


## Correlation and multivariate analysis across non-segregation and segregation generations in two cotton crosses

Waleed Mohamed Bassuny Yehia<sup>1</sup> and Essam Fathy El-Hashash<sup>2\*</sup> 

**Address:**

<sup>1</sup> Cotton Breeding Department, Cotton Research Institute, Agriculture Research Center, Egypt

<sup>2</sup> Department of Agronomy, Faculty of Agriculture, Al-Azhar University, Cairo, Egypt.

\*Corresponding author: **Essam F. El-Hashash**, [dressamelhashash@yahoo.com](mailto:dressamelhashash@yahoo.com)

Received: 20-06-2021; Accepted: 20-11-2021; Published: 27-11-2021

[10.21608/ejar.2021.81571.1117](https://doi.org/10.21608/ejar.2021.81571.1117)

### ABSTRACT

The present research uses Pearson's correlation coefficient and multivariate analysis to assess the interrelationships, similarities, and dissimilarities among non-segregation (P1, P2 and F1) and segregation (F2, BC1 and BC2) generations for seed cotton yield and yield components in the two crosses, Giza 92 x Pima S6 and Giza 93 x C.B. 58. For all variables evaluated in the two crossings, the analysis of variance revealed significantly substantial genetic variability across six generations. For all of the tested features in the two crossings, the F1 performed better than the other generations. Across all six generations of the two crossings, there were positive and highly significant correlations between seed cotton yield/plant, lint cotton yield/plant, and no. of bolls/plant characteristics. Seed cotton yield and yield components in the two crosses showed some favourable connections over the six generations. The UPGMA hierarchical clustering revealed a greater degree of similarity coefficients between the six generations and the attributes investigated. The similarity coefficients for the six generations and the examined features, respectively, ranged from 0.96 to 0.99 and 0.65 to 0.96. In the principal component analysis (PCA), the PCA1 extracted had an Eigenvalue of greater than 1 across six generations for all studied traits in the two crosses. The PCA displayed a total variation of 91.84% among the six generations contributed by PCA1 (79.47%) and PCA2 (12.38%) and mainly distinguished the generations into different groups. The PCA1 and PCA2 were dominated by F1 and segregation generations in the two crosses, respectively, showing high correlations with the first two PCAs. All studied traits, as well as boll weight and lint percentage traits, contributed to positive significant component loadings for the PCA1 and PCA2, respectively. The biplot analysis of the relationship between the six generations revealed that the most appropriate generations for selecting yield traits were F1 in the two crosses and BC1 and BC2 in the cross Giza 93 x C.B. 58. As a result, we recommend considering backcrossing for 2–5 cycles (BC2–BC5) at the C.B. 58 parent in the future to improve Egyptian cotton yield.

**Keywords:** Generations, Correlation, Multivariate Analysis, Cotton

### INTRODUCTION

The cotton crop is one of the most popular and most important natural fibers for textile manufacture in the old and new worlds (Akter et al., 2019). It is a flowering plant belonging to the family Malvaceae and the genus *Gossypium* (G). The G. genus is comprised of ~50 species. The most important cultivated species of the genus are *G. hirsutum*, *G. barbadense* (allotertraploid 2n = 52), *G. herbaceum*, and *G. arboreum* (diploid 2n = 26) (Rehman and Farooq, 2020). The annual world area harvested, yield and production of cotton were approximately 38640608 ha, 21374 hg/ha, and 82589031 tonnes in 2019, respectively. As for Egypt, the total area harvested, yield, and production during the 2019 growing season were 100000 ha, 30500 hg/ha, and 20500 tons, respectively (FAOSTAT, 2021). According to FAS/Cairo (Post), cotton area harvested in the market year 2021/22 will increase by 7% to 70,000 hectares (ha), up from 65,000 ha in the market year 2020/21. Post estimates production for the market year 2021/22 at 250,000 bales, compared to 215,000 bales in the market year 2020/21 (USDA, 2021).

The main goal of cotton breeders is looking and select the genotypes with high yield traits and take advantage of the permanent untapped genetic variation of Egyptian varieties, substantial work has been carried out to develop both yield and quality traits of cotton in Egypt (Yehia and El-Hashash, 2019). Through the process of breeding, progress in yield potential is produced by the gradual accumulation of genes that lead to higher yields or the elimination of unfavorable genes (Abdel-Monaem et al., 2020). The increase in yield can be possible if the existing genetic resources and information are properly utilized (Srou and El-Hashash, 2012).

A successful breeding program relies on comprehensive knowledge and understanding of relationships among yield traits and genetic diversity within and between the elite genetic materials of existing germplasm (Akter et al., 2019). The Pearson's correlation coefficient allows determining the magnitude of the relationships between essential and important traits (Abasianyanga et al., 2017). Thus, the correlation between traits can be useful in developing selection criteria for improving cotton yield, but correlation can also represent a swamp of interrelationships (Kloth, 1998). Cotton breeders study the associations between the yield and its components traits, because the selection of one trait may lead to improvement or deterioration of associated traits, depending on nature and size (Nandhini et al., 2018). Several cotton researchers have reported a seed cotton yield/plant was related positively significantly with lint cotton yield/plant, No. of bolls/plant, boll weight, lint percentage and seed index traits (for example Kerby et al., 1990; El-Hashash 2013; Latif et al., 2015; Jarwar et al., 2019; Yehia and El-Hashash 2019; Eldessouky et al., 2021).

Multivariate analysis such as principal component and cluster analysis are mostly used to evaluate the magnitude of genetic diversity among the germplasm. The principal component analysis (PCA) explains the contribution of the most

important traits that account for the total genetic variability. Cluster analysis allows the classification of individuals into groups in order to maximize their homogeneity within a group and show heterogeneity among groups. It is complementary to PCA (El-Hashash, 2016). Many researchers have used the PCA and cluster analysis to assess the relationship and diversity between several cotton germplasm, in addition to knowing the relationships between seed cotton yield and its components traits (Shah *et al.*, 2018; Akter *et al.*, 2019; Jarwar *et al.*, 2019; Muminov *et al.*, 2020; Abdel-Monaem *et al.*, 2020; Sarwar *et al.*, 2021). Scree plot is usually used for visual assessment of factor, which explains high amount of the changes in the data (Jarwar *et al.*, 2019), also it further highlighted the partitioning of the principal components (Abasianyanga *et al.*, 2017).

Due to the increasing demand for Egyptian cotton, this present investigation was conducted to study the interrelationships, similarities, and dissimilarities among non-segregation and segregation generations for seed cotton yield and yield attributes in the two crosses by adopting Pearson's correlation coefficient and multivariate analysis.

## MATERIAL AND METHODS

### Genetic material and field procedure:

Four genotypes were used in this study, namely Giza 92, Giza 93 (Egyptian varieties), Pima S6 (Egyptian American) and C.B. 58 (USA barbadense), which all belong to the species *Gossypium barbadense* L. The experiments were carried out during the three successive growing seasons from 2018 to 2020 at Sakha Agricultural Research Station, Kafr El-Shiekh Governorate, Egypt. In 2018 season, the four parental varieties were crossed to produce  $F_1$  hybrid seeds for the two crosses Giza 92 x Pima S6 and Giza 93 x C.B. 58. In 2019 season, each  $F_1$  was backcrossed to both parents to obtain  $BC_1$  and  $BC_2$ ; the parents were also crossed for more hybrid seeds and the  $F_1$  plants were selfed to obtain  $F_2$  seeds. The six populations *i.e.*,  $P_1$ ,  $P_2$ ,  $F_1$ ,  $F_2$ ,  $BC_1$  and  $BC_2$  for the two crosses were evaluated separately in a randomized complete blocks design with three replications during 2020 season. Each replicate consisted of 10 rows for  $F_2$ , 5 rows for  $BC_1$  and  $BC_2$  crosses (segregating generations), and 3 rows for each non-segregating generation  $P_1$ ,  $P_2$  and  $F_1$ . Each row was 4 meters in length and 0.60 m in width as well as comprised 10 hills. Hills were spaced at 40 cm apart and thinned to one plant per hill. All recommended agronomic practices of cotton were applied from sowing to harvesting to get a good and healthy plant population as usual.

