Principal component analysis for yield components in Greengram Accessions (*Vigna radiata* L.)

Srikanth Thippani[*], K.B. Eshwari and M.H.V. Bhave

College of Agriculture, Professor Jayashankar Agricultural University, Hyderabad
*Corresponding Author E-mail: thippanisrikanth@gmail.com

ABSTRACT

The present investigation was carried out to determine the relationship and genetic diversity among sixty greengram germplasm accessions using principal component for various morphological traits. In this study, Component 1 had the contribution from the traits viz., pod length, 100-seed weight and number of primery branches which accounted 21.64% to the total variability. Number of pods per cluster, number of pods per plant, number of seeds per pod and seed yield per plant has contributed 17.65% to the total variability in component 2. The remaining variability of 11.48%, 11.08%, 9.30% and 7.63% was console dated in component 3, component 4 and component 5 by various traits viz., Days to 50 flowering, Plant height, Number of pods per cluster, number of pods per plant, number of seeds per pod and seed yield per plant . The cumulative variance of 71.07% of total variation among eleven characters was explained by the first five axes. Thus the results of principal component analysis revealed, wide genetic variability exists in this greengram germplasm accessions.

Key words: *Principal component analysis, Greengram, Genetic diversity.*

INTRODUCTION

Studies on genetic diversity is the process by which variation among individuals or groups of individuals or populations is analysed. Data often involves numerical measurements and in many cases, combinations of different types of variables[9]. Evaluation of germplasm collection have the highest priority among germplasm functions. Germplasm enhancement embraces those activities required to aggregate useful genes and gene combinations into usable phenotypes. These aggregates could be considered as the feedstock for varietal development programmes[4].

Green gram (*Vigna radiata* (L.) Wilczek) is one of the important pulse crops of India. India alone accounts for 65% of its world acreage and 54% of the production. In India, it is grown on an area of 3.7 million hectares with production of 1.57 million tonnes and productivity of 406.98 Kg/ha (Anon., 2012). Mungbean has the potential to make up the gap of protein shortage, but its yield per hectare in the country is still low and there is a need for improvement.

Yield, being a quantitative trait is a complex character in any crop. Various morphological and physiological plant characters contribute to yield. These yield contributing components are inter-related with each other showing a complex chain of relationship and also highly influenced by the environmental conditions. In order to raise yield per unit area, new mungbean varieties must be developed along with the improved cultural practices. The crosses between the parents with maximum genetic divergence are generally the most responsive for genetic improvement[1]. To group the sets of germplasm accessions, Principal Component Analysis (PCA), cluster analysis[3,7] and Pearson's correlation analysis were used by taking into account several characters and relationship between them simultaneously. Recently PCA has been cited by various authors for the reduction of multivariate data into a few artificial varieties which can be further used for classifying material. This approach is especially valuable for screening large number of genetic resources by a large number of descriptor variables[2,5]. The aim of present study was to to identify better combinations as selection criteria for developing high yielding fine mungbean genotypes. Such type of findings may help mungbean breeders and it could provide new opportunities for promoting the production of mungbean with better yield.

MATERIALS AND METHODS

The material for the study comprised 60 sorghum germplasm accessions. The collection was undertaken in ANGRAU, Hyderabad during 2013. These genotypes were raised in a randomized block design with three replications in College farm, College of Agriculture, Rajendranagar. Five randomly selected plants from each genotype in each replication was taken for recording observations on 11 characters viz., Days to maturity, Number of primary branches, Number of clusters per plant, Number of pods per Cluster, Number of pods per plant, Number of seeds per pod, Pod length (cm), 100-seed weight (gm) and seed yield per plant.

