



International Journal of Research in Agronomy

E-ISSN: 2618-0618
P-ISSN: 2618-060X
© Agronomy
NAAS Rating (2026): 5.20
www.agronomyjournals.com
2026; 9(8): 597-603
Received: 19-05-2026
Accepted: 23-06-2026

Sandeep Varma
Department of Genetics and Plant
Breeding, CSA University of Agriculture
and Technology, Kanpur, Uttar Pradesh,
India

Shweta
Department of Genetics and Plant
Breeding, CSA University of Agriculture
and Technology, Kanpur, Uttar Pradesh,
India

Neha Yadav
Department of Genetics and Plant
Breeding, CSA University of Agriculture
and Technology, Kanpur, Uttar Pradesh,
India

Nitin Kumar
Department of Genetics and Plant
Breeding, CSA University of Agriculture
and Technology, Kanpur, Uttar Pradesh,
India

Parul Pandey
Department of Genetics and Plant
Breeding, CSA University of Agriculture
and Technology, Kanpur, Uttar Pradesh,
India

Majiga Revanth Kumar
Department of Genetics and Plant
Breeding, CSA University of Agriculture
and Technology, Kanpur, Uttar Pradesh,
India

Shivam Tripathi
Department of Genetics and Plant
Breeding, CSA University of Agriculture
and Technology, Kanpur, Uttar Pradesh,
India

Amit Kumar
Department of Genetics and Plant
Breeding, CSA University of Agriculture
and Technology, Kanpur, Uttar Pradesh,
India

Nikhil Kumar
Department of Genetics and Plant
Breeding, CSA University of Agriculture
and Technology, Kanpur, Uttar Pradesh,
India

Chandragupt Maurya
Department of Genetics and Plant
Breeding, CSA University of Agriculture
and Technology, Kanpur, Uttar Pradesh,
India

Corresponding Author:

Shweta
Department of Genetics and Plant
Breeding, CSA University of Agriculture
and Technology, Kanpur, Uttar Pradesh,
India

Elucidating genetics and morphological diversity in Indian mustard (*Brassica juncea* L.) through principal component and cluster analysis

**Sandeep Varma, Shweta, Neha Yadav, Nitin Kumar, Parul Pandey, Majiga
Revanth Kumar, Shivam Tripathi, Amit Kumar, Nikhil Kumar and
Chandragupt Maurya**

DOI: <https://www.doi.org/10.33545/2618060X.2026.v9.i8h.6297>

Abstract

Indian mustard (*Brassica juncea* (L.) Czern. & Coss.) is one of the most important oilseed crops cultivated worldwide, and the effective utilisation of genetic variability is essential for developing superior, high-yielding cultivars. The present investigation was conducted to assess genetic variability, character association, genetic divergence, and principal component structure among 36 mustard genotypes using 12 quantitative traits. Analysis of variance revealed highly significant differences among genotypes for all traits studied, indicating substantial genetic variability. The phenotypic coefficient of variation (PCV) was marginally higher than the genotypic coefficient of variation (GCV) for most traits, suggesting limited environmental influence on trait expression. High heritability coupled with high genetic advance as percent of the mean was observed for oil content, plant height, number of siliquae on the main raceme, and number of siliquae per plant, indicating the predominance of additive gene action and the effectiveness of direct selection. Correlation analysis revealed that seed yield per plant exhibited significant positive associations with 1000-seed weight (0.4765) and oil content (0.4563), while a significant negative association was observed with plant height (-0.4781). Cluster analysis grouped the genotypes into six distinct clusters, with the maximum inter-cluster distance recorded between Clusters VI and V (8.104), indicating substantial genetic divergence and the potential for exploiting heterosis through hybridization. Principal component analysis identified five principal components with eigenvalues greater than one, accounting for 74.23% of the total phenotypic variation. Plant height, number of siliquae on the main raceme, 1000-seed weight, seed yield per plant, and oil content contributed most prominently to the observed genetic diversity. The findings suggest that these traits can serve as reliable selection criteria in mustard breeding programmes, while genetically divergent genotypes from distant clusters may be utilized as promising parental lines for the development of high-yielding and genetically diverse mustard cultivars.

Keywords: *Brassica juncea*, genetic variability, heritability, correlation, cluster analysis, principal component analysis, genetic divergence, seed yield

1. Introduction

Indian mustard (*Brassica juncea* L.) is one of the most important oilseed crops cultivated in Asia and plays a significant role in the edible oil economy of India. It belongs to the family Brassicaceae (Cruciferae) and possesses an amphidiploid genome (AABB, $2n = 4x = 36$), originating from natural hybridization between *Brassica rapa* ($2n = 20$, AA) and *Brassica nigra* ($2n = 16$, BB) (Nagaharu, 1935) [20]. Indian mustard is predominantly self-pollinated, although a certain degree of cross-pollination may occur due to insect activity. The crop is well adapted to the semi-arid and subtropical regions of India because of its short duration, low water requirement, and ability to perform under diverse agro-climatic conditions. Mustard is the third most important oilseed crop in the world after soybean and oil palm and is widely cultivated in Asia, Europe, Canada, Africa, and Australia (Fu *et al.*, 2006) [7]. The seeds contain approximately 38–42% oil and 20–25% protein, making them an important source of edible oil and plant-based nutrition (Jat *et al.*, 2019) [9]. In addition to edible oil production, mustard is utilized as green leafy vegetables, condiments, spices, fodder, and livestock feed. The oil cake obtained after oil extraction is rich in protein and

