RESEARCH ARTICLE Legume Research- An International Journal, Volume 46 Issue 1: 10-17 (January 2023)

# Multivariate Analysis using $D^2$ and Principal Component Analysis in Mung bean [*Vigna radiata* (L.) Wilczek] for Study of Genetic Diversity

R.A. Jadhav[1], S.P. Mehtre[2], D.K. Patil[3], V.K. Gite[3]

10.18805/LR-4508


## ABSTRACT

**Background:** Seventy breeding lines of mung bean were evaluated for 20 different characters and mean values were worked for genetic diversity by Mahalanobis $D^2$ statistic.

**Methods:** The experiments included 70 mung bean breeding lines which were collected from Plant Breeding Unit, Agricultural Research Station, Badnapur. They were grown during *Kharif* 2016 at experimental research farm Badnapur of Vasantrao Naik Marathwada Krishi Vidyapeeth, Parbhani. The data were recorded for 20 different characters on 10 randomly selected plants. The statistical analysis were done by Mahalanobis $D^2$ statistics.

**Result:** The results of $D^2$ analysis indicated the presence of considerable genetic divergence among these breeding lines. In the present study, inter-cluster distances were worked out considering 20 characters and these distances ranged from 240.96 (between cluster V and cluster VII) to 1080.72 (between cluster V and cluster VIII). The inter-specific derivatives were grouped into eight clusters. The maximum inter-cluster distance was between cluster V and cluster VIII. The maximum inter-cluster distance was between cluster V and cluster VIII (1080.72), followed by cluster II and cluster III (932.25), cluster IV and cluster VIII (910.11), cluster VII and cluster VIII (738.30), cluster I and cluster VIII (732.61), cluster VI and cluster V (660.49) and cluster II and cluster VI (494.93). This suggested that there is wide genetic diversity between these clusters. Based on these studies, crosses can be made between breeding lines of these clusters to obtain desirable results either in transgressive breeding or in heterosis breeding. Cluster VIII and cluster IV showed high mean values for most of the yield contributing traits like100-seed weight, shelling %, harvest index, pod length, primary branches per plant, days to 50% flowering, days to maturity, leaf width and days to shattering. So the lines from cluster IV and cluster VIII can be used for mung bean yield improvement programme.


**Key words:** Genetic divergence, $D^2$ analysis, Mung bean, Principal component analysis.

## INTRODUCTION

Mung bean is an important grain legume, particularly in Asia. It is mainly utilized in making dhal, curries, soup, sweets and snacks. During the year 2019-20, estimated area and production of mung bean in Maharashtra is 19.28 lakh ha and 12.72 lakh tonnes with productivity of 441 kg/ha. Utilization of diverse genotypes is urgently needed for improvement of yield and quality as well as imparting disease and pest resistance. Information on genetic diversity is also valued for identification of diverse parental combinations to create segregating progenies with maximum genetic variability for further selection (Barrett and Kidwell, 1998), introgressions of desirable genes from diverse germplasm into the available genetic base (Thompson and Nelson, 1998) and the management of germplasm and for evolving conservation strategies.

Genetic diversity is of fundamental importance in the continuity of a species as it provides the necessary adaptation to the prevailing biotic and abiotic environmental conditions and enables change in the genetic composition to cope with changes in the environment. Genetic diversity in plants shows the magnitude of genetic variation existing in plant populations. The multivariate techniques, like $D^2$ analysis and principal component analysis may be an effective tool in the quantitative estimation of genetic diversity. Multivariate technique also plays a significant role in choice of divergent parents for hybridization to exploit maximum heterosis.

Department of Agricultural Botany, College of Agriculture, Vasantrao Naik Marathwada Krishi Vidyapeeth, Parbhani-431 402, Maharashtra, India.

[1]Department of Genetics and Plant Breeding, Vasantrao Naik Marathwada Krishi Vidyapeeth, Parbhani-431 402, Maharashtra, India.

[2]AICRP on Soybean, Vasantrao Naik Marathwada Krishi Vidyapeeth, Parbhani-431 402, Maharashtra, India.

[3]Agricultural Research Station, Badnapur-431 202, Maharashtra, India.

**Corresponding Author:** R.A. Jadhav, Department of Genetics and Plant Breeding, Vasantrao Naik Marathwada Krishi Vidyapeeth, Parbhani-431 402, Maharashtra, India.
Email: ranijadhav74@gmail.com

**How to cite this article:** Jadhav, R.A., Mehtre, S.P., Patil, D.K. and Gite, V.K. (2023). Multivariate Analysis using $D^2$ and Principal Component Analysis in Mung bean [*Vigna radiata* (L.) Wilczek] for Study of Genetic Diversity. Legume Research. 46(1): 10-17. doi: 10.18805/LR-4508.

**Submitted:** 16-09-2020 **Accepted:** 22-01-2021 **Online:** 08-03-2021

Grouping of genotypes by multivariate methods in the study is of great use to the mung bean breeders. Exploitation of the gene pool is of prime importance in order to develop high yield cultivars of mung bean. Therefore, the present study was scheduled to evaluate the 70 mung bean breeding lines on the basis of quantitative traits by using multivariate techniques to screen superior lines among them for the utilization as a parent in future hybridization program.

