

RESEARCH

Open Access



# Exploring agro-morphological and fiber traits diversity in cotton (*G. barbadense* L.)

Ehab A. A. Salama<sup>1</sup>, Mona A. Farid<sup>2</sup>, Youssef A. El-Mahalawy<sup>3</sup>, A. A. A. El-Akheder<sup>3</sup>, Ali A. Aboshosha<sup>2</sup>, Aysam M. Fayed<sup>4</sup>, W. M. B. Yehia<sup>3</sup> and Sobhi F. Lamlom<sup>5\*</sup>

## Abstract

Cotton (*Gossypium barbadense* L.) is a leading fiber and oilseed crop globally, but genetic diversity among breeding materials is often limited. This study analyzed genetic variability in 14 cotton genotypes from Egypt and other countries, including both cultivated varieties and wild types, using agro-morphological traits and genomic SSR markers. Field experiments were conducted over two seasons to evaluate 12 key traits related to plant growth, yield components, and fiber quality. Molecular diversity analysis utilized 10 SSR primers to generate DNA profiles. The Molecular diversity analysis utilized 10 SSR primers to generate DNA profiles. Data showed wide variation for the morphological traits, with Egyptian genotypes generally exhibiting higher means for vegetative growth and yield parameters. The top-performing genotypes for yield were Giza 96, Giza 94, and Big Black Boll genotypes, while Giza 96, Giza 92, and Giza 70 ranked highest for fiber length, strength, and fineness. In contrast, molecular profiles were highly polymorphic across all genotypes, including 82.5% polymorphic bands out of 212. Polymorphism information content was high for the SSR markers, ranging from 0.76 to 0.86. Genetic similarity coefficients based on the SSR data varied extensively from 0.58 to 0.91, and cluster analysis separated genotypes into two major groups according to geographical origin. The cotton genotypes displayed high diversity in morphology and genetics, indicating sufficient variability in the germplasm. The combined use of physical traits and molecular markers gave a thorough understanding of the genetic diversity and relationships between Egyptian and global cotton varieties. The SSR markers effectively profiled the genotypes and can help select ideal parents for enhancing cotton through hybridization and marker-assisted breeding.

**Keywords** Cotton, Yield, Genetic diversity, SSR, DNA fingerprinting, Fiber traits

## Introduction

Cotton (*Gossypium* L.) is an important economic and fiber crop grown annually in over eight countries including the USA, India, China, and Egypt. It provides over 95% of the raw natural fibers used in the textile industry and has value as a bioenergy and oilseed crop [1, 2]. Cotton belongs to the *Gossypieae* tribe in the *Malvaceae* family, comprising about 53 species including 46 diploids ( $2n=2x=26$ ) and 7 allotetraploids ( $2n=4x=52$ ) [1, 3]. The modern cotton cultivars are allotetraploid with 26 chromosomes ( $n=2x=26$ ), evolved through hybridization and domestication of A1-genome diploids *Gossypium herbaceum* ( $n=x=13$ ) with indigenous D5-genome

\*Correspondence:

Sobhi F. Lamlom  
sobhifaid@alexu.edu.eg

<sup>1</sup> Agricultural Botany Department, Faculty of Agriculture (Saba Basha), Alexandria University, Alexandria 21531, Egypt

<sup>2</sup> Genetics Department, Faculty of Agriculture, Kafr El-Sheikh University, Kafr El-Sheikh, Egypt

<sup>3</sup> Cotton Breeding Department, Agriculture Research Center, Cotton Research, Cotton Research Institute, Kafr El-Sheikh, Egypt

<sup>4</sup> Molecular Biology Department, Genetic Engineering and Biotechnology Institute, University of Sadat City, Sadat 32897, Egypt

<sup>5</sup> Plant Production Department, Faculty of Agriculture (Saba Basha), Alexandria University, Alexandria 21531, Egypt



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

diploids *G. raimondii*, a new world cotton species [4, 5] [6, 7].

Importantly, over 97% of the global cotton fiber production comes from the two main cultivated allotetraploid species, *G. hirsutum* L. and *G. barbadense* L. [8]. In addition, the genus *Gossypium* contains over 50 species (45 diploid and 5 allotetraploid) including the most widely grown *G. hirsutum* and *G. barbadense*. This has led to extensive phenotypic diversity in cotton crops across many geographic regions worldwide [9]. However, the Egyptian cotton (*G. barbadense*) known by its extra-long and staple fiber Pima cotton characteristics is a famous cotton source for textile industry worldwide due to its unique chemical composition properties.

Traditional cotton breeding methods aim to improve cotton quality and yield by identifying high-performing parent lines with desirable agronomic traits for crosses. However, these classical techniques have limitations. More advanced molecular breeding utilizing genetic markers and genotyping could overcome these limitations and accelerate cotton improvement [5, 10, 11]. It is an important goal for cotton breeders to predict genetic similarities/dissimilarities and assess genetic diversity among cotton germplasm including genotypes, cultivated species, and wild relatives [12]. This allows accurate selection of potential lines to accelerate cotton improvement programs and obtain satisfactory yield and quality by maintaining sufficient genetic variability in cotton gene pools [13, 14].

Therefore, the lack of suitable genetic diversity in breeding germplasm is a major constraint slowing cotton breeders' progress in developing new cultivars [15–19]. Using molecular markers as an alternative tool to identify and select superior parents early in breeding programs and incorporate them into marker-assisted selection (MAS) could significantly enhance genetics and reduce time and costs required to develop novel cotton cultivars [12, 20–22].

Taking together, one of the main tools for cotton breeders to study the genetic diversity is using the molecular markers system that have already overcome the obstacles and disadvantages of morphological markers/characters that have a limit number and are affected by different plant growth stages as well as various environmental conditions [23–25]. In this regard, there are many types of DNA molecular markers that have been extensively used for various genetic analyses of cotton crop species such as RFLPs, RAPD, AFLPs, ISSRs, SSRs, and SNPs [26–29]. SSR markers, which are small motifs consisting of one to six tandem repeats, have been extensively and effectively utilized in genetic diversity, DNA fingerprinting, and QTL mapping studies for cotton crops. This is due to their distinctive DNA-based markers, which possess

high specificity in amplifying genomic loci, a simple and easy operation system, a high degree of polymorphism, codominant nature, good reproducibility, and a wide distribution throughout the entire genome [26, 30, 31].

Given this context, the primary objective for cotton breeders is to create a publicly accessible database of molecular markers that can be used as a genetic diversity detection system. These markers, such as SSR markers, are closely associated with important agronomic and fiber quality traits. The aim is to expedite the process of selecting and breeding these traits to ensure sustainable cotton production [32, 33]. To date, genomic libraries contain over 1000 publicly available SSR designed primers from existing cotton DNA sequences generated by research groups worldwide [34, 35]. Many studies have used SSR markers to determine genetic diversity in diverse cotton germplasm. For example, Manonmani et al. [36] characterized genetic diversity of 12 Indian cotton genotypes using 55 SSR primer pairs. They found 40 pairs (25 polymorphic and 15 monomorphic) showed clear, scorable bands and an average of 1.8 alleles per locus [37].

