Genetic Divergence of Bt. Cotton (*Gossypium hirsutum* L.) Germplasm Based on Principle Component Analysis

Ali Imran[1], Muhammad Bilal Saleem[2], Aqib Anwer[2], Rao Saad Rehman[1*], Muhammad Naeem[3], Muhammad Farhan Hafeez[4], Annum Javaid Shah[3]

[1]Department of PB & G, University of Agriculture, Faisalabad Pakistan.
[2]Department of Entomology, University of Agriculture Faisalabad, Pakistan.
[3]Institute of Horticultural Sciences, University of Agriculture, Faisalabad, Pakistan.
[4]University College of Agriculture and Environmental Sciences, The Islamia University of Bahawalpur.
*Corresponding author's email: saad_rehman101@hotmail.com

Abstract: The basic aim of study was to evaluate the genetic diversity in Bt. germplasm of upland cotton for the improvement of existing cotton varieties or development of new cultivars. For this purpose, sixty locally adapted Bt cotton genotypes were gown in the experimental area of University of Agriculture, Faisalabad –Pakistan under randomized complete block design with three replications. Genetic variability was studied among 60 Bt. cotton genotypes using principle component analysis. Genetic advance, heritability and Coefficient of variance, for every trait were also estimated. ANOVA showed that there is significant relationship present among all genotypes which are under study. Highest genetic advance and heritability recorded for bolls and lint mass per boll, respectively. In Principal component analysis, first five PCs showed more than 1 Eigen value. In the first principal component, number of monopodial branches, seed cotton yield per plant bolls per plant and, sympodial branches were the most important traits contributing to variation that obtained about 26.15%. In PC- II, (18.87%) variation was obtained which was mainly by node number for first fruit bearing branch, node number for first effective boll formation, node height up to first fruit bearing branch, boll weight and lint mass per boll. High heritability, genetic variation and high to moderate genetic advance in different yield contributing traits indicates that germplasm can be used for further breeding to develop high yielding cultivars.

[Ali Imran, Muhammad Bilal Saleem, Aqib Anwer, Rao Saad Rehman, Muhammad Naeem, Muhammad Farhan Hafeez, Annum Javaid Shah . **Genetic Divergence of Bt. Cotton (*Gossypium hirsutum* L.) Germplasm Based on Principle Component Analysis.** *Nat Sci* 2017;15(11):114-118]. ISSN 1545-0740 (print); ISSN 2375-7167 (online). http://www.sciencepub.net/nature. 14. doi:10.7537/marsnsj151117.14.

Keywords: Environmental coefficient of variance (ECV), Genotypic coefficient of variance (GCV), Genetic advance, Heritability, Principal component analysis (PCA).

1. Introduction

Cotton is considered as the most important fiber crop of the world and also have importance as oilseed crop after soybean (Freeland *et al*., 2006). It is known as King of natural fibers (Shaukat *et al*., 2013). Cotton belongs to genus *Gossypium* and four species namely, *Gossypium. hirsutum* L., *Gossypium. barbadense* L., *Gossypium. arboreum* L., and *G. herbaceum* L., are grown for getting fiber for textile (Fryxell, 1992). Cotton production in Pakistan was 10.074 million bales (Anonymous, 2015-16).

In Pakistan, cotton yield is low in comparison to average production of other cotton growing countries. This is due to lack of resistant varieties, high temperature, cotton leaf curl virus (CLCuV) disease, pest attack and improper production technology (Panni *et al*., 2012). To overcome this situation several ways can be used including increasing inputs, pesticides use and varietal improvement and later one is most important among all. In future, sustainable production of cotton will directly depends on the development of new cotton varieties with high yield, better quality seed cotton and resistant to biotic stress and abiotic stress (Ahmad *et al*., 2012For varietal development in cotton, breeders use a small fraction of available germplasm and misleadingly assume that upland cotton has narrow genetic base (Bowman *et al*., 1996). Breeding program is successful when we have complete knowledge and understanding about the genetic diversity present in the crop germplasm, which allows the plant breeders in selection of best parental sources that will generate high yielding cultivars (Esmail *et al*., 2008). Different types of breeding methods like introduction of exotic germplasm, polyploidy and hybridization can be utilized for achieving the crop's desired genetic variability and segregating populations with diverse genotypes (Esmail *et al*., 2008).

