Multivariate analysis for CLCuD and various morphological traits in some advanced lines of cotton *Gossypium hirsutum* L.

[1]Jehanzeb Farooq, [1]Muhammad Rizwan, [1]Muhammad Anwar, [1]Muhammad Riaz, [1]Khalid Mahmood, [2,3,4]I. Valentin Petrescu-Mag

[1] Cotton Research Institute, AARI, Faisalabad, Pakistan; [2] University of Agricultural Sciences and Veterinary Medicine (USAMV), Faculty of Agriculture, Cluj-Napoca, Romania; [3] SC Bioflux SRL, Cluj-Napoca, Romania; [4] University of Oradea, Oradea, Romania. Corresponding author: J. Farooq, jehanzeb1763@hotmail.com

Abstract. To investigate genetic diversity in some advanced lines developed specifically for cotton leaf curl virus, multivariate analysis was employed. The study of principal components revealed that out of 08 PCs 03 components exhibited eigen value >1. These three PCs contributed 76.5% of the total variability amongst the genotypes assessed for CLCuD%, yield and agronomic traits. The remaining components contributed only 23.5% towards the total variability. The traits like monopodia per plant showed considerable negative factor loadings on PC I while bolls per plant, sympodia per plant, yield and boll weight had maximum positive loadings. Maximum positive loading on PC II was shown by plant height while maximum negative loadings were exhibited by nodes to first fruiting branch. Eighteen genotypes were grouped into 3 clusters based on various traits. Cluster analysis showed that cluster 1 comprised of 7, cluster 2, 2 and cluster 3 comprised of 9 genotypes. The results of the cluster analysis revealed that members of cluster 1 and 3 may be combined to obtain more useful results both in terms of yield and CLCuD.
Key Words: Multivariate, advanced lines, cotton, cluster, sympodia.

Introduction. Upland cotton is the chief source of natural fiber worldwide. In Pakistan, it is not only grown as an important crop but also play an important role in the textile industry (Riaz et al 2013). Though Pakistan is among one of the major cotton producing country but per acre yield is low when compared to the other cotton producing countries. A number of biotic and abiotic factors contribute this low productivity. Among these factors cotton leaf curl virus disease played a major part in yield reduction (Farooq et al 2011). This disease was reported in Pakistan in 1967 near Multan, since then it severely affected major cotton growing areas of Pakistan (Hussain & Ali 1975). To eliminate this disease effort has been made to develop resistant genotypes having higher yield potential. The researchers all over Pakistan are focusing this issue for last 23 years or so. The development of better cotton genotypes having desirable traits have been obtained by hybridizing distant parental lines (Akter et al 2009). To get particular information on nature and extent of genetic variability depends upon the various procedures exploited for its estimation, like plant characterization based on agronomical, morphological and physiological traits (Bajracharya et al 2006). Multivariate analysis based on principal component analysis (PCA) and principal coordinate analyses (PCoA) are mostly exploited to measure the degree of genetic diversity among the germplasm (Brown-Guedira et al 2000). Among these biometrical procedures the main advantage of principal component analysis (PCA) is that each genotype/variety can be assigned to only one group and it also reflects the significance of largest donor to the total variability at each axis of differentiation (Sharma 1998). Genetic variation for morphological traits has been estimated using principal component analysis, which leads to the detection of phenotypic variability in cotton (Esmail et al 2008; Li et al 2008). The objective of the present

investigations was to evaluate the genetic variation among locally developed cotton germplasms specifically for resistance against CLCuD and generally for morphological traits to recognize the model genotypes that can be exploited in future breeding program.

Material and Method

Plant material and site characteristics. Eighteen genotypes were evaluated during crop season 2013-2014 at experimental area of Cotton Research Institute, Faisalabad, Punjab, Pakistan which is present at an altitude of 184 m at 31° 21′ 52″ N 72° 59′ 40″ E with average rainfall of 300 mm. The experiment was sown on 15th of May to observe their tolerance ability regarding CLCuD and some yield related parameters. The characters nodes to 1st fruiting branch, monopodia per plant, sympodia per plant, bolls per plant, plant height, boll weight, CLCuD and yield were studied.

