RESEARCH ARTICLE

Assessment of genetic diversity of vegetable soybean [*Glycine max* (L.) Merrill] in R6 and R8 stages

Priya P. Pardeshi[*], Dhiraj R. Gangtire, Pravin V. Jadhav, Darasing R. Rathod, Priti A. Sonkamble, Sanjay B. Sakhare, Philips Varghese[1] and Ravindra S. Nandanwar

Abstract

India is a major soybean producer, but vegetable soybean (Edamame) is cultivated in very limited areas due to a lack of variety awareness and adaptive practices. To promote growth, a targeted breeding program is crucial, but inadequate research on the genetic diversity of edamame hinders progress. The analysis of 80 advanced soybean lines derived from crosses between grain and vegetable-type germplasm lines revealed significant correlations between seed yield, seed weight, and 100-seed weight at R6 and R8 stages. Principal component analysis further supported these findings, emphasizing traits crucial for identifying genotypes with hybridization potential. Promising lines, particularly from crosses MACS-1188 × AGS-459 and JS-SH-93-05 × Swarna Vasundhara, may be very useful in developing better-adapted, high-yield edamame varieties, fostering wider acceptance and consumption.

Keywords: Vegetable soybean, PCA, correlation analysis, cluster analysis

Introduction

India plays a great role in the global soybean market and it stands at fifth position on soybean production and export, after the United States, Brazil, Argentina, and China. Soybean is also the chief oilseed crop in India (Bhatia et al. 2018; Chandra et al. 2022) with production of 127.20 lakh tons for the 2021-22 year. Despite the remarkable growth in soybean production in India over the past few decades, there remains an underappreciated vegetable, soybean, which is also called as edamame. Edamame is a Japanese dish prepared with immature soybeans in the pod. It has not yet gained widespread popularity in the country, mainly because of a lack of adaptive accessions and farmers are unaware of the potential of this crop. Edamame is a highly nutritious food, rich in fiber, healthy protein, vitamins, and minerals while being low in fat and trypsin (Dhakal et al. 2021; Lu et al. 2022). Edamame is best harvested during the R6 to R7 growth stages, just before the pods start exhibiting a yellow hue. This timeframe relates to the peak levels of moisture and bean weight (Yu et al. 2022). It can be enjoyed as a whole seed, either as a snack or a vegetable in the human diet (Agyenim-Boateng et al. 2023). The increasing cultivation of high-yielding and adaptable varieties of vegetable soybeans by farmers indicates a positive trend in China (Agyenim-Boateng et al. 2023), North Europe (Zeipina et al. 2022) and Taiwan (Huang et al. 2022). This observation suggests that vegetable soybean has the capacity to become a pivotal crop in India's agricultural economy. Short-duration vegetable soybean types make them ideal for cultivating by small-scale farmers and the increasing demand for fresh, healthy options in India's growing vegetarian trend (Kumar et al. 2023a) and thus, vegetable soybeans have a promising future in India. The crop's potential has been recognized by researchers and breeders, indicating that it could be a very successful crop in the near future (Kumar et al. 2023a; Pardeshi et al. 2023). The present study, was, therefore, conducted to identify

Dr. Panjabrao Deshmukh Krishi Vidyapeeth, Akola 444 104, Maharashtra, India.

[1]Agharkar Research Institute, Pune 411 004, Maharashtra, India.

*Corresponding Author: Priya P. Pardeshi, Dr. Panjabrao Deshmukh Krishi Vidyapeeth, Akola 444 104, Maharashtra, India, E-Mail: priya_genetics@yahoo.com

