GMR

Assessment of genetic diversity of cotton genotypes for various economic traits against cotton leaf curl disease (CLCuD)

M. Javed[1], S.B. Hussain[2] and M. Baber[2]

[1]Department of Plant Breeding & Genetics,
Faculty of Agricultural Sciences & Technology, Bahauddin Zakariya University, Multan, Pakistan
[2]Institute of Molecular Biology & Biotechnology, Bahauddin Zakariya University, Multan, Pakistan

Corresponding author: M. Javed
E-mail: javedhassan@student.bzu.edu.pk / ch.javedhassan@gmail.com

Genet. Mol. Res. 16 (1): gmr16019446
Received October 18, 2016
Accepted December 13, 2016
Published February 8, 2017
DOI http://dx.doi.org/10.4238/gmr16019446

ABSTRACT. In Pakistan, cotton crop has been under enormous threat of cotton leaf curl disease (CLCuD) over the last four decades. In order to estimate genetic diversity in cotton germplasm CLCuD resistance, we assessed 100 cotton genotypes for their CLCuD resistance/tolerance and other related agronomical traits. Various statistical analytical tools, including correlation analysis, cluster analysis, and principal component analysis (PCA), were used to select the best genotypes. These genotypes can be used in future breeding programs to generate CLCuD resistant varieties. The same set of procedures could be utilized for other diseases in other crops. CLCuD incidence showed a significant negative genotypic correlation with yield-contributing traits followed by a significant negative association for phenotypic correlation. The seed cotton yield showed significant positive genotypic and phenotypic correlations with plant height, number of bolls per plant, and boll

weight. From the PCA we identified five principal components (PCs) that explained a significant amount of the variance among the variables, which may be used for selection of cotton genotypes with CLCuD resistance. Of the five PCs, the first four contributed more towards the total variability and had eigenvalues greater than one. The cluster analysis showed that the genotypes in one of the clusters performed particularly well with respect to CLCuD tolerance. These genotypes can be utilized for development of varieties with increased CLCuD tolerance.

Key words: CLCuD; *Gossypium*; Screening; Begomovirus; Cluster analysis

INTRODUCTION

About 26% of the farming community in Pakistan grows cotton on >3 million ha. Cotton and related commodities contribute to about 10% of total GDP, and account for 55% of the foreign exchange (Ahmed et al., 2010). In the textile sector, cotton products earned a foreign exchange worth US$10.22 billion during the 2014-2015 season (Anonymous, 2014-2015).

Cotton leaf curl disease (CLCuD) is a menace to cotton production in several African and Asian countries, including Pakistan, northwestern India. It was also recently reported in China. This disease is characterized by several whitefly transmitted begomoviruses (Family: Geminiviridae, Genus: *Begomovirus*) associated with specific satellite molecules (alpha- and betasatellites), which are responsible for symptom development (Sattar et al., 2013). The infected cotton plants (*Gossypium* L.) show a range of symptoms including vein thickening/swelling, leaf enations (which develop into leaf-like structures in extreme cases), and cup-shaped leaf curling. In some cases, CLCuD-affected cotton plants appear as lush and green as healthy plants, due to the proliferation of chloroplast-containing tissues (Sharma et al., 2005; Tahir et al., 2011). Plants infected soon after germination are usually severely stunted with compactly rolled leaves and produce no harvestable lint (Farooq et al., 2011). It has been shown that the leaf enations and vein thickening symptoms are due to the presence of cotton leaf curl Multan betasatellite (CLCuMuB) (Qazi et al., 2007; Tahir and Mansoor, 2011).

On the Indo-Pak subcontinent, cotton leaf curl virus (CLCuV) was identified for the first time in Pakistan near Multan in 1967. At that time, the disease was of minor importance and did not attract serious attention. A disease outbreak was recorded in 1988, causing damage to 22 ha of cotton field. After 1988, the geographic distribution of CLCuD increased greatly (Mahmood, 1999). Since its outbreak in 1988, the disease has been a major biotic constraint to cotton producing areas of Pakistan and north-western India (Briddon, 2003; Briddon and Markham, 2000; Kirthi et al., 2004). During the period 1988 through 2002 more than 7.7 million bales of cotton were lost due to CLCuD (Ahmad et al., 2002). The affected cotton crop was 60 ha in 1988-1989, which increased to 0.9 million ha in 1993-1994 (Harrison et al., 1997). During the late 1990s, due to an extensive breeding program, resistant cotton cultivars were developed and released to combat the disease, which helped reducing disease losses (Rahman et al., 2005). As a result, the disease incidence was reduced and it was not until 2001 that the disease reappeared in the vicinity of Burewala, in Vehari District (Mansoor et al., 2003). During the tranquil period, in the late 1990s and early 2000s, the virus types changed, resulting in susceptibility of all previously resistant cotton cultivars (Mansoor et al., 2003). A

newly emerged strain was reported to be cotton leaf curl Burewala virus (CLCuBuV), which was a recombinant of cotton leaf curl Multan virus and cotton leaf curl Kokhran virus. The betasatellite associated with CLCuBuV was also a new strain "Burewala strain", which was a recombinant of CLCuMuB and tomato leaf curl betasatellite (Amin et al., 2006).

