



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

A Padmavathi, M Lal Ahamed and P Anil Kumar

Department of Genetic and Plant Breeding, Agricultural College, Bapatla 522 101

ABSTRACT

An experiment was conducted to analyze the genetic diversity among 60 genotypes for 16 quantitative characters in upland cotton. The 60 genotypes were grouped into 15 clusters based on Mahalanobis D^2 analysis with cluster I being largest with 15 genotypes followed by 13 genotypes in cluster II, 11 genotypes in cluster XII and 10 genotypes in cluster VII. Based on Hierarchical cluster analysis all the genotypes were grouped into 8 clusters with cluster I being largest with 23 genotypes followed by 9 genotypes in cluster II and 8 genotypes in cluster IV. This random distribution of genotypes indicated that the genetic diversity and geographical diversity are not related. In Principal component analysis first six principal components with eigen value more than one contributed 87.315 per cent towards the total variability with PC_1 alone showing maximum of 22.27 per cent variability.

Key words: Cotton, Hierarchical cluster analysis, Mahalanobis D^2 analysis, Principal component analysis.

Cotton is the main economic and industrial crop of our country serving raw material for textile mills. Competition from synthetic fibres and challenges to improve fibre quality are the two major economic forces driving the current global market. 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 needed 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 60 cotton genotypes using multivariate Mahalanobis D^2 statistics, Hierarchical cluster analysis and Principal component analysis.

MATERIAL AND METHODS

The experiment was conducted during *kharif* 2007-2008 in Randomized Block Design with 60 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 90:45:45

N, P_2O_5 and K_2O kg ha⁻¹ were applied in split doses. Each plot consisted of two rows of 6m length and observations were recorded on five randomly selected competitive plants from each genotype per replication for 16 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), ginning out-turn (%), elongation (%), lint yield per plant (g) and seed cotton yield per plant (g). 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), were recorded on plot basis. Mahalanobis D^2 analysis (Tochers 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 significant differences among 60 genotypes for all the 16 quantitative traits (Table 1) indicating the presence of variability among genotypes for characters studied. The per cent contribution towards genetic divergence by all the 16 characters is presented in Table -2. Based on D^2 statistic all the 60 genotypes were grouped into 15 clusters (Table-3) using

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

Source	d.f.	Days to 50% flowering	Plant height (cm)	No. of monopodia / plant	No. of sympodia / plant	No. of bolls / plant	Boll weight (g)	Ginning-out-turn (%)	Seed index (g)	Lint index (g)
Replications	2	1.739	5.962	0.028	0.111	6.536	0.008	0.914	0.017	0.052
Treatments	59	27.812**	406.651**	1.551**	7.075**	112.220**	0.612**	19.514**	3.768**	1.210**
Error	118	2.355	10.904	0.025	0.156	6.805	0.043	1.138	0.070	0.044

Mean squares	
Source	d.f.
2.5% Span length (mm)	Micro-naire (10 ⁻⁶ g/in)
Bundle Strength (g/ tex)	Uniformity Ratio
Elongation (%)	Lint yield per plant (g)
eed cotton yield / plant (g)	

Mean squares								
Replications	2	1.113	0.048	1.280	1.193	0.002	31.747	154.260
Treatments	59	14.466**	0.718**	4.417**	19.320**	0.005*	250.887**	1710.056**
Error	118	1.437	0.069	1.474	3.538	0.001	16.884	121.262

* = Significant at 5% level, ** = Significant at 1% level, d.f = degrees of freedom

Together's method with the criterion that the intra-cluster average D² values should be less than the inter-cluster D² values. The distribution of 60 genotypes into fifteen clusters was at random with maximum number of genotypes in clusters I (15 genotypes each). Cluster II (13 genotypes) was the second largest followed by cluster XII (11 genotypes) and VII (10 genotypes), and remaining 11 clusters are with 1 genotype each. The mutual relationships between the clusters were represented diagrammatically by taking a average intra and inter cluster D values. The average intra- and inter- cluster D and D² values are presented in Table 4. The inter cluster distance was maximum between clusters XIII and XV followed by clusters XII and XV; VI and XV ; XIV and XIII ; IV and XIII; V and XV . Based on the inter- cluster distances among the groups, crosses after confirming the general combining ability of the genotypes from clusters XIII and XV may yield better and desirable segregants.

In hierarchical clustering (Ward's minimum variance) method 60 genotypes were grouped into Eight clusters. Among all the clusters, cluster I was largest containing 23 genotypes followed by cluster II with 9 genotypes, cluster IV with 8 genotypes, cluster V and VI each with 5 genotypes, cluster III with 4 genotypes and cluster VII and cluster VIII each with 3 genotypes (Table-5). The average intra and inter cluster distances are presented in Table-6. Maximum inter- cluster distance was observed between clusters III and VIII; clusters III and VII; clusters II and VIII.