Basic aim of study was to evaluate the genetic diversity in Bt. germplasm of upland cotton by using PCA (Principle Component Analysis). Genotypic, phenotypic, environmental variances, their coefficients of variance were considered to calculate the heritable and non-heritable part of variability. Genetic advance

and heritability were also under consideration. The knowledge thus obtained would be utilized in the improvement of existing cotton varieties or development of new cultivars.

2. Materials and Methods

Sixty locally adapted Bt cotton genotypes were gown in the experimental area of University of Agriculture, Faisalabad –Pakistan during 2015-16. Before sowing, the seeds of selected genotypes were treated with sulfuric acid to remove the fuzz. Cotton germplasm was sown in randomized complete block design with three replications. Plant to plant and row to row distances were 1 and 2.5 ft., respectively. The crop was raised to maturity with standard production practices. At the proper time, the data was recorded on morphological characters *viz* node number of first fruit bearing branch, number of first node with effective boll formation, node height up to first fruit bearing branch, plant height, monopodial branches/plant, sympodial branches/plant, bolls/plant, seed cotton yield/plant components as follows boll weight, seeds/boll, lint mass/boll, lint percentage (%), seed cotton/seed and lint/seed, from randomly selected five plants.

Data Analysis: Mean data of all the parameters were tested under analysis of variance (ANOVA) Steel *et al.* (1997). And then further analyzed by PCA (Principal component analysis) as described by Pearson and Neyman (1928) using statistical software Statistics.

ANOVA was performed in Statistics 8.1 software. The detailed procedure of all statistical data analysis has been given below:

Components of variances were estimated by the following formulas:

Genotypic variance (σ_g) = V_g = (MS_g – MS_e) / r

Environmental variance (σ_e) = V_e = (MS_e)

Phenotypic variance (σ_p) = V_p = V_g + V_e

The coefficients of variation were determined as described by Burton (1952) and revealed by Singh and Narayanan (2000).

Genotypic coefficient of variance (GCV) = (σ_g / trait mean) × 100

Environmental coefficient of variance (ECV) = (σ_e / trait mean) × 100

Phenotypic coefficient of variance (PCV) = (σ_p / trait mean) × 100

The PCV and GCV were classified as suggested by Sivasubramanian and Madhavamenon (1973) and are given below:

Low: below 10%

Moderate: 10-20%

High: more than 20%

Broad-sense heritability was estimated by the formula given by Lush (1940).

h^2 (b.s) % = (V_g / V_p) × 100

Categorization of broad-sense heritability was made according to Johnson *et al.* (1955) and has been given below.

Low: less than 30%

Moderate: 30-50%

High: more than 50%

Genetic advance and its percentage of mean was estimated by the following formula given by Johnson *et al.* (1955).

Genetic advance (G.A.) = K × {V_g / $_{(}V_{p)}^{1/2}$}

Genetic advance (G.A.) % = (Genetic Advance / Trait Mean) × 100

Where

$_{(}V_{p)}^{1/2}$ = phenotypic standard deviation

K = difference in selection having value at selection intensity of 10% is 1.76 Falconer and Mackay (1996).

The classification of percentage mean of genetic advance was made as suggested by Johnson *et al.* (1955) and is given below:

Low: below 10%

Moderate: 10-20%

High: more than 20%

3. Results and Discussion

Statistical analysis for each character showed that mean squares for all genotypes were significant, suggesting differences are present for all these characters. Earlier researchers also reported the significant differences for these characters (Imran *et al*., 2012; Iqbal *et al.*, 2013; Tang and Xiao, 2013; Baloch *et al*., 2014; Saeed *et al.*, 2014).

Genotypic, phenotypic and environmental variances including their coefficient of variances, broad sense heritability, genetic advance and percentage mean of genetic advance for all traits are given in Table 1.

For a particular trait, if phenotypic coefficient of variance is more than the genotypic coefficient of variance it shows that environmental influence is more than genetic component and vice versa. From Table 1 it is clear that highest genotypic coefficient of variance recorded for lint mass/ boll (34.37) while highest phenotypic coefficient of variance for the monopodial branches / plant (40.34). Estimates of heritability helps the breeders while making selection. However, Johnson *et al.* (1955) reported that alone heritability estimates do not give the clear idea about expected gain in the coming generation but only in conjunction with genetic advance. Genetic advance gives the magnitude of expected genetic gain obtained by one cycle of selection Idahosa *et al*. (2010). The estimation of genetic components gives higher heritability for seed cotton yield /plant (80%) and bolls/plant (83 %), respectively. The number of node having first

effective boll formation showed lower heritability about 46% as compared to other traits. Genetic advance for lint mass/ boll was recorded as highest 51% whereas the lint percentage exhibited lowest 6.9%.

