Journal of AGROBIOLOGY

J Agrobiol **30**(2): 107–115, 2013
DOI 10.2478/agro-2013-0011
ISSN 1803-4403 (printed)
ISSN 1804-2686 (on-line)
http://joa.zf.jcu.cz; http://versita.com/science/agriculture/joa

ORIGINAL ARTICLE

Estimation of genetic variability for Cotton leaf curl virus disease, fiber quality and some morphological traits using multivariate analysis in exotic *Gossypium* (Diploid and tetraploid) species

Frasat Saeed[1], Jehanzeb Farooq[1], Abid Mahmood[1], Muhammad Riaz[1], Tassawar Hussain[2], Abdul Majid[3]

[1]*Cotton Research Institute, Faisalabad, Pakistan*
[2]*Consultant (PAK-US Cotton Project) in International Center for Agricultural Research in Dry Areas (ICARDA), Pakistan Office*
[3]*Country manager in International Center for Agricultural Research in Dry Areas (ICARDA), Pakistan Office*

Received: 2nd January 2014
Revised: 1st October 2014
Published online: 30th March 2015

Abstract
This study was designed to evaluate the germplasm imported from USA for cotton leaf curl virus, fiber quality and yield components. Seventy nine cotton genotypes were evaluated using statistical procedures such as correlation analysis, cluster and principle components (PC). These evaluation procedures would result in obtaining parents that possess better tolerance against cotton leaf curl virus and having better fiber quality. Cotton leaf curl virus exhibited a significant negative association with plant height, monopodial and sympodial branches, and a significant positive correlation with fiber fineness but with other traits the relationship was non-significant. The number of days taken to first square has a significant positive correlation with days taken to first flower but a significant negative correlation with GOT and monopodial branches. Principal component (PC) analysis showed the first 4 PCs as having an Eigen value >1 explaining 65.4% of the total variation with CLCuD, plant height and sympodial branches, being the most important characters in PC1. Cluster analysis classified 79 accessions into four divergent groups. The genotypes in cluster 1 showed reasonable values of GOT, staple length and plant height but for other traits selection cannot be made in this cluster. Similarly,

✉ Frasat Saeed and Jehanzeb Farooq, Cotton Research Institute, Faisalabad, Pakistan
jehanzeb1763@hotmail.com
frasat_1436@yahoo.com

Abbreviations:
CLCuD, cotton leaf curl virus percentage; **DTFS**, days to 1st square; **DTFF**, days to 1st flower; **GOT**, ginning out turn; **SL**, staple length; **FF**, fiber fineness; **PH**, plant height; **NTFFB**, nodes to 1st fruiting branch; **MPP**, monopodia per plant; **SPP**, sympodia per plant; **BW**, boll weight

the 2nd cluster was comprised of genotypes having low values of fiber fineness and CLCuD and higher values for staple length. The members of the 3rd cluster were characterized by least values of days taken to first square and plant height and more staple length. Cluster 4 was characterized by maximum staple length and boll weight but minimum CLCuD. The genotypes in cluster 1 and 3 may be combined to obtain desirable traits related to earlier and better disease tolerance. Scatter plot and tree diagrams demonstrated sufficient diversity among the cotton accession for various traits and some extent of association between different clusters. The results indicated that the diversity among the genotypes could be utilized for the development of CLCuD resistant varieties with higher seed cotton yield and better fiber quality.

