third clade Claudio, Menderes 2005, Diva, Aydın 110, Lydia, and ST 373; and the fourth clade all the remaining genotypes analyzed in this study. Based on the dendrogram, the most similar genotype pairs were Ege 69 and Sayar 314, Barut 2005 and Nazilli M39, and Adana 98 and Cosmos. Maydos Yerlisi clustered with Özbek 142, based on their medium fiber quality, whereas Delcerro grouped with Giza 70, based on their high fiber quality (Figure 1).

3.2. Molecular characterization

Fourteen molecular markers linked to QTLs for fiber quality traits were analyzed in commercial cotton varieties of Turkey, including Egyptian cotton Giza 70 (*G. barbadense* L.) as an out-group control. Of these, 4 markers, CIR246, CIR381, JESPR65, and BNL4108, were found to be informative. EST-SSRs markers gave low polymorphism; marker NAU1369 showed the lowest PIC value (0.00). For fiber length, strength, and fineness traits, marker CIR246, which is located in the D2-1 linkage group of the cotton

Table 2. Information on molecular markers associated with QTL traits analyzed in 96 *Gossypium* varieties.

Trait name	Marker	Type*	Primer sequence (5'-3')*	Linkage group*	No. of alleles	PIC values**
FL (mm)	JESPR208	Genomic	CGCAACCAAACATATACTTCACAC	chrD9	1	0.022
	JESPR307	Genomic	CTTGGCCATGTATTCCTTCA	chrD6b	1	0.022
	JESPR65	Genomic	CCACCCAATTTAAGAAGAAATTG	A5	3	0.239
	NAU4024	EST	ACAAGCATCTTCATGGACCT	LG02	1	0.022
FS	JESPR127	Genomic	GATTTGGGTAACATTGGCTC	chrD8	1	0.022
	NAU1369	EST	TGGCAGAGATGAATGTAAGC	D6a	1	0.00
	BNL1231	Genomic	TAATAAAAGGGAAAGGAAAGAGTT	A11-1	1	0.022
	BNL3140	Genomic	CACCATTGTGGCAACTGAGT	D9-1	1	0.022
FF (mic)	BNL3259	Genomic	TTTTGAAATTCCAGCGAAGG	D2	1	0.022
	NAU2238	EST	TTTTTCATGGCTGAACTTTG	D6	1	0.022
FL, FS, FF	CIR246	Genomic	TTAGGGTTTAGTTGAATGG	D2-1	5	0.749
	CIR381	Genomic	TTTCCATCCTTTTGTGA	D2-1	2	0.042
FL, FF	NAU3260	EST	TTTTGCAGATGTTTGTAGGG	A10-1	1	0.022
FU	BNL4108	Genomic	TCCACCATTCCCGTAAATGT	chr6	4	0.583
DPL - SSR Markers	DPL35	Genomic	CATGGTTGTACCGGTTAGTATGTG	c06	2	0.098
	DPL68	Genomic	GTTCAACAGGTCTGTACCAGTTCC	c24	2	0.142
	DPL71	Genomic	GCAAACACCATCCTACCACAA	c19	5	0.329
	DPL75	Genomic	GAGGTCATTTCAGTCCAACTCTTT	c25	3	0.256
	DPL80	Genomic	GAACCAGAGGGATGATAATGACAC	c06	4	0.29
	DPL136	Genomic	TGCTCGTATCATAAGAACCCTAGC	c07	2	0.042
	DPL146	Genomic	ATATGTTGGAAGTTGGAACTGCTG	c24	2	0.189
	DPL212	Genomic	TGATAATGCTGATGTCATAGACGC	c19	2	0.042
	DPL220	Genomic	GTTGGCCTAAGCCTATAATGATGA	c08	3	0.256
	DPL253	Genomic	TCACTATCTCAAGACCACCTTCAA	c11	3	0.463
	DPL273	Genomic	ACCATTTCTTCCATAGACTTGCTG	c04	3	0.159
	DPL307	Genomic	CCTCTCTTAATTAATGCTCCTCCA	c23	3	0.219
	DPL322	Genomic	AAACCTCGTAGTCATAGGCTCAAA	c15	4	0.353
	DPL348	Genomic	AGAATGGTTGAAGTGATGGGTTAG	c18	2	0.022
	DPL395	Genomic	GTAACATCTTTCAATCTTGCTCCC	c09	2	0.042
	DPL405	Genomic	GAGATCCATGCTAACGTCTTACAAA	c17	4	0.656
	DPL431	Genomic	CTATCACCCTTCTCTAGTTGCGTT	c10	4	0.663
	DPL443	Genomic	ACGATGACGTCAAGGATGGTAT	c12	3	0.356
	DPL486	Genomic	CTTGATGCCTCTACTTATGCAACA	c20	3	0.403
	DPL490	Genomic	AGTATCGTCACTTGTCAAAGTCCA	c01	4	0.342
	DPL513	Genomic	AGACCCGGCTACTACATGTTATCTT	c01	4	0.724
	DPL752	Genomic	CACATCACCTAATTACCATTGAAGC	c01	4	0.497
	DPL866	Genomic	AGAGTCAACTTCGACGCCAA	c12	3	0.211
	DPL890	Genomic	ACAGCATTAGCAGGCACCTT	c26	3	0.209
	DPL901	Genomic	GATGTGGTTAGGTGAGAAAGCA	c03	2	0.342
Average					2.564	0.2337

