

GENETIC DIVERGENCE ON THE BASIS OF PRINCIPAL COMPONENT, CORRELATION AND CLUSTER ANALYSIS OF YIELD AND QUALITY TRAITS IN COTTON CULTIVARS

AMEER HUSSAIN JARWAR^{1,3}, XIAOYAN WANG², MUHAMMAD SHAHID IQBAL^{1,4}, ZAREEN SARFRAZ^{1,4}, LONG WANG¹, QIFENG MA.^{1*} AND FAN SHULI^{1*}

¹State Key Laboratory of Cotton Biology, Institute of Cotton Research of CAAS, Anyang 455000, China

²Anyang Institute of Technology, College of Biology and Food Engineering, Anyang, 455000, China

³Oil Seeds Section, Agriculture Research Institute, Tandojam, Sindh, Pakistan

⁴Cotton Research Institute, AARI, Faisalabad, Pakistan

*Corresponding author's email: Fsl427@126.com,

Abstract

The present research work was conducted to obtain genetic divergence regarding fiber and yield associated traits among 12 different Chinese cotton varieties (cultivars) of upland cotton. The phenological studies were accomplished via carrying out correlation, and cluster analyses. The outcomes from current study demonstrated strong positive correlation (0.69) between number of bolls per plant and cottonseed yield as well as positive significant correlation between lint index (0.94) and staple length (0.70); however, fiber strength exhibited significant negative correlation (0.65) with cottonseed yield per plant. Heat map of experimented genotypes depicted seven clusters are I, II, III, IV, V, VI and VII keeping cluster-III at extreme for highest individual genotypic scores concerning different fiber and yield associated traits. Principal component analysis revealed total variability of 74.38% among genotypes contributed by PC-1 (52.86%) and PC-2 (21.52%). Yield related traits i.e., number of bolls per plant, cotton seed yield and staple length contributed positive significant component loadings for the first two PCs. Cluster analysis (k) grouped studied cultivars into three discrete clusters. Cluster-2 and cluster-3 got optimum values for cottonseed yield, number of bolls per plant, seed index and staple length. The information was fetched together after applying different statistical tools in order to make the outcomes of genetic diversity fruitful for developing an improved cotton breeding program.

Key words: Correlation, Principal component analysis, Genetic diversity, Cotton cultivars.

Abbreviations; Symp= Sympodial branches, BW= Boll weight (g), GOT= Ginning out turn (percentage %), MIC= Micronaire value (microgram/inch), NOB= Number of boll, PH= Plant height (cm), SL= Staple length (mm), SI= Seed index, YLD= Yield (g)

Introduction

Cotton is the basic, most significant renewable natural textile fiber and the sixth largest source of oil in the world. It belongs to the genus *Gossypium*, which consists of five allotetraploid and 45 diploid species. Among them only four species are cultivated worldwide comprising of two diploids and two tetraploid also called old world and new world species respectively (Ulloa *et al.*, 2006). In crop plants genetic diversity is considered as a baseline for any breeding program as the abruptly changing climatic conditions are considered responsible for unanticipated effects on crop plants. It is therefore assumed that narrow genetic base of the crops is the most important reasons for decline of the crop. In many studies, scientists have established a narrow genetic base regarding upland cotton (May *et al.*, 1995; Mc Carty *et al.*, 2005; Satish *et al.*, 2009). Different cotton breeders and cotton growers (Li *et al.*, 2005; Rathinavel, 2017, Gopinath *et al.*, 2009,) have attributed yield reduction and deterioration in fiber quality of cotton to less diverse genotype availability.

Broadening genetic base and exploitation of genetic diversity is the basic objective in any hybrid-breeding program with an objective to crop improvement (Van-Esbroeck & Bowman, 1998; Shakeel *et al.*, 2015). Comprehensive and detailed information of genetic diversity available within the resources is a key requirement for a potential breeding program. It allows

the selection of desirable parents for crossing to develop a genetically diverse population for selection. For assessment of genetic diversity in cotton varieties phenotypic traits especially morphological or agronomic traits is considered as the most powerful tool by a plant breeder. Following this procedure can provide a baseline in the development of superior genotypes to be used as parents and possible hybrid combinations (Alishah, 2001; Ashok kumar *et al.*, 2011; Rathinavel, 2017).

