

Detecting Similarities and Differences of Promising Elite Cotton (*Gossypium hirsutum* L.) Genotypes Influenced by Seed Cotton Yield and Related Components

Mare Marco, Mubvekeri Washington

Department of Agricultural Research, Innovation and Development, Cotton Research Institute, Kadoma, Zimbabwe

Email address:

marcomare82@gmail.com (Mare Marco.), wmubvekerishumba@gmail.com (Mubvekeri Washington)

To cite this article:

Mare Marco, Mubvekeri Washington. (2024). Detecting Similarities and Differences of Promising Elite Cotton (*Gossypium hirsutum* L.) Genotypes Influenced by Seed Cotton Yield and Related Components. *Journal of Plant Sciences*, 12(1), 30-35.

<https://doi.org/10.11648/jps.20241201.15>

Received: May 11, 2023; **Accepted:** June 5, 2023; **Published:** February 1, 2024

Abstract: Cotton (*Gossypium hirsutum* L.) is a premier cash and natural fibre crop grown worldwide. Cotton Breeding Programme in Zimbabwe is driven by the basic principle of continuous improvement in genetics of available cotton germplasm for high production and productivity under varying agro-climatic conditions. An experiment was carried out at six sites namely Chitekete, Chizvirizvi, Masakadza, Muzarabani, Tokwane and Umguza over two seasons, 2020 and 2021. Ten treatments that comprised of eight promising lines and two checks were replicated three times and laid out in a Randomized Complete Block Design. Gross plot was 36m² and Net plot was 16m². Seed Cotton yield, lint yield, plant height, 100 boll weight, Earliness Index, Boll weight and Gin out turn were measured and data was analyzed using Genstat 18th version. Maximum variation as indicated by high values that ranged from 16.4% to 20.1% was observed for Earliness Index, Total Seed Cotton, Plant Height and Lint Yield whilst least variation was recorded for Gin-Out-Turn and 100 seed weight. Mean performance results revealed a yield difference of 539kg/ha between the highest and the lowest yielding candidate thus 22% yield advantage for 83-01-4 over 820-01-1. Of the seven computed principal components only PC1 and PC2 achieved eigenvalue >1 and contributed about 63.99% of the accumulative variation among the attributes under examination. The study results gave the breeder an option to target parents with strengths in some important traits for use in cotton yield improvement programmes. Key suggestions were also hinted for use in hybridization combinations.

Keywords: Principal Component Analysis, Correlation Coefficient, Hybridization, Breeding Programmes, Variation

1. Introduction

Cotton (*Gossypium hirsutum* L.) is a premier cash and natural fibre crop grown worldwide. Zimbabwe regards the crop as one of the main sources of foreign currency hence an essential and strategic crop for macroeconomic development of the country. Cotton Breeding Programme in Zimbabwe is driven by the basic principle of continuous improvement in genetics of available cotton germplasm for high production and productivity under varying agro-climatic conditions. Varied performance of genotypes can be attributed to many factors which include genetics-related, environmental-related and or the interaction of the two. The analytic studies performed with the aim of recommending the best genotypes suitable for production in particular environments have been

overshadowed by climatic uncertainties also known as climatic change. Hence there is need to employ some approaches that provide information which promotes clear understanding of the similarities and differences of genotypes being evaluated. Productivity is a universal performance attribute which makes a variety or genotype be highly or not adopted. That means comprehensive understanding about the crop nature, performance level, and association of numerous agronomic attributes with yield is necessary and becomes easy to tackle cotton yield limiting constraints [7].

Principal Component Analysis (PCA) is amongst multivariate biometrical techniques which have been frequently used to explore genetic diversity among genotypes on direct and indirect effects (Brown Guedira et al., 2000). The PCA technique makes it easy to estimate genetic

variation of morphological traits for easy recognition of phenotypic variation [7]. PCA removes dependency or redundancy in the data by dropping those features that contain the same information as given by other attributes and the derived components are independent of each other. The extent of genetic variation in variety development is very important and, in this study, PCA has been employed to workout similarities and differences between different genotypes regarding multiple traits under examination [6, 10, 12, 14].

