PLANT SCIENCE TODAY
Vol 12(4): 1-9
https://doi.org/10.14719/pst.8730

eISSN 2348-1900

RESEARCH ARTICLE

Morphological descriptor-based multivariate approaches for comprehensive genetic diversity analysis in soybean (*Glycine max* L. Merrill)

Kumar Jai Anand*, M K Shrivastava, Pawan K Amrate, Teena Patel, Yogendra Singh, Tasphiya Elahi, Vijay Kumar Katara & Shailendra Sagar Prajapati

Department of Plant Breeding and Genetics, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur 482 004, Madhya Pradesh, India

*Correspondence email - kumarjaianand@jnkvv.org

Received: 07 April 2025; Accepted: 05 October 2025; Available online: Version 1.0: 10 November 2025; Version 2.0: 21 November 2025

Cite this article: Kumar JA, Shrivastava MK, Pawan KA, Teena P, Yogendra S, Tasphiya E, Vijay KK, Shailendra SP. Morphological descriptor-based multivariate approaches for comprehensive genetic diversity analysis in soybean (*Glycine max* L. Merrill). Plant Science Today. 2025; 12(4): 1-9. https://doi.org/10.14719/pst.8730

Abstract

Assessment of genetic diversity in germplasm collections is crucial for effective genetic improvement in soybean (*Glycine max* L. Merrill). The present study evaluated 165 soybean germplasm lines, including five checks, at JNKVV, Jabalpur, during *Kharif* 2022, *Rabi*-Summer 2023 and *Kharif* 2023 to characterize morphological variability based on the DUS guidelines. A thorough statistical analysis involving phylogenetic assessment, hierarchical clustering and principal component analysis (PCA) was accomplished to elucidate the genetic divergence and trait-based associations among the genotypes. Phylogenetic analysis delineated the germplasm into six distinct groups, indicating substantial genetic differentiation. Genotypes NRC 181, VLS 89, JS 20-34 and NRC 138 formed an independent cluster, suggesting unique allelic variation. Hierarchical clustering established this genetic stratification, identifying distinct sub-clusters associated with differential morphological traits. PCA revealed five principal components explaining 55.85 % of the total phenotypic variation, with PC1 contributing 17.3 %, predominantly associated with flower colour and hypocotyl anthocyanin pigmentation. Genotypes such as EC 313915, GW 89, CAT 489A and GW 108 exhibited stable performance and significant contributions to multiple trait expressions. The PCA-biplot, based on the first two principal components and overlaid with six clusters, effectively distinguished genotype groups, confirming the robustness of classification. The Mantel test showed a strong correlation ($r = 0.6161$, $p = 0.001$) between distance matrices from PCA and hierarchical clustering, validating the consistency of the analysis. The investigation underscores the genetic variability within the evaluated soybean germplasm. It reinforces the importance of integrating morphological characterization with multivariate statistical approaches for strategic genotype selection in soybean improvement programs.

Keywords: genetic diversity; hierarchical clustering; Mantel test; morphological characterization; PCA; soybean

Introduction

Soybean (*Glycine max* (L.) Merrill), a member of the family *Fabaceae* and subfamily *Papilionoideae*, is one of the most economically and nutritionally significant legume crops globally (1, 2). Often referred to as the “golden bean”, as this is the essential source of dietary protein (40-42 %), oil (18-22 %), carbohydrates (30 %) and vital micronutrients, making it a cornerstone of food and feed systems worldwide (3-5). With the global population projected to exceed 9.7 billion by 2050, ensuring protein security is increasingly critical (6). Meeting the rising demand for sustainable protein sources is challenging, given limited land and water resources and the impacts of climate change (7, 8). Plant-based proteins have gained popularity and soybean stand out among legumes due to their high protein quality and balanced amino acid profile, making them essential for addressing the global protein gap while contributing to sustainable agriculture (9-11).

Global soybean cultivation covers over 140.01 million hectares, yielding 394.96 million metric tons, resulting in an average productivity of 2820 kg/ha. In India, the area allocated for soybean cultivation during the 2022-23 period amounted to 12.10 million hectares, significantly contributing to global production with 12.56 million metric tons (12). Soybean productivity in India is notably lower than the global average. This can be attributed to several factors, including the lack of improved high-yielding varieties, poor-quality seeds and a narrow genetic base in available germplasm. These constraints can be addressed by understanding the genetic diversity within the species, which is crucial to expanding the genetic base of cultivars, exploiting heterosis and generating productive recombinants (13-15). Employing methodologies like phylogenetic trees, hierarchical clustering and principal component analysis (PCA) serves to effectively assess genetic diversity among soybean germplasm lines based on qualitative morphological traits, which remain minimally influenced by environmental factors (16, 17). The phylogenetic tree illustrates evolutionary relationships among soybean genotypes based on qualitative morphological descriptors. Closely related genotypes are placed on shorter branches, while genetically distant

genotypes are positioned on longer branches. This structure identifies genotypic divergence based on their ancestral lineage, facilitating the development of superior soybean varieties suited to diverse agro-climatic conditions (16, 18). Cluster analysis is essential for classifying soybean genotypes based on genetic similarity and diversity (19). It groups accessions into distinct clusters, allowing the identification of genetically distant and closely related genotypes. This classification aids in selecting diverse breeding materials, optimizing hybridization strategies and maintaining genetic variability within breeding populations (20).

PCA is a powerful dimensionality reduction technique that enables researchers to visualize genetic diversity by transforming complex, high-dimensional trait data into a lower-dimensional space while preserving maximum variation. In plant breeding, PCA is widely used for germplasm characterization and selection by identifying significant sources of variation among genotypes (21). However, PCA alone does not classify, or group genotypes based on similarity. To enhance biological interpret-tation, overlaying cluster information on a PCA plot provides a more structured understanding of genetic relation-ships. This integration combines the strengths of PCA with hierarchical clustering, allowing clear differentiation between genetically similar and distinct groups (22). The Mantel test assessed the correlation between distance matrices derived from different analyses, providing insight into the relationship between different analytical approaches (23). Considering these deliberations, this investigation explored genetic divergence in soybeans by integrating phylogenetic assessment, cluster analysis and PCA, facilitating the identification of diverse lines for adequate use in breeding programs.

