AATCC Review

11 May 2025: Received
27 June 2025: Revised
07 July 2025: Accepted
08 August 2025: Available Online

https://aatcc.peerjournals.net/

Original Research Article
Open Access

Principal Component Analysis of Yield Traits in Mungbean (*Vigna radiata* L. Wilczek)

Saroj Verma[1], Deepak Gupta*[2], Sonu Jain[1], M. K. Sharma[1], Indu Bala Sethi[2] and Pushpa Lamba[2]

[1]S.K.N. College of Agriculture, Sri Karan Narendra Agriculture University, Jobner-303329, India
[2]College of Agriculture, Fatehpur-Shekhawati, Sikar Sri Karan Narendra Agriculture University, Jobner-303329, India

ABSTRACT

Mungbean (Vignaradiata L. Wilczek) is a vital pulse crop cultivated extensively across tropical and subtropical regions, notably in Asia. Despite its nutritional and agronomic value, mungbean productivity remains low due to biotic, abiotic, and genetic constraints. This study aimed to assess genetic variability and identify key yield-contributing traits among 45 mungbean genotypes using Principal Component Analysis (PCA). The experiment was conducted during Kharif 2024 at the Research Farm of SKN College of Agriculture, Jobner, using a Randomized Block Design with three replications. Thirteen quantitative traits including yield and yield components were recorded. PCA revealed that the first four principal components had eigenvalues greater than one and together explained 72.81% of the total variation. PC1 (36.58%) was associated with seed yield, number of pods per plant, pod length, clusters per plant, and plant height—traits critical for yield improvement. PC2 (15.35%) was mainly related to phenological traits like days to flowering and maturity, while PC3 (12.56%) and PC4 (8.32%) emphasized protein content, plant height, and 100-seed weight. The PCA biplot identified genotypes such as RMG 1249-1, RMG 1196, and HUM 1 as high-performing lines, suggesting their potential use in breeding programs. Overall, PCA proved effective in simplifying trait complexity and prioritizing traits and genotypes for mungbean yield enhancement.

Keywords: *Principal Component Analysis, Genetic Variability, Yield Traits, Multivariate Analysis, Genotype Evaluation, Biplot Analysis.*

Introduction

Mungbean (*Vigna radiata* L. Wilczek) is an important pulse crop extensively cultivated across tropical and subtropical regions, with significant production in Asia, Africa, and Oceania. Globally, mungbean is grown on about 6 million hectares, producing approximately 4.5 million tonnes annually, with an average productivity near 750 kg/ha (FAOSTAT, 2023). India is the largest producer of mungbean, accounting for over half of the global area and production. In India, mungbean is cultivated on around 3.5 million hectares, yielding about 1.5 million tonnes, with an average productivity of approximately 450 kg/ha (DES, 2023). Rajasthan, a leading mungbean-growing state, contributes significantly with an area of roughly 0.5 million hectares and productivity of around 600 kg/ha (Rajasthan Agriculture Department, 2023).

Despite its economic and nutritional importance, mungbean productivity remains constrained by biotic and abiotic stresses, along with genetic limitations (Singh *et al.*, 2016). Mungbean is valued for its short crop duration, high protein content, and its role in enhancing soil fertility via biological nitrogen fixation, making it crucial for food security and sustainable agriculture in cereal-based cropping systems (Nair *et al.*, 2012).

*Corresponding Author: ***Deepak Gupta***

DOI: https://doi.org/10.21276/AATCCReview.2025.13.03.589

Yield in mungbean is a complex quantitative trait influenced by multiple interrelated components such as pod number per plant, pod length, seeds per pod, 100-seed weight, and plant height. The intricate relationships among these traits complicate direct selection for yield improvement. Principal Component Analysis (PCA) is an effective multivariate technique to simplify such complexity by transforming correlated variables into uncorrelated principal components. This facilitates a better understanding of genetic variability and aids in identifying key traits contributing to yield differences (JolliffeandCadima, 2016).PCA has been widely used in legume breeding programs to analyze genetic diversity and support the selection of superior genotypes (Kumar *et al.*, 2020; Anitha*et al.*, 2022). Applying PCA to mungbean germplasm can reveal the major yield-contributing traits and streamline breeding efforts for higher productivity.This study aims to use PCA to evaluate variation in yield and related traits among diverse mungbean genotypes, with the objective of identifying principal components that capture most of the genetic variability, thereby assisting in the strategic development of improved mungbean varieties.