Table 1. Genotypic, phenotypic, environmental variances and their coefficients, broad sense heritability, genetic advance and its mean percentage.

Trait	G.V	GCV%	P.V	PCV%	E.V	ECV%	h^2	G.A	G.A%
Node number for 1st fruiting branch	0.6925	12.094	0.988	12.572	0.296	7.912	70.02	1.226	17.81
Node height up to 1st fruiting branch	1.5258	9.332	2.772	12.578	1.246	8.435	55.03	1.604	12.11
Node number for 1st effective boll formation	1.5178	13.322	3.263	19.534	1.745	14.287	46.51	1.47	15.89
Monopodial branches/plant	0.1687	31.555	0.276	40.337	1.107	25.12	61.19	0.562	43.19
Sympodial branches/ plant	6.8687	14.54	11.737	19.006	4.868	12.24	58.52	3.508	19.46
Bolls/plant	175.95	31.459	34.098	34.373	210.05	13.85	83.77	21.246	50.38
Plant Height	239.11	13.219	311.17	15.08	72.065	7.26	76.84	23.72	20.27
Seed cotton yield/ plant	414.29	22.31	517.54	24.945	103.24	11.14	80.05	31.86	34.94
Boll weight	0.0839	12.957	0.118	15.381	0.034	8.287	70.96	0.427	19.1
Seeds per boll	3.3106	8.799	4.775	10.569	1.465	5.853	69.33	2.65	12.82
Lint mass/ boll	0.0033	34.365	0.0045	40.267	0.001	20.99	72.84	0.086	51.33
Seed cotton per seed	0.0002	13.324	0.0004	18.119	0.0002	12.281	54.065	0.019	17.14
Lint %age	3.679	4.659	5.102	5.486	1.423	2.897	72.12	2.85	6.92
Lint per seed	0.0000001	14.8192	0.0000002	18.908	0.0000001	11.744	61.42	0.0005	20.32

3.1 Principal Component Analysis

The mean data was further analyzed by using PCA through STATISTICS software. The data matrix of 14 x 60 was prepared for the analysis. Out of fourteen PCs, first five PCs shows more than 1 Eigen value. The first PC showed maximum variation (26.15%) which was mainly due to number of sympodial and monopodial branches, number of bolls/ plant and seed cotton yield/ plant. In the second PC, node number for first fruit bearing branch, node number for first effective boll formation, node height up to first fruit bearing branch, boll weight and lint mass/ boll were the most important traits contributing to variation that obtained about 18.87%. The PC-III explained (12.94%) variation of the total variation mainly by lint percentage, lint/ seed and seeds/boll. In PC-IV, node number for first effective boll formation, node height up to first fruit bearing branch and seeds/boll were the most important traits contributing to variation that obtained about (10.25%). PC-V exhibited (6.94%) variation of the total variation mainly caused by node number for first effective boll formation, node height up to first fruit bearing branch and seeds/ boll (Table II).

Table 2: Principal component for 14 characters in 60 Bt germplasm lines of *G.hirsutum* L.

Variables	PC1	PC2	PC3	PC4	PC5
Plant Height	-0.563	-0.383	0.119	0.170	0.347
Sympodial branches/ plant	-0.673	0.187	0.225	-0.079	0.487
Monopodial branches/plant	-0.609	-0.032	0.053	0.130	-0.619
Bolls/plant	-0.736	0.345	0.171	0.355	-0.028
Seed cotton yield /plant	-0.836	-0.003	0.272	0.207	0.042
Node number for 1st fruiting branch	0.434	-0.572	0.236	0.362	-0.036
Node number for 1st effective boll formation	0.340	-0.526	-0.154	0.458	0.134
Node height up to 1st fruiting branch	0.119371	-0.648	0.160	0.464	0.182
Boll weight	-0.521	-0.666	0.147	-0.344	-0.121
Lint %age	0.079	0.101	0.588	0.355	-0.400
Seeds/ boll	-0.134	-0.356	0.647	-0.514	0.071
Seed cotton/seed	-0.598	-0.436	-0.449	-0.053	-0.206
Lint/seed	-0.423	-0.457	-0.698	-0.041	-0.033
Lint mass /boll	0.359	-0.581	0.244	-0.382	-0.138

3.2 Biplot

A principal component biplot (Figure 1) showed that variables were super imposed on the plot as vectors. Distance of every variable with respect to PC1 and PC2 showed the participation of this variable in the variation. In PC1 and PC2 together bolls per palnt, seed cotton yield/ plant, boll weight, node number of first fruit bearing branch and node height up to first fruit bearing branch showed more differences as represented in biplot, while lint percentage, sympodial branches per plant, seed/boll and lint/seed had minimum differences in PC1and PC2.

