



Genetic Divergence in Upland Cotton (*Gossypium hirsutum* L.)

G Eswara Rao, V Chenga Reddy, M Lal Ahamed, V Srinivasa Rao
C Panduranga Rao and V Bali Reddy

Department of Genetics and Plant Breeding, Agricultural College, Bapatla 522 101, Andhra Pradesh

ABSTRACT

Sixty genotypes of upland cotton (*Gossypium hirsutum* L.) collected from different geographic regions were subjected to Mahalanobis' D^2 statistic, cluster analysis and principal component analysis. On the basis of clustering methods, fourteen and eight clusters were obtained for Mahalanobis' D^2 statistic and cluster analysis, respectively. In PCA, six principal components were identified. The first six principal components with eigen values more than one contributed 78.21 per cent towards the total variability in cotton. The principal component analysis (PCA) enabled loading of similar type of variables on a common principal component

Key words : Cotton, D^2 analysis, Cluster Analysis, Principal Component Analysis.

A study on the nature and degree of divergence among the varieties of *Gossypium hirsutum* L. is desirable for an understanding of evolution and classifying them into various groups on the basis of their similarity or diversity. Mahalanobis' D^2 analysis, cluster analysis and principal component analysis are potent tools for estimating diversity. The PCA technique has been applied in cotton by Brown (1991). Therefore, an attempt was made to study the genetic divergence in 60 germplasm lines of tetraploid cotton through multivariate analysis.

MATERIAL AND METHODS

The present study was carried out with 60 genotypes of cotton (*Gossypium hirsutum*) obtained from different cotton growing zones of India. The experiment was conducted in randomized complete block design with three replications during *kharif* 2007. The inter- and intra-row spacing adapted was 105 x 60 cm. Each plot consisted of two rows of 6m length and observations were recorded on five competitive plants from each genotype per replication or on plot basis for the characters viz., plant height (cm), days to 50% flowering, number of monopodia per plant, 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 (%), elongation (%), lint yield per plant (g) and seed cotton yield per plant (g). D^2 analysis (Mahalanobis, 1928), principal component analysis (Jackson, 1991) and hierarchical cluster analysis (Anderberg 1993) were carried out using Indostat computer programme.

RESULTS AND DISCUSSION

The analysis of variance showed significant differences among genotypes for all the 16 characters. Simultaneous test of significance based on Wilks' criterion for the pooled effects of all the characters also showed significant differences among varieties ($\chi^2=2138.88$ for 944 degrees of freedom). On the basis of D^2 analysis and cluster analysis, the 60 genotypes were grouped into 14 and 8 clusters, respectively (Table 1 and 2).

The variation in the composition of individual cluster with regard to the number of genotypes indicated the presence of large amount of diversity in the population. On the basis of D^2 and cluster analysis, the scattering of genotypes from the same geographic region to different clusters might be due to the heterogeneity, genetic architecture of the population, past history of selection, developmental traits and degree of general combining ability (Murthy and Arunachalam, 1966). The results clearly indicated that there is no parallelism between the geographic diversity and genetic diversity in cotton in accordance with Altafer and Singh (2003).

On the basis of D^2 analysis the maximum intra-cluster distance was observed in cluster VI (19.09) and minimum intra-cluster distance was observed in cluster III, IV, VII, VIII, IX, X, XI, XII, XIII and XIV (0.00). While on the basis of cluster analysis maximum intra-cluster distance was observed in cluster II (57.29) and minimum intra-cluster distance was observed in cluster VIII (24.05). Inter-cluster distance was least between cluster VII and cluster VIII (8.24) and highest between cluster VIII and cluster X (78.25) on the basis of D^2 analysis while,

Table 1. Clustering pattern of 60 cotton (*Gossypium hirsutum* L.) genotypes by Tocher's method

Cluster number	Number of genotypes	Name of the genotypes
I	10	LH 2108, CPD 81, NH 644, LH 2107, F 2086, F 2164, P 57-6, CPD 1019, RS 2513, IH 65
II	21	L 781, CPD 731, L 389, SUMANGALA, TSH 0252, GISV 197, TCH 1390, RAH 111, COD 1050, RAH 61, L 798, H 1300, H 1316, NDLH 779, GSHV 97/612, BS 279, TSH 2005, ADL 903, KH 7139, GISV 103, SCC 404-1
III	1	RAH 59
IV	1	BS 277
V	9	F 1914, CSH 35, NDL 762, CCH 30-23, SURABHI, RAH 216, CPD 812, CCH LS 3, HAGH 101
VI	10	HS 278, RS 2524, GTHV 4/13, CSH 12, ADB 28, ABH 11, RHC 9854, ARBH 813, ARBH 225, ABH 29
VII	1	HAG 1055
VIII	1	CPD 787
IX	1	GJHV 374
X	1	CNHO 12
XI	1	TCH 1707
XII	1	GJHV 460
XIII	1	GISV 218
XIV	1	CSH 3047