The objective of this study was to explore and evaluate the molecular diversity and genetic polymorphisms among fourteen cotton genotypes using agronomic/morphological characters and genomic SSR markers. The goal was to identify suitable elite divergent genotypes that could be used as parents in future cotton breeding programs.

## Materials and Methods

### Experimental plant materials

A total of fourteen cotton (*Gossypium barbadense* L.) genotypes consisted of some Egyptian cotton genotypes and other foreign cotton cultivars were used for this investigation as experimental materials. The seed materials of these studied genotypes were obtained from Cotton Breeding and Genetics Department, Cotton Research Institute, Agricultural Research Center (ARC), Egypt. Details of these genotypes are presented in Table 1.

### Experimental design and field assay

The field experiment for this study was conducted in the Agricultural Research Station in Sakha, which is part of the Egyptian governate of Kafr El-Shaikh. During the 2015 and 2016 growing seasons, this study used a randomized complete block design (RCBD) using triplicates.

### Measurements of studied traits

Twelve morphological along with fiber traits were evaluated in the field as follows: 1. Position of first fruiting node (P.F.F.N); 2. Days to first flower (D.F.F); 3. Number of vegetative branches per plant (NO.V.B./P); 4. Number

**Table 1** Origin, and pedigree for the fourteen parental cotton genotypes utilize in this study

No.	Genotype	Pedigree	Origin
1.	Giza 86	G.75×G.81	Egypt
2.	Giza 68	G.36×G.56	
3.	Giza 96	G.84×(G.70×G.51B)×Pima62	
4.	Giza 94	10,229×G.86	
5.	Giza 92	G.84×(G.74×G.68)	
6.	Giza 70	G.59A×G.51B	
7.	Giza 93	G.77×PimaS6	
8.	Giza 45	G.28×G.7	
9.	Karchenky Branches	Unknown	Russia
10.	Suvin	Sujata×Vincent	India
11.	Pima 5 <sub>6</sub>	5934–23–2×615,903–98–4–4	USA
12.	Pima high percentage	Unknown	
13.	C.B. 58	Unknown	
14.	Big Black Boll (B.B.B)	Unknown	Greek

of fruiting branches per plant (NO.F.B./P); 5. Boll weight (B.W); 6. Seed cotton yield per plant (S.C.Y./P); 7. Lint yield per plant (L.Y./P); 8. Lint percentage (L.%); 9. Fiber length (F.L); 10. Fiber fineness (F.F); 11. Fiber strength (F.S); 12. Uniformity ratio (UR). All morphological, agronomical and fiber traits/properties tests were measured according to known cotton measurement standards.

#### DNA assay for diversity assessment

All molecular work related to this study was conducted in genetics and biotechnology laboratories, Faculty of Agriculture, Kafr El-Sheikh University and GEBRI, University of Sadat City, Egypt.

#### Genomic DNA samples collection, isolation, purification, and quantification

To assess the genetic diversity of these genotypes, the fresh leaves from each genotype were collected separately

during the seedling growth stage. Then, pre-weighted leave tissue samples from 0.2 to 0.5 g were immediately frozen in liquid nitrogen and fully grounded to fine powder using a pestle and mortar. The subsequent steps of total genomic DNA extraction and purification were carried out using the CTAB method [38] with a few modifications. The isolated DNA samples were measured quantitatively using UV Mass spectrophotometer at a specific optical density ( $A_{260}$  and  $A_{280}$ ) as well as were qualitatively checked 1.5% agarose gel along with the standard DNA marker/ladder. The DNA samples were stored at  $-20^{\circ}\text{C}$  in a final concentration of 50 ng per microliter for further downstream steps.

#### PCR amplification, electrophoresis detection and polymorphism analysis protocols

The isolated genomic DNA samples from the 14 cotton genotypes were screened using 10 BNL series SSR primers/markers. These examined primers were obtained and designed based on available sequence information in the Cotton Marker Database (CMD) as summarized in Table 2. The DNA samples were amplified using polymerase chain reaction (PCR) in a 20  $\mu\text{l}$  final reaction volume according to the method described by Saif et al. [39]. The PCR amplification reactions were conducted using 20 ng of DNA in a 25- $\mu\text{L}$  reaction volume, comprising 0.3  $\mu\text{M}$  of each primer, 200  $\mu\text{M}$  of dNTPs, 5  $\mu\text{L}$  (1X) of Taq polymerase buffer, 1.5 mM  $\text{MgCl}_2$ , and 0.5 U Taq DNA polymerase. For SSR reactions, a Touchdown PCR program was employed. The primary program involved 9 cycles at  $94^{\circ}\text{C}$  for 1 min,  $54^{\circ}\text{C}$  for 1 min (with a  $1^{\circ}\text{C}$  decrease in every cycle), and  $72^{\circ}\text{C}$  for 1 min. Subsequently, 28 cycles were executed at  $94^{\circ}\text{C}$  for 1 min,  $45^{\circ}\text{C}$  for 1 min, and  $72^{\circ}\text{C}$  for 1 min. The initial cycles were preceded by a denaturation step at  $94^{\circ}\text{C}$  for 5 min and followed by an extension step at  $72^{\circ}\text{C}$  for 5 min. Then, the PCR products (amplicons) were stored at  $4^{\circ}\text{C}$  for the next step of gel electrophoresis. The amplified PCR products were

**Table 2** Detailed summary of SSR primers involved in the molecular analysis of present study

Primer code	Forward sequence (5'-3')	Reverse sequence (5'-3')
BNL2823	ATATTCATGCCTCTGCAGCC	GTTTTAGTTTTGGACTTAGAGGC
BNL2827	ATCGCGGCATTAATGAATA	AATACATCCGCTCATTTCCG
BNL1044	TGCTCTTTTTGGGGGACTA	ATTGGCTTTGGTTGGTTGAG
BNL1440B	CCGAAATATACTGTGCATCTAAACG	CCCCCGGACTAATTTTTCAA
BNL3408A	ATCCAAACCATTGCACCACT	GTGTACGTTGAGAAGTCATCTGC
BNL3408B	ATCCAAACCATTGCACCAAGC	GTGTACGTTGAGAAGTCATCTAT
BNL2634A	AACAACATTGAAAGTCGGGG	CCCAGCTGCTTATTGGTTTC
BNL193	TGTGAGCCATTGCTGTTAGC	TAAAGTCTGGCATTGTGAGC
BNL1047	GCTTGTCATCTCCATTGCTG	TAGCCCGTTTCATGTTCTTC
BNL1061	GCTTGTCATCTCCATTGCTG	TAGCCCGTTTCATGTTCTTC

separated by gel electrophoresis on 3% agarose gel [39]. The gel photos were visualized and taken under UV light using Gel Documentation System, and the bands were scored as 1 while the absence of a band was recorded as 0, and the [0,1] binary data matrix was constructed.

