

Article

# Sensitivity of Cotton *Gossypium barbadense* L. to Various Selection Criteria for Yield, Yield Components, and Earliness Attributes

Mohamed S. Abdel-Aty<sup>1</sup>, Fouad. A. Sorour<sup>1</sup>, Haitham. M. K. Kotb<sup>2</sup>, Omar M. Ibrahim<sup>3</sup>, Ammar AL-Farga<sup>4\*</sup> and Amira M. El-Tahan<sup>3\*</sup>

<sup>1</sup> Agronomy Department, Faculty of Agriculture, Kafrelsheikh University, 33516, Egypt.

<sup>2</sup> Cotton Breeding Department, Cotton Research Institute, Agriculture Research Center, Egypt.

<sup>3</sup> Plant Production Department, Arid Lands Cultivation Research Institute, the City of Scientific Research and Technological Applications, SRTA-City, Borg El Arab, Alexandria, Egypt.

<sup>4</sup> Department of Biochemistry, College of Science, University of Jeddah, Jeddah, Saudi Arabia.

\* Correspondence: amalfarga@uj.edu.sa (A.A-F); aeltahan@srtacity.sci.eg (A.M.E-T)

**Abstract:** Competition between cotton, winter, and summer crops is the greatest obstacle for plant breeders in Egypt. Breeders are thus seeking early maturing, high-yielding cultivars that are suited for planting following winter crops. This study was conducted during the 2016 and 2018 growing seasons at the Sakha agriculture research station, Egypt's agricultural research facility. This research used breeding material from the F<sub>2</sub>, F<sub>3</sub>, and F<sub>4</sub> generations that resulted from crossing Giza 94 and Uzbekistan. During the 2016 growing season, the F<sub>2</sub> population was cultivated, and the top 5% of the F<sub>2</sub> population (75 plants) were chosen based on the selection criteria. In the summer of 2017, the F<sub>3</sub> families were cultivated, and 5% of the superior F<sub>3</sub> families were chosen to become the F<sub>4</sub> families, which were planted in the summer of 2018 at the single plant and bulk studies in a randomised full block design with three replications. The average performance of the characteristics boll weight, seed cotton yield, lint percentage, and lint index increased as generations passed from F<sub>2</sub> to F<sub>4</sub>. In the meanwhile, the characteristic value of the first flower to open decreased, showing the accumulation of advantageous alleles. For all examined characteristics, the phenotypic coefficient of variance was greater than the genotypic coefficient of variation. In addition, the data demonstrated that heritability in a broad sense was greater in F<sub>4</sub> populations than in F<sub>3</sub> and F<sub>2</sub> populations. With a few exceptions, the phenotypic and genotypic associations for all investigated characteristics were positive and very significant. According to the data, initial opening flowering and boll weight had a substantial impact on seed cotton output per feddan. In general, five superior families No. 9, 13, 18, 19, and 20 were chosen from the F<sub>4</sub> population, which were more significant for yield and yield component than the commercial variety, and these genotypes may be promising for the breeding programme to develop new genotypes with superior yield.

**Keywords:** cotton; crossing; genotypes; phenotypic coefficient

## 1. Introduction

Cotton is one of the essential commercial fibers produced and is one of the most critical commodities imported and exported worldwide due to its uses in the textile industry, mattress, bedding industry, oil industry, soap industry, etc. [1]. Cotton belongs to the genus *Gossypium* L. (Malvaceae), containing 52 species; 45 are diploid  $2n = 2x = 26$  chromosomes, and seven are allotetraploid:  $2n = 4x = 52$  chromosomes [2]. Only four species are the most widely cultivated. Two of them named old-world diploid species (*G. arboreum* L. and *G. herbaceum* L.), representing 2% of cotton production [3-5]. The other two are new world allotetraploid species (*G. barbadense* L. and *G. hirsutum* L.) [6]. The cultivated area of cotton is going lower year after year, despite its importance for the national economy [7-10]. The most critical challenges facing the plant breeder in Egypt are competition

among cotton, winter and summer crops. Therefore breeders are looking for early maturity and high-yielding varieties adapted for the late planting after winter crops [3, 11-13]. Traditional breeding methods such as pedigree selection have successfully developed improved varieties. Most Egyptian cotton varieties were developed by this method. All of Egyptian cotton varieties developed and produce by using selection method in a commercial varieties or by hybridization between two parents and selected by pedigree methods during the period 1907-2022. More than 97 cotton varieties were created and approved at the Egyptian cotton research institute. Various breeding methods were used to increase the earliness and productivity, and to improve the quality of cotton fiber. To this end, hybridization of *G. barbadense* L. and its combination with intraspecific hybridization were used. A new stage was a step forward in the cotton selection for increasing the productivity and fiber lint percentage.

This study was conducted to select the best genotypes which possess early maturity and high yield throughout the improvement of seed cotton yield. It was also to correlate traits using pedigree selection and study the performance and (phenotypic and genotypic) correlation in the genotypes.

## 2. Materials and Methods

The present study was carried out at Sakha Agriculture Research Station, Agriculture Research Center, during three successive summer seasons of 2016, 2017, and 2018. The breeding materials used in this study were the F<sub>2</sub>, F<sub>3</sub>, and F<sub>4</sub> generations from crossing between the two varieties Giza 94 (characterized by high yield, high lint percentage, and earliness) and Uzbekistan is boling to *Gossypium barbadense* L. (characterized by high boll weight values and lint percentage).

In the 2016 season, the F<sub>2</sub> population and their parents were grown in rows 4.5m. long and the spacing between rows was 70 cm, and plant to plant was 70 cm the number of F<sub>2</sub> populations was 800 plants and every parent the number of plants was 50 plants to every parent, one plant per hill was maintained. As usual, all agricultural practices were carried out for ordinary cotton fields in the area. Self-pollination was practiced for all F<sub>2</sub> plants at the harvest 375 guarded plants were picked up separately. First opening flower (F.O.F.) (I1), boll weight (B.W.) (I2), seed cotton yield (S.C.Y.) (I3), lint Percentage (L.%) (I4), and lint index (L.I.) (I5) using the selection criteria. As well as using 5% selection intensity, the plants having the highest performance in each selection criteria were saved; these gave a total of 75 F<sub>3</sub> selected progenies (fifteen superior progenies from each selection criteria).