The extent of genetic variation can be either estimated using univariate analysis or by multivariate analysis. Conventionally univariate analysis has been the most accepted biometrical method to work out genetic variations. In recent years, multivariate analysis has gained popularity and principal component analysis (PCA) and cluster analysis and Principal coordinate analysis (PCoA) have been employed to workout similarities and differences between different genotypes regarding multiple traits under examination (Sharma, 1998; Jian *et al.*, 2006). Among all other biometrical methods, PCA has the edge to assign only one group to each genotype, at the same time it also depicts the importance of major contributor toward total diversity at each axis of differentiation (Sharma, 2006). It is therefore, for the assessment of genetic diversity, breeders have widely used principal component analysis, which helps in exploration of promising genotypes for future breeding programmes (Quailing & Zhe, 2011; Rehman *et al.*, 2015).

The aims of the current study were to appraise different cotton cultivars for correlation variability and diversity based on fiber and yield quality related traits to understand the mechanism and procedure for improvement of cotton cultivars with improved yield and quality parameters.

Materials and Methods

Twelve Chinese cotton genotypes having wider adaptability were examined in the field conditions (2017). Delinted cottonseeds were grown in a complete randomized block design (RCBD) with three replications at the experimental field of Cotton Research institute, Anyang Henan, China. Each entry/treatment was a plot of 600 x 450 cm size composed 80 cm apart six rows. The space among plants within rows was 28 cm. All the agronomical and cultivation exercises such as fertilizer application, weeding, irrigation and hoeing were adopted at the same time respectively.

Measurement of survey traits: For the measurements of traits, ten representatives spoiled and fit staying and healthy plants were selected randomly from each genotype and later they marked with labels for recognition. Genotypes were estimated for number of bolls per plant (NB), lint index (LI), seed index (SI), seed cotton yield (SCY), and fiber quality, and other boll weight BW measurements including fiber length (FL), and fiber fineness (FF), Plant height (cm), Ginning out turn%, were calibrated through Spin lab HVI-900.

Data analyses: The average data results were subjected to basic statistics (descriptive), correlation investigation, cluster analysis (Anderson & Boyes, 1957) and other most important elements estimates (PCA) through the statistical procedures by Sokal & Sneath, (1973). Cluster analyses were performed using K means clustering.

Results and Discussion

Correlation investigation: The main statistics of studies traits exhibited enough amounts of variability between 12

cotton genotypes (Table 1). The data of simple correlation coefficient exposed positive affiliation between traits analysed in (Table 1). Correlation Matrix showed positive correlations in sky blue colour while negative associations in red color. The color intensity and the size of the circle are proportional to the correlation coefficients. Meanwhile, in the other right side of the correlogram, in the diagram it showing that the Fable color for different associations (Fig. 2). Lint index showed positive relationship with seed index while an positive significant association with cotton seed yield. However, for other parameters that demonstrated non-significant association, boll weight displayed negative association with number of bolls and staple length. Whereas, the numbers of bolls were detected to be in extreme positive association with cottonseed, while it exhibited negative correlation with staple length. Fiber length and cottonseed yield were highly associated; however, fiber strength displayed negative association with cottonseed. Pearson's coefficient of correlation was performed for 12 qualitative traits and correlation matrix was calculated for the comparison of different traits using XLSTAT version 2016.05 software. The standardized values of 26 qualitative traits were used to perform principal component analysis (PC-A) using PAST-3 (Sharma *et al.*, 2006) to estimate the close relationship between different traits and to capture the existing genetic variation. A scree plot was drawn from the Eigen values associated with a component or factor in descending order viz., the number of the components. Scree plot is usually used for visual assessment of factor, which explains high amount of the changes in the data. Shannon and Weiner described the diversity index (Shannon & Weaver, 1949) which was estimated for the measurement of phenotypic diversity for each trait (Fig. 3).

To describe the variation among different traits; the contributions of PCS play a vital role in identifying the changes. The PC-1 and PC-2 relationship was estimated tentatively for the association between the lines using marker data, the function as the matrix for delegate familiar connectedness. The schematic heat model diagram showed association among 275 RILs of Set 1.

Table1. Eigen values and cumulative variability in different PCs for yield and fiber quality related traits in cotton cultivars.

Sr. No.	F1	F2	F3	F4	F5	F6	F7	F8	F9
Symp	0.866	-0.108	0.073	0.320	-0.314	0.126	-0.116	0.058	0.001
BW	0.555	0.816	0.120	0.018	0.027	-0.077	-0.060	-0.013	-0.015
GOT	-0.762	0.214	0.446	0.233	0.329	0.007	-0.076	0.079	0.003
MiC	0.530	-0.301	0.730	-0.225	-0.146	-0.127	0.080	0.039	0.000
Bno	0.881	0.337	-0.164	0.055	0.169	0.109	0.192	0.065	0.000
PHt	0.654	-0.491	0.098	0.527	0.171	-0.091	0.040	-0.069	-0.001
SL	0.736	-0.386	-0.462	-0.166	0.169	-0.160	-0.093	0.072	0.000
SI	0.806	-0.273	0.267	-0.315	0.256	0.173	-0.085	-0.052	0.000
YLD	0.666	0.738	0.043	-0.011	-0.010	-0.079	-0.037	-0.040	0.015
Eigenvalue	4.758	1.937	1.075	0.615	0.380	0.120	0.085	0.030	0.000
Variability (%)	52.862	21.521	11.946	6.832	4.227	1.335	0.941	0.330	0.005
Cumulative %	52.862	74.382	86.329	93.161	97.388	98.723	99.664	99.995	100.000