Therefore, the aim of this study was to explore variability among some promising elite cotton germplasm for yield and related components and improve the cotton breeding system.

2. Material and Methods

2.1. Experimental Sites and Season Information

A field Experiment was carried out at six different sites namely Chitekete, Chizvirizvi, Masakadza, Muzarabani, Tokwane and Umguza (see Table 1 for more site information). The experiment was conducted over two seasons, thus 2020

and 2021 season.

2.2. Experimental Candidates and Records

The experiment involved ten genotypes, whereby 2 were checks (see Table 2 for more information). Seven parameters were measured and recorded and these included Seed Cotton yield, lint yield, plant height, 100 boll weight, Earliness Index, Boll weight and Gin out turn.

2.3. Experimental Design and Layout

The experiment was laid out in Randomized Complete Block Design replicated 3 times. Gross plot was 36m² and Net plot was 16m².

2.4. Data Analysis

The observations recorded were subjected to Genstat Statistical Package 18th version. The output included mean performance for all recorded observations, Coefficient of Variance, Principal component analysis for Eigen vectors, Correlations among Variates, and Intergroup Distances.

Table 1. Experimental Sites Information.

Location	Latitude	Longitude	Altitude (m)	Av. Annual Rainfall (mm)	Max Temp °C
Chitekete	17°25' South	28° 56' East	914	450-500	45
Masakadza	17°25' South	16°28' East	914	450-650	45
Tokwane	25°47' South	31°15' East	1105	350-650	37
Umguza	20° 03' South	28°34' East	1374	450-500	34
Muzarabani	16° 23' South	31° 00' East	432	600-800	42
Chizvirizvi	29°59' South	32°10' East	410	450-600	40

Table 2. Experimental Genotypes information.

Genotype	Status	Source
812-01-3	Experimental Line	Cotton Research Institute
820-01-1	Experimental Line	Cotton Research Institute
83-01-4	Experimental Line	Cotton Research Institute
830-01-3	Experimental Line	Cotton Research Institute
830-01-7	Experimental Line	Cotton Research Institute
831-01-3	Experimental Line	Cotton Research Institute
85-01-1	Experimental Line	Cotton Research Institute
89-01-2	Experimental Line	Cotton Research Institute
CRIMS1	Commercial Variety	Cotton Research Institute
CRIMS2	Commercial Variety	Cotton Research Institute

3. Results and Discussions

3.1. Coefficient of Variance (CV)

Maximum variation as indicated by high values was observed for parameters Earliness Index-20.1%, Total Seed Cotton-18.4%, Plant Height-18.8% and Lint Yield-16.4%, whilst least variation was recorded for Gin-Out-Turn-6.5%. These high values of variation indicated that the traits are strongly affected by environmental factors which occurred from one site to another as well as across the two seasons. The characters with low variation are typically not easily affected by changes in environments. The results are in agreement with what was obtained by Vinodhana and

Gunasekaran 2019 [7], where the two researchers carried out an experiment with the aim of exploring genetic diversity among different collected germplasm. In their experiment, maximum variation of around 17% was recorded on Seed Cotton yield, plant height and number of bolls. This would require further analysis for identifying the number of genes controlling the traits.

3.2. Mean Performance for All Recorded Parameters

3.2.1. Total Seed Cotton

Among the ten tested genotypes, 83-01-4 and 89-01-2 yielded the highest with 2427kg^{ha}⁻¹ and 2304kg^{ha}⁻¹ respectively. 820-01-1 recorded the lowest yield of 1888kg^{ha}⁻¹. This reveals a yield difference of 539kg^{ha}⁻¹

between the highest and the lowest yield thus 22% yield advantage for 83-01-4 over 820-01-1.