In Africa, the disease is associated with only one single species of cotton leaf curl Gezira virus, which is associated with the cotton leaf curl Gezira betasatellite (Idris and Brown, 2002). The situation in southern Asia is more complex than in Africa: during the two CLCuD epidemics in Pakistan and India, at least six distinct begomovirus species have been identified in cotton (Kirthi et al., 2004; Sattar et al., 2013; Brown et al., 2015).

In addition to conventional approaches, such as controlling the whitefly insect vector, eradication of non-cultivated weed hosts, various seed treatments, and agronomical approaches, the best strategy is the development of CLCuD resistant varieties (Farooq et al., 2011). In order to meet this challenge, new germplasms continuously need to be included in breeding programs. The present study was conducted to identify CLCuD resistant/tolerant cotton germplasms to be used in current breeding programs.

MATERIAL AND METHODS

Plant material and field testing

Plant seeds of approximately 100 cotton genotypes were obtained from the Central Cotton Research Institute Multan and Punjab seed corporation, Khanewal, Pakistan. All the germplasms were sown at an experimental area of the Faculty of Agricultural Science & Technology, Bahauddin Zakariya University, Multan, Pakistan.

The genotypes were cultivated in triplicates in a randomized complete block design. Row length for each genotype was maintained at 3.05 m with plant to plant and row to row distances of 30 and 75 cm, respectively. All necessary conventional agronomic and cultural practices were adopted. No pesticides were applied against whitefly throughout the experiment, to ensure maximum inoculum pressure.

Virus inoculum and measurement of CLCuD incidence

Due to the absence of any artificial inoculation technique for CLCuD virus inoculum, all the germplasms were exposed directly to the natural field inoculum. The response of all the genotypes to CLCuD exposure was visually observed by adopting a disease rating scale (Table 1). A CLCuD severity index (SI) and CLCuD severity incidence were calculated using the formula described by Akhtar et al. (2010):

$$CLCuD\ incidence\ (\%) = Sum\ of\ all\ disease\ ratings / Total\ number\ of\ plants\ x\ 16.66 \quad \text{(Equation 1)}$$

Measurement of morphological and agronomical characteristics

Ten plants were selected randomly from each genotype and tagged for measurement of various morphological and agronomical traits (Table 2). Data collected included number of nodes to 1st monopodium, plant height (cm), number of monopodia per plant, number of sympodial branches per plant, leaf length (cm), leaf width (cm), petiole length (cm), percent

chlorophyll, number of flowers per plant, number of bolls per plant, 100-seed weight (g), boll weight (g), ginning out turn (%), and seed cotton yield (g).

Table 1. Rating scale for cotton leaf curl disease symptoms described by Akhtar et al. (2010) and Iqbal et al. (2014).

Symptom	Disease rating	Disease index (%)	Disease reaction
No disease, absence of symptoms.	0	0	Immune
Thickening of a few small veins or the presence of leaf enations on 10 or fewer leaves	1	0.1-10	Highly resistant
Thickening of a small group of veins	2	10.1-20	Resistant
Thickening of all veins but no leaf curling	3	20.1-30	Moderately resistant
Severe vein thickening and leaf curling on the top one third of the plant	4	30.1-40	Moderately susceptible
Severe vein thickening and leaf curling on half of the plant	5	40.1-50	Susceptible
Severe vein thickening, leaf curling, leaf enation, and stunting of the plant with reduced fruit production	6	>50	Highly susceptible

Table 2. Summary statistics of 100 cotton genotypes for the 14 studied traits.

Serial No.	Parameter	Minimum	Maximum	Mean	SE	SD	Variance
1	CLCuD SI	0.65	5.63	2.18	0.09	0.86	0.75
2	Plant height (cm)	23.33	195.67	124.50	4.31	43.14	1861.13
3	Nodes to 1st monopodium	0.33	8.33	4.26	0.15	1.51	2.27
4	Monopodia/plant	0.44	5.22	1.57	0.09	0.86	0.74
5	Sympodial branches/plant	4.89	31.44	13.86	0.51	5.09	25.88
6	Leaf length (cm)	1.70	19.78	5.53	0.26	2.61	6.82
7	Leaf width (cm)	1.82	11.51	6.21	0.25	2.50	6.23
8	Petiole length (cm)	1.29	8.44	4.84	0.20	2.01	4.03
9	Chlorophyll (%)	33.71	53.32	43.13	0.37	3.72	13.86
10	Bolls/plant	2.22	60.00	29.24	1.23	12.33	152.02
11	100-seed weight (g)	1.28	39.69	8.45	0.63	6.28	39.39
12	Boll weight (g)	1.20	6.05	2.19	0.05	0.55	0.30
13	Seed cotton yield/plant (g)	11.78	154.06	58.92	3.12	31.20	973.28
14	Ginning out turn (%)	30.88	46.47	40.73	0.33	3.34	11.18