Material and Methods

Location, Climate, and Weather Conditions: The present study entitled "Principal Component Analysis of Yield Traits in Mungbean (*Vigna radiata* L. Wilczek)" was conducted during the Kharif season of 2024 at the Research Farm of Sri Karan Narendra College of Agriculture, Jobner, Jaipur, Rajasthan.

The experimental site is located at 450 meters above mean sea level, positioned at latitude 26°5' N and longitude 75°28' E, falling under Agro-Climatic Zone III A (Semi-arid Eastern Plain Zone) of Rajasthan. The region experiences a semi-arid climate characterized by extreme temperature variations, with summer temperatures rising up to 48°C and winter temperatures dropping below freezing point.

Experimental Material : The experimental material consisted of 45 green gram genotypes procured from the All India Coordinated Research Project on MULLaRP, Rajasthan Agricultural Research Institute, Durgapura, Jaipur, Rajasthan.

Experimental Design and Procedure : The experiment was conducted using a Randomized Block Design (RBD) with three replications during Kharif 2024. Each genotype was sown in two paired rows per plot, with each row measuring 3.0 meters in length. Row-to-row spacing was maintained at 30 cm, and plant-to-plant spacing at 10 cm. Standard agronomic practices were followed for crop management.

Characters Studied: The observations for quantitative traits were recorded on five randomly selected plants from each genotype in each replication, except for days to 50% flowering, days to maturity, and 100seed weight, which were recorded on a whole-plot basis. The traits studied included days to 50% flowering, days to maturity, plant height (cm), number of primary branches per plant, number of clusters per plant, number of pods per plant, pods per cluster, pod length (cm), number of seeds per pod, 100-seed weight (g), seed yield per plant (g), chlorophyll content at flowering stage (SPAD values), and protein content in seeds (%).

Statistical Analysis : The data collected for all quantitative traits were subjected to Principal Component Analysis (PCA) to assess the extent of genetic variability and to identify the most discriminating traits contributing to genetic divergence among the 45 green gram genotypes. PCA was performed to reduce the dimensionality of the multivariate dataset by transforming the correlated variables into a smaller number of uncorrelated variables known as principal components. These principal components were computed based on eigenvalues and eigenvectors derived from the correlation matrix. Components with eigenvalues greater than one were considered significant, and the traits with high loading values on the first few principal components were interpreted as the major contributors to the total genetic variation. The biplot and scree plot were used to visualize the relationships among genotypes and traits, and to interpret the clustering patterns. All statistical computations and graphical visualizations were performed using excel, Agri Analyze, and R software.

Results and Discussion

Principal Component Analysis (PCA) was employed to dissect the multivariate genetic variation among 45 green gram genotypes using 13 quantitative traits. The analysis revealed that the first four principal components (PCs) exhibited eigenvalues greater than 1.0, thus satisfying Kaiser's criterion and justifying their retention for further interpretation. These four PCs cumulatively accounted for 72.81% of the total phenotypic variation, indicating that they capture the majority of genetic diversity present within the studied genotypes (Table 1 and Fig. 1).

Pc1 had the highest eigenvalue (2.1806) and contributed 36.58% of the total variance. This indicates that PC1 encapsulates a substantial portion of the trait variability and is likely associated with agronomic attributes of significant breeding relevance, such as yield components, plant architecture, or phenological traits. PC2, with an eigenvalue of 1.4126, explained 15.35% of the total variance, followed by PC3 and PC4, which contributed 12.56% and 8.32%, respectively. The remaining components exhibited eigenvalues less than 1 and collectively contributed lesser proportions of the variability, hence were not considered further for interpretation.The sharp reduction in variance contribution beyond the fourth component suggests that most of the important genetic information is contained within these first few dimensions. Similar patterns have been reported in mungbean and other pulse crops where the first three to five PCs are often sufficient to describe more than 70% of the total variability (Kumar *et al*., 2020; Sharma *et al.,* 2018). The high proportion of variance explained by PC1 highlights the presence of key differentiating traits, which could be efficiently exploited in breeding programs for trait improvement and parental selection.

