RESEARCH ARTICLE Legume Research- An International Journal, Volume 49 Issue 1: 15-24 (January 2026)

Principal Component and Cluster Analysis in Mungbean [*Vigna radiata* (L.) Wilczek]

Sayantan Das[1], Ashutosh Sawarkar[1], Samita Saha[1], R.B. Raman[2], T. Dasgupta[1]

10.18805/LR-5305

ABSTRACT

Background: Mungbean is a warm-season crop and its ability to grow quickly, allowing adaptation to multiple cropping systems and crop rotations. Principal component analysis (PCA) and cluster analysis are the useful tool to explain the diversity among the genotypes, identification of superior genotypes and characters forfuture breeding programs.

Methods: The experiment was conducted at the Agricultural Experimental Farm, (AEM), Division of Genetics and Plant Breeding, RKMVERI, Narendrapur, West Bengal, for summer, 2021 and 2022 in a randomized block design.

Result: Principal components *i.e.* PC_1 to PC_3 which eigen value more than one extracted from the data contributed 67.45% from the total variation. Cluster analysis revealed that the maximum contribution for the total divergence was recorded for number of branches per plant (25.7%), seed yield per plant (14.5%), plant height (13.8%) and number of pods per plant (12%). Genotypes had classified into six clusters and their intra and inter cluster distance showed the genetic diversity between them. The highest genotypes *i.e.* nine had in the cluster 1 followed by cluster 5 contributed six genotypes, cluster 2, 4 and 6 represented four genotypes each and the cluster 3 had three genotypes. The identified diverse genotypes like AKM 96-2, IPDI-539, PRATIKSHA NEPAL, SUKETI-1, TMB 96-2, SAMRAT, SIKHA and VIRAT showed high contribution to total diversity for number of branches per plant, plant height, number of pods per plant and number of primary branches per plant could be used in future breeding program.

Key words: Cluster analysis, Greengram, Multivariate analysis, Principal component analysis, Yield.

INTRODUCTION

Mungbean [*Vigna radiata* (L.) Wilczek] is also known as green gram, gold gram, moong. It is a diploid (2n=2x=22), self-pollinating grain legume belonging to *Fabaceae* family. Mungbean is a warm-season crop that provides a good opportunity for crop diversification due to its ability to grow quickly, allowing adaptation to multiple cropping systems and rotations of crops. In the economic year 2023, India produced over 3 million metric tonnes [Statista (2023) www.statista.com]. Despite numerous fluctuations, the trend has been upward. Currently, India leads the production and it is followed by China and Myanmar (Nair *et al.,* 2014). Mungbean is an excellent source of high-quality protein, fiber, minerals andvitamins.Mungbean's drought tolerance, low input requirements andshort growth cycle make it a worthwhile crop for cultivation, especially in areas with environmental stress. Furthermore, it is more adaptable and digestible than other pulses (Parida and Das, 2005).

There are significant constraints hampering the development of the mungbeanprogramme. For example, the limited genetic diversity that hinders mungbean productivity. Needless to say, there is a need to broaden the genetic base of the released varieties (Win *et al.,* 2020). Another, yield, a complex quantitative trait, is the result of multiple morphological and physiological traits are always influenced by environmental factors (Gupta *et al.,* 2023). Therefore, to develop new varieties of mung bean, an assessment of phenotypic diversity or characterization of morphological and agronomical characters plays an important role in the selection of desirable parents in the genetic improvement or plant breeding programme.

Multivariate techniques are an efficient tool for interpreting the genetic variation present in the germplasm, selecting it anddevelop the strategies to incorporate useful diversity in their breeding programmes. Several multivariate techniques are used to estimate genetic diversity, including metroglyph analysis, D^2 statistics, cluster analysis, principal component analysis, principal coordinate analysis, canonical analysis, factor analysis andcorrespondence analysis. Amon them, multivariate, cluster andprincipal component analyses are commonly used to estimate quantitative trait variability and identify superior genotypes.

[1]Division of Genetics and Plant Breeding, IRDM Faculty Centre, Ramakrishna Mission Vivekananda Educational and Research Institute, Narendrapur-700 103, Kolkata, West Bengal, India.
[2]Nalanda College of Horticulture, Bihar Agricultural University, Noorsarai, Nalanda-803 113, Bihar, India.

Corresponding Author: Ashutosh Sawarkar, Division of Genetics and Plant Breeding, IRDM Faculty Centre, Ramakrishna Mission Vivekananda Educational and Research Institute, Narendrapur-700 103, Kolkata, India.
Email: annu.sawarkar@gmail.com

How to cite this article: Das, S., Sawarkar, A., Saha, S., Raman, R.B. and Dasgupta, T. (2026). Principal Component and Cluster Analysis in Mungbean [*Vigna radiata* (L.) Wilczek]. *Legume Research.* **49(1):** 15-24. doi: 10.18805/LR-5305.

Submitted: 15-02-2024 **Accepted:** 09-07-2024 **Online:** 12-08-2024

(Jeberson *et al.,* 2017). Therefore, the present investigation to assess the genetic diversity using principal component analysis, on the basis of agronomical characters, classify the germplasm in the similar group and identify the superior genotypes which could be used as parents in breeding program.

