The Andhra Agric. J 55(3):286-292, 2008

Multivariate Analysis in Paprika Chilli (*Capsicum annuum* L.)

K Vijay Kumar, V Chenga Reddy, K V Siva Reddy, J Satish Babu, P V Rama Kumar and R Srinivasulu

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

ABSTRACT

Genetic divergence was assessed among forty genotypes of paprika chilli for 15 traits using Mahalanobis' D^2, principal component and cluster analyses. On the basis of these clustering methods eleven clusters were obtained in Mahalanobis'D^2 and seven clusters in hierarchical cluster analysis. Cluster VIII was the largest comprising eight genotypes in D^2 analysis and cluster IV was the largest comprising 11 genotypes in cluster analysis. In principal component analysis five principal components were identified which accounted for 88.41 per cent of the variability. PC_1 contributed 29.07 per cent of the total variability.

Key words : Cluster Analysis, Paprika Chilli, Principal Component Analysis and Ward's Minimum Variance

Thin and large skinned peppers commonly known as paprika are rich in carotenoid pigment capsanthin which imparts red colour embedded with low concentration of pungent principle capsaicin. Paprika is in great demand in international market for the fact that oleoresin extracted permits better distribution of colour and flavour in food industry. Little breeding efforts were made to reap benefits of paprika chilli. Hence, the amount of genetic diversity among the paprika chilli genotypes was estimated for planning the future crossing programme.

MATERIAL AND METHODS

The present investigation was undertaken to assess the nature and magnitude of diversity among the 40 genotypes of paprika chilli (Table1). The genotypes were planted in *kharif*, 2006 in randomized block design at Lam Farm, Guntur. Each genotype was raised following the intra- and inter-row spacing of 60 x 30 cm. The biometrical observations were recorded for plant height (cm), plant spread (cm), fruit length (cm), fruit girth (cm), number of fruits per plant, number of seeds per fruit, number of branches per plant, days to 50% flowering, days to maturity, 100-dry fruit weight (g), 1000-seed weight (g), oleoresin (%), capsanthin (EOA colour value), capsaicin (%) and dry fruit yield per plant (g).

Cluster analysis classifies a set of observations into two or more mutually exclusive unknown groups based on combinations of interval variables. Agglomerative heirarchial clustering technique (Ward's minimum variance) was followed for cluster analysis as given by Anderberg (1993). In the present study, the clusters were performed as per D^2 values (Mahalanobis, 1936). PCA was performed by using SPSS Software on the correlation of matrix of traits there by removing the effect of scale (Jackson, 1991).

RESULTS AND DISCUSSION

The analysis of variance revealed highly significant differences among the genotypes for 15 characters. The differences indicated the existence of lot of variability among the genotypes for the characters studied. On the basis of D^2 analysis and cluster analysis, the 40 genotypes were grouped into 11 (Table-2) and 7 clusters respectively (Table-3& Fig-1). 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 analyses, the scattering of genotypes from the same geographic region to different clusters might be due to the heterogeneity, genetic architecture of the general combining ability (Murthy and Arunachalam, 1966). The results have clearly indicated that there is no parallelism between the geographic diversity and genetic diversity as also reported by Chatterjee *et al.,* 2006.

In D^2 analysis, maximum intra-cluster distance was observed in cluster VIII (186.59) and minimum intra-cluster distance was observed in clusters IV, VI, VII, X and XI (0.00) (Table 4), while in cluster analysis maximum intra-cluster distance was observed in cluster VI (374.35) and minimum

Table1. Accession number and source of 40 paprika chilli (*Capsicum annuum* L.) genotypes

LCA-451	Pedigree selections from open pollinated hybrids (Hong Kong, Thailand, Korea and Japan)
LCA-1	-do-
LCA-431	-do-
LCA-444	-do-
LCA-439	-do-
LCA-440	-do-
LCA-443	-do-
LCA-437	-do-
LCA-445	-do-
LCA-446	-do-
LCA-448	-do-
LCA-447	-do-
LCA-452	-do-
LCA-424	-do-
LCA-442	-do-
LCA-450	-do-
LCA-418	-do-
LCA-453	-do-
LCA-452	-do-
LCA-454	-do-
LCA-455	-do-
LCA-456	-do-
LCA-457	-do-
LCA-458	-do-
LCA-459	-do-
LCA-460	-do-
LCA-2	-do-
GP-5	Native germplasm collection from Andhra Pradesh (Warangal, Khammam, Adilabad and West Godavari)
GP-6	-do-
GP-7	-do-
GP-8	-do-
GP-9	-do-
GP-10	-do-
GP-11	-do-
GP-12	-do-
GP-13	-do-
GP-14	-do-
GP-15	-do-
GP-17	-do-
GP-18	-do-

