Research Article
Genetic diversity analysis in American cotton (*Gossypium hirsutum* L.)

B. Mansingh Naik, Y. Satish* and D. Ratna Babu
Regional Agricultural Research Station, Lam,
Guntur, Andhra Pradesh – 522034
E-mail: tej.satish1@gmail.com

Abstract
Genetic divergence was studied among 50 genotypes of American cotton (*Gossypium hirsutum* L.) using Mahalanobis D^2 statistic and principal component analysis. On the basis of D^2 statistic, 50 genotypes were grouped into 8 clusters and PCA identified six principal components which explained 81.99% of the total variability among the genotypes. Based on intra- and inter-cluster distance among the groups, it is suggested to make crosses between the genotypes of cluster VI (ARBH 1402) and cluster VII (HYPS 152), cluster I (SSGR 105) and cluster VIII (L 788), cluster I (SSGR 105) and cluster IV (L 799) and between cluster V (GJHV 497) and cluster VIII (L 788) for heterosis breeding. Divergence studies indicated that geographical diversity is always not necessarily associated with the genetic diversity. Multivariate analysis revealed the presence of maximum divergence among L 788, HS 294, MCU 5, HYPS 152 and GISV 267 reporting their parental eligibility in exploitation of hybrid vigour.

Introduction
Cotton (*Gossypium* spp.) is the most important renewable natural fiber crop of global importance and continues to be the predominant and sustainable fiber in the Indian textile scenario (Asha *et al*., 2015). It plays a key role in the national economy in terms of its contribution to trade, industrial activities, employment and foreign exchange earnings (Gite *et al*., 2006). Despite stiff competition from the manmade synthetic fiber cotton retains its reputation as "queen of the fiber plants". In Andhra Pradesh, cotton is a crop of great economic value and playing a significant role in the socio-economic status of the farmers. It occupies an area of about 6.63 lakh hectares with an annual production of 19.10 lakh bales (170 kg each) and productivity of 641 kg lint ha^{-1} (AICRP, Annual Report, 2015-16).

To broaden the genetic base through hybrid breeding program, the genetic divergence among available germplasm is prerequisite. In order to develop hybrids in cotton successfully, choice of suitable parents through careful and critical evaluation is of paramount importance. The quantification of the degree of divergence present in the population is of immense value in identifying diverse genotypes for recombination breeding programmes. Mahalanobis D^2 statistic is a powerful tool for quantifying genetic divergence in germplasm collections with respect to the characters considered together. Hence, genetic divergence was estimated using Mahalanobis D^2 statistic in 50 genotypes of American cotton in order to initiate a systematic hybrids or recombinants development programme.

Materials and methods
The present investigation was conducted during *kharif*, 2015-16 in randomized block design with 50 genotypes with three replications following 105 cm x 60 cm spacing at Regional Agricultural Research Station, Lam, Andhra Pradesh. The observations were recorded on five randomly selected plants from each genotype per replication for 15 characters *viz*., plant height (cm), days to 50% flowering, number of monopodia, number of sympodia per plant, number of bolls per plant, boll weight (g), ginning out-turn (%), seed index (g), lint index (g), 2.5% span length (mm), micronaire (10^{-6}g/in), bundle strength (g/tex), uniformity ratio, seed cotton yield per plant (g) and lint yield per plant (g). The quality parameters were studied at Central Institute for Research on Cotton Technology (CIRCOT), RARS, Lam, Guntur, Andhra Pradesh. The data were statistically analyzed to study diversity by Mahalanobis D^2 statistic as per Rao (1952) and principal component analysis (PCA) as described by Jackson (1991).

Result and discussion
Based upon the D^2 analysis, all the 50 genotypes grouped into 8 clusters (Table1 & Fig.1), out of which cluster II possessed the largest number of genotypes (22) followed by cluster III (12), cluster I (8), cluster VII (4) and clusters IV, V, VI and VIII were solitary clusters with single genotype and with zero intra-cluster D^2 values.

The per cent contribution towards genetic divergence was maximum by seed index (21.80) followed days to 50 per cent flowering, boll weight (17.71), 2.5 % span length (6.53), micronaire (4.98), ginning out turn (4.57), plant height (3.35), number of bolls per plant (0.90), bundle strength

(0.82), seed cotton yield per plant (0.57), number of bolls per plant (0.33), lint yield per plant (0.33), number of sympodia per plant (0.16) and uniformity ratio (0.00) (Table 2).