FL: Fiber length, FS: Fiber Strength, FF: Fiber Fineness, FU: Fiber Uniformity

*available at: Cotton Marker Database (CMD)

***PIC: Polymorphism Information Content*

Table 3. The scale of fiber quality traits used in this study.

Fiber length (mm)*		Fiber strength (g/tex)**	
Medium	22–25	Average	25–27
Medium-long	26–28	High	28–30
Long	29–34	Very high	>30
Extra-long	>34		
Fiber fineness (mic)**		Fiber Uniformity (%)**	
Fine	3.1–3.9	Average	80–82
Average	4.0–4.9	High	83–85
Coarse	5.0–5.9	Very high	>86

* Bradow and Davidonis (2000)
** USTER Technologies (Switzerland, 2010)

genome, was very informative (Supplementary data, on journal's website), showing a PIC value of 0.749. The only tested marker for fiber uniformity, BNL4108, showed a medium (0.583) PIC value (Table 2).

One out of the 26 DPL SSR analyzed markers, selected from microsatellite-enriched genomic libraries of *G. hirsutum* (data provided by Dr David Fang, Delta and Pine Land Company, Winterville, MS, USA), DPL348 was found to be the least informative marker (PIC value = 0.022). Markers DPL513 and DPL431 showed 4 polymorphic alleles and were the most informative markers, with 0.724 and 0.663 PIC values, respectively. The DPL markers located on chromosome 1 (DPL490, DPL513, DPL752) were very informative, with PIC values > 0.490 (Table 2).

Thirty out of 40 SSR loci were highly polymorphic, rendering a total of 103 alleles with an average of 2.57 alleles per locus ranging from 80 bp to 300 bp, with an average PIC value of 0.233 (Table 2). The UPGMA analysis of the molecular marker data resulted in more than 8 clades (Figure 2). Egyptian cotton Giza 70 (*G. barbadense* L.) and the commercial cotton variety Maydos Yerlisi (*G. herbaceum* L.) were distinctly separated from all genotypes based on their taxonomic adscription and chromosomal structure. Within the remaining varieties, the most similar cultivar pairs were Sandra and DP5111, GSN12 and Cosmos, and DP419 and Furkan.

3.3. Combined data analysis

The more informative cluster analysis of the combined data revealed a cluster pattern similar to both the separate morphological and molecular dendrograms. The combined UPGMA analysis resulted in more than 8 clades (Figure 3). The upland cottons were also separated from lowland cotton Maydos Yerlisi and, within the upland cottons, Egyptian cotton Giza 70 (*G. barbadense* L.) was also distinctly separated from commercial cotton cultivars of Turkey (*G. hirsutum* L. species).

The PCA derived from combined data showed similar relationships between genotypes as revealed by the UPGMA dendrogram. Egyptian cotton Giza 70 and the commercial cotton variety Maydos Yerlisi were distinctly separated from the other genotypes. When these 2 distinct genotypes (Giza 70 and Maydos Yerlisi) were excluded to increase the resolution of genotypes, Özbek 142 (19), Dicle 2002 (58), Nazilli M-342 (60), DP499 (50), Flash (84), DP5614 (45), and SG1001 (40) were distinctly separated as revealed by UPGMA of combined morphological and molecular data (Figure 4).