Key words: cotton leaf curl virus; germplasm; biplot; cluster analysis

INTRODUCTION

The cotton leaf curl virus is the most destructive disease of cotton crops and affects a vast area of cotton in Pakistan. It was first reported in Nigeria in 1912 (Farquharson 1912), later in Tanzania (Jones and Mason 1926) and in Sudan (Bailey 1934). In Pakistan, this disease was reported in 1967 near the Multan region (Hussain and Ali 1975). The virus infected plants may show a variable range of symptoms depending on the intensity of the disease; characteristic symptoms include yellowing and thickening of the small veins on the lower surface of young leaves and thickening of veins sometimes more pronounced in the form of upward and downward curling of leaves (Farooq et al. 2011). Under severe attack the infected plants sometimes develop leaf enations (oval or cuplike foliar worth) on the underside of the leaf, and plants become stunted and significant reduction in yield occurs (Ahmed et al. 2010). Currently institutes like the Cotton Research Institute, Faisalabad, Nuclear Institute for Agriculture and Biology and Central Cotton Research Institute, Multan etc. in Pakistan are working on the development of CLCuD resistant varieties. One major project namely "Enhancing Cotton Germplasm, improving resistance to CLCuD and supporting cotton and best management practices for small Farmers" funded by USDA (The United States Department of Agriculture) is running at the Cotton Research Institute Faisalabad and at different research institutes of Pakistan. For this project, a lot of germplasm was imported from United States for screening against CLCuD.

Multivariate analysis based on Mahalonobis's D^2 statistics (MDS), principal component analysis (PCA) and principal coordinate analysis (PCoA) are mostly used to evaluate the magnitude of genetic diversity in the germplasm (Brown-Guedira 2000). Among biometrical procedures the main advantage of principal component analysis (PCA) is that each genotype can be assigned to only one group and it also reflects the importance of the largest contributor to the total variation at each axis of differentiation (Sharma 1998). Genetic variation for morphological traits has been estimated using principal component analysis, which leads to the recognition of phenotypic variability in cotton (Esmail et al. 2008, Li et al. 2008). The objective of the present research was to evaluate the genetic diversity among cotton germplasms specifically for resistance against CLCuD and generally for morphological and fiber traits to identify the ideal genotypes that can be exploited in future breeding programs.

MATERIALS AND METHODS

The experiment was conducted at the Cotton Research Institute, Faisalabad, Punjab, Pakistan. Seventy six genotypes of *G. hirsutum*, two genotypes of *G. arboretum* (AKA-8401, A2-101) and one genotype of *G. herbaceum* (A1-03) imported from United States of America were tested during these studies. The studies were carried out during the cropping seasons 2012–13 and sowing of the experiment was done on 19th of June. The experiment was sown late so that the genotypes would receive virus infestation, as during late sowing the virus infestation is more severe due to frequent attack of white fly (*Bemisia tabaci* G.). For each entry, the plot size measured 6.09 × 0.763 m, comprising 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. For measuring the traits 10 representative, undamaged plants were selected in each line and marked for identification. The number of Nodes to 1st fruiting branch counted from zero node (cotyledonary node) to the node at which the first flower was appeared. Data on plant height in centimeters were recorded from the base of the plant to the tip of the plant. Data on monopodia and sympodia were taken by counting the number of vegetative and fruiting branches. Boll weight was calculated by averaging the weight of five bolls. Cleaned and dry samples of seed cotton were weighed and then ginned separately with a single roller electric ginning machine. The lint obtained from each sample was weighed and ginning outturn % was calculated by the following formula

$$\text{Ginning outturn (\%)} = \text{Weight of lint} / \text{Weight of seed cotton} \times 100$$

Fiber characteristics such as staple length, and the fiber fineness of each guarded plant were measured by using spin lab HVI-900. CLCuD (%) and the reaction of the cultivars was determined using the disease scale described by Akhtar et al. (2010) and Farooq et al. (2011). The disease scale is given in Table 1. The % age of CLCuD incidence was calculated by using the following formula;

CLCuV disease incidence (%) = Sum of all disease ratings/total number of plants × 25

Statistical analysis

The average was subjected to basic statistics, correlation analysis, cluster analysis and principal component analysis (PCA) using statistical software packages of SPSS version 19 and STATISTICA version 5.0 (Sneath and Sokal 1973). Cluster analysis was performed using K-means clustering while a 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.