Principal Component Analysis (PCA) was used as a data reduction technique to summarise the information of different phenotypic observations. The mean observation for each accession was standardized by subtracting from each observation the mean value of the character and subsequently dividing by its respective standard deviation. This resulted in standardized values for each trait with average 0 and standard deviation of 1 or less[10]. The standardized values were used to perform principal component analysis (PCA) using XLSTAT to know the importance of different traits in explaining multivariate polymorphism. The principal component analysis was performed in order to confirm the diversity pattern brought out by clusteranalysis.

RESULTS AND DISCUSSION

Analysis of variance for all the 11 characters studied is furnished in the Table-1. Significant differences among the entries for all the characters were observed, indicating the existence of a considerable amount of variation among the genotypes. The mean and range for different characters are given in Table 2. Days to 50% flowering and days to maturity varied from 30 to 40.33 and 54.00 to 68.50, average 35.2 and 61.30 days, respectively. MGG-360 was the earliest maturing variety of 54 days. Whereas genotypes COGG-912 and CO-6 were the latest maturing (68.50 days). The mean value recorded for the plant height was 36.45 cm with range from 23.50 cm (TMV-37) to 56.00 cm (LGG-479). Primary branches ranged from 0.27 (MGG-295) to 3.00 (RM-9-128) with a mean of 1.19. Number of clusters per plant ranged between 4.50 (MGG-360) to 18.20 (LGG-479) with a mean of 6.25. Number of pods per cluster varied from 2.60 (TLM-24) to 6.10 (LGG-549) with a mean of 3.82. Number of pods per plant varied from 7.53 (NM-1) to 30.40 (MGG-347) with a general mean of 14.30. The number of seeds per pod varied from 9.00 (DGGC-4) to 12.20 (MGG-341) with a mean of 10.90. The variability for pod

length ranged from 5.67 cm (PM-723) to 8.90 cm (TMV-37) with a mean value of 6.80 cm. Highest 100-seed weight of 4.81 g was recorded by TMV-37. Whereas, lowest value of 2.65 g was associated with CO-6 with a mean value of 3.42 g. Seed yield per plant ranged from 3.97 g (IPM-2K-14-19) to 14.89 g (RM-9-136) with a mean of 9.85 g. Besides, high GCV (24.94%), PCV (25.13%) and high heritability (98%) coupled with high genetic advance as per cent of mean (50.98%) were recorded for yield

Principal component analysis

The eigen values (variances), per cent variability, cumulative per cent variability and component loading of the different characters are given in Table 3. In the present studies, the first six components with eigen values more than one contributed 78.70 per cent of the cumulative variability amongst 60 greengram genotypes evaluated for 11 quantitative traits. The principal components with eigen values less than one were considered as non-significant. The first principal component contributed maximum towards variability (21.64%). Characters *viz.,* plant height (-0.52), pod length (0.47), number of pods per cluster (0.39), days to maturity (-0.34) and clusters per plant (-0.23) explained the maximum variance in first principal component (PC_1). This suggest that this component reflects high plant height and pod length of each accession. The second principal component (PC_2), which described 17.65 per cent of the total variance reflected significant loadings of number of pods per plant (0.42) and number of seeds per pod (0.59), which were positively correlated. This suggest that this component reflects number of pods per plant and number of seeds per pod of each accession. The third principal component (PC_3) was characterized by conspicuously high loading of days to 50% flowering (0.58) and clusters per plant (0.18) that were positively correlated. This suggest that this component reflects high days to 50% flowering and clusters per plant of each accession. The fourth principal component (PC_4) was characterized by conspicuously high

loading of days to 50% flowering (0.46) and 100-seed weight (0.48) that were positively correlated. This suggest that this component reflects high days to 50% flowering and 100-seed weight of each accession. The fifth principal component (PC_5) was characterized by conspicuously high loading of primary branches per plant (0.47) and days to 50% flowering (0.39) that were positively correlated. This suggest that this component reflects high primary branches per plant and days to 50% flowering of each accession.