#### Data and genetic diversity analysis

The observed field data was analyzed to estimate the mean performance differences between studied cotton genotypes based on 12 agro-morphological and fiber collected traits using SPSS 21.0 software (<https://www.ibm.com/products/spss-statistics>). While the genotypic data based on SSR markers screening assay was analyzed as follows; For each cotton genotype, the amplified gel bands of each target SSR primer representing different alleles were scored, and the allelic bands reflecting the allelic variation were compared with 100 bp DNA ladder. Quantity one software (Gel Doc, Bio-Rad Laboratory, Inc.) was used for capturing gel images and the length of generated DNA fragments were estimated. Then, its data was converted into [0,1] binary matrix subjected to multivariate analysis. The polymorphism information content (PIC) analysis was estimated for all cotton genotypes based on their SSR markers allelic frequency according to the described method of Anderson et al. [40] which showed various PIC values [41] indicating different informative potential of each used SSR marker (High: more than 0.5; Moderate: between 0.5 and 0.25; Slightly below: 0.25.). The cluster dendrogram was constructed

using UPGMA method based on the pairwise genetic distances [42] between the cotton genotypes using Numerical Taxonomy System, NTSYS-PC and NTSYS Pc 2.1 software [43]. Finally, the similarity matrices based on Jaccard similarity coefficients [44] were estimated by NTSYS-PC and NTSYS Pc 2.1 software.

## Results

### Agronomic and morphological characters

Table 3 displays the mean values of agro-morphological traits derived from the genotypes that were examined in the field. In general, the initial fruiting node location character showed the highest mean values from the Egyptian cotton genotypes; G.96 (8.33), G.68 (8.33), and G.93 (8.17), in that order. Conversely, the C.B. 58 genotype had the lowest mean value (5.42), while the Karchenky Branches and Suvin genotypes came in at 5.92 and 6.08, respectively. But in the genotypes that were considered, this feature varied between 5.42 and 8.33. While G.45 (25.58) had the best mean value for fruiting branches per plant, G.68 (22.17) and Suvin (20.08) were next ideal. On the other hand, genotypes G.93 (17.58), G.86 (17.75), and G.94 (17.92) yielded the lowest mean values.

The highest mean values for the number of vegetative branches per plant were obtained from genotypes G.68 (5 branches), G.92 (4 branches), and G.45 (5 branches), as shown in Table 3. However, genotypes C.B. 58 (2.08), Pima high percentage (2.08), and G.94 (2.25), in that

**Table 3** The mean performances of fourteen genotypes for earliness, growth habit, yield, and fiber quality traits for two years

Genotype	P.F.F.N	No.F.B.P	NO.V.B.P	D.F.F	B.W	S.C.Y.P	L.Y.P	L.%	FL	F.S	F.F	U.R
<b>G.86</b>	7.92	17.75	3.58	74.05	3.05	108.70	42.61	0.3919	33.97	10.37	4.10	87.27
<b>G.68</b>	8.33	22.17	4.00	72.11	3.13	100.01	34.85	0.3489	35.93	10.67	3.63	86.70
<b>Pima s6</b>	8.08	17.92	3.83	70.71	3.06	96.08	34.78	0.3618	35.50	10.57	4.07	85.67
<b>Suven</b>	6.08	20.08	2.50	68.52	2.99	78.91	29.54	0.3744	34.50	10.37	4.03	84.23
<b>G.96</b>	8.33	18.08	3.17	70.16	3.31	97.71	38.28	0.3907	36.00	11.07	4.10	87.37
<b>G.94</b>	7.17	17.92	2.25	68.58	3.31	111.59	45.14	0.4047	34.53	11.20	3.90	87.57
<b>C.B. 58</b>	5.42	19.00	2.08	67.13	2.82	90.16	34.25	0.3800	34.63	10.50	4.10	84.67
<b>P.H.P</b>	6.58	18.33	2.08	66.59	2.85	80.79	32.01	0.3955	34.03	11.00	3.50	85.87
<b>G.92</b>	7.67	18.75	4.08	69.40	3.03	81.39	27.82	0.3433	33.83	11.50	4.00	86.87
<b>K.B</b>	5.92	18.67	2.33	67.79	3.05	68.75	24.67	0.3604	35.43	10.40	4.23	84.47
<b>G.70</b>	8.00	19.50	3.75	74.79	2.57	106.64	38.04	0.3567	36.80	11.40	3.90	88.10
<b>G.93</b>	8.17	17.58	3.33	70.18	2.84	102.40	35.07	0.3425	37.07	11.43	3.33	87.27
<b>G.45</b>	8.17	25.58	5.00	70.04	2.54	92.41	31.51	0.3412	36.83	11.17	3.27	86.97
<b>B.B.B</b>	6.58	19.25	2.42	67.00	3.15	101.14	40.81	0.4041	34.07	11.17	4.13	85.53
<b>L.S.D<sub>0.05</sub>*</b>	0.59	1.83	0.70	2.01	0.19	14.81	5.86	2.12	0.82	0.28	0.25	0.98
<b>L.S.D<sub>0.01</sub>**</b>	0.78	2.43	0.92	2.66	0.25	19.60	7.76	2.81	1.09	0.38	0.338	1.30

P.F.F.N/ First Fruiting Node, No. FB/P Number of fruiting branches per plant, No.V.B.P Number of vegetative branches per plant, D.F.F Days to first flower, B.W. Boll weight, LCY.p. Lint cotton yield per plant, SCY.p Seed cotton yield per plant, L% Lint percentage, FL. Fiber length, FF Fiber fineness, FS Fiber strength, UR% Uniformity ratio

\* LSD0.05 = least significant differences of means ( $p < 0.05$ ), \*\*LSD0.01 = least significant differences of means ( $p < 0.01$ )

order, produced the lowest mean values. Curiously, the Egyptian genotypes G.70 (74.79), G.86 (74.05), and G.68 (72.11), in that order, had the highest mean values for the days to first flower trait. In contrast, the genotype Pima high percentage (66.59), B.B.B. (67), and C.B. 58 (67.13), in that order, had the lowest mean values.