Experimental design. For each entry, plot size measured 4.572 m × 6.096 m, comprising six rows set 75 cm apart. Distance between plants within rows was 30 cm. Normal agronomic and cultural practices (irrigation, weeding, hoeing, and fertilizer applications) were adopted as and when required.

Measurement of traits. For measuring the traits, the 10 representative and undamaged plants were selected in each line and marked for identification. Data were collected for number of first fruiting branch, monopodia per plant, sympodia per plant, bolls per plant, plant height, boll weight, CLCuD and yield.

CLCuD incidence (%). Cotton leaf curl virus disease incidence (%) and reaction of the cultivars was determined by using disease scale (Table 1) modified from the scale described by Akhtar et al (2010). Then percentage of CLCuD incidence was calculated by using the following formula.

CLCuV disease incidence (%) = sum of all disease ratings/total number of plants × 16.16

Table 1

Rating scale for cotton leaf curl virus (CLCuD) symptoms

Symptoms	Disease rating y	Disease index (%) z	Disease reaction
Absence of symptoms	0	0	Immune
Thickening of a few small veins or the presence of leaf enations on 10 or fewer leaves of a plant	1	0.1- 1	Highly resistant
Thickening of a small group of veins	2	1.1- 5	Resistant
Thickening of all veins but no leaf curling	3	5.1-10	Moderately resistant
Severe vein thickening and leaf curling on the top third of the plant	4	10.1 – 15	Moderately susceptible
Severe vein thickening and leaf curling on the half of the plant	5	15.1 – 20	Susceptible
Severe vein thickening, leaf curling, and stunting of the plant with reduced fruit production	6	>20	Highly susceptible

Statistical analysis. The average data of all the traits were subjected to basic statistics, cluster analysis and PCA using statistical software packages of SPSS version 19 and Statistica version 5.0 (Sneath & Sokal 1973). Cluster analysis was performed using K-means clustering while tree diagram based on elucidation distances was developed by

Ward's method. First two principal components were plotted against each other to find out the patterns of variability among genotypes and association between different clusters using SPSS version 19.

Results and Discussion

Principal component analysis. The basic statistics of different studied traits revealed considerable amount of variability among 18 cotton genotypes (Table 2). The conservation and exploitation of genetic resources could be made by dividing the total variance into its components. It also provides an opportunity for utilization of suitable germplasm in crop improvement for particular plant traits (Pecetti et al 1996). PCA is a powerful tool to obtain parental lines for successful breeding programmes (Akter et al 2009).

Table 2

Basic statistics for various traits of 18 cotton genotypes

Trait	Minimum	Maximum	Mean± S.E.	Std. Deviation	Variance
No. of first fruiting branch	4.730	8.200	6.555±0.248	1.051	1.105
Monopodia per plant	1.470	3.270	2.379±0.103	0.436	0.190
Sympodia per plant	16.80	31.00	24.45±1.08	4.57	20.91
Bolls per plant	18.33	35.93	28.09±1.23	5.24	27.42
Plant height	120.80	164.53	137.97±2.69	11.24	130.43
Boll weight (g)	2.70	3.900	3.240±0.0834	0.3540	0.1253
CLCuD%	0.57	22.77	6.87±1.66	7.03	49.42
Yield (kg/ha)	1650	3407	2674±127	537	288714

In this study, out of total 08, three principal components (PCs) were extracted having eigen value >1. These three PCs contributed 76.5% of the total variability amongst the genotypes assessed for CLCuD%, yield and agronomic traits (Table 3).