Table 1. Rating scale for cotton leaf curl virus disease (CLCuD) symptoms

Symptoms	Disease rating	Disease index (%)	Disease reaction
No symptom	0	0	Immune
Thickening of only secondary and tertiary veins	1	0.1–10	Highly tolerant
Thickening of tertiary veins, secondary and primary veins	2	10–30	Tolerant
Vein thickening, leaf curl or enation or both	3	30–50	Susceptible
Stunting alone with, vein thickening leaf curl/enation	4	>50	Highly susceptible

(Sum of all disease ratings/total # of plants) × 25

RESULTS AND DISCUSSIONS

The basic statistics of the traits studied demonstrated considerable variability among the 79 USA cotton genotypes (Table 2). Simple correlation coefficients revealed some significant associations among 10 traits (Table 3). The correlation between different traits is an important feature for the instigation of any breeding program as it offers probabilities for the selection of genotypes having desirable traits simultaneously (Ali et al. 2009). Cotton leaf curl virus disease exhibited a negative and significant association with plant height, monopodial and sympodial branches but a significant and positive correlation with fiber fineness although with other traits it showed non-significant results. However, Farooq et al. (2013) found a negative association of CLCuD with GOT%. Days taken to first square has a significant and positive correlation with days taken to first flower but a significant and negative correlation with GOT and monopodial branches. The significant positive association between 1st flower and square under virus intensive conditions was found in the studies of Farooq et al. (2013). Days taken to first flower showed a significant and negative correlation with GOT. The GOT had a significant positive correlation with fiber fineness but a significant and negative correlation with nodes to first monopodia. For fiber fineness, a significant and negative correlation was observed with plant height. Plant height showed a positive association with sympodia per plant and monopodia per plant exhibited significant and positive association with sympodia per plant. Farooq et al. (2013) found a positive correlation

among yield contributing traits. Similarly, performance and a positive association of seed cotton yield and its components was observed in *G. hirsutum* cultivars (Méndez-Natera et al. 2012). These associations may be considered while selecting parents for future breeding programmes especially under late sown conditions.

Table 2. Basic statistics for various traits of 79 cotton genotypes

Variable	Mean	SE Mean	StDev	Variance	Coef Var	Minimum	Maximum
CLCuD (%)	65.71	2.26	20.07	402.67	30.54	0.00	92.00
DTFS	49.519	0.691	6.143	37.740	12.41	38.000	61.000
DTFF	66.975	0.686	6.100	37.204	9.11	53.000	76.000
GOT (%)	34.152	0.375	3.336	11.130	9.77	24.000	40.000
SL (mm)	27.176	0.0849	0.755	0.570	2.78	24.000	29.000
FF (µg/inch)	4.7582	0.0593	0.5274	0.2781	11.08	2.9000	5.8000
PH (cm)	56.65	1.77	15.78	248.87	27.85	35.00	151.00
NTFFB	5.924	0.114	1.010	1.020	17.05	3.000	9.000
MPP	1.9114	0.0904	0.8037	0.6459	42.05	1.0000	4.0000
SPP	9.532	0.388	3.445	11.868	36.14	5.000	26.000
BW (g)	2.7291	0.0609	0.5409	0.2926	19.82	1.4000	4.5000

Table 3. Simple correlation coefficients of various quality and morphological traits in cotton

	CLCuD (%)	DTFS	DTFF	GOT (%)	SL (mm)	FF (µg/inch)	PH (cm)	NTFFB	MPP	SPP
DTFS	–0.142									
DTFF	–0.093	0.938*								
GOT	–0.010	–0.198*	–0.243*							
SL	–0.031	0.002	–0.065	0.166						
FF	0.204*	–0.143	–0.153	0.302*	0.095					
PH	–0.553*	0.033	0.040	–0.101	–0.015	–0.273*				
NTFFB	–0.011	–0.118	–0.065	–0.213*	0.050	–0.114	0.051			
MPP	–0.220*	–0.214*	–0.181	0.077	–0.061	0.049	0.128	0.055		
SPP	–0.556*	–0.105	–0.090	0.055	–0.020	–0.057	0.697*	0.071	0.365*	
BW	0.035	–0.047	–0.036	0.080	0.088	–0.082	0.078	0.180	0.333	0.034

* Statistically significant correlation at P≤0.05 and 0.01.