Symp= Sympodial branches, BW= Boll weight (g), GOT= Ginning out turn (%age), MIC= Micronaire value (microgram/inch), NOB= Number of boll, PH= Plant height (cm), SL= Staple length (mm), SI=Seed index, YLD= Yield (g)

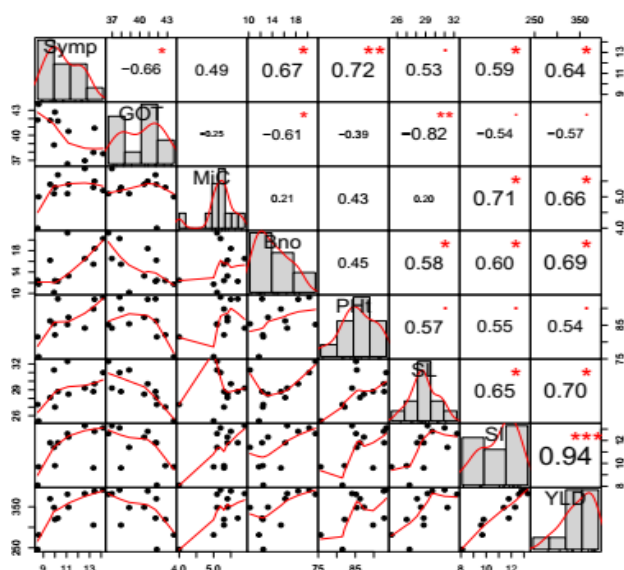


Fig. 1. Schematic model of plot showing the letters are significant. Symp= Sympodial branches, BW= Boll weight (g), GOT= Ginning out turn (percentage %), MIC= Micronaire value (microgram/inch), NOB= Number of boll, PH= Plant height (cm), SL= Staple length (mm), SI=Seed index, YLD=Yield (g).

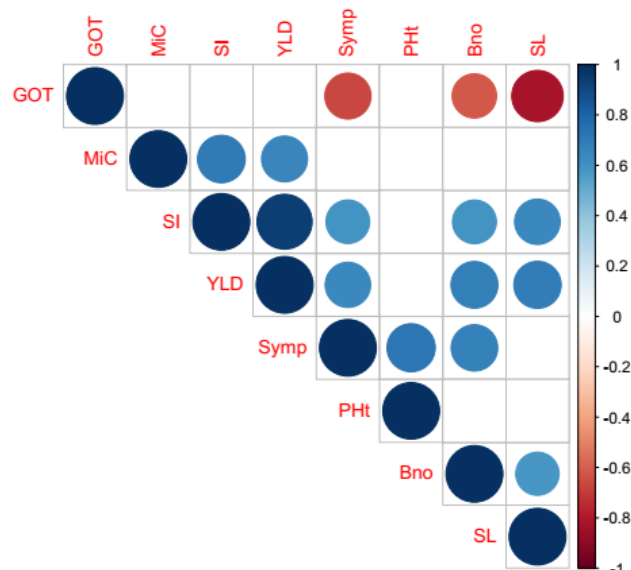


Fig. 2. Correlation Matrix showing characters correlation with each other. A positive correlation is shown in blue and negative correlations in shown in red color. The color intensiveness and the size of the circle are relatively proportional to the correlation.

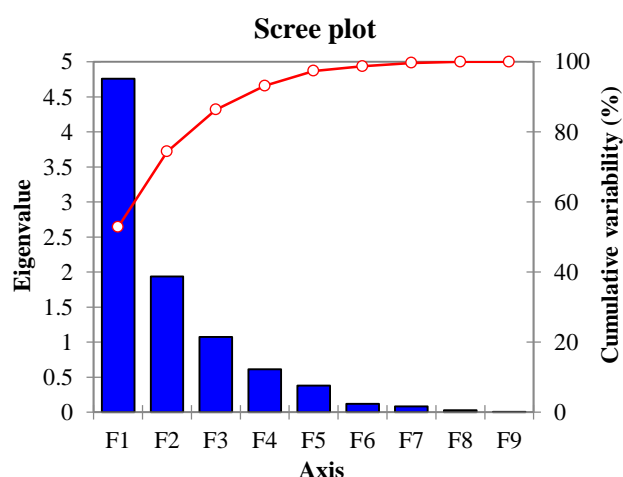


Fig. 3. Eigenvalue for yield and yield contributing factors of cotton showing the cumulative variability percentage (%) (line).