4. Discussion

Assessments of phenotypic and genotypic variability among plant varieties are essential for plant breeding and genetic diversity studies. Although phenotypic diversity has been characterized for decades, genotypic characterization studies depend on the recent challenges in molecular marker technology. Regarding phenotypic aspects, 96 commercial cotton varieties were analyzed for fiber length, strength, fineness, and uniformity by HVI analysis. The analysis demonstrated that the quality of Turkish cotton fibers is high and it can be concluded that some varieties can have high potential for the textile industry (Table 1). From 1964 to the present, critical improvements in fiber quality have been observed; for example, Maydos Yerlisi is one of the oldest varieties, released in 1964, with medium quality and is not cultivated nowadays. The most qualified genotypes detected in this study were Delcerro, Aydın 110, Diva, Menderes 2005, ST 373, Claudio, and Lydia (released in 1977, 2001, 2002, 2005, 2006, 2010, and 2012, respectively), after Egyptian cotton Giza 70. While

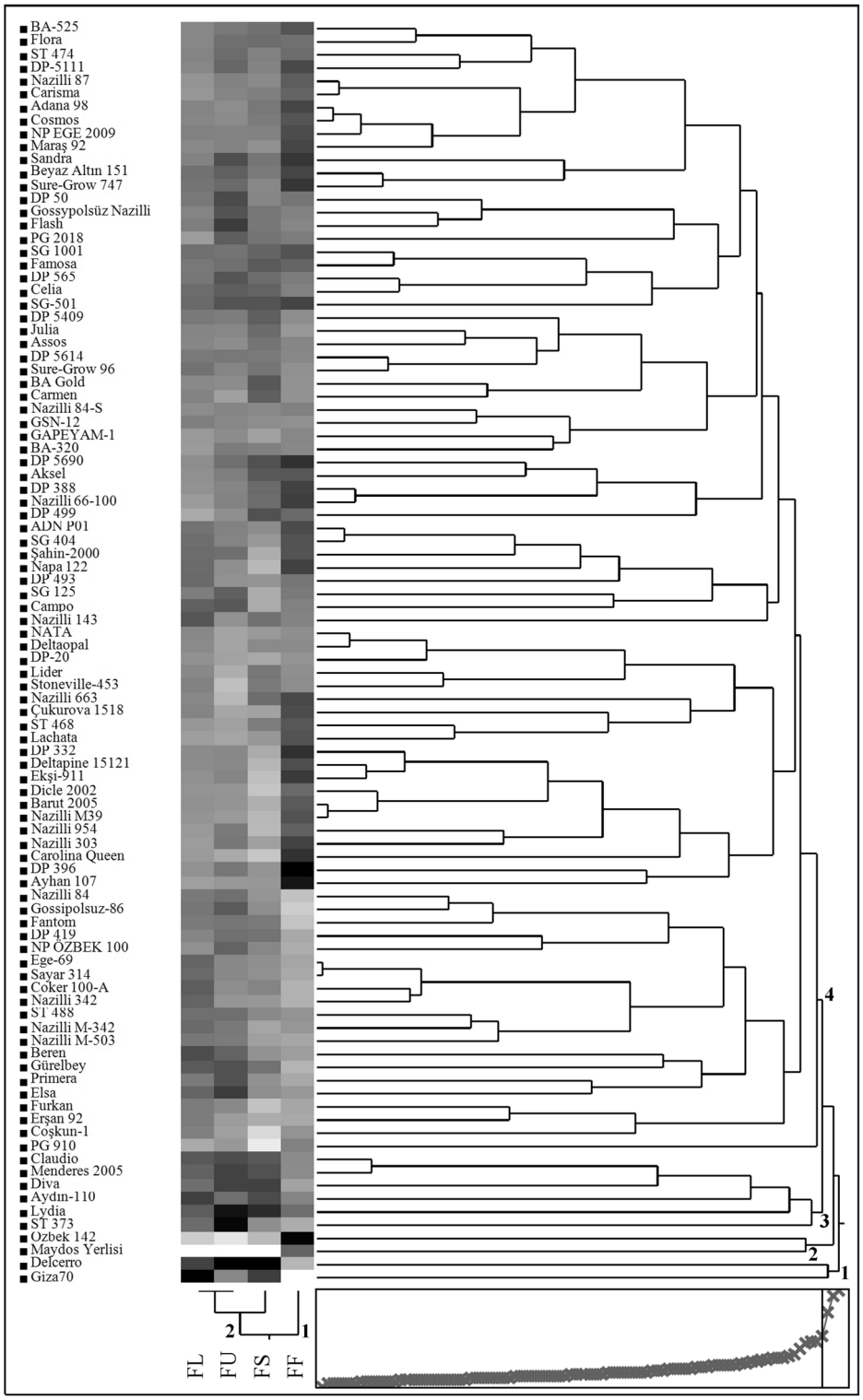

Figure 1. Two-way dendrogram based on UPGMA analysis of morphological data with distance graph, showing relationships among 96 cotton varieties for fiber length, strength, fineness, and uniformity. FL: Fiber length, FU: Fiber uniformity, FS: Fiber strength, FF: Fiber fineness. Numbers indicate the clades and grayscale cells indicate the relative similarity of each variety in each fiber trait, with the gradient from white to black increasing intervals.