serves as a valuable feed supplement. Furthermore, mustard possesses medicinal and therapeutic properties, as its seeds and oil are known to aid digestion, stimulate gastric secretions, and exhibit antimicrobial and antioxidant activities (Mitiku *et al.*, 2022) ^[18]. India is one of the largest producers and consumers of mustard worldwide. During the crop year 2024–25, mustard occupied approximately 8.63 million hectares and contributed substantially to the country's oilseed production. Despite its economic importance, the productivity of Indian mustard remains relatively low compared with its genetic potential. Yield improvement is constrained by several factors, including a narrow genetic base, susceptibility to biotic and abiotic stresses, and inadequate exploitation of available genetic variability. Therefore, the identification and utilization of genetically diverse and superior genotypes are essential for developing high-yielding cultivars with improved oil quality and stress tolerance.

Genetic variability is the fundamental prerequisite for the success of any crop improvement programme, as the efficiency of selection largely depends on the magnitude and nature of genetic variation available within the breeding population (Burton and DeVane, 1953; Falconer and Mackay, 1996) ^[3, 6]. Estimation of genetic parameters, including the phenotypic coefficient of variation (PCV), genotypic coefficient of variation (GCV), heritability, and genetic advance, provides valuable insights into the relative contributions of genetic and environmental factors to trait expression and helps predict the effectiveness of selection (Johnson *et al.*, 1955; Burton and DeVane, 1953) ^[10, 3]. Furthermore, correlation analysis reveals the degree of association among economically important traits, whereas path coefficient analysis partitions these associations into direct and indirect effects, thereby identifying traits with the greatest influence on seed yield and facilitating the formulation of effective selection strategies (Wright, 1921; Dewey and Lu, 1959) ^[31, 5]. The assessment of genetic divergence using multivariate techniques such as cluster analysis enables the identification of genetically diverse parental lines, which can be exploited in hybridization programmes to maximize heterosis and generate superior transgressive segregants (Mahalanobis, 1936; Rao, 1952) ^[14, 25]. Similarly, principal component analysis (PCA) serves as an efficient multivariate statistical approach for reducing data dimensionality, identifying the major sources of variation, and determining the traits contributing most significantly to genetic diversity among genotypes (Jolliffe, 2002; Hair *et al.*, 2019) ^[11, 8]. In view of these considerations, the present investigation was undertaken to assess the extent of genetic variability, character association, genetic divergence, and principal component structure among Indian mustard (*Brassica juncea* L.) genotypes. The study aimed to identify superior yield-contributing traits and genetically diverse genotypes that can be effectively utilized in future breeding programmes for the development of high-yielding and improved mustard cultivars.

2. Materials and Methods

The present experiment was conducted on Indian mustard (*Brassica juncea* L.) at the Student Instructional Farm (SIF), Chandra Shekhar Azad University of Agriculture and Technology, Kalyanpur, Kanpur, Uttar Pradesh, during two consecutive Rabi seasons, 2024–25. The experimental material

comprised 36 distinct mustard genotypes, including three check varieties. These genotypes were obtained from the National Bureau of Plant Genetic Resources (NBPGR), New Delhi.

The experiment was laid out in a Randomized Block Design (RBD) with three replications. Each genotype was grown in two rows of 3 m length, maintaining a row-to-row spacing of 30 cm and a plant-to-plant spacing of 10 cm. Observations were recorded on twelve quantitative traits, namely: days to 50% flowering, days to maturity, plant height (cm), length of main raceme (cm), number of primary branches per plant, number of secondary branches per plant, number of siliquae on the main raceme, number of siliquae per plant, average number of seeds per silique, 1000-seed weight (g), oil content (%), and seed yield per plant (g).

Statistical analyses were performed to assess genetic variability, character association, genetic divergence, and principal component analysis (PCA) for seed yield and its component traits. Principal Component Analysis (PCA) is a widely used multivariate technique that reduces a large number of correlated variables into a smaller set of uncorrelated components while retaining most of the original information (Massey, 1965; Jolliffe, 1986) ^[11]. In the present study, PCA was employed to identify the most important traits and genotypes before initiating a hybridization programme, thereby facilitating the selection of superior parental lines for breeding purposes. The statistical analyses were carried out using the factextra package in RStudio (Kassambara and Mundt, 2020) ^[12], which provides effective tools for visualizing and interpreting multivariate analyses.

3. Results and Discussion

3.1 Analysis of Variance

The analysis of variance (ANOVA) presented in Table 1 revealed highly significant differences among the 36 mustard genotypes for all the traits studied, indicating the presence of substantial genetic variability within the experimental material. The existence of significant variation among genotypes suggests ample opportunities for the selection and improvement of desirable traits through breeding programmes (Singh and Chaudhary, 1985; Falconer and Mackay, 1996) ^[27, 6].