#### **Yield and its component characters**

The mean values of yield and its associated parameters for all cotton genotypes are detailed in Table 3. In terms of boll weight (B.W), the data indicates that the genotypes G.96, G.94, and B.B.B exhibited the highest mean values at 3.31, 3.31, and 3.15, respectively, while the lowest mean values were observed in G.45, G.70, and C.B.58 at 2.54, 2.57, and 2.82, respectively. Analysis of seed cotton yield per plant (S.C.Y./P) in Table 3 reveals that the top-performing international cotton genotypes were Karchenky Branches, Suvin, and Pima high percentage, with mean values of 68.75, 78.91, and 80.79, respectively. Considering the lint yield per plant (L.Y./P), the data in Table 3 demonstrates that the genotypes G.94, G.86, and B.B.B exhibited the highest mean values at 45.14, 42.61, and 40.81, respectively. Conversely, the cotton genotypes Karchenky Branches, G.92, and Suvin recorded the lowest mean values at 24.67, 27.82, and 29.54, respectively. In terms of lint percentage (L.%), the results indicated that the genotypes G.94, B.B.B, and Pima high percentage had the highest mean values at 40.47, 40.41, and 39.55, respectively. On the other hand, the Egyptian cotton genotypes G.45, G.93, and G.92 exhibited the lowest mean values at 34.12, 34.25, and 34.33, respectively.

#### **Fiber quality properties traits**

In the case of cotton quality characteristics and properties, it was stated that the highest mean values for fiber length (F.L) trait were acquired from the Egyptian cotton genotypes, G.93, G.45 and G.70 (37.07, 36.83 and 36.8) respectively. Conversely, the lowest mean values were found from genotypes G.92, G.86 and P.H.P (33.83, 33.97 and 34.03), respectively. On the other side, the results in Table 3 exhibited that the highest mean values for the traits of fiber strength (F.S) were (11.5, 11.43 and 11.4) on the Egyptian genotypes G.92, G.93 and G.70, respectively whereas the lowest mean values were obtained from genotypes Suvin, G.86 and Karchenky Branches (10.37, 10.37 and 10.4), respectively.

In addition, the fiber fineness (F.F) character showed higher mean values from the cotton genotypes; Karchenky Branches, B.B.B and G.86 (4.23, 4.13, 4.1), respectively. While the lowest mean values were obtained from genotypes G.45, G.93 and P.H.P (3.27, 3.33 and 3.5), respectively. Finally, almost all the tested genotypes showed high ratio of uniformity but the

highest mean values for uniformity ratio (U.R) trait were obtained from The Egyptian genotypes; G.70, G.94 and G.96 (88.1, 87.57 and 87.37) respectively. While the lowest mean values were attained from the Indian genotype (Suvin), the Russian genotype (Karchenky Branches), and one of American cottons (C.B. 58). Taking together, most of the Egyptian cotton genotypes showed high agronomic performance compared to the international genotypes.

#### **Molecular diversity revealed by SSR markers.**

Ten SSR markers, revealing a notably high level of polymorphism (polymorphic DNA), as outlined in Table 4 and Fig. 1. Across all tested cotton genotypes, the results indicated that the primer pairs designed for SSR analysis generated a total of 212 bands, with 175 of them being polymorphic. This accounted for 82.54% of the total bands, with an average of 17.5 polymorphic bands per marker. The number of bands varied between 5 and 8 for the primer pairs BNL2827 and BNL2823. The polymorphic bands percentage ranged from 72.4% for the primer pair BNL1440B to 100% for the primer pair BNL193. The Polymorphic Information Content (PIC) values for the SSR primer pairs ranged from 0.76 for BNL193 to 0.86 for BNL2827, with the latter recording the highest PIC value among the ten SSR markers. In summary, all examined SSR markers were deemed informative, collectively revealing an average PIC of 0.815.

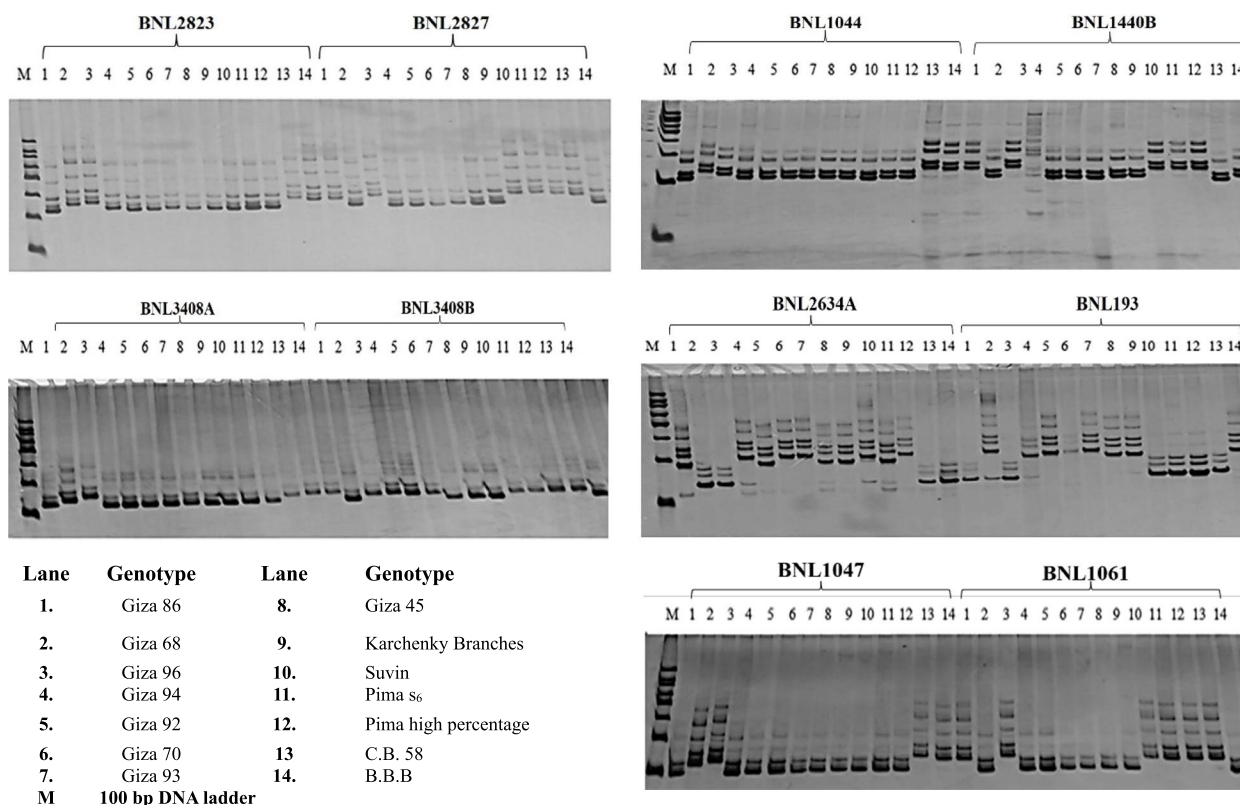
#### **Similarity coefficient assessment**

The genetic similarity co-efficient matrix of cotton genotypes used in this study (Table 5) showed that the similarity index (SI) values ranged from 0.5824 to 0.9066 with