## MATERIALS AND METHODS

The material for present study comprised of 70 mung bean breeding lines collected from Plant Breeding Unit, Agricultural Research Station, Badnapur. The material for experiment included 70 breeding lines of mung bean [*Vigna radiata* (L.) Wilczek] including 12 were derivatives of three way crosses [(BM-4 × BWM-29) × BM-2003-2], 12 were derivatives of four way crosses ([(BM-4 × BWM-29) × BM-2003-2] × BM-2002-1), 41 were derivatives from inter-specific cross (BM-4 × BWM-29) and five were checks. The checks used for the study are BM-2003-2, AKM-4, BM-4, BM-2002-1 and BPMR-145. They were grown during *Kharif* 2016 at experimental research farm Badnapur of Vasantrao Naik Marathwada Krishi Vidyapeeth, Parbhani. The material was planted in a randomized block design with two replications. Each genotype was grown in two rows of 3 meter length with row to row and plant to plant spacing of 45 cm and 15 cm, respectively. The observations on 20 characters were recorded on five randomly selected plants in each of the two replications for days to 50% flowering, days to maturity, days to shattering, plant height, primary branches per plant, leaf length, leaf width, clusters per plant, pods per cluster, pods per plant, seeds per pod, pod length, 100-seed weight, shelling %, biological yield per plant, harvest index, protein content, calcium content, seed hardness and seed yield per plant. The replicated data were subjected to genetic divergence analysis using Mahalanobis' $D^2$ statistic (Mahalanobis, 1928) as described by Rao (1952). The set of breeding lines are grouped into different clusters following $D^2$ statistic (using Tocher's scheme).

## RESULTS AND DISCUSSION

The data collected on 20 yield and yield contributing characters from 70 mung bean breeding lines were subjected to genetic divergence analysis employing Mahalanobis' $D^2$ statistic, principal component and hierarchical cluster analysis. The magnitude of values suggested that there is considerable variability in the genotypes studied, which led to genetic diversity.

The per cent contribution of all 20 characters in 70 genotypes towards genetic divergence is presented in Table 1. Among the characters studied, calcium content (43.40%) showed maximum contribution followed by pods per plant (17.35%), plant height (17.27%), clusters per plant (6.09%), seed yield per plant (5.22%), harvest index (3.64%), seeds per pod (2.15%), seed hardness (1.95%), 100-seed weight (0.83%), pods per cluster (0.66%), primary branches per plant (0.62%), days to 50% flowering (0.25%), shelling % (0.21%), biological yield per plant (0.21%), protein content (0.12%) and leaf width (0.04%). While the contribution of other traits like days to 50% flowering, days to maturity, leaf length and pod length was zero per cent. Parthasarathy *et al.,* (2005) reported that days to maturity contributed maximum towards total genetic divergence followed by seed yield per plant. Vivekanandan *et al.,* (2005) observed 98% contribution of 100-seed weight. Gadakh *et al.,* (2013) found maximum contribution of protein content towards divergence. Singh *et al.,* (2015) showed the contribution of biological yield per plant, seed yield and pods per cluster.

### Grouping of genotypes into various clusters

The 70 inter-specific derivatives were grouped into eight clusters using the Tocher's method (Table 2). The distribution of 70 derivatives into eight clusters was at random with maximum number of genotypes in cluster I (26 genotypes) followed by cluster III having 21 genotypes. Cluster IV consists of eight genotypes. Cluster II and cluster VIII with six genotypes, whereas cluster V, cluster VI and cluster VII are monogenotypic consisting of one genotype.

### Average intra and inter-cluster $D^2$ value

The average intra and inter-cluster $D^2$ values were estimated as per the procedure given by Singh and Chaudhary (2010) and are presented in Fig 1. The nearest and farthest cluster for each of the eight clusters is indicated in Table 3. Inter-cluster divergence expresses the diversification among clusters (group of genotypes resembling each other, hence with low intra-cluster divergence). Intra-cluster $D^2$ values

**Table 1:** Contribution of different characters towards genetic divergence in 70 mung bean [(*Vigna radiata* (L.) Wilczek)] breeding lines.

| Source | No. of times ranked first | Per cent contribution |
|---|---|---|
| Days to 50% flowering | 6 | 0.25 |
| Days to maturity | 0 | 0.00 |
| Days to shattering | 0 | 0.00 |
| Plant height (cm) | 417 | 17.27 |
| Primary branches/plant | 15 | 0.62 |
| Leaf length (cm) | 0 | 0.00 |
| Leaf width (cm) | 1 | 0.04 |
| Clusters/plant | 147 | 6.09 |
| Pods/cluster | 16 | 0.66 |
| Pods/plant | 419 | 17.35 |
| Seeds/pod | 52 | 2.15 |
| Pod length (cm) | 0 | 0.00 |
| 100-Seed weight (gm) | 20 | 0.83 |
| Shelling % | 5 | 0.21 |
| Biological yield/plant (gm) | 5 | 0.21 |
| Harvest index (%) | 88 | 3.64 |
| Protein content (%) | 3 | 0.12 |
| Calcium content (mg/100 gm) | 1048 | 43.40 |
| Seed hardness (kg/cm$^2$) | 47 | 1.95 |
| Seed yield/plant (gm) | 126 | 5.22 |