Table 1. Pedigree of 80 advance lines of vegetable soybean

Lines	Name of the genotypes	Pedigree
1-16	1. VS-1-6-17, 2.VS-1-13-17, 3.VS-1-17-17, 4.VS-1-20-17, 5. VS-1-21-17, 6.VS-1-28-17, 7.VS-1-30-17, 8.VS-1-34-17 9. VS-1-38-17, 10. VS-1-41-17, 11. VS-1-47-17, 12. VS-1-52-17, 13.VS-1-60-17, 14. VS-1-67-17, 15. VS-1-75-17 and 16. VS-1-80-17	Himso-1563 × NRC-55
17-43	17. VS-2-114-17, 18. VS-2-116-17, 19. VS-2-122-17,20. VS-2-123-17, 21.VS-2-124-17, 22.VS-2-125-17, 23. VS-2-128-17, 24.VS-2-129-17, 25. VS-2-130-17, 26. VS-2-131-17, 27. VS-2-132-17, 28. VS-2-133-17, 29.VS-2-135-17, 30. VS-2-136-17, 31. VS-2-137-17, 32. VS-2-139-17, 33. VS-2-141-17, 34. VS-2-146-17, 35. VS-3-98-17, 36.VS-3-99-17, 37. VS-3-103-17, 38. VS-3-104-17, 39. VS-3-105-17, 40.VS-3-106-17, 41.VS-3-108-17, 42.VS-3-154-17 and 43. VS-3-152-17	JS-SH-93-37 × Swarna Vasundhara
44 –70	44. VS-4-178-17, 45. VS-4-182-17, 46.VS-4-183-17, 47.VS-4-184-17, 48. VS-4-186-17, 49. VS-4-187-17, 50. VS-4-190-17, 51. VS-4-192-17, 52. VS-4-195-17, 53. VS-4-198-17, 54. VS-4-200-17, 55. VS-4-201-17, 56. VS-4-205-17, 57. VS-4-209-17, 58.VS-4-213-17, 59. VS-4-219-17, 60. VS-4-223-17, 61.VS-4-226-17,62. VS-4-227-17, 63. VS-4-238-17, 64. VS-4-239-17, 65. VS-4-241-17, 66. VS-4-244-17, 67. VS-4-245-17, 68. VS-4-247-17, 69. VS-4-250-17 and 70. VS-4-254-17	MACS-1188 × AGS-459
71–80	71.VS-5-267-17, 72. VS-5- 255-17, 73. VS-5-265-17, 74. VS-5-266-17, 75. VS-5-268-17, 76. VS-5-272-17, 77. VS-5-273-17, 78. VS-5-274-17, 79. VS-5-275-17 and 80. VS-5-276-17	JS-SH-93-05 × Swarna Vasundhara

promising and diverse lines of vegetable soybeans suitable for farming in Indian conditions. Since vegetable soybean is still in its initial stage of establishment in the country, the research aimed to pinpoint varieties that hold great potential for agricultural growth.

Material and methods

Plant material and experimental conditions

In the present study, a set of 80 advanced lines was selected for the analysis. These lines were carefully chosen based on desired characters from 300 advanced recombinant lines (F_7) generated from the four crosses, namely, Himso-1563 × NRC-55, JS-SH-93-37 × Swarna Vasundhara, MACS-1188 × AGS-459 and JS-SH-93-05 × Swarna Vasundhara involving popular grain and vegetable type soybean varieties obtained from Agharkar Research Institute, Pune, Maharashtra state of India (Table 1). During *kharif* 2021, a set of 80 lines was raised at the experimental field of Dr. Panjabrao Deshmukh Krishi Vidyapeeth, Akola, Maharashtra, India (2042 26.0208″ N, 77° 0.10.6560″ E). The genotypes were planted in a randomized complete block design with three replications. The single-row plot of 50 cm × 20 cm was the experimental unit with planting density was 35 plants per plot. The soil pH ranged between 7.6 and 8.1, with moderate organic carbon levels. The recommended package of practices, including the recommended fertilizer dose, was executed to raise the good crop.

Morpho-physiological and biochemical characterization

The data were recorded on days for 50% flowering (DF), days to maturity (DM) on pot basis, plant height (PH), number of branches per plant (NB), number of pods per plant (NP), number of seeds per pod (NS), chlorophyll content (CC), leaf area index (LAI), protein content in green seed (PR6 in %), protein content in mature seed (PR8 in %), green pod yield per plant (GPY in g), 100-green seed weight (GSW in g), 100- mature seed weight (MSW in g), seed yield at R6 stage (GY/Pl in g), seed yield at R8 stage (MY/Pl in g), carbohydrate content in R6 stage (CHOR6 per 100g), carbohydrate content in R8 stage (CHOR8 per 100 g), oil content in R6 stage (OilR6 in %) and oil content in R8 stage (OilR8 in %). The various yield and quality-contributing traits in the lines were assessed at two different growth stages for better indulgent, which is a green stage (R6), characterized by pods containing green seeds filling the pod volume at one of the four topmost nodes on the central stem, and mature stage (R8), when 95% of the pods have attained their complete coloration.

To measure chlorophyll content, the SPAD 502 device was used, and the results were expressed in SPAD units at the flowering stage on the fifth leaf. For the evaluation of the leaf area index (LAI), a leaf area meter (HS 1395) was applied, which was available at the Physiology Laboratory, Dr. PDKV, Akola. The oil percentage was determined using a nuclear magnetic resonance (NMR) machine at the Instrumental Cell of the Oilseeds Research, Dr. PDKV, Akola. Additionally, carbohydrate and protein content were estimated as percentages following the methods suggested by Pardeshi et al. (2023).