Table 1. Eigenvalues, percentage of variance, and cumulative variance explained by principal components in 45 green gram genotypes

Principal Component	Eigenvalues	Proportion of Variance (%)	Cumulative Variance (%)
PC1	2.1806	36.578	36.578
PC2	1.4126	15.350	51.928
PC3	1.2779	12.562	64.490
PC4	1.0401	8.321	72.811
PC5	0.9296	6.648	79.459
PC6	0.8234	5.215	84.674
PC7	0.7734	4.601	89.275
PC8	0.7108	3.886	93.161
PC9	0.6237	2.992	96.153
PC10	0.4805	1.776	97.929
PC11	0.3897	1.168	99.097
PC12	0.3359	0.868	99.965
PC13	0.0678	0.035	100.000

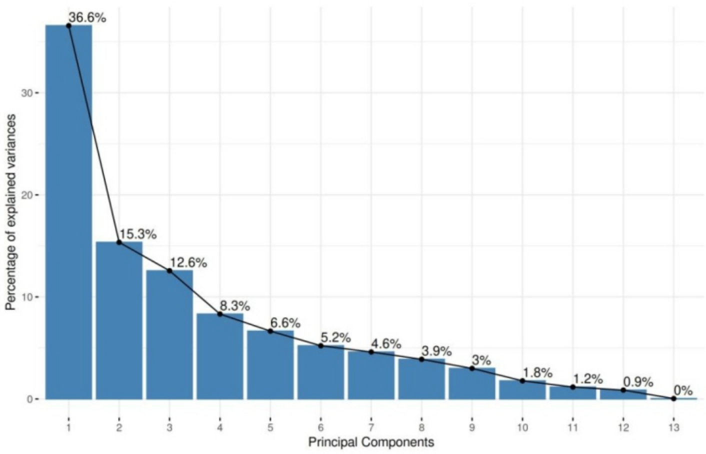

Figure 1. Scree Plot Showing Variance Explained by Principal Components in 45 Green Gram Genotypes

Principal Component Analysis (PCA) revealed meaningful associations (Table 2 and Fig. 2) between traits and the first four principal components (PCs). PC1 (36.58% variation) was positively associated with yield-related traits such as seed yield (SY: 0.3857), number of pods per plant (0.3789), pod length (0.2936), clusters per plant (0.3052), and plant height (0.299), indicating its importance in selecting high-yielding genotypes. These findings are consistent with those of Singh et al. (2016), who found yield and yield-contributing traits as major contributors to PC1 in mungbean.

PC2 (15.35%) showed strong positive loadings for days to flowering (0.5136) and days to maturity (0.5316), suggesting its role in governing phenological traits. Similar trends were observed in cowpea PCA studies by Meena*et al.* (2018), where flowering and maturity traits clustered in an independent component.

PC3 (12.56%) emphasized traits like protein content (0.3723) and plant height (0.3203), reflecting the potential for nutritional and vegetative improvement. Meanwhile, PC4 (8.32%) exhibited a high positive correlation with protein content (0.4898) and 100-seed weight (TW: 0.4338), identifying these traits as targets for enhancing seed quality and size.These results highlight the multidimensional trait contributions towards genetic divergence. Traits loading heavily on early PCs are instrumental for selection in breeding programs aimed at improving yield, earliness, and nutritional quality. PCA thus serves as a robust statistical tool for trait prioritization and genotype classification (Kumar *et al.*, 2020; Sharma *et al.,* 2017).

