



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

K Venkateswarlu, V Chenga Reddy, J S V Samba Murthy, V Srinivasa Rao, C Panduranga Rao,
K V Siva Reddy and J Sateesh Babu

Department of Genetics and Plant Breeding, Agricultural College, Bapatla 522101, Andhra Pradesh

ABSTRACT

Fifty genotypes of upland cotton (*Gossypium hirsutum* L.) collected from different research centers across the country were subjected to Mahalanobis' D^2 statistic, cluster analysis and principal component analysis based on sixteen characters. Eight and 8 clusters were obtained for D^2 statistic and cluster analysis, respectively. Divergence studies indicated that ginning out-turn, 2.5% span length, oil content, number of monopodia plant⁻¹ and seed index contributed maximum to genetic diversity. Multivariate analysis revealed that wider genetic diversity existed among the genotypes GSHV-155, GJHV-448 and LK-861, revealing the scope for exploitation of heterosis.

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

Traditionally Mahalanobis' (1936) D^2 statistic to measure genetic divergence as suggested by Rao (1952) has been used by different workers in cotton (Singh and Gill, 1984 and Ekbote and Khoragade, 1986). The present study was carried out with different methods of clustering based on D^2 analysis, hierarchical cluster analysis and principal component analysis in 50 genotypes of upland cotton.

MATERIAL AND METHODS

The present study was carried out with 50 genotypes of cotton (*Gossypium hirsutum* L.) obtained from different research centres across the country (Table 1). The experiment was conducted in randomized block design with three replications at Regional Agricultural Research Station, Lam, Guntur during *kharif* 2008-09. The inter- and intra-row spacing adapted was 120cm x 60cm. Each plot consisted of one row of 6m length and observations were recorded on five randomly selected plants from each genotype per replication for characters viz., plant height (cm), number of monopodia plant⁻¹, number of sympodia plant⁻¹, number of bolls plant⁻¹, boll weight (g), seed index (g), lint index (g), lint yield plant⁻¹ and seed cotton yield plant⁻¹. Days to 50% flowering, ginning out-turn (%), 2.5% span length (mm), micronaire (10^{-6} g/in), bundle strength (g/tex), uniformity ratio and oil content (%) were recorded on plot basis. The fibre quality characters were analyzed at CIRCOT Regional Unit at Regional Agricultural Research Station, Lam, Guntur. The data were statistically analyzed to study diversity by Mahalanobis' D^2 statistic as per Rao (1952),

principal component analysis (PCA) as described by Jackson (1991) and cluster analysis as described by Anderberg (1993).

RESULTS AND DISCUSSION

On the basis of D^2 values and cluster analysis the fifty genotypes were grouped into eight clusters each (Table 2). Cluster III was the biggest cluster with 17 genotypes, followed by clusters I with 13 genotypes, cluster II with 12 genotypes and cluster V with four genotypes. The remaining clusters *i.e.*, cluster IV, cluster VI, cluster VII and cluster VIII consisted of single genotype, each.

Based on Ward's minimum variance dendrogram, the clustering pattern revealed that cluster III was the largest and consisted of 13 genotypes, followed by cluster VIII with 12 genotypes, cluster IV with 8 genotypes, cluster V with 6 genotypes and cluster VI with 4 of genotypes. Whereas cluster II, cluster I comprised 3 genotypes, each. Cluster VII was the smallest cluster with single genotype as shown in Table 2 and dendrogram (Fig. 1).

Based on D^2 values intra-cluster D^2 values ranged from zero (cluster IV, cluster VI cluster VII and cluster VIII) to 39.678 (cluster V). Whereas the maximum inter-cluster distance was observed between clusters VII and VIII, followed by clusters VI and VIII. The least inter-cluster distance was observed between clusters I and IV (Table 3).

D^2 statistic revealed that cluster II recorded the highest mean values for lint yield plant⁻¹ (65.61), seed cotton yield plant⁻¹ (187.13), number of sympodia plant⁻¹ (19.11), number of bolls plant⁻¹

Table 1. Name of the cotton (*Gossypium hirsutum* L.) genotypes employed in the present investigation and their source.