MATERIALS AND METHODS

The research investigation was conducted at the Agricultural Experimental Farm, (AEM), Division of Genetics and Plant Breeding, RKMVERI, Narendrapur, West Bengal, over a span of two years, from 2021 and 2022 in summer season. The AEM, Narendrapur lies at 22°43′N latitude and 88°40′E longitude belongs to New Alluvial Zone at an elevation of 8 m above mean sea level.

A total 30 diverse germplasm obtained from the ICAR-Indian Institute of Pulse Research, Kanpur which are presented in Table 1. The experiment was laid out in the randomized complete block design (RCBD) with three replications. The size of unit plot was 3 m×2.2 m. The distances of 45 cm between rows, 15 cm within the plants, 75 cm between plots and 1 m between replications. The experimental field was prepared; genotypes were sown and followed various cultural operations for better growth of seedlings. The thinning was followed by 12 days after sowing and left one plant on a hill. Five randomly selected plants were harvested per plot when colour of almost all the pods become black or brown. Pods of each plant were kept separately in paper bags and sun dried. Threshing was done by hand and strict care was taken to avoid mechanical mixture of seeds.

11 yield attributing characters were recorded and presented in Table 2. The recorded data were analyzed for descriptive statistics (mean, maximum, minimum, Standard error of difference, coefficient of variation). Pearson's simple correlation coefficient was studied among yield attributing characters. The traits were analyzed as multivariate by using procedures of principal component analysis (PCA) and cluster analysis (CA) with the help of R software version 4.3.2. PCA was carried out using a correlation matrix to reveal the relationship between quantitative traits that are correlated by converting them into uncorrelated traits known as PCs (Johnson and Wichern, 1988). PCA provided information about the relative importance of each trait for screening germplasm in this research. The other multivariate analysis, most widely used clustering technique, D^2 statistics were employed suggested by Mahalanobis (1936) and Rao (1952).

RESULTS AND DISCUSSION

The simple descriptive statistics including mean value, maximum and minimum value and coefficient of variation (CV) of the 11 morphological characters studiedpresented in Table 3. The statistical analysis showed that wide range of variation indicated in the germplasm for the characters. High variation was recorded in DTF(36.50-53.16), DTM (63.50-71.83), PH (43.12-76.06 cm), NPB1.480-4.34), BPP (5.47-9.91), NCP (5.37-10.65), NPP (16.20-42.18), LOP (6.76-8.72 cm), NSP(9.33-14.11), HSW(2.55-4.50 g) andSYP.(4.32-13.76 g). Among genotypes, VIRAT, SIKHA, SAMRAT and IPDI-539 showed highest SYP and NPP. Early flowering was recorded inSIKHA, IPM2K-14-9 and MH2-15. TMB 96-2 had consistently good performer for NPB,

Table 1: List of mungbean germplasm.

Code	Germplasm
1	AKM 96-1
2	AKM-96-2
3	CO8
4	COGC-912
5	EC5200-34
6	IPDI-539
7	IPM2-17
8	IPM2-5-8
9	IPM2K-14-9
10	MH2-15
11	MH-521
12	MH-805
13	MI-1464
14	ML-1059
15	ML-2056
16	ML-5
17	NM-1
18	PDM 04-123
19	PDM-281
20	PDM-54
21	PRATIKSHA NEPAL
22	PUSA-0672
23	PUSA-0891
24	PUSA-9972
25	SML-1808
26	SUKETI-1
27	TMB 96-2
28	SAMRAT
29	SIKHA
30	VIRAT

Table 2: List of morphological characters of mungbean.

Characters	Code
Days to 50% flowering	DTF
Days to maturity	DTM
Plant height (cm)	PH
Number of primary branches per plant	NPB
Number of branches per plant	BPP
Number of clusters per plant	NCP
Number of pods per plant	NPP
Pod length (cm)	LOP
Number of seed per plant	NSP
100 seed weight (g)	HSW
Seed yield per plant (g)	SYP

Table 3: Variations in quantitative traits of munbean genotypes on pooled basis.

	Minimum	Lowest five genotypes	Maximum	Top five genotypes	Mean±SEd	CV
DTF	36.500	SIKHA, IPM2K-14-9, MH2-15, PRATIKSHA NEPAL, IPDI-539	53.167	ML-1059, PUSA-9972, PDM 04-123, SML-1808, PDM-281	43.273±1.198	4.789
DTM	63.500	IPDI-539, SUKETI-1,EC5200-34, MH2-15,ML-2056	71.833	ML-1059, PDM 04-123, MH-521, TMB 96-2, PUSA-9972	67.955±1.048	2.672
PH	43.120	ML-5, SML-1808, MH-521, PUSA-0891, EC5200-34	76.067	VIRAT, SIKHA, AKM-96-2, IPM2-5-8, ML-1059	57.048±1.517	4.604
NPB	1.480	MH-805, PDM 04-123, NM-1, EC5200-34, ML-1059	4.342	TMB 96-2, SIKHA, SAMRAT, SUKETI-1, VIRAT	2.483±0.135	9.491
BPP	5.475	PRATIKSHA NEPAL, ML-1059, ML-5, MH-805AKM-96-2	9.916	PUSA-0891, TMB 96-2, VIRAT, PUSA-0672, PDM 04-123	7.648± 0.204	4.637
NCP	5.371	MH2-15, ML-2056, ML-5, PUSA-9972, IPM2K-14-9	10.650	SAMRAT, MH-521, VIRAT, TMB 96-2, AKM-96-2	6.596±0.388	10.189
NPP	16.200	ML-2056, PDM-54, ML-1059, COGC-912,PDM-281	42.185	SIKHA, VIRAT, SAMRAT, IPDI-539, AKM-96-2	23.218±1.154	8.611
LOP	6.768	PDM 04-123, MH2-15, TMB 96-2, IPDI-539, AKM-96-2	8.727	SUKETI-1, EC5200-34, PRATIKSHA NEPAL, IPM2-17, IPM2-5-8	7.481±0.1915	4.432
NSP	9.330	MH-521, MH2-15, AKM 96-1, PDM-54, IPDI-539	14.117	EC5200-34, SAMRAT, IPM2-17, COGC-912, SUKETI-1	11.084±0.539	8.435
HSW	2.557	IPM2-5-8, COGC-912, MI-1464, PDM-54, MH-521	4.501	TMB 96-2, SAMRAT, SUKETI-1, VIRAT, IPM2-17	3.445±0.1315	6.614
SYP	4.327	COGC-912,PDM-54, MI-1464, PDM-281, ML-1059	13.766	VIRAT, SIKHA, SAMRAT, IPDI-539, SUKETI-1	7.554±0.4655	10.672