**Table 2:** Clustering pattern of 70 mung bean [*Vigna radiata* (L.) Wilczek] breeding lines by Tocher's method.

| Cluster number | No. of genotypes | Name of genotype (s) |
|---|---|---|
| I | 26 | Vaibhav, AKM-4, BM-4, BM-2003-2, SPS-42-1-11-1, SPS-45-1, SPS-6-4-2, SPS-6-13-6, SPS-45-2, SPS-40-1-11, SPS-45-3, SPS-42, SPS-42-1-10, SPS-6-44-13, SPS-10, SPS-6-39-19, SPS-6-53-3, SPS-6-49-2, SPS-7, SPS-6-1-6, SPS-6-44-12, SPS-6-67-3, SPS-6-20-3, SPS-6-39-21, SPS-6-67-5, SPS-6-53-14 |
| II | 6 | SPS-6-42-1, SPS-6-42-9, SPS-6-42-18, SPS-6-52-5, SPS-7-4-10, SPS-8-20-4 |
| III | 21 | SPS-1-2, SPS-2-4, SPS-27, SPS-40-1-12, SPS-42-1-11-2, SPS-42-1-12-2, SPS-6-1-4, SPS-6-1-25, SPS-6-1-26, SPS-6-11-7, SPS-6-29-9, SPS-6-30-2, SPS-6-44-10, SPS-6-47-1, SPS-6-47-3, SPS-6-49-9, SPS-6-52-1, SPS-6-53-13, SPS-6-67-14, SPS-7-4-16, SPS-8-20-7 |
| IV | 8 | SPS-1, SPS-50, SPS-40-1-2, SPS-40-1-3, SPS-42-1-8, SPS-6-1-11, SPS-6-23-6, BPMR-145 |
| V | 1 | SPS-6-1-9 |
| VI | 1 | SPS-40-1-13 |
| VII | 1 | SPS-6-23-5 |
| VIII | 6 | SPS-40-1-7, SPS-42-1-12-1, SPS-6-52-6, SPS-6-52-15, SPS-6-53-4, SPS-6-67-8 |

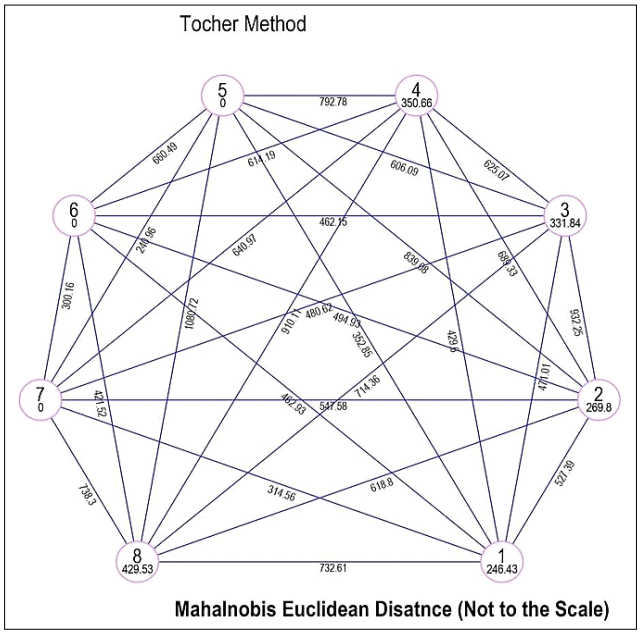



**Fig 1:** Inter and Intra-cluster distance of 70 mung bean breeding lines in eight clusters based on Euclidean distance.

ranged from 0.00 (cluster V, cluster VI and cluster VII) to 429.53 (cluster VIII). The intra-cluster distance indicated the diversity among genotypes grouped in that clusters. Genotypes grouped into the same cluster presumably differ little from one another as the aggregate of characters measured. General notion exists that the larger is the divergence between the parental genotypes, the higher will be the heterosis in crosses (Falconer, 1964). Therefore, it would be desirable to attempt crosses between genotypes belonging to distant clusters for getting highly heterotic crosses which are likely to yield a wide range of segregants on which selection can be practiced.

In the present study, inter-cluster distances were worked out considering 20 characters and these distances ranged from 240.96 (between cluster V and cluster VII) to 1080.72 (between cluster V and cluster VIII). Wide diversity was also reported by earlier workers where Ramana and Singh (1987) grouped 39 genotypes into eight clusters. Similarly Manivannan *et al.,* (1998) grouped 30 genotypes into eight clusters.

**Table 3:** The nearest and farthest cluster from each cluster based on $D^2$ values using Tocher's method in mung bean [*Vigna radiata* (L.) Wilczek] breeding lines.

| Cluster no. | Nearest cluster with $D^2$ values | Farthest cluster with $D^2$ values |
|---|---|---|
| I | VII (314.56) | VIII (732.61) |
| II | III (932.25) | VI (494.93) |
| III | VI (462.15) | II (932.25) |
| IV | I (429.60) | VIII (910.11) |
| V | VII ( 240.96) | VIII (1080.72) |
| VI | VII (300.16) | V (660.49) |
| VII | V (240.96) | VIII (738.30) |
| VIII | VI (421.52) | V (1080.72) |

Cluster I comprising eleven genotypes was nearest to cluster VII (314.56) followed by cluster V (352.85) and was farthest from cluster VIII (732.61) followed by cluster II (527.39). Cluster II consists of five genotypes. It was nearest to cluster VI (494.93) followed by cluster II (527.39) and was farthest from cluster III (932.25) followed by cluster V (839.08). Cluster III is second largest cluster which constitutes of 12 genotypes. It was nearest to cluster VI (462.15) followed by cluster I (471.01). It was farthest from cluster II (932.25) followed by cluster VIII (714.36). Four genotypes were grouped into cluster IV. It was nearest to cluster I (429.60) followed by cluster VI (614.19) and it was farthest from cluster VII (910.11) followed by cluster VI (792.78).