Materials and Methods

Experimental site and plant material

The field investigation was conducted during *Kharif* 2022 (July-October), *Rabi-Summer* 2023 (January-May) and *Kharif* 2023 (July-October) at JNKVV, Jabalpur, Madhya Pradesh. This study meticulously evaluated 165 diverse lines of soybean germplasm, including five reference checks: JS 20-34, JS 20-69, JS 20-98, JS 20-116 and NRC 138. The experimental materials were sourced from the All India Coordinated Research Project (AICRP) on Soybean within the Department of Genetics and Plant Breeding at JNKVV Jabalpur.

Experimental design and agronomic management

Utilizing an Augmented Complete Block Design (ACBD), each genotype was planted in two rows measuring 3 m, with a spacing of 40 cm between rows and 5 cm between plants. The experimental site featured uniform landscape medium black soil with a pH of 7.5 and was free from waterlogging. Any excess plants were thinned out 10 days after seeding to maintain appropriate spacing. Throughout the growing period, all recommended agricultural practices were strictly followed to ensure optimal care and nourishment of the plants.

Observations recorded

Observations were recorded for various traits, including hypocotyl anthocyanin pigmentation, plant growth type, days to 50 % flowering, leaf shape, leaf colour, plant growth habit, flower colour, plant height (cm), pod pubescence, pod pubescence colour, pod colour, days to maturity, seed size (100 seeds weight), seed shape, seed colour, seed hilum colour and seed cotyledon colour. These observations were made as per the DUS guideline given by the Protection of Plant Varieties and Farmers' Rights Authority (PPV & FRA), Govt. of India, New Delhi. Days to 50 % flowering, plant height (cm), days to maturity, seed size (100 seeds weight) were recorded only in *Kharif* 2022 and *Kharif* 2023, while the rest of the traits were recorded in all three seasons.

Statistical analysis

Morphological characterization-based hierarchical clustering and phylogeny tree were created using R Studio version 4.2.2. Trait columns were converted to factors, encoded numerically and standardized. A dissimilarity matrix was computed using Euclidean distance, followed by hierarchical clustering using Ward's D^2 method. A Minimum Spanning Tree (MST) was constructed using the igraph package to visualize phylogenetic relationships. Genotypes were colour-coded based on cluster membership to enhance interpretation. Hierarchical clustering was performed using ggtree, ggplot2, ggdendro, ape and NbClust, package. PCA analysis was performed as proposed in the previous studies (24, 25) with the packages Factoextra and FactomineR. The Mantel test employed Pearson's correlation in the vegan package in R to estimate the correlation between distance matrices and assess the structural conformity among PCA-derived and hierarchical clustering distances.

Results and Discussion

Phylogenetic analysis

The phylogenetic analysis utilizing morphological descriptors classified soybean genotypes into six groups, each illustrated by distinct colours, thereby displaying considerable genetic divergence among specific accessions (Fig. 1). Notably, the genotypes NRC 181, VLS 89, JS 20-34 and NRC 138 were grouped within the red cluster, indicating a substantial divergence from the primary germplasm pool and underscoring their unique genetic characteristics. Genotypes GW 108, GW 89, CAT 489A and EC 313915 constitute an independent subgroup shown in dark green colour. They exhibit considerable divergence, suggesting the presence of rare phenotypic traits. Additionally, the genotypes DCB 137, JS 20-98, AMS 77, NRC 137, JS 22-44, RVSM 2012-19, JS 22-41, JS 22-04, JS 21-75, JSM 283, NRC 125, PS 1589, RVSM 2012-4, RVS 13-15, RVS 13-20, SL 955 and SL 1213, located at the terminal branches, also represent a distinct genetic lineage.

This analysis highlights the potential of these genotypes as valuable genetic resources for widening the genetic base and enhancing soybean breeding programs. Thirty chickpea genotypes were explored through the construction of a phylogenetic tree that revealed six distinct groups and highlighted their evolutionary relationships (16). Furthermore, the importance of phylogenetic clustering in recognizing genetic lineages, underlining its significance in identifying distinct genotypes (26).

Cluster analysis

The hierarchical clustering further validated the phylogenetic assessment, representing six distinct clusters and delivering