### Traits measurement:

The studied traits on a ten individual guarded plant basis were taken of the six non-segregation and segregation generations in the two crosses studied. The data were recorded for boll weight in grams (B.W., g), seed cotton yield/plant in grams (S.C.Y./P, g), lint cotton yield/plant in grams (L.C.Y./P, g), lint percentage (L.%), number of bolls/plant (No. of B./P) and seed index (S.I., g) traits. The collected data were statistically analyzed.

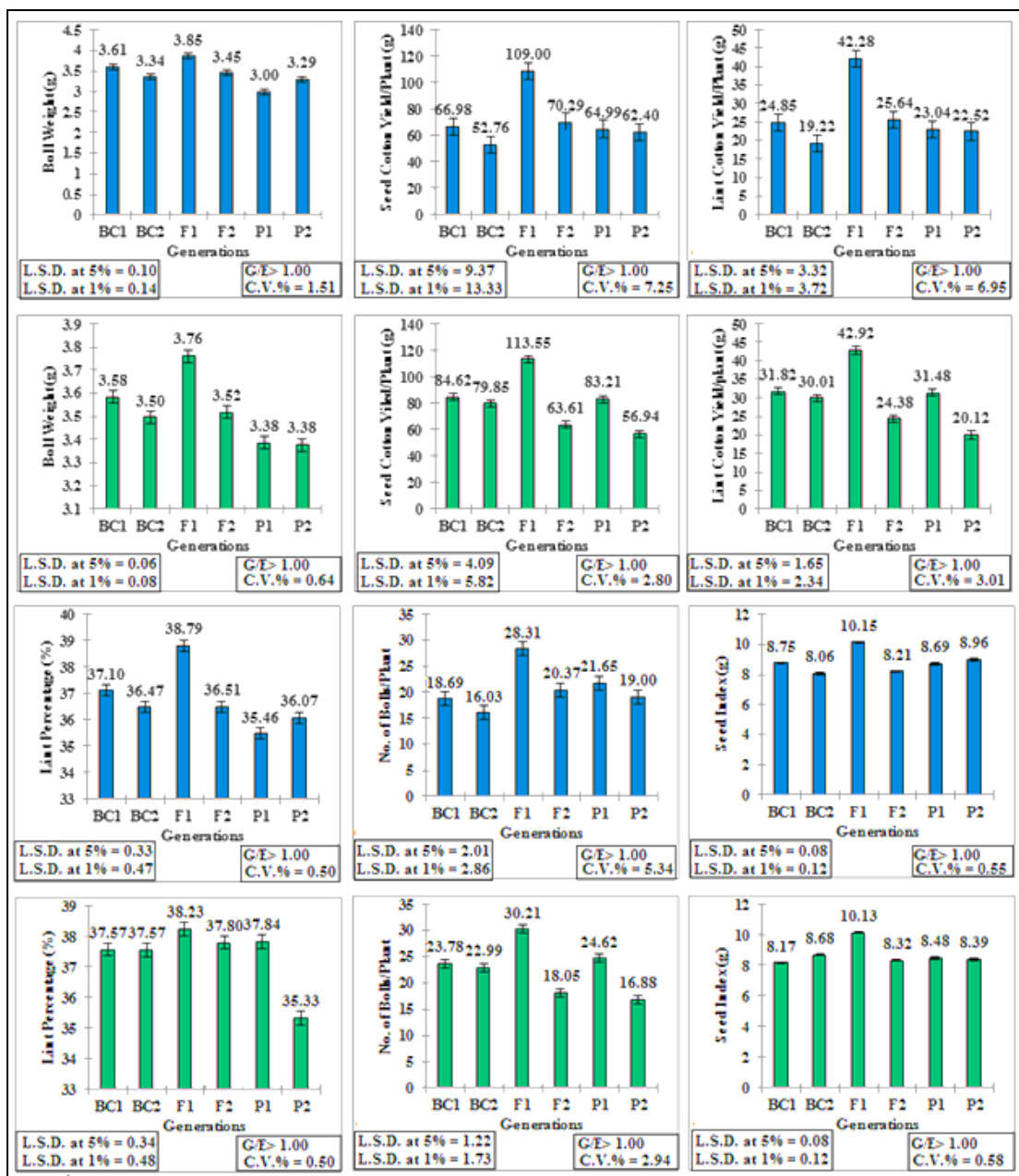
### Statistical approaches:

The data of studied traits were subjected to a one-way ANOVA test following the method of Steel and Torrie (1997) to determine the significant differences and the coefficient of variation (CV%) among six generations using XLSTAT software as described by Addinsoft (2021). According to Gomes (2009), the estimates of CV% were classified as very high ( $CV \geq 21.0\%$ ), high ( $15.0\% \leq CV \leq 21.0\%$ ), moderate ( $10\% < CV \leq 20\%$ ) and low ( $CV < 10\%$ ). The significance test was done with the least significant difference test (L.S.D) at 0.05 and 0.01 levels of probability according to Steel and Torrie (1997). Quantity and plot Pearson's correlation coefficient, as well as multivariate analysis (principal component and cluster analysis) were performed for a better understanding of the relationship among studied traits across six generations using the computer software program PAST version 4.03 (Hammer *et al.*, 2020).

## RESULTS

### ANOVA and mean performances:

The results of one-way ANOVA showed statistically significant differences ( $P < 0.01$ ) among non-segregation ( $P_1$ ,  $P_2$  and  $F_1$ ) and segregation ( $F_2$ ,  $BC_1$  and  $BC_2$ ) generations for yield and yield component traits studied in the two crosses Giza 92 x Pima S6 and Giza 93 x C.B. 58 (Fig. 1). The values of generations/error variance ratio were greater than unity (G/E ratio  $> 1$ ) for all traits studied in the two crosses. According to classification by Gomes (2009), the values of CV% recorded for studied traits across the six generations in the two crosses were low ( $CV < 10\%$ ). The lint percentage and seed index traits recorded the lowest values of CV % in the two crosses (Fig. 1). The variety Giza 93 exhibited the best performance for all traits studied compared with the other three varieties, except for the seed index trait (variety C.B. 58). The  $F_1$  mean performance was determined to have a high value compared with the assessed respective parents and three segregation generations for all the studied traits in the two crosses. When comparing within the segregation generations, the  $BC_1$  generation showed superiority in most studied traits in the two crosses (Fig. 1).



**Fig. 1.** The values of means and LSD for studied traits during the non-segregation and segregation generations in the two hybrids G. 92 x Pima S6 (blue columns) and Giza 93 x C.B. 58 (green columns). P<sub>1</sub> and P<sub>2</sub>: First and second parents; F<sub>1</sub> and F<sub>2</sub>: First and second generations; BC<sub>1</sub> and BC<sub>2</sub>: First and second backcrosses, respectively; G/E ratio: mean squares of generations/ mean squares of error; C.V%: Coefficient of variation; LSD values denote highly significant differences between the six generations were analyzed by ANOVA test.

**Correlation analysis:**

The Pearson's correlation coefficient of cotton yield and its components are separated for an evident understanding of relationships between these traits in each generation as presented in **Table (1)**. Out of 180 Pearson correlations between studied traits, 56 (ranged from 6 in P<sub>1</sub> to 13 in BC<sub>1</sub>) and 57 (ranged from 7 in P<sub>1</sub> to 13 in F<sub>2</sub>) positive correlation coefficients were seen within the six generations in the two crosses Giza 92 x Pima S6 and Giza 93 x C.B. 58, respectively. Seed cotton yield/plant, lint cotton yield/plant and No. of bolls/plant traits had strong positive and highly significant correlations ( $p < 0.01$ ) across all six generations in the two crosses. Boll weight was significantly and positively correlated ( $p < 0.01$ ) with seed and lint cotton yields/plant in F<sub>2</sub> generation during the two crosses studied, and with lint percentage trait in BC<sub>1</sub> generation under the cross Giza 93 x C.B. 58. The seed index shows a positive and significant correlation with seed cotton yield/plant, lint cotton yield/plant and lint percentage traits in BC<sub>1</sub> generation across the cross Giza 92 x Pima S6. Lint percentage displayed a positive association with seed index in P<sub>2</sub> and BC<sub>1</sub> generations ( $p < 0.05$ ), as well as with lint cotton yield/plant in BC<sub>1</sub>, BC<sub>2</sub> ( $p < 0.05$ ), P<sub>2</sub>, F<sub>1</sub> and F<sub>2</sub> ( $p < 0.01$ ) generations during the cross Giza 93 x C.B.58.