The PCA scores for 60 genotypes in the first three principal components with eigen value more than one were computed and presented in Table 4. These three PCA scores for 60 genotypes plotted in graph to get the 2D (PCA I as X axis and PCA II as Y axis) and 3D (PCA I as X axis, PCA II as Y axis and PCA III as Z axis) scatter diagram (Fig. 1 and Fig. 2). The genotypes of divergent clusters like UPM-84-178, TMV-37 and LGG-479 scattered far apart in the 2D (Fig 1) and 3D (Fig 2) plots. The combinations between these genotypes would give better recombinants to exploit the heterosis. Ghafoor *et al.*[6] have takenfirst four components of PCA with eigen values >1 contributed 78.7percent of the total variance amongst 40 mungbean genotypes studied. Yimram *et al.*[11] evaluated 9 qualitative and 21 quantitative traits in 340 diverse cultivated mungbean accessions collected from AVRDC and principal component analysis revealed that the first three PCs explained 74.9 percent of the total variation. Eighteen quantitative and 37 qualitative characters for 646 greengram accessions were subjected to multivariate analysis and 63.79% variation was explained by first 3 principal components by Pandiyan *et al*[8].

The principal component scores of genotypes were used as input for clustering procedures in order to group the genotypes into various clusters and to confirm the results of principal component analysis. To group the genotypes into various clusters, hierarchical

cluster analysis was followed. Principal component scores were used as variables instead of attributes for clustering procedures, making the results equivalent to those from initially standardized data as the correlation matrix was used for principal component analysis.

Table 1: Analysis of variance for yield and yield components in sixty genotypes of greengram (*Vigna radiata* (L.) Wilczek)

S.No.	Character	Mean sum of squares		
		Replications (d.f=2)	Treatments (d.f=59)	Error (d.f=118)
1.	Days to 50% flowering	38.61	15.14**	4.14
2.	Days to maturity	66.15	46.63**	5.07
3.	Plant height (cm)	84.25	119.86**	7.33
4.	Number of primary branches per plant	0.00	0.57**	0.03
5.	Number of clusters per plant	27.47	10.49**	0.56
6.	Number of pods per Cluster	9.84	1.39**	0.35
7.	Number of pods per plant	0.86	69.15**	0.90
8.	Number of seeds per pod	10.29	1.39**	0.37
9.	Pod length (cm)	0.07	0.81**	0.09
10.	100-seed weight (gm)	0.01	0.38**	0.06
11.	Seeds yield per plant (gm)	0.15	18.19**	0.09

** Significant at 1 per cent level

Table 2: Mean and Range for seed yield and its components in greengram (*Vigna radiata* (L.) Wilczek)

S.No.	Character	Mean	Range	
			Minimum	Maximum
1.	Days to 50% flowering	35.02	30.00	40.33
2.	Days to maturity	61.30	54.00	68.50
3.	Plant height (cm)	36.45	23.50	56.00
4.	Number of primary branches per plant	1.19	0.27	3.00
5.	Number of Clusters per plant	6.25	4.50	18.20
6	Number pods of per Cluster	3.82	2.60	6.10
7.	Number of pods per plant	14.30	7.53	30.40
8.	Number of seeds per pod	10.90	9.00	12.20
9.	Pod length (cm)	6.80	5.67	8.90
10.	100-seed weight (g)	3.42	2.65	4.81
11.	Seed yield per plant (g)	9.85	3.97	14.89

Table 3: Character loading of six principal components for different characters of greengram
(*Vigna radiata* (L.) Wilczek)