S.No.	Name of the genotype	Source
1	CSH 612	CICR, Sirsa
2	TSH 9908	TANU, Srivilliputtore.
3	RS252	RAU, Sriganaganagar.
4	TCH 1716	TANU, Coimbatore.
5	CPD 801	UAS, Dharwad.
6	GJHV 448	JAU, Junagadh.
7	KH 140	JNKV, Khandwa.
8	F 2168	PAU, Faridkot.
9	CNH 1104	CICR, Nagpur.
10	NDLH 1938	RARS, Nandyal.
11	LH 2123	PAU, Ludhiana.
12	L 604	RARS, Lam.
13	BS 51	AICCP.
14	CCH 2623	CICR, Coimbatore.
15	CSH 3129	CICR, Sirsa.
16	CCH 4474	CICR, Coimbatore.
17	F 2170	PAU, Faridkot.
18	GISV 216	NAU, Surat.
19	HS 283	HAU, Hissar.
20	RAH 336	UAS, Raichur.
21	RS 2527	RAU, Sriganaganagar.
22	BS 41	AICCP.
23	RS 2013	RAU, Sriganaganagar
24	LH 2132	PAU, Ludhiana.
25	P 1752	IARI, New Delhi.
26	IHANI 251	AICCP.
27	SCS 415	UAS, Siriguppa.
28	L 801	RARS, Lam.
29	HAG 1015	UAS, Hagari.
30	F 1861	PAU, Faridkot.
31	GSHV 155	NAU, Surat.
32	H 1360	HAU, Hissar.
33	CA 7	CICR, Nagpur.
34	TCH 1715	TNAU, Coimbatore.
35	CANDTS 55	AICCP
36	NH 630	MAU, Nanded.
37	RAH 61	UAS, Raichur.
38	BS 277	AICCP.
39	CPD 1050	UAS, Dharwad.
40	GISV 218	NAU, Surat.
41	RHC 9854	AICCP.
42	BS 279	AICCP.
43	MCU 5	TNAU, Coimbatore.
44	ARBH 225	UAS, Arbhavi.
45	CPD 1019	UAS, Dharwad
46	ARBH 813	UAS, Arbhavi.
47	P 57-6	AICCP.
48	RAH 216	UAS, Raichur.
49	SURABHI	CICR, Coimbatore.
50	LK 861	RARS, Lam.

Table 2. Genotypes included in each cluster based on Mahalanobis' D² analysis and Ward's minimum variance method

Based on (Mahalanobis' D ² analysis) D ² value			Based on cluster analysis (Ward's minimum variance method)	
Cluster No.	No. of genotypes	Name of the genotype	No. of genotypes	Name of the genotype
I	13	F-2170, LH-2132, CSH-3129, HAG-1015, LH-2123, SCS-415, F-2168, RS-2525, HIHANI-251, L-604, BS-51, NH-630, CPD-801.	3	TSH-9908, TCH-1716, TCH-1715.
II	12	RHC-9854, ARBH-813, RAH-61, SURABHI, RAH-216, CPD-1050, CPD-1019, ARBH-225, BS-279, GISV-218, F-1861, P-57-6.	3	CCH-2623, RAH-336, GJHV-448.
III	17	HS-283, P-1752, GSHV-155, CNH-1104, CANDTS-55, CCH-4474, CCH-2623, RS-2527, ND LH-1938, RAH-336, CA-7, KH-140, CSH-612, TCH-1715, L-801, BS-41, GISV-216.	13	KH-140, CA-7, GISV-216, F-2168, F-1861, LH-2123, L-604, BIHANI-251, CSH-612, BS-51, HAG-1015, NH-630, L-801.
IV	1	RS-2013	8	RS-2525, SCS-415, F-2170, LH-2132, CSH-3129, CPD-801, H-1360, BS-277.
V	4	TSH-9908, TCH-1716, MCU-5, BS-277.	6	CNH-1104, CANDTS-55, HS-283, P-1752, RS-2527, GSHV155.
VI	1	H-1360.	4	ND LH-1938, CCH-4474, BS-41, RS-2013.
VII	1	GJHV-448.	1	LK-861.
VIII	1	LK-861.	12	RAH-61, SURABHI, RHC-9854, ARBH-813, ARBH-225, CPD-1050, CPD-1019, GISV-218, RAH-216, BS-279, P-57-6, MCU-5.