**Table 1.** Pearson correlation coefficient among studied traits in the four cotton cultivars and their F<sub>1</sub>, F<sub>2</sub> and backcross generations.

Generations		Giza 92 x Pima S6						Giza 93 x C. B. 58					
		P <sub>1</sub>	P <sub>2</sub>	F <sub>1</sub>	F <sub>2</sub>	BC <sub>1</sub>	BC <sub>2</sub>	P <sub>1</sub>	P <sub>2</sub>	F <sub>1</sub>	F <sub>2</sub>	BC <sub>1</sub>	BC <sub>2</sub>
B.W.	S.C.Y./P	0.22	0.01	-0.13	0.28**	0.09	-0.02	0.11	-0.11	-0.31	0.38**	0.02	-0.04
	L.C.Y./P	0.22	0.01	-0.23	0.28**	0.12	0.02	0.07	-0.17	-0.29	0.38**	0.06	-0.11
	L.%	-0.13	-0.07	-0.12	-0.04	0.20	0.17	-0.23	-0.15	-0.10	0.11	0.35**	-0.28*
	No.B./P	-0.01	-0.11	-0.74**	0.01	-0.33**	-0.68**	-0.52**	-0.48**	-0.74**	0.04	-0.35**	-0.43**
	S.I.	0.27	-0.23	-0.06	0.07	0.11	0.09	-0.02	-0.20	0.14	-0.05	-0.07	-0.01
S.C.Y./P	L.C.Y./P	0.99**	1.00**	0.84**	0.98**	0.99**	0.97**	0.98**	0.81**	0.91**	0.99**	0.99**	0.98**
	L.%	-0.46*	0.23	-0.56**	-0.02	0.01	-0.25*	-0.10	0.04	0.28	0.11	0.11	0.06
	No.B./P	0.97**	0.99**	0.76**	0.95**	0.91**	0.73**	0.79**	0.93**	0.86**	0.93**	0.93**	0.92**
	S.I.	-0.19	0.21	0.02	0.04	0.23*	-0.46**	-0.04	-0.19	0.33	-0.02	-0.10	0.04
L.C.Y./P	L.%	-0.36	0.31	-0.03	0.14	0.18	-0.01	0.11	0.61**	0.64**	0.26**	0.27*	0.25*
	No.B./P	0.97**	0.99**	0.73**	0.93**	0.89**	0.69**	0.80**	0.78**	0.79**	0.92**	0.90**	0.93**
	S.I.	-0.20	0.22	0.14	0.02	0.27*	-0.44**	-0.12	0.09	0.32	0.01	-0.04	-0.05
L.%	No.B./P	-0.44*	0.23	-0.30	-0.04	-0.05	-0.27*	0.06	0.09	0.25	0.10	0.01	0.19
	S.I.	-0.01	0.26	0.17	-0.08	0.25*	0.17	-0.39**	0.42*	0.15	0.10	0.25*	-0.45**
No.B./P	S.I.	-0.25	0.22	0.06	0.01	0.17	-0.42**	-0.02	-0.09	0.16	0.01	-0.06	0.01

B.W.: Boll weight; S.C.Y./P: Seed cotton yield/plant; L.C.Y./P: Lint cotton yield/plant; L.%: Lint percentage; No.B./P: Number of bolls/plant; S.I.: Seed index; P<sub>1</sub> and P<sub>2</sub>: First and second parents; F<sub>1</sub> and F<sub>2</sub>: First and second generations; BC<sub>1</sub> and BC<sub>2</sub>: First and second back crosses, respectively. According to the Pearson correlation coefficient, the asterisks (\* and \*\*) denote a significant difference at 0.05 and 0.01 level of significance, respectively.

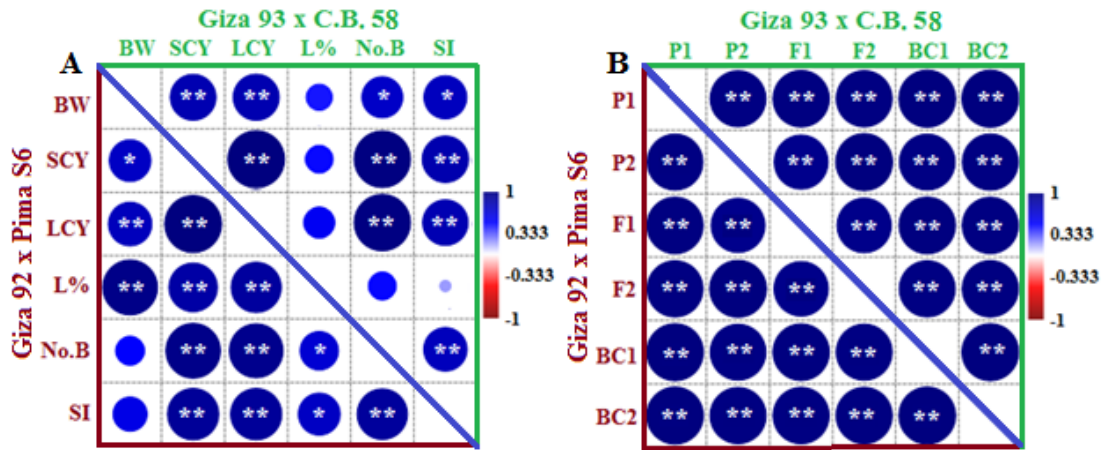
The intergeneration correlations for yield and its components traits are shown in **Table (2)**. Several positive correlations were observed across both of them generations comparisons for studied traits in the two crosses Giza 92 x Pima S6 (54) and Giza 93 x C.B. 58 (36). The P<sub>1</sub> generation positively correlated with P<sub>2</sub> generation for lint percentage and seed index traits (p < 0.05 and 0.01) in the cross Giza 93 x C.B. 58 and Giza 92 x Pima S6, respectively, with F<sub>2</sub> generation for lint cotton yield/plant (p < 0.05) in the cross Giza 92 x Pima S6 and with BC<sub>2</sub> generation for No. of bolls/plant (p < 0.01) in the cross Giza 93 x C.B. 58. The P<sub>2</sub> generation had a positive and significant correlation with F<sub>1</sub> generation for No. of bolls/plant (p < 0.01), with BC<sub>1</sub> (p < 0.01) and BC<sub>2</sub> (p < 0.05) generations for lint percentage and with BC<sub>2</sub> generation for boll weight (p < 0.05) in the cross Giza 92 x Pima S6; as well as with F<sub>2</sub> and BC<sub>1</sub> generations for boll weight (p < 0.05) and seed index (p < 0.01) traits in the cross Giza 93 x C.B. 58, respectively. The F<sub>1</sub> generation had a significant positive association with F<sub>2</sub> generation for lint percentage (p < 0.05) and with BC<sub>1</sub> generation for lint cotton yield/plant (p < 0.01), No. of bolls/plant and seed index (p < 0.05) traits across the cross Giza 92 x Pima S6; while with F<sub>2</sub> generation for seed index (p < 0.01), with BC<sub>1</sub> for boll weight (p < 0.05) and with BC<sub>2</sub> for seed cotton yield/plant (p < 0.01) under the cross Giza 93 x C.B. 58. The BC<sub>1</sub> generation showed a significant and positive correlation with BC<sub>2</sub> generation for lint percentage (p < 0.01) in the two crosses and for lint cotton yield/plant (p < 0.05) traits in the cross Giza 92 x Pima S6.

**Table 2.** Pearson correlation coefficient between the non-segregation and segregation generations for studied traits in the two hybrids of cotton.