Statistical analysis

The data were subjected to basic statistics, including correlation analysis, principal component analysis (PCA), and cluster analysis, using SPSS v. 19 and STATISTICA v. 5.0 (Sneath and Sokal, 1973). The cluster and dendrogram analyses were carried out using K-means clustering. Differences in disease index among the genotypes were tested using the Duncan multiple range test (Steel and Torrie, 1986). The above-mentioned statistical software were used to identify the models of variability between genotypes and the relationship between different clusters of particular traits (Akhtar et al., 2010; Iqbal et al., 2014; Saeed et al., 2014). In order to estimate the genetic diversity for breeding of different traits in various crops, similar statistical methods have been used successfully in many breeding programs (Coser et al., 2016).

RESULTS

Correlations with CLCuD incidence

In any breeding program, genotypes with desirable traits are required, thus the selection of genotypes is made on the basis of their association with particular traits of interest (Ali et al., 2009). High genetic variability was observed among all cotton genotypes (Table 2). The genotypic and phenotypic correlation coefficients revealed significant associations among 14 traits studied in this experiment (Table 3).

Table 3. Genetic (G) and phenotypic (P) correlation coefficients of CLCuD incidence (%), plant height (PH), number of bolls, plant yield, and other agronomical traits of 100 cotton genotypes.

Trait	Type	CLCuD	PH	NFM	MPP	SPP	LL	LW	PL	CPP	BPP	SI	BW	YPP	GOT
CLCuD	G	1.000	-0.1472*	0.0688	0.1165	-0.2721*	-0.2650*	-0.2221*	-0.1945*	0.1120*	-0.1194*	-0.0581*	-0.2403*	0.0792	-0.2590*
	P	1.000	-0.1355*	0.0692	0.0912	-0.2369**	-0.2502**	-0.2040**	-0.1801**	0.0841	-0.1039	-0.0526	-0.2175**	0.0628	-0.2311**
PH	G		1.000	0.4787*	-0.1972*	0.5795*	0.8745*	0.8804*	0.8665*	-0.3992*	0.1038*	-0.4328*	0.3424*	0.6050*	0.2982*
	P		1.000	0.4565**	-0.1723**	0.5579**	0.8581**	0.8604**	0.8404**	-0.3448**	0.1122	-0.4168**	0.3123**	0.5528**	0.2949**
NFM	G			1.000	-0.3543*	0.0829*	0.3300*	0.3051*	0.2889*	-0.2442*	-0.1580*	-0.2840*	0.0835	0.3434*	-0.0453*
	P			1.000	-0.3193	0.0683	0.3123**	0.2890**	0.2736**	-0.1926**	-0.1478**	-0.2585**	0.0543	0.3127**	-0.0423
MPP	G				1.000	-0.2770*	-0.2139*	-0.1996*	-0.2227*	-0.0524	0.3279*	0.4253*	-0.0924	-0.3198*	0.1228*
	P				1.000	-0.2095**	-0.1848**	-0.1672**	-0.1954**	-0.0610	0.3175**	0.3744**	-0.0521	-0.2772	0.1417*
SPP	G					1.000	0.4925*	0.4962*	0.5512*	-0.2135*	0.5204*	-0.2406*	05413*	0.4897*	0.6882*
	P					1.000	0.4625**	0.4632**	0.5065**	-0.2019	0.5378**	-0.2182**	0.4892**	0.4160**	0.6746**
LL	G						1.000	0.9866*	0.9695*	-0.3787*	-0.0829*	-0.4260*	0.2579*	0.4567*	0.1397*
	P						1.000	0.9732**	0.9537**	-0.3370**	-0.0680	-0.4088**	0.2413**	0.4201**	0.1421*
LW	G							1.000	0.9601*	-0.3898*	-0.0544*	-0.4145*	0.2574*	0.4899*	0.1646*
	P							1.000	0.9503**	-0.3404**	-0.0432	-0.3965**	0.2388**	0.4506**	0.1630**
PL	G								1.000	-0.4122*	-0.0152	-0.3999*	0.2955*	0.5104*	0.2057*
	P								1.000	-0.3642**	-0.0089	-0.3826**	0.2702**	0.4687**	0.1982**
CPP	G					.				1.000	-0.0481*	0.1421*	-0.2676*	-0.3479*	-0.1322*
	P									1.000	-0.0753	0.1265*	-0.2247**	-0.2687**	-0.1336*
BPP	G										1.000	0.2299*	0.4172*	0.2685*	0.8951*
	P										1.000	0.2280**	0.3939**	0.2260**	0.8801**
SI	G											1.000	0.1524*	-0.3959*	0.1174*
	P											1.000	0.1482*	-0.3677**	0.1222*
BW	G												1.000	0.2562*	0.6483*
	P												1.000	0.2221**	0.6008**
YPP	G													1.000	0.3134*
	P													1.000	0.2624**
GOT	G														1.000
	P														1.000

*,**Significant and highly significant, respectively.