In the 2017 season, 75 self-selected plants from F<sub>2</sub> were sown to produce plants of F<sub>3</sub> families. Two rows represented each plot for each plant: 4.5 m long and 0.7 m wide. Seeds were planted on one side of the ridge at 70 cm between hills. The recommended cultural practices were done through the growing season. The best 44 plants of the 75 F<sub>3</sub> families (selection intensity=5%) were chosen for the five selection criteria.

In the 2018 season, the seeds of 44 F<sub>3</sub> families (selfed seeds) were planted to represent the plants of the F<sub>4</sub> family and the two parents and commercial variety Giza 86. A randomized complete block design with three replications was used; plot size one row, 4.5 m. long and 0.7 m. wide with 70 cm. hill spacing. Hills were thinned to one plant per hill. The normal cultural practices for cotton production were performed at the proper time. The studied characters were; first opening flower, boll weight, seed cotton yield, lint cotton yield, lint Percentage, and lint index.

The ordinary analysis of variance for randomized complete block design was done according to [15]. Estimates of phenotypic (P.C.V. %)

$$\text{The genotypic coefficient of variation (GCV)} = \frac{\sqrt{\sigma_g^2}}{\bar{x}} \times 100$$

and genotypic (G.C.V. %)

$$\text{The phenotypic coefficient of variation (PCV)} = \frac{\sqrt{\sigma_{ph}^2}}{\bar{x}} \times 100$$

coefficient of variations, phenotypic ( $\sigma^2_{ph}$ ) =  $\frac{\sigma^2_g + \sigma^2_e}{r}$

and genotypic ( $\sigma^2_g$ ):

$$\sigma^2_g = (M_F - M_E)/r$$

variances were determined according to [16]. The calculation of phenotypic and genotypic correlation coefficients was calculated as outlined by [17]

$$\text{Phenotypic correlation } (r_{ph}) = \frac{\sigma_{pij}}{\sqrt{\sigma^2_{pi} \cdot \sigma^2_{pj}}}$$

$$\text{Genotypic correlation } (r_g) = \frac{\sigma_{gij}}{\sqrt{\sigma^2_{gi} \cdot \sigma^2_{gj}}}$$

$$h_b^2 = \frac{\sigma^2_g}{\sigma^2_{ph}} \times 100$$

Heritability in the broad sense was obtained as described by [18].

The data collected and were subjected to statistical analysis using analysis of variance technique and LSD (least significant test and Duncan values to describe the highest of treatment means using MSTAT-C computer statistical software (51)

### 3. Results

Estimates of means, phenotypic and genotypic coefficient of variations (PCV and GCV), and heritability in the broad sense ( $H_b^2$  %) for  $F_2$ ,  $F_3$ , and  $F_4$  generations for all the studied traits are presented in table 1. The results showed that the mean values of earliness F.O.F. (first opening flowe) were 71.81, 71.39, and 69.77 days; the values of PCV were 4.54, 1.61, and 3.62, while the GCV values were 3.27, 1.36, and 3.34 for  $F_2$ ,  $F_3$ , and  $F_4$  generations respectively. The mean values were 3.25, 3.37, and 3.77 gm concerning boll weight. The phenotypic coefficient of variation values were 14.06, 7.42, and 6.13, while the values of GCV were 7.00, 6.19, and 5.79 for  $F_2$ ,  $F_3$ , and  $F_4$  generations, respectively. Regarding the seed cotton yield per plant (S.C.Y.), the mean values were 270.08, 304.61, and 317.33 gm./plant, these results appeared increase the potential yield from  $F_2$  generation to  $F_4$  generation and stable genes performance as well as increase the additive variance and more accumulation from its. The PCV values were 40.2, 17.85, and 15.19, while the GCV values were 16.40, 12.87, and 14.98 for  $F_2$ ,  $F_3$ , and  $F_4$  generations respectively. For lint percentage (L %), the mean values were 38.64, 39.91, and 41.37 %, the PCV values were 5.2, 4.6, and 3.31, and the GCV values were 4.68, 4.44, and 3.27 for  $F_2$ ,  $F_3$ , and  $F_4$  generations respectively. Concerning the lint index, the mean values were 7.19, 7.12, and 8.32, and the PCV values were 14.27, 10.15, and 6.21, while the GCV values were 11.99, 9.37, and 6.05 for  $F_2$ ,  $F_3$ , and  $F_4$  generation, respectively. The results showed that the values of heritability in the broad sense ( $h_b^2$ %) in  $F_4$  were more significant than the  $F_3$ . Also the values of  $F_3$  were greater than the  $F_2$  generation for all the studied traits. These results agree with those reported by [4, 5, 19-21].

**Table 1.** Range and mean values, phenotypic and genotypic coefficients of variations, and heritability in the broad sense for F<sub>2</sub>, F<sub>3</sub>, and F<sub>4</sub> generations for all the selected traits.

Traits	Gen.	Parameters					
		Range		X <sup>-</sup>	PCV	GCV	H <sup>2</sup> b %
		Min.	Max.				
F.O.F	F <sub>2</sub>	63.00	81.00	71.81	4.54	3.27	51.89
	F <sub>3</sub>	56.00	75.00	71.39	1.61	1.36	71.02
	F <sub>4</sub>	64.00	73.00	69.77	3.62	3.34	85.49
B.W	F <sub>2</sub>	1.45	4.04	3.25	14.06	7.00	24.80
	F <sub>3</sub>	2.43	4.35	3.37	7.42	6.19	69.65
	F <sub>4</sub>	3.31	4.24	3.77	6.13	5.79	89.26
S.C.Y	F <sub>2</sub>	12.43	649.20	270.08	40.20	16.40	51.02
	F <sub>3</sub>	112.53	707.46	304.61	17.85	12.87	52.03
	F <sub>4</sub>	190.10	445.00	317.33	15.19	14.98	97.21
L%	F <sub>2</sub>	31.64	44.62	38.64	5.20	4.68	80.92
	F <sub>3</sub>	34.77	45.82	39.91	4.60	4.44	93.04
	F <sub>4</sub>	37.73	45.37	41.37	3.31	3.27	97.20
L.I	F <sub>2</sub>	3.09	9.48	7.19	14.27	11.99	70.62
	F <sub>3</sub>	4.20	9.06	7.12	10.15	9.37	85.23
	F <sub>4</sub>	6.99	9.97	8.32	6.21	6.05	94.85

- First opening flower (F.O.F), boll weight (B.W.), seed cotton yield (S.C.Y.), lint Percentage (L%) and lint index (L.I.).