Table 3

(Continuation) Factor loadings by various traits

Variable	PC I	PC II	PC III
Nodes to 1st first fruiting branch	0.140	-0.458	-0.560
Monopodia per plant	-0.122	-0.164	0.775
Sympodia per plant	0.502	0.087	0.214
Bolls per plant	0.508	0.010	0.093
Plant height	0.075	0.631	-0.098
Boll weight	0.456	-0.150	0.049
CLCuD%	0.072	0.576	-0.137
Yield (kg/ha)	0.488	-0.052	0.028

However, the remaining components contributed only 23.5% towards the total diversity. The PC I contributed maximum towards the variability (43.5%) followed by PC II (18.6%) and PC III (14.4%). Saeed et al (2014) reported important contribution of first PCs in total variability while studying different traits. The traits like monopodia per plant showed considerable negative factor loadings on PC I while bolls per plant, sympodia per plant, yield and boll weight had maximum positive loadings. Maximum positive loading on PC II was shown by plant height while maximum negative loadings were exhibited by nodes to first fruiting branch. The negative loadings on PC I and PC II were shown by monopodia per plant on both components while the traits like nodes to first fruiting branch, boll weight and yield showed negative loadings only on PC II. The PC III was

elucidated by diversity among the genotypes for monopodia per plant with maximum positive and no. of first fruiting branch with maximum negative loadings. PC analysis confirmed the extent of variation for the traits among the material studied which could be utilized in designing a breeding programme aimed at improving CLCuD tolerance, boll weight, bolls per plant and ultimately seed cotton yield as it is generally assumed that maximum variation yield maximum heterotic effects (Nazir et al 2013). Malik et al (2011) and Ashokkumar & Ravikesavan (2011) mentioned the presence of ample amount of variation in colored cotton genotypes offer huge scope for characterization of colored cotton genotypes. A PC biplot showed that variables and genotypes are super imposed on the plot as vectors (Figure 1). Distance of each variable with respect to PC 1 and PC 2 showed the contribution of these variables in the variation of genotypes used. The findings are in accordance with the findings of Saeed et al 2014 who found major role of first two components while studying different genotypes. The biplot exhibited that as whole plant height, CLCuD, monopodia per plant and nodes to 1st fruiting branch contributed maximum towards variation in the genotypes studied. The scatter plot of yield vs CLCuD showed that the genotypes FH 323, MNH 868, FH 142, FH 332, FH 334 and FH 337 showed minimum values for CLCuD (Figure 2).
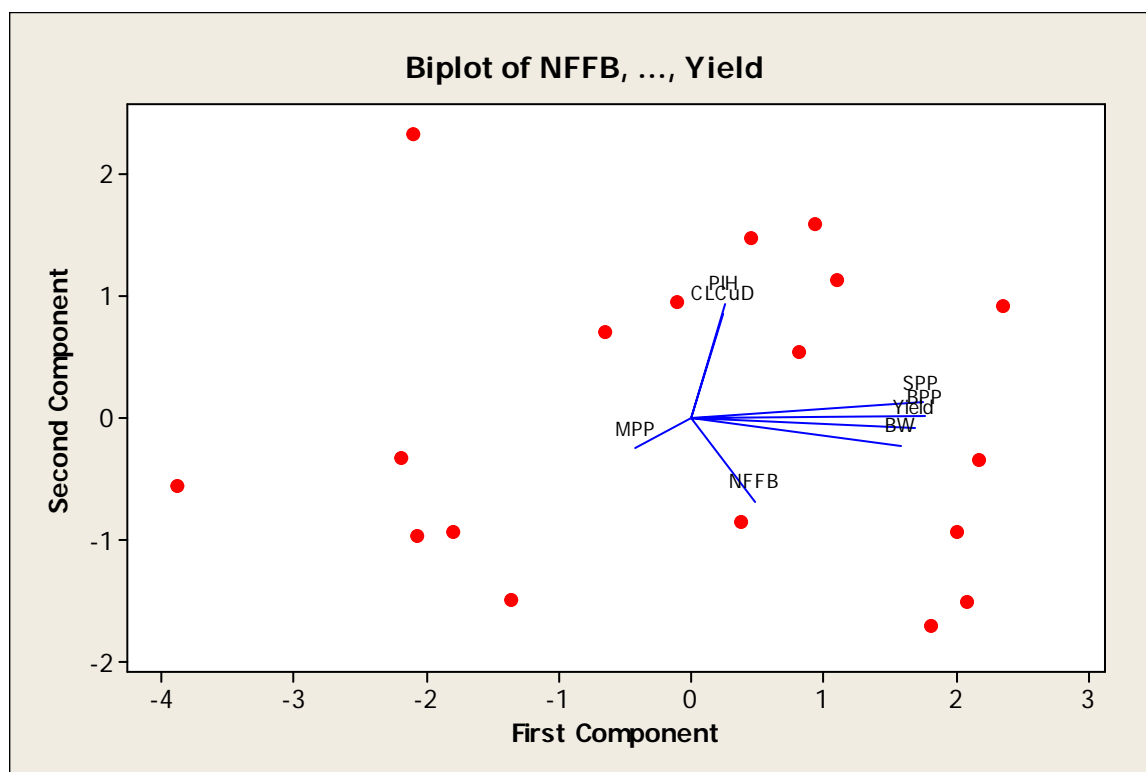