Table 2. Clustering of 40 paprika chilli (*Capsicum annuum* L.) genotypes by Tocher's method

Cluster No.	No. of genotypes	Name of the genotype(s)
I	2	LCA-444, LCA-439
II	4	LCA-445, LCA-459, GP-9, GP-12
III	7	LCA-1, LCA-452, LCA-447, LCA-458, LCA-446, LCA-455, LCA-457
IV	1	LCA-437
V	7	LCA-442, LCA-456, LCA-460, LCA-459, LCA-453, LCA-443, LCA-454
VI	1	LCA- 418
VII	1	LCA-451
VIII	8	LCA-424, LCA-448, LCA-452, LCA-2, GP-5,GP-7, GP-8, GP-15
IX	7	LCA-440, GP-14, GP-17, GP-18, LCA-440, GP-10,GP-11, GP-13
X	1	GP-6
XI	1	LCA-431

Table 3. Clustering of 40 paprika chilli (*Capsicum annuum* L.) genotypes by Ward's minimum variance method

Cluster No.	No. of genotypes	Name of the genotype(s)
I	9	LCA -451, LCA-445, LCA-446, LCA-450,LCA-454, GP-9, GP-12, LCA-424, LCA-2,
II	5	LCA-444, LCA-439, LCA-443,LCA-453, LCA-452
III	4	LCA-442, LCA-456, LCA-460, LCA-459
IV	11	LCA-1, LCA-452, LCA-447, LCA-458, LCA-455, LCA-418, LCA-457, LCA-437, LCA-431, GP-13, GP-11
V	4	LCA-440, GP-18, GP-14, GP-17
VI	4	LCA-448, GP-10, GP-6, GP-7
VII	3	GP-5, GP-15, GP-8

Table 4. Average intra- and inter-cluster D^2 values among seven cluster in 40 paprika chilli *Capsicum annuum* L.) genotypes

Cluster	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
I	**23.425**	70.728	91.011	103.63	106.09	154.256	106.708	127.238	196.84	270.27	217.75
II		**49.56**	116.64	132.02	102.616	139.71	60.84	149.32	189.33	317.55	318.625
III			**63.20**	108.16	142.086	93.315	101.20	245.23	185.77	457.10	170.30
IV				**0.00**	109.83	190.716	148.35	223.80	159.51	306.60	114.49
V					**92.736**	144.72	155.50	207.93	243.98	390.45	284.93
VI						**0.00**	133.17	301.369	298.94	627.00	214.62
VII							**0.00**	171.61	154.256	324.36	287.98
VIII								**186.59**	292.75	245.23	400.00
IX									**174.50**	325.80	305.55
X										**0.00**	548.496
XI											**0.00**

Bold diagonal values indicating intra-cluster distance

Table 5. Intra - and Inter -cluster squared Euclidean distance between seven clusters formed by Ward's minimum variance method in 40 paprika chilli (Capsicum annuum L.) genotypes.

	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V	Cluster VI	Cluster VII
Cluster I	**253.889**	385.716	416.928	493.921	545.548	705.625	509.850
Cluster II		**268.706**	429.139	500.636	723.931	663.739	642.005
Cluster III			**179.745**	521.294	904.791	911.358	641.209
Cluster IV				**351.039**	636.733	1082.954	901.842
Cluster V					**326.423**	756.563	843.090
Cluster VI						**374.351**	645.119
Cluster VII							**320.799**

Table 6. Per cent contribution of each character towards genetic divergence in as per D^2 analysis 40 genotypes of paprika chilli (*Capsicum annuum* L.)

Source	Times ranked first	Contribution (%)
Days to 50% flowering	7	0.90%
Plant height (cm)	45	5.77%
Plant spread (cm)	1	0.13%
Days to maturity	32	4.10%
No. of branches plant^{-1}	0	0.00%
Fruit length (cm)	73	9.36%
Fruit girth (cm)	36	4.62%
No.of fruits plant^{-1}	177	22.69%
100-dry fruit weight (g)	6	0.77%
No.of seeds fruit^{-1}	69	8.85%
1000-seed weight (g)	5	0.64%
Oleoresin (%)	1	0.13%
Capsanthin (EOA colour value)	90	11.54%
Capsaicin (%)	231	29.62%
Dry fruit yield plant^{-1} (g)	7	0.90%

Table 7. Eigen values per cent and cumulative variance in paprika chilli (*Capsicum annuum* L.) for six PC's.