BPP, NCP and TSW. While, EC5200-34 had maximum NSP and LOP, it was followed by Suketi-1. CV is the parameter to measure the genetic variability for the characters. High CV (>10%) was observed for SYP and it was followed by NCP represent the substantial amount of genetic variability and with the accordance of the findings of Gupta *et al.* (2023). Moderate CV (5-10%) was observed for NPB and it was followed by NPP, NSP and HSW. Low variability was found in DTF and it was followed by BPP and PH. Similar findings were corroborated by Sarkar and Kundagrami (2016) and Kaur *et al.* (2017).

The correlation coefficient among all yield attributing characters isanalyzed by Pearson's correlation are represented in Fig 1. The dark blue and red cell indicates highest and lowest significant positive and negative correlationcoefficient respectively. The highest significant positive correlation had revealed between SYP and NPP, it was followed in between DTF and NPB, betweenSYPand NPB also with HSW, between NPB and NCP, between DTF and DTM and between PH and NPP respectively. These characters could be considered for the future breeding programme in mungbean. On the other side, DTF had significant negative correlation coefficient between NPP and also exhibited with SYP. Any successful plant breeding program involves careful selection of suitable parents. Parents with greater genetic distance are likely to produce more variation. The genotypes with wide variation provide more scope for selection (Canci and Toker, 2014). According to Bos and Caligari (1995), more genetic variation in traits leads to greater genetic gain. Yimram *et al.* (2009) reported that significant variation in mungbean correlated with growth, phenology, yield components and grain yield.

Cluster analysis was conducted to explain the genetic relationships among various genotypes and to identify suitable parents for the breeding programs. The importance of genetic diversity in parental genotypes is emphasized for the enhancement of breeding initiatives. Utilizing a hierarchical classification approach with Ward linkage clustering method based on 11 yield attributing traits, 30 germplasm were categorized into six distinct and clearly defined clusters in Table 4. Among these, the highest genotypes namely, nine had in cluster 1 followed by cluster 5 contributed six genotypes, cluster 2, 4 and 6 represented four genotypes each and the cluster 3 had only three

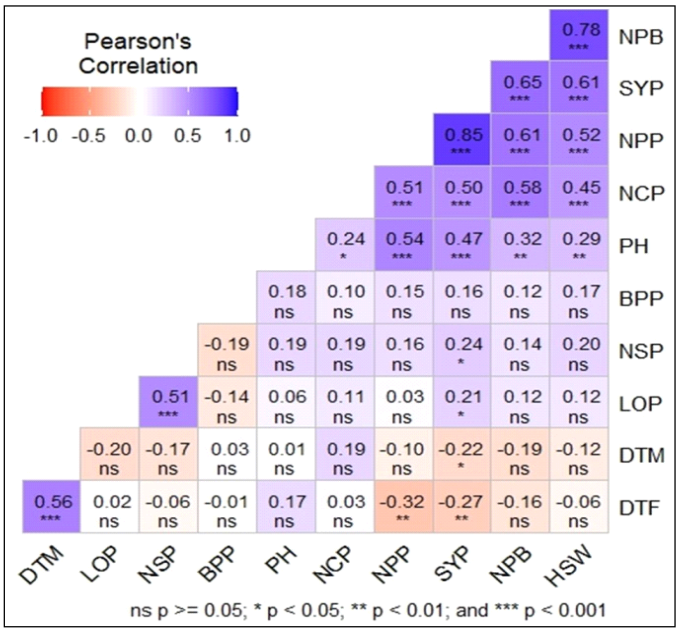


Fig 1: Correlation coefficients among 11 quantitative traits in mungbean genotypes.

Table 4: Distribution of 30 mungbean germplasm in different clusters related to agronomic characters over two years.