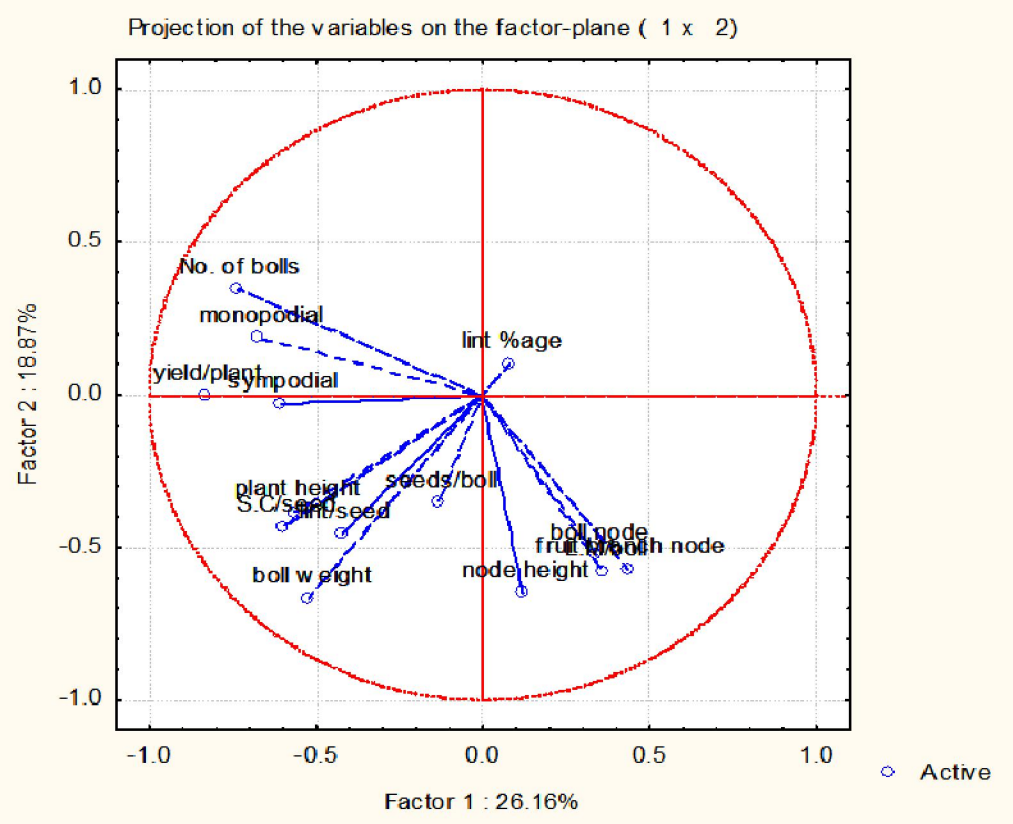

Figure 1: Principal component biplot of 60 cotton *G.hirsutum* L. genotypes.

Ahmed *et al*. (2012) used PCA to find the extent of genetic variation in cotton germplasm originated before and after 1975 which might be helpful for successful breeding program. Mohammadi and Prasanna (2003) focused on use of statistical tools and methods of genetic ranges. They analyzed that cluster analysis and PCA are the mostly employed and seemed predominantly valuable.

4. Conclusion

Results shows that the germplasm had variation for all traits and showed high broad sense heritability and moderate to high genetic advance which indicates that it might be used for the improvement of the entire earliness and yield related traits in future.

***Corresponding Author:**
Rao Saad Rehman,
Department of Plant Breeding & Genetics,
University of Agriculture, Faisalabad, Pakistan.
E-mail: saad_rehman101@hotmail.com