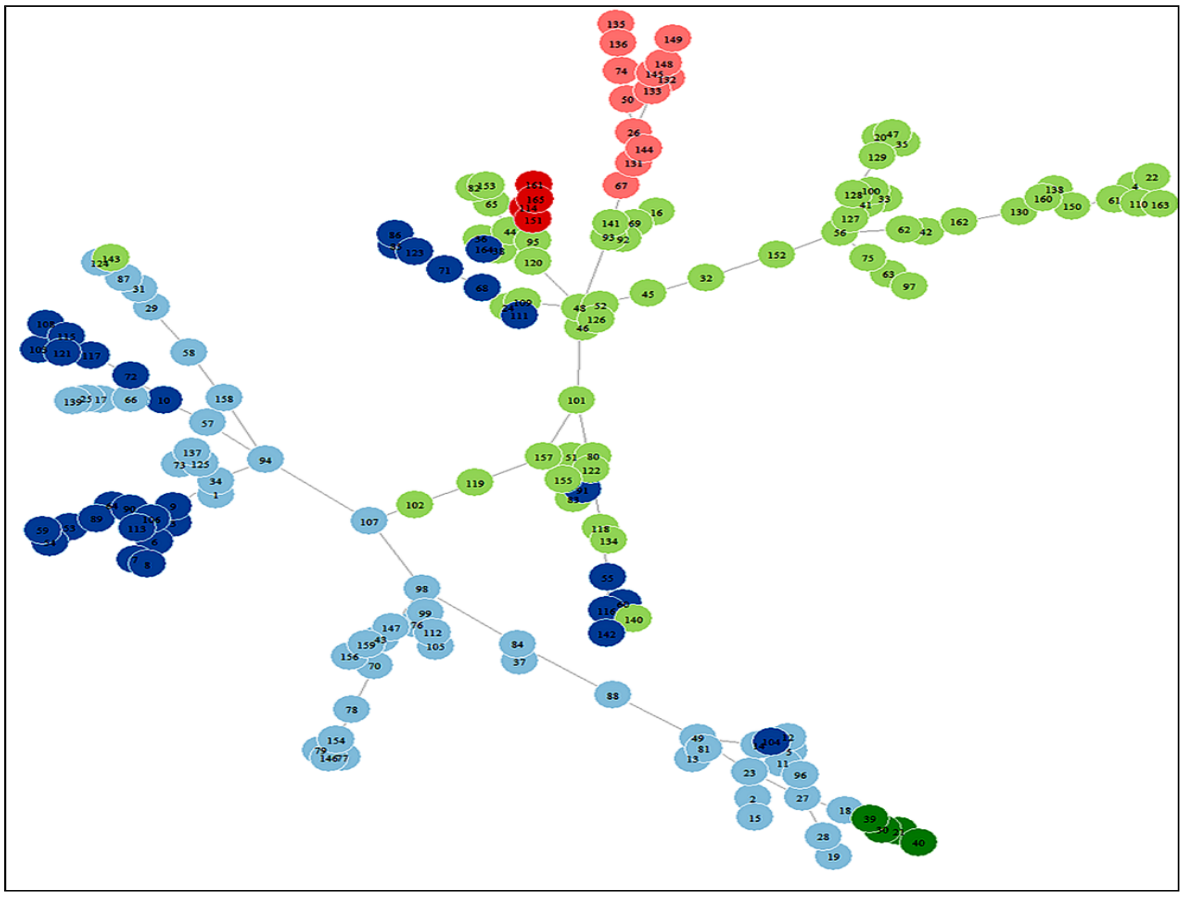

Fig. 1. Phylogenetic relationships of soybean genotypes using morphological traits.

valuable insights into genetic diversity, heterotic grouping and trait-based selection. In the circular dendrogram (Fig. 2), Cluster 6 (comprising NRC 138, VLS 89, NRC 181 and JS 20-34) and Cluster 4 (including CAT 489A, EC 313915, GW 89 and GW 108) consisted of only four genotypes, exhibiting substantial genetic divergence and unique morphological characteristics. Additionally, Cluster 5 (depicted in green) includes genotypes SL 955, SL 1213, RVS 13-15, JS 22-35, RVS 13-20 and NRC 178, while Cluster 2 (represented in orange) comprises JS 22-12, JS 23-10, JS 24-23 and MAUS 791, positioned at opposite ends of the dendrogram. The considerable genetic distance between these two clusters highlights their potential for breeding programs to enhance genetic diversity and broaden the genetic base. Strategic hybridization between genotypes from these clusters could facilitate the development of improved soybean cultivars with superior adaptive traits.

Cluster analysis was conducted using Euclidean distance, focusing on qualitative morphological traits, which resulted in the classification of 320 soybean germplasm lines into four primary clusters (27). Similarly, 120 amaranth genotypes were categorized into three basic clusters based on qualitative data from 20 traits (28). Additionally, an analysis of 28 vegetable soybean genotypes based on their morphological traits revealed eight distinct clusters (17).

Principal component analysis

The PCA analysis revealed that out of 16 principal components (PCs), only five showed more than one eigenvalue, apprehending 55.85 % of the variability among the investigated traits (Table 1). The scree plot illustrated the variance associated with each PC (Fig. 3). The first principal component (PC1), responsible for the highest variation (17.33 %) was primarily linked to flower colour and hypocotyl anthocyanin pigmentation (Fig. 4). PC2 was associated with seed-related traits, including cotyledon colour, seed colour, hilum colour and seed size (100-seed weight). PC3 was influenced by plant growth type, habit, seed shape and pod colour, whereas PC4 mainly reflected plant height and days to 50 % flowering (Fig. 5). Finally, PC5 addressed leaf shape, colour, pod pubescence and plant maturity days (Table 2 and 3).

Genotype EC 313915 exhibited the highest RC score, along with GW 89, CAT 489A and GW 108, highlighting their stability and significant contribution to multiple traits. Other promising selections included SL 955, VLS 89, NRC 181, RVS 2012-19 and RVSM 2012-11, all appearing in four RCs. Several genotypes viz., EC 313915, JS 22-53, RSC 9, SL 1213, RVS 2011-12, NRC 138, JS 20-116, NRC 178, PS 1682, DS 3124, JS 24-23, MAUS 791 and JS 20-34 were noted in at least three RCs, showcasing their genetic potential (Supplementary Table S1).

The PCA-clustering biplot further categorizes (Fig. 6) genotypes into six clusters. The red cluster (Cluster 1) focuses on seed traits, the yellow (Cluster 2) on maturity and leaf characteristics, while the green cluster (Cluster 3) correlates with plant growth habits and seed size. The blue cluster (Cluster 5) highlights pigmentation traits and the pink cluster (Cluster 6) relate to growth attributes. The isolated cyan cluster (Cluster 4) represents unique traits, underscoring the genetic diversity within the soybean genotypes evaluated. Clustering patterns show that closely positioned genotypes share similar phenotypic attributes, while those widely dispersed exhibit distinct combinations.