Cluster	No. of genotypes	Percentage	Genotypes
I	9	30	AKM 96-1, CO8, IPM2-5-8, MI-1464, PDM-281, PDM-54, PUSA-0672, PUSA-0891, SML-1808
II	4	13.33	AKM-96-2, IPM2-17, PRATIKSHA NEPAL, SUKETI-1
III	3	10	COGC-912, EC5200-34, ML-5
IV	4	13.33	IPDI-539, IPM2K-14-9, MH2-15, ML-2056
V	6	20	MH-521, MH-805, ML-1059, NM-1, PDM 04-123, PUSA-9972
VI	4	13.33	TMB 96-2, SAMRAT, SIKHA,VIRAT

genotypes. Genotypes within the same cluster are indicative of a closer genetic relationship compared to those belonging to different clusters. Fig 2 showed a dendrogram that the distribution of genotypes across clusters revealed significant genetic variability. The average intra and inter-cluster D^2 values were estimated among 30 Germplasm (Table 5). The least intracluster distance was observed in cluster I (2.625) it indicates minimum difference among the germplasm occupied in this cluster. The maximum intra cluster D^2 value was observed in cluster VI (4.985) followed by cluster III (3.883), cluster II (3.381), cluster IV (3.192) and cluster V (2.912) indicating that maximum differences occur among the genotypes that fall in these clusters. Furthermore, the maximum intercluster distance was between cluster III and cluster VI (7.263) followed by cluster V and cluster VI (6.750), cluster I and cluster VI (6.239) and cluster IV and cluster VI (6.087). It indicates that genotypes lied in these clusters are genetically diverse. According to Falconer (1964), the greater divergence between parental genotypes parallels to increased heterosis in crosses. It is advantageous to undertake crosses between genotypes originating from distant clusters. Therefore, it would be desirable to select a diverse array of segregants to produce highly heterotic crosses, which could be suitable for subsequent selective breeding. Therefore, it would be desirable to select genotypes present in these diverse clusters for the breeding programme. Improvement in crop yield and its associated traits constitutes the fundamental goal of any breeding program. Therefore, the evaluation of cluster diversity related to seed yield and its contributing attributes is essential for the purpose of genotype selection. (Gupta *et al.,* 2023). In the present experiment, significant variations were observed among the clusters for the majority of the traits.

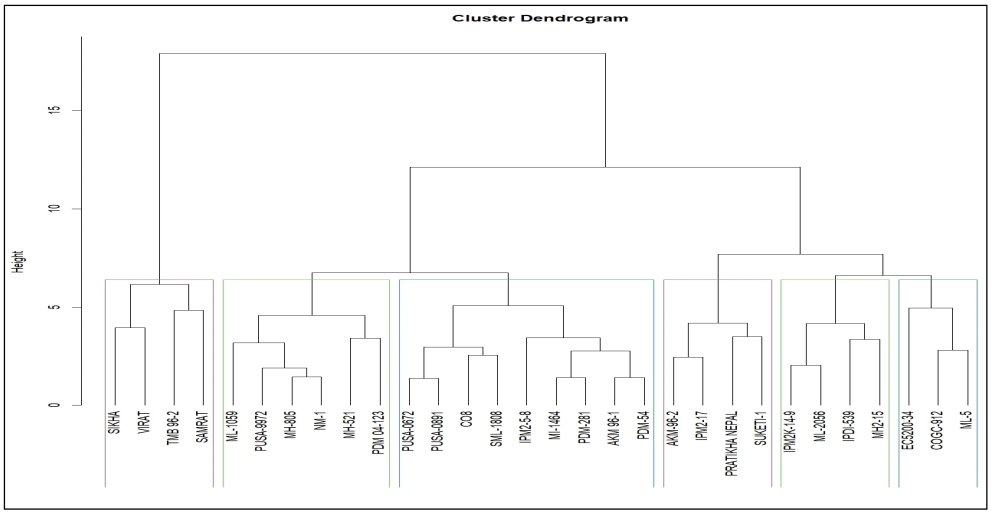


Fig 2: Dendrogram based on 11 different agronomical characters of 30 mungbean genotypes.

Table 5: Intra and inter cluster average distances among clusters of mungbean genotypes following ward method.

Cluster no.	I	II	III	IV	V	VI
I	2.625	4.232	4.001	3.563	3.317	6.239
II		3.381	4.718	4.317	4.932	5.286
III			3.883	4.419	4.580	7.263
IV				3.192	4.761	6.087
V					2.912	6.750
VI						4.985

Table 6: Mean values of seed yield and its components in various clusters.

Cluster no.	DTF	DTM	PH	NPB	BPP	NCP	NPP	LOP	NSP	HSW	SYP
I	44.050	67.815	55.305	2.212	8.278	6.280	21.023	7.533	10.633	3.194	6.317
II	43.250	66.708	58.589	3.203	6.481	6.999	24.962	8.010	11.648	3.883	8.277
III	39.799	67.500	49.490	1.917	6.684	6.097	20.875	7.698	12.604	3.014	6.640
IV	38.125	64.958	55.378	2.491	7.857	5.825	22.160	7.177	10.721	3.603	7.912
V	47.973	70.889	57.147	1.862	7.137	6.479	20.112	7.224	10.536	3.188	6.223
VI	42.249	68.458	66.617	3.706	8.683	8.222	33.884	7.360	11.583	4.121	11.936