The first six principal components with eigen values more than one contributed 87.315 per cent towards the total variability (Table -7). The hierarchical cluster analysis and PCA confirmed findings of each other. The above results are supported by Altaher and Singh (2003).

Table 2. Contribution of different characters towards genetic divergence in 60 genotypes of cotton (*Gossypium hirsutum* L.).

Source	Times ranked first	Contribution % towards divergence
1 Days to 50% Flowering	13	0.73
2 Plant Height (cm)	236	13.33
3 No.of monopodia/ Plant	420	23.73
4 No. of sympodia/ Plant	376	21.24
5 No.of bolls/ Plant	59	3.33
6 Boll Weight (g)	56	3.16
7 Ginning Out Turn (%)	86	4.86
8 Seed Index (g)	390	22.03
9 Lint Index (g)	8	0.45
10 2.5%span Length(mm)	36	2.03
11 Micronaire (10^{-6} g/In)	71	4.01
12 Bundle Strength (g/tex)	3	0.17
13 Uniformity Ratio	9	0.51
14 Elongation (%)	3	0.17
15 Lint Yield/ Plant (g)	1	0.06
16 Seed Cotton Yield/ Plant (g)	3	0.17

Table 3. Clustering of 60 genotypes of cotton (*Gossypium hirsutum* L.) by Tocher's method.

Cluster No.	No. of genotypes	Name of genotype(s)
I	15	Tashkent 3,MCU 11,D6, NISD 2, G.Cot100, PeeDee011, ECV early,CPD 420, TSH 9904, NA 1584, RS 875, L 614, L 606, NH 577, F 2089
II	13	CNH 120 MB, SA-53-1, 010-1, ARB 9009, GISV-97/016, F 2020, TSH 9907, HLS 72, Dunn 118, Deltopine 66, IC 356932, CPD 478, ARB 815
III	1	IC 357103
IV	1	Tx lama
V	1	A 7262
VI	1	L 713
VII	10	Tidewater 53, TCH 1599, L604 sp, Empire glandless, Rai 123, G. Cot 16, CRH 71, CNH 7-94-2, IH 07, TSH 333
VIII	1	L 766
IX	1	TCH 1218
X	1	Gageti
XI	1	NA 777
XII	11	CCH 18, TSH 9819, KDCAKD, WGP 8, GSHY 97/13, L 741, AKH 9331, GSHY-01/1338, TSH 332, GJHV-01/35, KH 121
XIII	1	G204-13
XIV	1	CCH-05-1
XV	1	TXORHY-1-78

Table 4. Average intra-and inter-cluster D² values among 15 clusters in 60 genotypes of cotton (*Gossypium hirsutum* L.).

Cluster No	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	X11	X111	XIV	XV
I	26.00 (5.10)	38.08 (6.17)	44.24 (6.65)	89.57 (9.46)	65.93 (8.12)	112.74 (10.62)	62.06 (7.88)	40.85 (6.39)	33.85 (5.82)	39.71 (6.30)	37.71 (6.14)	89.59 (9.47)	102.41 (10.12)	63.96 (8.00)	119.84 (10.95)
II		35.01 (5.92)	57.09 (7.56)	61.70 (7.85)	86.91 (9.32)	146.66 (12.11)	73.26 (8.56)	51.75 (7.19)	46.95 (6.85)	52.67 (7.26)	62.07 (7.88)	133.57 (11.56)	133.54 (11.56)	57.48 (7.58)	84.10 (9.17)
III			0.000 (0.0)	91.73 (9.58)	17.37 (4.17)	56.21 (7.50)	71.50 (8.46)	70.32 (8.39)	48.21 (6.94)	73.12 (8.55)	46.75 (6.84)	94.17 (9.70)	125.23 (11.19)	88.93 (9.43)	157.80 (12.56)
IV				0.00 (0.0)	115.32 (10.74)	172.37 (13.13)	103.88 (10.19)	139.24 (11.80)	107.28 (10.38)	104.83 (10.24)	129.95 (11.40)	224.41 (14.98)	226.66 (15.06)	59.52 (7.71)	40.44 (6.36)
V					0.00 (0.0)	17.41 (4.17)	60.94 (7.81)	114.26 (10.69)	64.93 (8.06)	75.89 (8.71)	73.65 (8.58)	79.53 (8.92)	100.50 (10.02)	139.50 (11.81)	212.67 (14.58)
VI						0.00 (0.0)	88.08 (9.39)	167.85 (12.96)	105.82 (10.29)	108.83 (10.43)	123.97 (11.13)	95.00 (9.75)	132.22 (11.50)	199.13 (14.11)	286.27 (16.92)
VII							52.57 (7.25)	107.01 (10.34)	77.59 (8.81)	76.16 (8.73)	81.07 (9.00)	105.23 (10.26)	77.30 (8.79)	129.31 (11.37)	176.83 (13.30)
VIII								0.00 (0.0)	30.77 (5.55)	42.60 (6.53)	64.00 (8.00)	113.57 (10.66)	149.64 (12.23)	56.96 (7.74)	136.55 (11.69)
IX									0.00 (0.0)	31.62 (5.63)	73.68 (8.58)	84.94 (9.22)	133.90 (11.57)	69.87 (8.36)	132.75 (11.52)
X										0.00 (0.0)	75.65 (8.70)	91.84 (9.58)	119.22 (10.92)	75.07 (8.66)	123.15 (11.10)
XI											0.00 (0.0)	95.30 (9.77)	98.27 (9.91)	83.88 (9.16)	169.22 (13.01)
X11												80.92 (9.0)	103.38 (13.93)	193.92 (17.06)	291.17 (17.06)
X111														227.27 (15.08)	332.13 (18.22)
XIV														0.00 (0.0)	64.22 (8.01)
XV															0.00 (0.0)