The upper-right quadrant features genotypes (RVS 13-20, JS 22-35, JS 21-08, PS 1670, RVS 2012-19, PS 1682, SL 1213 and SL 955) with early flowering and pigmentation traits, while the lower-right quadrant (B 1667, JS 22-53, PS 1092, JS 20-50, G 11, JS 20-98, PS 1661, PS 1660 and JS 20-78) is associated with seed

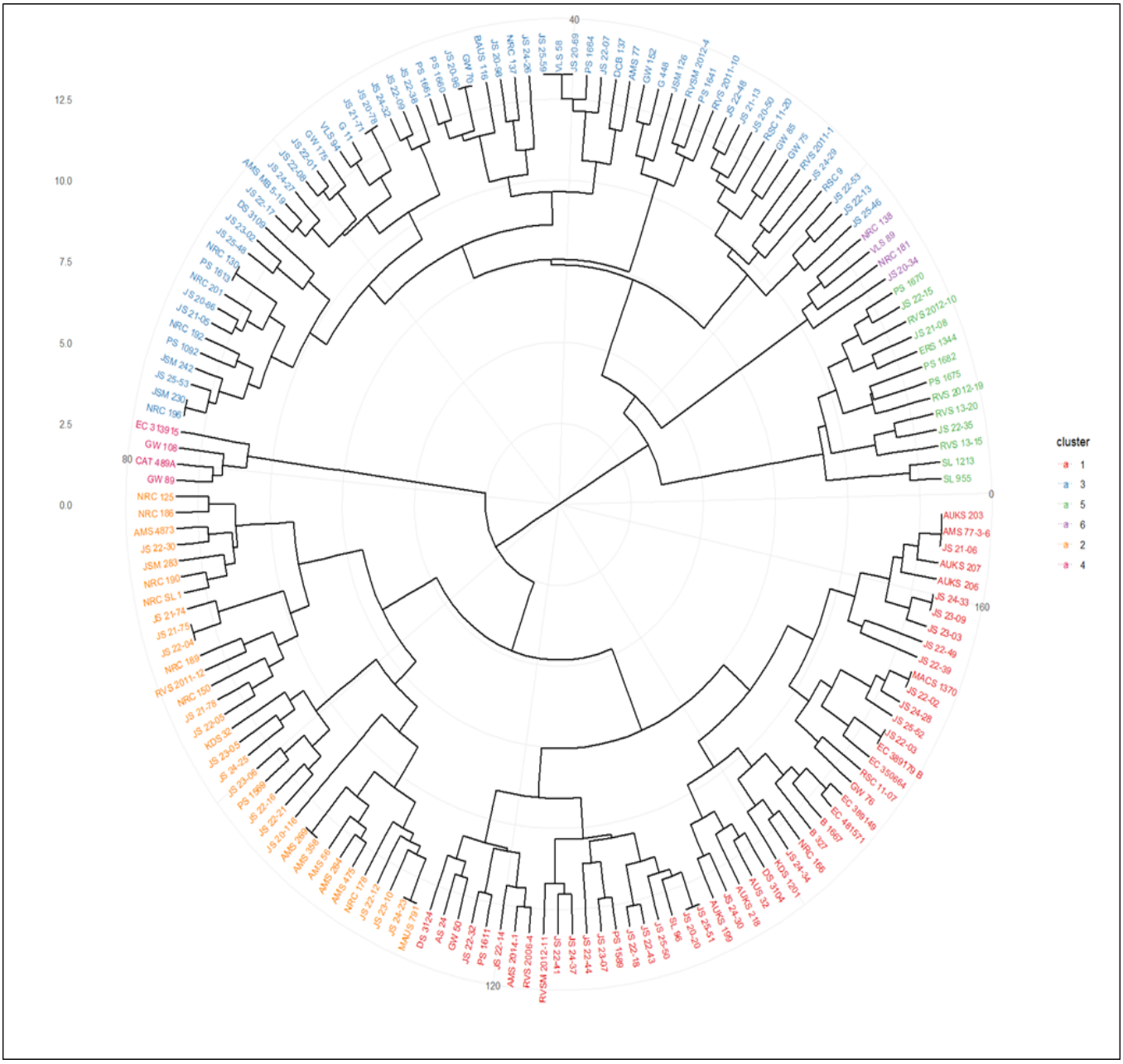

Fig. 2. Dendrogram of hierarchical clustering depicting the genetic relationships among soybean genotypes based on Ward's method with Euclidean distance. The dendrogram classifies the genotypes into six distinct clusters, with each cluster represented by a unique colour.

Table 1. Eigenvalues, variance contribution and cumulative variance in principal component analysis

Principal components	Eigen value	Percentage of variance	Cumulative percentage of variance
PC1	2.773	17.331	17.331
PC2	2.064	12.898	30.229
PC3	1.591	9.941	40.169
PC4	1.278	7.987	48.156
PC5	1.231	7.694	55.849
PC6	0.998	6.233	62.081
PC7	0.976	6.099	68.18
PC8	0.897	5.601	73.781
PC9	0.836	5.225	79.005
PC10	0.738	4.607	83.612
PC11	0.639	3.99	87.602
PC12	0.569	3.555	91.156
PC13	0.515	3.215	94.37
PC14	0.487	3.044	97.414
PC15	0.406	2.536	99.949
PC16	0.009	0.052	100