#### *Selection criteria*

The mean values and phenotypic and genotypic coefficient of variations for yield, its component and earliness traits for the five different selection criteria from F<sub>4</sub> families are presented in Table 2. Data showed that the mean value of the first opening flower as a selection criterion was the lowest and desirable value (65.86 days). Meanwhile, the mean values of the first opening flower when selection based on boll weight, seed cotton yield per plant, lint percentage. Lint indexes were 70.96, 67.11, 70.80, and 71.20 days respectively. Also, the results cleared that the values of PCV and GCV were 1.84 and 1.41, respectively, and lower than the seed cotton yield and lint percentage as selection criteria. Concerning selection criterion boll weight (I<sub>2</sub>), the mean value was moderate (3.77 gm) compared to the first opening flower and seed cotton yield as selection criteria were low (3.57 and 3.63 gm). When lint percentage and lint index were used as selection criteria, the boll weight were high (3.85 and 3.93 gm. respectively). Also PCV and GCV for boll weight in selection criteria I<sub>2</sub> had the highest values (7.61 and 7.26%) compared to the other indices except for the lint index. For the lint percentage criterion, the results cleared that the mean value of L% was high (41.94%).

**Table 2.** mean phenotypic and genotypic coefficient of variations for earliness and yield and its components for the five different selection criteria from F<sub>4</sub> selected families.

Selection criteria	Traits	Parameters		
		Mean (X̄)	PCV %	GCV %
I 1	First opening Flower (F.O.F)	65.86	1.84	1.41
	Boll weight (B.W)	3.57	3.63	3.09
	Seed cotton yield / plant (S.C.Y./P.)	106.32	13.17	12.88
	Lint percentage (L%)	40.12	1.34	1.25
	Lint index (L.I)	7.74	2.83	2.46
I 2	First opening Flower (F.O.F)	70.96	1.45	0.67
	Boll weight (B.W)	3.77	7.61	7.26
	Seed cotton yield / plant (S.C.Y./P.)	119.02	16.75	16.42
	Lint percentage (L%)	40.33	1.74	1.63
	Lint index (L.I)	8.43	5.56	5.31
I 3	First opening Flower (F.O.F)	67.17	3.93	3.86
	Boll weight (B.W)	3.63	4.2	3.53
	Seed cotton yield / plant (S.C.Y./P.)	116.51	4.92	4.38
	Lint percentage (L%)	39.56	3.52	3.45
	Lint index (L.I)	7.56	2.83	2.42
I 4	First opening Flower (F.O.F)	70.8	2.61	2.13
	Boll weight (B.W)	3.85	5.33	5
	Seed cotton yield / plant (S.C.Y./P.)	119.74	15.43	15.28
	Lint percentage (L%)	41.94	2.53	2.48
	Lint index (L.I)	8.43	4.71	4.5
I 5	First opening Flower (F.O.F)	71.2	1.59	1.23
	Boll weight (B.W)	3.93	7.7	7.35
	Seed cotton yield / plant (S.C.Y./P.)	122.07	11.5	11.14
	Lint percentage (L%)	41.27	2.16	2.11
	Lint index (L.I)	8.86	2.66	2.33

In contrast, the mean values of L% were low (40.12, 40.33, 39.56, and 41.27%) when the selection was based on first opening flower (F.O.F), boll weight( B.W.), seed cotton yield per plant( S.C.Y.), and lint index(L.I.) as criteria, respectively. Also, the PCV and GCV coefficient of variations were higher with 4.71 and 4.50 % values than the other criteria except for the boll weight criterion. The results also cleared that the mean values of lint index as a selection criterion were high (8.86, 8.43, 7.56, and 8.43) when first opening flower, boll weight, seed cotton yield, and lint percentage were used as selection criteria.

*Mean performance of the 44 selected families of the F<sub>4</sub> generation.*

The mean performance of the 44 selected families of F<sub>4</sub> generation from the population (G94 X Uzbekistan) for all the studied traits was calculated (Table 3). For the first opening, flowers ranged from 64 to 73 days, and the families number 6, 7, 8, 17, 22, 23, and 34 were shown highly significant values compared with the better parent.