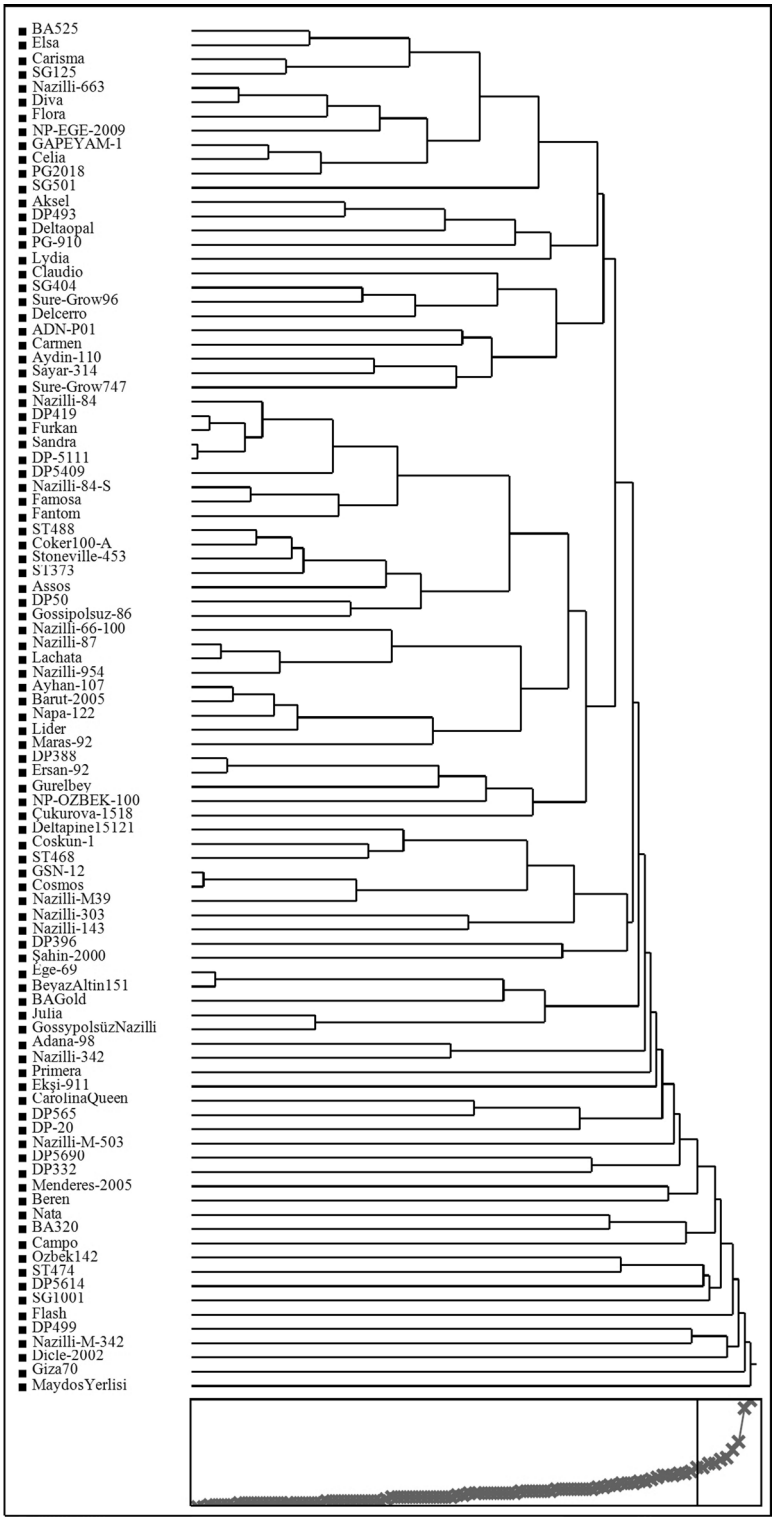

Figure 2. UPGMA dendrogram based on molecular data from 30 SSR markers, showing relationships among 96 cotton varieties.