Principle component analysis

The conservation and exploitation of genetic resources could be made by dividing the total variance into a number of components. It also provides an opportunity for the utilization of suitable germplasm in crop improvement for specific plant traits (Pecetti and Damania 1996). The principal components analysis is a very reliable tool for obtaining parental lines for successful breeding programs (Nazir et al. 2013).

In this study, out of a total of 11, four principal components (PCs) having an Eigen value of >1 were extracted. These four PCs contributed 65.4% of the total variability amongst the cotton genotypes assessed for CLCuD and other yield related traits (Table 4a). However, the remaining components contributed only 34.6% towards the total diversity for this set of cotton genotypes. The PC I contributed most towards the variability (21.9%) followed by PC II (20.3%), PC III (12.4%), and PC IV (10.7%). Nazir et al. (2013) studying various yield related traits found the contribution of first two PCs important in the total variation. Traits like plant height,

monopodial branches and sympodial branches showed considerable positive factor loadings on PC I while CLCuD had maximum negative loadings (Table 4b). The 2nd PC was related to diversity among cotton genotypes due to GOT, monopodial branches and fiber fineness. The PC III was explained by variation among genotypes due to CLCuD, the number of nodes to the 1st monopodia and boll weight with their positive loadings and negative loadings exhibited by GOT and fiber fineness. The PC IV was elucidated by diversity among the genotypes for the number of monopodial branches with positive loadings but the staple length exhibited negative loadings. PC analysis confirmed the magnitude of variability for the traits among the material studied which could be manifested in designing a hybridization program aimed at improving CLCuD tolerance, fiber quality, earliness and ultimately seed cotton yield, as it is generally assumed that maximum variability can yield maximum heterotic effects (Nazir et al. 2013). Malik et al. (2011) and Ashokkumar and Ravikesavan (2011) observed that the presence of a sufficient amount of variability in colored cotton genotypes offer an enormous capacity for characterization of colored cotton genotypes. A PC biplot Fig. 1 showed that variables and genotypes are super imposed on the plot as vectors. The 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 results are in agreement with the result obtained in the studies of Nazir et al. (2013). The biplot showed that whole days to 1st square, flower, sympodia per plant and monopodia per plant contributed the most towards variability in the germplasms studied.

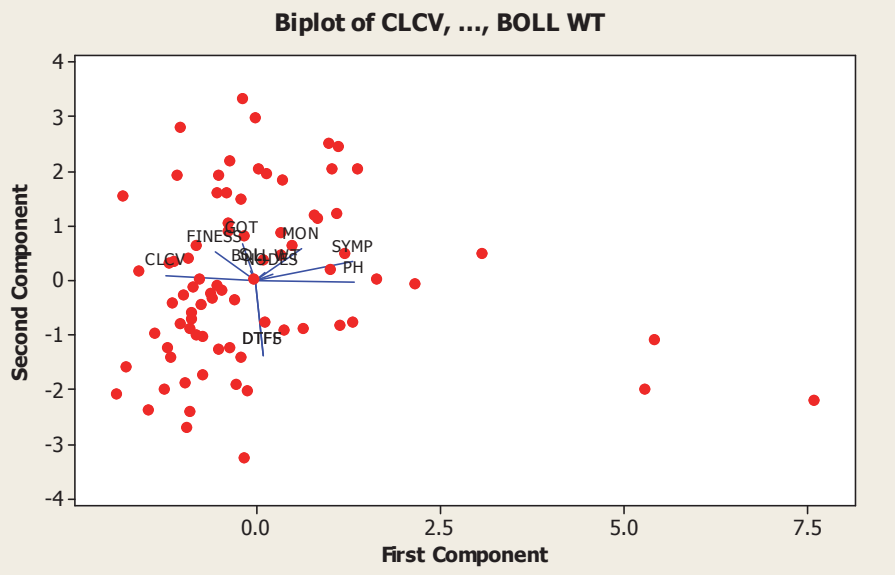

Fig. 1. Biplot between PC-1 and PC-2 showing contribution of various traits in variability of germplasm