Cluster V also consist of four genotypes as that of cluster IV. It was nearest to the cluster VII (240.96) followed by cluster I (352.85) and it was farthest from cluster VIII (1080.72) followed by cluster II (839.08). Cluster VI comprised of three genotypes. It was nearest to the cluster I

(280.93) followed by cluster II (755.34) and was farthest from cluster VII (4451.42) followed by cluster VI (3163.07). Cluster VI consists of five genotypes. It was closest to the cluster VII (300.16) followed by cluster VIII (421.52) and was farthest from cluster V (660.49) followed by cluster IV (614.19). Cluster VII consists of nine genotypes. It was closest to the cluster V (240.96) followed by cluster VI (300.16) and was farthest from cluster VIII (738.30) followed by cluster IV (640.97). Cluster VIII is the largest cluster consisting of 20 genotypes. It was nearest to cluster VI (421.52) followed by cluster II (618.80) and was farthest from cluster (1080.72) followed by cluster IV (910.11).

The maximum inter-cluster distance was between cluster V and cluster VIII (1080.72), followed by cluster II and cluster III (932.25), cluster IV and cluster VIII (910.11), cluster VII and cluster VIII (738.30), cluster I and cluster VIII (732.61), cluster VI and cluster V (660.49) and cluster II and cluster VI (494.93). This suggested that there is wide genetic diversity between these clusters. Based on these studies, crosses can be made between genotypes of these clusters to obtain desirable results either in transgressive breeding or in heterosis breeding.

### Cluster mean values

The cluster means of all the 20 characters had presented in Table 4. Cluster means indicate average performance of all genotypes present in a particular cluster. The data indicated a wide range of mean values between the clusters.

For days to 50% flowering, least cluster mean was recorded by cluster II (36.83) followed by cluster II (36.90) and cluster VI (37.00) and high for cluster VIII (42.50). For days to maturity cluster II and cluster III (67.33) showed least cluster mean and cluster IV (73.44) showed high mean performance. For days to shattering, least cluster distance was recorded by cluster I (77.19) followed by cluster III (77.29) and high for cluster VIII (83.50). For plant height highest mean performance was recorded for cluster IV (57.95) followed by cluster II (44.08) and least for cluster V (30.84). For primary branches per plant lowest cluster mean was recorded for cluster V (3.06) followed by cluster VII (3.51) and cluster III (3.53) and highest for cluster II and cluster VIII (4.26). For leaf length cluster VI (15.31) showed highest cluster mean followed by cluster IV (12.23) and cluster V (12.08) and low mean performance for cluster VII (9.92). High mean performance for leaf width was recorded for the cluster IV (11.53) followed by cluster VI (11.24) and lowest for cluster VII (6.72).

The trait number of pod clusters per plant has recorded maximum cluster mean performance for the cluster II (8.28) followed by cluster VIII (7.84) and minimum for cluster V (4.86). The character number of pods per cluster had showed highest cluster mean for cluster II (5.53) followed by cluster VII (5.25) and least cluster mean for cluster IV (4.61). Number of pods per plant showed maximum cluster mean for cluster VI (48.90) followed by cluster II (45.95) and minimum cluster mean for cluster V (20.58). The trait seeds per pod had showed highest cluster mean for cluster VI (17.17) followed by cluster V (16.23) and least cluster mean for cluster IV (13.94). The character pod length had showed maximum cluster mean for cluster VI (9.17) followed by cluster V (8.73) and minimum cluster mean for cluster VII (6.32). In case of 100 seed weight, the highest cluster mean

**Table 4:** Mean values of eight clusters estimated by Tocher's method from mung bean [*Vigna radiata* (L.) Wilczek] breeding lines.

| Cluster no. | Days to 50% flowering | Days to maturity | Days to shattering | Plant height | Primary branches /plant | Leaf length | Leaf width | Clusters /plant | Pods/ cluster | Pods/ plant |
|---|---|---|---|---|---|---|---|---|---|---|
| I | 37.50 | 68.06 | 77.19 | 42.55 | 3.62 | 11.32 | 9.23 | 5.86 | 4.85 | 25.74 |
| II | 36.83 | 67.33 | 78.00 | 44.08 | 4.26 | 11.71 | 9.51 | 8.28 | 5.53 | 45.95 |
| III | 36.90 | 67.33 | 77.29 | 43.79 | 3.53 | 11.19 | 9.33 | 5.77 | 4.72 | 22.44 |
| IV | 42.31 | 73.44 | 83.25 | 57.95 | 3.94 | 12.23 | 11.53 | 5.28 | 4.61 | 27.03 |
| V | 41.00 | 71.00 | 81.00 | 30.84 | 3.06 | 12.08 | 8.97 | 4.86 | 4.69 | 20.58 |
| VI | 37.00 | 67.50 | 78.00 | 43.72 | 3.97 | 15.31 | 11.24 | 5.91 | 4.56 | 48.90 |
| VII | 37.50 | 68.00 | 78.50 | 31.38 | 3.51 | 9.92 | 6.72 | 5.24 | 5.25 | 37.13 |
| VIII | 42.50 | 73.25 | 83.50 | 38.93 | 4.26 | 10.93 | 9.61 | 7.84 | 5.03 | 41.60 |