Character	PC1	PC 2	PC 3	PC 4	PC 5	PC 6
Days to 50 flowering	0.06	0.10	0.57	0.46	0.39	0.01
Day to Maturity	-0.34	-0.33	0.03	0.23	-0.26	-0.53
Plant height (cm)	-0.52	0.02	0.05	0.00	0.27	-0.28
Number of primary branches per Plant	0.16	-0.14	0.08	-0.62	0.47	-0.18
Number of clusters per plant	-0.30	-0.37	0.18	-0.14	0.38	0.02
Number of pods per cluster	-0.39	0.17	-0.10	0.19	0.28	0.57
Number of pods per plant	-0.05	0.57	0.15	-0.26	-0.04	0.00
Number of seeds per pod	-0.09	0.59	0.13	0.01	0.01	-0.45
Pod length (cm)	0.47	-0.03	-0.06	0.07	0.33	-0.18
100 seed wt (g)	0.25	0.04	-0.39	0.48	0.32	-0.21
Seed yield per plant (g)	-0.24	0.16	-0.65	-0.07	0.24	-0.09
Eigene Value (Root)	2.38	1.94	1.25	1.22	1.02	0.84
% Var. Exp.	21.64	17.65	11.40	11.08	9.30	7.63
Cum. Var. Exp.	21.64	39.29	50.69	61.77	71.07	8.69

Table 4: The PCA scores of genotypes of 60 genotypes of greengram (*Vigna radiata* (L.) Wilczek)

Genotype	PCA-1	PCA- 2	PCA-3
Asha	-2.827	7.367	-17.902
CO-6	-6.802	3.782	-9.045
COGG-912	-1.288	5.379	-11.004
DGGC-4	-1.298	4.722	-10.299
GM-06-08	-1.260	5.712	-9.475
HUM-1	-2.742	7.297	-12.545
IPM-2K-14-19	-1.432	8.204	-6.739
IPM-302-02	-2.178	7.142	-15.019
KM-11-564	-2.866	8.611	-12.308
KM-200	-3.575	6.751	-16.490
KM-8-651	-5.502	6.578	-13.309
KM-8-657	-4.750	7.996	-13.912
LGG-407	-5.679	8.219	-19.764
LGG-410	-7.349	8.994	-19.522
LGG-450	-4.753	7.855	-15.848
LGG-460	-6.692	8.630	-18.199
LGG-479	-12.435	2.404	-16.223
LGG-491	-7.548	8.013	-18.638
LGG-521	-6.300	7.642	-17.310
LGG-522	-5.106	5.744	-14.292
LGG-541	-5.155	7.095	-17.978
LGG-544	-5.395	9.288	-13.677
LGG-547	-6.212	6.790	-17.142
LGG-549	-6.582	5.710	-16.343