Cluster analysis

Seventy nine cotton genotypes were grouped into 4 clusters based on various traits (Table 5). Cluster analysis (Fig. 2) showed that cluster 1 comprised 22 genotypes, cluster 2, 1 while cluster 3 had 54 and cluster 4 contained 2 genotypes (Table 6). The genotypes in cluster 1 showed reasonable values of GOT, staple length and plant height but for other traits, selection cannot be made in this cluster. Similarly, the 2nd cluster was composed of genotypes having low values of fiber fineness and CLCuD and higher values for staple length (Table 5). The members of the 3rd cluster were characterized by least numbers of days taken to first square and plant height and more staple length. Cluster 4 is characterized by maximum staple length and boll weight but minimum CLCuD. The researchers Amurrio et al. (1995) and Rabbani et al. (1998) found a lack of association among different clusters based on agronomic traits and origin of genotypes in peas (*Pisum sativum*) and mustard (*Brassica juncea*)

respectively. Similarly, extensive variations in clusters have been reported by Nazir et al. (2013). The presence of wide variation between the clusters is of great genetic value in obtaining genotypes aimed at cotton selection for adaptation to CLCuD hit areas. Similar results associated with germplasms grouping have been reported by Ayana and Bekele (1998) and Grenier et al. (2001).

Table 4a. Principle component analysis of different quality and morphological traits in cotton

	PC I	PC II	PC III	PC IV
Eigen value	2.4101	2.2370	1.3665	1.1779
% of total variance	21.9	20.3	12.4	10.7
Cumulative variance %	21.9	42.2	54.7	65.4

Table 4b. Factor loadings by various traits

Variable	PC1	PC2	PC3	PC4
CLCuD (%)	–0.509	0.043	0.192	0.034
DTFS	0.037	–0.610	–0.223	–0.150
DTFF	0.041	–0.613	–0.165	–0.110
GOT (%)	–0.077	0.305	–0.495	–0.278
SL (mm)	–0.033	0.081	–0.155	–0.668
FF (µg/inch)	–0.231	0.231	–0.410	–0.050
PH (cm)	0.551	–0.014	–0.013	–0.037
NTFFB	0.092	0.049	0.591	–0.242
MPP	0.254	0.259	–0.060	0.158
SPP	0.546	0.152	–0.133	0.011
BW (g)	0.053	0.070	0.280	–0.594

Table 5. Calculation of means for each cluster depending on trait variability

Traits	Clusters			
	Cluster 1	*Cluster 2*	*Cluster 3*	*Cluster 4*
CLCuD (%)	18.00	0.00	88.00	0.00
DTFS	45.00	58.00	39.00	58.00
DTFF	60.00	75.00	55.00	73.00
GOT (%)	38.00	29.00	34.00	31.00
SL (mm)	27.00	27.20	27.50	28.00
FF (µg/inch)	4.80	4.30	5.30	3.00
PH (cm)	57.00	151.00	47.00	101.00
NTFFB	5.00	5.00	5.00	5.00
MPP	1.00	1.00	4.00	3.00
SPP	8.00	26.00	13.00	15.00
BW (g)	1.80	2.30	2.60	3.20