(49.42) and boll weight (3.94). The genotypes falling under this cluster may be tried for direct adoption and hybridization for improvement of yield. Cluster IV had recorded the highest mean values for plant height (131.93), ginning-outturn (33.35) and oil content (21.23). Cluster VI recorded the highest mean value for microinaire (4.70) and uniformity ratio (52.56). Cluster VIII recorded the highest mean values days to 50% flowering (61.00), bundle strength (28.20) and 2.5% span length (36.41).

These genotypes falling under respective clusters may be tried for improving the fibre characters. Cluster VII recorded the highest mean values for monopodia plant¹ (2.56), ginning out-turn (38.14) and seed index (7.56) (Table 5).

Based on the 6 principal components, a cumulative of (77.41%) variation formed the basis for the divergence into different clusters. Main principal components are given in Table 7.

Table 3. Average intra- and inter-cluster D^2 values among eight clusters based on Mahalanobis' D^2 values

Cluster No.	II	III	IV	V	VI	VII	VIII
I	28.13	29.34	27.66	40.01	33.47	61.00	133.50
II		42.79	38.93	38.91	55.02	62.92	151.93
III			43.54	51.90	55.64	55.99	153.15
IV				38.39	64.01	85.65	105.36
V					66.59	64.92	96.47
VI						108.44	175.43
VII							180.30

Table 4. Average intra- and inter-cluster Euclidian² values among eight clusters based on cluster analysis

Cluster No.	II	III	IV	V	VI	VII	VIII
I	139.86	101.99	156.79	135.00	146.51	248.16	136.78
II		120.09	174.46	122.18	166.90	556.96	148.22
III			86.97	86.93	126.91	405.80	96.74
IV				92.85	117.58	428.33	113.32
V					109.11	454.16	142.29
VI						438.34	135.75
VII							452.66

The grouping based on the cluster analysis resulted in eight clusters. Agglomerative cluster analysis revealed that wide genetic distance exists between cluster II (CCH-2623, RAH-336 and GJHV-448) and VII (LK-861) followed by cluster V (CNH-1104, CANDTS-55, HS-283, P-1752, RS-2527 and GSHV-155) and VII (LK-861). Therefore, selection of parents (CCH-2623, GJHV-448 and LK-861) from these clusters will produce superior segregants through hybridization for improvement of seed cotton, lint yield and oil content (Table 4). The distribution of genotypes based on PCA values are shown in 3D plot where the genotypes (GSHV-155 and HS-283) and (MCU-5 and CPD-1050) were away from other genotypes on the basis of PCA I and PCA II scores. D^2 statistic showed that the characters contributing towards divergence in Table 7 revealed that the maximum contribution towards divergence is by ginning out-turn (%) (15.92) followed by 2.5% span length (14.86), oil content (%) (14.12),

number of monopodia plant⁻¹ (13.71) and seed index (10.12). Thus, these five characters altogether account for more than 65% of the total divergence in 50 genotypes studied (Table 6).

In general, multivariate analysis revealed that divergence existed among GSHV-155, GJHV-448 and LK-861 indicating the scope for exploitation of heterosis with high seed cotton yield and fibre quality. For obtaining hybrids with high level of heterosis the question generally arises regarding the ideal distance (degree of divergence) at phenotypic level. Arunachalam and Bhandopadhyay (1984) have proved experimentally more number of heterotic combinations with higher level of heterosis was from the parents grouped into moderate divergent groups. Hence, selection of varieties should depend more on genetic diversity than the geographical diversity. Similar findings were also reported by Muraleedhar *et al.*, (2005) and Karunakar Raju *et al.*, (2005).