Data analysis

Coefficients of variation (CV) and correlations between traits using Indo Stat software were estimated. Subsequent analysis was conducted using the R software. To construct the correlation network, we utilized the 'Q graph' package (Epskamp et al. 2012) and interpreted the results to assess the relative importance of each trait in relation to genetic diversity among the studied lines. For principal component analysis (PCA), the average trait values were utilized. To

Table 2. Analysis of variance for various characters in advanced soybean lines

Source of variation	Degree of freedom	DF	DM	PH	NBP	N_P	NSP	CC	LAI	PR6	PR8	GPY	GSW	MSW	GY_PI	MY_PI	CHO_R6	CHO_R8	Oil_R6	Oil_R8
Replication	2	3.48	0.19	0.12	1.03	20.20	0.26	6.59	0.04	1.98	2.10	197.27	2.02	2.79	0.58	360.52	0.78	4.23	0.06	1.76
Genotypes	79	20.96**	21.93**	190.04**	8.44**	288.29**	0.38**	77.00**	0.18**	60.32**	63.78**	6570.41**	283.57**	143.03**	3917.77**	1437.48**	18.18**	46.27**	3.38b**	1.60**
Error	158	1.38	2.59	7.16	1.01	17.52	0.10	1.70	0.04	1.07	3.99	184.94	6.18	3.23	111.98	47.85	1.79	6.32	1.17	0.46

** 1% level of significance

DF = Days for 50% flowering, DM = Days for maturity, PH = Plant height, NBP = Number of branches per plant, N_P = Number of pods per plant, NSP = Number of seeds per pod, GPY = Green pod yield per plant, GSW = 100-green seed weight, MSW = 100- mature seed weight, MY_PI = Seed yield at R6 stage, GY_PI = Seed yield at R6 stage, CC = Chlorophyll content, LAI = Leaf area index, PR6 = Protein content in green seed, PR8 = Protein content in mature seed, CHO_R6 = Carbohydrate content in R6, CHO_R8 = Carbohydrate content in R8, Oil_R6 = Oil content in R6, Oil_R8 = Oil content in R8 stage

mitigate potential biases arising from variations among the variables, we standardized the data by centering them to have a mean of zero and a standard deviation of one. The R packages 'FactoMineR' and 'factoextra' (Le et al. 2008; Kassambara and Mundt 2017; Peter et al. 2022) were employed to visualize the results of the PCA. To determine the optimal number of clusters, the K-means algorithm and Bayesian Information Criterion were applied using the 'adegenet' R package.

Results and discussion

Analysis of variance

Analysis of variance (ANOVA) indicated that there were significant differences among the lines under investigation for all the nineteen causal traits related to yield and quality traits (p < 0.005) (Table 2). Notably, genotypes VS-4-254-17, VS-4-255-17, and VS-4-268-17 exhibited superior mean performance in terms of green pod yield. Additionally, these genotypes displayed commendable green seed yield per plant (Supplementary Table S1).

Study on genetic parameters

The analysis depicted in Fig. 1 revealed that for several traits, such as DF, DM, PH NSP and (Oil at R6) the phenotypic coefficient of variation (PCV) values exceeded the genotypic coefficient of variation (GCV) values. This observation suggests a more significant influence of environmental factors on these traits. In contrast, traits like GPY, GSW, MSW, GY/PI and MY/PI displayed a more balanced contribution from both genetic and environmental factors, indicating a greater genetic influence on their performance. These findings corroborate earlier research conducted by Pardeshi et al. (2023). Further, reports of Kumar et al. (2023a) has also revealed similar trend of variation while studying on vegetable soybean. Additionally, all the yield-related traits in the R6 stage demonstrated both high heritability and substantial genetic advance over the mean (GAM). These results underscore the potential for a robust response to selection in these vegetable soybean traits due to their high heritability.

Principal component analysis (PCA) and rotational matric component (RCM)