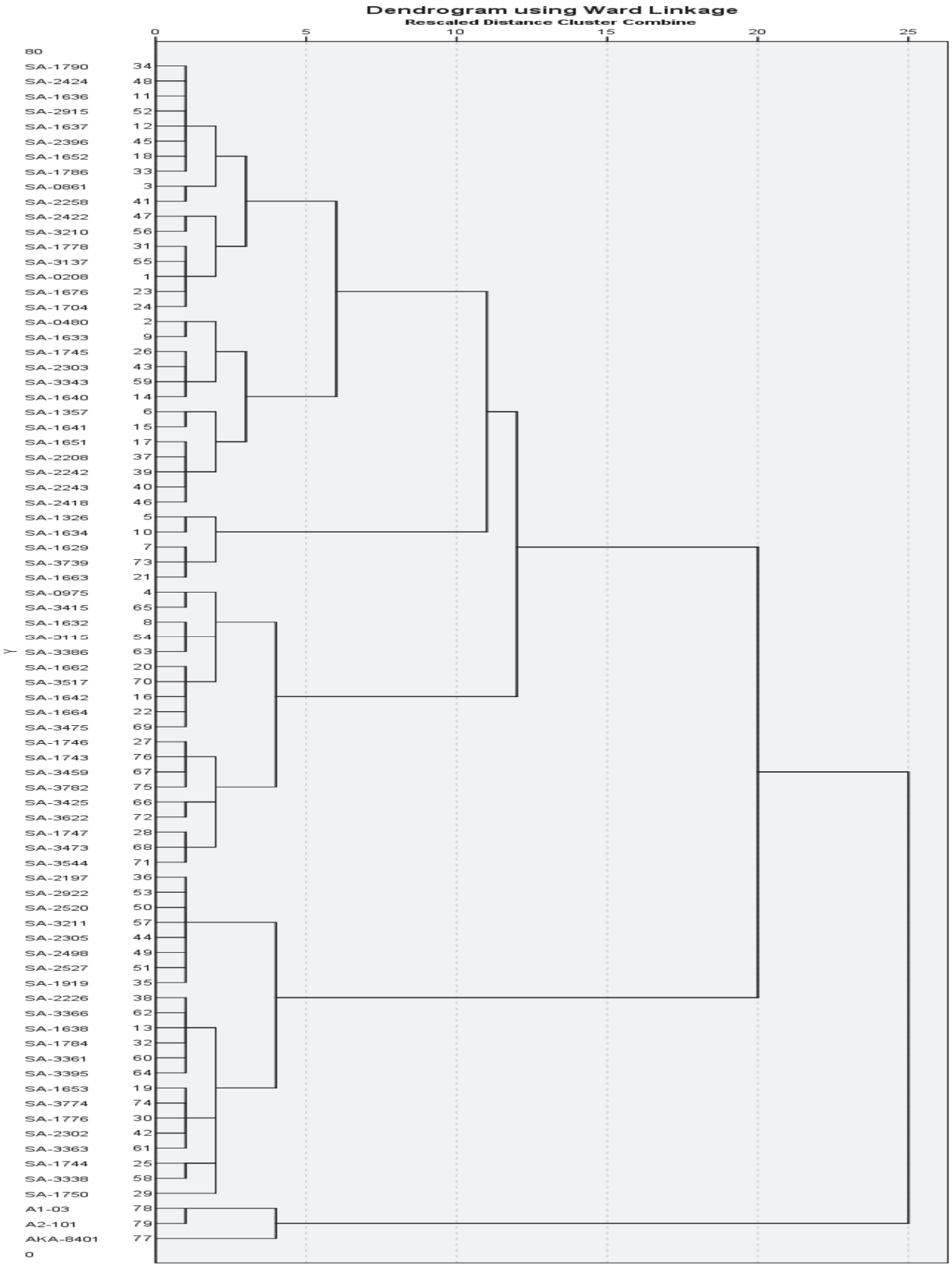

Fig. 2. Tree diagram of 79 cotton genotypes based on different fiber quality and morphological traits

Table 6. Cluster membership of various genotypes

Cluster 1	22	SA-0480,SA-0975,SA-1326,SA-1357,SA-1629,SA-1632,SA-1633,SA-1634,SA-1640,SA-1641,SA-1651, SA-1663,SA-1745,SA-2242,SA-2303,SA-2418,SA-3343,SA-3415,SA-3425,SA-3544,SA-3622,SA-3739
Cluster 2	1	AKA-8401
Cluster 3	54	SA-0208,SA-0861,SA-1636,SA-1637,SA-1638,SA-1642,SA-1652,SA-1653,SA-1662, SA-1664,SA-1676,SA-1704,SA-1744,SA-1746,SA-1747,SA-1750,SA-1776,SA-1778,SA-1784, SA-1786,SA-1790,SA-1919,SA-2197,SA-2208,SA-2226,SA-2243,SA-2258,SA-2302,SA-2305,SA-2396, SA-2422,SA-2424,SA-2498,SA-2520,SA-2527,SA-2915,SA-2922 ,SA-3115,SA-3137,SA-3210, SA-3211,SA-3338,SA-3361,SA-3363,SA-3366,SA-3386,SA-3395,SA-3459,SA-3473 ,SA-3475, SA-3517,SA-3774,SA-3782,SA-1743
Cluster 4	2	A1-03,A2-101

The use of various statistical procedures made it possible to recognize the genotypes having tolerance to CLCuD, better fiber quality and possessing earliness. Useful correlations and information generated from cluster and PC analysis will be helpful in designing breeding programmes aimed at obtaining high yielding genotypes possessing a high degree of CLCuD tolerance and showing better fiber quality.