Yield and yield components traits												
Crosses	B.W.		S.C.Y./P		L.C.Y./P		L.%		No.B./P		S.I.	
	I	II	I	II	I	II	I	II	I	II	I	II
r(P <sub>1</sub> , P <sub>2</sub> )	0.06	0.16	0.23	-0.52**	0.25	-0.17	-0.25	0.37*	0.30	0.04	0.46**	-0.67**
r(P <sub>1</sub> , F <sub>1</sub> )	-0.50**	-0.54**	-0.12	-0.11	-0.10	-0.03	-0.11	-0.18	0.32	-0.24	0.33	-0.02
r(P <sub>1</sub> , F <sub>2</sub> )	0.25	0.15	0.30	0.19	0.36*	0.02	-0.45*	-0.45*	0.24	0.09	-0.62**	-0.03
r(P <sub>1</sub> , BC <sub>1</sub> )	0.34	-0.06	-0.28	0.04	-0.10	-0.02	0.02	0.21	-0.30	-0.54**	0.14	-0.24
r(P <sub>1</sub> , BC <sub>2</sub> )	-0.06	-0.05	0.11	0.11	0.24	0.27	0.13	0.17	0.01	0.70**	-0.04	-0.17
r(P <sub>2</sub> , F <sub>1</sub> )	-0.49**	-0.20	0.33	-0.58**	0.08	-0.73**	-0.16	-0.65**	0.47**	-0.53**	0.08	-0.30
r(P <sub>2</sub> , F <sub>2</sub> )	-0.12	0.44*	-0.53**	-0.36*	-0.48**	-0.36*	-0.30	-0.55**	-0.59**	-0.34	-0.25	-0.10
r(P <sub>2</sub> , BC <sub>1</sub> )	-0.37*	0.11	-0.11	0.14	-0.06	0.01	0.59**	0.19	0.06	-0.02	-0.01	0.54**
r(P <sub>2</sub> , BC <sub>2</sub> )	0.40*	-0.29	0.27	-0.29	0.30	-0.06	0.42*	0.28	0.32	-0.24	-0.13	-0.25
r(F <sub>1</sub> , F <sub>2</sub> )	0.35	0.13	-0.33	-0.09	-0.28	-0.05	0.45*	0.24	-0.19	-0.21	0.15	0.53**
r(F <sub>1</sub> , BC <sub>1</sub> )	0.06	0.38*	0.34	-0.51**	0.49**	-0.34	0.01	-0.55**	0.41*	-0.17	0.37**	0.32
r(F <sub>1</sub> , BC <sub>2</sub> )	0.09	0.22	0.04	0.52**	0.06	0.15	0.09	-0.83**	0.28	0.22	0.26	-0.43*
r(F <sub>2</sub> , BC <sub>1</sub> )	0.23	0.26	-0.14	-0.06	-0.18	-0.05	-0.30	-0.04	-0.08	-0.09	-0.36*	0.10
r(F <sub>2</sub> , BC <sub>2</sub> )	0.17	-0.08	-0.29	0.08	-0.29	0.12	0.05	0.12	-0.38*	0.17	0.32	-0.55**
r(BC <sub>1</sub> , BC <sub>2</sub> )	0.12	-0.14	0.34	-0.90	0.42*	-0.83**	0.73**	0.68**	0.15	-0.70**	0.25	-0.51**

B.W.: Boll weight; S.C.Y./P: Seed cotton yield/plant; L.C.Y./P: Lint cotton yield/plant; L.%: Lint percentage; No.B./P: Number of bolls/plant; S.I.: Seed index; I: G. 92 x Pima S6; II: Giza 93 x C. B. 58. P<sub>1</sub> and P<sub>2</sub>: First and second parents; F<sub>1</sub> and F<sub>2</sub>: First and second generations; BC<sub>1</sub> and BC<sub>2</sub>: First and second back crosses, respectively. According to the Pearson correlation coefficient, the asterisks (\* and \*\*) denote a significant difference at 0.05 and 0.01 level of significance, respectively.

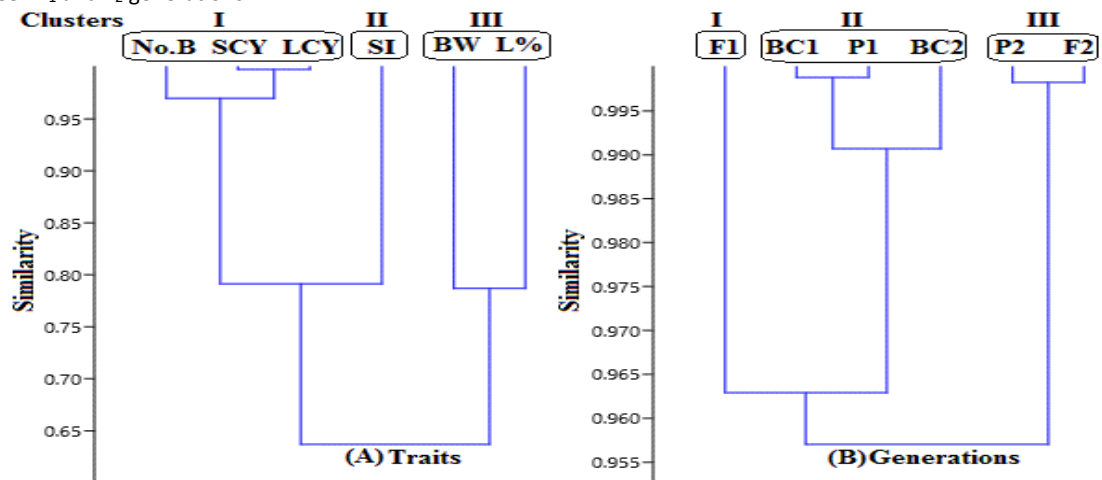
The results in Fig. (2A) cleared that 13 and 10 Pearson’s correlation coefficients exhibited significantly ( $p < 0.05$  or  $0.01$ ) among the yield and its components traits across six generations in the two crosses Giza 92 x Pima S6 and Giza 93 x C.B. 58, respectively. Positive and significant correlations ( $p < 0.05$  or  $0.01$ ) between all possible pairs for investigated traits were found during the cross Giza 92 x Pima S6, except boll weight had a positive and insignificant correlation with the number of bolls/plant and seed index traits. In the cross, Giza 93 x C.B. 58, except for the correlations that comprise the lint percentage trait, all correlations between yield and its components traits were positive and significant ( $p < 0.05$  or  $0.01$ ). During the two crosses, there was a high, positive and significant correlation ( $p < 0.01$ ) among all combinations of the six generations across studied traits (Fig. 2B).



**Fig. 2.** Plot describing Pearson’s correlation between yield and yield components traits (A) as well as between six generations (B) in the two hybrids of cotton. BW: Boll weight; SCY: Seed cotton yield/plant; LCY: Lint cotton yied/plant; L%: Lint percentage; No.B: Number of bolls/plant; SI: Seed index; P<sub>1</sub> and P<sub>2</sub>: First and second parents; F<sub>1</sub> and F<sub>2</sub>: First and second generations; BC<sub>1</sub> and BC<sub>2</sub>: First and second back crosses, respectively. The large and medium blue circles indicate a positive and significant (\*  $p < 0.05$ ) or highly significant (\*\*  $p < 0.01$ ), while the small blue circles indicate a positive and non-significant correlation.