**Table 4** The diversity analysis results generated by ten simple sequence repeats (SSR) markers used in the study

No.	Marker name	Chr. No.	No. of allele	Polymorphic bands	PIC
1	BNL2823	6	8	16	0.84
2	BNL2827	1	8	12	0.86
3	BNL1044	4	5	18	0.81
4	BNL1440B	25	4	14	0.79
5	BNL3408A	3	3	12	0.80
6	BNL3408B	3	2	18	0.78
7	BNL2634A	7	6	19	0.83
8	BNL193	18	3	24	0.76
9	BNL1047	25	6	22	0.85
10	BNL1061	25	5	20	0.83
Total			<b>50</b>	<b>175</b>	<b>8.15</b>
Average			<b>5</b>	<b>17.5</b>	<b>0.82</b>





**Fig. 1** DNA fingerprints showed the polymorphism of fourteen cotton genotypes with ten SSR primers

an average of 0.7473 as well as a high dissimilarity coefficient of 0.5824 and 0.6044 for the Egyptian genotype Giza 68 with genotypes Pima s<sub>6</sub> and Pima high percentage, respectively. On the other hand, the highest similarity coefficient was recorded for the genotypes Giza 70 and Giza 93 (0.9066) and the genotypes Pima s<sub>6</sub> and Pima high percentage respectively (0.9011). In addition to that, the Indian genotype (Suvin) recorded the higher genetic similarity index with the American genotypes; Pima high percentage (0.8901 followed by Pima s<sub>6</sub> (0.8681) cotton respectively.

**Cluster analysis**

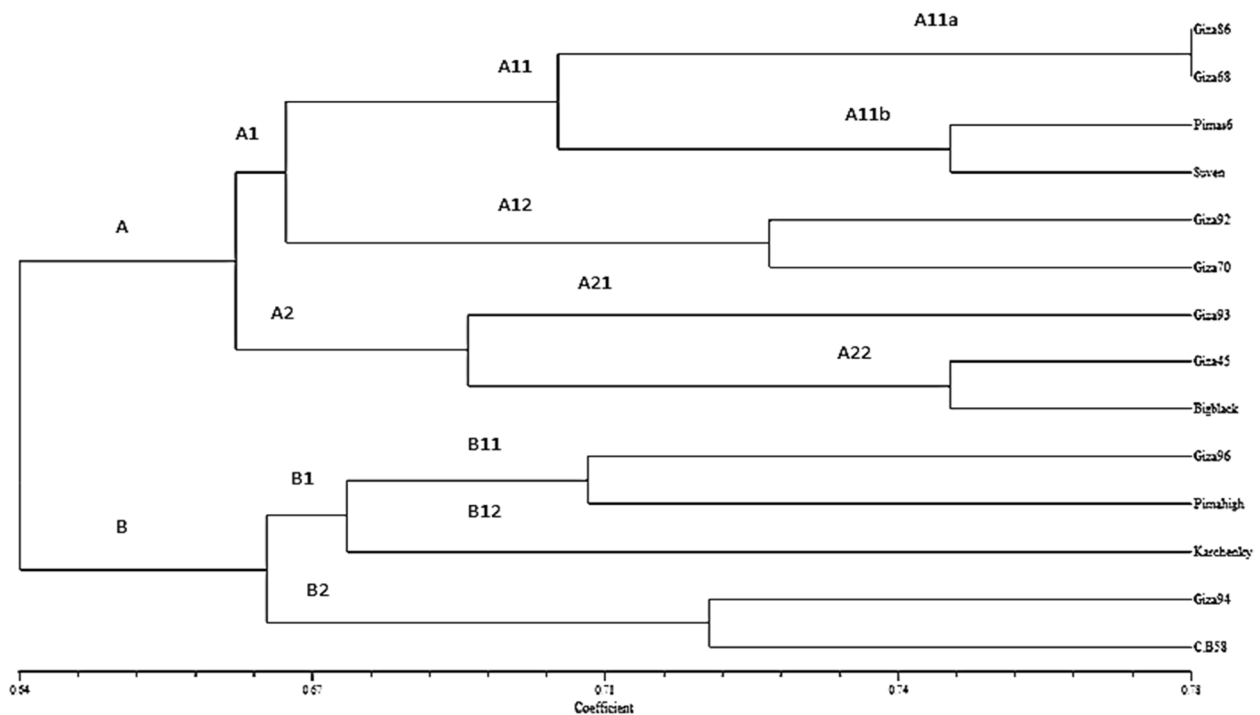
For this study, the fourteen cotton genotypes were scored based on the presence and absence of amplified band for each SSR marker and its specific alleles. Thus, genetic distance analysis showed that for each genotype combination, the genetic distance ranged from 0.64 to 0.78 and according to the cluster analysis of combined SSR data, all 14 genotypes used in this study were separated into two major clusters (Fig. 2). The constructed dendrogram has grouped the used genotypes into two distinguished clusters namely, A and B. According to the phylogenetic tree, it shown in Fig. (2)

that the genetic similarity of 0.66 was the start separation point for main cluster to two sub clusters A1 and A2, the first sub cluster consisted of A11 and A12 at genetic similarity of 0.70, the A12 sub cluster included Giza 92 and Giza 70 at genetic similarity of 0.73. The A12 sub cluster separated to A11a and A11b, the A11a included G.68 and G.86 at genetic similarity 0.78 while the A11b included Pima s<sub>6</sub> and Suvin at genetic similarity 0.75. The A2 sub cluster separated to A21 and A22, the A21 included individual cultivars G.93, while A22 included G.45 and B.B.B genotypes at genetic similarity 0.75. On the other side, the second main cluster was separated into two sub main clusters namely, B1 and B2 at the genetic similarity of 0.66. Then, the first sub cluster B1 was subsequently divided into another sub clusters of B11 and B12 at genetic similarity 0.71 while, The B2 included two genotype G.94 and C. B. 58 at genetic similarity 0.72. The B11 included G.96 and Pima high percentage at genetic similarity 0.71, while B12 included only one cultivar Karchenk Branches.

**Discussions**

One of the essential goals for cotton breeders, is to develop modern varieties with promising characteristics in terms of fiber quality as well as agronomical





**Fig. 2** Cluster analysis dendrogram constructed from the studied cotton genotypes through ten SSR primers

economic traits to increase the farmers profitability under its current cultivation system. Unfortunately, the achievement of this goal is hindered by the poor and tapered of genetic base of modern crop varieties due to the continued extensive selection process during its progress course that eventually has led to a lack of genetic variability amongst the core cotton genotypes [17, 37, 45]. Therefore, the current investigation was aimed to estimate the genetic diversity/variability among different Egyptian and international cotton genotypes using important agro-morphological traits and DNA based SSR markers.