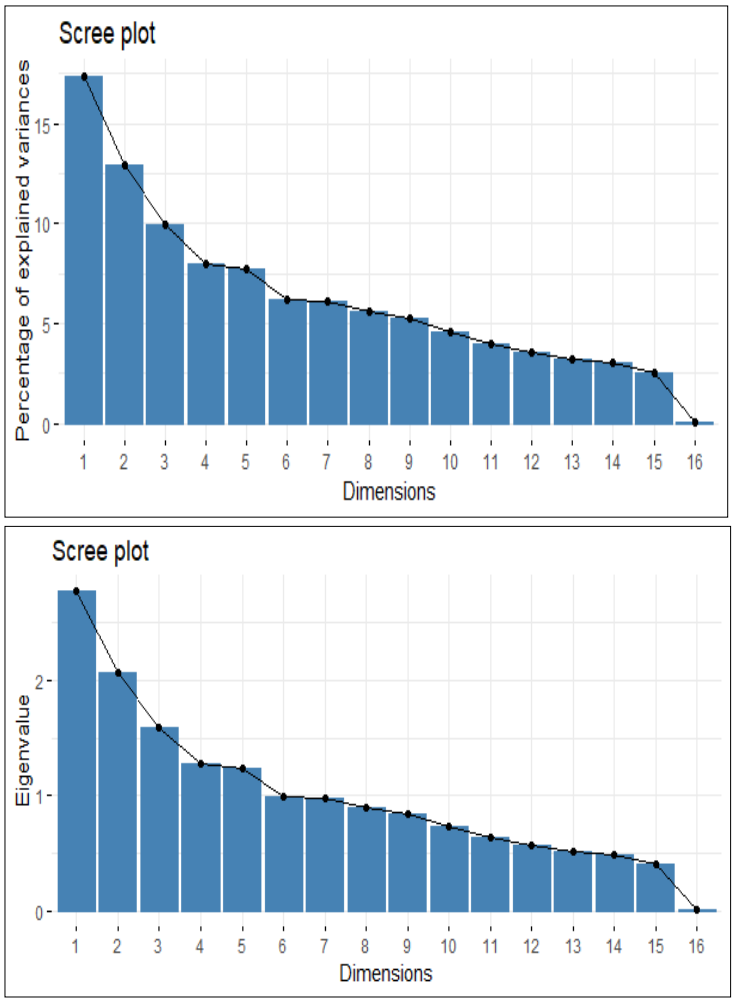

Fig. 3. Scree plot depicting principal component eigenvalues for dimensional reduction.

Fig. 4. Morphological markers used for distinctness in soybean genotypes.

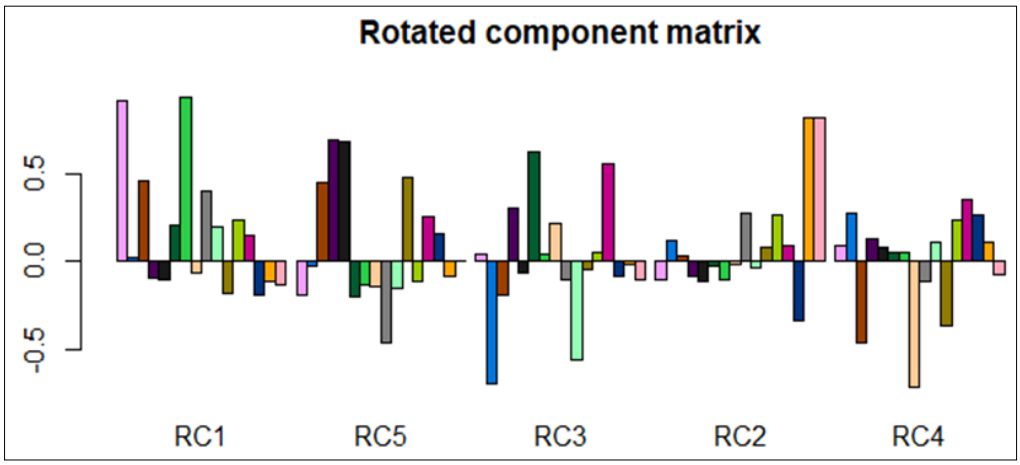

Fig. 5. Factor loadings of principal components in the rotated component matrix.

Table 2. Principal component values (>0.500) of rotation component matrix of soybean genotypes in pooled analysis

Traits	Rotated Components				
	RC1	RC2	RC3	RC4	RC5
Hypocotyl anthocyanin pigmentation	0.924	-0.106	0.045	0.090	-0.192
Plant growth Type	0.025	0.124	-0.697	0.280	-0.026
Days to 50 % flowering	0.460	0.033	-0.195	-0.460	0.450
Leaf shape	-0.092	-0.086	0.304	0.132	0.693
Leaf colour	-0.104	-0.115	-0.065	0.078	0.686
Plant growth habit	0.208	-0.027	0.629	0.050	-0.203
Flower colour	0.939	-0.107	0.041	0.049	-0.135
Plant height (cm)	-0.061	-0.013	0.221	-0.714	-0.139
Pod pubescence	0.403	0.273	-0.098	-0.116	-0.460
Pod colour	0.195	-0.033	-0.558	0.114	-0.151
Days to maturity	-0.185	0.079	-0.040	-0.365	0.485
Seed size (100-seed weight)	0.241	0.270	0.057	0.242	-0.108
Seed shape	0.150	0.090	0.557	0.356	0.257
Seed hilum colour	-0.194	-0.334	-0.079	0.269	0.159
Seed colour	-0.110	0.825	-0.013	0.112	-0.085
Seed cotyledon colour	-0.134	0.827	-0.099	-0.070	0.003

Table 3. Rotated matrix results of different traits (showed eigenvalue >0.500)

	RC1	RC2	RC3	RC4	RC5
Traits	Flower colour	Seed cotyledon colour	Plant growth type	Plant height (cm)	Leaf shape
	Hypocotyl anthocyanin pigmentation	Seed colour	Plant growth habit	Days to 50 % flowering	Leaf colour
		Seed hilum colour	Seed shape		Pod pubescence
		Seed size (100-seed weight)	Pod colour		Days to maturity