Table 2. Clustering pattern of 60 cotton (*Gossypium hirsutum* L.) genotypes by Ward's minimum variance method

Cluster number	Number of genotypes	Name of the genotypes
I	16	RAH 61, TSH 2005, CPD 1050, ARBH 813, ARBH 225, ABH 29, H 1316, GSHV 97/612 L 798, H 1300, Sumangala, L 389, L 781, CPD 731, TCH 1390, RAH 111
II	4	TCH 1707, GJHV 374, GJHV 460, GISV 218
III	2	HAG 1055, CPD 787
IV	9	TSH 0252, KH 7139, CCH 03-23, Surabhi, NDLH 779, BS 277, BS 279, GISV 197, CSH 12
V	5	HS 278, RS 2524, GTHV 4/13, ADB 28, CSH 3047
VI	13	P 57-6, F 2164, LH 2107, F 2086, NH 644, LH 2108, CPD 814, RAH 216, IH 65, RS 2513, RHC 9854, ABH 11, CNHO 12
VII	6	CPD 1019, CPD 812, RAH 59, SCC 404-1, HAGH 101, CCH LS 3
VIII	5	ADL 903, GISV 103, F 1914, CSH 35, NDL 762

Table 3: Average intra-and inter-cluster D² values among fourteen clusters in 60 cotton (*Gossypium hirsutum* L.) genotypes.

Cluster	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV
I	7.91	20.98	14.57	24.02	16.07	19.06	46.03	60.36	48.20	18.51	27.64	42.18	49.78	28.20
II		11.19	16.68	14.75	19.14	17.75	18.90	26.80	20.21	27.53	16.37	18.33	24.16	32.18
III			0.00	16.58	14.77	20.32	31.11	36.42	36.60	28.72	19.55	38.39	35.89	30.79
IV				0.00	21.11	20.26	16.20	18.85	28.06	48.41	19.97	20.14	41.01	620.39
V					14.83	25.01	36.79	48.42	36.83	30.22	23.88	38.91	44.95	35.93
VI						19.09	31.58	42.46	35.65	24.98	25.62	25.76	36.06	31.01
VII							0.00	8.24	18.74	62.56	19.60	17.17	30.79	45.52
VIII								0.00	22.27	78.25	26.93	22.15	35.07	44.39
IX									0.00	54.22	14.52	14.61	14.57	43.12
X										0.00	45.90	49.34	40.48	55.61
XI											0.00	23.83	25.36	28.15
XII												0.00	21.70	34.03
XIII													0.00	67.40
XIV														0.00

Table 4. Average intra-and inter-cluster Euclidian² values among eight clusters in 60 cotton (*Gossypium hirsutum* L.) genotypes.

Cluster	I	II	III	IV	V	VI	VII	VIII
I	34.06	61.12	72.00	46.54	64.74	65.89	70.81	46.38
II		57.29	72.27	70.18	93.72	126.70	116.26	86.41
III			24.73	68.82	105.41	165.05	145.17	98.92
IV				34.16	61.90	74.74	57.63	46.21
V					48.35	66.39	81.30	69.45
VI						31.94	49.56	49.09
VII							39.46	47.02
VIII								24.05

Table 5. Contribution of different characters towards genetic divergence in 60 cotton (*Gossypium hirsutum* L.) genotypes .

Characters	%Contribution towards divergence	Times ranked first
Plant height (cm)	7.34	130
Days to 50% flowering	0.68	12
Number of monopodia per plant	39.72	703
Number of sympodia per plant	2.09	37
Number of bolls per plant	2.99	53
Boll weight(g)	2.26	40
Ginning out-turn(%)	4.75	84
Seed index(g)	2.37	42
Lint index(g)	3.16	56
2.5% span length (mm)	13.62	241
Micronaire(10 ⁻⁶ g/in)	4.97	88
Bundle strength (g/tex)	3.62	64
Uniformity ratio(%)	2.26	40
Fibre elongation(%)	1.24	22
Lint yield per plant (g)	8.36	148
Seed cotton yield per plant(g)	0.57	10