**Table 3.** Mean performances of the 44 selected families from the cross G.94 x Uzbekistan for all the studied traits.

Family Number	F.O.F	B.W	S.C.Y	L%	L.I
1	72.00	4.03	110.86	40.86	8.84
2	71.00	3.87	109.64	40.28	8.43
3	67.00	3.54	118.41	39.63	7.77
4	67.00	3.36	119.46	39.74	7.43
5	71.00	3.58	121.69	41.05	7.83
6	66.00	3.68	123.99	40.29	7.58
7	65.00	3.76	108.38	39.98	7.55
8	66.00	3.60	116.10	40.29	7.71
9	71.67	4.00	159.60	38.87	7.76
10	71.00	3.17	106.36	40.43	8.66
11	71.00	3.84	136.53	40.76	8.73
12	69.00	3.95	122.40	41.16	8.89
13	70.00	3.93	153.66	41.01	9.00
14	71.00	3.94	122.21	40.88	8.81
15	70.00	3.62	105.84	40.30	8.47
16	72.00	3.65	100.91	39.94	7.68
17	65.00	3.78	112.57	37.36	7.47
18	72.00	3.80	156.58	42.30	8.58
19	70.00	3.83	156.38	42.00	8.26
20	72.00	3.70	151.70	41.80	8.13
21	70.00	3.37	116.35	38.39	7.21
22	64.00	3.47	94.04	40.09	7.94
23	65.00	3.59	81.70	39.87	7.71
24	67.00	3.66	106.16	41.24	8.07
25	72.00	3.73	118.99	43.86	8.59
26	71.00	3.57	103.74	42.83	8.29
27	73.00	3.55	106.91	42.59	8.08
28	73.00	3.89	109.54	43.20	8.39
29	72.00	3.48	107.04	42.62	8.13
30	73.00	3.81	99.91	43.73	8.36
31	73.00	4.26	123.27	42.35	9.31
32	72.00	4.02	119.03	41.86	8.89
33	72.00	4.25	116.77	43.09	9.03
34	66.00	3.68	115.99	42.65	7.91
35	68.00	4.00	141.49	40.26	8.06
36	69.33	3.87	122.40	41.73	8.41
37	70.00	3.86	148.25	40.78	8.70
38	70.00	4.02	117.08	41.30	8.88
39	69.00	3.64	97.89	41.54	7.83
40	70.00	3.94	103.19	40.63	8.70
41	71.00	4.12	119.55	40.69	8.74
42	69.00	3.95	108.29	40.03	8.26
43	72.00	3.88	110.26	41.49	8.56
44	69.00	3.77	99.73	41.19	7.77
<b>overall mean</b>	69.77	3.77	118.20	41.07	8.26
<b>better parent</b>	70.00	3.97	117.26	39.57	8.47
<b>commercial variety</b>	72.33	3.76	102.92	38.30	7.62
<b>L.S.D0.05</b>	2.69	0.21	8.39	0.64	0.33
<b>L.S.D0.01</b>	3.55	0.28	11.09	0.84	0.43

- First opening flower( F.O.F, boll weight (B.W.), seed cotton yield(S.C.Y.), lint Percentage ( L.%) and lint index ( L.I).

Moreover, eleven families had highly significant and desirable for first opening flower compared with commercial variety these families are 3, 4, 6, 7, 8, 17, 22, 23, 24, 34 and 35. Thus the cotton breeders could utilize these families in breeding programs to produce new earlier promising varieties than parents and commercial variety. The results for boll weight (B.W.) cleared that the boll weight for F<sub>4</sub> selected families ranged from 3.17 to 4.26 gm. And the families no. 31, 33, and 41 were differed significantly compared to the better parent. At the same time, 24 families from 44 studied families were higher than the commercial variety (G.86). For seed, cotton yield per plant ranged from 81.7 to 159.6 gm./plant for families no. 23 and 9, respectively. Families No. 9, 11, 13, 18, 19, 20, 35, and 37 were higher than the better parent for seed cotton yield per plant. Also, the results cleared that 38 families from 44 selected F<sub>4</sub> generation families were higher than the commercial variety for seed cotton yield per plant, and 24 families had highly significant differences compared to the commercial variety. For lint percentage (L%) for F<sub>4</sub> selected families ranged from 37.36 to 43.86% for families no. 17 and 25, respectively. the results showed that 43 from 44 selected families were higher than the better parent and the families no. 1, 5, 10, 11, 12, 13, 14, 18, 19, 20, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43 and 44 were highly significant compared to better parent for the both traits. On the other hand, for commercial variety (G.86), the results indicated that all selected F<sub>4</sub> families were higher than commercial variety for lint percentage. All selected families had highly significant differences compared to commercial variety except families no. 9 and 21, which did not differ significantly. For the lint index (L.I), the value ranged from 7.21 for family no. 21 to 9.31 for family no. 31. Mean values of 16 families from 44 selected families had higher significant values than the better parent, these families were: 1, 10, 11, 12, 13, 14, 18, 25, 31, 32, 33, 37, 38, 40, 41 and 43. On the other hand, the commercial variety, the results recorded that 39 from 44 selected families were wider commercial variety and 30 from these 39 higher selected families were differed highly significant from the commercial variety with except families no. 3,5,8, 9, 16, 22, 23, 34, 39 and 44.

#### *Phenotypic and genotypic correlation coefficients.*

Phenotypic and genotypic correlation coefficients were calculated among all the studied traits for the F<sub>4</sub> selected families, and the results are presented in Table 4. A significant positive phenotypic correlation coefficient ( $r_{ph}$ ) was found between the first opening flower and each boll weight, lint percentage, and lint index. The boll weight showed a highly significant positive phenotypic correlation with seed cotton yield and lint index and a significant positive correlation between lint percentage and lint index. To genotypic correlation coefficients ( $r_g$ ), the results showed a highly substantial positive genotypic correlation between first opening flower and each of boll weight, seed cotton yield, lint percentage, lint index, and boll weight correlated significant positive with lint index. The same relationship was found between lint percentage and lint index. The results showed that the two characters i.e., first opening flower and boll weight, significantly affected seed cotton yield per feddan. These results are in agreement with those obtained by [22-26].

**Table 4.** Phenotypic ( $r_{ph}$ ) (above diagonal) and genotypic ( $r_g$ ) (below diagonal) correlation coefficients among all the studied traits for the F<sub>4</sub> selected families.