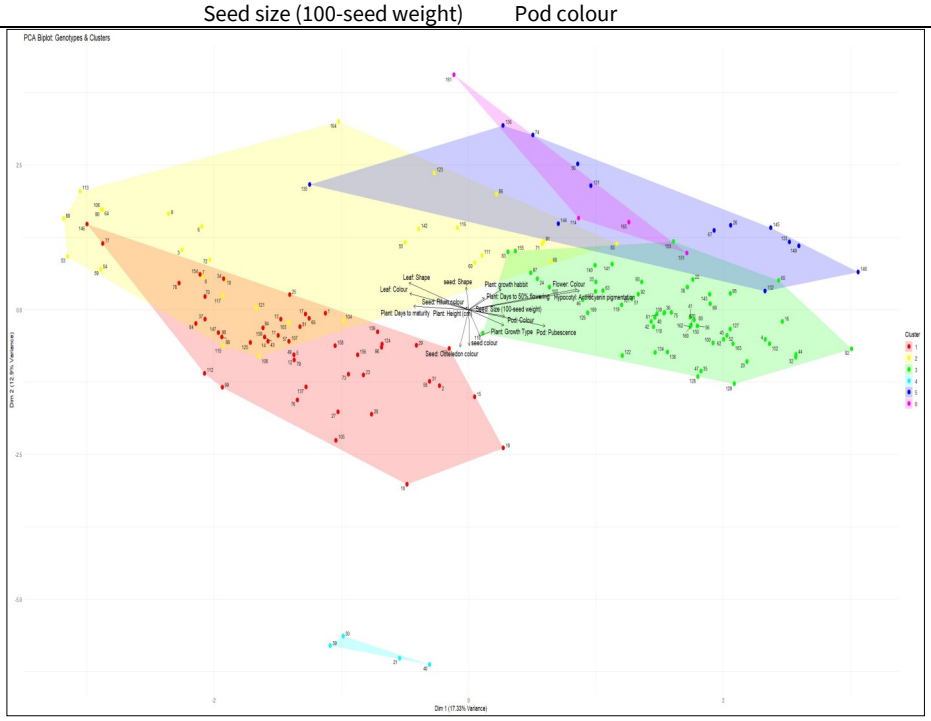

Fig. 6. Principal component analysis (PCA) biplot showing the distribution of soybean genotypes based on morphological traits. The colors indicate different clusters identified by hierarchical clustering. PCA captures the major axes of variation, explaining 29.6 % of the total variance in the first two components.

and pod characteristics. The upper-left side of the plot contains genotypes (NRC 178, JS 23-10, JS 24-23, RVSM 2012-11, JS 22-41, RVS 13-15, JS 20-116, JS 20-34 and PS 1569) with distinct growth attributes. Genotypes (B 327, CAT 489A, EC 313915, GW 89, GW 108 and KDS 1201) in the bottom-left quadrant are outliers, representing critical genetic resources for enhancing desirable traits in soybean breeding programs and should be prioritized for further evaluation and selection.

A total of 20 M6 mutants of the rice cultivar 'Improved White Ponni', along with one control of Improved White Ponni, were analyzed using morphological data. A biplot was created based on the first two principal components, which were overlaid with three clusters identified through hierarchical clustering (29). PCA was conducted on 20 qualitative traits across 320 soybean germplasm lines, revealing seven principal components (PC1 to PC7) with eigenvalues exceeding one. Collectively, these components accounted for approximately 59.10 % of the total variation, demonstrating their effectiveness in summarizing the original 20 variables (27). In another study, six principal components were identified through eigenvalue analysis and scree plots in the PCA of 19 qualitative traits across 120 amaranth genotypes, explaining nearly 72 % of the total variance (28). Additionally, PCA on cultivated yam bean (*Pachyrhizus* spp.) indicated that the first ten principal components accounted for 90 % of the overall variation (30). For 45 soybean genotypes based on eleven morphological traits, the first four principal components explained 77.25 % of the variation (31).

Mantel test-based correlation analysis of PCA and hierarchical clustering

The Mantel test revealed a significant correlation (r = 0.6161, p = 0.001) between the distances derived from PCA and those obtained from hierarchical clustering (Fig. 7). This suggests a moderate to strong relationship between the two distance matrices. This indicates that the clustering structure established through PCA is primarily preserved in hierarchical clustering, reinforcing these methods' consistency in capturing genotype similarities. The statistical significance (p = 0.001) confirms that this correlation is not due to random chance. Additionally, the results from the permutation test indicate that the null distribution's 99 % quantile is 0.0325, which is considerably lower than the observed correlation coefficient of 0.6161. This strong association validates the robustness of the clustering structure and emphasizes the reliability of the identified diverse genotypes for breeding applications. In their review, "Mantel Test in Population Genetics", the authors emphasized the significance of the Mantel test in assessing associations between two matrices (23). Furthermore, a substantial similarity was observed between the results of cluster analysis and principal component analysis (30).

Conclusion

This study employs qualitative morphological traits for genetic divergence assessment, revealing substantial genetic variation through PCA, hierarchical clustering and phylogenetic analysis. The genotypes EC 313915, GW 89, CAT 489A, GW 108, SL 955, SL 1213, VLS 89, NRC 181, RVS 13-15, RVS 13-20, NRC 138 and JS 20-34 showed significant genetic divergence across all analyses making them valuable candidates for hybridization with the potential to generate transgressive segregants for key agronomic traits in future breeding programme. Furthermore, the Mantel test established a significant correlation (0.6161) between the distances derived from PCA and those from hierarchical clustering, thereby reinforcing the reliability of these methods in capturing genetic relationships.

Acknowledgements

I sincerely acknowledge the Department of Genetics and Plant Breeding, College of Agriculture, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur, India Madhya Pradesh that have provided research materials, technical and financial support for the research work.