The maximum intra cluster distance was observed for cluster VII (70.13), followed by cluster III (47.79), cluster II (38.88) and cluster I (30.59), while, it was zero for clusters IV, V, VI and VIII (Table 3) single the clusters had one member each. The high intra-cluster distance in cluster VII indicates the presence of wide genetic diversity among the genotypes present within this cluster.

The inter-cluster distance was maximum between cluster VI and VII (296.37) followed by cluster I and VIII (275.78), cluster I and IV (195.67) and cluster V and VIII (194.82) (Table 3) indicating the wide genetic diversity between these clusters. Based on these studies crosses can be made between the genotypes of these clusters to obtain promising hybrids or desirable transgressive segregants after testing for the combining ability and gene action.

The inter cluster distance was found to be minimum (133.11) between cluster III and VII suggesting a close relationship between members of them and a low degree of diversity among the genotypes. Since, the magnitude of heterosis largely depends on degree of genetic diversity among parents and hence, selection of parents from these clusters should be avoided for combination breeding.

Intra-and inter-cluster distances obtained by Mahalanobis D^2 statistic indicated that crosses between the genotypes of cluster VI (ARBH 1402) and cluster VII (HYPS 152), between the genotypes of cluster I (SSGR 105) and VIII (L 788), between cluster I (SSGR 105) and IV (L 799), between cluster V (GJHV 497) and VIII (L 788) might be resulted with promising hybrids followed by segregants with variability.

Principal component analysis (PCA) identified six principal components (PCs), which contributed 81.99 per cent of cumulative variance. The first principal component (PC_1) contributed maximum towards variability (27.33) followed by PC_2 (19.12), PC_3 (13.70), PC_4 (9.37), PC_5 (6.78) and PC_6 (5.71) (Table 4). The significant factors loaded in PC1 towards maximum genetic divergence were seed cotton yield per plant, bundle strength, lint index, micronaire, number of bolls per plant, number of sympodia per plant, plant height and uniformity ratio. This PC analysis revealed the presence of wide divergence between L 788, MCU 5, HYPS 152, GISV 267, TSH 0499, SSGR 105, ARBH 1402 and HS reporting their parental eligibility in exploitation of hybrid vigour.

General notation exists that the larger is the divergence between the genotypes, the higher will be the heterosis (Falconer, 1964). Therefore, it would be desirable to attempt crosses between the genotypes belonging to distant clusters for getting highly heterotic crosses which are likely to yield a wide range of segregants on which selection can be practiced. However, *per se* performance of the identified line also should be taken into consideration for selecting lines for any hybridization programme. The genotypes L 788, HS 294, MCU 5, HYPS 152 and GISV 267 showed maximum inter-cluster distance in Mahalanobis D^2 analysis, principal component analysis and also have better per se performance in sympodia per plant, number of bolls per plant, boll weight, seed index, lint index and quality characters. So they can be exploited for the development of heterotic hybrids in future breeding programmes.

References

AICRP Annuual Report. 2015-16. All India Coordinated Research Project on cotton. Coimbatore, Tamilnadu, India.

Asha, R., Ahamed, M.L., Babu, D.R and Kumar, P.A. 2015. Character association and path coefficient analysis for yield and component traits in upland cotton. *J. Cotton Res. and Devl.*, **29**(1): 31-35.

Falconer, D.S. 1964. An Introduction to Quantitative Genetics. Second Edition. Oliver and Boyd, Edinburgh, London., 312-324.

Gite, V.K., Misal, M. B and Kalpande, H.V. 2006. Correlation and path analysis in cotton (*Gossypium hirsutum* L.). *J. Cotton Res. and Devl.*, **20**: 51-54.

Jackson J.E. 1991 A User's Guide to Principal Components. John Wiley and Sons, NewYork.

Mahalanobis, P.C. 1928. A statistical study at Chinese head measurement. *J. Asiatic Soc. Bengal*, **25**: 301-307.

Rao, C.R. 1952. Advanced Statistical Methods in Biometrical Research. John Wiley and Sons, New York., 236-272.