Figure 1. Biplot between PC-1 and 2 showing contribution of various traits in variability among different genotypes.
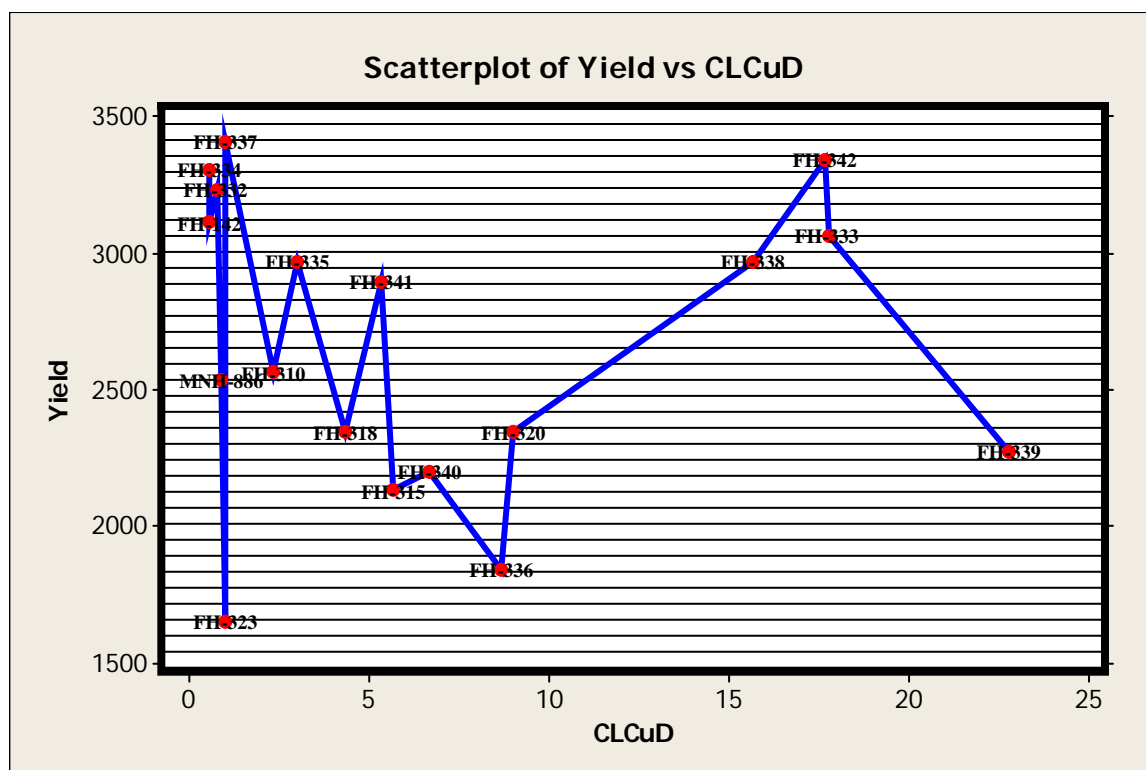
Figure 2. Scatter plot of yield Vs CLCuD of the genotypes studied.