The mean values of seed yield and its components in various clusters are shown in Table 6. Cluster I revealed the low to moderate range of mean values for most of the characters. DTF (44.05), NPB (2.21), NCP (6.28), NPP (21.02), LOP (7.53 cm), NSP (10.63), HSW (3.19 g) and SYP (6.31g) were recorded minimum. Cluster II had the lowest mean values for BPP (6.48) and highest mean values for LOP. (8.01 cm). Cluster III had lowest mean values for most of the characters. This cluster possessed the accessions showed less DTF, least mean values for PH (49.49 cm), NPB (1.91) and HSW (3.01 g). It had highest mean values for NSP (12.60). Cluster IV showedlowest mean values for DTM (64.95), NCP (5.82) and LOP (7.17 cm). Cluster V showed highest mean values for DTF (47.93), DTM (70.88) and lowest mean values for NPP (20.11) and NSP (10.53). Among the clusters, the Cluster VI had highest mean values for most of the characters like PH (66.61 cm), NPB (3.70), BPP (8.68), NCP (8.22), NPP (33.88), HSW (4.12 g) and SYP (11.93 g). From the above comparison, cluster II, IV and VI had better cluster mean for SYP and its attributing characters. This will be very useful for future plant breeding programme and improvement of a new variety. Fig 3 revealed the percentage of contribution of each character towards the total diversity. The minimum contribution (<5%) was recorded in NSP (1.4%), NCP (2.5%), DTM (3.2%) and LOP (3.9%). The maximum

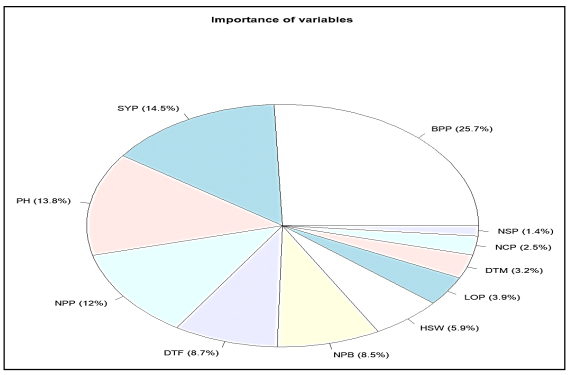


Fig 3: Contribution of agronomical characters in total variation.

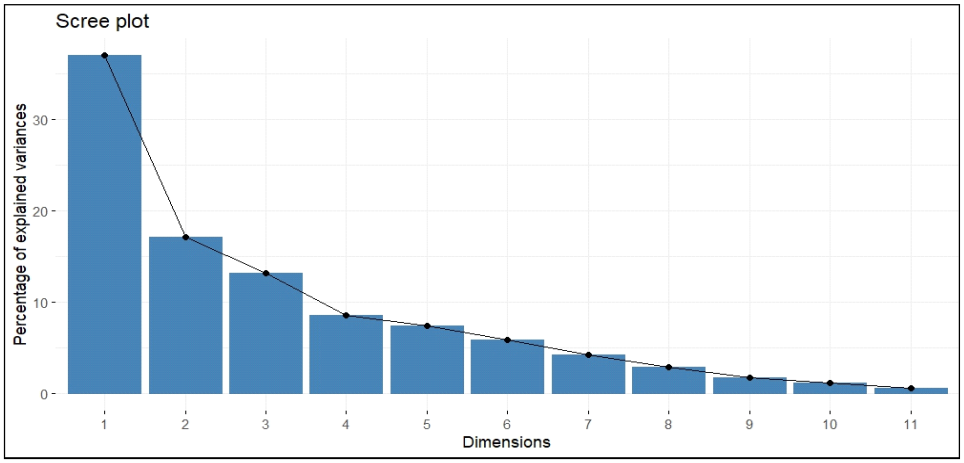


Fig 4: Scree plot showing 11 principal components.

Table 7: Eigen values, percentage of variance and cumulative variance of first 5 principal components of mung bean.

Principal components	PC1	PC2	PC3	PC4	PC5
Eigen value	4.076	1.890	1.453	0.940	0.817
Percentage of variance	37.050	17.180	13.210	8.542	7.426
Cumulative percentage	37.050	54.240	67.450	75.988	83.414
Individual characters PC score					
DTF	-0.152	0.380	0.542	-0.138	0.259
DTM	-0.146	0.526	0.362	0.138	-0.147
PH	0.269	0.183	0.157	-0.691	-0.236
NPB	0.422	0.075	-0.019	0.297	0.178
BPP	0.089	0.288	-0.375	-0.396	0.664
NCP	0.333	0.217	0.247	0.324	0.038
NPP	0.426	0.106	-0.117	-0.101	-0.390
LOP	0.114	-0.473	0.383	-0.111	0.369
NSP	0.180	-0.408	0.431	-0.143	0.005
HSW	0.396	0.094	-0.004	0.286	0.282
SYP	0.450	0.000	-0.039	-0.098	-0.119

contribution for the total divergence was recorded from BPP (25.7%), SYP (14.5%), PH (13.8%) and NPP (12%). Similar results were the findings of Ajay *et al.* (2012), Gokulakrishnan *et al.* (2012) and Jadhav *et al.* (2023).