Table 5. Mean values of clusters estimated from 50 genotypes of cotton (*Gossypium hirsutum* L.) based on D² analysis and cluster analysis.

Cl/No	Plant height (cm)	Days to 50% flowering	No. of monopodias/plant ⁻¹	No. of sympodiabolls/plant ⁻¹	No. of bolls/plant ⁻¹	Boll weight (g)	Ginning out-turn (%)	Seed index (g)	Lint index (g)	2.5% span length (mm)	Micronaire (10g/inch)	Bundle strength (g/Tex)	Uniformity ratio	Oil content (%)	Seed lint yield/plant ⁻¹ (g)	
I	123.39 (125.00)	57.38 (58.33)	0.97 (2.11)	14.16 (13.68)	39.11 (34.36)	3.93 (3.60)	35.72 (36.12)	9.10 (9.03)	5.08 (5.10)	27.19 (30.33)	4.09 (3.64)	22.37 (23.10)	50.44 (49.18)	20.65 (19.94)	52.34 (41.61)	146.47 (151.13)
II	129.54 (128.20)	58.00 (57.33)	1.32 (1.87)	19.11 (14.68)	49.42 (33.38)	3.94 (4.04)	35.03 (37.96)	9.01 (9.61)	4.83 (6.04)	26.54 (25.85)	4.23 (4.64)	22.99 (21.82)	48.42 (51.07)	20.69 (21.03)	65.61 (45.93)	187.13 (121.25)
III	121.85 (120.60)	56.72 (56.43)	1.29 (1.17)	14.37 (14.39)	37.22 (40.82)	3.77 (3.69)	36.29 (36.10)	10.22 (8.64)	5.93 (4.88)	27.32 (27.06)	4.27 (4.16)	22.37 (22.14)	50.49 (50.02)	20.60 (20.30)	47.45 (51.48)	130.85 (142.60)
IV	131.93 (12170)	57.00 (57.70)	1.40 (0.75)	13.93 (14.53)	40.83 (37.67)	3.33 (4.09)	33.35 (35.01)	9.06 (9.17)	4.53 (4.99)	31.37 (26.46)	3.69 (4.20)	23.53 (22.12)	46.66 (50.89)	21.23 (20.58)	45.60 (51.13)	136.66 (146.14)
V	122.43 (119.30)	58.58 (57.94)	1.83 (1.11)	16.45 (13.89)	41.74 (35.40)	3.46 (3.59)	34.55 (36.00)	8.53 (11.00)	4.49 (6.47)	29.11 (26.20)	3.82 (4.27)	22.83 (21.88)	48.11 (50.88)	19.82 (20.33)	48.01 (43.17)	140.06 (120.06)
VI	109.86 (123.20)	57.00 (56.58)	0.33 (1.25)	13.33 (15.06)	39.00 (40.74)	3.91 (3.78)	33.66 (34.69)	8.33 (10.99)	4.50 (5.84)	24.47 (30.49)	4.70 (4.05)	20.00 (23.78)	52.56 (48.91)	19.67 (21.49)	49.50 (51.63)	147.08 (148.75)
VII	107.06 (113.60)	58.33 (61.00)	2.56 (1.66)	13.66 (16.53)	29.23 (19.80)	3.65 (3.33)	38.14 (35.20)	7.56 (8.03)	4.73 (4.36)	24.53 (35.41)	4.62 (2.52)	21.40 (28.20)	51.20 (42.83)	21.09 (18.35)	34.02 (20.83)	89.58 (59.16)
VIII	113.60 (129.80)	61.00 (58.05)	1.66 (1.43)	16.53 (19.37)	19.80 (49.96)	3.33 (3.91)	35.20 (34.86)	8.03 (8.98)	4.36 (4.80)	35.41 (26.84)	2.52 (4.22)	28.20 (23.10)	42.83 (48.24)	18.35 (20.73)	20.83 (65.76)	59.16 (188.36)

Figures given in parentheses are based on cluster analysis

Table 6. Contribution of different characters towards genetic divergence in 50 cotton (*Gossypium hirsutum* L.) genotypes using D² analysis.