For the yield-contributing traits, like plant height, number of sympodial branches per plant, leaf length, leaf width, seed index, and number of bolls per plant, we found negative genotypic and phenotypic correlations with CLCuD incidence (Table 3). Plant height was negatively correlated with CLCuD incidence for both types of association, indicating that genotypes with a high CLCuD intensity had a reduced plant height. The physiological parameters such as leaf length, leaf width, and petiole length were also negatively correlated with CLCuD intensity, whereas percent chlorophyll showed a positive association with CLCuD.

Among trait correlations

The yield-contributing traits, like seed cotton yield, showed significantly positive genotypic and phenotypic correlations with plant height, number of nodes to 1st monopodium, number of sympodial branches per plant, leaf length, leaf width, number of bolls per plant, and boll weight (Table 3). The genotypic and phenotypic correlations of plant height were positively associated with number of sympodial branches per plant, number of bolls per plant, boll weight, and seed cotton yield. This confirmed that genotypes with greater plant height also had other desired yield-contributing traits. The number of monopodia per plant showed negative genotypic and phenotypic correlations with number of sympodial branches per plant, leaf length, leaf width, boll weight, and seed cotton yield. On the other hand, number of sympodial branches per plant was positively correlated (genotypically and phenotypically) with plant height, number of bolls per plant, boll weight, and seed cotton yield. Typically, seed cotton yield is directly proportional to the number of sympodial branches in cotton plants (Farooq et al., 2014b). Thus, our results confirmed the direct contribution of the number of sympodial branches to high seed cotton yield. The number of bolls per plant had a significant positive genotypic correlation with plant height, 100-seed weight, boll weight, seed cotton yield, and ginning out turn, followed by positive phenotypic correlation, but a negative association was found with CLCuD. A similar association was observed by Ashokkumar and Ravikesavan (2010). Boll weight and seed cotton yield had significant positive genotypic and phenotypic correlations with seed cotton yield, ginning out turn, 100-seed weight, number of bolls per plant, petiole length, leaf width, leaf length, number of sympodial branches per plant, number of nodes to 1st monopodium, and plant height.

PCA results

The PCA produced 14 principal components (PCs) of which we selected the first five (PC1-PC5). PC1-PC4 had an eigenvalue >1 explaining 74.26% of the total variation, whereas PC5 explained only 6.49% of the variability. PC1-PC5 thus explained a total of 80.75% of the variance observed in the cotton genotypes under study. The remaining PCs explained only 19.25% of the total variability. PC1 explained the maximum share of variability (37.71%) with an eigenvalue of 5.28, followed by PC2 (19.50%) with eigenvalue 2.73, PC3 (9.06%) and an eigenvalue of 1.26, PC4 (7.99%) had an eigenvalue of 1.12, and PC5 (6.49%) had an eigenvalue of 0.90 (Table 4).

The correlations between the variables and factors are called factor loadings. They indicate the percentage of variance in inventive variable. Traits like plant height, number of nodes to 1st monopodium, number of sympodial branches per plant, number of bolls per plant,

boll weight, seed cotton yield, and ginning out turn had significant negative factor loadings (variance percentage) on PC1.

Table 4. Principle component analysis of CLCuD incidence, plant height, number of bolls per plant, plant yield, and other agronomical traits of 100 cotton genotypes.

	PC1	PC2	PC3	PC4	PC5
Eigenvalue	5.2807	2.7304	1.2679	1.1181	0.9086
Cumulative eigenvalue	5.2808	8.0112	9.2791	10.3972	11.3058
Total variance (%)	37.7199	19.5029	9.0564	7.9867	6.4900
Cumulative variance (%)	37.72	57.223	66.2794	74.2662	80.7562
Factor loadings by different traits					
Variables	PC1	PC2	PC3	PC4	PC5
CLCuD SI	0.2449	0.2176	0.4234	0.6705	0.2010
Plant height (cm)	-0.9190	0.1097	-0.0399	0.1120	0.0333
Nodes to 1st monopodium	-0.3972	0.3563	0.3664	0.1382	-0.5128
Monopodia/plant	0.2718	-0.4028	-0.5164	0.5778	0.1618
Sympodial branches/plant	-0.7032	-0.4499	0.2022	-0.1917	0.1699
Leaf length (cm)	-0.8968	0.2406	-0.3036	-0.0151	0.0929
Leaf width (cm)	-0.8973	0.2222	-0.2873	0.0307	0.1217
Petiole length (cm)	-0.9053	0.1851	-0.2488	0.0325	0.1133
Chlorophyll (%)	0.4385	0.0232	0.1992	-0.4168	0.4962
Bolls/plant	-0.2051	-0.8837	0.1843	0.1554	0.1462
100 seed weight (g)	0.4490	-0.4905	-0.3083	0.0435	-0.3276
Boll weight (g)	-0.4450	-0.5545	0.0395	-0.1464	-0.3644
Seed cotton yield/plant (g)	-0.6377	0.0165	0.4626	0.2065	0.0539
Ginning out turn (%)	-0.4249	-0.8415	0.1559	-0.0162	0.0468