**Cluster analysis:**

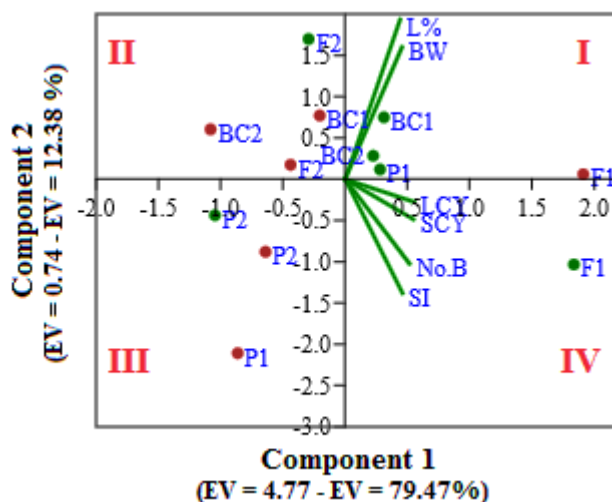
For comparison and determination of the similarities and differences between traits and generations studied on the basis of the data in the two crosses, UPGMA hierarchical clustering with correlation similarity index was used. Based on the six generations data in the two crosses (Fig. 3A), the cluster analysis separated the studied traits into three main clusters at 80% similarity. The first cluster contains No. of bolls/plant, seed cotton yield/plant and lint cotton yield/plant traits with 96% similarity. While the second (seed index) and third (lint percentage and boll weight traits) clusters comprised of the rest studied traits at 80% similarity. The maximum similarity occurred between seed and lint cotton yield/plant traits and the minimum similarity occurred between seed index and lint percentage traits. Fig. (3B) comprising of six generations could be further divided into three clusters based on the data of yield and yield component traits in the two crosses. The first cluster consists of one generation (F<sub>1</sub>) at 96% similarity. The second cluster consisted of three generations (P<sub>1</sub>, BC<sub>1</sub> and BC<sub>2</sub>) with 99% similarity. While the third cluster is comprised of two generations (P<sub>2</sub> and F<sub>2</sub>) at above 99% similarity. The greatest similarity occurred between P<sub>1</sub> and BC<sub>1</sub> generations followed by between P<sub>2</sub> and F<sub>2</sub> generations, while the lowest similarity occurred between F<sub>1</sub> and F<sub>2</sub> generations.



**Fig. 3.** Tree diagram constructed using UPGMA hierarchical clustering with correlation similarity index. BW: Boll weight; SCY: Seed cotton yield/plant; LCY: Lint cotton yied/plant; L%: Lint percentage; No.B: Number of bolls/plant; SI: Seed index; P<sub>1</sub> and P<sub>2</sub>: First and second parents; F<sub>1</sub> and F<sub>2</sub>: First and second generations; BC<sub>1</sub> and BC<sub>2</sub>: First and second back crosses, respectively.

**Principal component analysis:**

Principal component analysis (PCA) is a multivariate statistical technique. The PCA simplifies the complex data by transforming the number of correlated variables into a smaller number of variables in the data collection in a breeding and selection process called principal components (PCAs). Consequently, it has been used to estimate the similarities and dissimilarities relationships between the studied traits across six generations of variables in the two crosses of cotton, which are graphically displayed in a biplot of PCA1 and PCA2 (Fig. 4). Out of five PCAs, the first main PCA1 extracted had eigenvalues larger than one (Eigenvalue >1) with a value of 4.77 (Fig. 4). While the rest four PCAs had eigen values less than one (Eigen value < 1). The first two PCAs contributed 91.84% of the total variation existing among six generations of the studied traits in the two crosses. The contributions of PCA1 to the total variance were higher than that of the other components, with PCA1 describing only about 79.47% of the measured data total variability (Fig. 4).



**Fig. 4.** Biplot diagram among first two PCAs display contribution of studied traits in the variability of six generations across the two crosses. EV: Eigen value; EV: Explained variance; BW: Boll weight; SCY: Seed cotton yield/plant; LCY: Lint cotton yied/plant; L%: Lint percentage; No.B: Number of bolls/plant; SI: Seed index; P<sub>1</sub> and P<sub>2</sub>: First and second parents; F<sub>1</sub> and F<sub>2</sub>: First and second generations; BC<sub>1</sub> and BC<sub>2</sub>: First and second back crosses, respectively.

The five PCAs for six generations based on the studied traits during the two crosses are shown in Table (3). The PCA1 was dominated by F<sub>1</sub> generation in the two crosses and by P<sub>1</sub>, BC<sub>1</sub> and BC<sub>2</sub> generations in the cross Giza 93 x C.B. 58. On the other hand, the segregation generations influenced the PCA2 in the two crosses, which explained 12.38% of the total variability. The PCA3 was controlled by P<sub>2</sub> and F<sub>1</sub> generations across the two crosses as well as by BC<sub>1</sub> and BC<sub>2</sub> generations in the cross Giza 92 x Pima S6. As for, the PCA4 was affected by non-segregation generations in the cross Giza 92 x Pima S6 as well as by P<sub>1</sub> and F<sub>2</sub> generations in the cross Giza 93 x C.B. 58.

**Table 3.** Results of principal component analysis (PCAs) for six generations based on the studied traits during the two crosses.

PCAs Generations	Giza 92 x Pima S6					Giza 93 x C.B. 58				
	PCA1	PCA2	PCA3	PCA4	PCA5	PCA1	PCA2	PCA3	PCA4	PCA5
P <sub>1</sub>	-1.89	-1.82	-0.55	0.32	0.00	0.61	0.10	-1.05	0.20	0.00
P <sub>2</sub>	-1.40	-0.76	0.49	0.35	0.01	-2.28	-0.38	0.70	-0.53	-0.02
F <sub>1</sub>	4.18	0.05	0.49	0.28	-0.03	4.01	-0.89	0.27	-0.23	0.00
F <sub>2</sub>	-0.96	0.15	-0.16	-0.46	-0.01	-0.64	1.46	-0.02	0.59	-0.02
BC <sub>1</sub>	-0.45	0.66	0.74	-0.02	0.06	0.68	0.65	-0.75	-0.63	0.00
BC <sub>2</sub>	-2.36	0.52	0.17	0.06	-0.03	0.49	0.24	-0.32	0.06	0.03

P<sub>1</sub> and P<sub>2</sub>: First and second parents; F<sub>1</sub> and F<sub>2</sub>: First and second generations; BC<sub>1</sub> and BC<sub>2</sub>: First and second back crosses, respectively.

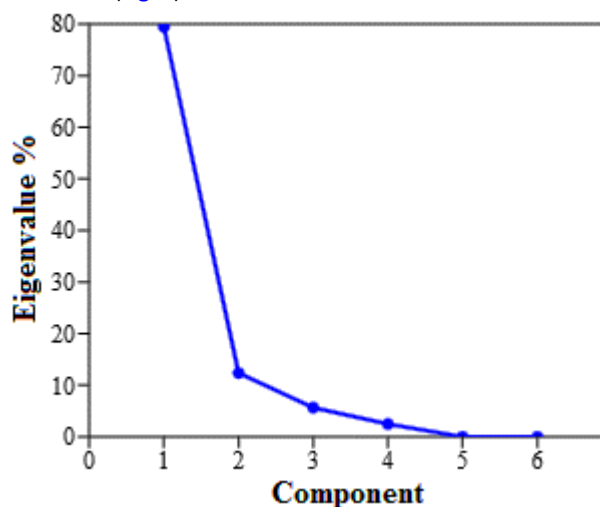
Based on the six generations of variables in the two crosses (Table 4), high positive correlations of factor loadings were recorded for seed and lint cotton yields/plant traits with PCA1 (0.45), for lint percentage with PCA2 and PCA4 (0.63), for number of bolls/plant with PCA5 (0.68), and for seed index with PCA3 (0.63). The six generations in PCA1 presented positive loading values for all studied traits. Also, all the variables of the yield and its components contributed to the PCA1. The PCA2 has identified boll weight and lint percentage traits as possessing positive loading factors and contributing to the variables. While, the boll weight and seed index traits have positive loadings on PCA3, PCA4 was dominated by lint percentage and seed index traits with positive factor loadings, while all studied traits in PCA5 obtained positive factor loadings, except seed and lint cotton yields/plant traits.

**Table 4.** The correlation coefficient between studied traits and five main PCAs generated using six generations in the two crosses.

Traits \ PCAs	PCA1	PCA2	PCA3	PCA4	PCA5
B.W.	0.37	0.52	0.53	-0.50	0.23
S.C.Y./P	0.45	-0.16	-0.17	-0.23	-0.27
L.C.Y./P	0.45	-0.09	-0.17	-0.15	-0.63
L.%	0.36	0.63	-0.28	0.63	0.05
No.B./P	0.42	-0.33	-0.43	-0.14	0.68
S.I.	0.38	-0.45	0.63	0.51	0.05

B.W.: Boll weight; S.C.Y./P: Seed cotton yield/plant; L.C.Y./P: Lint cotton yield/plant; L.%: Lint percentage; No.B./P: Number of bolls/plant; S.I.: Seed index.