Numerous researchers widely recommend principal component analysis to reduce
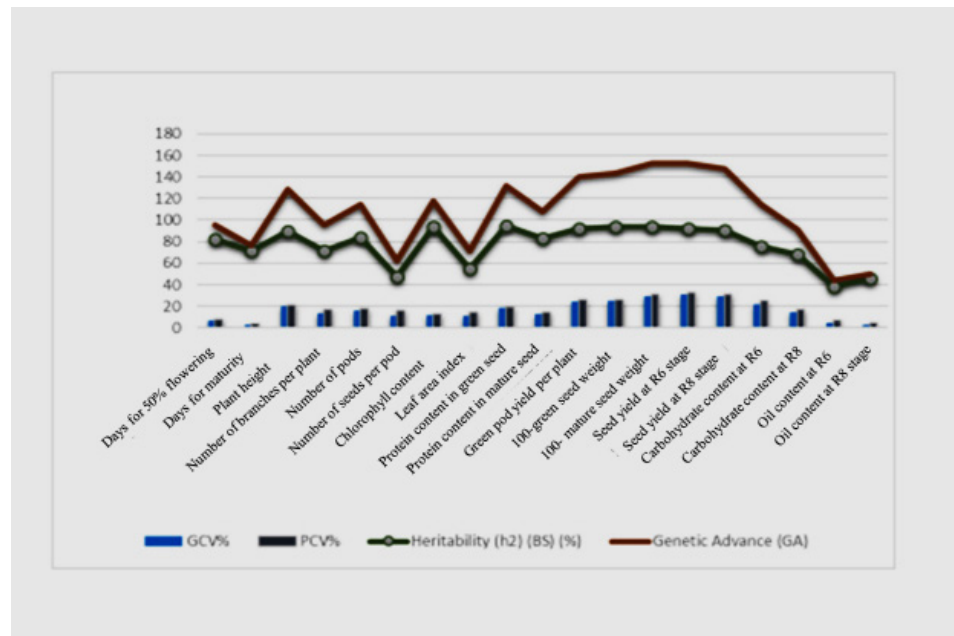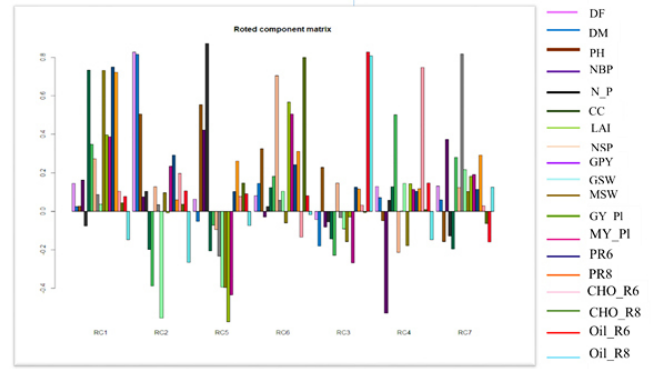

Fig. 1. Estimates of genetic parameters for the traits evaluated in the study

Table 3. Eigenvalues, per cent variance and cumulative percentage of variance of promising lines

Principal component	Eigen value percentage	Percentage of variance	Cumulative percentage of variance
PC1	4.353	22.910	22.910
PC2	3.119	16.415	39.325
PC3	1.769	9.312	48.638
PC4	1.364	7.178	55.815
PC5	1.268	6.672	62.488
PC6	1.084	5.706	68.194
PC7	0.907	4.771	72.965
PC8	0.783	4.124	77.089
PC9	0.744	3.917	81.006
PC10	0.629	3.311	84.317
PC11	0.592	3.114	87.431
PC12	0.555	2.920	90.351
PC13	0.420	2.209	92.560
PC14	0.401	2.111	94.670
PC15	0.303	1.594	96.265
PC16	0.225	1.184	97.449
PC17	0.214	1.127	98.576
PC 18	0.173	0.910	99.487
PC19	0.098	0.513	100.000

DF: Days for 50% flowering, DM Days for maturity, PH: Plant height, NBP: Number of branches per plant, N_P: Number of pods per plant, NSP: Number of seeds per pod, GPY: Green pod yield per plant, GSW: 100-green seed weight, MSW: 100- mature seed weight per, GY-PI: Seed yield at R6 stage, MY/PI: Seed yield at R8 stage, CC: Chlorophyll content, LAI: Leaf area index, PR6: Protein content in green seed, PR8:Protein content in mature seed, CHO_R6: Carbohydrate content in R6 stage, CHO_R8: Carbohydrate content in R8 stage, Oil_R6: Oil content in R6 stage, Oil_R8: Oil content in R8 stage

Fig. 2. Rotation component matrix PC for 19 variables in 80 advanced vegetable soybean lines

several related variables into a smaller set of variables. In the context of assessing diversity and correlation in soybean genotypes, researchers such as El-Hashash et al. (2016) and Siddquie et al. (2023) have endorsed its utility. In our study, PCA was conducted on the economic traits of vegetable soybeans. Out of the nineteen initial variables, only six principal components (PCs) exhibited eigenvalues exceeding 1.00 (Table 3). These six components collectively accounted for approximately 68.194% of the variability observed among the studied traits. Notably, the first principal component contributed the most to this variability, with a substantial proportion of 22.19%.

The obtained scores from PCA serve as valuable indicators for designing precise selection indices, with the intensity of each index determined by the variability explained by each principal component. A high PC score for a specific line within a particular component signifies high values for the variables associated with that line. Additionally, the rotated component matrix sheds light on the variability of all the components and their influence on the performance of traits, (Fig. 2). For instance, RC1 primarily encompasses yield-related traits such as NS, GPY, GY/PI and MY/PI, Fig. 3 illustrates the superior performing lines for each principal component. In the case of PC1, line VS-5-276-17

recorded the highest score, followed by VS-5-268-17, VS-4-200-17, VS-5-272-17, and VS-5-266-17, suggesting these lines as potential selections for yield-related traits.