3.2.2. Plant Height

Tall plants were recorded for CRIMS2, 131.9cm, whilst short plants were recorded for 89-01-2. The variation on height was above 5cm which reflected directly to the maximum variation recorded for the trait.

3.2.3. Gin Out Turn

The highest ginning percentage was recorded for 831-01-3 that revealed 43.43% followed by 89-01-2 that recorded 43.13%. That was above the mean performance of the trait 42.60%. Candidate 83-01-4 recorded the lowest ginning percentage of 42.03%.

3.2.4. Earliness Index

Though all the genotypes recorded an index above 75%,

candidate 831-01-3 achieved 79.12% of its total yield from first pick, whilst 830-01-7 followed with a record of 78.64%. CRIMS1 recorded the lowest Earliness Index of 75.78%.

3.2.5. Boll Weight

Overall mean performance on boll weight was 5.716g, and this was surpassed by the highest record of 5.875g and 5.803g by 812-01-3 and 820-01-1 respectively. Small bolls were recorded for 831-01-3 that recorded 5.639g and 830-01-7 that recorded 5.650g respectively.

3.2.6. 100 Seed Weight

812-01-3 recorded the highest seed weight of 11.050g, while 831-01-3 and CRIMS2 recorded the lowest seed sizes of 10.389g and 10.442g respectively which are even below the overall mean of 10.688g.

Table 3. Mean Performance on Measured Attributes for the 10 Promising Elite Cotton Genotypes.

Genotype	TSC (kg ha ⁻¹)	PH (m)	LY (kg ha ⁻¹)	GOT (%)	EI (%)	BW (g)	100 SW (g)
812-01-3	2063	130.2	862	42.47	76.88	5.875	11.05
820-01-1	1888	127.2	791	42.08	76.91	5.678	10.66
83-01-4	2427	130	1020	42.03	77.77	5.803	10.56
830-01-3	2174	127.5	925	42.59	77.85	5.661	10.84
830-01-7	2095	127.6	878	42.59	78.64	5.650	10.61
831-01-3	2126	131.1	911	43.43	79.12	5.639	10.39
85-01-1	2145	128.9	904	42.06	77.79	5.775	10.84
89-01-2	2304	124.9	982	43.13	77.12	5.686	10.68
CRIMS1	2148	126.9	918	42.86	75.78	5.719	10.83
CRIMS2	2142	131.9	916	42.81	77.16	5.675	10.42
Mean	2151	128.6	911	42.6	77.50	5.716	10.69
CV	18.4	18.8	16.4	6.5	20.1	10.8	6.9

TSC-Total Seed Cotton yield, LY-lint yield, PH-plant height, 100 SW-100 seed weight, EI-Earliness Index, BW-Boll weight and GOT-Gin out turn.

3.3. Principle Component Analysis

Pattern of genetic variability in the evaluated population of 10 genotypes grouped the measured traits into components through Principal Component Analysis. Seven PCs were realized but only two described eigenvalue >1 and contributed about 63.99% of the accumulative variation among the attributes under examination (Table 4 & Figure 1). That means the two PCs had greater influence in explaining the total variation and the classification of the PCs.

The results revealed that the first PC explained the highest variability of 48.08% (Figure 1) and this was as an effect of some traits which included Total Seed Cotton (0.49682), Lint yield (0.494), 100 seed weight (0.42141) and Boll Weight (0.41541). Such information also signified their importance in divergence. The second PC explained variability of 15.91% and was more associated with Got% (0.77311) and Boll weight (0.12347). Minimum variability was observed in PC6 and PC7 with the eigenvalues of 0.266 and 0.003, respectively (Table 4). The indication is in agreement with what Vinodhana and Gunasekaran 2019 [7] got in a related experiment that they carried out.