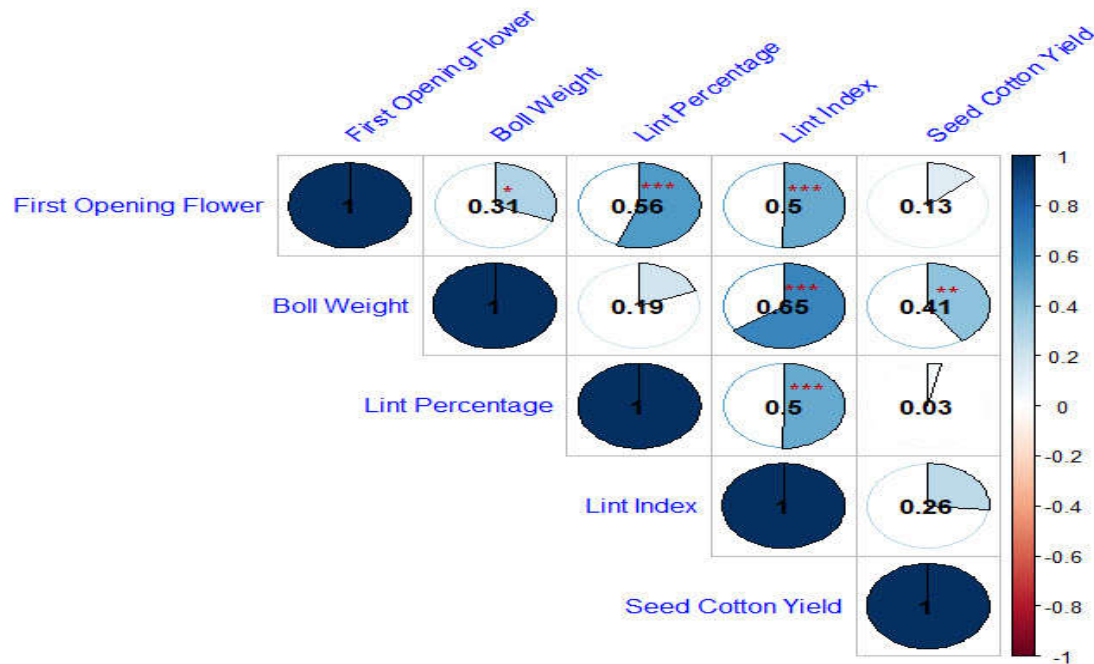
Traits	F.O.F	BW	SCY	L%	LI
F.O.F		0.321*	0.228	0.542**	0.574**
BW	0.711**		0.32*	0.21	0.623**
SCY	0.407**	0.165		0.002	0.218
L%	0.756**	0.059	-0.125		0.501**
LI	0.99**	0.538**	0.02	0.421**	

\*and \*\* significant and highly significant at 0.05 and 0.01 probability levels, respectively.

- First opening flower (F.O.F), boll weight (B.W.), seed cotton yield (S.C.Y.), lint Percentage (L%) and lint index (L.I).

### Correlation analysis:

Figure (1) illustrates the Spearman correlation matrix among the studied traits. The figure reveals that Seed Cotton Yield (SCY) has a moderate and positive significant association (0.41) with Boll Weight (BW). However, no association was found between SCY and First Opening Flower (FOF), Lint Percentage (LP), or Lint Index (LI). Both LI and LP have a moderate and significant association with FOF (0.5 and 0.56), respectively. At the same time, BW has a weak but significant association with FOF (0.31).

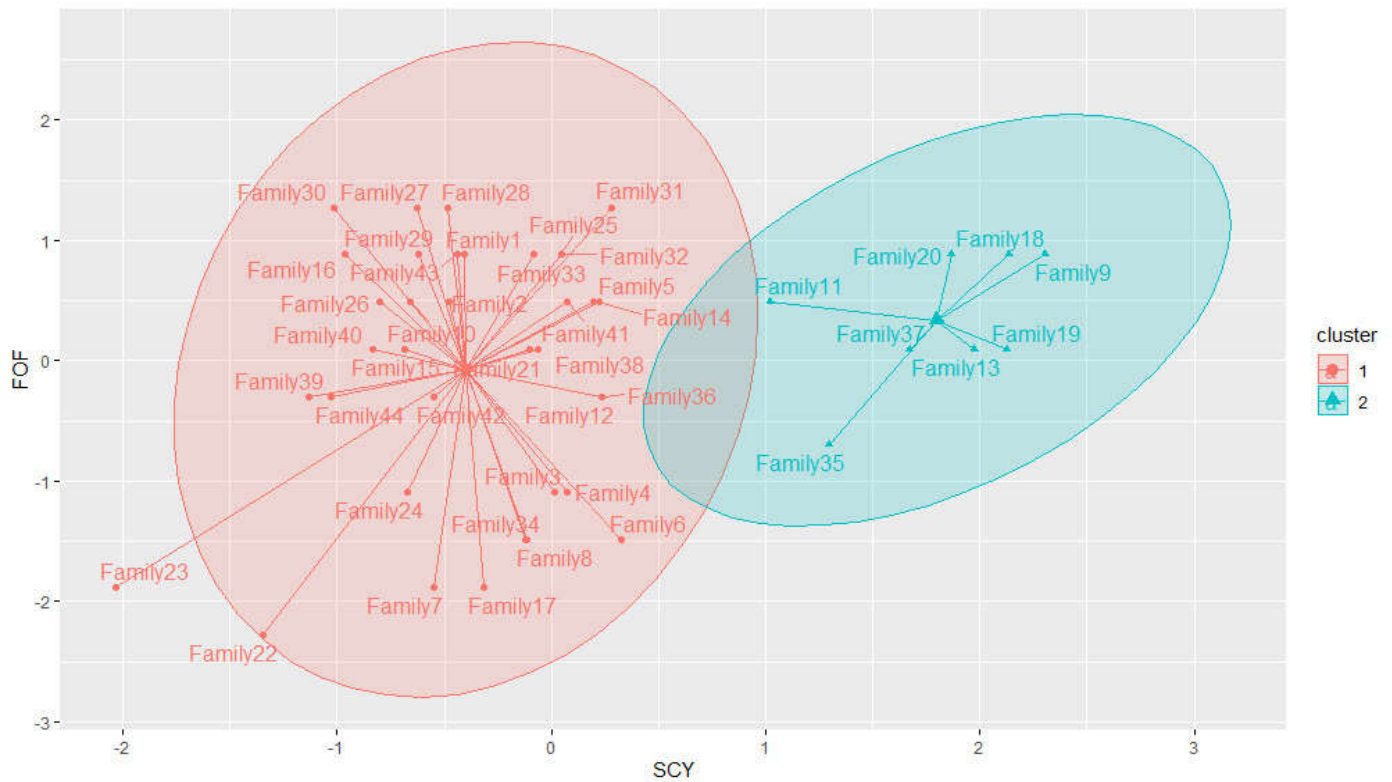


**Figure 1.** Spearman correlation matrix among the studied traits.

On the other hand, LI has a moderate and significant association with BW and LP (0.65 and 0.5). An indirect association between SCY and LI via BW exists since SCY is associated with BW and LI is associated with BW. Also, an indirect association between SCY and FOF via BW is observed; BW directly affects SCY and an indirect effect via FOF.