Cluster analysis. This is another useful statistical procedure to obtain genotypes from various clusters having desirable traits. Eighteen genotypes were grouped into 3 clusters based on various traits (Table 4).

Table 4

Cluster analysis of various traits in different genotypes of cotton

Variable	Cluster-1	Cluster-2	Cluster-3
Nodes to 1st first fruiting branch	6.73	6.40	6.30
Monopodia per plant	2.33	2.44	2.37
Sympodia per plant	27.89	22.04	17.40
Bolls per plant	31.87	25.54	20.00
Plant height (cm)	141.18	136.45	128.84
Boll weight (g)	3.43	3.11	2.84
CLCuD%	6.93	7.39	4.84
Yield kg/ha	3140.19	2339.81	1745.34

Cluster analysis showed that cluster 1 comprised, 7; cluster 2, 2; and cluster 3, 9; genotypes (Table 5). The genotypes in cluster 1 only showed reasonable values of plant height, no. of bolls per plant, sympodia per plant. Amurrio et al (1995) and Rabbani et al (1998) reported lack of relationship between various clusters based on agronomic traits and origins of genotype in peas (*Pisum sativum*) and mustard (*Brassica juncea*) respectively. Similarly wide variations in clusters have been reported by Nazir et al (2013). The occurrence of wide variation between the clusters is of great genetic value in providing genotypes aimed at cotton selection for adaptation to CLCuD hit areas. Similar kind of results associated to germplasm grouping has been reported (Grenier et al 2000; Saeed et al 2014). The results of the cluster analysis revealed that the genotypes in

cluster 1 and 3 may be combined to get practical benefits as genotypes in cluster 1 had more high yielding genotypes and in cluster 3 the strains showed better CLCuD tolerance. Both principle component and cluster analysis provided excellent information which may be exploited to obtain useful results.

Table 5

Cluster membership of various studied genotypes

Cluster-1	7	FH-310, FH-315, FH-318, FH-320, MNH-886, FH-339, FH-340
Cluster-2	2	FH-323, FH-336
Cluster-3	9	FH-332, FH-333, FH-334, FH-142, FH-335, FH-337, FH-338, FH-341, FH-342

Conclusions. The results of the cluster analysis revealed that the genotypes in cluster 1 and 3 may be combined to get practical benefits as genotypes, in cluster 1 had more high yielding genotypes and in cluster 3 the strains showed better CLCuD tolerance. Both principle component and cluster analysis provided excellent information which may be exploited to obtain useful results.

References

Akhtar K. P., Haider S., Khan M. K. R., Ahmad M., Sarwar N., Murtaza M. A., Aslam M., 2010 Evaluation of gossypium species for resistance to leaf curl Burewala virus. Ann Appl Biol 157:135-147.

Akter A., Hasan M. J., Paul A. K., Mutlib M. M., Hossain M. K., 2009 Selection of parent for improvement of restorer line in rice (*Oryza sativa* L.). SAARC J Agric 7:43-50.

Amurrio J. M., de Ron A. M., Zeven A. C., 1995 Numerical taxonomy of Iberian pea landraces based on quantitative and qualitative characters. Euphytica 82:195-205.

Ashokkumar K., Ravikesavan R., 2011 Morphological diversity and per se performance in upland cotton (*Gossypium hirsutum* L.). J Agric Sci 3:107-113.

Bajracharya J., Steele K. A., Jarvis D. I., Sthapit B. R., Witcombe J. R., 2006 Rice landrace diversity in Nepal: Variability of agro-morphological traits and SSR markers in landraces from a high-altitude site. Field Crops Res 95:327-335.

Brown-Guedira G. L., Thompson J. A., Nelson R. L., Warburton M. L., 2000 Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. Crop Sci 40:815-823.

Esmail R. M., Zhang J. F., Abdel-Hamid A. M., 2008 Genetic diversity in elite cotton germplasm lines using field performance and RAPD markers. World J Agric Sci 4:369-375.