The relationships between yield and its component traits across six generations in the two crosses studied are graphically displayed by the biplot of the first two PCAs (Fig. 4). The two cross performances during the six generations displayed a positive correlation among yield and its component variables, but, they differed in their degree of consistency in quantity. Positive and significant correlations ( $p < 0.05$  or  $0.01$ ) between all possible pairs of investigated traits were found across six generations in the two crosses, except between lint percentage and seed index traits, which had a positive and insignificant correlation. When comparing the six generations, the PCA1 and PCA2 showed that the yield and yield component variables were distributed in different regions and formed different groups (Fig. 4). The  $F_1$  generation in the two crosses produced the highest yield and yield components. The biplot diagram depicted the contribution of yield and its component traits in creating a variability of six generations. The  $F_1$  generations in the two crosses, as well as  $P_1$ ,  $BC_1$  and  $BC_2$  generations in the cross Giza 93 x C.B. 58, occupied Quadrants first (I) and fourth (IV) of the diagram, and which is strongly correlated with PCA1. On the other hand, the rest of the generations in the two crosses were positively associated with PCA2 and occurred in the second (II) and third (III) quarters. Lint percentage and boll weight traits were located in quarter I with  $P_1$ ,  $BC_1$  and  $BC_2$  in the cross Giza 93 x Pima S6C.B. 58 and with  $F_1$  generation Giza 92 x Pima S6. As for, the other traits, they were located near the  $F_1$  generation in quarter IV across the cross Giza 93 x C.B. 58. The  $F_2$  generation in the two crosses, as well as  $BC_1$  and  $BC_2$  generations in the cross Giza 92 x Pima S6, were located near the most studied traits. The scree plot of the PCA for six generations on yield and yield components during the two crosses showed that the first two eigenvalues correspond to the whole percentage of the variance in the dataset (Fig. 5).

**Fig. 5.** Scree plot of PCA between respective eigenvalues and components number.

## DISCUSSION

Highly significant differences among the two parents and their  $F_1$ ,  $F_2$  and backcross generations were observed for all studied traits in the two crosses, indicating the presence of genetic variability for these traits in those studied materials. The generation/error variance values were higher than unity, indicating that the genetic variances were higher than the environmental variances among the six generations for those traits across the two crosses. Similar types of results in the six populations were reported by Hussain *et al.* (2008) in cotton and Omoigui *et al.* (2019) in cowpea. Significant differences were observed among the  $F_2$ ,  $F_3$  and  $F_4$  generations for most studied traits by Srour and El-Hashash (2012), demonstrating the existence of inherent variability among the two crosses of Egyptian cotton. The analysis of variance depicted highly significant genetic variability among cotton germplasm for yield and its components (Shah *et al.*, 2018; Yehia and El-Hashash, 2019).

Based on the CV% values, the environmental influence was low for lint percentage and seed index traits when compared to those referring to the other traits, so this trial would be considered to have high precision. The magnitude of CV% indicated that the non-segregation and segregation generations had exploitable genetic variability during the selection of traits examined in cotton. These results are different from those identified by Raza *et al.* (2016) and Li *et al.*, (2020) in cotton, where the CV% values were more than 10%.

The results of mean performances indicated that the relationship between non-segregation and segregation generations revealed that there was different behaviour for studied traits in the study materials. Thus, it is possible to benefit from the selection in the segregation generations in future breeding programmes to improve these traits in Egyptian cotton. It has been reported that the mean values of non-segregation generations were better than the mean values of segregation generations in some cases, and conversely, in another case, in the two crosses of Egyptian cotton.

The correlation coefficients among different pairs of plant traits for the two crosses across generations indicated that cotton yield can be improved and increased by increasing most of the yield component traits. The result could be a possibility of plants with desirable attributes to cotton yield in the next segregating generations. In this connection, Srour and El-Hashash (2012) reported similar results. Positive intergeneration correlations for studied traits in the two crosses indicated that selection for the increased value of one trait will result increase in value of other. The finding of Barman and Borah (2012) in mutant rice strain and Srour and El-Hashash (2012) in cotton revealed that the correlation coefficient among  $F_2$ ,  $F_3$  and  $F_4$  generations were significant or highly significant for studied traits. A significant correlation ( $p < 0.05$  or  $0.01$ ) between cotton yield and yield components traits across the single and double-cross hybrids were observed by El-Hashash (2013).

Generally, the seed cotton yield/plant positively and significantly correlated with all studied traits in the two crosses, except the lint percentage in the cross Giza 93 x C.B. 5. This implies that cotton yield is effective for selection in a later generation and the two crosses may be used in improving cotton yield in Egyptian cotton. Like present findings, the significant positive correlation for yield and its components traits were reported by Khokhar *et al.* (2017), Jarwar *et al.*, (2019), Kumar (2020), Rehman *et al.*, (2020) and Sarwar *et al.*, (2021). Srour and El-Hashash (2012) and Khokhar *et al.*, (2017) stated that more than one trait can be used as a selection standard in the next segregating generations of cotton.

Using UPGMA hierarchical clustering, the tree diagram exhibited the highest correlation between the traits or generations inside each cluster. While the lowest correlation of traits or generations was found among the clusters, in amphidiploid cotton hybrid plants, Muminov *et al.*, (2020) mentioned that the cluster analysis grouped the parents and  $F_1$ - $F_6$  generations into four different clusters. According to the report of Abdel-Monaem *et al.* (2020), the parents and  $F_1$  generation plants were divided into four and seven major clusters, respectively, and the clusters of the  $F_1$  generation were widely divergent from the parental populations in Egyptian cotton. Based on 100 cotton genotypes of *Gossypium hirsutum* L., the mean performance of yield and yield component traits was grouped into six different clusters (Akter *et al.*, 2019).

The PCA1 was kept for the final analysis (Eigenvalue  $>1$ ), in which, the PCA1 explains variance more than an individual attribute (Sharma, 1996) and it expresses more variability and support to select the trait with a positive loading factor. The proportions explained by PC1 and PC2 indicate that the first two PCAs may be used to summarize the original variables in any further analysis of the data, as well as to explain the total variation and the grouping of the PCAs. Based on all measured data in the two crosses, the first two PCAs had mainly distinguished the generations in different groups. Therefore, the first two PCAs were employed to draw a biplot. According to Abasianyanga *et al.*, (2017), Shah *et al.*, (2018), Akter *et al.*, (2019), Abdel-Monaem *et al.*, (2020) and Sarwar *et al.*, (2021), the eigenvalues had higher than one for the first six, five, four, eight and four PCAs, and which contributed for 70.39%, 78.66%, 53.23%, 85.28% and 79.90% of the accumulative variation of the original variables under examination in cotton, respectively. The PCA1 and PCA2 explained 99.77% and 69.30% of the total variation between six generations in sugar beet (Bayomi *et al.*, 2019) and between 230  $F_2$  populations in cotton (Nandhini *et al.*, 2018) based on all studied traits, respectively.

Selection during generations that have high PCA1 and PCA2 for studied traits are suitable and effective in the two crosses. In practice, the choice of  $F_2$  vs. backcross-based populations in "second cycle" breeding is complicated by the fact that the breeder regards not only a single trait but several characters simultaneously (Meichinger, 1987). These components can be interpreted as a response related to the yield and its components traits and which possess positive and negative contributions to the variables. PCA1 is considered very important in breeding programs to improve cotton yield. Similarly, Abasianyanga *et al.*, (2017), Shah *et al.*, (2018), Rathinavel (2019), Jarwar *et al.*, (2019) reported that all or some yield and yield components traits showed a positive correlation of factor loadings on the five PCAs, especially PCA1 that divulged great contribution towards total variability.

El-Hashash and EL-Agoury (2019) stated that the acute angle (below 90 degrees) between the two variables indicates the significance of these variables. Positive and significant correlations for most studied traits indicate that selection based on these traits would result in an increasing cotton yield in both crosses. Hence, emphasis must be placed on these materials in a breeding program to improve the Egypt cotton. The biplot showed the degree of correlation amongst cotton traits (Sarwar *et al.*, 2021). The PCA1 and PCA2 results indicate that there are differences between yield and yield components variables across the six generations in the two crosses. The respective variable distances from the first two PCAs demonstrate the contribution of different variables at the total variability (Sarwar *et al.*, 2021).