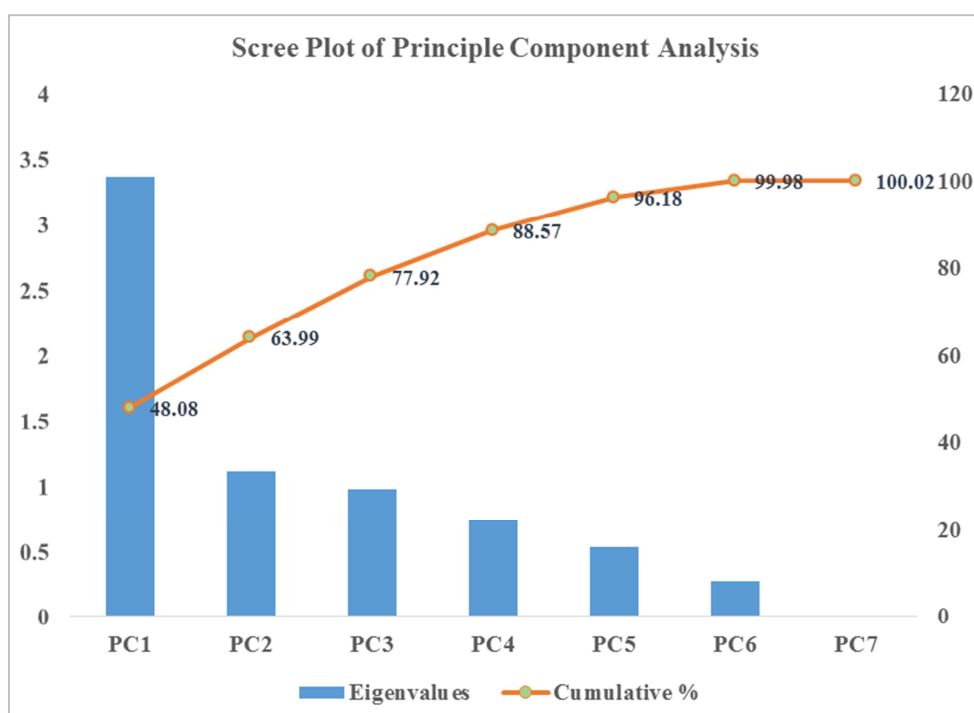
Under PC1, the germplasm displayed positive effects for Total Seed Cotton (0.49682), Lint yield (0.494), 100 seed weight (0.42141) and Boll Weight (0.41541), Plant Height (0.39258), whilst negative effects were revealed for Earliness Index (-0.03887) and Gin Out Turn (-0.05784).

Germplasm in PC2 showed positive effects for Got% (0.77311), Boll weight (0.12347), 100 Seed Weight (0.01705), Boll Weight (0.12347) and Lint Yield (0.03764) whilst Earliness Index, Plant Height and Total Seed cotton were negative. PC1 had maximum variability so the genotypes in PC1 should be opted for selection, followed by those in PC2. The attributes in PC1 and PC2 that contributed to much of the variation (63.99%) gained much and special emphasis on cotton improvement in future breeding work, and this is in line with similar results by Kaleri *et al.*, 2015, Latif *et al.*, 2015 and Shah *et al.*, 2018 [6, 8, 13]. Vinodhana and Gunasekaran 2019 [7] also through their study on PCA, identified few characters that they concluded to play prominent role in classifying the variation existing in the germplasm set, and in the list included Seed Cotton yield, Number of bolls, boll weight, micronaire value, number of sympodia and span length.

Table 4. Eigenvalues and Cumulative Variability in Different PCs for Yield and Related Attributes in Cotton Cultivars.