Farooq A., Farooq J., Mahmood A., Batool A., Rehman A., Shakeel A., Riaz M., Shahid M. T. H., Mehboob S., 2011 An overview of cotton leaf curl virus disease (CLCuD) a serious threat to cotton Productivity. Aust J Crop Sci 5(12):1823-1831.

Grenier C., Deu M., Kresovich S., Bramel-Cox P. J., Hamon P., 2000 Assessment of genetic diversity in three subsets constituted from the ICRISAT sorghum collection using random vs. nonrandom sampling procedures B. Using molecular markers. Theor App Genet 101:197-202.

Hussain T., Ali M., 1975 A review of cotton diseases in Pakistan. Pak Cottons 19:71-86.

Li Z., Wang X., Yan Z., Guiyin Z., Wu L., Jina C., Ma Z., 2008 Assessment of genetic diversity in glandless cotton germplasm resources by using agronomic traits and molecular markers. Front Agric China 2:245-252.

Malik W., Iqbal M. Z., Khan A. A., Noor E., Qayyum A., Hanif M., 2011 Genetic basis of variation for seedling traits in *Gossypium hirsutum* L. Afr J. Biotechnol 10:1099-1105.

Nazir A., Farooq J., Mahmood A., Shahid M., Riaz M., 2013 Estimation of genetic diversity for CLCuV, earliness and fiber quality traits using various statistical procedures in different crosses of Gossypium hirsutum L. Vestnik Orel Gau 43:2–9.

Pecetti L., Annicchiario P., Damania A. B., 1996 Geographic variation in tetraploid wheat (*Triticum turgidum* spp. *Turgidum* convar. Durum) landraces from two provinces in Ethiopia. Genet Resour Crop Evol 43:395-407.

Rabbani M. A., Iwabuchi A., Murakami Y., Suzuki T., Takayanagi K., 1998 Phenotypic variation and the relationships among mustard (*Brassica juncea* L.) germplasm from Pakistan. Euphytica 101:357-366.

Riaz M., Farooq J., Sakhawat G., Mahmood A., Sadiq M. A., Yaseen M., 2013 Genotypic variability for root/shoot parameters under water stress in some advance lines of cotton (*Gossypium hirsutum* L.). Genet Mol Res 12(1):552-561.

Saeed F., Farooq J., Mahmood A., Riaz M., Hussain T., Majeed A., 2014 Assessment of genetic diversity for Cotton leaf curl virus (CLCuD), fiber quality and some morphological traits using different statistical procedures in *Gossypium hirsutum* L. Aust J Crop Sci 8(3):442-447.

Sharma J. R., 1998 Statistical and biometrical techniques in plant breeding. New Age International (P) Limited Publishers, New Delhi, 432 pp.

Sneath P. H. A., Sokal R. R., 1973 Numerical taxonomy: the principles and practice of numerical classification. W. F. Freeman & Co., San Francisco, 573 pp.

Authors:
Jehanzeb Farooq, Cotton Research Institute, AARI, Pakistan, Faisalabad, 38040, e-mail: jehanzeb1763@hotmail.com
Muhammad Rizwan, Cotton Research Institute, AARI, Pakistan, Faisalabad, 38040, e-mail: muhammadrizwan29@gmail.com
Muhammad Anwar, Cotton Research Institute, AARI, Pakistan, Faisalabad, 38040, e-mail: manwar196@yahoo.com
Muhammad Riaz, Cotton Research Institute, AARI, Pakistan, Faisalabad, 38040, e-mail: riaz_1964@yahoo.com
Khalid Mahmood, Cotton Research Institute, AARI, Pakistan, Faisalabad, 38040, e-mail: crifsd@gmail.com
I. Valentin Petrescu-Mag, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, Faculty of Agriculture, Department of Environment and Plants Protection, Romania, Cluj, Cluj-Napoca 400372, 3-5 Calea Mănăştur Street; SC Bioflux SRL, Romania, Cluj, Cluj-Napoca 400488, 54 Ceahlău Street; University of Oradea, Romania, Oradea 410087, 1 Universităţii Street, e-mail: zoobiomag2004@yahoo.com