References

1. Ahmad, M.Q., S.H. Khan and F.M. Azhar. 2012. Decreasing level of genetic diversity in germplasm and cultivars of upland cotton (*Gossypium hirsutum* L.) in Pakistan. J. Agric. Soc. Sci. 89: 92-96.
2. Anonymous. 2015-16. Pakistan Economic Survey, Ministry of Finance, Economic Advisor's Wing, Islamabad.
3. Baloch, M.J., N. U. Khan, M. A. Rajput, W. A. Jatoi, S. Gul, I. H. Rind and N. F. Veesar. 2014. Yield related morphological measures of short duration cotton genotypes. J. Anim. Plant Sci. 24: 1198-1211.
4. Bowman, D.T., O.L. Mayand and D.S. Cahoun. 1996. Genetic base of upland cotton cultivars

released between 1970 and 1990. Crop Sci. 36: 577-581. http://dx.doi.org/10.2135/cropsci1996.0011183X003600030008x.
5. Burton, G.W. 1952. Quantitative inheritance in grasses. Proceedings of the 6th International Grassland Congress, 1: 227-283.
6. Esmail, R.M., J.F. Zhang and A.M. Abdel-Hamid. 2008. Genetic diversity in the elite cotton germplasm lines using field performance and RAPDs markers. World J. Agri. Sci. 4: 369-375.
7. Falconer, D. S., and T.F.C. Mackay. 1996. Introduction to Quantitative Genetics. Ed 4. Longmans Green, Harlow, Essex, UK.
8. Freeland, J.T.B., B. Pettigrew, P. Thaxton and G.L. Andrews. 2006. Agro meteorology and cotton production. In: (3rd Ed). Guide to agricultural meteorological practices, pp. 1-17.
9. Fryxell, P. 1992. A revised taxonomic interpretation of *Gossypium* L. (*Malvaceae*). Rheedea. 2:108-165.
10. Idahosa, D. O., J. E. Alika and A. U. Omoregie. 2010. Genetic variability, heritability and expected genetic advance as indices for yield and yield components selection in cowpea (*Vigna unguiculata* (L.) walp.
11. Imran, M., A. Shakeel, F.M. Azhar, J. Farooq, M.F. Saleem, A. Saeed, W. Nazeer, M. Naeem and A. Javaid. 2012. Combining ability analysis for within –boll yield components in upland cotton (*Gossypium hirsutum* L.). Genet. Mol. Res. 11: 2790-2800.
12. Iqbal, M., M. Naeem, M. Rizwan, W. Nazeer, M.Q. Shahid, U. Aziz, T. Aslam, and M. Ijaz. 2013. Studies of genetic variation for yield related traits in upland cotton. Am-Euras. J. Agric. & Environ. Sci., 13: 611-618.
13. Johnson, H.W, H.F Robinson and R.E Comstock. 1955. Estimation of genetic and environmental variability in soybeans. Agronomy J. 47: 314–318.
14. Lush, J.L. 1940. Intra-sire correlations or regressions of offspring on dam as a method of estimating heritability of characteristics. Proc. Am. Soc. Anim. Prod. 33: 293-301.
15. Mohammadi, S.A. and B.M. Prasanna. 2003. Analysis of genetic diversity in crop plants, salient statistical tools and considerations. Crop Sci. 43: 1235-1248.
16. Neyman, J. and Pearson, E. S. 1928. The use and interpretation of certain test criteria for purpose of statistical inference. Biometrika 20: 175-240.
17. Panni, M.K., N.U. Khan, Fitmawati, S. Batool and M. Bibi. 2012. Heterotic studies and inbreeding depression in F_2 populations of upland cotton. Pakistan J. Bot. 44: 1013-1020.
18. Saeed, F., J. Farooq, A. Mahmood, T. Hussain, M. Riaz and S. Ahmad. 2014. Genetic diversity in upland cotton for cotton leaf curl virus disease, earliness and fiber quality. Pakistan J. Agric. Res. 27: 226-236.
19. Shaukat, S., T.M. Khan, A. Shakeel and S. Ijaz. 2013. Estimation of best parents and superior cross combinations for yield and fiber quality related traits in upland cotton (*Gossypium hirsutum* L.). Sci. Tech. Dev. 32: 281-284.
20. Singh, P. and S.S. Narayanan. 2000. Biometrical Techniques in plant breeding. 2nded. Kalyani Publ. Ludhiana, India.
21. Sivasubramanian, V. and P. Madhavamenon. 1973. Path analysis of yield and yield components in rice. Madras Agric. J. 60: 1217-21.
22. Steel, R.G.D., J.H. Torrie and D.A. Dickey. 1997. Principles and procedures of statistics. A biometrical approach. Singapore: 3rd Ed. McGraw Hill Book Co. Inc.
23. Tang, F. and W. Xiao. 2013. Genetic effects and heterosis of within-boll yield components in upland cotton (*Gossypium hirsutum* L.). Euphytica 194: 41–51.

11/13/2017