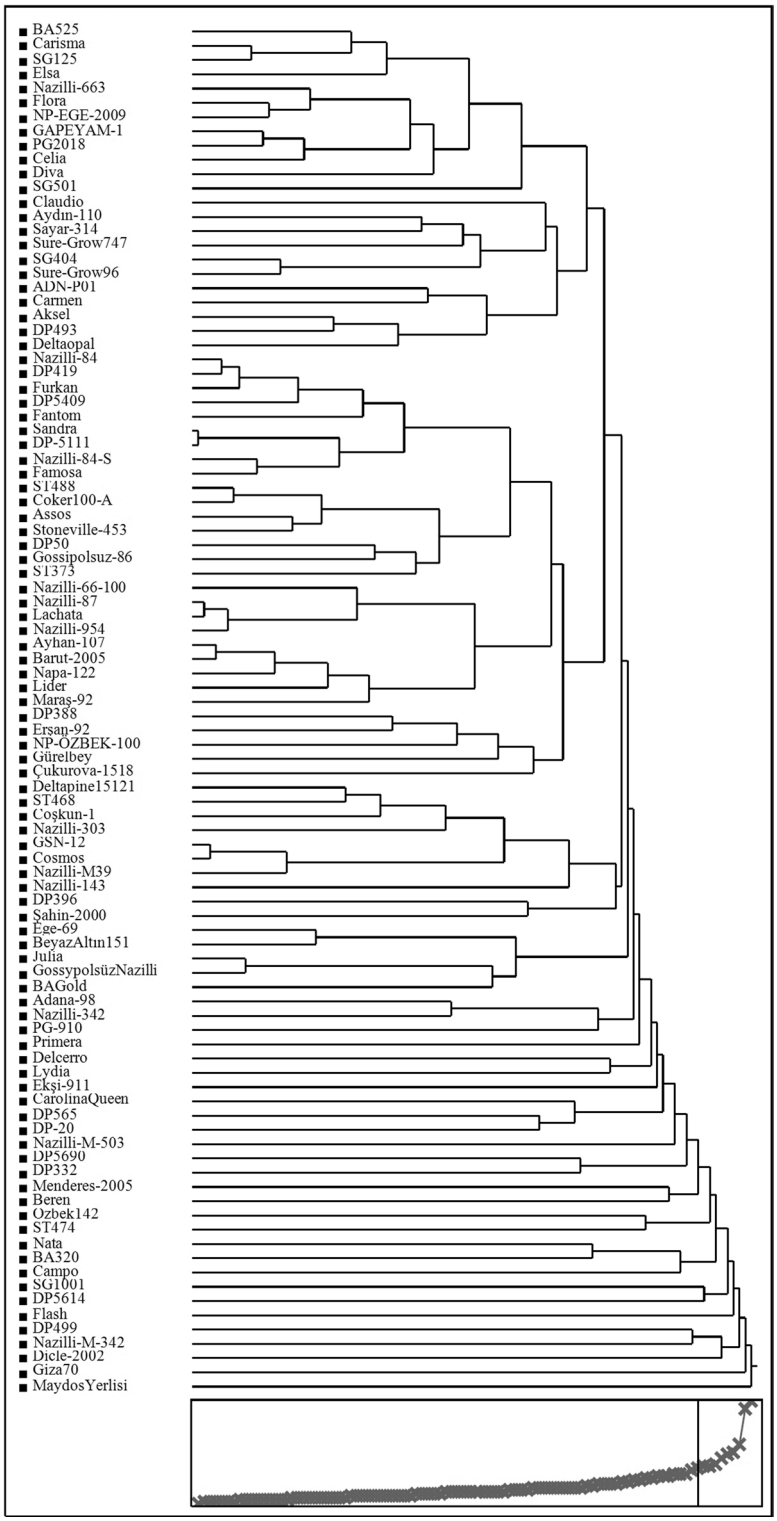

Figure 3. UPGMA dendrogram of both morphological and molecular data using 107 characters, showing relationships among 96 cotton varieties.