Table 1. Clustering pattern of 50 genotypes of American cotton

Cluster No.	Number of genotypes	Name of the genotype
I	8	SSGR 105, RAH 1033, PBH 10, F 2501, PBH 13, SCS 1061, TCH 1806, SCS 1207
II	22	TCH 1741, SCS 1214, TSH 0533-2, HS 294, RAH 1066, CNH 1118, GJHV 510, GTHV 13/37, CSH 2838, L 604, CNH 5, TSH 0533-1, CCH 14-1, H 1442, ARBH 1401, GSHV 177, NA 1325, LRK 516, NDLH 1938, L 1060, RS 2765, BS 23
III	12	CCH 14-2, L 1008, H 1471,RS 2767, F 2493, CPD 1401, LH 2253, SURAJ, BS 26, CPD 1402, L 389, L 770
IV	1	L 799
V	1	GJHV 497
VI	1	ARBH 1402
VII	4	HYPS 152, TSH 0499, GISV 267, MCU 5
VIII	1	L 788

Table 2. Contribution of different characters towards genetic divergence in 50 cotton (*G. hirsutum* L.) genotypes

S. No.	Character	Contribution towards divergence (%)	Times ranked first
1	Plant height (cm)	3.35	41
2	Days to 50% flowering	21.80	267
3	Number of monopodia per plant	0.33	4
4	Number of sympodia per plant	0.16	2
5	Number of bolls per plant	0.90	11
6	Boll weight (g)	17.71	217
7	Ginning out-turn (%)	4.57	56
8	Seed index (g)	37.80	463
9	Lint index (g)	0.16	2
10	2.5% span length (mm)	6.53	80
11	Micronaire (10^{-6}g/in)	4.98	61
12	Bundle strength (g/tex)	0.82	10
13	Uniformity ratio	0.00	0
14	Seed cotton yield per plant (g)	0.57	7
15	Lint yield per plant (g)	0.33	4

Table 3. Average intra- and inter-cluster D^2 values of 8 clusters in 50 cotton genotypes

Cluster No	I	II	III	IV	V	VI	VII	VIII
I	**30.59**	71.15	130.37	195.67	95.59	138.46	175.90	275.78
II		**38.88**	68.25	86.96	60.21	77.86	145.89	177.85
III			**47.79**	80.07	85.97	92.85	133.11	80.02
IV				**0.00**	58.37	151.40	124.10	136.14
V					**0.00**	181.93	75.38	194.82
VI						**0.00**	296.37	170.16
VII							**70.13**	202.06
VIII								**0.00**

Note: Bold and diagonal values indicate intra-cluster D^2 distance

Table 4. Eigen values, proportion of the total variance represented by first seven principal components, cumulative per cent variance and component loading of different characters in cotton (*G. hirsutum* L.)

| Traits	PC$_1$	PC$_2$	PC$_3$	PC$_4$	PC$_5$	PC$_6$
Eigene Value (Root)	4.10	2.87	2.06	1.41	1.02	0.86
% Var. Exp.	27.33	19.12	13.70	9.37	6.78	5.71
Cum. Var. Exp.	27.33	46.44	60.14	69.51	76.29	81.99
Plant Height (cm)	0.04	0.30	0.16	0.54	0.08	0.20
Days to 50% flowering	-0.24	0.31	-0.30	0.16	-0.33	-0.10
No. of monopodia/ Plant	-0.09	-0.26	-0.37	-0.31	-0.17	0.03
No. of sympodia/ Plant	0.05	0.37	0.26	0.12	0.40	-0.29
No. of bolls/ Plant	0.12	-0.18	-0.19	0.60	-0.28	0.28
Boll Weight (g)	-0.11	0.29	0.31	-0.36	-0.13	0.47
Ginning out-turn (%)	-0.35	0.21	-0.00	-0.07	-0.08	0.32
Seed Index (g)	-0.38	-0.25	0.22	0.18	0.16	0.24
Lint Index (g)	0.22	0.38	-0.19	-0.08	-0.08	-0.03
2.5% Span length (mm)	-0.35	0.20	-0.17	-0.06	0.25	0.23
Micronaire (10^{-6} g/Inch)	0.20	-0.04	-0.41	0.02	0.27	0.45
Bundle Strength (g/tex)	0.30	-0.06	-0.20	-0.08	0.56	0.20
Uniformity Ratio	0.01	-0.43	0.39	0.03	0.03	0.12
Seed Cotton Yield/ Plant(g)	0.46	0.06	0.08	-0.04	-0.18	0.05
Lint Yield/ Plant (g)	-0.35	-0.10	-0.26	0.15	0.28	-0.32