In multivariate statistics, principal components analysis (PCA) possesses the capability to convert a set of potentially correlated variables into a reduced set of uncorrelated variables referred to as principal components (PC). Eigenvalues are commonly employed to ascertain the number of factors to retain. If the eigen value is less than one, it indicates that the explanatory efficacy of the principal components is inferior to the average explanatory efficacy of the original variables (Jadhav *et al*., 2023). PCA, only the pertinent principal components were extracted from the 11 characters. The identification of key contributors to variance was facilitated by examining the characters with high loading values on PC1. Eigen values greater than one can serve as an inclusion criterion. Principal components with eigen values less than unity were considered non-significant. In the present investigation, out of 11 principal components first three PC *i.e.* PC_1 to PC_3 which eigen value more than one extracted from the data contributed 67.44% from the total variation (Table 7 and Fig 4). PC_1 showed 37.05% followed by PC_2 (17.18%), PC_3 (13.21%). The eigen value and variance associated with each principal component gradually decreased, while cumulative variability increased gradually. Table 7 shows the principal components results along with each character loading score. Scree plot graph depicted in Fig 4 explained the percentage of variance associated with each PCs obtained by drawing a graph between percentage of variances and principal component numbers. In PC_1, the high positive component loading from SYP(0.450), NPP (0.426), NPB (0.422), HSW (0.396) and NCP (0.333). Whereas, high negative loading components was observed in PC1 with DTF and DTM. The high positive

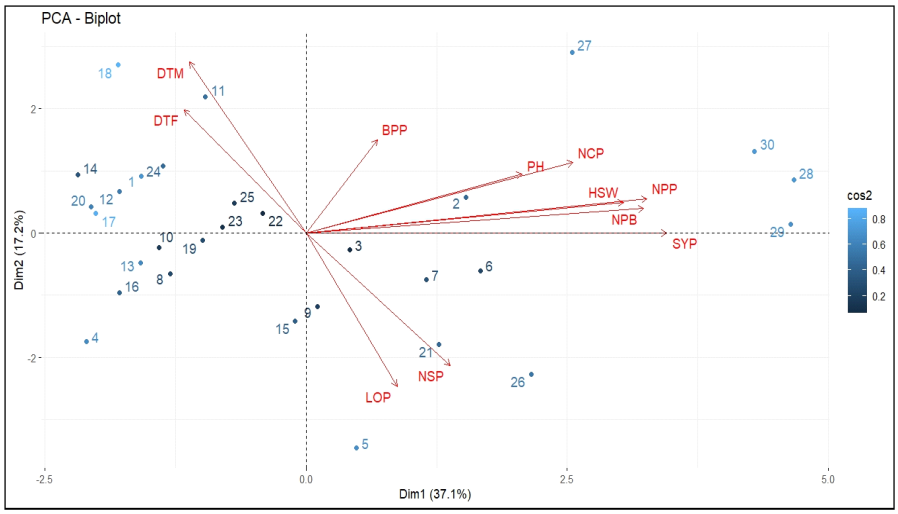


Fig 5: Biplot showing the distribution of genotypes and contribution of all the characters in first two principal components.

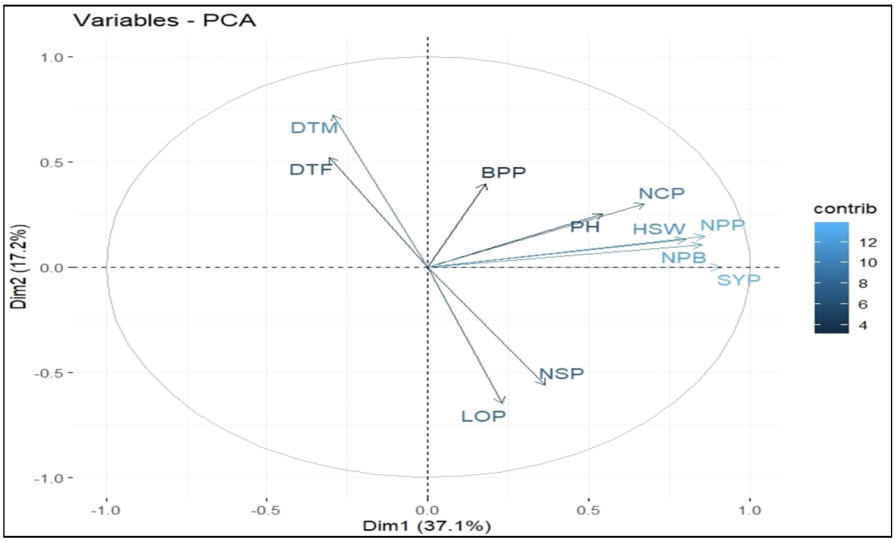


Fig 6: Loading plot showing contribution of variables towards the first two PCs.

Table 8: PC scores of 30 mung bean genotypes in each PC.