Moving on to RC2, it encompasses traits like, DF, DM, PH and NBP, with line VS-1-20-17 scoring the highest, followed by VS-3-104-17. RC3 is primarily subject by PH and NP number of pods per plant, while the fourth component is closely allied to LAI, GSW, MSW and CHO at R8. Among the lines, VS-4-223-17 followed by VS-4-195-17 exhibited the highest scores for this component. RC5 is closely associated with, Oil at R6: Oil content at R6 and Oil at R8: Oil content at R8 stage, with line VS-3-98-17 achieving the highest scores. RC6 is characterized by CC and CHO at R6, and the top scorers for this component are line VS-4-192-17, followed by VS-5-267-17 and VS-1-30-17. RC7 is dominated by a single trait, which is PR6, lines VS-5-274-17 and VS-5-276-17 scoring the highest among all the lines, as depicted in Fig. 3. These findings align with those observed by El Hashish et al. (2016), who noted that PC1 and PC2, with eigenvalues greater than unity, explained a significant portion of the total variability among soybean genotypes, mainly related to seed yield.

The principal component biplot, as illustrated in Fig. 4, further exemplifies the relationships between various traits and genotypes within their respective principal components. This figure delineates four groups of traits against four groups of lines, providing valuable insights into their associations. In summary, the results of PCA offer valuable insights into the relationships between different traits and genotypes, guiding us toward more informed selections and breeding strategies for soybean improvement.
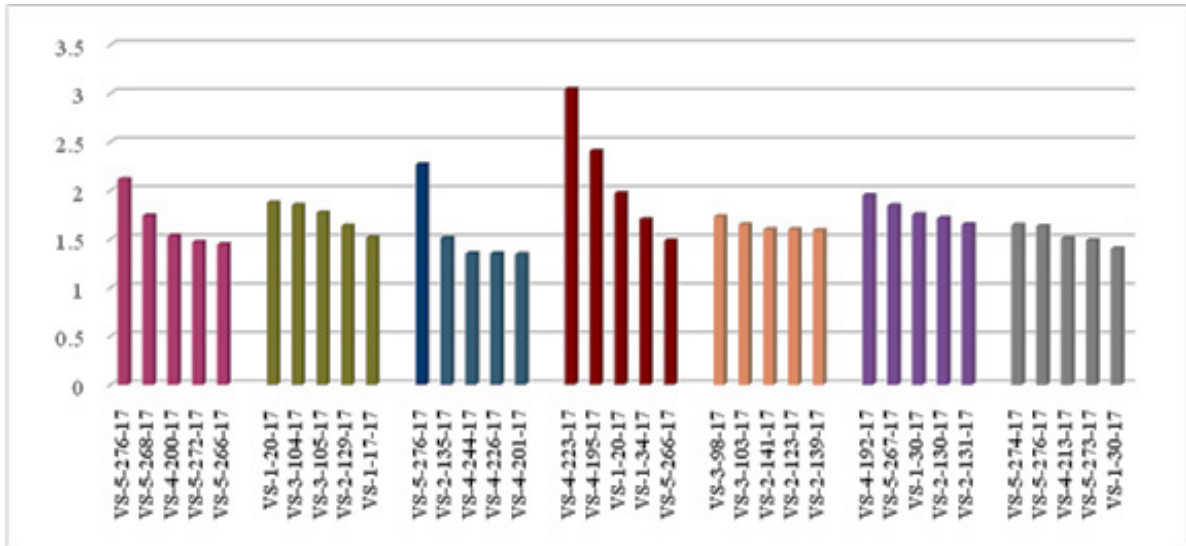
Fig. 3. Five genotypes are selected on the basis of PC score in decreasing order in each component

Fig. 4. Biplot illustrating the relationship between 80 lines and 19 traits of vegetable soybean

Minor angles concerning length of vectors specifying in a similar direction proposed a high degree of significant correlation of the different characters in terms of discriminating lines (Al-Naggar et al. 2020) as GY/Pl with MY/Pl: Seed yield at R8 stage and GPY with GSW and MSW was highly correlated with each other. These conclusions align with the findings obtained from the correlation analysis.