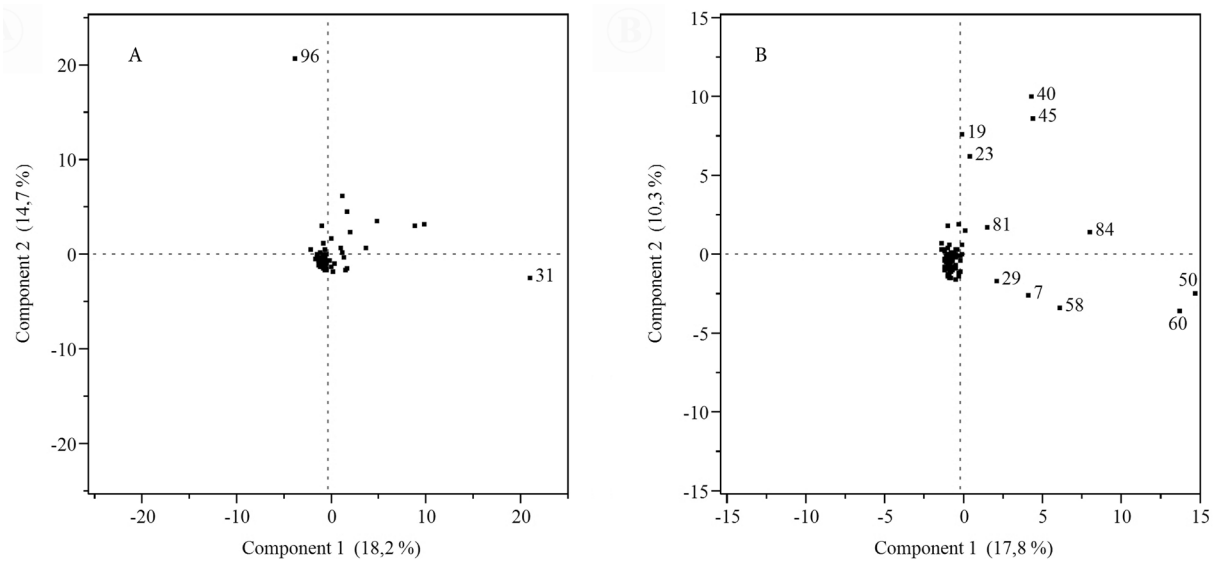

Figure 4. Principal component analysis (PCA) of 96 cotton varieties, based on both morphological (fiber length, strength, fineness, and uniformity) and molecular (SSR) data. A. Graph with all genotypes, including 96: Giza 70 and 31: Maydos Yerlisi. B. Graph, excluding the genotypes 96 and 31 to increase the resolution. Numbers indicate the genotype's codes given in Table 1.

Delcerro (parent of Aydın 110) and Aydın 110 are the oldest cultivars ST 373, Claudio and Lydia are new cultivars with high fiber quality; therefore, they are promising varieties for fiber quality improvements in breeding programs and the textile industry. Akışcan (2012) also reported favorable improvements in the fiber quality traits of 44 commercial Turkish cultivars, released from 1980 to 2008. However, their agronomically important characteristics should be analyzed, i.e. adaptation ability, yield, and disease resistance, before they can be chosen for breeding.

The UPGMA analysis of commercial cotton varieties, based on combined morphological and molecular data analysis, underlined that the cluster pattern is in agreement with morphological and molecular dendrograms. The old variety Maydos Yerlisi was distinctly separated from all the genotypes and this is not surprising since it is the only lowland genotype and is visibly distinct from other upland genotypes. Although upland cotton has a very narrow genetic structure resulting from its evolutionary history, domestication, and breeding (Paterson et al., 2004), significant genetic diversity was observed among the analyzed commercial Turkish cotton varieties, revealed by both PCA and UPGMA analysis (Figures 3 and 4). PCA indicated the relationships of genotypes in a more meaningful form showing that PCA should be used along with the dendrogram to gain a better understanding of relationships among genotypes. A recent study on the establishment of genetic diversity, population structure, and identification of core sets of allelic richness in US Upland cotton indicated average genetic distance between *G. hirsutum* accessions with low levels of genetic diversity in the Upland cotton germplasm pool (Tyagi et al., 2014). Although there are many studies on genetic diversity of cotton cultivars, only some of the recent studies have analyzed Turkish cotton varieties for genetic variability. Erkılınç and Karaca (2005) analyzed the genetic variation in 36 Turkish cotton varieties using microsatellites and identified 2 distinct genotypes, Delcerro and Aydın 110; our results confirm those findings. Bardak and Bölek (2012) used 7 commercial Turkish cotton genotypes to analyze the genetic diversity of diploid and tetraploid cottons, and reported that genetic distance among *G. hirsutum* L. genotypes was between 0.04 and 0.23. Surgun et al. (2012) also analyzed 9 Turkish cotton varieties by RAPD markers and detected the rate of polymorphism among the genotypes to be 18.1%. To the best of our knowledge, this is the only in-depth study for the genetic diversity assessment of commercial cotton varieties released in Turkey since 1964 using both fiber quality traits and molecular markers. These findings can help cotton breeders to rapidly screen genotypes for fiber quality traits in a laboratory setting, and intensify the rapid improvement for new diverse cotton cultivars.

Regarding our SSR markers and DPL series SSR markers analysis the average of alleles per locus of 30 SSRs (2.56) is similar to that observed by Bertini et al. (2006) using 31 SSR primers to characterize 53 cultivars (2.13 alleles per locus). PIC values provide an estimation of marker power by considering both the number of alleles at a locus and the relative frequencies of those alleles in the population under study (Pei et al., 2010). Our observed