### Cluster analysis:

Clustering involves observations grouping based on two criteria: (i) external separation where the observations within a certain cluster are well-separated from entities in another cluster, and (ii) compactness or cohesion that describes the similarity between observations within the same cluster [27]. Cluster analysis was performed using Squared Euclidian distance dissimilarity measure and Fuzzy C-means clustering algorithm [28, 29]. Before computing the Squared Euclidian distance, the data were standardized due to the different scales of the studied traits by subtracting each value from the mean and dividing by standard deviation. All the studied traits were used to construct the distance matrix using the Euclidian Squared coefficient. The distance matrix data based on SCY and FOF was used to visualize clustering results among all the families. To identify the optimal number of clusters, 42 cluster validity indices were used [30]. These indices effectively partition observations into homogeneous groups [31, 32]. We can also observe that all the families were grouped into 2 clusters (Figure 2). Families' numbers 9,11,13,18,19,20,35, and 37 belong to cluster 2, while the rest of the families occupied cluster 1.



**Figure 2.** Results of cluster analysis using Fuzzy C-means and Squared Euclidean distance.

The averages of the studied traits are shown in Table (5). The characteristic that differentiates easily between the two clusters is SCY, where Cluster 2 (150.15) outperforms cluster1 (110.69). These results emphasize that selection from cluster2 is better than selection from cluster1 in breeding programs. A slight difference between the two clusters was found in FOF, while BW, LP, and LI were almost similar in the two clusters.

**Table 5.** Average of first opening flower (FOF), boll weight (BW), seed cotton yield (SCY), lint percentage (LP), and lint index (LI) for the two clusters.

	FOF	BW	SCY	LP	LI
<b>Cluster1</b>	69.64	3.75	110.69	41.09	8.22
<b>Cluster2</b>	70.59	3.87	150.15	41.01	8.39

- First opening flower (F.O.F), boll weight (B.W.), seed cotton yield (S.C.Y.), lint Percentage (L.%) and lint index (L.I.)

Every family has a membership value in Table (6) ranging from 0 to 1. These values explain the degree of belonging a particular family to a specific cluster. A membership value close to 1 means high belonging; during a membership close to 0 means lower belonging, the membership of all families to their cluster was distinct, over 0.90, except for a few families such as 5,14,23,31,36 which have membership between 0.8 and 0.9. In contrast, families 6 and 11 have membership less than 0.8.

Table 6. Membership of the families to the clusters.

Family no.	Cluster 1	Cluster 2	Family no.	Cluster 1	Cluster 2	Family no.	Cluster 1	Cluster 2
Family1	0.9961	0.0039	Family16	0.9594	0.0406	Family31	0.8089	0.1911
Family2	0.9978	0.0022	Family17	0.9736	0.0264	Family32	0.9272	0.0728
Family3	0.9368	0.0632	Family18	0.0209	0.9791	Family33	0.9594	0.0406
Family4	0.9171	0.0829	Family19	0.0189	0.9811	Family34	0.9644	0.0356
Family5	0.8681	0.1319	Family20	0.003	0.997	Family35	0.0796	0.9204
Family6	0.7869	0.2131	Family21	0.9659	0.0341	Family36	0.8484	0.1516
Family7	0.9842	0.0158	Family22	0.9115	0.0885	Family37	0.0029	0.9971
Family8	0.9645	0.0355	Family23	0.8453	0.1547	Family38	0.9634	0.0366
Family9	0.0385	0.9615	Family24	0.9861	0.0139	Family39	0.9432	0.0568
Family10	0.9889	0.0111	Family25	0.9226	0.0774	Family40	0.9749	0.0251
Family11	0.2170	0.7830	Family26	0.9759	0.0241	Family41	0.9204	0.0796
Family12	0.8484	0.1516	Family27	0.9854	0.0146	Family42	0.9959	0.0041
Family13	0.007	0.9930	Family28	0.9898	0.0102	Family43	0.9962	0.0038
Family14	0.8525	0.1475	Family29	0.9887	0.0113	Family44	0.9547	0.0453
Family15	0.9878	0.0122	Family30	0.9497	0.0503			

Figure (3) showed the relationship between the families and the studied traits based on scaled data using a color scale. Families number 9,11, 13, 18, 19, 20, 35, and 37 have higher SCY (red color). It is apparent from the figure that their high seed yield is attributed to all the other traits where these traits for those families also have higher values (color adjacent to red); this demonstrates that all these traits contribute to the seed yield of these families where any selection criteria in these families would lead to select high yield genotype. Families number 31 and 33 have higher values of BW and LI; however, these high values didn't lead to a high seed yield. Also, family numbers 25, 28, and 30 have high LP. This high value didn't lead to high seed yield; this contradiction reveals that selection based on BW, LI, or LP criteria would not lead to selecting a high seed yield genotype.

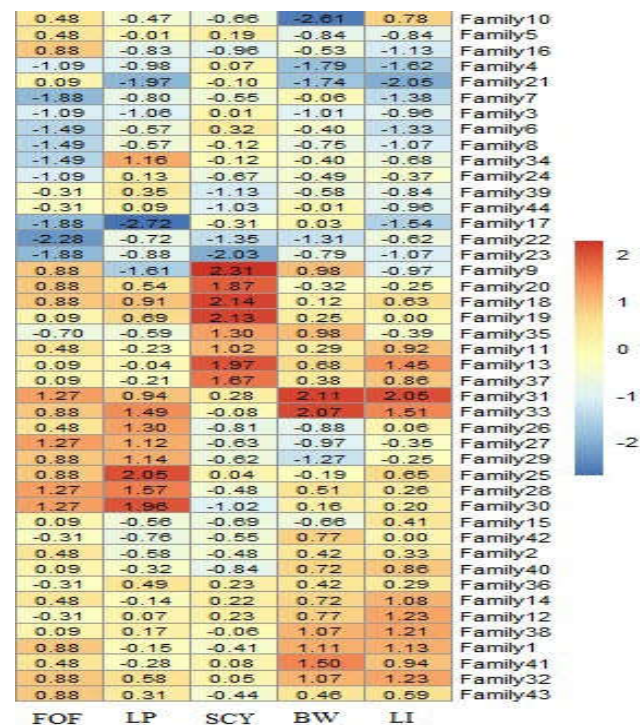


Figure 3. Heatmap for the relationship between families and the studied traits based on scaled data.