Correlation analysis

Correlation coefficients were employed to assess the degree of association between various traits in vegetable soybeans. Given the limited available literature on trait associations in vegetable soybeans in India, our objective was to determine the relative contributions of different traits to yield at both the R6 and R8 stages, providing a clearer understanding for effective selection. It is well-known that grain yield is
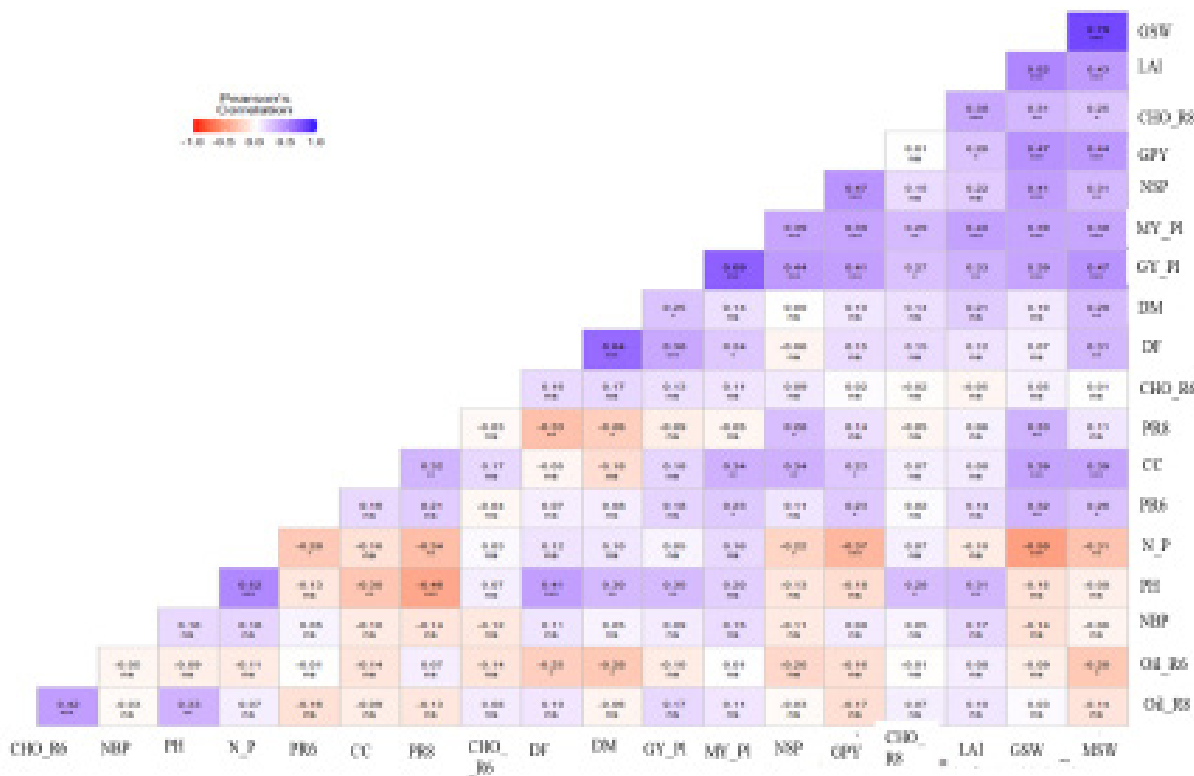
Fig. 5. Association analysis of 19 yield and quality governing traits in advanced lines of vegetable soybean

significantly influenced by environmental factors, making it a complex trait influenced by multiple related traits (Kumar et al. 2023a). To explore these associations, we calculated Pearson correlation coefficients (Fig. 5) among physio-morphological traits, another notable finding was the strong and positive correlation between GY_Pl and various traits, including GSW (0.39, $p < 0.001$), MSW (0.47, $p < 0.001$), and GPY (0.41, $p < 0.001$). The significant positive connection observed may be due to the pleiotropic action of the same gene or the presence of genes governing both characteristics on the same chromosome with tight linkage. This shows that positively associated traits can be addressed concurrently if they both contribute to yield. Seed weight is an important component of soybean produce and is commonly associated with seed yield. Seed size was also found to be associated with seed yield (Li et al. 2019; Kumar et al. 2023b).

A moderate negative correlation between the NP and GPY (-0.37, $p < 0.001$) was recorded, which suggested that selecting plants with a greater number of pods may lead to a reduction in the green pod weight per plant. This phenomenon may be due to compensation mechanisms in certain traits related to yield. However, it's important to note that the external appearance of pods and their substantial size are crucial criteria for selecting high-quality vegetable soybeans (Casas-Leal et al. 2022). Additionally,

it was also observed that GY/Pl and MY/Pl, as well as Oil at R6 and Oil at R8 stages, showed high heritability and strong correlations with each other. This suggests that these traits are mostly influenced by genetics rather than the environment, making them reliable candidates for selection. The association indicates that gains in one trait are likely to result in improvements in others, simplifying breeding efforts. This also suggests that if these traits are controlled by a limited number of genes, selecting for their combination could be an effective approach for vegetable soybean improvement. Kumar et al. (2023a) has identified 11 distinct QTLs were mapped for the yield and yield-related traits using 105 polymorphic microsatellite markers across 20 chromosomes of soybean.