Traits	PC1	PC2	PC3	PC4	PC5	PC6	PC7
100 SW (g)	0.42141	0.01705	0.15447	-0.58885	0.0342	0.67106	0.00494
BW (g)	0.41541	0.12347	0.39033	-0.40181	-0.06009	-0.70332	-0.01023
EI (%)	-0.03887	-0.61408	0.72341	0.27139	0.11464	0.10572	0.01052
GOT (%)	-0.05784	0.77311	0.48154	0.31506	0.16865	0.17299	0.09718
LY (kg ha^{-1})	0.494	0.03764	-0.0467	0.40613	-0.30409	0.0766	-0.69935
PH (m)	0.39258	-0.07026	-0.22968	0.17535	0.86641	-0.0822	0.00496
TSC (kg ha^{-1})	0.49682	-0.05766	-0.11681	0.35099	-0.33242	0.03607	0.70796
Statistical Variables							
Eigenvalues	3.366	1.114	0.975	0.745	0.532	0.266	0.003
Variability %	48.08	15.91	13.93	10.65	7.61	3.8	0.04
Cumulative %	48.08	63.99	77.92	88.57	96.18	99.98	100.02

TSC (kg ha^{-1})=Total Seed Cotton Yield, BW (g)=Boll weight (g), GOT (%)=Gin Out Turn percentage, EI (%)=Earliness Index, LY (kg ha^{-1})=Lint Yield, PH=Plant height (cm), 100 SW (g)=100 Seeds weight

**Figure 1.** Eigenvalue for Yield and Yield Contributing Factors of Cotton Showing the Cumulative Variability Percentage (%) (line).

3.4. Intergroup Distances-Mahalanobis (D-squared)

The maximum intergroup distances among tested genotypes was between CRIMS2 and 830-01-7 (6), followed by CRIMS2 and 89-01-2 (4.554), and CRIMS1 and 830-01-3 (4.553), Table 5. Intergroup distances of above 2 were recorded between CRIMS1 and 89-01-2 (3.479), CRIMS1 and 85-01-1 (2.78), CRIMS2 and 831-01-3 (2.616), 89-01-2 and 83-01-1 (2.604) and 830-01-7 and 830-01-3 (2.726). This indicates that the genotypes are far diverse or distant relationship. Some genotypes indicated close relationships

such as 820-01-1 and 812-01-3 (0.193), 85-01-1 and 89-01-2 (0.305), 830-01-7 and 820-01-1 (0.346), and CRIMS1 and 830-01-3 (0.355). The results indicated that those genotypes with distant relationship can be used in future breeding programs and have huge promise for developing high variations that are important in developed segregated. This was also indicated by Ali *et al.*, 2012 and Siddique *et al.*, 2010 [1, 13, 15] in their study on the contribution of characters towards divergence of the genotypes). The results can also be as a baseline for developing superior genotypes [2, 3, 9].

Table 5. Intergroup Distances-Mahalanobis (D-squared).

812-01-3	0			
820-01-1	0.193	0		
83-01-4	1.149	0.905	0	
830-01-3	1.738	1.511	3.772	0

830-01-7	0.41	0.346	0.327	2.726	0				
831-01-3	0.974	0.533	1.301	0.951	0.935	0			
85-01-1	0.681	0.479	2.021	0.358	1.334	0.451	0		
89-01-2	1.415	1.149	2.604	0.205	1.93	0.419	0.305	0	
CRIMS1	0.971	1.237	0.665	4.553	0.355	2.255	2.78	3.478	0
CRIMS2	1.926	1.807	0.442	6	0.714	2.616	3.774	4.554	0.458
	812-01-3	820-01-1	83-01-4	830-01-3	830-01-7	831-01-3	85-01-1	89-01-2	CRIMS1

3.5. Correlations Between DATA Variates (Correlation Coefficient Analysis)

Correlation analysis showed that Total seed cotton yield was positively correlated to all related attributes. Its relationship to plant height (0.5872) was the highest, followed by its relationship to Boll weight (0.5107, then to Lint yield (0.4685) and to 100 Seed Weight, whilst weak positive correlation was recorded to Gin out turn (0.00292) and to Earliness Index (0.0494). Height had a string positive contribution to total seed cotton as it directly influenced to more fruiting branches hence more fruiting bodies which resulted in more yield, as well as boll weight abs 100 seed

weight, which are known to be yield factors. These results were in agreement with what Sahar et al., 2021 [12, 11] reported after an experiment on genetic variability where their results revealed positive correlation for yield with the similar attributes in this present study. Such attributes will be main and important target parameters when selecting parents for use in hybridization programmes where high total seed cotton is to be achieved.