The results of our investigation suggest that the Egyptian cotton genotypes demonstrated the most elevated mean performance values across all assessment criteria for growth performance. On the other hand, the genotypes G.96, G.94, B.B.B, and G.70 demonstrated the highest average values for both yield and its constituent components. The cotton genotypes G.96, G.92, Karchenky, G.94, G.93, B.B.B, and G.70 had the greatest average values for fiber qualities when compared to the other cotton genotypes. Likewise, a multitude of previous studies have demonstrated comparable patterns in the assessed agro-morphological parameters [46, 47].

The results obtained from the analysis of fourteen cotton genotypes using ten SSR/microsatellite molecular

markers are of significance. It is crucial to note that the effectiveness of various DNA-based markers in assessing genetic variation in crops can vary based on genetic principles and the rationale behind using each molecular marker [48]. In our genetic diversity analysis, as depicted in Table 4, the total and average number of polymorphic bands for the studied SSR markers were found to be higher compared to the findings of Kurt et al. [30]. In another study, they analyzed twenty-nine genotypes, including interspecific hybrid cotton, using twelve genomic SSR markers. They observed a different number of amplified alleles ranging from 2 to 4 for each locus, with an average of 2.53 alleles per locus. [12, 41, 49–51].

Moreover, the investigation carried out by Dongre et al. [52] found that out of twenty-five SSR markers tested in their study, 17 markers were able to produce 56 polymorphic bands in addition to four SSR markers that showed a monomorphic pattern while the remaining markers were non-scorable and non-reproducible bands. Taking together, the similar findings by using genomic SSR markers have been reported by various researchers such as [26, 31, 53–55]. In our investigation, the Polymorphic Information Content (PIC) values for all analyzed SSR markers varied, ranging from 0.76 for the primer pair BNL193 to 0.86 for the primer pair BNL2827, with an average PIC of 0.82. Additionally, it is crucial to



emphasize that the discerned genetic diversity in the examined germplasm materials is not solely indicated by the varying number of amplified alleles for each marker. It also correlates with other factors, such as the type of marker system utilized, the separation technique of PCR products, and the resolution power of the analysis [56].

On the other side, the genetic similarity co-efficient and phylogenetic analysis results were figured out the genetic relationships amongst the studied cotton genotypes. These results were based on the molecular profiling data of examined cotton genotypes and it might be help to design as well as to conduct a hybridization-based breeding programs with the wide clustered related genotypes [57]. For example, as per our results, the hybridization between the Egyptian cotton genotypes such as Giza 68 with genotypes Pima s6 or Pima high percentage is a suitable parental combination in next breeding schemes due to the high dissimilarity coefficient between these genotypes. With this respect, SSR markers are a highly preferable tool to characterizes different crop genotypes to describe their expansion regarding its genetic diversity as well as it is the suitable choice marker system to assess DNA-based fingerprinting for the major crop improvement schemes [58, 59]. In addition, the higher genetic variability in cotton genotypes was recorded through the implementation of SSR based markers system in cotton genetic diversity analysis and marker assisted selection studies [60]. On the other hand, according to Ditta et al. [41] It was asserted by the individual that a PIC value exceeding 0.5 for each SSR marker indicated the informative capacity of said marker. The findings of our investigation revealed that the polymorphism information content (PIC) results indicated a PIC value over 0.5 for all SSR markers that were evaluated. Thus, in summary, these SSR markers can serve as a valuable tool for cotton crop breeders to investigate the genetic diversity and expand the genetic resources of cotton. This will help identify appropriate parental lines and establish a strong foundation for future marker assisted selection (MAS) schemes aimed at enhancing new modern cotton genotypes.

## Conclusion

In conclusion, the superior performance of Egyptian cotton genotypes, particularly G.96, G.68, and G.93, in key agronomic traits underscores their potential for cultivation and breeding programs. The robust fiber quality traits exhibited by these genotypes further highlight their significance in contributing to high-quality cotton production. The molecular analysis, using SSR markers, not only revealed a substantial level of genetic polymorphism but also facilitated the identification of distinct genetic relationships among the studied genotypes. This

comprehensive understanding of both phenotypic and genotypic characteristics provides valuable insights for cotton breeders and farmers, aiding in the selection and development of improved cotton varieties with enhanced agronomic performance and fiber quality.

## Acknowledgements

Authors would like to acknowledge their universities for supporting the research.

## Authors' contributions

Mona A. Farid, Youssef A. El-Mahalawy, A. A. A. El-Akher, Ali A. Aboshosha, Aysam M. Fayed, W. M. B. Yehia, Sobhi F. Lamloom and Ehab A. A. Salama, design idea, methodology, data analysis and wrote the main manuscript text. Sobhi F. Lamloom and Ehab A. A. Salama, editing and reviewing. All authors reviewed the manuscript.

## Author's information

Not Applicable (NA).

## Funding

Not Applicable (NA).

## Availability of data and materials

All data generated or analysed during this study are included in this published article.

## Declarations

### Ethics approval and consent to participate

This article does not contain any studies with human or animal subjects. The current experimental research and field study including the collection of plant material, is complying with relevant institutional, national, and international guidelines and legislation and used for research and development.

### Consent for publication

Not applicable (NA).

### Competing interests

The authors declare no competing interests.

Received: 20 November 2023 Accepted: 15 March 2024

Published online: 15 May 2024

## References

- Kranthi KR. Cotton production practices: snippets from global data 2017. *The ICAC Recorder*. 2018;36:4–14.
- Zhang Z-S, Xiao Y-H, Luo M, Li X-B, Luo X-Y, Hou L, Li D-M, Pei Y. Construction of a genetic linkage map and QTL analysis of fiber-related traits in upland cotton (*Gossypium hirsutum* L.). *Euphytica*. 2005;144:91–9.
- Chen ZJ, Scheffler BE, Dennis E, Triplett BA, Zhang T, Guo W, Chen X, Stelly DM, Rabinowicz PD, Town CD. Toward sequencing cotton (*Gossypium*) genomes. *Plant Physiol*. 2007;145(4):1303–10.
- Sanamyan MF, Bobohujayev SU, Abdulkarimov SS, Makamov AK, Silkova OG. Features of Chromosome Introgression from *Gossypium barbadense* L. into *G. hirsutum* L. during the Development of Alien Substitution Lines. *Plants*. 2022;11(4):542.
- Sarode D, Sharma K, Shinde N, Gaikwad A, Chavhan R, Pimpale P. Molecular Characterization of Cotton (*Gossypium hirsutum* L.) Genotypes Using SSR Markers for Fiber Quality Traits. *Int J Curr Microbiol App Sci*. 2021;10(03):1748–59.
- Grover CE, Zhu X, Grupp KK, Jareczek JJ, Gallagher JP, Szadkowski E, Seijo JG, Wendel JF. Molecular confirmation of species status for the allopolyploid cotton species, *Gossypium ekmanianum* Wittmack. *Genet Resour Crop Evol*. 2015;62(1):103–14.