Cluster analysis

The examination of clustering holds significance in evaluating the variability within populations. These studies facilitate the monitoring of germplasm banks and the identification of highly promising crosses. Additionally, they provide crucial insights into the genetic distance among genotypes, offering valuable information for leveraging genetic variability in the creation of new cultivars (Cantelli et al. 2016). A dendrogram was constructed using 19 traits influencing desirable features in vegetable soybeans. The genotypes were classified into three main clusters. Cluster I with 38 lines derived from Himso-1563 × NRC-55 and a few

from the JS-SH-93-05 × Swarna Vasundhara. Although they had lower yields and yield-related traits compared to other lines, however, they excelled in the number of pods per plant. Notably, 16 lines from the cross, Himso-1563 × NRC-55, had the highest average protein content in the R8 stage. Cluster II comprised 18 progenies resulting from the JS-SH-93-37 × Swarna Vasundhara cross. These lines matured late, had numerous branches and pods per plant, and yielded an average of 161.8 g of green pods.

Cluster III contained 24 lines with exceptional characteristics as they displayed the highest average protein content in the green stage and excelled in yield-related traits such as GPY, GSW, MSW, GY/Pl, MY/Pl. Previous studies also stressed the importance of producing a large number of fresh pods in vegetable soybeans. A study conducted by Chandra et 1l. (2022) on genetic divergence analysis using k-means clustering and principal component analysis reported seed longevity, which may be useful in vegetable soybeans. Their study successfully grouped all 125 genotypes from three species of soybean into three distinct clusters. Selecting divergent parents based on genetic variance is crucial for breeding superior progenies with specific traits. The lines chosen in this study exhibited significant diversity and heritable desirable traits, making them valuable resources for India's vegetable soybean breeding program to develop improved varieties with enhanced characteristics.

PCA condensed the initial 19 trait parameters into seven principal components, effectively capturing the dataset's essential variation. This reduced set of components is invaluable for further analysis, as it maximizes variance explanation within a lower-dimensional space. Genetic divergence analysis was then conducted, precisely estimating diversity among selected accessions. This information empowers plant breeders to select diverse parents for hybridization strategically. In conclusion, multivariate analysis, including CA and PCA, grouped accessions and streamlined data dimensions. Genetic divergence analysis provided insights into genetic diversity, can help breeders for selection of potential parents in developing improved plant varieties. The identified better performing lines, VS-4-276-17, VS-4-255-17, VS-4-268-17, VS-4-200-17, VS-5-272-17, and VS-5-266-17, represent novel genotypes resulting from the cross of vegetable and grain type soybean genotypes. These genotypes exhibit promising traits and offer potential for further utilization in vegetable breeding programs in India.

Authors' contributions

Conceptualization of research (PVJ, SBS, RSN); Designing of the experiments (PPP, PVJ, DRG); Contribution of experimental materials (PV); Execution of field/lab experiments and data collection (PPP, DRG, DRR, PAS); Analysis of data and interpretation (PPP); Preparation of the manuscript (PPP, SBS).

References

Agyenim-Boateng K.G., Zhang S., Zhang S., Khattak A.N., Shaibu A., Abdelghany A.M., Qi J., Azam M., Ma C., Feng Y. and Feng H. 2023. The nutritional composition of the vegetable soybean (Maodou) and its potential in combatting malnutrition. Front. Nutri., **9**: 1034115. https://doi.org/10.3389/fnut.2022.1034115

Al-Naggar A.M.M., Shafik M.M and Musa R.Y.M. 2020. Genetic diversity based on morphological traits of 19 maize genotypes using principal component analysis and GT biplot. Annu. Res. Rev. Biol., pp.68- 5. https://doi.org/10.9734/ARRB/2020/v35i230191

Bhartiya A., Aditya J.P., Kumari V., Kishore N., Purwar J.P., Agrawal A., Kant L. and Pattanayak A. 2018. Stability analysis of soybean [*Glycine max* (L.) Merrill] genotypes under multi-environments rainfed condition of North Western Himalayan hills. Indian J. Genet. Plant Breed., **78**(3): 342-347.

Cantelli D.A.V., Hamawaki O.T., Rocha M.R., Nogueira A.P.O., Hamawaki R.L., Sousa L.B. and Hamawaki C.D.L. 2016. Analysis of the genetic divergence of soybean lines through hierarchical and Tocher optimization methods. Genet. Mol. Res., **15**(4): 1-13. http://dx.doi.org/10.4238/gmr.15048836

Casas-Leal N.E., Pereira F.A.C. and Vello N.A. 2022. Improvement of vegetable soybean: genetic diversity and correlations of traits between immature and mature plants. Crop Breed. Appl. Biotech., **22**: e40052218. https://doi.org/10.1590/1984-70332022v22n1a08

Chandra S., Talukdar A., Taak Y., Yadav R.R., Saini M. and Sipani N.S. 2022. Seed longevity studies in wild type, cultivated and inter-specific recombinant inbred lines (RILs) of soybean [*Glycine max* (L.) Merr.]. Genet. Resour. Crop Evol., 1-11. https://doi.org/10.1007/s10722-021-01240-2

Dhakal K., Zhu Q., Zhang B., Li M. and Li S. 2021. Analysis of shoot architecture traits in edamame reveals potential strategies to improve harvest efficiency. Front. Plant Sci., **12**: 614926. https://doi.org/10.3389/fpls.2021.614926