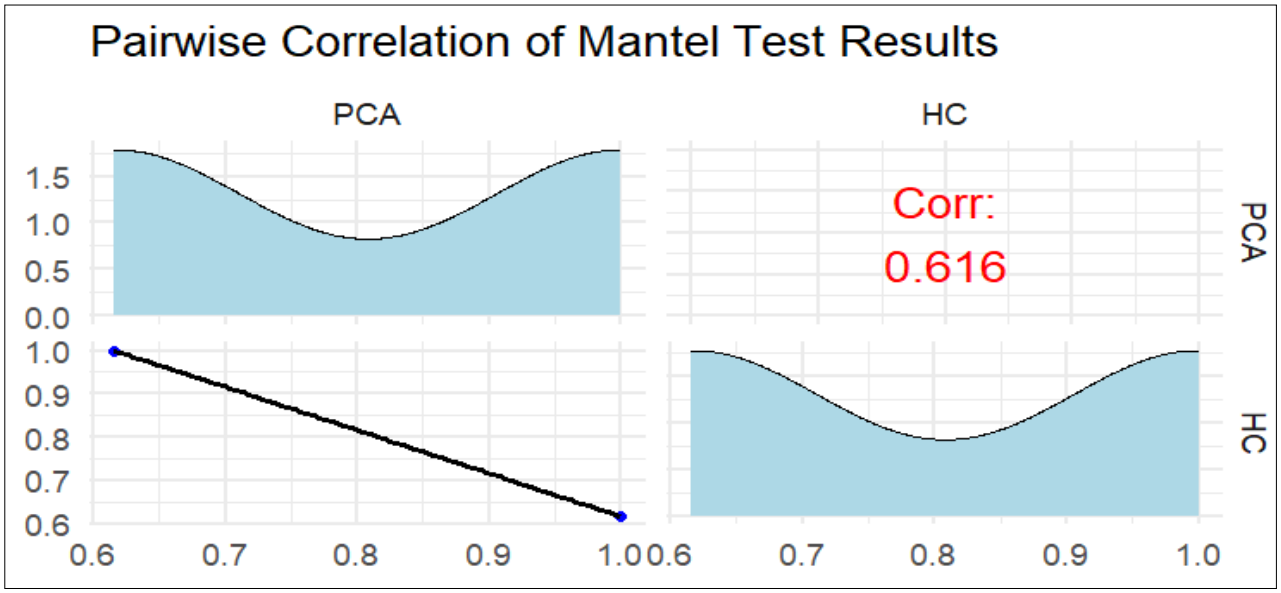

Fig. 7. Pairwise correlation between PCA and HC matrices using Mantel test (This figure illustrates the Mantel test correlation between distance matrices obtained from PCA and hierarchical clustering for the evaluated soybean germplasm. The lower-left scatter plot displays the relationship between the two distance matrices, showing a negative trend. The upper-left and lower-right density plots represent the distribution of distances within PCA and HC, respectively. The upper-right panel highlights the computed Mantel correlation coefficient (r = 0.616, p = 0.001) in red, indicating a moderate positive correlation between the two approaches despite the inverse trend in the scatter plot. This suggests that PCA and HC capture genetic divergence in different ways, reinforcing the robustness of genotype classification).

Authors' contributions

KJA drafted the manuscript. MKS provided materials, guidance and research planning. PKA, YS and TE contributed to the final drafting of the manuscript. TP and SSP performed the statistical analysis. VKK carried out grammar checking and proofreading. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interest to declare.