Table 2: Trait Loadings on Principal Components (PC1–PC4)

Traits	PC1	PC2	PC3	PC4
Days to 50% flowering	-0.129	0.5136	0.4133	0.0812
Days to maturity	-0.1217	0.5316	0.381	-0.0092
Plant height (cm)	0.299	-0.1846	0.3203	-0.0068
Primary branches per plant	0.3274	0.071	-0.0829	-0.0139
Clusters per plant	0.3052	-0.288	0.2691	-0.2331
Pods per plant	0.3789	0.1313	-0.2204	0.0406
Pods per cluster	0.1935	0.3952	-0.4471	0.2284
Pod length (cm)	0.2936	0.0524	0.2483	-0.5084
Seeds per pod	0.2977	0.206	-0.1547	0.1328
100-seed weight	0.3185	0.0972	0.1442	0.4338
Chlorophyll content (SPAD values)	0.1459	-0.2879	-0.0288	-0.4188
Protein content (%)	0.2269	-0.1392	0.3723	0.4898
Seed yield per plant (g)	0.3857	0.0149	-0.0886	0.0729

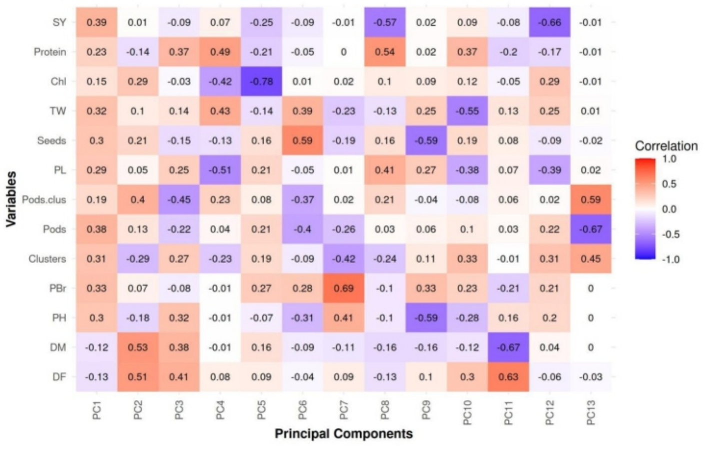

Fig. 2: Correlation Matrix of Phenotypic Traits with Principal Components

DF : Days to 50% flowering, DM : Days to maturity, PH : Plant height, Pbr : Primary branches per plant, Clusters : Clusters per plant, Pods : Pods per plant, Pods/clus : Pods per cluster, PL : Pod length, Seeds : Seeds per pod, TW : 100-seed weight, Chl : Chlorophyll content and Protein : Protein content in seed

In the PCA biplot (Figure 3), genotypes RMG 1249-1, RMG 1196, and HUM 1 were distinctly located on the positive side of PC1, demonstrating strong associations with high yield and related traits, and thus serve as ideal candidates for selection and improvement programs (Singh *et al*., 2020). Conversely, genotypes Pusa-9531 and RMG 1166, aligned along the positive PC2 axis, were characterized by early flowering and maturity traits, suggesting their potential for short-duration environments (Kumar *et al.,* 2019).

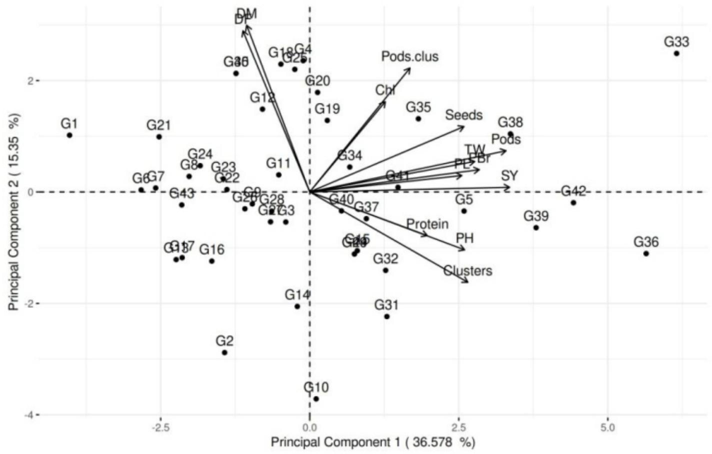

Fig. 3: Distribution of genotypes and trait contributions based on PC1 and PC2 (51.93% variance)

