



Multivariate Analysis in Upland Cotton (*Gossypium hirsutum* L.)

R. Asha, M. Lal Ahamed*, D. Ratna Babu and P. Anil Kumar

Department of Genetics and Plant Breeding
Acharya N.G. Ranga Agricultural University, Bapatla-522101

An experiment was conducted to analyze the genetic diversity among 40 genotypes for 15 quantitative characters which were grouped into 7 clusters based on Mahalanobis D_2 analysis with cluster I and III being the largest each with 8 genotypes followed by 7 genotypes in cluster VII. Based on Hierarchical cluster analysis all the genotypes were grouped into 7 clusters with cluster VII being the largest each with 13 genotypes followed by 8 genotypes in cluster IV. This random distribution of genotypes indicated absence of parallelism between geographical and genetic diversity. In Principal component analysis first seven principal components with eigen value more than one contributed 87.98 per cent towards the total variability with PC₁ alone showing maximum of 22.27 per cent variability.

Key words: Cotton, Mahalanobis D_2 analysis, Hierarchical cluster analysis, Principal component analysis

Cotton is the most important natural textile fibre and sixth largest vegetable oil source in the world, and is the cornerstone of textile industries worldwide

.Because of the revolution of textile technology, the cotton industry needs raw cotton with higher quality and quantity. So, it is imperative to develop high yielding cotton cultivars with low cost of production. For this a better knowledge about the genetic diversity of cotton is warranted for exploitation of existing variability. It also plays an important role in the manifestation of heterosis. Hybrids between genotypes of diverse origin display a greater heterosis than those hybrids involving closely related parents. Multivariate analysis has been extensively used as a quantitative measure to identify diverse genotypes. The objective of this study was to assess the genetic diversity and relationship among the 40 cotton genotypes using multivariate Mahalanobis D_2 statistics, Hierarchical cluster analysis and Principal component analysis.

Materials and Methods

The experiment was conducted during *kharif* 2011-12 in randomized block design with 40 genotypes obtained from all over India with three replications following 120 x 60 cm spacing at Agricultural College Farm, Bapatla, Andhra Pradesh. The soils are black cotton type with clay texture. Recommended doses of fertilizers were applied. Each plot consisted of two rows of 6m length and observations were recorded on ten randomly selected plants from each genotype per replication for 15 characters *viz.*, plant height (cm), number of monopodia plant per plant, number of sympodia plant per plant, number of bolls per plant, boll weight (g), seed index (g), lint index (g), ginning out-turn (%). The data on days to 50 per cent flowering, ginning out-turn (%), bundle strength (g/tex), uniformity ratio, 2.5 per cent span length

(mm), micronaire (10⁻⁶g/in), uniformity ratio and fibre elongation (%) were recorded on plot basis. Mahalanobis D_2 analysis (Tocher's method) was worked out as given by Rao (1952). Agglomerative hierarchical clustering technique (Ward's minimum variance) was followed for cluster analysis as given by Anderberg (1993). PCA was performed as per Jackson (1991).

Results and Discussion

The analysis of variance revealed highly significant differences among 40 genotypes for 15 quantitative traits (Table 1) indicating the existence of variability among genotypes for characters studied. The per cent contribution towards genetic divergence by all the 15 characters is presented in Table -2. Based on D_2 statistic all the 40 genotypes were grouped into 7 clusters (Table-3) using Tocher's method with the criterion that the intra-cluster average D_2 values should be less than the inter-cluster D_2 values. The distribution of 40 genotypes into seven clusters was at random with maximum number of genotypes in clusters I and III (8 genotypes each). Cluster VI (7 genotypes) was the second largest followed by cluster IV and VI (5 genotypes), cluster II with 4 genotypes and cluster V with 3 genotypes. The mutual relationships between the clusters were represented diagrammatically by taking average intra and inter cluster D values. This was in accordance with the results Eswara Rao *et al.* (2009), Gopinath *et al.* (2009), Venkateswarulu *et al.* (2010) and Srinivasulu *et al.* (2010). The average intra- and inter-cluster D and D_2 values are presented in Table 4. The inter cluster distance was maximum between clusters IV and VII followed by clusters V and IV; III and VII ; IV and VI ; I and VII . Based on the inter- cluster distances among the groups, crosses after confirming the general combining ability of the genotypes from clusters VII and IV may yield better and desirable segregants.

*Corresponding author email: lalahamed@gmail.com

Table 1. Analysis of variance for yield and yield components in cotton (*Gossypium hirsutum* L.)