**Table 4: Continue…**

| Cluster no. | Seeds /pod | Pod length | 100-Seed weight | Shelling % | Biological yield/plant | Harvest index | Protein content | Calcium content | Seed hardness | Seed yield /plant |
|---|---|---|---|---|---|---|---|---|---|---|
| I | 14.77 | 8.49 | 5.15 | 53.17 | 28.95 | 32.64 | 22.91 | 117.88 | 2.84 | 9.95 |
| II | 14.19 | 7.69 | 4.83 | 52.80 | 29.72 | 33.95 | 22.12 | 119.68 | 3.17 | 11.20 |
| III | 14.94 | 8.17 | 4.95 | 59.09 | 29.04 | 31.18 | 22.60 | 102.14 | 3.03 | 10.24 |
| IV | 13.94 | 8.53 | 5.33 | 51.97 | 25.96 | 36.18 | 22.58 | 115.68 | 2.61 | 9.86 |
| V | 16.23 | 8.73 | 4.24 | 28.93 | 17.95 | 25.45 | 21.73 | 120.63 | 4.53 | 5.60 |
| VI | 17.17 | 9.17 | 5.74 | 63.83 | 45.68 | 26.06 | 24.27 | 104.72 | 2.96 | 12.49 |
| VII | 14.71 | 6.32 | 3.48 | 47.77 | 13.32 | 30.33 | 25.00 | 113.84 | 2.63 | 4.48 |
| VIII | 14.54 | 8.37 | 6.20 | 63.86 | 42.92 | 39.42 | 20.29 | 103.84 | 3.01 | 18.09 |

was observed in VIII (6.20) followed by cluster VI (5.74) while least for cluster VII (3.48). For shelling % highest cluster mean was found in cluster VIII (6.20) followed by cluster VI (63.83) and least cluster mean was found for cluster V (28.93). Biological yield has showed maximum cluster mean for cluster VI (45.68) followed by cluster VIII (42.92) and minimum for cluster VII (13.32). For harvest index maximum cluster mean was observed in cluster VIII (39.42) followed by cluster IV (36.18) and minimum cluster mean was observed in cluster V (25.45). For protein content the highest cluster mean was for cluster VII (25.00) followed by cluster VI (24.27) and least cluster mean was found in cluster VIII (20.29). For calcium content the highest cluster mean was for cluster V (120.63) followed by cluster II (119.68) and least cluster mean was found in cluster III (102.14). For seed hardness highest cluster mean was found in cluster V (4.53) followed by cluster II (3.17) and least cluster mean was found in cluster I (2.61). For seed yield per plant highest cluster mean was observed in cluster VIII (18.09) followed by cluster VI (12.49) and least cluster mean was found in cluster VII (4.48).

Thus, cluster VIII and cluster IV showed high mean values for most of the yield contributing traits like100-seed weight, shelling %, harvest index, pod length, primary branches per plant, days to 50% flowering, days to maturity, leaf width and days to shattering. So the genotypes from cluster IV (SPS-1, SPS-2-4, SPS-6-1-11, BPMR-145) and cluster VIII (SPS-6-4-2, SPS-45-1, SPS-42-1-10, SPS-40-1-11, SPS-40-1-13, SPS-6-20-3, SPS-6-44-13, SPS-6-47-3, SPS-6-53-13, SPS-6-67-3, SPS-7-4-16, SPS-8-20-7, SPS-42-1-12-1, SPS-6-42-9, SPS-6-52-5, SPS-6-52-6, SPS-6-52-15, SPS-6-67-8, SPS-7-4-10, SPS-8-20-4) can be used for mung bean yield improvement programme. Similar results were the findings of Raje and Rao (2001), Haritha and Reddy (2002), Gupta *et al.,* (2004), Ahmed *et al.,* (2006), Yimram *et al.,* (2009), Ajay *et al.,* (2012), Gokulakrishnan *et al.,* (2012), Gadakh *et al.,* (2013), Sarkar and Kundagrami (2016) and Razzaque *et al.,* (2016).

**Principal component analysis**

PCA confirms the group constellations obtained by $D^2$ analysis. It determines the effective number of axis of differentiation primary and secondary or so based on number of canonical vectors. As a multivariate statistical technique, the principal components analysis (PCA) has the ability to transform a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components. The eigen values are often used to determine how many factors to retain. If the characteristic value is lower than one, it explains that the explanatory power of principal components is lower than the average explanatory power of the original variables.