A strong positive relationship was recorded between Boll weight and 100 seed weight (0.7075), whereas strong negative relationship was recorded between Gin-Out-Turn and Earliness index (-0.104), and between plant height and Gin out turn (-0.1282).

Table 6. Correlations Between DATA Variates (Correlation Coefficient Analysis).

100 SW (g)	1							
BW (g)	0.7075	1						
EI (%)	-0.0533	0.0334	1					
GOT (%)	-0.0968	0.083	-0.104	1				
LY (kg ha^{-1})	0.5438	0.5554	-0.0586	-0.0189	1			
PH (m)	0.4561	0.3863	-0.0795	-0.1282	0.5757	1		
TSC (kg ha^{-1})	0.4191	0.5107	0.0494	0.0292	0.4685	0.5872	1	
	100 SW (g)	BW (g)	EI (%)	GOT (%)	LY (kg ha^{-1})	PH (m)	TSC (kg ha^{-1})	

4. Conclusion

The investigation which included Principal component analysis, intergroup distance analysis, correlation analysis between variates, coefficient of variance analysis, multivariate mean performance analysis, deduced comprehensive evaluation. It is in this context that similarities and differences of the tested elite genotypes were obtained such that distant varieties were notched for use in future hybridization programmes [4, 16]. Such information was recorded through identifying parameters with low variation implying that they are not easily affected by change of environments, also genotypes with good performance under different environments were identified. Important traits which had a strong positive correlation with seed cotton yield were also noted, such as plant height, Boll weight, Lint yield and 100 Seed Weight and these gave the breeder an option to target parents with strengths in such traits for use as parents in cotton yield improvement programmes. In the context of distant related genotypes, which imply their usefulness in hybridization combinations, identified were CRIMS2 and 830-01-7, CRIMS2 and 89-01-2, CRIMS1 and 830-01-3, CRIMS1 and 89-01-2, CRIMS1 and 85-01-1, CRIMS2 and 831-01-3, 89-01-2 and 83-01-1 as well as 830-01-7 and 830-01-3. The results indicate that the genotypes are far diverse or of distant relationship. Such outcomes enable plant breeders to identify promising genotypes as parental sources

that will generate diverse populations for selection and for the development of improved cotton varieties. Concerning Principal Component Analysis, PC1 displayed positive effects for Total Seed Cotton, Lint yield, 100 seed weight and Boll Weight. This concluded that the parameters are influential in discriminating tested genotypes, a key aspect in cotton improvement programme. The project brought up significant results and are recommended for future research in building genetic diversity [5], broadening genetic base, and the development of superior, climate resilient varieties. In the end, production and productivity of cotton in Zimbabwe will improve in the face of climate variability.

Funding

The work was funded by the Government of Zimbabwe.

Acknowledgments

The Authors would like the Government of Zimbabwe through the Department of Agricultural Research, Innovation and Development (ARID) under the Ministry of Lands, Agriculture, Fisheries, Water and Rural Development for providing financial support in carrying out the research trials across all sites, and data collection. Lastly, authors want to thank the Breeding Department Team for its support during data collection.

Conflicts of Interest

All the authors do not have any possible conflicts of interest.