average PIC value was 0.23, whereas the average PIC value of a core set of SSR markers for *Gossypium* species ranged from 0.0 to 0.82 (Yu et al., 2012). One of the findings of the present study is that the most informative markers are CIR246 with 0.749 PIC value among fiber quality linked SSR markers and DPL513 with 0.724 PIC value among tested DPL series primers. The PIC value result of marker CIR246 supports its relationship with yield and yield components in cotton (Wang et al., 2007). In addition to the polymorphic information of marker CIR246 revealed in this study, Silva et al. (2014) reported that CIR246, which confers resistance to race 18 of *Xanthomonas axonopodis* pv. *malvacearum* in cotton, was also associated with other bacterial blight resistance gene complexes. By contrast, Mishra et al. (2013) found CIR246 to be the lowest informative marker for the validation of fiber quality linked markers in the diploid species *G. arboreum*. All this research indicates the importance of marker CIR246 for breeding programs through marker assisted selection (MAS) for some traits. By contrast, our assessed EST-SSR markers showed the lowest information, probably caused by technical low resolution problems. The present study shows that SSR markers can be successfully applied to studying genetic diversity and relationships in *Gossypium* species. It is strongly recommended that both morphological and molecular assays be used in tandem for analysis of cotton genotypes. Our analysis has estimated Turkish cotton diversity; this information would be useful in the selection of parental varieties for breeding. The plant materials used in the current study are also very important for further characterization studies since they consist of a pool of commercial cotton varieties released in Turkey since 1964.

Acknowledgments

The authors would like to thank biologist Tuğçe Hançer and agricultural engineer Deniz Can for their valuable help during the laboratory and field experiments. This study was partly supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK, project no. TEYDEB/7090874) and the Technology Development Foundation of Turkey (TTGV, project no. 920/D61).

References

Ali MA, Khan IA, Awan SI, Ali S, Niaz S (2008). Genetics of fiber quality traits in cotton (*Gossypium hirsutum* L.). Australian J Crop Sci 2: 10–17.

Akışcan Y (2012). Determination of genetic improvement on fiber quality traits in some cotton cultivars released between 1980 and 2009 in Turkey. SDU Ziraat Fak Derg 7: 32–40 (article in Turkish with an abstract in English).

Anderson JA, Churchill GA, Autrique JE (1993). Optimizing parental selection for genetic linkage maps. Genome 36: 181–186.

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

Beasley JO (1942). Meiotic chromosome behaviour in species, species hybrids, haploids and induced polyploids of *Gossypium*. Genetics 27: 25–54.

Bertini CHCD, Schuster I, Sediyama T, Barros G, Moreira MA (2006). Characterization and genetic diversity analysis of cotton cultivars using microsatellites. Genet Mol Biol. 29: 321–329.

Bradow JM, Davidonis GH (2000). Quantitation of fiber quality and the cotton production - processing interface. J Cotton Sci 4: 34–64.

Chee PW, Campbell T (2009). Bridging classical and molecular genetics of cotton fiber quality and development. In: Paterson AH editor. Genetics and Genomics of Cotton, Plant Genetics and Genomics: Crops and Models. Vol. 3. New York, NY, USA: Springer Science-Business Media, pp. 283–311.

Chen ZJ, Scheffler BE, Dennis E, Triplett BA, Zhang TZ, Guo WZ, Chen XY, Stelly DM, Rabinowicz PD, Town CD et al. (2007). Toward sequencing cotton (*Gossypium*) genomes. Plant Physiol 145: 1303–1310.

Cotton Marker Database: CMD. Website: http://www.cottonmarker.org [accessed 22 May 2014].

Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19: 11–15.

Endrizzi JE, Turcotte EL, Kohel RJ (1984). Qualitative genetics, cytology, and cytogenetics. In: Koheland RJ, Lewis CF, editors. Cotton, American Society of Agronomy, Madison, WI, USA, pp. 81–129.

Erkılınç A, Karaca M (2005). Assessment of genetic variation in some cotton varieties (*Gossypium hirsutum* L.) grown in Turkey using microsatellite. Akdeniz Üniv Ziraat Fak Derg 18: 201–206.

Franco J, Crossa J, Villasenor J, Taba S, Eberhart SA (1997). Classifying Mexican maize accessions using hierarchical and density search methods. Crop Sci 37: 972–980.

Fryxell PA, Craven LA, Stewart JMD (1992). A revision of *Gossypium* sect. *grandicalyx* (*Malvaceae*), including the description of six new species. Syst Bot 17: 91–114.

Gower JC (1971). A general coefficient of similarity and some of its properties. Biometrics 27: 857–874.

Jaccard P (1908). Nouvelles recherches sur la distribution lorale. Bulletin de la Société vaudoise des sciences naturelles 44: 223–270 (in French).

Kumar P, Gupta VK, Misra AK, Modi DR, Pandey BK (2009). Potential of molecular markers in plant biotechnology. Plant Omics J 2: 141–162.