Table 6. Eigene values, proportion of the total variance represented by first six 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
Eigene value (Root)	208.138	78.684	76.313	45.874	39.98	33.753
% Var.Exp	33.785	12.772	12.387	7.446	6.346	5.479
Cumulative Variance Explained	33.785	46.557	58.944	66.390	72.737	78.215
Plant height (cm)	0.009	0.280	0.262	0.709	0.010	0.302
Days to 50% flowering	0.005	0.015	0.137	0.067	0.259	-0.23
Number of monopodia/plant	0.867	-0.116	0.094	0.145	0.112	-0.318
Number of sympoida /plant	-0.021	0.195	-0.018	-0.036	-0.345	-0.065
Number of bolls / plant	-0.044	0.313	0.056	-0.155	0.409	0.016
Boll weight(g)	-0.101	0.096	-0.086	0.134	-0.203	-0.128
Ginning out-turn(%)	0.012	0.239	0.057	-0.115	-0.627	-0.266
Seed index(g)	0.216	0.210	0.233	-0.085	-0.068	0.259
Lint index(g)	0.087	-0.145	-0.057	-0.468	0.012	0.221
2.5% span length (mm)	-0.168	0.019	0.766	-0.163	0.009	-0.259
Micronaire(10 ⁻⁶ g/in)	0.144	0.388	-0.189	-0.087	-0.034	0.490
Bundle strength (g/tex)	0.329	0.189	0.003	-0.183	-0.232	0.123
Uniformity ratio(%)	0.074	0.077	-0.333	0.085	0.071	-0.261
Fibre elongation(%)	0.128	0.053	0.284	-0.274	0.057	0.234
Lint yield per plant (g)	-0.055	0.663	-0.099	-0.107	0.240	-0.383
Seed cotton yield /plant(g)	-0.013	0.075	-0.088	-0.165	0.273	-0.086

PC=Principal component

it was least between cluster IV and cluster VIII (46.21) and maximum between cluster III and cluster VI (165.05) on the basis of cluster analysis (Table 3 and 4).

Based on D² analysis genotypes were grouped into fourteen clusters. Cluster II had the maximum number (21) of genotypes followed by cluster I and VI (10), cluster V (9) and other clusters had only one cluster. While based on cluster analysis, 60 genotypes are grouped into eight clusters. Cluster I comprised the maximum number (16) of genotypes followed by cluster VI (13), cluster IV (9), cluster VII (6), cluster V and VIII (5), cluster II (4) and cluster III (2). The findings are in agreement with Muraleedhar and Murthy (2005) and Vijayalaxmi (2008).

The contribution of individual characters to the divergence is presented (Table 5). The characters viz., number of monopodia per plant (39.72%) followed by 2.5% span length (13.62%), lint yield per plant (8.36%), plant height (7.34%), micronaire (4.97%), ginning out turn (4.75%), bundle strength (3.62%) and and number of bolls per plant (2.99%) contributed maximum to the total divergence in the material. The results are in accordance with

Karunakar Raju *et al.* (2005) and Vijayalaxmi (2008).

In the principal component method, first 6 principal components altogether explained 78.2 % of the variability (Table 6). The first component explained 33.7% of total variability and remaining ones accounted for progressively lesser and lesser amount of variation.

The PCA scores of 60 cotton genotypes (Table 7) were plotted on graph to get the 2D scatter diagram (Fig.1). All the genotypes were plotted for PC₁, PC₂ which clearly indicated the grouping of genotypes of same clusters falling nearer to each other for example the genotypes GISV 218 (cluster XIII) and GJHV 460 (cluster XII) fall nearer to each other on the positive axis of PC₁. While genotypes P 57-6 (Cluster I) and GJHV 374 (Cluster IX) fall away from each other indicating their specificity in clustering. The PC₁ reflected high component loading of number of monopodia per plant and bundle strength while PC₂ reflected significant loading of lint yield per plant. PC₃ reflected positive significant loadings of plant height and number of monopodia per plant. The results of hierarchical cluster analysis and PCA confirmed the findings of each other. Genotypes belonging to a common cluster have

Fig 1. Two dimensional graph showing relative position of 60 cotton (*Gossypium hirsutum* L.) genotypes based on PCA scores.