CLCuD incidence, number of monopodial branches per plant, percent chlorophyll, and 100-seed weight loaded positively (Table 4). PC2 had high diversity of genotypes with positive loadings for number of nodes to 1st monopodium, CLCuD incidence, leaf length, leaf width, and plant height, whereas number of bolls per plant, 100-seed weight, ginning out turn, boll weight, number of monopodial branches per plant, and number of sympodial branches per plant loaded negatively onto that same PC. The highest positive loadings in PC3 were seed cotton yield and CLCuD incidence, followed by the number of nodes to 1st monopodium and ginning out turn and ultimately low diversity. PC4 had the highest positive loading for CLCuD incidence, number of monopodial branches per plant, seed cotton yield, and number of bolls per plant, whereas traits like number of sympodial branches per plant, percent chlorophyll, and boll weight loaded negatively. Percent chlorophyll, and number of bolls per plant both loaded strongly and positively on PC5.

In the PC biplot (Figure 1) the contribution of genotypes on the observed variation is illustrated as distance between the variables with respect to PC1 and PC2. The biplot showed that seed cotton yield, number of bolls per plant, boll weight, plant height, percent chlorophyll, CLCuD incidence, and 100-seed weight contributed the most in variability among the germplasms.

In a PC scatter plot, genotypes located near each other are considered to have a similar contribution with respect to the different variables studied. By contrast, genotypes located far from each other are more diverse. Based on the PCA, biplot, and scatter plot analyses, genotypes FH-900, MNH-886, SAHARA-120, CIM-557, S-12, K-68/9, CIM-20, AGC-555, FH-142, and VH-303 had the maximum variability for CLCuD resistance and can be used in breeding programs successfully (Figure 2).

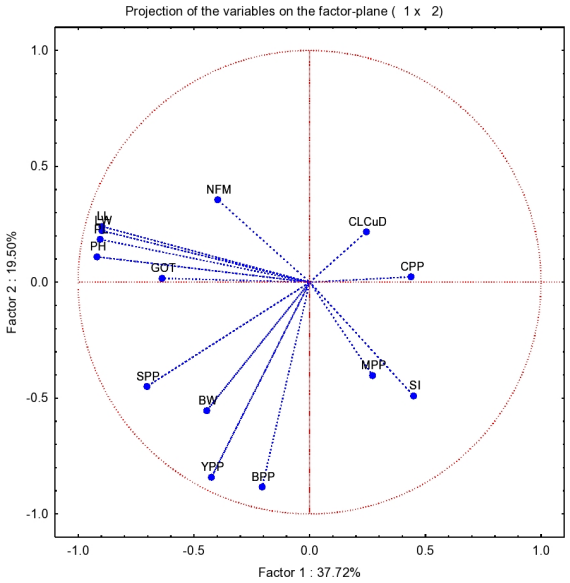

Figure 1. Biplot between PC1 and PC2 showing involvement of different traits in variability. G = genotypic correlation, P = phenotypic correlation, CLCuD = cotton leaf curl virus disease, PH = plant height, NFM = nodes to 1st monopodial, MPP = monopodial branches/plant, SPP = synpodial branches/plant, LL = leaf length, LW = leaf width, PL = petiole length, CPP = chlorophyll %, BPP = number of bolls per plant, SI = seed index, BW = boll weight, YPP = seed cotton yield/plant, GOT = ginning out turn percentage.

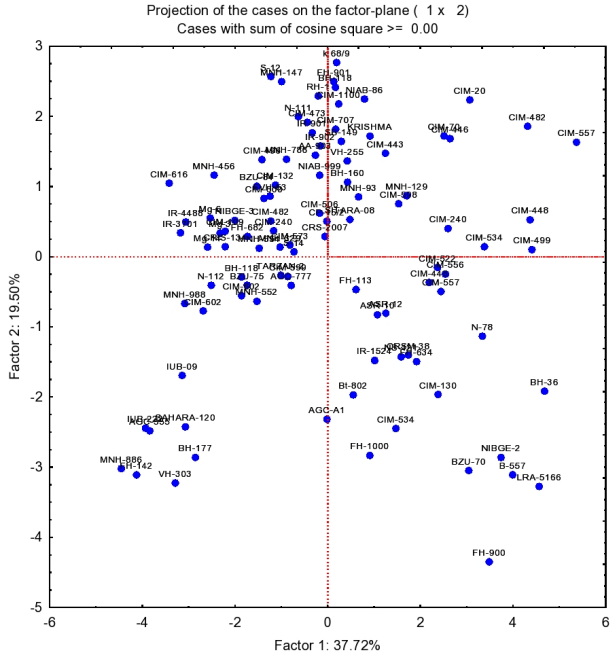

Figure 2. PC scatterplot showing classification of cotton genotypes on PC1 and PC2.