In (Fig. 1). The distribution of each variable is exhibited on the diagonal. On the lower side: the vicariate schematic model plots with a line of best fit are shown.

Above the diagonal, the values of the correlation are marked with significance level as stars,

The significance level is connected to a symbol: p-values (0, 0.001, 0.01, 0.05, 0.1, 1) \Leftrightarrow symbols (“***”, “**”, “*”, “ ”).

The correlation matrix is used for to investigate the dependence between multiple variables at the same time. The result containing the correlation coefficients between each variable and the others. There are different methods for correlation analysis; Pearson parametric correlation test. Spearman and Kendall rank based correlation analysis. The positive correlations are displayed in the blue and the negative correlations in red color. In the red diagonal color, express high association within line with

itself, and the symmetrical off-diagonal components, that represents the association of the quantity measures in this case identity by declension for set of lines. There is an apparently color on the diagonal which express a cluster of nearly connected lines. The tree diagram model on the right side showing that the results of a cluster analysis on the IBD matrix.

In context of high variance, as presented earlier by (Nazir *et al.*, 2013) the superior cultivars against number of bolls per plant were selected from PC-1, that has the optimum variability and high Eigen value. PC-1 is important and authentic to inquire manifestation of wide range of variation among different traits. The information could be helpful in the selection of better parents for breeding techniques to create superior genotypes for desired traits (Malik *et al.*, 2011).

Principle components analysis (PCA): The sustainable use of genotypic resources could made potential by partitioning of total variance into its component factors (Fig. 3). It also provides a opportunity to allow the genetic resources for crop improvement for a particular plant trait (Pecetti *et al.*, 1996, Sokal & Sneath, 1973). The most important factor is to distribute the total variance into the different factors, which is considered as a potential tool for applicable parental lines for a useful breeding programme (Akter *et al.*, 2009).

Four out of nine component factors (PCs) were discovered with an Eigenvalue of >1 . These principal factors (PCs) contributed 92.2% towards the total variability existing among cotton varieties regarding different fiber and yield associated related traits (Table 2). Contributes of PC-I was superlative towards the variability (37%) complied by PC-2 (37.5%), PC-3 (18.1%) and PC-4 (11.2%). The characters including cottonseed yield, fiber length, and number of bolls per plant displayed positive correlation of factor loadings on PC-1 while fiber strength

and boll weight have significantly negative loadings (Table 1). The PC-2 was described only by cottonseed yield, and number of bolls with positive factor loadings, meanwhile the remaining traits in that PC obtained negative loadings. PC-3 was compiled with fiber uniformity, and fiber fineness and fiber length with their positive loading and negative loading factors were discovered for seed index, lint index, and number of bolls. The PC-4 was explained by variance due to staple length and fiber strength having positive factors loading, while fiber uniformity, cottonseed yield, and boll weight occur negative loadings. A PC biplot (Fig. 4) demonstrated genotypes and variables are superb and inflict as a vector on plot. The gap within traits with regard to PC-1 and PC-2 depicted contribution of these traits in creating variation of genotypes. However, the overall biplot diagram demonstrated that cotton yield, number of bolls and staple length contributed significantly towards the diversity in germplasm while ginning outturn percentage was found varying regarding all characteristics.

The PCA was conducted to dispense the aggregate variance into the PCs to select the superior germplasm based on average amount of different traits. In the PC-A, genetic resources can be divided and furthermore used for improvement of crop and advancement for economic traits, Thus the variability of traits in the study have a bearing on the eigenvalue of the PC for the specific traits, such as the maximum Eigenvalue is responsible for wide variance for the trait under demonstration. The principal factor PC is a powerful practical method that provides easiness for obtaining suitable parental genotypes in order to generate useful breeding programs (Nazir *et al.*, 2013). Many scientists have applied this method to formalize the genotypic divergence studies. Our present research also testifies the robustness of similar methods and in line with the previous findings obtained by (Saeed *et al.*, 2015), for catering easy recognition of superior fiber characteristics of cotton. The researchers also described high amount of first two principal factors to the total variance. Ashokkumar & Ravikesavan (2011), and Malik *et al.*, (2011) also described enormous quantity of genetic diversity in different upland cotton cultivars. These conclusions are also in correspondence with that of Nazir *et al.*, (2013) who concluded that the contribution between first two PCs are very important in the total variation. Exploitation of genotypic resources could be made by assigning the total

variation into number of factors, however it also provides a chance to used very important and appropriate germplasm in future for the crop improvement program to amend the yield production for specific plant traits.