Lacape JM, Llewellyn D, Jacobs J, Arioli T, Becker D, Calhoun S, Al-Ghazi Y, Liu S, Palai O, Georges S et al. (2010). Meta-analysis of cotton fiber quality QTLs across diverse environments in a *Gossypium hirsutum* x *G. barbadense* RIL population. BMC Plant Biol 10: 132–156.

Lusas EW, Jividen GM (1987). Glandless cottonseed: a review of the west 25 years of processing and utilization research. J Am Oil Chem Soc 64: 839–854.

Mishra KK, Fougat RS, Ballani A (2013). Validation of fiber quality linked SSR markers derived from allotetraploid (*Gossypium hirsutum*) in diploid (*Gossypium arboreum*). Int J Sci R Know 1: 349–357.

Mondini L, Noorani A, Pagnotta MA (2009). Assessing plant genetic diversity by molecular tools. Diversity 1: 19–35.

Paterson AH, Boman RK, Brown SM, Chee PW, Gannaway JR, Gingle AR, May OL, Smith CW (2004). Reducing the genetic vulnerability of cotton. Crop Sci 44: 1900–1901.

Pei Z, Gao J, Chen Q, Wei J, Li Z, Luo F, Shi L, Ding B, Sun S (2010). Genetic diversity of elite sweet sorghum genotypes assessed by SSR markers. Biol Plantarum 54: 653–658.

Percival AE, Wendel JF, Stewart JM (1999). Taxonomy and germplasm resources. In: Smith CW, Cothren JT, editors. Cotton: Origin, History, Technology, and Production. New York, NY, USA: John Wiley and Sons Inc., pp. 33–63.

SAS Institute Inc. (2012). JMP 10 Discovering JMP. Cary, NC, USA: SAS Institute Inc.

Shen X, Cao Z, Singh R, Lubbers EL, Xu P, Smith CW, Paterson AH, Chee PW (2011). Efficacy of qFL-chr1, a quantitative trait locus for fiber length on cotton (*Gossypium spp.*). Crop Sci 51: 2005–2010.

Silva RA, Barroso PAV, Hoffmann LV, Giband M, Coutinho WM (2014). A SSR marker linked to the B12 gene that confers resistance to race 18 of *Xanthomonas axonopodis* pv. *malvacearum* in cotton is also associated with other bacterial blight resistance gene complexes. Australas Plant Path 43: 89–91.

Surgun Y, Çöl B, Bürün B (2012). Genetic diversity and identification of some Turkish cotton genotypes (*Gossypium hirsutum* L.) by RAPD-PCR analysis. Turk J Biol 36: 143–150.

Tohumluk Tescil ve Sertifikasyon Merkez Müdürlüğü: TTSM. Website: http:// www.ttsm.gov.tr/TR/belge/1-248/tescilli-cesitler-listesi-registered-varieties.html [accessed 22 May 2014].

Tyagi P, Gore MA, Bowman DT, Campbell BT, Udall JA, Kuraparthy V (2014). Genetic diversity and population structure in the US Upland cotton (*Gossypium hirsutum* L.). Theor Appl Genet 127: 283-295.

Uster Technologies (2010). Website: http://www.uster.com/fileadmin/customer/Knowledge/Textile_Know_How/Yarn_testing/U_LabSystems_Description_of_al_quality.pdf [accessed 22 May 2014].

Wang B, Guo W, Zhu X, Wu Y, Huang N, Zhang T (2007). QTL mapping of yield and yield components for elite hybrid derived-RILs in Upland cotton. J Genet Genomics 34: 35–45.

Wendel JF, Brubaker CL, Percival AE (1992). Genetic diversity in *Gossypium hirsutum* and the origin of Upland cotton. Am J Bot 79: 1291–1310.

Wendel JF, Crohn RC (2003). Polyploidy and the evolutionary history of cotton. Adv Agron 78: 139–186.

Yu JZ, Fang DD, Kohel RJ, Ulloa M, Hinze LL, Percy RG, Zhang J, Chee P, Scheffler BE, Jones DC (2012). Development of a core set of SSR markers for the characterization of *Gossypium* germplasm. Euphytica 187: 203–213.

Zhang ZS, Rong J, Waghmare VN, Chee PW, May OL, Wright RJ, Gannaway JR, Paterson AH (2011). QTL alleles for improved fiber quality from a wild Hawaiian cotton, *Gossypium tomentosum*. Theor Appl Genet 123: 1075–1088.

Supplementary Data: SSR profiles of amplified products for the analysis of 96 cotton varieties using marker CIR246. **M**: 100 bp marker (Thermo Fisher Scientific, Waltham, MA, USA). The parts include 32 genotypes per gel with given order listed in Table 1. Numbers indicate the groups.