References

- [1] Ali CR, Arshad M, Khan MI, Fzal M. Study of earliness in commercial cotton (*G. hirsutum* L.) genotypes. *J Res Sci.* 2003; 14(2): 153–7.
- [2] Ashokkumar, K. and R. Ravikesavan. 2011. Morphological diversity and performance in upland cotton (*Gossypium hirsutum* L.). *J. Agri. Sci.*, 3: 107-113.
- [3] Baloch A, Baloch M, Jatoi S, et al. Genetic diversity analysis in genetically modified cotton (*Gossypium hirsutum* L.) genotypes. *Sindh Uni Res J.* 2015; 47(3): 527–30 doi: <https://doi.org/10.20546/ijemas.2019.804.304>
- [4] Brown-Guedira, G. L. (2000) Evaluation of Genetic Diversity of Soybean Introductions and North American Ancestors Using RAPD and SSR Markers. *Crop Science*, 40, 815-823. <https://doi.org/10.2135/cropsci2000.403815x>
- [5] Jian, C., L. Wei, L. Ruili and L. F. W. Feng. 2006. Genetic diversity detected by cluster analysis North of Anhui major wheat cultivars. *Chinese Agri. Sci. Bull.*, 11: 031.
- [6] Kaleri, A. A., S. Y. Rajput, G. A. Kaleri and J. A. Marri. 2015. Analysis of Genetic diversity in genetically modified and non-modified cotton (*Gossypium hirsutum* L.)
- [7] Kumari Vinodhana, N. and Gunasekaran, P. 2019. Evaluation of Genetic Diversity in Cotton (*Gossypium barbadense* L.) Germplasm for Yield and Fibre Attributes by Principle Component Analysis. *Int. J. Curr. Microbiol. App. Sci.* 8(04): 2614-2621.0.1111/pbi.12688.
- [8] Latif, A., M. Bilal, S. B. Hussain and F. Ahmad. 2105. Estimation of genetic divergence, association, direct and indirect effects of yield with other attributes in cotton (*Gossypium hirsutum* L.) using biplot correlation in and path coefficient analysis. *Tropical Plant Res.* 2(2): 120- 126.
- [9] Quailing, W. and L. Zhe. 2011. Principal Component analysis of F2 individual selection in upland cotton (*Gossypium hirsutum* L.). *J. Henan Inst. Sci. & Technol.*, (Natural Sci. Ed.), 5: 004.
- [10] Rathinavel, K. 2017. Exploration of genetic diversity for qualitative traits among the extant upland cotton (*Gossypium hirsutum* L.) varieties and parental lines. *Int. J. Curr. Microbiol. App. Sci.*, 6(8): 2407-2421.
- [11] Rehman, S. U., M. A. Abid, M. Bilal, J. Ashraf, S. Liaqat, R. I. Ahmed and G. Qambar. 2015. Genotype by trait analysis and estimates of heritability of wheat (*Triticum aestivum* L.) under drought and control conditions. *Basic Res. J. Agric. Sci. Rev.*, 4(4): 127-134.
- [12] SAHAR, A., ZAFAR, M. M., RAZZAQ, A. *et al.* Genetic variability for yield and fiber related traits in genetically modified cotton. *J Cotton Res* 4, 19 (2021). <https://doi.org/10.1186/s42397-021-00094-4>
- [13] Shah, A. S., Khan, J. S, Ullah. K1 and Sayal, O. U. 2018. Genetic Diversity in Cotton Germplasm using Multivariate Analysis. *Sarhad Journal of Agriculture*, 130-135.
- [14] Sharma, J. R. 1998. Statistical and Biometrical Techniques in Plant Breeding. New Age International (P) Limited Publishers, New Dehli. 432p.
- [15] Siddique MA, Rashid ESMH, Khalequzzaman M, et al. Genetic diversity of local rainfed rice (*Oryza sativa* L.). *Bangladesh J Plant Breed Genet.* 2010; 23(2): 41–6. <https://doi.org/10.3329/bjpb.v23i2.9324>.
- [16] Ullah A, Sun H, Yang X, Zhang X. Drought coping strategies in cotton: increased crop per drop. *Plant Biotechnology J.* 2017; 15(3): 271–84. <https://doi.org/1>