Association between different traits is a most important phenomenon to demonstrate cotton improvement strategies (Ali *et al.*, 2009). In this investigation, simple correlation exhibits correlation between the study traits. Lint index showing a direct positive correlation with cottonseed index, and have positive correlation with cottonseed yield. Boll weight, whereas on the other side, it demonstrates negative correlation with number of bolls, and have negative correlation with staple length. And the number of boll was established to be positive correlated with some quantitative traits as shown in (Table 1) that are confirming the findings of (Salahuddin *et al.*, 2010), who describes positive correlation between quantitative traits of cotton, such as ginning outturn percentage %, cotton seed, and number of bolls. Clement *et al.*, (2011) described the negative correlation between fiber and yield quality traits in cotton. It also demonstrates that fiber strength and fiber length had negative correlation with cottonseed, while the positive correlation was exhibit with fiber maturity in respective manners.

Cluster analysis: Cluster analysis is a method that aims to clarify and classify a sample of objects (or subject) based on a set of deliberately chosen variables into a number of different groups keeping alike objects in the similar group. An example used in the field of psychopathology, where the characterization of patients on the basis of clusters of symptoms can be used to identified or classified of an appropriate form of the ropy. The cluster analysis has no mechanism for distinguishing irrelevant and relevant variables. Using cluster analysis, 12 cultivars were divided into three clusters based on different parameters. Cluster-1 genotypes with functional values for fiber strength, boll weight, and fiber uniformity. In similar manner cluster-2 composed of genotypes holding significant values for cotton yield, lint index, and fiber strength while smaller values were recorded for fiber uniformity, and boll weight (Alishah, 2001). While cluster-3 was obtained with higher values for number of bolls, seed index, fiber length, and cottonseed yield, and fiber uniformity. However, the other traits from this cluster could not be extracted.

Table 2. Squared cosines for yield and yield contributing factors of cotton.

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Symp	0.7495	0.0116	0.0053	0.1023	0.0986	0.0160	0.0134	0.0034	0.0000
BW	0.3082	0.6663	0.0144	0.0003	0.0007	0.0060	0.0036	0.0002	0.0002
GOT	0.5807	0.0456	0.1992	0.0544	0.1080	0.0000	0.0058	0.0062	0.0000
MiC	0.2811	0.0903	0.5327	0.0507	0.0212	0.0161	0.0064	0.0015	0.0000
Bno	0.7754	0.1133	0.0269	0.0030	0.0285	0.0120	0.0367	0.0042	0.0000
PHt	0.4280	0.2413	0.0097	0.2773	0.0291	0.0083	0.0016	0.0047	0.0000
SL	0.5416	0.1491	0.2139	0.0275	0.0285	0.0255	0.0087	0.0052	0.0000
SI	0.6491	0.0747	0.0714	0.0992	0.0658	0.0030	0.0071	0.0027	0.0000
YLD	0.4439	0.5445	0.0019	0.0001	0.0001	0.0063	0.0014	0.0016	0.0002

Values in bold correspond for each variable to the factor for which the squared cosine is the largest