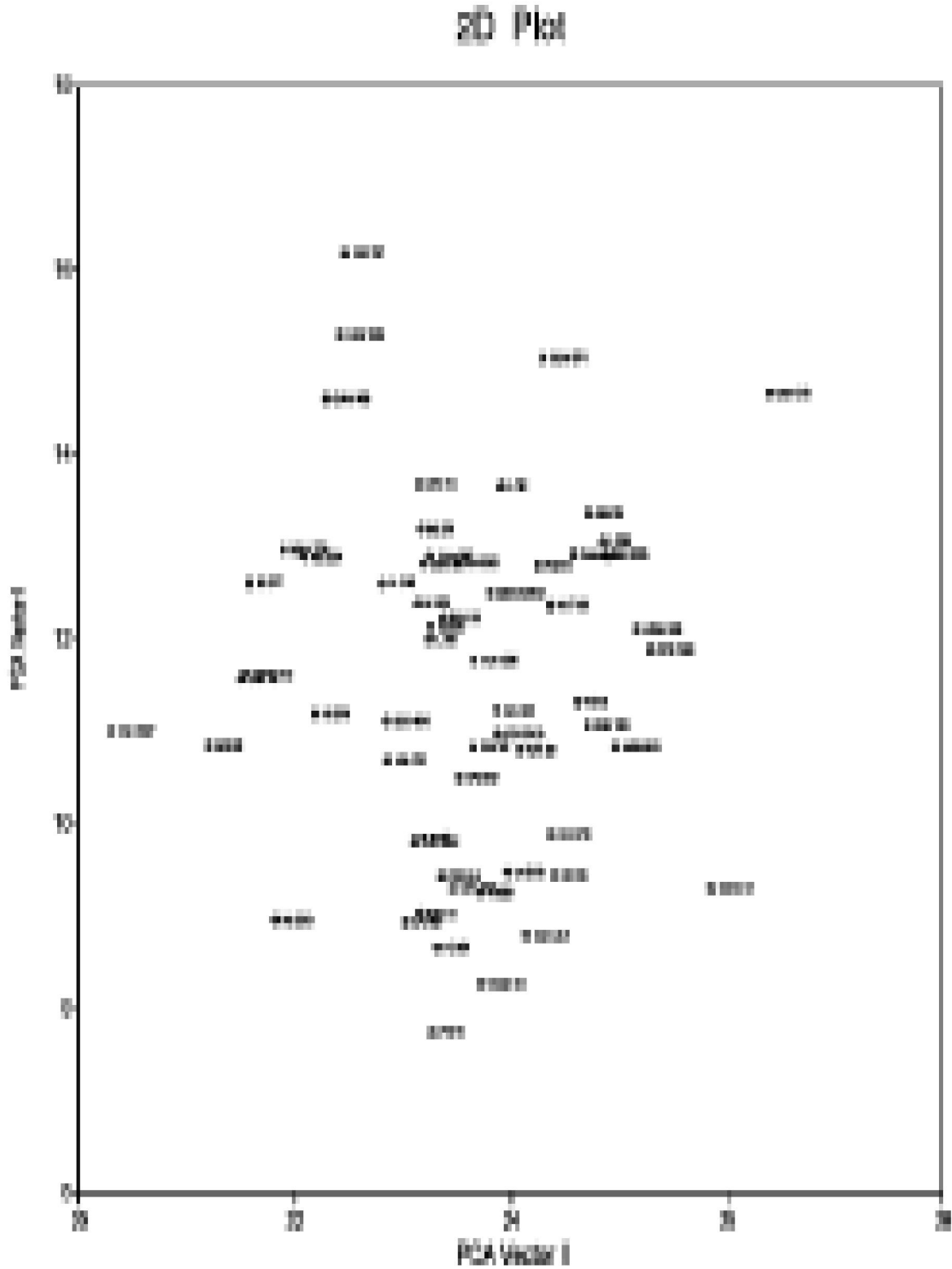


Table 7. PCA scores of 60 genotypes of cotton (*Gossypium hirsutum* L.).