The eigen values (variances), per cent variability, cumulative per cent variability and component loading of the different characters are given in Table 5. The principal

**Table 5:** Eigen values, proportion of the total variance represented by first seven principal components, cumulative per cent variance and component loading of 20 characters in mung bean breeding lines.

| | $PCA_1$ | $PCA_2$ | $PCA_3$ | $PCA_4$ | $PCA_5$ | $PCA_6$ | $PCA_7$ |
|---|---|---|---|---|---|---|---|
| Eigen value (Root) | 5.02 | 2.83 | 2.07 | 1.80 | 1.41 | 1.17 | 1.04 |
| % var. exp. (Variability %) | 25.12 | 14.13 | 10.33 | 8.99 | 7.04 | 5.88 | 5.21 |
| Cum. var. exp. (Cumulative %) | 25.12 | 39.25 | 49.58 | 58.57 | 65.61 | 71.48 | 76.69 |
| Days 50% flowering | 0.228 | 0.251 | 0.064 | 0.424 | 0.198 | 0.031 | 0.050 |
| Days to maturity | 0.159 | 0.285 | 0.022 | 0.372 | 0.231 | -0.275 | 0.019 |
| Days to shattering | 0.212 | 0.124 | 0.032 | 0.508 | -0.100 | -0.065 | -0.094 |
| Plant height (cm) | 0.115 | 0.484 | 0.006 | -0.255 | -0.028 | -0.129 | -0.035 |
| Primary branches/plant | 0.257 | 0.160 | -0.113 | 0.021 | -0.438 | 0.111 | 0.029 |
| Leaf length (cm) | -0.081 | -0.306 | -0.244 | 0.263 | 0.176 | 0.395 | -0.142 |
| Leaf width (cm) | 0.171 | 0.283 | -0.181 | -0.037 | -0.275 | 0.283 | -0.206 |
| Clusters/plant | 0.273 | -0.278 | 0.094 | -0.080 | -0.152 | -0.271 | -0.025 |
| Pods/cluster | 0.205 | -0.273 | -0.026 | -0.107 | 0.139 | -0.389 | -0.385 |
| Pods/plant | 0.333 | -0.223 | -0.194 | 0.031 | -0.073 | -0.032 | -0.086 |
| Seeds/pod | -0.145 | -0.125 | 0.153 | 0.167 | -0.526 | -0.202 | -0.002 |
| Pod length (cm) | 0.289 | 0.053 | -0.097 | -0.358 | -0.154 | 0.218 | 0.026 |
| 100-seed weight (gm) | 0.401 | -0.027 | -0.071 | -0.021 | -0.023 | 0.064 | -0.059 |
| Shelling % | -0.056 | 0.118 | 0.451 | -0.210 | 0.147 | -0.093 | -0.423 |
| Biological yield/plant (gm) | 0.169 | -0.149 | 0.408 | 0.008 | -0.067 | 0.129 | 0.423 |
| Harvest index (%) | 0.326 | 0.103 | 0.036 | -0.176 | 0.277 | -0.057 | -0.035 |
| Protein content (%) | -0.196 | 0.191 | -0.160 | -0.015 | -0.222 | -0.460 | 0.240 |
| Calcium content (mg/100 gm) | 0.023 | 0.066 | -0.390 | -0.189 | 0.305 | -0.106 | 0.444 |
| Seed hardness (kg/cm$^2$) | 0.123 | -0.259 | -0.346 | 0.046 | -0.037 | -0.290 | 0.017 |
| Seed yield/plant (gm) | 0.277 | -0.176 | 0.369 | -0.010 | 0.052 | 0.024 | 0.385 |

PC = Principal component.

component analysis sorted only significant principal components out of the total 20 attributes. The contribution of the main characters for variance easily identified by the characters loaded on the $PC_1$ with high loading values.

Generally, if eigen value is higher than one, it can be used as an inclusion criterion. The principal components with eigen values less than one were considered as non-significant. In present investigation seven principal components $PC_1$ to $PC_7$, with eigen values more than one which are extracted from the original data, contributed 76.70% of the total variation. However, remaining 13 components contributed only 23.30% of total variation for this set of 70 breeding lines evaluated for 20 quantitative traits. These principal component scores might be used to summarize the original 20 variables in any further analysis of the data. PCA scores for the first seven principal components Table 5.

$PC_1$ to $PC_7$ showed 25.12%, 14.13%, 10.33%, 8.99%, 7.04%, 5.88% and 5.21% variability among different traits under advance breeding lines. Among seven principal components first two components revealed major contribution (39.25%) towards total variation. Eigen value and variance associated with each principal component decreased gradually and cumulative variability increased gradually. The result of principal components is depicted as a character loading score Fig 1.

The characters effective in the first factor had a high level of loading coefficients. Here, the first principal component ($PC_1$) had high positive component loading from 100-seed weight (0.401) followed by pods per plant (0.333), harvest index (0.326), pod length (0.289) and seed yield per plant (0.277). On the other hand, high negative component loading was observed in $PC_1$ with protein content (-0.196) followed by seeds per pod (-0.145), shelling % (-0.056) and leaf length (-0.081). Here positive component loading indicated that selection for yield associated traits were effective from $PC_1$. This revealed that, $PC_1$ contributed maximum variability of yield associated traits.