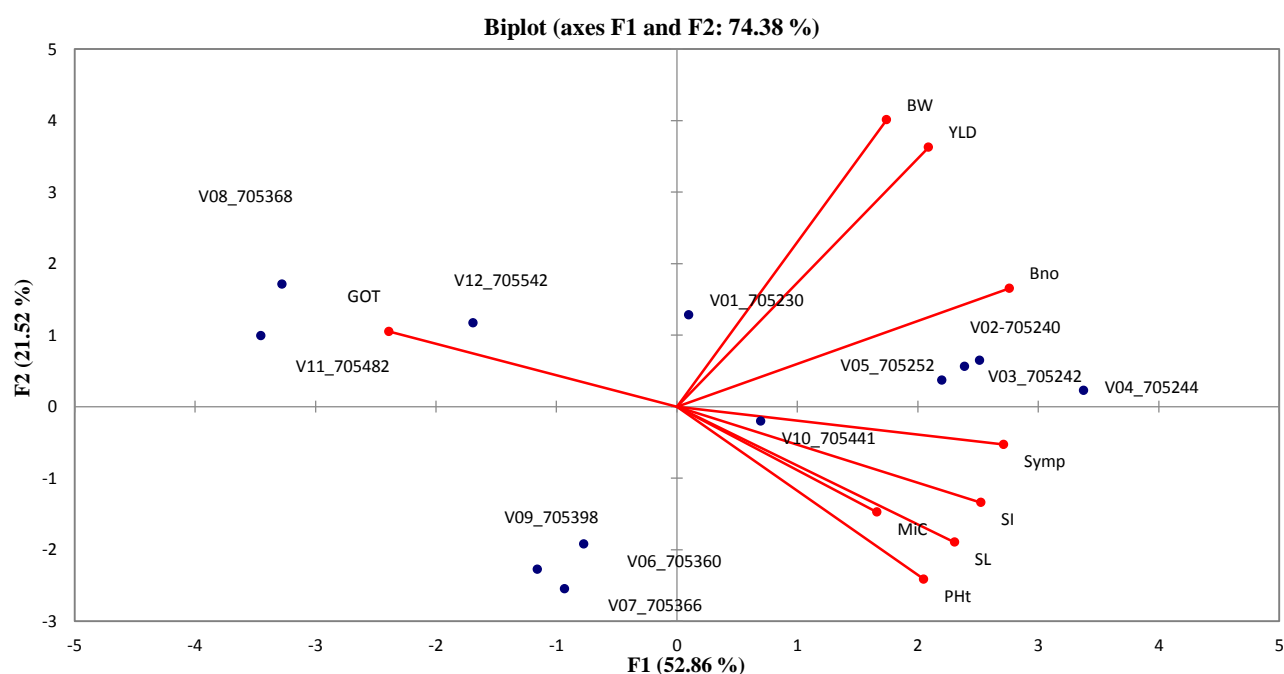


Fig. 4. Biplot among PC-1 and PC-2 display contribution of different traits in variability of upland cotton (*Gossypium hirsutum* L.).

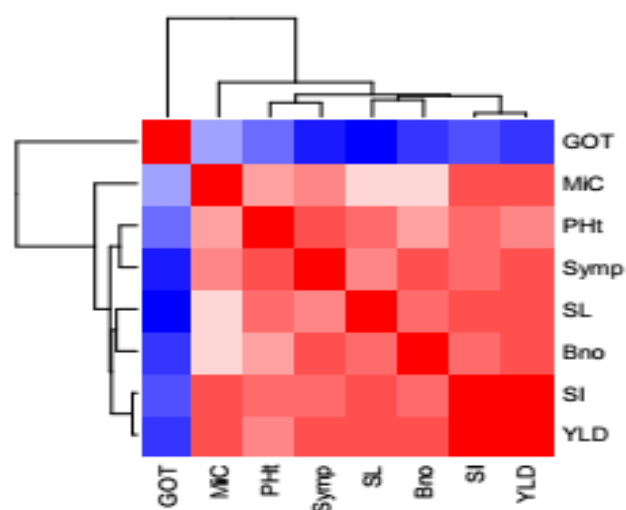


Fig. 5. A heat map of 12 Chinese cotton cultivars showing the relatedness between yield and fiber related traits.

To calculate the appropriate genotypic variability existing among all studies clusters, a Ward's dendrogram was constructed in the (Fig. 6) as represented previously by (Nazir *et al.*, 2013; and Grenier *et al.*, 2000). The tree diagram showed the presence of wide variation among clusters. Based on clusters and tree diagram analysis, the cultivars are including 705230, 705240, 705242, 705244, 705252, 705360, 705366, 705368, 705398, 705441, 705482, and 705542 are highly recommended to be used for the evaluation and identification of relationship between genotypes. The statistical implements could be used for the designation of other potential resources. The dendrogram resulting from Wards clustering demonstrates variation among the clusters, is represented in Fig. 4.

A heat model map is a diagrammatically represents the data values in different colors. Actually the heat map was originated by trademarked software, namely designer

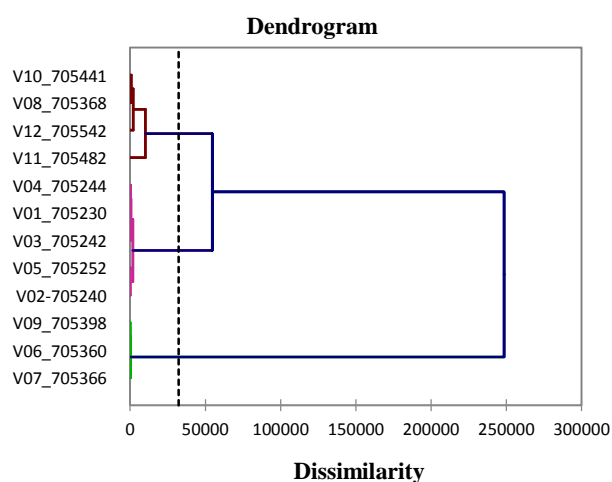


Fig. 6. Dendrogram resulting from Ward's clustering exhibits the presence of high variation between the clusters suggesting high genetic variability among genotypes.