|			
Madhira mung	-2.358	7.929	-14.660
MGG-295	-5.011	7.987	-19.997
MGG-330	-4.833	8.997	-20.923
MGG-332	-5.537	7.774	-15.320
MGG-335	-3.392	9.176	-18.642
MGG-341	-7.142	15.365	-20.072
MGG-347	-7.097	16.743	-18.289
MGG-348	-4.785	8.942	-17.815
MGG-351	-4.881	8.327	-18.510
MGG-359	-7.204	11.575	-18.718
MGG-360	-5.735	13.825	-16.408
MGG-361	-4.081	7.290	-13.298
MGG-477	-6.738	6.611	-10.971
ML-1628	-3.325	9.502	-12.355
NDMK-10-35	-5.382	9.371	-10.693
NM-1	-5.131	4.988	-14.222
OUM-145	-4.231	10.259	-10.185
PDM-54	-5.370	7.922	-12.296
PDM-993	-4.640	11.091	-11.562
PM-10-13	-2.687	6.622	-13.070
PM-11-3	-1.023	5.598	-11.548
PM-723	-4.146	6.672	-14.858
RM-8-665	-4.585	6.440	-15.017
RM-9-122	-1.828	7.096	-14.072
RM-9-126	-3.798	11.179	-16.876
RM-9-128	-3.505	8.020	-11.675
RM-9-136	-7.723	11.438	-21.413
TJM-3	-2.425	8.142	-15.257
TLM-10-109	-3.227	5.873	-13.839
TLM-24	-2.056	7.470	-13.996
TM-96-2	-2.437	5.998	-10.564
TMV-37	3.388	8.341	-16.426
TRM-1	-6.026	10.332	-11.336
UPM-84-178	-6.860	13.710	-7.072
WGG-2	-5.502	8.968	-10.054
WGG-37	-6.916	8.575	-18.573
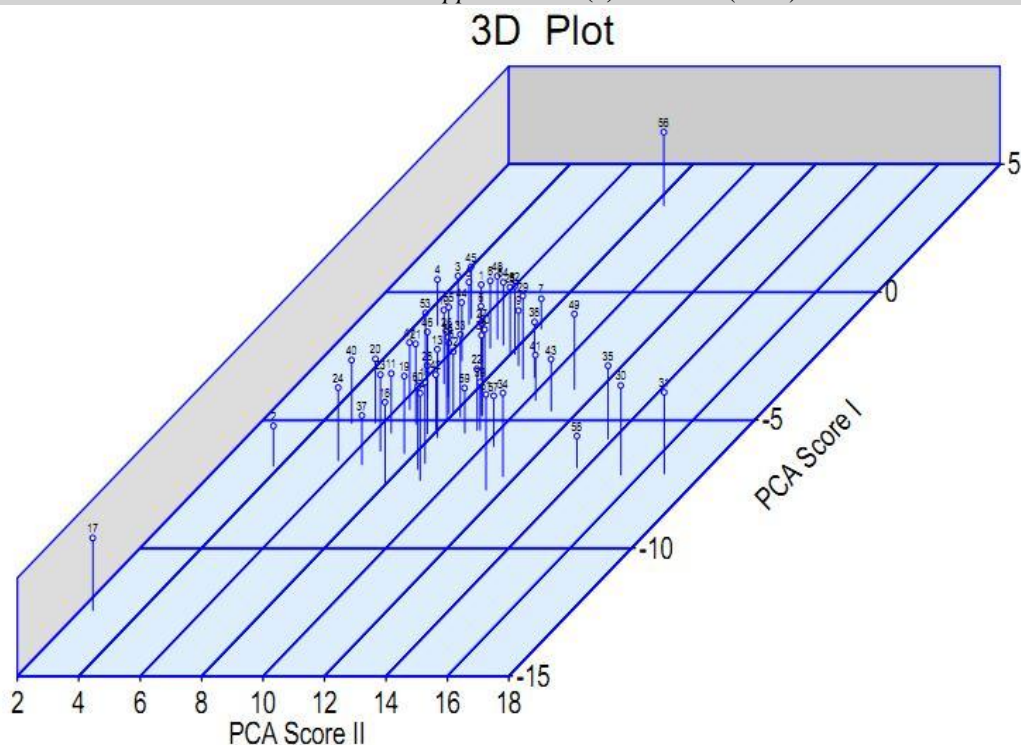
Fig 4.2. Three dimensional graph showing relative position of genotypes of greengrem based on PCA scores
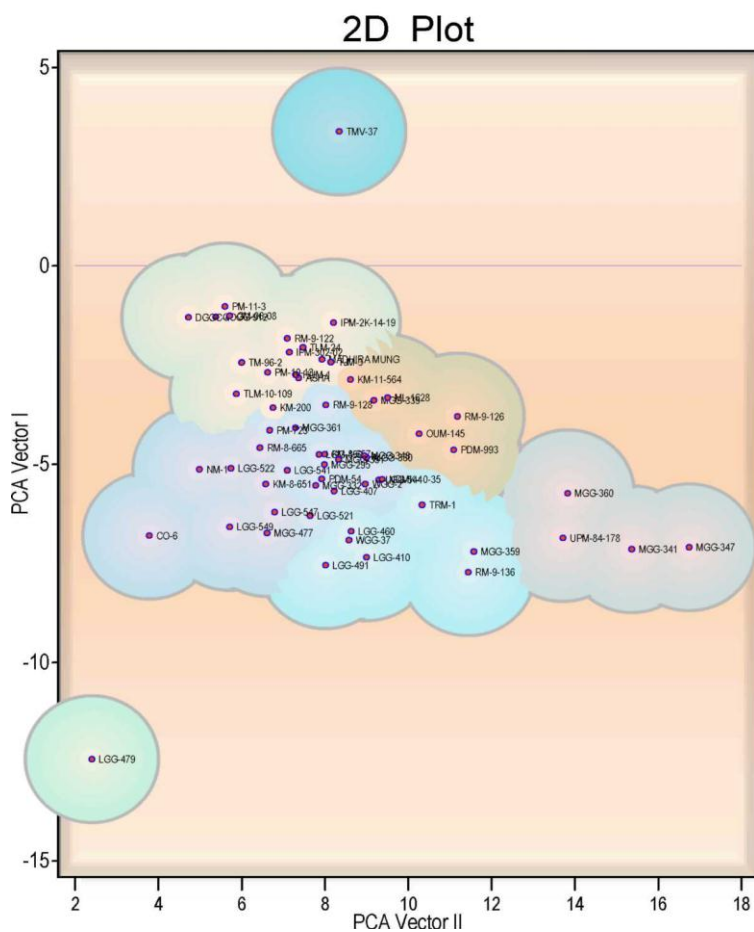