PC = Principal component
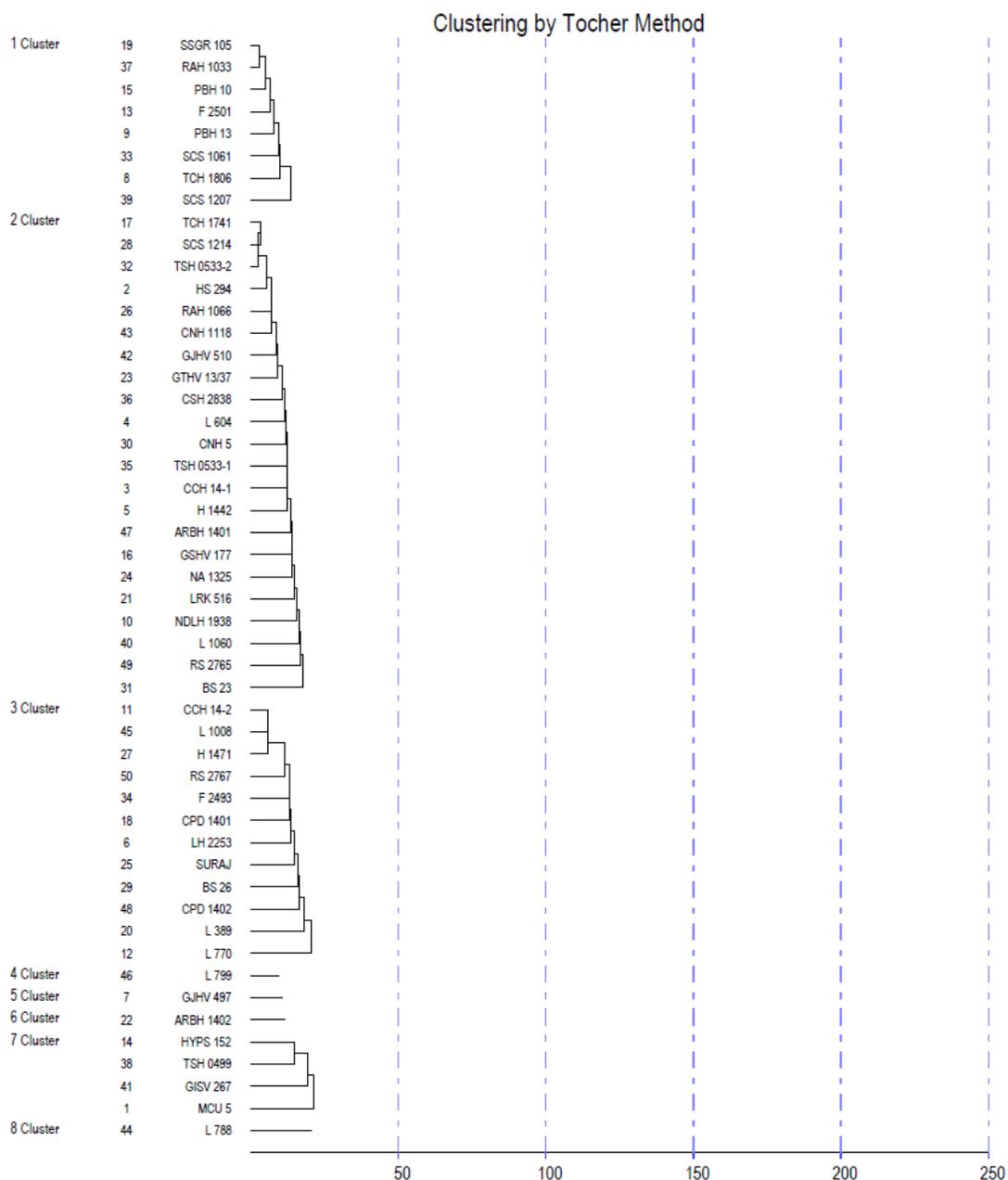
Fig. 1. Dendrogram showing relationship among 50 cotton (*G. hirsutum* L.) genotypes in eight clusters based on Mahalanobis D² values