	Canonical Roots Analysis (P. C. A.)					
	PC_1	PC_2	PC_3	PC_4	PC_5	PC_6
Eigen value (Root)	1080.028	810.7183	491.7891	395.3633	284.1245	222.1824
% Variance explained	29.07376	21.82409	13.23869	10.64296	7.64848	5.98103
Cumulative variance explained	29.07376	50.89785	64.13654	74.77952	82.42799	88.40903

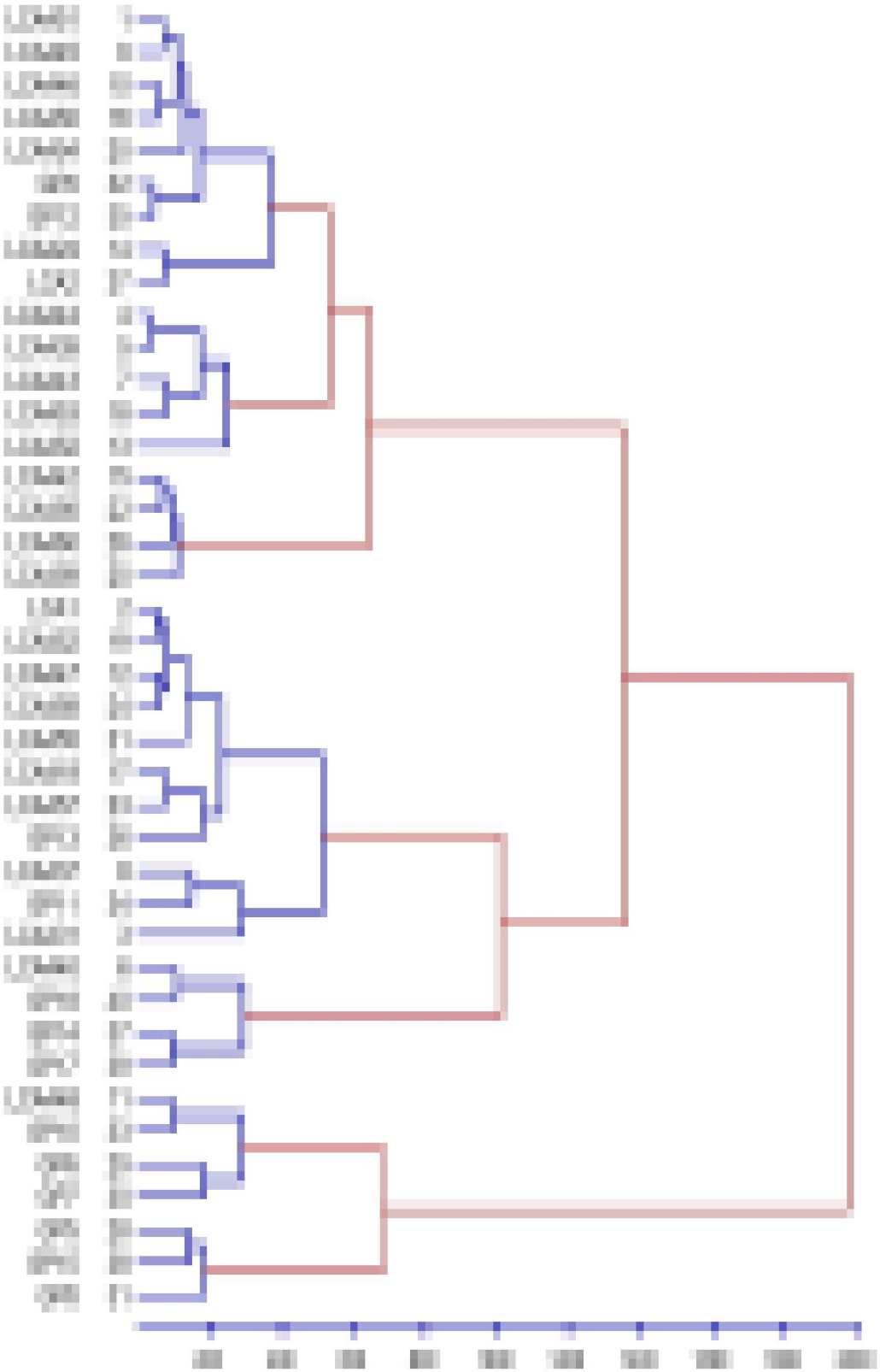

Fig. 1 :- Dendrogram showing relationshop of 40 paprika chilli (Capsicum annuum L.) genotypes in VII clusters.