Cluster analysis

The cluster analysis grouped the 100 cotton genotypes into five clusters based on different traits and their level of variability (Table 5). Clusters 1-5 comprised 31, 29, 25, 8, and 7 genotypes, respectively (Table 5). The genotypes in Cluster 1 had higher values compared to all other clusters for all traits investigated in this study. Members of Cluster 2 showed the ideal value of ginning out turn, had a significant chlorophyll percent and seed cotton yield. The members of Clusters 3 and 4 had better tolerance against CLCuD intensity, but otherwise showed an overall poor performance in terms of plant height, number of sympodial branches per plant, petiole length, and seed cotton yield. Cluster 5 consisted of genotypes that had considerable tolerance against CLCuD as well as maximum seed cotton yield. Members of Cluster 5 also showed best value for plant height, number of sympodial branches per plant, number of bolls per plant, and boll weight.

Table 5. Estimation of trait means for all clusters depending on variability.

Trait	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
CLCuD SI	2.22	2.23	2.24	2.41	1.36
Plant height (cm)	168.30	118.89	40.64	79.67	176.56
Nodes to 1st monopodium	4.89	4.60	2.02	3.34	4.07
Monopodia/plant	1.34	1.44	1.57	2.14	1.64
Sympodial branches/plant	15.53	12.39	7.40	13.42	24.21
Leaf length (cm)	7.23	5.36	2.01	2.78	7.61
Leaf width (cm)	8.24	5.93	2.51	3.31	8.72
Petiole length (cm)	6.63	4.63	1.86	2.59	7.13
Chlorophyll (%)	41.22	43.81	44.48	45.59	40.67
Bolls/plant	28.93	22.77	19.01	45.86	50.57
100-seed weight (g)	6.45	7.28	9.69	16.92	6.42
Boll weight (g)	2.24	2.06	1.83	2.31	2.63
Seed cotton yield/plant (g)	59.19	43.67	29.73	85.46	131.83
Ginning out turn (%)	42.39	40.70	35.76	39.51	42.70

DISCUSSION

Genetic diversity always offers prospects to plant breeders for development of advanced genotypes with desirable characteristics, including resistance against biotic and abiotic stressors. Our correlation analyses showed similar results as those found in previous studies, with significant negative correlations between plant height and number of sympodial and monopodial branches with CLCuD incidence (Saeed et al., 2014). CLCuMB associated with helper begomovirus causes cell proliferation in chloroplasts and leads to higher chlorophyll contents in the infected cotton plants (Sharma et al., 2005; Ajmal et al., 2011; Tahir et al., 2011). In our study, we found a similar pattern in which percent chlorophyll was positively associated with CLCuD incidence.

The earliness related traits, like number of nodes to 1st monopodium, had significant negative genotypic and phenotypic correlations with number of monopodial branches per plant, percent chlorophyll, number of bolls per plant, 100 seed weight, and ginning out turn, as has been found in previous studies (Shah et al., 2010). Similar associations among these traits were also confirmed by Farooq et al. (2014a).

A PCA was used to distribute the total variance into PCs to select the best performing germplasms on the basis of average values of various traits. In the PCA, genetic resources

can be partitioned and further utilized in crop improvement for the desired trait, like CLCuD resistance in this case (Pecetti and Damania, 1996). The variability of the trait under study depends on the eigenvalue of the PC for that particular trait, i.e., the higher the eigenvalue the more variability for the trait under investigation.

In order to identify the level of variation between different traits, the contribution of the first two PCs play an important role in estimating the variability. PC1 and PC2 explained the maximum variance, as reported previously by Nazir et al. (2013). Thus, the best performing genotype against CLCuD incidence should be selected from PC1, which has the maximum variability and highest eigenvalue. A PCA is very useful for investigating evidence of extensive variation in different traits. This information could be used for selection of parents for breeding programs to produce CLCuD resistant cultivars as well as varieties with other desired traits (Malik et al., 2011). Grouping of germplasms with great variation between the clusters is of great genetic value for the selection of CLCuD resistant material (Grenier et al., 2000).

In order to estimate the genetic variation present among all studied clusters, a Ward's dendrogram was constructed (Figure 3) as described previously (Grenier et al., 2000; Nazir et al., 2013). The dendrogram showed the presence of wide variation among the clusters suggesting high genetic variability among genotypes. Based on the cluster and Ward's dendrogram analyses, the members of Cluster 5, including FH-142, VH-303, MNH-886, BH-177, AGC-555, IUB-222, and SAHARA-120, are recommended to be used for the development of CLCuD resistant cultivars. These statistical tools could be used for the identification of other potential sources, for example, screening of bread wheat has been done to discover resistance against stem rust in wheat (Nzuve et al., 2012).