Character	Contribution towards divergence (%)	Times ranked first
Plant Height (cm)	5.80	71
Days to 50% flowering	2.94	36
Monopodia Plant ⁻¹	13.71	168
Sympodia Plant ⁻¹	6.86	84
Bolls Plant ⁻¹	3.59	44
Boll Weight (gm)	1.96	24
Ginning Out Turn(%)	15.92	195
Seed Index (g)	10.12	124
Lint Index (g)	3.27	40
2.5% span Length(mm)	14.86	182
Uniformity Ratio	0.00	0
Micronaire (10-6g/Inch)	3.10	38
Bundle Strength (g/Tex)	1.06	13
Oil Content (%)	14.12	173
Lint Yield Plant ⁻¹ gm	0.08	1
Seed Cotton Yield Plant ⁻¹	2.61	32

Table 7. Eigen 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)	3.533	3.275	1.986	1.506	1.168	0.917
% Var. Exp.	22.081	20.470	12.415	9.412	7.300	5.733
Cum. Var. Exp.	22.081	42.551	54.965	64.377	71.677	77.410
Plant Height (cm)	0.243	0.165	0.082	0.359	0.198	0.001
Days to 50% flowering	0.100	-0.257	0.141	0.252	0.086	-0.685
No. of monopodia Plant ⁻¹	0.192	-0.135	-0.404	0.184	-0.458	0.030
No. of sympodia Plant ⁻¹	0.420	-0.094	0.013	0.175	0.086	-0.158
No. of bolls Plant ⁻¹	0.352	0.293	0.065	-0.112	-0.199	-0.173
Boll Weight (g)	0.117	0.286	0.310	0.151	0.338	0.300
Ginning out-turn (%)	-0.103	-0.080	-0.530	0.315	0.054	0.237
Seed Index (g)	-0.304	0.165	0.110	0.328	-0.120	-0.259
Lint Index (g)	-0.453	0.001	-0.119	0.237	-0.014	-0.258
2.5% Span length (mm)	-0.011	-0.403	0.119	0.143	-0.289	0.250
Micronaire (10g/Inch)	-0.010	0.438	-0.261	0.165	0.074	-0.175
Bundle Strength (g/Tex)	0.373	-0.225	0.075	0.304	0.064	0.103
Uniformity Ratio	-0.064	0.412	-0.020	-0.071	-0.225	0.008
Oil content (%)	-0.031	0.242	0.165	0.522	-0.235	0.271
Lint Yield Plant ⁻¹ (g)	-0.354	-0.214	0.291	0.166	0.232	0.130
Seed Cotton Yield Plant ⁻¹ (g)	-0.041	0.025	0.450	-0.013	-0.558	-0.030

Fig 1. Dendrogram showing relationship of 50 cotton (*Gossypium hirsutum* L.) genotypes in eight clusters using Ward's minimum variance method.