Cormac Kinney in (1991), to demonstrate a 2D matrix portraying market information, through the likely plots viz., as shading matrices have subsist for over a century.

In the heat map of correlation, coefficient seed index was widely correlated with yield, and number of bolls per plant (Méndez-Natera, 1996), staple length with number of bolls, whereas fiberfineness, sympodial branch, and plant height were significantly correlated with each others, as well as ginning out turn percentage% and plant height both were found significantly correlated with each other's (Fig. 5). Ginning out turn percentage %, micronaire value, plant height, and sympodial branches were highly significant and discovered to be in correlation with each others, whereas the seed index, number of bolls per plant, fiber length, and sympodial branches were positively correlated with fiber length, and plant height.

Conclusion

In conclusion, the characters showing direct positive correlation between each other may be further assessed through early generation selection. Based on biplot and correlation analysis number of bolls per plant, boll weight, cottonseed yield, seed and lint-index can be used as selection criteria for breeding program methods of high yielding cotton varieties. Overall, characters including cottonseed yield, number of bolls, staple length, and seed index exhibited higher values among different clusters and can be further used as selection criteria to increase the cottonseed yield in different traits.

Acknowledgements

I am very thankful to my Professor Dr. Fan Shuli and Ma Qifeng they teach me and make me able to write this Research article, and I am also thankful to my all my lab fellows, they help me where i need their help. We are also thankful to Cotton Research Institute Anyang; to provide us seed for this experiment. This Research project was financially supported by the National Natural Science Foundation of China (Grant No. 31701474), Henan Province key R & D and extension Project (182102110048) and Key Scientific Research projects in Henan Colleges and Universities (17B180001). We are grateful to an anonymous referee for improvement of the manuscript.