Genotypes	SYP	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
AKM 96-1	5.256	-1.549	0.902	-0.855	0.052	0.239	-0.125	-0.487	-0.275	0.010	-0.005	0.421
AKM-96-2	8.397	1.505	0.560	0.694	0.210	-1.080	-0.856	0.723	-0.233	0.121	-0.199	0.133
CO8	9.501	0.409	-0.264	-0.013	0.418	-0.277	1.009	-0.964	0.572	-0.157	-0.244	-0.080
COGC-912	4.327	-2.067	-1.720	0.163	0.126	-0.999	0.892	1.107	-0.340	-0.679	-0.311	-0.287
EC5200-34	8.796	0.476	-3.401	0.864	-0.744	0.430	1.557	0.854	0.464	0.815	-0.264	0.209
IPDI-539	10.825	1.644	-0.596	-2.440	0.681	-0.992	-0.712	-0.353	-0.210	0.982	-0.305	-0.102
IPM2-17	6.825	1.131	-0.739	1.580	0.126	0.601	-0.615	0.413	-0.043	-0.331	-0.022	0.523
IPM2-5-8	6.288	-1.281	-0.639	0.763	-1.791	-0.209	0.565	-0.731	-0.736	-0.265	-0.464	0.011
IPM2K-14-9	7.154	0.110	-1.161	-1.550	-0.441	0.613	0.008	1.033	0.414	-0.233	0.879	-0.048
MH2-15	5.986	-1.384	-0.227	-2.920	0.218	-0.571	-0.513	0.182	-0.179	0.174	0.310	0.086
MH-521	6.914	-0.949	2.154	0.021	1.537	0.174	1.152	-0.892	-0.848	0.046	-0.258	-0.410
MH-805	6.465	-1.756	0.659	0.674	0.440	-1.166	0.105	-0.189	0.221	0.379	0.613	0.004
MI-1464	4.761	-1.557	-0.466	0.189	-0.248	0.426	-0.450	-0.160	-0.851	0.034	-0.446	0.144
ML-1059	5.021	-2.150	0.919	2.959	-0.530	-0.867	-0.837	0.281	0.166	-0.046	-0.007	-0.079
ML-2056	7.682	-0.103	-1.391	-0.485	-0.099	0.670	-1.081	0.632	0.050	-0.112	0.013	-0.794
ML-5	6.797	-1.758	-0.938	0.179	1.666	-0.868	0.373	-0.155	1.282	-0.157	-0.340	-0.075
NM-1	5.030	-1.980	0.309	0.595	-0.263	-0.246	-0.127	0.192	-0.148	-0.047	0.321	-0.008
PDM 04-123	5.857	-1.772	2.658	-0.111	-1.000	-0.082	0.199	0.348	0.352	0.122	0.055	0.147
PDM-281	4.879	-0.974	-0.116	0.275	0.068	0.302	-1.138	0.723	-0.630	0.126	-0.155	0.096
PDM-54	4.586	-2.023	0.415	-1.103	-0.624	0.505	-0.248	0.070	-0.927	0.220	0.071	0.022
PRATIKSHA NEPAL	7.906	1.251	-1.758	0.195	1.343	-0.776	-0.264	-1.239	-0.396	-0.938	0.649	0.291
PUSA-0672	7.004	-0.407	0.315	-0.841	-0.705	0.732	1.199	-0.082	0.132	-0.551	0.447	0.029
PUSA-0891	6.879	-0.788	0.093	-1.118	-0.187	1.519	1.123	-0.313	0.229	-0.165	-0.266	0.057
PUSA-9972	8.049	-1.348	1.060	1.020	-0.313	-0.512	-0.579	-0.592	1.011	0.279	0.254	-0.253
SML-1808	7.702	-0.674	0.469	-0.122	0.435	1.083	-0.257	-0.402	0.868	0.737	-0.117	0.316
SUKETI-1	9.980	2.119	-2.230	1.233	0.621	1.827	-1.284	-1.149	-0.059	0.104	0.019	-0.187
TMB 96-2	8.414	2.503	2.858	-0.369	1.322	1.552	-0.427	1.110	0.511	-0.735	-0.450	0.015
SAMRAT	11.878	4.591	0.839	1.269	1.608	-0.559	1.426	0.860	-0.846	0.575	0.380	-0.002
SIKHA	13.686	4.562	0.143	-1.476	-1.562	-1.992	-0.409	-0.235	0.470	-0.512	-0.518	0.099
VIRAT	13.766	4.221	1.294	0.729	-2.367	0.524	0.314	-0.584	-0.023	0.205	0.360	-0.277

loading components signified the effectiveness of the selection process for traits associated with yield on PC1. This observation illuminates that PC1 accounted for the predominant variability in traits linked to yield. Similarly, in second principal component, DTM (0.526), DTF (0.380), BPP (0.288), NCP (0.217) and PH (0.183) were contributed major variation. High negative loading components were LOP (-0.473) and NSP (-0.408). Furthermore, the important characters, DTF (0.542), NSP (0.431), LOP (0.383), DTM (0.362) and NCP (0.247) were contributed more variation in PC3. The characters like BPP (-0.375) and it was followed by NPP (-0.117) and SYP (-0.039) showed negative loading. Principal component analysis proves advantageous for breeders in preparing the targeted breeding programs based on informed insights into the specific groups where particular traits hold greater significance. From the above loading scores, high positive loading components and the highest cluster mean for the various traits were found common and get confirmation their diversity possesses by the germplasm.

The biplot (Fig 5) and the loading plot (Fig 6) showed the SYP, NPB, NPP, HSW and DTM are far from the origin and had higher loading and great influence on the variation. Therefore, the loading plot reflecting the contributions of the characters to PC1 and PC2. Genotypes that were proximate and overlapped on the biplot exhibited similar properties, while those that were distant and remote from the origin demonstrated genetic variation. Genotypes from divergent clusters like TMB 96-2, SAMRAT, SIKHA, VIRAT, SUKETI-1, PDM 04-123, EC5200-34 and COGC-912 were scattered far apart on the plot.