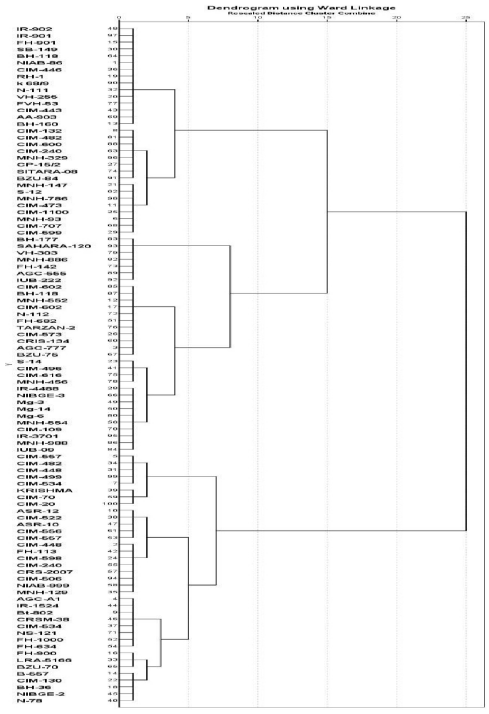

Figure 3. Ward's dendrogram of cotton genotypes showing five clusters.

To conclude, the uses of different statistical methods, like correlation coefficients, PCA, and cluster analysis, provide information that can be used to identify and classify genotypes with high CLCuD resistance. These statistical instruments made it possible to select CLCuD tolerant genotypes that also showed high seed cotton yield and other valuable agronomical traits associated with increased production. Cotton breeding programs intended for the development of CLCuD resistance may use our results for a comprehensive approach in selection of the best performing genotypes to be used in developing resistance against CLCuD.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by the Higher Education Commission (HEC), Pakistan, Department of Plant Breeding and Genetics and Institute of Molecular Biology & Biotechnology, B.Z.U Multan. This research article is part of a PhD study by M. Javed. The authors are thankful to Dr. M.N. Sattar for his valuable comments and suggestions through the course of this manuscript.

REFERENCES

Ahmad G, Malik SA, Mahmood Z, Iqbal MZ, et al. (2002). Effect of cotton leaf curl virus disease severity on morphology, yield and fibre characteristics of susceptible lines/cultivars of cotton (*G. hirsutum* L.). *Asian J. Plant Sci.* 1: 705-707. http://dx.doi.org/10.3923/ajps.2002.705.707

Ahmed MZ, Ren SX, Mandour NS, Maruthi MN, et al. (2010). Phylogenetic analysis of *Bemisia tabaci* (Hemiptera: Aleyrodidae) populations from cotton plants in Pakistan, China, and Egypt. *J. Pest Sci.* 83: 135-141. http://dx.doi.org/10.1007/s10340-009-0279-4

Ajmal S, Perveen R, Chohan S, Yasmin G, et al. (2011). Role of secondary metabolites biosynthesis in resistance to cotton leaf curl virus (CLCuV) disease. *Afr. J. Biotechnol.* 10: 18137-18141.

Akhtar KP, Haidar S, Khan MKR, Ahmad M, et al. (2010). Evaluation of *Gossypium* species for resistance to leaf curl Burewala virus. *Ann. Appl. Biol.* 157: 135-147. http://dx.doi.org/10.1111/j.1744-7348.2010.00416.x

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

Amin I, Mansoor S, Amrao L, Hussain M, et al. (2006). Mobilisation into cotton and spread of a recombinant cotton leaf curl disease satellite. *Arch. Virol.* 151: 2055-2065. http://dx.doi.org/10.1007/s00705-006-0773-4

Anonymous (2014-2015). Pakistan economic survey. Pakistan Bureau of Statistics, Chapter 2.

Ashokkumar K and Ravikesavan R (2010). Genetic studies of correlation and path coefficient analysis for seed oil, yield and fibre quality traits in cotton (*G. hirsutum* L.). *Aust. J. Basic. Appl. Sci.* 4: 5496-5499.

Briddon RW (2003). Cotton leaf curl disease, a multicomponent begomovirus complex. *Mol. Plant Pathol.* 4: 427-434. http://dx.doi.org/10.1046/j.1364-3703.2003.00188.x

Briddon RW and Markham PG (2000). Cotton leaf curl virus disease. *Virus Res.* 71: 151-159. http://dx.doi.org/10.1016/S0168-1702(00)00195-7

Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, et al. (2015). Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. *Arch. Virol.* 160: 1593-1619. http://dx.doi.org/10.1007/s00705-015-2398-y

Coser SM, Motoike SY, Corrêa TR, Pires TP, et al. (2016). Breeding of *Acrocomia aculeata* using genetic diversity parameters and correlations to select accessions based on vegetative, phenological, and reproductive characteristics. *Genet. Mol. Res.* 15: http://dx.doi.org/10.4238/gmr15048820.

Farooq A, Farooq J, Mahmood A, Shakeel A, et al. (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, Farooq A, et al. (2014a). Correlation and path coefficient analysis of earliness, fiber quality and yield contributing traits in cotton (*Gossypium hirsutum* L). *J. Anim. Plant Sci.* 24: 781-790.