7. Wendel J, Brubaker C. RFLP diversity in *Gossypium hirsutum* L. and new insights into the domestication of cotton. *Am J Bot.* 1993;80(6):71.
8. Zhang H-B, Li Y, Wang B, Chee PW. Recent advances in cotton genomics. *International Journal of plant genomics.* 2008;2008:742304.
9. Wendel JF, Brubaker C, Alvarez I, Cronn R, Stewart JM. Evolution and natural history of the cotton genus. *Genet Genomics Cotton.* 2009;3–22.
10. Guiamba HDSS, Zhang X, Sierka E, Lin K, Ali MM, Ali WM, Lamolom SF, Kalaji HM, Telesiński A, Yousef AF. Enhancement of photosynthesis efficiency and yield of strawberry (*Fragaria ananassa* Duch.) plants via LED systems. *Frontiers Plant Science.* 2022;13:918038.
11. H Ren F, Zhang SF, Lamolom B, Zhang. Manipulating rhizosphere microorganisms to improve crop yield in saline-alkali soil: a study on soybean growth and development. *Front Microbiol.* 2023;14:1233351.
12. Ditta A, Zhou Z, Cai X, Shehzad M, Wang X, Okubazghi KW, Xu Y, Hou Y, Sajid Iqbal M, Khan MKR. Genome-wide mining and characterization of SSR markers for gene mapping and gene diversity in *Gossypium barbadense* L. and *Gossypium darwinii* G. walt accessions. *J Agronomy.* 2018;8(9):181.
13. Lacape J-M, Llewellyn D, Jacobs J, Arioli T, Becker D, Calhoun S, Al-Ghazi Y, Liu S, Palai O, Georges S. Meta-analysis of cotton fiber quality QTLs across diverse environments in a *Gossypium hirsutum* x *G. barbadense* RIL population. *BMC Plant Biol.* 2010;10(1):1–24.
14. Lamolom SF, Irshad A, Mosa WF. The biological and biochemical composition of wheat (*Triticum aestivum*) as affected by the bio and organic fertilizers. *BMC Plant Biol.* 2023;23(1):111.
15. Rafiq M, Saqib M, Jawad H, Javed T, Hussain S, Arif M, Ali B, Bazmi MSA, Abbas G, Aziz M. Improving quantitative and qualitative characteristics of Wheat (*Triticum aestivum* L.) through nitrogen application under semiarid conditions. *Phyton.* 2023;92(4):1001–17.
16. Kandil EE, Lamolom SF, Gheith E-SM, Javed T, Ghareeb RY, Abdelsalam NR, Hussain S. Biofortification of maize growth, productivity and quality using nano-silver, silicon and zinc particles with different irrigation intervals. *J Agric Sci.* 2023;161(3):339–355.
17. Abdel-Aty M, Sorour F, Yehia W, Kotb H, Abdelghany AM, Lamolom SF, Shah AN, Abdelsalam NR. Estimating the combining ability and genetic parameters for growth habit, yield, and fiber quality traits in some Egyptian cotton crosses. *BMC Plant Biol.* 2023;23(1):1–21.
18. Ibrahim IA, Yehia W, Saleh FH, Lamolom SF, Ghareeb RY, El-Banna AA, Abdelsalam NR. Impact of plant spacing and nitrogen rates on growth characteristics and yield attributes of Egyptian cotton (*Gossypium barbadense* L.). *Frontier Plant Science.* 2022;13:916734.
19. Ochar K, Su B-h, Zhou M-m, Liu Z-x, Gao H-w, Lamolom SF, Qiu L-j. Identification of the genetic locus associated with the crinkled leaf phenotype in a soybean (*Glycine max* L.) mutant by BSA-Seq technology. *J Integr Agric.* 2022;21(12):3524–39.
20. Ren H, Zhang F, Zhu X, Lamolom SF, Liu X, Wang X, Zhao K, Wang J, Sun M, Yuan M. Cultivation model and deficit irrigation strategy for reducing leakage of bundle sheath cells to CO<sub>2</sub>, improve 13C carbon isotope, photosynthesis and soybean yield in semi-arid areas. *J Plant Physiol.* 2023;285:153979.
21. Abdel-Aty M, Youssef-Soad A, Yehia W, El-Nawsany R, Kotb H, Ahmed GA, Hasan ME, Salama EA, Lamolom SF, Saleh FH. Genetic analysis of yield traits in Egyptian cotton crosses (*Gossypium barbadense* L.) under normal conditions. *BMC Plant Biology.* 2022;22(1):462.
22. Zhang B, Zhao K, Ren H, Lamolom SF, Liu X, Wang X, Zhang F, Yuan R, Wang J. Comparative study of isoflavone synthesis genes in two wild soybean varieties using transcriptomic analysis. *Agriculture.* 2023;13(6):1164.
23. Kohel RJ, Bird LS. Inheritance and linkage analysis of the yellow pulvinus mutant of cotton. *J Cotton Sci.* 2002;6:115–8.
24. AbdElgalil MAS, Hefzy M, Sas-Pasz L, Ali HM, Lamolom SF, Abdelghany AM. Unraveling the influence of water and nitrogen management on quinoa (*Chenopodium quinoa* Willd.) agronomic and yield traits. *Water.* 2023;15(7):1296.
25. Ren H, Zhang B, Zhang F, Liu X, Wang X, Zhang C, Zhao K, Yuan R, Lamolom SF, Abdelghany AM. Integration of physiological and transcriptomic approaches in investigating salt-alkali stress resilience in soybean. *Plant Stress.* 2024;11:100375.
26. L Siu S, Saha D, Stelly B, Burr R, Cantrell. 2000 Chromosomal assignment of microsatellite loci in cotton. *J Hered.* 91 4 326–332.
27. EE Hakki C, Savaskan MS, Akkaya. 2001 Genotyping of Anatolian doubled-haploid durum lines with SSR markers. *Euphytica.* 122 257–262.
28. Lamolom SF, Zhang Y, Su B, Wu H, Zhang X, Fu J, Zhang B, L-j QIU. Map-based cloning of a novel QTL qBN-1 influencing branch number in soybean [*Glycine max* (L.) Merr.]. *The Crop Journal.* 2020;8(5):793–801.
29. Hayatu NG, Liu Y-R, Han T-F, Daba NA, Zhang L, Zhe S, Li J-W, Muazu H, Lamolom SF, Zhang H-M. Carbon sequestration rate, nitrogen use efficiency and rice yield responses to long-term substitution of chemical fertilizer by organic manure in a rice–rice cropping system. *J Integr Agric.* 2023;22(9):2848–64.
30. Kurt Y, Bilgen B, Kaya N, IŞIK K. Genetic comparison of *Pinus brutia* Ten. populations from different elevations by RAPD markers. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca.* 2011;39(2).
31. O Gutierrez S, Basu S, Saha J, Jenkins D, Shoemaker C, Cheatham J, McCarty Jr. 2002 Genetic distance among selected cotton genotypes and its relationship with F<sub>2</sub> performance. *Crop Sci.* 42 6 1841–1847.
32. M Shaukat A, Abbasi K, Ramzan H, Aiman. 2024 MEMON SQ, MAQSOOD Z, GAAFAR A-RZ, HODHOD MS, LAMLLOM SF: Ameliorating heat stressed conditions in wheat by altering its physiological and phenotypic traits associated with varying nitrogen levels. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca.* 52 1 13471–13471.
33. Ibrahim MF, Ali MM, Lamolom SF, Kalaji HM, Yousef AF. Climate change necessitates a change in the cultivation date of caraway (*Carum carvi* L.). *J Water Land Dev.* 2022;54.
34. Chen ZJ. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *J Annu Rev Plant Biol.* 2007;58:377–406.
35. Elmahdy AM, Ahmed YM, Bakr AA, Abdallah AM, Abdelghany AM, El-Sorady GA, et al. Revolutionizing Maize Farming with Potassium Silicate Foliar Spray and Water Management Techniques. *Silicon.* 2023;15(16):7121–35.
36. Manonmani K, Mahalingam L, Malarvizhi D, Premalatha N, Sritharan N. Molecular diversity studies in cotton (*Gossypium hirsutum* L.) using SSR markers. *Int J Curr Microbiol App Sci.* 2019;8(7):1731–7.
37. Ren H, Zhao K, Zhang C, Lamolom SF, Liu X, Wang X, Zhang F, Yuan R, Gao Y, Cao B. Genetic analysis and QTL mapping of seed hardness trait in a soybean (*Glycine max*) recombinant inbred line (RIL) population. *Gene.* 2024;905:148238.
38. Dellaporta SL, Wood J, Hicks JB. A plant DNA miniprep: version II. *Plant Mol Biol Report.* 1983;1:19–21.
39. Saif I, MA S, Riad S, Elbagoury M. Molecular characterization of some Egyptian cotton varieties. *Alexandria Sci Exch J.* 2017;38:44–52.
40. Anderson JA, Churchill G, Autrique J, Tanksley S, Sorrells M. Optimizing parental selection for genetic linkage maps. *Genome.* 1993;36(1):181–6.
41. Ditta A, Zhou Z, Cai X, Wang X, Okubazghi KW, Shehzad M, Xu Y, Hou Y, Sajid Iqbal M, Khan MKR. Assessment of genetic diversity, population structure, and evolutionary relationship of uncharacterized genes in a novel germplasm collection of diploid and allotetraploid *Gossypium* accessions using EST and genomic SSR markers. *Int J Mol Sci.* 2018;19(8):2401.
42. Nei M, Li W-H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci.* 1979;76(10):5269–73.
43. Rohlf F. NTSYS-pc: microcomputer programs for numerical taxonomy and multivariate analysis. *Am Stat.* 1987;41:330.
44. Jaccard P. Nouvelles recherches sur la distribution florale. *Bull Soc Vaud Sci Nat.* 1908;44:223–70.
45. Obaid-ur-Rehman SHK, Sajjad M. Genetic diversity among cotton (*Gossypium hirsutum* L.) genotypes existing in Pakistan. *Am Eurasian J Sustain Agric.* 2009;3(4):816–23.
46. Yehia W, El-Hashash EF. Estimates of genetic parameters for cotton yield, its components, and fiber quality traits based on line x tester mating design and principal component analysis. *Egypt J Agric Res.* 2022;100(3):302–15.
47. Abdel-Aty MS, Youssef-Soad A, Yehia WMB, El-Nawsany RTE, Kotb HMK, Ahmed GA, Hasan ME, Salama EAA, Lamolom SF, Saleh FH, et al. Genetic analysis of yield traits in Egyptian cotton crosses (*Gossypium barbadense* L.) under normal conditions. *BMC Plant Biol.* 2022;22(1):462.
48. Ambreen I, Sadia A, Usman I, Tayyaba S. Molecular characterization of cotton using simple sequence repeat (SSR) markers and application of genetic analysis. *Int J Genet Mol Biol.* 2013;5(4):49–53.
49. Medini M, Hamza S, Rebai A, Baum M. Analysis of genetic diversity in Tunisian durum wheat cultivars and related wild species by SSR and AFLP markers. *Genet Resour Crop Evol.* 2005;52:21–31.