Ethical issues: None

References

1. Amrate PK. Survey and present status of soybean diseases in central India. International Journal of Bioresource and Stress Management. 2024;15(5):1-10. https://doi.org/10.23910/1.2024.5299
2. Mehra SH, Shrivastava MK, Amrate PK, Yadav RB. Studies on variability, correlation coefficient and path analysis for yield associated traits in soybean (*Glycine max* (L.) Merrill). Journal of Oilseeds Research. 2020;37(1):56-9. https://doi.org/10.56739/jor.v37i1.136390
3. Agarwal DK, Billore SD, Sharma AN, Dupare BU, Srivastava SK. Soybean: introduction, improvement and utilization in India—problems and prospects. Agricultural Research. 2013;2(4):293-300. https://doi.org/10.1007/s40003-013-0088-0
4. Uikey S, Sharma S, Amrate PK, Shrivastava MK. Identification of rich oil-protein and disease resistance genotypes in soybean (*Glycine max* (L.) Merrill). International Journal of Bioresource and Stress Management. 2022;13(5):497-506. https://doi.org/10.23910/1.2022.2478
5. Banerjee J, Shrivastava MK, Singh Y, Amrate PK. Estimation of genetic divergence and proximate composition in advanced breeding lines of soybean (*Glycine max* (L.) Merrill). Environment and Ecology. 2023;41(3C):1960-8. https://doi.org/10.60151/envec/VYWE5744
6. Malila Y, Owolabi IO, Chotanaphuti T, Sakdibhornssup N, Elliott CT, Visessanguan W, et al. Current challenges of alternative proteins as future foods. NPJ Science of Food. 2024;8(1):53. https://doi.org/10.1038/s41538-024-00291-w
7. Aschemann-Witzel J, Gantriis RF, Fraga P, Perez-Cueto FJ. Plant-based food and protein trend from a business perspective: markets, consumers and the challenges and opportunities in the future. Critical Reviews in Food Science and Nutrition. 2021;61(18):3119-28. https://doi.org/10.1080/10408398.2020.1793730
8. Porritt J, McCarthy M. The global protein challenge. Stockholm Resilience Centre; 2015.
9. Richter CK, Skulas-Ray AC, Champagne CM, Kris-Etherton PM. Plant protein and animal proteins: do they differentially affect cardiovascular disease risk? Advances in Nutrition. 2015;6(6):712-28. https://doi.org/10.3945/an.115.009654
10. Qin P, Wang T, Luo Y. A review on plant-based proteins from soybean: health benefits and soy product development. Journal of Agriculture and Food Research. 2022;7:100265. https://doi.org/10.1016/j.jafr.2021.100265
11. Messina M. Perspective: soybeans can help address the caloric and protein needs of a growing global population. Frontiers in Nutrition. 2022;9:909464. https://doi.org/10.3389/fnut.2022.909464
12. Anonymous. ICAR-IISR Indore statistics; 2024.
13. Ghiday T, Sentayehu A. Genetic divergence analysis on some soybean (*Glycine max* (L.) Merrill) genotypes grown in Pawe, Ethiopia. American-Eurasian Journal of Agriculture and Environmental Science. 2015;15(10):1927-33.
14. Ramteke R, Murlidharan P, Shivakumar M, Gireesh C, Ramesh SV. Morpho-agronomic characterization of Indian soybean for grouping and varietal protection. Indian Journal of Genetics and Plant Breeding. 2015;75(3):382-5. https://doi.org/10.5958/0975-6906.2015.00060.7
15. Pachori S, Tiwari S, Shrivastava MK, Amrate PK, Thakur S, Gupta S. Unveiling genetic diversity in soybean germplasm: a comprehensive analysis through DUS characteristics. Electronic Journal of Plant Breeding. 2023;14(4):1541-8. https://doi.org/10.37992/2023.1404.173
16. Patel T, Babbar A, Behera K, Anand KJ, Patel M, Kujur MJ, et al. Integrating diversity analysis and morphological characterization for strategic trait selection in advanced breeding lines of chickpea (*Cicer arietinum* L.). Legume Research-An International Journal. 2025; In press. https://doi.org/10.18805/LR-5468
17. Shilpashree N, Devi SN, Manjunathagowda DC, Muddappa A, Abdelmohsen SA, Tamam N, et al. Morphological characterization, variability and diversity among vegetable soybean (*Glycine max* L.) genotypes. Plants. 2021;10(4):671. https://doi.org/10.3390/plants10040671
18. Rani R, Raza G, Tung MH, Rizwan M, Ashfaq H, Shimelis H, et al. Genetic diversity and population structure analysis in cultivated soybean (*Glycine max* [L.] Merr.) using SSR and EST-SSR markers. PLoS One. 2023;18(5):e0286099. https://doi.org/10.1371/journal.pone.0286099
19. Malik MF, Ashraf MU, Qureshi AS, Khan MR. Investigation and comparison of some morphological traits of the soybean populations using cluster analysis. Pakistan Journal of Botany. 2011;43(2):1249-55.
20. Khan NA, Islam MS, Bhuiyan SH, Hasan KM, Hasan MK. Evaluation of yield contributing characters and cluster analysis of soybean genotypes. Algerian Journal of Biosciences. 2022;3(1):27-32. https://doi.org/10.57056/ajb.v3i1.52
21. Hegay S, Geleta M, Bryngelsson T, Asanaliev A, Garkava-Gustavsson L, Persson Hovmalm H, et al. Genetic diversity analysis in *Phaseolus vulgaris* L. using morphological traits. Genetic Resources and Crop Evolution. 2014;61:555-66. https://doi.org/10.1007/s10722-013-0056-3
22. Koh KY, Ahmad S, Lee JI, Suh GH, Lee CM. Hierarchical clustering on principal components analysis to detect clusters of highly pathogenic avian influenza subtype H5N6 epidemic across South Korean poultry farms. Symmetry. 2022;14(3):598. https://doi.org/10.3390/sym14030598
23. Diniz-Filho JA, Soares TN, Lima JS, Dobrovolski R, Landeiro VL, Telles MP, et al. Mantel test in population genetics. Genetics and Molecular Biology. 2013;36:475-85. https://doi.org/10.1590/S1415-47572013000400002
24. Massy WF. Principal components regression in exploratory statistical research. Journal of the American Statistical Association. 1965;60(309):234-56. https://doi.org/10.1080/01621459.1965.10480787
25. Jolliffe IT. Mathematical and statistical properties of sample principal components. New York: Springer; 1986. p. 23-49. https://doi.org/10.1007/978-1-4757-1904-8_3
26. Getahun T, Tesfaye K, Fikre A, Haileslassie T, Chitikineni A, Thudi M, et al. Molecular genetic diversity and population structure in Ethiopian chickpea germplasm accessions. Diversity. 2021;13(6):247. https://doi.org/10.3390/d13060247
27. Thakur R, Kumari V, Kumari J, Khajuria D, Kumari A, Kapoor R. DUS characterization of indigenous and exotic soybean (*Glycine max* (L.) Merrill) germplasm in north western Himalayas. Himachal Journal of Agricultural Research. 2024:190-200.
28. Yeshitila M, Gedebo A, Tesfaye B, Degu HD. Agro-morphological genetic diversity assessment of *Amaranthus* genotypes from Ethiopia based on qualitative traits. CABI Agriculture and Bioscience. 2024;5(1):95. https://doi.org/10.1186/s43170-024-00288-1

29. Andrew-Peter-Leon MT, Ramchander S, Kumar KK, Muthamilarasan M, Pillai MA. Assessment of efficacy of mutagenesis of gamma-irradiation in plant height and days to maturity through expression analysis in rice. PLoS One. 2021;16(1):e0245603. https://doi.org/10.1371/journal.pone.0245603

30. Zanklan AS, Becker HC, Sørensen M, Pawelzik E, Grüneberg WJ. Genetic diversity in cultivated yam bean (*Pachyrhizus* spp.) evaluated through multivariate analysis of morphological and agronomic traits. Genetic Resources and Crop Evolution. 2018;65:811-43. https://doi.org/10.1007/s10722-017-0582-5

31. Kachare S, Tiwari S, Tripathi N, Thakur VV. Assessment of genetic diversity of soybean (*Glycine max*) genotypes using qualitative traits and microsatellite markers. Agricultural Research. 2020;9(1):23-34. https://doi.org/10.1007/s40003-019-00412-y

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.