El-Hashash E. 2016. Genetic diversity of soybean yield based on cluster and principal component analyses. J. Adv. Biol. Biotech., **10**(3): 1-9. https://doi.org/10.9734/JABB/2016/29127

Epskamp S., Cramer A.O., Waldorp L.J., Schmittmann V.D. and Borsboom D. 2012. qgraph: Network visualizations of relationships in psychometric data. J. Stat. Soft, **48**: 1-18. https://doi.org/10.18637/jss.v048.i04

Huang Y.H., Ku H.M., Wang C.A., Chen L.Y., He S.S., Chen S., Liao P.C., Juan P.Y. and Kao C.F. 2022. A multiple phenotype imputation method for genetic diversity and core collection in Taiwanese vegetable soybean. Front. Plant Sci., **13**: 948349. https:// doi. org/ 10. 3389/ fpls.

Kumar R., Saini M., Taku M., Debbarma P., Mahto R.K., Ramlal A., Sharma D., Rajendran A., Pandey R., Gaikwad K., Lal S.K. and Talukdar Akshay. 2023b. Identification of quantitative trait loci (QTLs) and candidate genes for seed shape and 100-seed weight in soybean [*Glycine max* (L.) Merr.]. Front. Plant Sci.,

13: 1074245. https://doi/10.3389/fpls.2022.1074245.

Kassambara A. and Mundt F. 2017. Factoextra: Extract and visualize the results of multivariate data analyses. R package version 1.0.4

Kumar R, Saini M., Taku M., Debbarma P., Sharma D., Pandey R., Gaikwad K., Lal S.K. and Talukdar A. 2023a. Genetic studies and mapping of quantitative trait loci (QTL) for yield and yield-related traits in vegetable and seed soybean (*Glycine max* L.) -derived populations. Indian J. Genet. Plant Breed., **83**(03): 301-309. DOI: https://doi.org/10.31742/ISGPB.83.

Le S., Josse J. and Husson F. 2008. FactoMineR: An R package for multivariate analysis. J. Stat Soft., **25**: 1-18. https://doi.org/10.18637/jss.v025

Li X., Zhang X., Zhu L., Bu Y., Wang X., Zhang X., Zhou Y., Wang X., Guo N., Qiu L. and Zhao J. 2019. Genome-wide association study of four yield-related traits at the R6 stage in soybean. BMC Genet., **20**(1): 1-15. https://doi.org/10.1186/s12863-019-0737-9

Lu W., Sui M., Zhao X., Jia H., Han D., Yan X. and Han Y. 2022. Genome-wide identification of candidate genes underlying soluble sugar content in vegetable soybean (*Glycine max* L.) via association and expression analysis. Front. Plant Sci., **13**: 930639. https://doi.org/10.3389/fpls.2022.930639

Pardeshi P., Jadhav P., Sakhare S., Zunjare R., Rathod D., Sonkamble P., Saroj R. and Varghese P. 2023. Morphological and microsatellite marker-based characterization and diversity analysis of novel vegetable soybean [*Glycine max* (L.) Merrill]. Mol. Bio. Rep., **50**(5): 4049-4060. https://doi.org/10.1007/s11033-023-08328-1

Peter S., Le Provost G., Mehring M., Müller T and Manning P. 2022. Cultural worldviews consistently explain bundles of ecosystem service prioritization across rural Germany. People Nat., **4**(1): 218-230.https://doi.org/10.1002/pan3.10277

Shilpashree N., Devi S.N., Manjunathagowda D.C., Muddappa A., Abdelmohsen S.A., Tamam N., Elansary H.O., El-Abedin T.K.Z., Abdelbacki A.M. and Janhavi V. 2021. Morphological characterization, variability and diversity among vegetable soybean (*Glycine max* L.) genotypes. Plants, **10**(4): 671. https://doi.org/10.3390/plants10040671

Siddquie M.N.E.A and Hoque M.A. 2023. Genetic diversity based on principal component and cluster analysis for various characters in spring wheat (*Triticum aestivum* L.) genotypes under drought condition. Fundam. Appl. Agric., **8**(1&2): 435-446. https://doi.org/10.5455/faa.146884

Yu D., Lord N., Polk J., Dhakal K., Li S and Yin Y. 2022. Duncan S.E., Wang H., Zhang B., Huang H. Physical and chemical properties of edamame during bean development and application of spectroscopy-based machine learning methods to predict optimal harvest time. Food Chem., **368**: 130799. https://doi.org/10.1016/j.foodchem.2021.130799

Zeipina S., Vagen I.M. and Lepse L. 2022. Possibility of vegetable soybean cultivation in North Europe. Horticulturae, **8**(7): 593. https://doi.org/10.3390/horticulturae8070593