Source of Variation	d.f.	Plant height (cm)	Days to 50% flowering	No. of monopodia plant ⁻¹	No. of sympodia plant ⁻¹	No. of bolls plant ⁻¹	Boll weight (g)	Ginning-out-turn (%)	Seed index (g)
Mean sum of squares									
Replications	2	222.1830	1.1083	0.0001	3.2480	20.4016	0.1341	1.7510	0.0731
Treatment	39	646.1204**	10.5638**	0.3341**	17.2260**	162.6393*	1.0978**	14.5549**	3.4096**
Error	78	109.4772	0.3561	0.0053	3.0174	9.5989	0.0540	2.2439	0.0808

Source of variation	d.f.	Lint index (g)	2.5% span length (mm)	Micronaire (10 ⁶ g/in)	Bundle strength (g/tex)	Uniformity ratio	Elongation (%)	Seed cotton yield plant ⁻¹ (g)
Mean sum of squares								
Replications	2	0.0306	0.1874	0.0861	0.9836	5.5466	0.0432	51.2666
Treatment	39	1.0704**	16.8824**	1.0115**	10.4506**	23.2764**	0.1460**	1387.4840**
Error	78	0.0631	0.9509	0.0613	0.5016	2.3496	0.0298	92.7402

** = Significance at 1% level, d.f. = Degrees of freedom

In hierarchical clustering (Ward's minimum Table 2.variance) method 40 genotypes were grouped into seven clusters. Among all the clusters, cluster VII was largest containing 13 genotypes followed by cluster

Table 2. Contribution of different characters towards genetic divergence in 40 genotypes of cotton (*Gossypium hirsutum* L.)

Source	Times Ranked first	Contribution %
Plant height cm	42	5.38
Days to 50% flowering	67	8.59
No. of monopodia plant ⁻¹	170	21.79
No. of Sympodia plant ⁻¹	3	0.38
No. of Bolls plant ⁻¹	20	2.56
Boll weight (gm)	39	5.00
Ginning out-turn %	4	0.51
Seed index(g)	198	25.38
Lint index(g)	28	3.59
2.5 %Span length(mm)	33	4.23
Micronaire(10 ⁶ g/in)	43	5.51
Bundle strength (g/tex)	55	7.05
Uniformity ratio	18	2.31
Fibre Elongation %	4	0.51
Seed cotton Yield plant ⁻¹	56	7.18

IV with 8 genotypes, cluster VI with 6 genotypes, cluster III and I each with 5 genotypes, cluster II with 2 genotypes and cluster V with 1 genotype (Table-5).

Table 3. Clustering of 40 genotypes of cotton (*Gossypium hirsutum* L.) by Tocher's method

Cluster No.	No. of genotypes	Name of genotype(s)
I	8	G.cot 12, HLS 772, TSH 9974, L763, L 765, L603 ,L604, L788
II	4	G. ageti, KDCKAD, TCH 1716, L 761
III	8	HS 271, JK 276-4, NDLH 1938, NISC 40, NDLH 1939, RAH 216, Narasimha, L 766 HAG 812, JK-5, KH 121 ,KGL 54620, Pee Dee 0113
IV	5	H 492, NA 1584, RS 2557
V	3	GJHV 338, JK 276-10-5 ,P 403, RS 419, VIKRAM
VI	5	GSHV 97/291, GBHV 164, NA 1290 BP, RAH 902, RAH 912, L 769 ,L 770
W	7	

The average intra and inter cluster distances are presented in Table-6. Maximum inter- cluster distance was observed between clusters VI and V; clusters II and V; clusters III and V.

Table 4. Average intra-and inter-cluster D2 values among eight clusters in 40 cotton genotypes (*Gossypium hirsutum* L.)

Cluster No.	I	II	III	IV	V	VI	VII
I	24.69 (4.96)	40.29	74.26	60.72	84.48	228.59	175.39
II		32.65 (5.71)	60.73	67.39	81.16	185.92	127.63
III			52.17 (7.22)	80.76	107.3	121.18	103.77
IV				52.07 (7.21)	138.86	244.74	156.50
V					81.17 (9.00)	192.70	151.18
VI						0.00 (0.00)	87.81
VII							0.00 (0.00)

Note: Bold and diagonal values indicate intra-cluster D2 distance; figures in parentheses are D values

towards the total variability (Table .7). The hierarchical cluster analysis and PCA confirmed findings of each

Table 5. Clustering of 40 cotton (*Gossypium hirsutum* L.) genotypes by Ward's minimum variance method