Fig 4.1. Two dimensional graph showing relative position of genotypes of greengrem based on PCA scores

REFERENCES

1. Arunachalam, V., Genetic distance in plant breeding. *Indian Journal of Genetics and Plant Breeding* **41:** 226-236 (1981).

2. Beiragi, M.A., Ebrahimi, M., Mostafavi, K., Golbashy, M.M., Khorasani, K.S., A Study of Morphological Basis of corn (*Zeamays* L.) yield under drought stress condition using Correlation and Path Coefficient Analysis. *Journal of Cereals and Oilseeds* **2(2):** 32-37 (2011).

3. Dillon, W.R., Goldstein, M., Multivariate Analysis:Methods and applications. John Wiley and Sons, New York (1984).

4. Evgenidis, G., Traka-Mavrona, E., Koutsika-Sotiriou, M., Principal Component and Cluster Analysis as a Tool in the Assessment of Tomato Hybrids and Cultivars. *International Journal of Agronomy* 1-7 (2011).

5. Golbashy, M., Ebrahimi, M., Khorasani, K.S., Choucan, R., Evaluation of drought tolerance of some corn (*Zea mays* L.) hybrids in Iran. *African Journal of Agricultural Research* **5(19):** 2714-2719 (2010).

6. Ghafoor, A. A., Sharif, Z., Ahmad, M.A., Zahid and Rabbani A. A., Genetic diversity in blackgram (*Vigna mungo* (L.) Hepper). *Field Crops Research*, **69:** 183-190 (2001).

7. Hair, J.F., Anderson, Jr. R.E., Tatham, R.L., Multivariate Data Analysis (2nd edn). Macmillan, New York (1987).

8. Pandiyan, M., Senthil, N., Packiaraj, D., ThangaPandian, R., Suresh, R., Jagadeesh, S., Characterization and evaluation of 646 greengram (Vigna radiate L.) genotypes for constituting core collection. *Wudpecker Journal of Agricultural* Research **1(8):** 294 –301 (2012).

9. Sneath, P.H.A., Sokal, R.R., Numerical Taxonomy: The Principles and Practice of Numerical Classification, W. H. Freeman and Co., San Francisco, California, USA (1973).

10. Upadhyaya, H.D., Ortiz, R., Bramel, P.J. and Singh, S., Phenotypic diversity for morphological and agronomic characteristics in chickpea core collection. *Euphytica* **123:** 333-342 (2002).

11. Yimram, T., Somta, P., Srinives, P., Genetic variation in cultivated mungbeangermplasm and its implication in breeding for high yield. *Field Crops Research* **112:** 260-266 (2009).