Bold and diagonal values indicate intra-cluster distance, figures in parentheses are D² values

Table 5. Clustering of 60 cotton (*Gossypium hirsutum* L.) genotypes by Ward's minimum variance method.

Clster No	No. of genotypes	Name of genotype(s)
I	23	Dunn 118, L 614, NA 777, NH 577, RS 875, F 2089, ARB 815, TSH 9904, CNH 7-94-2, Pee Dee 011, ECV early, NA 1584, D6, F 2020, IC 356932, G Cot 100, NISD2, Tashkent 3, MCU 11, L 766, TCH 1218, Gageti, CPD 420
II	9	010-1, GISV-97/016, CNH 120MB, SA-53-1, ARB 9009, HLS 72, L 606, Deltopine 66, TSH 9907
III	4	CPD 478, Tx lama, TXORHY1-78, CCH-05-1
IV	8	Tidewater 5-3, TCH 1599, Rai 123, CRH 71, Empire glandless, A 7262, IC 357103, L 604 sp
V	5	TSH 333, L 741, CCH 18, TSH 9819, GSHY-01/1338
VI	5	WGP 8, KDCAKD, GSHY 97/13, GCot 16, AKH 9331
VII	3	IH 07, G 204-13, KH 121
VIII	3	TSH 332, L 713, GJHV-01/35

Table 6. Average intra – and inter- cluster Euclidian² values among the eight clusters in 60 cotton (*Gossypium hirsutum* L.) genotypes.

Cluster No	I	II	III	IV	V	VI	VII	VIII
I	105.35	123.76	254.50	189.34	239.18	182.15	324.13	434.14
II		91.66	179.60	230.37	350.87	287.13	432.16	604.62
III			130.82	324.46	588.86	452.33	737.00	855.32
IV				137.67	307.98	234.72	298.03	340.55
V					160.75	208.23	305.17	268.87
VI						137.44	318.98	322.70
VII							145.77	324.53
VIII								200.78

Bold and diagonal values indicate intra-cluster distance

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.) .

	PC1	PC2	PC3	PC4	PC5	PC6
Eigene Value (Root)	896.448	397.671	287.270	179.292	156.021	104.776
% Variance explained	38.721	17.177	12.408	7.744	6.739	4.526
Cum. Variance explained	38.721	55.898	68.306	76.051	82.790	87.315

All the three methods of grouping revealed a single concept of non-correspondence of genetic divergence and geographic diversity. The genotypes TXORHY-1-78, GJHV-01035, CCH-05-1 and GSHY-01/1338 showed maximum inter-cluster distance in the multivariate analysis and can be exploited for the development of heterotic hybrids after testing for their combining ability.

LITERATURE CITED

- Anderberg M R 1993** *Cluster Analysis for Application*. Academic Press, New York.
- Altaher A F and Singh R P 2003** Genetic diversity studies in upland cotton (*Gossypium hirsutum* L.) *Journal of Indian Society Cotton Improve.*, 28: 158-163.
- Gururajan K N and Manickam 2002** Genetic divergence in Egyptian cotton (*Gossypium barbadense* L.). *Journal of Indian Society Cotton Improve.*, 27(2): 77-83.
- Jain P P and Yadav I S 2001** Studies on genetic divergence in American cotton (*Gossypium hirsutum* L.). *Journal of Cotton Research Developement*, 15 (1) : 81-83.
- Jackson J E 1991** A User's Guide to Principal Components. John Wiley and Sons Inc., New York.
- Kalsy H S and Garg H R 1995** Genetic divergence and heterosis in American cotton. *Crop Improve.*, 22(2): 232-236.
- Karunakar Raju G B, Chenga Reddy V and Panduranga Rao C 2005** Multivariate analysis of genetic diversity in upland cotton. *Andhra Agricultural Journal* 52 (1&2): 66-74.
- Kumar P R, Raveedran T S and Krishnan V 2000** Genetic diversity and association studies for fibre characters in cotton (*Gossypium hirsutum* L.). *Journal of Indian Society Cotton Improve.*, 14(1): 111-118.
- Rao C R 1952** *Advanced Statistical Methods in Biometrical Research*. John Wiley and Sons Inc., New York pp. 236-272.