S.No	GENOTYPE	PCAI X Vector	PCAI Y Vector	PCA III Z Vector
1	RAH 61	12.789	24.247	26.842
2	CSH 3047	10.998	20.299	27.221
3	TSH 0252	12.829	23.493	28.304
4	P 57-6	7.729	22.274	27.247
5	ABH 11	9.016	23.156	25.279
6	CPD 1050	11.876	25.288	26.706
7	CPD 1019	9.290	23.463	28.601
8	LH 2108	9.468	23.967	26.805
9	IH 65	9.823	23.190	26.395
10	TCH 1707	12.884	22.251	28.483
11	F 2164	8.646	23.315	26.812
12	HAGH 101	8.257	23.726	29.632
13	DSH 12	11.567	21.507	27.586
14	H 1316	12.591	22.807	26.606
15	ARBH 225	12.102	25.163	26.797
16	GISV 218	14.664	26.383	26.086
17	NDLH 779	12.963	21.893	27.456
18	ABH 29	13.358	24.716	25.836
19	RAH 59	10.794	24.082	28.264
20	L 798	11.981	23.243	26.315
21	HS 278	11.179	22.186	25.923
22	RS 2524	12.887	22.085	25.156
23	ADB 28	10.840	21.200	26.509
24	GJHV 460	14.597	22.287	25.960
25	LH 2107	8.926	23.024	26.943
26	NH 644	9.433	24.403	26.840
27	RHC 9854	9.797	23.105	25.425
28	KH 7139	12.355	24.365	28.786
29	SCC 404-1	11.104	22.849	27.831
30	RS 2513	8.952	21.819	26.511
31	F2086	9.245	23.720	26.798
32	GTHV 4/13	11.608	21.548	25.969
33	CCH LS 3	8.777	24.144	29.200
34	TSH 2005	12.907	24.870	26.283
35	CCH 03-23	10.958	23.877	29.607
36	BS 277	12.594	21.578	28.178
37	SURABHI	12.120	23.258	29.650
38	BS 279	13.189	23.165	27.793
39	GJHV 374	15.032	24.295	27.953
40	ARBH 813	10.814	24.969	26.408
41	CPD 812	10.488	23.521	28.430
42	Sumangala	11.912	24.582	27.967
43	H 1300	12.381	23.135	25.875
44	CNHO 12	9.290	25.854	24.696
45	RAH 216	9.872	24.975	27.397
46	CPD 814	9.417	23.351	27.112
47	ADL 903	11.221	23.866	27.383
48	HAG 1055	15.290	22.420	27.857
49	F 1914	11.307	24.616	28.199
50	CPD 787	16.171	22.461	28.028
51	GISV 103	11.058	24.705	27.470
52	GSHV 97/612	12.490	23.801	25.589
53	NDL 762	10.674	22.858	27.810
54	TCH 1390	11.763	23.658	26.875
55	CSH 35	10.817	23.650	28.982
56	RAH 111	12.215	23.366	27.390
57	L 781	13.648	23.913	26.660
58	L 389	13.047	24.846	27.817
59	CPD 731	13.666	23.140	27.331
60	ISV 197	12.087	23.188	27.516

fallen nearer to each other and vice-versa and this confirms the results of cluster analysis.

All the three methods of grouping revealed a single concept of non-correspondence of genetic and geographical diversities. All the three methods of classifying genotypes into different groups are equally useful but hierarchical cluster analysis gave an additional advantage of identifying sub-cluster of the major groups at different levels so that each small group can be critically analysed.

LITERATURE CITED

- Altaher A F and Singh R P 2003.** Genetic diversity studies in upland cotton (*Gossypium hirsutum* L.). Journal of the Indian Society for Cotton Improvement 28(3): 158-163.
- Anderberg M R 1993.** Cluster Analysis for Application. Academic Press, New York.
- Brown J S 1991.** Principal component and cluster analysis of cotton cultivars variability across the U.S cotton belt. Crop Science 31: 915-922.
- Jackson J E 1991.** A User's Guide to Principal Component. John Wiley and sons, New York.
- Karunakar Raju G B , Chenga Reddy V and Panduranga Rao C 2005.** Multivariate analysis of genetic diversity in upland cotton. The Andhra Agricultural Journal, 52(1&2):66-74.
- Mahalanobis P C 1928.** A statistical study at Chinese head measurement. Journal of Asiatic Society of Bengal 25: 301-307.
- Muraleedhar S. Aski and Murthy J V S 2005.** Mahalanobis' D² and principal component analysis of genetic diversity in tetraploid cotton. The Andhra Agricultural Journal, 52 (1&2): 99-104.
- Murthy B R and Arunachalam V 1966.** The nature of divergence in relation to breeding systems in some crop plants. Indian Journal of Genetics and Plant Breeding 26:188-198.
- Vijayalaxmi G 2008.** Genetic divergence in cotton (*Gossypium hirsutum* L.). M.Sc. (Ag) thesis, Acharya N.G. Ranga Agricultural University, Rajendranagar, Hyderabad.