2D PLOT

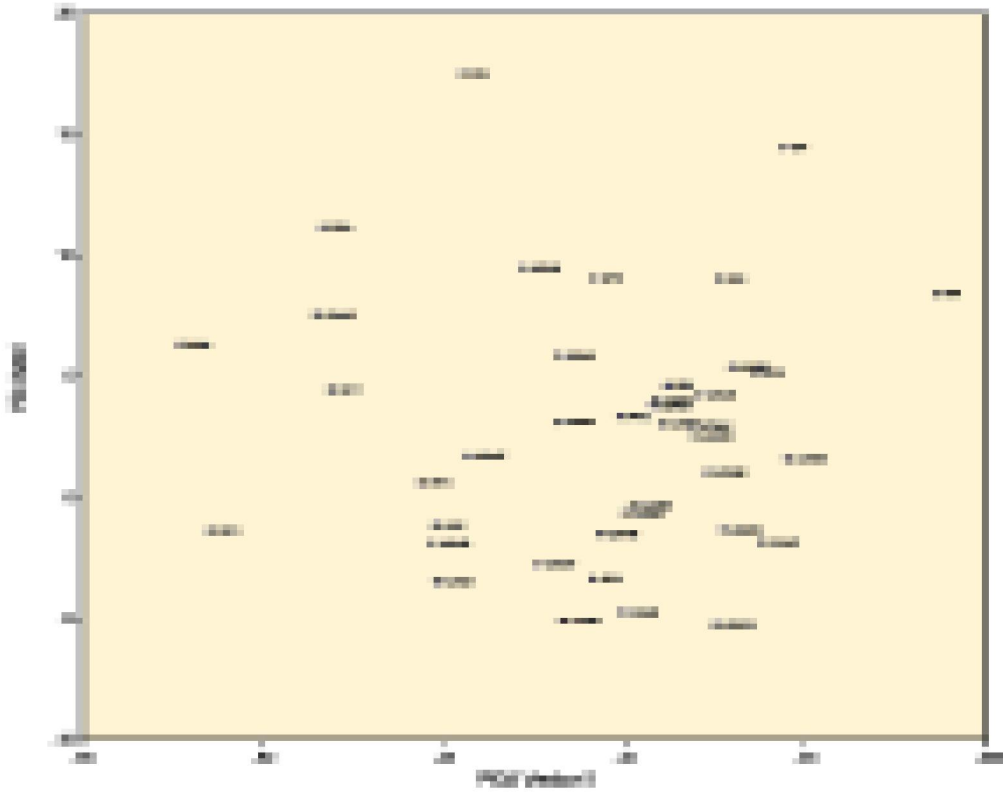

Fig. 2 :- Two dimenslonal graph showing relative position of paprika chilli (Capsicum annuum L.) genotypes based on PCA scores