- First opening flower (F.O.F, boll weight (B.W.), seed cotton yield (S.C.Y.), lint Percentage (L.%) and lint index (L.I.).

#### 4. Discussion

The results cleared that the genetic variability within F<sub>2</sub> plants was more significant than that within F<sub>3</sub> and F<sub>4</sub> progenies, which may be due to the presence of possible considerable dominant gene action expressed in the F<sub>2</sub> population, which is confirmed with those finding previous by [1, 33, 34].

The results revealed that the mean values for the traits were improving from generation to generation; this trend could be due to the additive gene action. Also, mean values of heritability in the broad sense were high; it may be due to the close relationship between PCV and GCV values. The values of genotypic and phenotypic coefficients were narrow for traits that indicated less environmental influence in the expression of these traits. Similar results were obtained by [6, 35-37].

The high PCV and GCV and high heritability values are good indicators for a genetic advance through the amount of genetic variance expected from the selection. Therefore, selection based on the first opening flower was more effective for improving this trait and selecting the best genotypes with the earliest first opening flower date. The results showed that direct selection for lint percentage would be favorable for breeding programs for this trait, these results agree with those obtained by [2, 38-43].

The results showed that when used as a selection criterion, the lint index was more effective for improving boll weight and seed cotton yield per plant and selecting the best genotypes with the highest lint index value. These results indicated using the lint index selection to improve boll weight and seed cotton yield. Similar results were previously reported [33, 36, 44-46]

The differences between generations may be due to the various genotypes scored by each selection criteria. The analysis of variance showed highly significant differences among families of F<sub>4</sub> for all the studied characters meaning that selection criteria would be effective. The genetic advance for the selected families was different and more than better parent and commercial variety. Also, the results indicated that mean performance for most selected families differed; these results could be due to gene expression for these families; therefore, selection for adaptable families and higher than means for better parent and commercial variety well be a help to produce high yielding recombination and earlier to grow it often winter cropping without a decrease in yield. These results are in harmony with those previously reported [33, 46-50]. It appeared that most of the yield and its component traits were positive and either significant or highly significant correlated with each other, indicating that these traits are essential components of cotton yield. A positive association between major yield components is very substantial to the breeder because component breeding would be very effective under such a situation. Such correlations indicate that improving one or more of these traits would enhance the other.

**Author Contributions:** Conceptualization, M.S.A.A., F.A.S., H.M.K.K., O.M.I., A.A.F., and A.M.E.T.; methodology, M.S.A.A., F.A.S, and H.M.K.K.; formal analysis, M.S.A.A., F.A.S, H.M.K.K., O.M.I, and A.M.E.T.; investigation, M.S.A.A., F.A.S, H.M.K.K., O.M.I., and A.M.E.T ; data curation, M.S.A.A., F.A.S, H.M.K.K., O.M.I., and A.M.E.T; writing—original draft preparation, M.S.A.A., F.A.S, H.M.K.K., O.M.I, and A.M.E.T ; writing—review and editing M.S.A.A., F.A.S, H.M.K.K., O.M.I, A.A.F., and A.M.E.T. All authors have read and agreed to the published version of the manuscript.

**Conflict of interest:** No conflict of interest exists in the submission of this manuscript. The authors declare that the work described has not been published previously, and is not under consideration for publication elsewhere, in whole or in part.

**Consent for publication:** The MANUSCRIPT has been approved for publication by all authors. All the authors listed have approved the manuscript enclosed herewith