The major contributing characters for variation in the second principal component ($PC_2$) were plant height (0.484) followed by days to maturity (0.285), leaf width (0.283), days to 50% flowering (0.251), protein content (0.191) and calcium content (0.066). High negative component loading was showed by characters *viz.,* leaf length (-0.306) followed by clusters per plant (-0.278), pods per cluster (-0.273), seed hardness (-0.259) and pods per plant (-0.223).

Likewise, the important traits *viz.*, shelling % (0.451) followed by biological yield per plant (0.408), seed yield per plant (0.369), seeds per pod (0.153) and clusters per plant (0.094) contributed more variation for $PC_3$. The characters like calcium content (-0.390) followed by seed hardness (-0.346), leaf length (-0.244), pods per plant (-0.194) and leaf width (-0.181) showed high negative loading.

Similarly, days to shattering (0.508) followed by days to 50% flowering (0.424), days to maturity (0.372), leaf length (0.263) and seeds per pod (0.167) had more variation in $PC_4$ and characters had high negative loading in $PC_4$ are pod length (-0.358) followed by plant height (-0.255), shelling % (-0.210), calcium content (-0.189) and harvest index (-0.176).

$PC_5$ had high positive loading for the characters *viz.,* calcium content (0.305) followed by harvest index (0.277), days to maturity (0.231), days to 50% flowering (0.198) and leaf length (0.176), whereas high negative loading for components seeds per pod (-0.526) followed by primary branches per plant (-0.438), leaf width (-0.275), protein content (-0.222) and pod length (-0.154).

$PC_6$ had high positive loading for the characters leaf length (0.395) followed by leaf width (0.283), pod length (0.218), biological yield per plant (0.129) and primary branches per plant (0.111). High negative loading was showed by components like, protein content (-0.460) followed by pods per cluster (-0.389), seed hardness (-0.290), days to maturity (-0.275) and clusters per plant (-0.271).

$PC_7$ had high positive loading for characters like, calcium content (0.444) followed by biological yield per plant (0.423), seed yield per plant (0.385), protein content (0.240) and days to 50% flowering (0.050). Again in the same principal component, characters which showed high negative loading are shelling % (-0.423) followed by pods per cluster (-0.086), leaf width (-0.206), leaf length (-0.142) and days to shattering (-0.094).

Usually, only one variable was selected from these identified groups. Hence, for the first group 100-seed weight was best choice, which had the largest loading from $PC_1$. Likewise, plant height, shelling %, days to shattering, calcium content, leaf length and calcium content was the best choice for second, third, fourth, fifth, sixth and seventh principal component, respectively.

The PCA scores for 70 breeding lines in the first three principal components with eigen value more than one were computed. These three PCA scores for 70 genotypes plotted in graph to get the 3D (PCA I as X axis, PCA II as Y axis and PCA III as Z axis) scatter diagram (Fig 2).

The genotypes of divergent clusters like SPS-42-1-8, SPS-6-1-26 and SPS-7-4-10 scattered far apart in the 3D (Fig 2) plots. The combinations between these genotypes would give better recombinants which will be utilized for development of elite cultivars in mung bean.

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. Principal component scores were used as characters 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.

From 3D graph, it can be concluded that the breeding lines which placed far apart on graph are SPS-42-1-8, SPS-6-1-26 and SPS-7-4-10. These lines are also grouped in