Farooq J, Farooq A, Riaz M, Shahid MR, et al. (2014b). Cotton leaf curl virus disease a principle cause of decline in cotton productivity in Pakistan (a mini review). *Can. J. Plant Prot.* 2: 9-16.

Grenier C, Deu M, Kresovich S, Bramel-Cox PJ, et al. (2000). Assessment of genetic diversity in three subsets constituted from the ICRISAT *Sorghum* collection using random vs. non-random sampling procedures B. using molecular markers. *Theor. Appl. Genet.* 101: 197-202. http://dx.doi.org/10.1007/s001220051469

Harrison BD, Liu YL, Khalid S, Hameed S, et al. (1997). Detection and relationships of cotton leaf curl virus and allied whitefly-transmitted geminiviruses occurring in Pakistan. *Ann. Appl. Biol.* 130: 61-75. http://dx.doi.org/10.1111/j.1744-7348.1997.tb05783.x

Idris AM and Brown JK (2002). Molecular analysis of Cotton leaf curl virus-Sudan reveals an evolutionary history of recombination. *Virus Genes* 24: 249-256. http://dx.doi.org/10.1023/A:1015380600089

Iqbal M, Naeem M, Aziz U, Afzal J, et al. (2014). An overview of cotton leaf curl virus disease, persistant challeng for cotton production. *Bulg. J. Agric. Sci.* 20: 405-415.

Kirthi N, Priyadarshini CGP, Sharma P, Maiya SP, et al. (2004). Genetic variability of begomoviruses associated with cotton leaf curl disease originating from India. *Arch. Virol.* 149: 2047-2057.

Mahmood T (1999). Cotton leaf curl virus disease and its present status in Punjab. *Pak. Cotton Grower* 3: 17-18.

Malik W, Iqbal MZ, Khan AA, Noor E, et al. (2011). Genetic basis of variation for seedling traits in *Gossypium hirsutum* L. *Afr. J. Biotechnol.* 10: 1099-1105.

Mansoor S, Amin I, Iram S, Hussain M, et al. (2003). Breakdown of resistance in cotton to cotton leaf curl disease in Pakistan. *Plant Pathol.* 52: 784. http://dx.doi.org/10.1111/j.1365-3059.2003.00893.x

Nazir A, Farooq J, Mahmood A, Shahid M, et al. (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.

Nzuve F, Bhavani S, Tusiime G and Njau P (2012). Field screening of bread wheat for partial sources of resistance to stem rust. Third RUFORUM Biennial Meeting, Entebbe, 1419-1431.

Pecetti L and Damania AD (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.

Qazi J, Amin I, Mansoor S, Iqbal MJ, et al. (2007). Contribution of the satellite encoded gene betaC1 to cotton leaf curl disease symptoms. *Virus Res.* 128: 135-139. http://dx.doi.org/10.1016/j.virusres.2007.04.002

Rahman M, Hussain D, Malik TA and Zafar Y (2005). Genetics of resistance to cotton leaf curl disease in *Gossypium hirsutum*. *Plant Pathol.* 54: 764-772. http://dx.doi.org/10.1111/j.1365-3059.2005.01280.x

Saeed F, Farooq J, Mahmood A, Hussain T, et al. (2014). Genetic diversity in upland cotton for cotton leaf curl virus disease, earliness and fiber quality. *Pak. J. Agric. Res* 27: 226-236.

Sattar MN, Kvarnheden A, Saeed M and Briddon RW (2013). Cotton leaf curl disease - an emerging threat to cotton production worldwide. *J. Gen. Virol.* 94: 695-710. http://dx.doi.org/10.1099/vir.0.049627-0

Shah MKN, Malik SA, Murtaza N, Ullah I, et al. (2010). Early and rapid flowering coupled with shorter boll maturation period offers selection criteria for early crop maturity in upland cotton. *Pak. J. Bot.* 42: 3569-3576.

Sharma P, Rishi N and Malathi VG (2005). Molecular cloning of coat protein gene of an Indian cotton leaf curl virus (CLCuV-HS2) isolate and its phylogenetic relationship with others members of Geminiviridae. *Virus Genes* 30: 85-91. http://dx.doi.org/10.1007/s11262-004-4585-x

Sneath PHA and Sokal RR (1973). Numerical taxonomy: The principles and practice of numerical classification. W.F. Freeman & Co., San Francisco.

Steel RGD and Torrie JH (1986). Principles and procedure of statistics. A biometrical approach. 2nd edn. McGraw Hill Book Company, New York.

Tahir MN and Mansoor S (2011). βC1 of chili leaf curl betasatellite is a pathogenicity determinant. *Virol. J.* 8: 509. http://dx.doi.org/10.1186/1743-422X-8-509

Tahir MN, Amin I, Briddon RW and Mansoor S (2011). The merging of two dynasties--identification of an African cotton leaf curl disease-associated begomovirus with cotton in Pakistan. *PLoS One* 6: e20366. http://dx.doi.org/10.1371/journal.pone.0020366