Cluster No	No. of genotypes	Name of genotype(s)
I	5	GSHV 97/291,GBHV 164,TSH 9974,JK-5,HAG 812
II	2	L 761,L 763
III	5	L 765,L 788,L 603,L 604,HLS 72
IV	8	G. ageti, KDCKAD,G. cot 12,VIKRAM,KH 121,NA 1290 BP,RAH 912 ,Pee Dee 0113
V	1	TCH 1716
VI	6	GJHV 338,JK-276-10-5,HS 271,P 403,H 492,NA 1584
VII	13	JK-276-4,NISC 40,NDLH 1939,NARASIMHA,KGL 54620,NDLH 1938, RS 2557,RAH 216,L 770,RS 419,L 766, RAH 902,L 769

other. The above results are supported by Vijaya Lakshmi et al. (2009), Srinivasulu et al. (2010) and

Table 6. Average intra and inter- cluster Euclidian2 values among the seven clusters in 40 cotton genotypes (*Gossypium hirsutum* L.)

Cluster No	I	II	III	IV	V	VI	VII
I	183.56	403.89	278.53	271.90	617.59	279.43	259.68
II		263.45	291.35	460.61	710.56	551.61	557.07
III			149.39	280.03	632.03	229.40	256.95
IV				151.80	374.83	283.30	198.15
V					0.00	773.77	532.58
VI						104.81	171.95
VII							113.53

Bold and diagonal values indicate intra-cluster distance

Table 7. Eigen values, proportion of the total variance represented by first eight principal components, cumulative per cent variance and component loading of different characters in cotton (*Gossypium hirsutum* L.)

	PC ₁	PC ₂	PC ₃	PC ₄	PC ₅	PC ₆	PC ₇
Eigene Value (Root)	3.341	2.738	2.152	1.557	1.313	1.086	1.011
% Var. Exp.	22.271	18.253	14.346	10.380	8.751	7.239	6.739
Cum. Var. Exp.	22.271	40.523	54.869	65.248	74.000	81.238	87.977

Altaher and Singh (2003).

All the three methods of grouping revealed a single concept of non-correspondence of genetic divergence and geographic diversity. The genotypes Vikram, Pee Dee 0113 and KH-121 showed maximum inter-cluster distance in the multivariate analysis and can be exploited for the development of heterotic hybrids after testing their combining ability.

Acknowledgement

First author is highly grateful to Acharya N.G. Ranga Agricultural University for providing financial assistance during the M.Sc (Ag.) programme.

References

- Altaher, A.F. and Singh, R.P. 2003. Genetic diversity studies in upland cotton (*Gossypium hirsutum* L.). *J. Ind. Soc. for Cotton Improvement*, **28**: 158-163.
- Anderberg, M.R. 1993. *Cluster Analysis for Application*. Academic Press, New York.
- Eswara Rao, G., Chenga Reddy, V., Lal Ahamed, M., Srinivasa Rao, V., Panduranga Rao, C. and Bali Reddy, V. 2009. Genetic divergence in upland cotton (*Gossypium hirsutum* L.). *The Andhra Agric. J.*, **56**: 309-314.
- Gopinath, M., Rajamani, S., Naik, R.K. and Rao, C.M. 2009. Genetic divergence for lint characters in upland cotton (*Gossypium hirsutum* L.). *J. Cotton Res. and Dev.*, **23**: 46-48.
- Jackson, J.E. 1991. *A User's Guide to Principal Components*. John Wiley and Sons Inc., New York.
- Rao, C.R. 1952. *Advanced Statistical Methods in Biometrical Research*. John Wiley and Sons Inc., New York. pp. 236-272.
- Srinivasulu, P., Samba Murthy, J.S.V., Rama Kumar, P.V. and Srinivasa Rao, V. 2010. Multivariate analysis of genetic diversity in upland cotton (*Gossypium hirsutum* L.). *The Andhra Agric. J.*, **57**: 148-155.
- Venkateswarulu, K., Chenga Reddy V., Samba Murthy, J.S.V., Srinivasa Rao, V., Panduranga Rao, C., Siva Reddy, K.V. and Sateesh Babu, J. 2010. Multivariate analysis of genetic diversity in upland cotton (*Gossypium hirsutum* L.). *The Andhra Agric. J.*, **57**: 240-247.
- Vijaya Lakshmi, G., Chenga Reddy, V., Panduranga Rao, C., Satish Babu, J. and Srinivasulu, R. 2009. Multivariate analysis of genetic diversity in upland cotton (*Gossypium hirsutum* L.). *The Andhra Agric. J.*, **56**: 8-15.