The phenotypic expression of individual genotypes was explained through the principal component (PC) scores, as delineated in the Table 8. On the basis of PC scores, it is possible to suggest exact selection indices, the strength of which can be determined by the variability each principal component can explain. According to Singh and Chaudhary (1977), high values for variables within a specific genotype are represented by a correspondingly high PC score attributed to that genotype within the associated component. The highest PC scores of positive values >1.5 in each PC could be used them as selection indices.

The genotypes such as AKM 96-2, IPDI-539, PRATIKSHA NEPAL, SUKETI-1, TMB 96-2, SAMRAT, SIKHA and VIRAT had high yielders with high PC scores. These genotypes were also good performer for the other associated yield traits. The characters with high variability areemphasized by PC analysis. Therefore, rigorous selection can be designed to quickly increase yield.

CONCLUSION

From the experiment, a significant amount of genetic diversity was revealed and it was highlighted in clusters using hierarchical cluster analysis. The descriptive analysis of yield and its attributing traits had pointed out some important characters like BPP, PH, NPP and NPB as the important characters to identify and classify the diversity from the germplasm. Through PCA, it was helpful to identify traits and genotypes from similarities and differences among the genotypes. According to the finding of the research the identified diverse genotypes from the clusters like AKM 96-2, IPDI-539, PRATIKSHA NEPAL, SUKETI-1, TMB 96-2, SAMRAT, SIKHA and VIRAT were superior in yield and associated traits. Therefore, the information of this result will be very useful with these characters for future mungbean breeding program.

ACKNOWLEDGEMENT

All the authors acknowledge the Indian Institute of Pulses Research (ICAR-IIPR), Kanpur for providing the germplasm for research.

Conflict of interest

All authors declare that they have no conflict of interest.

REFERENCES

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

Bos, I. and Caligari, P. (1995). Selection Methods in Plant Breeding. Published by Chapman and Hall, 2-6 Boundary Row, London, SE1 8HN, UK.

Canci, H. and Toker, C. (2014). Yield components in mung bean [*Vigna radiata* (L.) Wilczek]. *Turkish Journal of Field Crops*. **19(2):** 258-261.

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

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

Gupta, D., Muralia, S., Gupta, N.K., Gupta, S., Jakhar, M.L. and Sandhu, J.S. (2023). Genetic diversity and principal component analysis in mungbean [*Vigna radiata* (L.) Wilczek] under rainfed condition. *Legume Research*. **46(3):** 265-272. doi: 10.18805/LR-4568.

Jadhav, R.A., Mehtre, S.P., Patil, D.K. and Gite, V.K. (2023). Multivariate analysis using D2 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.

Jeberson, M.S., Shashidhar, K.S., Wani, S.H. and Singh, A.K. (2017). Multivariate analysis in mungbean [*Vigna radiata* (L.) Wilczek] for genetic diversity under acidic soils of Manipur, India. *International Journal of Current Microbiology and Applied Sciences*. **6(7):** 760-769.

Johnson, R.A. and Wichern, D.W. (1988). Applied Multivariate Statistical Analysis. 2nd Edition, John Wiley and Sons Inc., New York.

Kaur, S., Bains, T.S. and Singh, P. (2017). Creating variability through interspecific hybridization and its utilization for genetic improvement in mungbean [*Vigna radiata* (L.) Wilczek]. *Journal of Applied and Natural Science*. **9(2):** 1101-1106.

Mahalanobis, P.C. (1936). On the Generalized Distance in Statistics. Proceedings National Academy of Science India. pp 249-255.

Nair, R., Schafleitner, R., Easdown, W., Ebert, A., Hanson, P., D'arros, H.J., Donough, H.K.J. (2014). Legume improvement program at AVRDC-The World Vegetable Center: Impact and future prospects. *Ratarstvo i povrtarstvo*. **51(1):** 55-61.

Parida, A.K. and Das, A.B. (2005). Salt tolerance and salinity effects on plants: A review. *Ecotoxicology and Environmental Safety*. **60:** 324-349.

Rao, C.R. (1952). Advance Statistical Method in Biometrical Research. John Wiley and Sons, Inc. New York.

Sarkar, M. and Kundagrami, S. (2016). Multivariate analysis in some genotypes of mungbean [*Vigna radiata* (L.) Wilczek] on the basis of agronomic traits of two consecutive growing cycles. *Legume Research.* **39:** 523-527. doi: 10.18805/lr.v0iOF.11037.

Singh, R.K., Chaudhary, B.D. (1977). Biometrical Methods in Quantitative Genetic Analysis. Kalyani Publishers, New Delhi. p 1-304.

Statista (2023). India: Production Volume of Moong Beans from Financial Year 2014 to 2023 (in million metric tons). (2023, March 1). https://www.statista.com/statistics/1140259/india-production-volume-of-moong (Accessed May 14, 2024).

Win, K.S., Win, S., Htun, T.M. and Win, N.K.K. (2020). Characterization and evaluation of mungbean [*Vigna radiata* (L.) Wilczek] germplasm through morphological and agronomic characters. *Indian Journal of Agricultural Research.* **54(3):** 308-314. doi: 10.18805/IJARe.A-520.

Yimram, T., Somta, P. and Srinives, P. (2009). Genetic variation in cultivated mungbean germplasm and its implication in breeding for high yield. *Field Crops Research.* **112(2-3):** 260-266.