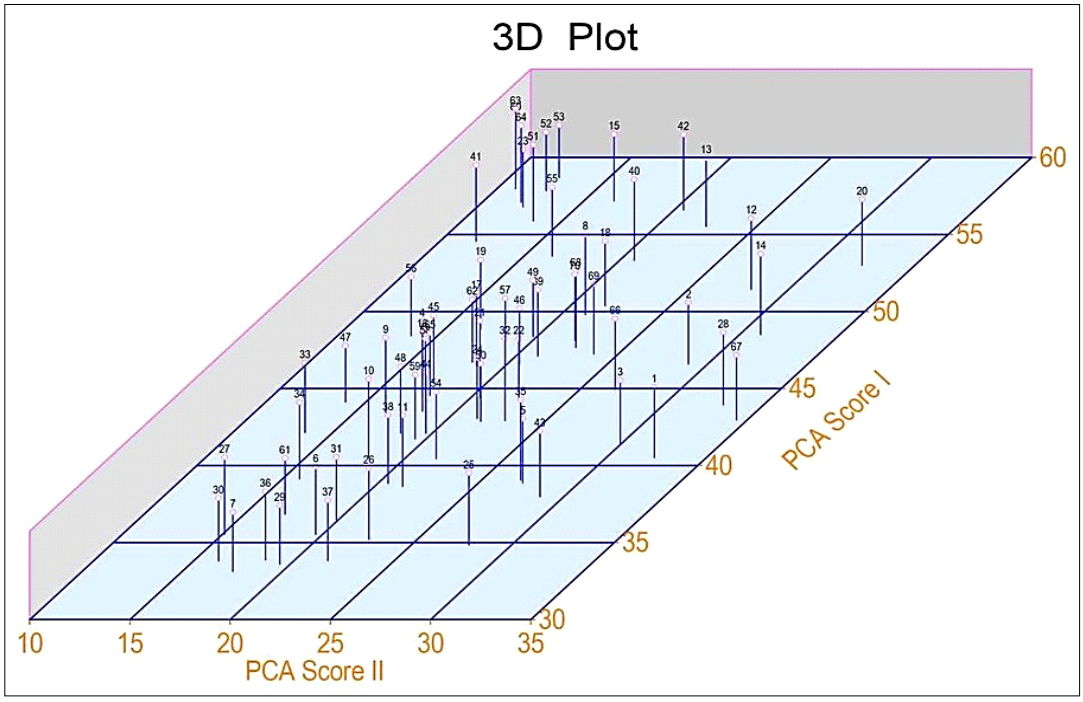



**Fig 2:** Three dimensional graph showing relative position of 70 Mung bean breeding lines based on PCA scores.

different clusters as per $D^2$ analysis. So these lines are quiet divergent from each other and in future can be utilized for developing elite cultivars in mung bean breeding programme. The results of $D^2$ analysis are confirmed with this PCA analysis and it can be proved that this genotypes are divergent form each other with maximum inter-cluster distance.

## CONCLUSION

It is inferred that with respect to seed yield per plant and number of pods per plant which contributed maximum towards genetic diversity and clusters VI and cluster VIII were the superior clusters. The genotypes in these clusters could be widely used in crossing programme for generation of wide spectrum of variability of yield.

## REFERENCES

Ahmed, N., Lavanya, G.R. and Kole, C.R. (2006). Estimation of genetic divergence in green gram. Journal of Maharashtra University. 32(3): 430-432.

Ajay, T., Tiwari, J.K. and Mishra, S.P. (2012). Genetic divergence analysis in mung bean [*Vigna radiata* (L.) Wilczek]. International Journal of Food, Agriculture and Veterinary Science. 2(3): 64-70.

Anonymous, (2019). Dpd.gov.in/pulses-kharif%20 prospects.

Barrett, B.A. and Kidwell, K.K. (1998). AFLP based genetic diversity assessment among wheat cultivars from Pacific Northwest. Crop Sci. 38: 1261-1271.

Falconer, D.S. (1964). An Introduction to Quantitative Genetics. 2nd ed. Oliver and Boyd Publishing Co. Pvt. Ltd. Edinburgh. 312-324.

Gadakh, S.S., Dethe, A.M., Kathale, M.N. and Kahate, N.S. (2013). Genetic diversity for yield and its component traits in green gram [*Vigna radiata* (L.) Wilczek]. Journal of Crop and Weed. 9(1): 106-109.

Gokulakrishnan, J., Kumar, S.B. and Prakash, M. (2012). Studies on genetic diversity in mung bean (*Vigna radiata* L.). Legume Research. 35(1): 50-52.

Gupta, S.K., Singh, K. and Rathore, P. (2004). Multivariate analysis in green gram [*Vigna radiata* (L.) Wilczek]. J. Res. Punjab Agric. Univ. 41: 306-309.

Haritha, S. and Reddy, S.M. (2002). Clustering of mung bean genotypes by complete linkage dendrogam. Legume Research. 25(4): 288-291.

Mahalanobis, P.C. (1928). A statistical study at Chinese head measurement. Journal of Asiatic Society of Bengal. 25: 301-307.

Mannivannan, N., Murugan, E., Viswanathan P.L. and Dhanakodi, C.V. (1998). Genetic divergence in greengram. Legume Research. 21(2): 131-133.

Parthasarathy, K., Eswari, K.B., Durga Rani, C.H. and Rao, S.V. (2005). Genetic divergence in green gram [*Vigna radiata* (L.) Wilczek]. The Andhra Agric. J. 52: 451-458.

Raje, R.S. and Rao, S.K. (2001). Genetic diversity in germplasm collections of mung bean [*Vigna radiata* (L.) Wilczek]. Indian Journal of Genetics. 61(1): 50-52.

Ramana, M.V. and Singh, D.P. (1987). Genetic parameters and character associations in green gram. Indian J. Agric. Sci. 57(9): 661-663.

Rao, C.R. (1952). Advanced Statistical Methods in Biometrical Research. John Wiley and Sons. New York. 236-272.

Razzaque, M.A., Haque, M.M., Rahman, M.M., Bazzaz, M.M. and Khan, M.S.A. (2016). Screening of mung bean [*Vigna radiata* (L.) Wilczek] genotypes under nutrient stress in soil. Bangladesh J. Agril. Res. 41(2): 377-386.

Sarkar, M. and Kundagrami, S. (2016). Multivariate analysis in some genotypes of mung bean [*Vigna radiata* (L.) Wilczek] on the basis of agronomic traits of two consecutive growing cycles. Legume Research. 39(4): 523-527.

Singh, C.M., Mishra, S.B., Pandey, A. and Arya, M. (2015). Multivariate analysis in mung bean [*Vigna radiata* (L.) Wilczek] to identify the genetic donors for pubescence and agro-morphological traits. Legume Research. 38(6): 767-771.

Singh, R.K. and Chaudhary, B.D. (2010). Biometrical Methods in Quantitative Genetic Analysis. Kalyani Publishers, New Delhi. 293-301.

Thompson, J.A. and Nelson, R.L. (1998). Core set of primers to evaluate genetic diversity in soybean. Crop Sci. 38: 1356-1362.

Vivekanandan, P., Gnanamalar, R.P. and Nanickavelu, A. (2005). Department of Plant Breeding and Genetics, Agril. College and Research Institute, Madurai. Research on Crops. 6: 343-393.

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