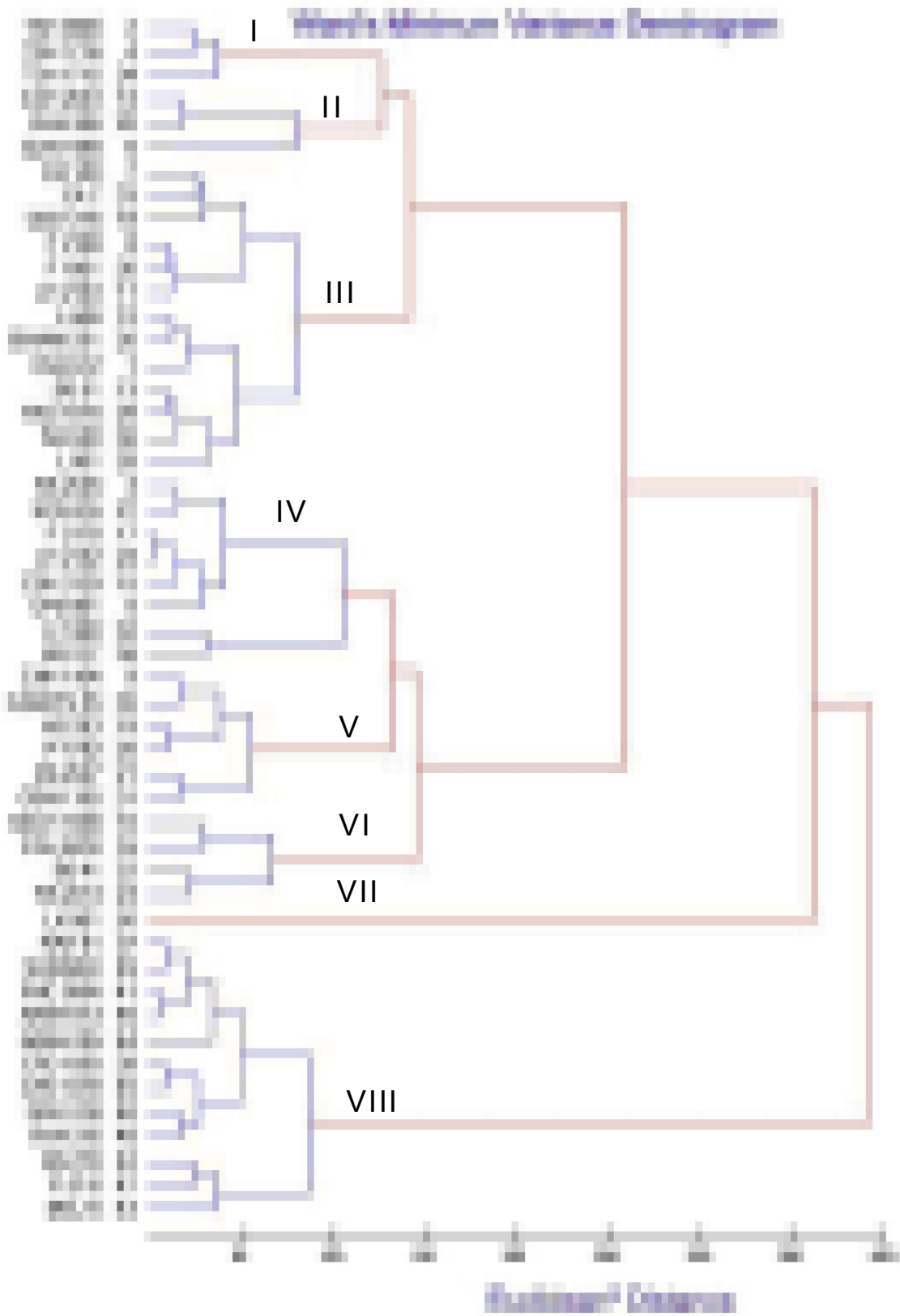
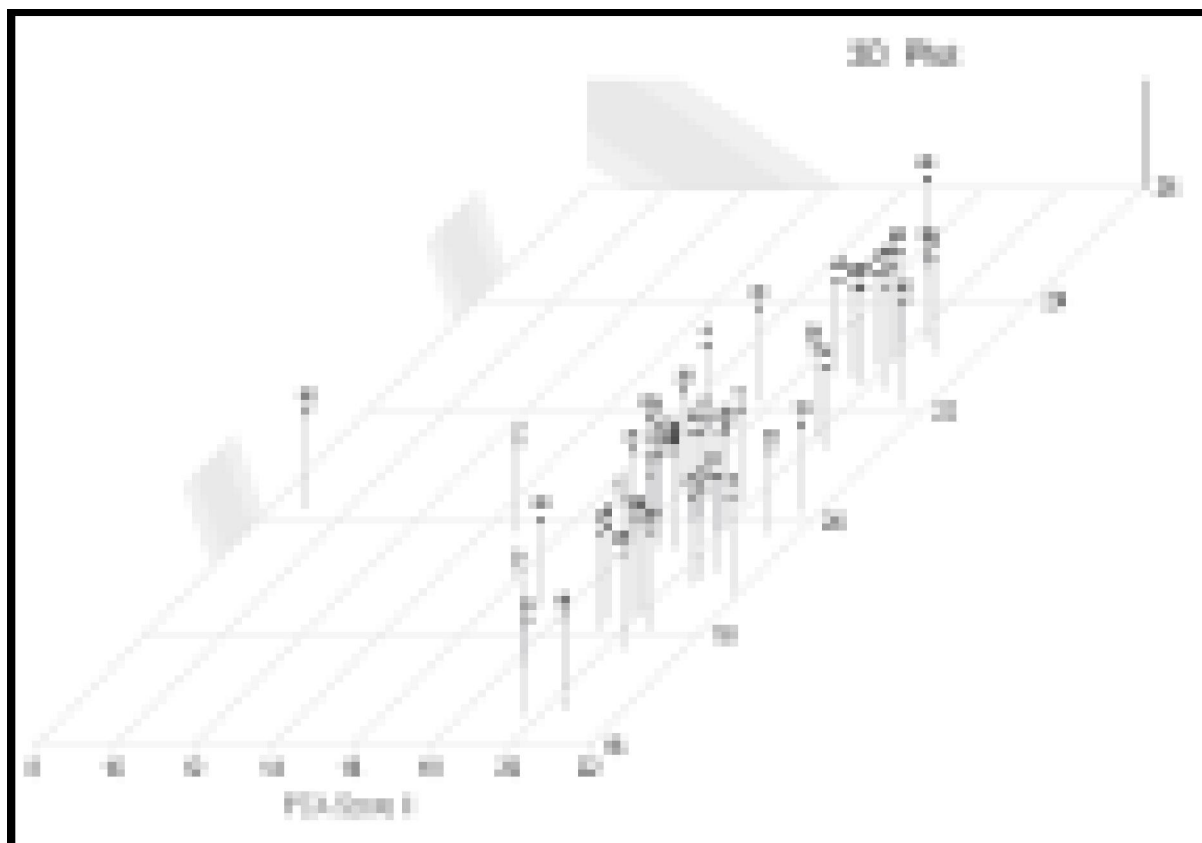