References

- Akter, A., M.J. Hasan, A.K. Paul, M. Mutlib and M.K. Hossain. 2009. Selection of parent for improvement of restorer line in rice (*Oryza sativa* L.). *SAARC J. Agri.*, (7): 43-50.
- Ali, M.A., I.A. Khan and N.N. Nawab. 2009. Estimation of genetic divergence and linkage for fibre quality traits in upland cotton. *J. Agri. Res., Lahore*, 47(3):229-236.
- Ali, M.A., N.N. Nawab, A. Abbas and M. Sajjad. 2009. Evaluation of selection criteria in *Cicer arietinum* L. using correlation coefficients and path analysis. *Aust. J. Crop Sci.*, 3: 65-70.
- Alishah, E. 2001. Study of morphological traits and genetic variation in different genotypes of upland cotton (*Gossypium hirsutum* L.) in Iran. *Seed & Pl. Imp. J.*, (17): 44-60.
- Anderson, E. and J.W. Boyes. 1957. An experimental investigation of judgments concerning genera and species. *Evolution*, 11(2): 260-263.
- Ashok, K., R. Ravikesavan and J.K.S. Sivas. 2010. Combining ability estimates for yield and fibre quality traits in Line X Tester crosses of upland cotton (*Gossypium hirsutum*). *Int. J. Biol.*, 2(1): 179-183.
- Ashokkumar, K. and R. Ravikesavan. 2011. Morphological diversity and performance in upland cotton (*Gossypium hirsutum* L.). *J. Agri. Sci.*, 3: 107-113.
- Clement J.D., G.A. Constable, W.N. Stiller and S.M. Liu. 2012. Negative associations still exist between yield and fibre quality in cotton breeding programs in Australia and USA. *Field Crops Res.*, 128: 1-7.
- Everitt, B.S., S. Landau and M. Leese. 2001. Cluster Analysis. 4th Edition, Arnold, London.
- Gopinath, M., S. Rajamani, R.K. Naik and C.H.M. Rao. 2009. Genetic divergence for lint characters for upland cotton (*G. hirsutum*). *J. Cotton Res. Devel.*, 23: 18-22.
- Jian, C., L. Wei, L. Ruili and L.F.W. Feng. 2006. Genetic diversity detected by cluster analysis North of Anhui major wheat cultivars. *Chinese Agri. Sci. Bull.*, 11: 031.
- Li, R.L., Z. Long and J. Chen. 2005. Comparison of fibre quality characteristics of different varieties of hybrid cotton. *J. Hunan Agri. Univ.*, 31: 492.
- Malik, W., M.Z. Iqbal, A.A. Khan, E. Noor, A. Qayyum and M. Hanif. 2011. Genetic basis of variation for seedling traits in *Gossypium hirsutum* L. *Afr. J. Biotech.*, 10: 1099-1105.
- May, O.L., D.T. Bowman and D.S. Calhoun. 1995. Genetic diversity of US Upland cotton cultivars released between 1980 and 1990. *Crop Sci.*, 35: 1570-1574.
- McCarty, J.C., J.N. Jenkins and J. Wu. 2004. Primitive accession derived germplasm by cultivar crosses as sources for cotton improvement: II. Genetic effects and genotypic values. *Crop Sci.*, 44: 1231-1235.
- McCarty, J.C., J.N. Jenkins and J. Wu. 2005 Primitive accession derived germplasm by cultivar crosses as sources for cotton improvement. *Crop Sci.*, 44: 1231-1235.
- Méndez-Natera, J.R. 1996. Analysis de los coeficientes de correlacion lineal y de los coeficientes de trayectoria en algodón (*G. hirsutum* L.). *SABER.*, 8(1): 56-62.
- Natera, M.J.R., A. Rondon, J. Hernandez and J.F.M. Pinto. 2012. Genetic studies in upland cotton. III. Genetic parameters, correlation and path analysis. *SABRO J. Breed. Genet.*, 44: 112-128.
- Nazir, A., J. Farooq, A. Mahmood, M. Shahid and M. Riaz. 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., P. Annicchiaro and A.B. Damania. 1996. Geographic variation in tetraploid wheat (*Triticum turgidum* spp. *Turgidum* var. *Durum*) landraces from two provinces in Ethiopia. *Genet. Resour. Crop Evol.*, 43:395-407.
- Quailing, W. and L.Zhe. 2011. Principal Component analysis of F2 individual selection in upland cotton (*Gossypium hirsutum* L.). *J. Henan Inst. Sci. & Technol.*, (Natural Sci. Ed.), 5: 004.
- Rahman, M. T. Shaheen, N. Tabbasam, M.A. Iqbal, M. Ashraf, Y. Zafar and A.H. Paterson. 2012. Genetic resources in cotton and their improvement. *Agron. Sustain. Devel.*, 32: 419-432.
- Rathinavel, K. 2017. Exploration of genetic diversity for qualitative traits among the extant upland cotton (*Gossypium hirsutum* L.) varieties and parental lines. *Int. J. Curr. Microbiol. App. Sci.*, 6(8):2407-2421.
- Rehman, S.U., M.A. Abid, M. Bilal, J. Ashraf, S. Liaqat, R.I. Ahmed and G. Qambar. 2015. Genotype by trait analysis and estimates of heritability of wheat (*Triticum aestivum* L.) under drought and control conditions. *Basic Res. J. Agric. Sci. Rev.*, 4(4):127-134.
- Saeed, F., R.H. Shabbir, J. Farooq, M. Riaz and K. Mahmood. 2015. Genetic diversity analysis for earliness, fiber quality and cotton leaf curl virus in *Gossypium hirsutum* L. *Accessions. Cotton. Genom. Genet.*, 6: 1-7.
- Salahuddin, S., S. Abro, M.M. Kandhro, L. Salahuddin and S. Laghari. 2010. Correlation and path coefficient analysis of yield components of upland cotton (*Gossypium hirsutum* L.). *World Appl. Sci. J.*, 8: 71-75.
- Satish, Y., P.P. Jain and B.S. Chhabra. 2009. Studies on genetic diversity in upland cotton (*G. hirsutum* L.). *J. Cotton Res. Devel.*, 23: 18-22.
- Shakeel, A., I. Talib, M. Rashid, A. Saeed, K. Ziaf and M.F. Saleem. 2015. Genetic diversity among upland cotton genotypes for quality and yield related traits. *Pak. J. Agri. Sci.*, 52(1): 73-77.
- Shannon-Weaver. 1949. Diversity Index. In: Kennish, M.J. (Ed.), *Encyclopedia of Estuaries*. Springer Netherlands, Dordrecht, pp. 572-573.
- Sharma, J. R. 2006. Statistical and biometrical techniques in plant breeding. *New Age Inter. Publi.*, 432 pp.
- Sharma, J.R. 1998. Statistical and Biometrical Techniques in Plant Breeding. New Age International (P) Limited Publishers, New Delhi. 432p.
- Sneath, P.H.A. and R.R. Sokal. 1973 Numerical taxonomy: the principles and practice of numerical classification. W. F. Free-Man and Co. San Francisco, 573 pp.
- Ulloa, M., R.B. Hutmacher, R.M. Davis, S.D. Wright, R. Percy and B. Marsh. 2006. Breeding for fusarium wilt race 4 resistances in cotton under field and greenhouse conditions. *J. Cotton Sci.*, 10: 114-127.
- Van-Esbroeck, G.A. and D.T. Bowman. 1998. Cotton germplasm diversity and its importance to cultivar development. *J. Cot. Sci.*, 2: 121-129.

(Received for publication 25 March 2018)