Figure 4 further illustrates the trait contributions to PC1, where seed yield, pods per plant, pod length, and seeds per pod exhibit strong positive loadings, indicating their synergistic influence on productivity (Meena*et al.,* 2018). These traits collectively enhance yield performance and are crucial targets for selection. Conversely, days to maturity and biomass showed negative loadings, suggesting an inverse relationship with yield traits. This reflects a trade-off where early maturity and reduced vegetative growth may favor reproductive efficiency. Such relationships are significant for selecting genotypes suitable for short-season environments (Meena*et al.,* 2018). The PCA thus aids in trait prioritization for targeted breeding programs.

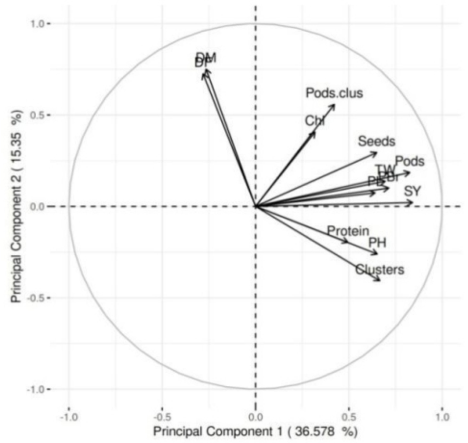

Fig. 4: PCA Biplot of green gram genotypes based on yield and morphological traits

Figure 5 demonstrates the spatial distribution of all 45 genotypes, revealing a clear separation along PC1 and PC2, which aids in distinguishing high-performing genotypes from low-performing ones based on multivariate trait expressions (Ramesh *et al.,* 2018). The acute angles between trait vectors such as seed yield, pods per plant, and test weight suggest strong positive correlations, while obtuse angles between days to maturity and yield traits indicate inverse relationships (Singh *et al.,* 2016). Genotypes near the origin of the biplot displayed minimal variation and can be considered genetically less diverse, while those dispersed toward the extremes signify high divergence useful for hybridization strategies (Kumar *et al.,* 2020).

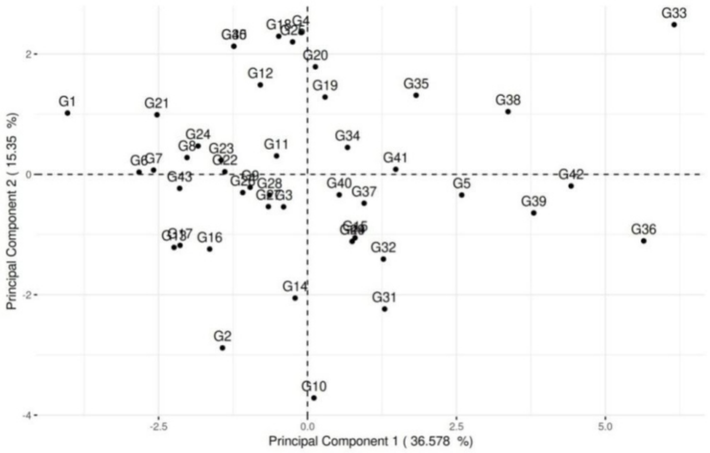

Fig. 5: PCA Scatter Plot of Genotypes Based on Principal Components 1 and 2

The breeding strategies adopted in this study, guided by multivariate analysis, enabled the identification of genetically diverse and agronomically superior green gram genotypes. The integration of PCA facilitated the prioritization of yield-related, phenological, and quality traits for effective selection. Genotypes showing favorable combinations of traits such as high seed yield, early flowering, and enhanced protein content were identified as potential parents for future hybridization programs. These findings underscore the value of strategic trait-based selection and genotype classification in strengthening genetic improvement efforts. The targeted utilization of these genotypes in breeding pipelines will enhance the development of high-yielding, early-maturing, and nutritionally enriched varieties, contributing significantly to sustainable green gram cultivation.