3D PLOT

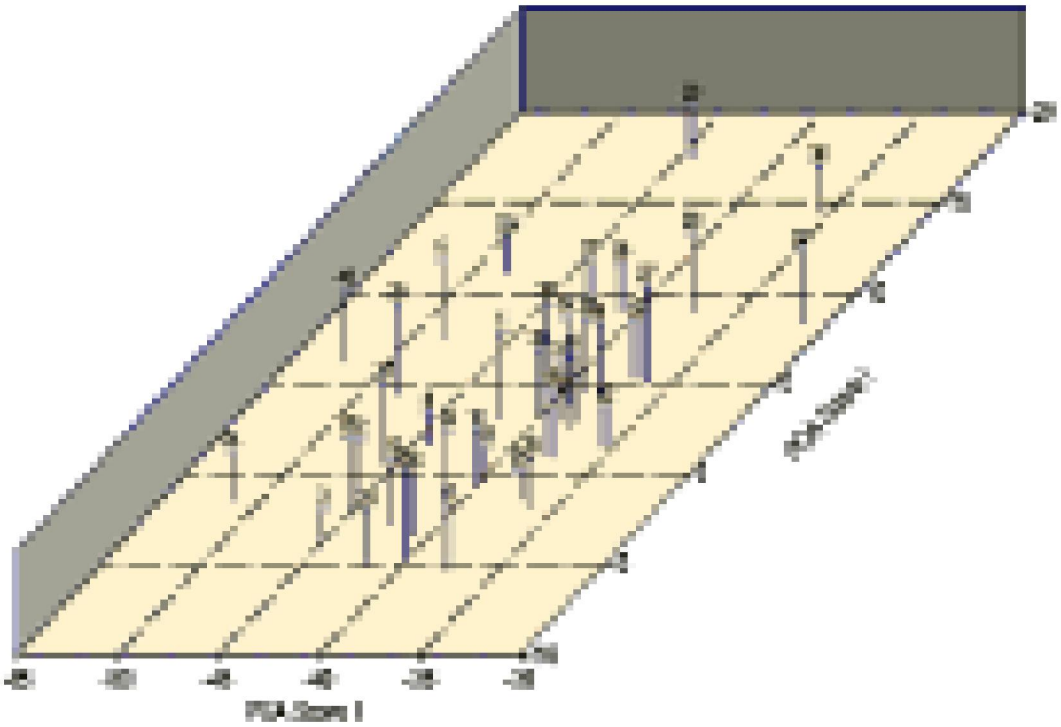

Fig. 3 :- Three dimenslonal graph showing relative position of paprika chilli (Capsicum annuum L.) genotypes based on PCA scores

intra-cluster distance was observed in cluster III (179.75).

Inter-cluster distance was least between clusters II and VII (60.84) and highest between clusters VI and X (627.00) on the basis of D^2 analysis (Table 4), while it was least between clusters I and II (385.72) and maximum between IV and VI (1082.95) on the basis of cluster analysis (Table-5). The hybridization between genotypes from the clusters with medium inter-cluster distances should give rise to heterotic hybrids. Cluster VIII was the largest comprising of eight genotypes in D^2 analysis (Table 2) and cluster IV was the largest comprising of 11 genotypes in cluster analysis (Table 3).

The additional advantage of D^2 analysis is estimation of the contribution of various characters towards the expression of the genetic divergence (Table-6). This analysis indicated that capsaicin content (29.62%) contributed maximum to the total divergence, followed by number of fruits per plant (22.69%) and capsanthin content (11.54%) while, rest of the characters contributed less significantly towards the divergence.Results of cluster analysis based on PCA scores were compared with the results of the principal component analysis on a visual aid in desecrating clusters in the 2D and 3D scattered diagrams. The genotypes falling in the same cluster are closer to each other in scattered diagrams. Brown (1991) and Altaher and Singh (2003) also studied the utilization of principal component analysis combined with clustering of Ward's method in genetic divergence studies on cotton.

In the present investigation, principal component (PC) method was used to extract the principal factor (PF) as it does not require the assumption of normal distribution of population. The PC's with eigen values ›1 were retained and ‹ 1 were considered as non-significant (Legendre and Legendre, 1984). The five principal components explained 88.41% of the variability (Table-7). The first PC explained 29.07% of the total variability in the set of all variables and remaining ones accounted for progressively lesser and lesser amount of variation.

The results of hierarchical cluster analysis and PCA confirmed the findings of each other. The plot of genotypes on 2D and 3D diagrams showed clear differential of genotypes according to their cluster membership in each cluster (Fig2 and 3). Genotypes belonging to a common cluster have fallen nearer to each other and vice-versa there by confirming the results of cluster analysis.

All the three methods of grouping revealed a single concept of non-corresponding of genetic divergence and geographic diversity. In a broad sense all the three methods of classifying the genotypes into different groups are equally useful but heirarchial cluster analysis gave an additional advantage of identifying sub-clusters 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.) using two methods of clustering. *Indian Society of Cotton Improvement,* 158-163.

Anderberg M R 1993 Cluster analysis of applicatioins. *Academic Press, New York.*

Brown S J 1991 Principal component and cluster analysis of cotton cultivar variability across the US cotton's belt. *Crop science,* 31:915-922.

Chatterjee B, Chenga Reddy V, Satish Babu J, Ramana J V and Pandurranga Rao C 2006 Application of principal component and cluster analyses in chilli (*Capsicum annuum*.L). *The Andhra Agricultural Journal.* 53 (3&4): 141-146

Jackson J E 1991 A user's guide to principal components. *John Wiley and Sons, New York.*

Legendre L and Legendre P 1984 Ecologia numerique. *Presses de Universite de Qubee, Canada,* Vol 1-2.

Mahalanobis P C 1936. On the genesalized distance in statistics. Proceddings of National Institute of Science 2: 49-55.

Murthy B R and Arunachalm V 1966 The nature of divergence in relation to breeding systems in some crop plants. *Indian Journal of*

(Received on 11.10.2007 and revised on 11.12.2007)