The biplot analysis of the relationship between the six generations revealed that the most appropriate generations for selecting yield traits were  $F_1$  in the two crosses and  $BC_1$  and  $BC_2$  in the cross Giza 93 x C.B. 58. Meichinger (1987) declared that  $F_2$ ,  $BC_1$  and  $BC_2$  offer equal alternatives with respect to time, work,



inbreeding level, and the amount of genetic variation released within lines in subsequent selfing generations if linkage and epistasis are of small importance. Therefore, the choice to separate a population can be based on the characteristics of the first segregating generations. During the biplot diagram of the first two PCAs, the plants close to the ideal type would be selected (Abdel-Monaem *et al.*, 2020). The extent of variation in each trait between genotypes showed greater divergence (Rathinavel, 2019) and most traits contributed to more variance (Nandhini *et al.*, 2018), thus these traits could be considered essential for selection in large populations as well as which may be useful for an effective program to cotton breeding. The results of scree plot were harmonic with El-Hashash (2016) who reported that there is a break in the plot that separates the meaningful components from the trivial components. Also, the results were consistent with Abdel-Monaem *et al.*, (2020) among parental cotton genotypes and their F<sub>1</sub>S cross combinations in Egyptian cotton. Most researchers would agree that the first two PCAs are probably meaningful. This finding is in agreement with those reported by Abasianyanga *et al.*, (2017), Jarwar *et al.*, (2019) and Sarwar *et al.*, (2021). These results show that there is a divergence among the six populations, thus this diversity can be used to improve the yield and its components in cotton.

## CONCLUSION

Significant divergences among the six generations for all studied traits in the two crosses were observed by ANOVA. The F<sub>1</sub> performance was higher than the other generations for all the studied traits in the two crosses. The results of Pearson's correlation coefficient and multivariate analysis from our study could be useful and used in breeding programmes for cotton yield improvement. As a result, we recommend considering backcrossing for 2–5 cycles (BC<sub>2</sub>–BC<sub>5</sub>) at the C.B. 58 parent in the future to improve Egyptian cotton yield.

**Funding:** Not applicable.

**Conflict of Interest:** The authors declare no conflict of interest

## REFERENCES

- Abasianyanga, I., Balu, P. A. & Ramakrishnan, P. (2017). Association and principal component analysis of yield and its components in cultivated cotton. *Electronic Journal of Plant Breeding*, 8(3), 857-864. DOI: 10.5958/0975-928X.2017.00140.5
- Abdel-Monaem, M.A., Abido, W. A. E. Hadházy, A., Ghoneima, M. H. EL-Mansy Y. M. & EL-Shazly, M. W. (2020). Genetic divergence among Egyptian cotton genotypes under water deficit conditions, *Acta Ecologica Sinica*, <https://doi.org/10.1016/j.chnaes.2020.11.007>
- Addinsoft, (2021). XLSTAT statistical and data analysis solution. New York, USA. <https://www.xlstat.com>.
- Akter, T., Islam, A. K. M. A., Rasul, M. G., Kundu, S., Khalequzzaman & JAhmed J. U. (2019). Evaluation of genetic diversity in short duration cotton (*Gossypium hirsutum* L.). *Journal Cotton Research*, 2, 1. <https://doi.org/10.1186/s42397-018-0018-6>
- Barman, D. & Borah S. P. (2012). Effect of Selection Response on F<sub>3</sub> and F<sub>4</sub> Generation for Yield and Yield Component Characters in Mutant Rice Strain (*Oryza sativa* L.), *APCBEE Procedia*, 4, 183-187, <https://doi.org/10.1016/j.apcbee.2012.11.031>.
- Bayomi, K. E. M., El-Hashash, E. F., & Moustafa, E. S. A. (2019). Comparison of Genetic Parameters in Non-segregating and Segregating Populations of Sugar Beet in Egypt. *Asian Journal of Research in Crop Science*, 3(4), 1-12. <https://doi.org/10.9734/ajrcs/2019/v3i430055>
- Eldessouky, S. E. I., El-Fesheikawy, A. B. A. & Baker, K. M. A. (2021). Genetic variability and association between oil and economic traits for some new Egyptian cotton genotypes. *Bull Natl Res Cent* 45, 43. <https://doi.org/10.1186/s42269-020-00438-1>
- El-Hashash, E. F. (2013). Heterosis and Gene Action among Single and Double-Cross Hybrids Performances in Cotton. *American-Eurasian Journal of Sustainable Agriculture*, 13 (4), 505-516. DOI:10.5829/idosi.aejaes.2013.13.04.1955
- El-Hashash, E. F. (2016). Genetic Diversity of Soybean Yield Based on Cluster and Principal Component Analyses. *Journal of Advances in Biology & Biotechnology*, 10(3), 1-9. <https://doi.org/10.9734/JABB/2016/29127>
- El-Hashash, E. F. (2017). Comparison of various methods for estimating of heritability and number of genes affecting the quantitative traits in cotton. Egypt. *Journal Plant Breeding*, 21(5), 612-628.
- El-Hashash, E. F. & EL-Agoury R. Y. (2019). Comparison of grain yield-based drought tolerance indices under normal and stress conditions of rice in Egypt. *Scholars Journal of Agriculture and Veterinary Sciences*, 6(1), 41-54.
- FAOSTAT, (2021) Food and agriculture organization of the United Nations. <http://faostat.fao.org>. FAO Statistics Division. Crops Primary data last updated April 20, 2021.
- Gomes, F. P. (2009). Curso de estatística experimental. 15.ed. Piracicaba: Esalq, 477p.
- Hammer, Ø., Harper, D. A. T. & Ryan P. D. (2020). Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, 4(1), 9 pp.
- Hussain M., Azhar F. M. & Khan A. A. (2008). Genetic basis of variation in leaf area, petiole length and seed cotton yield in some cotton (*Gossypium hirsutum*) genotypes. International. *Journal Agriculture Biology*, 10, 705–708.
- Jarwar, A. H., Wang, X., Iqbal M., Sarfraz, Z., Wang, L. Qifeng M. A. & Shuli, F. (2019). Genetic divergence on the basis of principal component, correlation and cluster analysis of yield and quality traits in cotton cultivars. *Pakistan Journal of Botany*, 51(3), 1143-1148. DOI: 10.30848/PJB2019-3(38).