ACKNOWLEDGEMENT

This material is based upon work supported by the US Department of Agricultural Research Service under agreement No. 58-6402-178F. Any opinion, findings, conclusions or recommendations expressed in this publication are those of author (s) and do not necessarily reflect the views of US Department of Agriculture.

REFERENCES

Ahmed S, Islam NU, Mahmood A, Ashraf F, Hayat K, Hanif M (2010): Screening of cotton germplasm against cotton leaf curl virus. Pak J Bot 42: 3327–3342.

Akhtar KP, Haider S, Khan MKR, Ahmad M, Sarwar N, Murtaza MA, Aslam M (2010): Evaluation of gossypium species for resistance to leaf curl burewala virus. Ann Appl Biol 157: 135–147.

Ali MA, Nawab NN, Abbas A, Zulkiffal M, Sajjad M (2009): Evaluation of selection criteria in *Cicer arietinum* L. using correlation coefficients and path analysis. Aust J Crop Sci. 3: 65–70.

Amurrio JM, de Ron AM, Zeven AC (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.

Ayana A, Bekele E (1999): Multivariate analysis of sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. Genet Resour Crop Evol 46: 273–284.

Bailey MA (1934): Sudan empire cotton. Gr Rev 11: 280–288.

Brown-Guedira GL (2000): Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. Crop Sci 40: 815–823.

Esmail RM, Zhang JF, Abdel-Hamid AM (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 MTH, Mehboob S (2011): An overview of cotton leaf curl virus disease (CLCuD) a serious threat to cotton productivity. Aust J Crop Sci 5: 1823–1831.

Farooq J, Anwar M, Riaz M, Mahmood A, Farooq A, Iqbal MS, Iqbal MS (2013): Association and path analysis of earliness, yield and fiber related traits under cotton leaf curl virus (CLCuV) intensive conditions in *G. hirsutum* L. Plant Knowledge J 2: 43–50.

Farquharson CO (1912): A report of the mycologist. A report Agric. Deptt. Nigeria. In Siddique MA, Hungus LC (eds.). Cotton growth in Gezira environment.W Haffer and Sons Ltd. Cambridge England.

Grenier C, Deu M, Kresovich S, Bramel, Cox PJ, Hamon P (2001): Assessment of genetic diversity in three subsets constituated from the ICRISAT sorgham collection using random vs. non-random sampling procedures by 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.
Jones GH, Mason TG (1926): On two obscure diseases of cotton. Ann Bot 160: 759–772.
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 MZ, Khan AA, Noor E, Qayyum A, Hanif M (2011): Genetic basis of variation for seedling traits in *Gossypium hirsutum* L. Afr J Biotechnol 10: 1099–1105.
Méndez-Natera, JR, Rondón A, Hernández J, Merazo-Pinto JF (2012): Genetic studies in upland cotton. III. Genetic parameters, correlation and path analysis. Sabrao J Breeding Genet 44: 112–128.
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, Damania AB (1996): Geographic variation in tetraploid wheat (*Triticum turgidum* spp. *turgidum* con var. *durum*) landraces from two provinces in Ethiopia. Genet Resour Crop Ev 43: 395–407.
Rabbani MA, 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.
Sharma JR (1998): Statistical and biometrical techniques in plant breeding. New Age International (P) Limited Publishers, New Delhi.
Sneath PHA, Sokal RR (1973): Numerical Taxonomy: The principles and practice of numerical classification. W. F. Free-Man and Co. San Francisco.