Fig 2: Three dimensional graph showing relative position of 50 cotton (*Gossypium hirsutum* L.) genotypes based of PCA scores (Genotype numbers as per Table 1)



LITERATURE CITED

- Anderberg M R 1993.** Cluster Analysis for Application. Academic Press, New York.
- Arunachalam V and Bhandopadhyay V A 1984.** Limits to genetic divergence for occurrence of heterosis- experimental evidence from the crop plants. Indian Journal of Genetics and Plant Breeding 44(3):548-554.
- Ekbote AP and Khorgade PW 1986.** Genetic divergence in upland cotton (*Gossypium hirsutum* L.). Punjabrao Krishi Vidyapeeth Research Journal 8(2):1-4.
- Jackson J E 1991.** A User's Guide to Principal Components. John Wiley and Sons Inc., New York.
- Karunakar Raju G B, Chenga Reddy V and Panduranga Rao C 2005.** Multivariate analysis of genetic diversity in upland cotton(*Gossypium hirsutum* L.). The Andhra Agricultural Journal 52(1&2): 66-74.
- Mahalanobis PC 1936.** On the generalized distance in statistics. Proceedings of National Academy of Sciences (India) 2:49-55.
- Muraleedhar S A, Murthy J S V S, Rama Kumar P V and Srinivasa Rao V 2005.** Mahalanobis' D^2 and principal component analysis of genetic diversity in tetraploid cotton(*Gossypium hirsutum* L.). The Andhra Agricultural Journal 52 (1&2): 99-104.
- Rao CR 1952.** Advanced Statistical Methods in Biometric Research. John Willey and Sons Inc., New York: 390.
- Singh TH and Gill SS 1984.** Genetic diversity in upland cotton under different environments. Indian Journal of Genetics and Plant Breeding 44(3):506-513.