Conclusion

The application of Principal Component Analysis (PCA) effectively unraveled the complex genetic architecture underlying 13 quantitative traits across 45 green gram genotypes. The first four principal components, cumulatively explaining 72.81% of the total variation, highlighted key agronomic, phenological, and nutritional traits contributing to genetic diversity. Traits such as seed yield, pods per plant, and protein content emerged as primary drivers of divergence, providing valuable insights for trait-based selection. The spatial distribution and trait associations observed in the PCA biplots further facilitated the identification of promising genotypes RMG 1249-1, RMG 1196, and HUM 1 for yield enhancement, and Pusa-9531 and RMG 1166 for early maturity. This multidimensional approach not only simplifies genotype classification but also streamlines breeding decisions aimed at improving productivity, adaptability, and nutritional quality. Overall, PCA proves to be a powerful statistical tool in green gram breeding programs, enabling precise trait prioritization and informed parental selection for genetic improvement.

References

1. Anitha, V., Kumar, P. and Reddy, B. V. S. (2022). Genetic diversity analysis in mungbean (*Vigna radiata* L.) using multivariate approaches. *Journal of Pulse Research*, 35(1), 45–53.
2. DES (2023). *Agricultural Statistics at a Glance 2023*. Directorate of Economics & Statistics, Ministry of Agriculture, Government of India.
3. FAOSTAT (2023). *FAOSTAT Statistical Database*. Food & Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/faostat/en/
4. Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: a review & recent developments. *Philosophical Transactions of the Royal Society A*, 374(2065), 20150202.
5. Kumar, A., Verma, S., Yadav, R. K. and Singh, D. (2020). Genetic divergence & principal component analysis in mungbean (*Vigna radiata* L. Wilczek) for yield & its contributing traits. *Legume Research*, 43(3), 391–396.
6. Kumar, P., Yadav, R. K. and Soni, D. K. (2019). Multivariate analysis in green gram (*Vigna radiata* L.) for yield & maturity traits. *Journal of Pharmacognosy & Phytochemistry*, 8(3), 1188–1191.
7. Kumar, S., Singh, A. and Sharma, M. (2020). Application of principal component analysis in mungbean improvement programs. *Plant Breeding*, 139(5), 721–730.
8. Meena, R. P., Kumawat, S. M., Singh, P., Meena, B. L. and Meena, A. K. (2018). Genetic diversity & principal component analysis in cowpea [*Vigna unguiculata* (L.) Walp.]. *International Journal of Current Microbiology & Applied Sciences*, 7(1), 2360–2370.
9. Nair, R. M., Basu, P. S. and Singh, A. K. (2012). Breeding for improved mungbean productivity. *Indian Journal of Genetics & Plant Breeding*, 72(1), 1–7.
10. Rajasthan Agriculture Department (2023). *Annual Agricultural Statistics Report*. Government of Rajasthan.
11. Ramesh, R., Sundaram, S. and Rajanna, M. P. (2018). Principal component analysis & genetic diversity in green gram (*Vigna radiata* L. Wilczek). *Electronic Journal of Plant Breeding*, 9(4), 1292–1299.
12. Sharma, P., Kumar, A. and Yadav, S. K. (2018). Multivariate analysis for yield & its contributing traits in mungbean (*Vigna radiata* L. Wilczek). *Journal of Food Legumes*, 31(1), 67–70.
13. Sharma, S. K., Kumar, R. and Singh, S. K. (2017). Assessment of genetic diversity in mungbean [*Vigna radiata* (L.) Wilczek] through multivariate analysis. *Journal of Pharmacognosy & Phytochemistry*, 6(5), 1317–1321.
14. Singh, A., Singh, B. and Singh, C. (2016). Genetic divergence & principal component analysis in mungbean (*Vigna radiata* L.). *Indian Journal of Agricultural Sciences*, 86(6), 748–751.
15. Singh, D., Kumar, V. and Meena, M. C. (2020). Assessment of genetic diversity in green gram [*Vigna radiata* (L.) Wilczek] using principal component & cluster analysis. *Legume Research*, 43(1), 124–128.
16. Singh, S., Singh, B. D. and Kumar, A. (2016). Mungbean: Nutritional quality & productivity improvement. *Legume Research*, 39(5), 797–805.