**Data Availability Statement:** The data that support the findings of this study are available on the manuscript.

## References

1. Chiurciu, I.A., et al., *Cotton Market Trends in the World and in the European Union*. 2016.
2. Montes, E., et al., *Assessment of Gene Flow Between Gossypium hirsutum and G. herbaceum: Evidence of Unreduced Gametes in the Diploid Progenitor*. G3 (Bethesda, Md.), 2017. 7(7): p. 2185-2193.
3. Abo Sen, E.Z.F., et al., *Evaluation of genetic behavior of some Egyptian Cotton genotypes for tolerance to water stress conditions*. Saudi Journal of Biological Sciences, 2022. 29(3): p. 1611-1617.
4. Aripov, K. and M. Ioelovich, *COMPARATIVE STUDY OF SUPRAMOLECULAR STRUCTURE OF CELLULOSE IN COTTON FIBERS OF GOSSYPIUM HIRSUTUM AND GOSSYPIUM BARBADENSE*. Cellulose Chemistry and Technology, 2020. 54: p. 635-641.
5. Brown, J. and Z. Khan, *Breeding Cotton for Cotton Leaf Curl Disease Resistance*. 2022. p. 171-197.
6. Hu, Y., et al., *Gossypium barbadense and Gossypium hirsutum genomes provide insights into the origin and evolution of allotetraploid cotton*. Nature Genetics, 2019. 51(4): p. 739-748.
7. Hoffmann, L., et al., *Genetic diversity of Gossypium barbadense from the central Brazilian Amazon*. Acta Amazonica, 2018. 48: p. 1-9.
8. Hu, G., et al., *Evolution and Diversity of the Cotton Genome*, in *Cotton Precision Breeding*. 2021, Springer. p. 25-78.
9. Iqbal, M., et al., *Cotton Genetics and Genomics*. 2022. p. 15-27.
10. Mahmood, T., et al., *Insights into drought stress signaling in plants and the molecular genetic basis of cotton drought tolerance*. Cells, 2020. 9(1): p. 105.
11. Abdellatif, K., et al., *Molecular diversity of Egyptian cotton (Gossypium barbadense L.) and its relation to varietal development*. Journal of Crop Science and Biotechnology, 2012. 15.
12. Abou-Elwafa, S., *Efficiency of Single and Multiple Traits Selection for Yield and Its Components in Varietal Maintenance of Giza 90 Egyptian Cotton Variety*. World Journal of Agricultural Research, 2016. 4: p. 166-172.
13. Areej, J., et al., *Drought Stress Modified Genetic Components and Combining Ability of Cotton Genotypes*. 2021.
14. Miller, P.A. and J.O. Rawlings, *Selection for Increased Lint Yield and Correlated Responses in Upland Cotton, Gossypium hirsutum L.1*. Crop Science, 1967. 7(6): p. crops1967.0011183X000700060024x.
15. Snedecor, G.W. and W.G. Cochran, *Statistical Methods*. 7th Edition, Iowa State University Press, Ames., 1980.
16. Burton, G.W., *Quantitative inheritance in grasses*. . Proceedings of 6th International Grassland Congress, 1, 277-283., 1952.
17. Singh, R.K. and B.D. Chaudhary, *Biometrical methods in quantitative genetic analysis*. 1979, New Delhi, India: Kalyani Publishers.
18. Burton, G.W. and E.H. DeVane, *Estimating Heritability in Tall Fescue (Festuca Arundinacea) from Replicated Clonal Material1*. Agronomy Journal, 1953. 45(10): p. 478-481.
19. Ahmad, I., et al., *Genetically Modified Cotton*. 2022. p. 271-289.
20. Qiu, T., et al., *GbAt11 gene cloned from Gossypium barbadense mediates resistance to Verticillium wilt in Gossypium hirsutum*. Journal of Cotton Research, 2020. 3.
21. Imtiyazahmed, S., et al., *Development and hybridity confirmation of F1 interspecific hybrids between Gossypium barbadense and Gossypium anomalum*. 2020.
22. Ahuja, S.L., L.S. Dhayal, and R. Prakash, *A Correlation and Path Coefficient Analysis of Components in G. hirsutum L. Hybrids by Usual and Fibre Quality Grouping*. Turkish Journal of Agriculture and Forestry, 2006. 30: p. 317-324.
23. Azeem, I. and S. Azhar, *Estimates of Heritability and pattern of association among different characters of Gossypium hirsutum L*. Pak. J. Agri. Sa, 2006. 37(1-2).
24. Chakravarthy, A., M. Naik, and M. T N, *9 Arthropods on Cotton: A Comparison Between Bt and Non-Bt Cotton*. 2022.
25. Chaturvedi, G., A. Singh, and R. Bahadur, *Screening techniques for evaluating crop germplasm for drought tolerance*. Plant Archives, 2012. 12(1): p. 11-18.
26. Dube, S., *Gossypium barbadense (Gallini cotton)*. 2020.
27. Cormack, R.M., *A Review of Classification*. Journal of the Royal Statistical Society. Series A (General), 1971. 134(3): p. 321-367.
28. Bezdek, J., *Cluster Validity with Fuzzy Sets*. Journal of Cybernetics, 1973. 3.
29. Bezdek, J., *Pattern Recognition With Fuzzy Objective Function Algorithms*. 1981.
30. Desgraupes, B., *clusterCrit*. . 2018.
31. Everitt, B., *Cluster analysis*. 1993, London; New York: E. Arnold ; Halsted Press.
32. Halkidi, M. and M. Vazirgiannis, *Clustering Validity Assessment: Finding the optimal partitioning of a data set*. 2001. 187-194.
33. Chattha, W.S., et al., *A novel parent selection strategy for the development of drought-tolerant cotton cultivars*. Plant Genetic Resources, 2021. 19(4): p. 299-307.
34. Çiçek, S., et al., *Maintenance Breeding On Some Cotton Varieties (Gossypium hirsutum L.)*. 2015.
35. Ganesan, S., *Selection parameters for improving the seed cotton yield and fibre quality traits in American Cotton (Gossypium hirsutum L.)*. Electronic Journal of Plant Breeding, 2010. 1.
36. Gonzales, H., J. Cavero, and L. Soto, *Gossypium barbadense: revaloración y sustentabilidad*. Cuadernos del Centro de Estudios de Diseño y Comunicación, 2021.
37. Iqbal, M., et al., *Correlation and Path Coefficient Analysis for Earliness and Yield Traits in Cotton ( G. hirsutum L . )*. Asian Journal of Plant Sciences, 2006. 5.
38. López, E., et al., *In vitro propagation of Gossypium barbadense L. "native cotton" of brown fiber*. Agroindustrial Science, 2020. 10: p. 235-239.
39. Mehdi, M., M. Raza, and M. Azeem, *Cotton Breeding and Entrepreneurship*. 2022. p. 311-321.

40. Milligan, G.W. and M.C. Cooper, *An examination of procedures for determining the number of clusters in a data set*. Psychometrika, 1985. **50**(2): p. 159-179.
41. Nazmey, M., A. Eissa, and Y. W.M.B, *GENETICAL ANALYSIS FOR F1 AND F2 GENERATION IN SOME EGYPTIAN COTTON CROSSES (Gossypium barbadense L.)*. Journal of Agricultural Chemistry and Biotechnology, 2010. **1**: p. 341-350.
42. Nidagundi, J., et al., *Combining ability studies in cotton (Gossypium barbadense L.)*. 2018. **Vol. 7**.
43. Punia, S., R.s. Sangwan, and S. Nimbal, *Studies on Association, Path Analysis and Genetic Parameters for Seed Cotton Yield and It's Contributing Characters in Desi Cotton (Gossypium arboreum L.)*. International Journal of Current Microbiology and Applied Sciences, 2017. **6**: p. 104-111.
44. Dhivya, R., et al., *Variability, heritability and genetic advance in upland cotton (Gossypium hirsutum L.)*. African Journal of Plant Science, 2014. **8**: p. 1-5.
45. Ergashovich, K.A., et al., *Effects of microelements on drought resistance of cotton plant*. International Journal of Psychosocial Rehabilitation, 2020. **24**(2): p. 643-648.
46. Wang, P., et al., *Introgression from Gossypium hirsutum is a driver for population divergence and genetic diversity in Gossypium barbadense*. The Plant Journal, 2022.
47. Yang, J., et al., *Gossypium barbadense RVd mRNA, complete cds*. 2015.
49. Zhong, W., et al., *Development and Characterization of SSR Markers in the Gossypium Barbadense Genome*. 2021.
50. Zhu, Q.-H., et al., *Genetic dissection of the fuzzless seed trait in Gossypium barbadense*. Journal of experimental botany, 2018. **69**.
51. M-stat c (1986). A microcomputer program of design management and analysis of agronomic research experiments. Michigan State., Univ. USA.