50. Code C. The efficiency of AFLP and SSR markers in genetic diversity estimation and gene pool classification of common bean (*Phaseolus vulgaris* L.). *J Acta Agric Slov*. 2008;91:87–96.
51. Kuang Z, Xiao C, Ilyas MK, Ibrar D, Khan S, Guo L, Wang W, Wang B, Huang H, Li Y. 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*. 2022;12(7):1513.
52. Dongre A, Bhandarkar M, Banerjee S. Genetic diversity in tetraploid and diploid cotton (*Gossypium* spp.) using ISSR and microsatellite DNA markers. NIScPR Online Periodicals Repository. 2007.
53. CH Bertini I Schuster T Sedyama 2006 Barros EGd, Moreira MA: Characterization and genetic diversity analysis of cotton cultivars using microsatellites *Genetics Molecular Biology* 29 321–329
54. Khan AI, Fu Y-B, Khan IA. Genetic diversity of Pakistani cotton cultivars as revealed by simple sequence repeat markers. *Commun Biometry Crop Sci*. 2009;4(1):21.
55. Ma Q, Zhao J, Lin H, Ning X, Liu P, Deng F, Si A, Li J. Association between SSR markers and fibre traits in sea island cotton (*Gossypium barbadense*) germplasm resources. *J Genet*. 2017;96:55–63.
56. Lacape J-M, Nguyen T-B, Thibivilliers S, Bojinov B, Courtois B, Cantrell RG, Burr B, Hau B. A combined RFLP SSR AFLP map of tetraploid cotton based on a *Gossypium hirsutum* × *Gossypium barbadense* backcross population. *Genome*. 2003;46(4):612–26.
57. Mishra KK, Fougat RS, Ballani A. Validation of fiber quality linked SSR markers derived from allotetraploid (*Gossypium hirsutum*) in diploid (*Gossypium arboreum*). *Int J Sci R Know*. 2013;1:349–57.
58. Bligh HFJ, Blackhall NW, Edwards KJ, McClung AM. Using amplified fragment length polymorphisms and simple sequence length polymorphisms to identify cultivars of brown and white milled rice. *Crop Sci*. 1999;39(6):1715–21.
59. Jeung JU, Hwang HG, Moon HP, Jena KK. Fingerprinting temperate japonica and tropical indica rice genotypes by comparative analysis of DNA markers. *Euphytica*. 2005;146(3):239–51.
60. Li Z, Wang X, Zhang Y, Zhang G, Wu L, Chi J, Ma Z. Assessment of genetic diversity in glandless cotton germplasm resources by using agronomic traits and molecular markers. *Front Mech Eng China*. 2008;2:245–52.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.