- Kerby, T. A., Cassman, K. G. & Keeley, M. (1990). Genotypes and plant densities for narrow-row cotton systems. I. Height, nodes, earliness, and location of yield. *Crop Science*, 30(3), 644-649. <https://doi.org/10.2135/cropsci1990.0011183X003000030034x>
- Khokhar, E. S., Shakeel, A., Maqbool, M. A., Anwar, M. W., Tanveer, Z. & Irfan, M. F. (2017). Genetic study of cotton (*Gossypium hirsutum* L.) genotypes for different agronomic, Yield and quality traits. *Pakistan Journal of Agricultural Research*, 30(4), 363-372. <http://dx.doi.org/10.17582/journal.pjar/2017/30.4.363.372>
- Kloth, R. H. (1998) Analysis of commonality for traits of cotton fiber. *Journal Cotton Science*, 2, 17-22.
- Kumar, K. B. (2020). Inter Relationship between Seed Cotton Yield and Yield Contributing Characters in American Cotton (*Gossypium hirsutum* L.). *International Journal of Current Microbiology and Applied Sciences*, 9(6), 2276-2279. <https://doi.org/10.20546/ijcmas.2020.906.278>
- Latif, A., Bilal, M., Hussain, S. B. & Ahmad, F. (2105). Estimation of genetic divergence, association, direct and indirect effects of yield with other attributes in cotton (*Gossypium hitsutum* L.) using biplot correlation in and path coefficient analysis. *Tropical Plant Research*, 2(2), 120-126.
- Li, B., Tian, Q., Wang, X., Han, B., Liu, L., Kong, X., Si, A., Wang, J., Lin, Z., Zhang, X., Yu, Y. & Yang, X. (2020). Phenotypic plasticity and genetic variation of cotton yield and its related traits under water-limited conditions. *The Crop Journal*, 8(6), 966-976. <https://doi.org/10.1016/j.cj.2020.02.003>.
- Meichinger, A. E. (1987). Expectation of means and variances of testcrosses produced from F2 and backcross individuals and their selfed progenies. *Heredity*, 59, 105-115. DOI:10.1038/hdy.1987.101
- Muminov, K. H., Ernazarova, Z. & Amanov, B. (2020). Cluster analysis of valuable economic traits in amphidiploid cotton hybrid plants. *EurAsian Journal of BioSciences*, 14, 4973-4981.
- Nandhini, K., Amala Balu, P. & Isong, A. (2018). Genetic Analysis and Inheritance Studies in F2 Population of Upland Cotton (*G. hirsutum* L.). *International Journal of Pure & Applied Bioscience*, 6(2), 1499-1505. doi: <http://dx.doi.org/10.18782/2320-7051.6472>
- Omoigui O. L., Arrey, M. O., Kamara, A. Y., Danmaigona, C. C., Ekeruo, G., & Timko, M. P. (2019). Inheritance of resistance to Cercospora leaf spot disease of cowpea [*Vigna unguiculata* (L.) Walp]. *Euphytica*, 215, 101. <https://doi.org/10.1007/s10681-019-2420-7>
- Rathinavel, K. (2019). Agro-morphological Characterization and Genetic Diversity Analysis of Cotton Germplasm (*Gossypium hirsutum* L.). *International Journal of Current Microbiology and Applied Sciences*, 8(2), 2039-2057. <https://doi.org/10.20546/ijcmas.2019.802.237>
- Rehman, A. & Farooq, M. (2020). Morphology, Physiology and Ecology of Cotton Chapter 2. Cotton Production, First Edition. Edited by Khawar Jabran and Bhagirath Singh Chauhan. John Wiley & Sons Ltd.
- Rehman, A., Mustafa, N., DU, X. & Azhar M.T. (2020). Heritability and correlation analysis of morphological and yield traits in genetically modified cotton. *Journal Cotton Research*, 3(23), 1-9. <https://doi.org/10.1186/s42397-020-00067-z>
- Raza, H., Khan, N. U., Khan, S. A., Gul, S., Latif, A., Hussain, I., Khan, J., Raza, S., & Baloch, M. (2016). Genetic variability and correlation studies in F4 populations of upland cotton. *The Journal of Anim. & Plant Science*, 26(4), 1048-1055.
- Sarwar, G., Nazir, A., Rizwan, M., Shahzadi E. & Mahmood, A., (2021). Genetic diversity among cotton genotypes for earliness, yield and fiber quality traits using correlation, principal component and cluster analyses. *Sarhad Journal of Agriculture*, 37(1), 307-314. <http://dx.doi.org/10.17582/journal.sja/2021/37.1.307.314>
- Shah, S. A. I., Khan, S. J. Ullah, K. & Sayal, O. U. (2018). Genetic diversity in cotton germplasm using multivariate analysis. *Sarhad Journal of Agriculture*, 34(1):,130-135. <http://dx.doi.org/10.17582/journal.sja/2018/34.1.130.135>
- Sharma, S. (1996). Applied Multivariate Techniques; Wiley: New York, NY, USA.
- Srouf, M. S. M. & El-Hashash, E. F. (2012). Phenotypic and Genotypic Correlation Coefficients for Some Yield and Fiber Quality Traits of Segregating Populations (F<sub>2</sub>, F<sub>3</sub> and F<sub>4</sub>) in Egyptian Cotton Crosses. *Journal of King Abdulaziz University-Meteorology Environment and Arid Land Agriculture Sciences*, 23(1), 19-35. DOI: 10.4197/Met. 23-1.2.
- Steel, R. G. D., Torrie, J. H. & Dickey, D. A. (1997). Principles and procedures of statistics: a biometrical approach. 3rd ed. New York: McGraw Hill.
- USDA, (2021). United States Department of Agriculture Foreign Agricultural Service, Office of Global Analysis. Cotton and Products Annual, Egypt Cotton Production Rebounds. March 31, 2021, Report Number: EG2021-0005. 1-17.
- Yehia, W. M. B. & El-Hashash, E. F. (2019). Combining ability effects and heterosis estimates through line x tester analysis for yield, yield components and fiber traits in Egyptian cotton. *Elixir Agriculture*, 131, 53238-53246.



**Copyright:** © 2021 by the authors. Licensee EJAR, EKB, Egypt. EJAR offers immediate open access to its material on the grounds that making research accessible freely to the public facilitates a more global knowledge exchange. Users can read, download, copy, distribute, print or share a link to the complete text of the application under [Creative Commons BY-NC-SA 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

## الارتباطات والتحليل متعدد المتغيرات عبر الأجيال غيرالإنعزالية والإنعزالية في هجينين من القطن

وليد محمد يحيى<sup>1</sup> و عصام فتحى الحشاش<sup>2\*</sup>

<sup>1</sup> قسم بحوث تربية القطن، معهد بحوث القطن، مركز البحوث الزراعية، الجيزة، مصر

<sup>2</sup> قسم المحاصيل، كلية الزراعة، جامعة الأزهر، القاهرة، مصر

\* بريد المؤلف المراسل: [dressamelhashash@yahoo.com](mailto:dressamelhashash@yahoo.com)

### الملخص

في هذا البحث تم استخدام معامل ارتباط بيرسون وتحليل متعدد المتغيرات لتقييم العلاقات المتبادلة والتشابهات والاختلافات بين الأجيال غير الإنعزالية ( $F_1$ ،  $P_2$ ،  $P_1$ ) والإنعزالية ( $BC_2$ ،  $BC_1$ ،  $F_2$ ) لمحصول القطن الزهر ومكوناته في الهجينين (Pima S6 x Giza 92) و (C.B. 58 x Giza 93). أظهر تحليل التباين تبايناً وراثياً عالي المعنوية بين الستة أجيال لجميع الصفات المدروسة في الهجينين. كان أداء الجيل الأول ( $F_1$ ) أعلى من الأجيال الأخرى لجميع الصفات المدروسة في الهجينين. أظهرت النتائج ارتباطات موجبة وعالية معنوية بين صفات محصول القطن الزهر/نبات، محصول القطن الشعير/نبات وعدد اللوز/نبات عبر الأجيال الستة في الهجينين. لوحظ عدد من الارتباطات الموجبة بين الأجيال الستة لمحصول القطن الزهر ومكوناته في الهجينين. أظهر تحليل UPGMA hierarchical clustering مستوى أعلى من معاملات التشابه بين الأجيال الستة وبين الصفات المدروسة. تراوحت معاملات التشابه للأجيال والصفات المدروسة من 0.96 الى 0.99 ومن 0.65 الى 0.96، على التوالي. في تحليل (PCA) principal component analysis، كان للمكون الرئيسي الأول (PCA1) قيمة  $Eigenvalue > 1$  عبر الستة أجيال لجميع الصفات المدروسة في الهجينين. أظهر PCA تبايناً إجمالياً بنسبة 91.84% بين الأجيال الستة التي ساهم بها PCA1 (79.47%) و PCA2 (12.38%) والتي ميزت بشكل أساسي الأجيال في مجموعات مختلفة. كما بينت النتائج سيطرت الجيل الأول والأجيال الإنعزالية على المكونين PCA1 و PCA2 في الهجينين، على التوالي، مما يُظهر ارتباطات عالية مع أول مكونين رئيسيين (PCA1 و PCA2). ساهمت كل الصفات المدروسة بالإضافة الى وزن اللوز والنسبة المئوية للتصافي في تحميل مكونات معنوية موجبة لكل من PCA1 و PCA2 على التوالي. كشف تحليل biplot للعلاقة بين الأجيال الستة أن الأجيال الأكثر ملاءمة لاختيار صفات المحصول كان الجيل غير الانعزالي  $F_1$  في الهجينين والجيلين الانعزاليين  $BC_1$  و  $BC_2$  في الهجين (C.B. 58 x Giza 93). ذلك، نوصي بالنظر في إمكانية إجراء التهجين العكسي لمدة تتراوح من 2 الى 5 دورات ( $BC_5 - BC_2$ ) مع الأب C.B. 58 لتحسين محصول القطن المصري في المستقبل.

**الكلمات المفتاحية:** الاجيال، الارتباط، التحليل متعدد المتغيرات، القطن