The effect of replication was found to be non-significant for most of the traits, indicating uniform experimental conditions and minimal environmental variation among replications. However, significant differences due to replication were observed for the number of siliquae on the main raceme and the number of siliquae per plant, suggesting a slight environmental influence on the expression of these traits. The highly significant mean sum of squares due to genotypes for all the characters demonstrated the presence of considerable genetic diversity among the studied mustard genotypes. Such diversity is a prerequisite for the effective selection of superior genotypes and the development of improved cultivars with enhanced yield potential and other desirable agronomic traits (Burton and DeVane, 1953; Johnson *et al.*, 1955) ^[3, 10]. Similar results indicating significant genetic variability among Indian mustard genotypes have been reported by Yadava *et al.* (2011) ^[32], Kumar *et al.* (2017) ^[13], Meena *et al.* (2020) ^[16], and Verma *et al.* (2022) ^[30], who emphasized the importance of genetic variation for the improvement of yield and its contributing traits in mustard.

Table 1: Analysis of variance of augmented block design for 12 characters in *Brassica juncea* genotypes

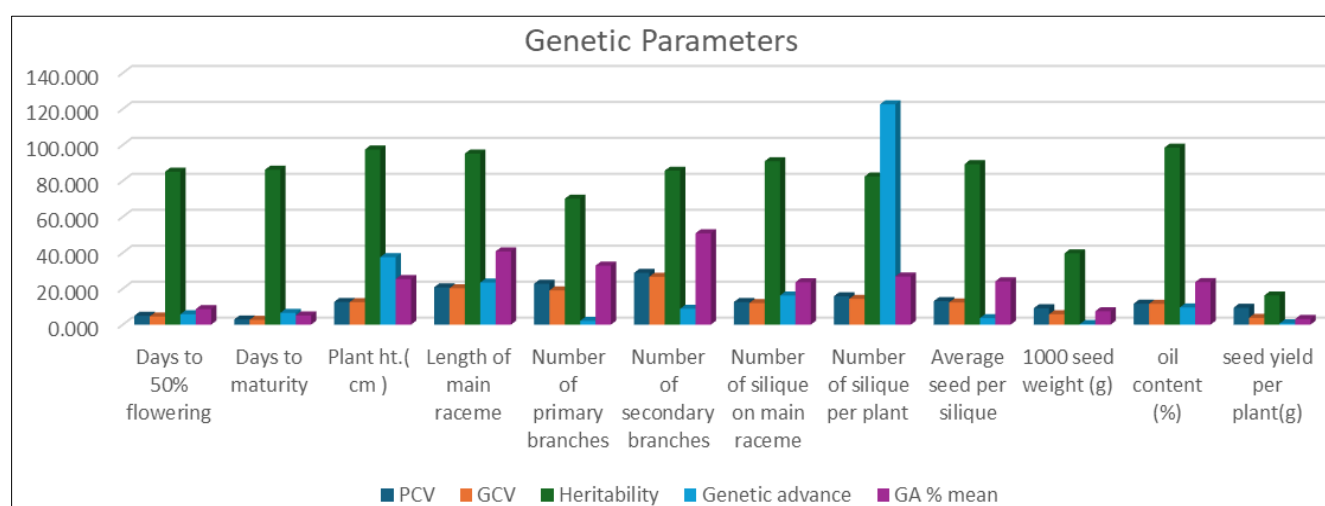
Source of variance	Genotype	Replication	Error
Degree of freedom	35	2	70
Days to 50% flowering	29.673**	11.843	1.633
Days to maturity	37.101**	16.565	1.879
Plant height (cm)	1031.839**	203.245	8.882
Length of main raceme	413.965**	12.128	6.814
Number of primary branches	5.184**	0.065	0.646
Number of secondary branches	67.523**	24.287	3.563
Number of silique on the main raceme	210.754**	92.898	6.774
Number of silique per plant	13795.463**	804.454	911.997
Average seed per silique	10.873**	2.481	0.415
1000 seed weight (g)	0.346**	0.029	0.116
oil content (%)	65.874**	1.549	0.332
seed yield per plant (g)	6.637**	3.926	4.193

**Indicates significant at 1%

3.2 Genetic Variability, Heritability and Genetic Advance

The estimates of phenotypic coefficient of variation (PCV), genotypic coefficient of variation (GCV), heritability, genetic advance (GA), and genetic advance as percent of mean (GAM) for twelve quantitative traits are presented in Figure 1. In general, PCV values were slightly higher than the corresponding GCV values for all the traits, indicating a relatively small influence of environmental factors on trait expression and suggesting that most of the observed variability was genetically controlled. Similar observations were reported by Burton and DeVane (1953) [3] and Johnson *et al.* (1955) [10]. High PCV and GCV estimates were observed for the number of secondary branches per plant, average seeds per silique, number of siliques per plant, plant height, and length of main raceme, indicating the presence of substantial genetic variability for these traits. The close proximity between PCV and GCV values for most characters suggested limited environmental influence and greater scope for improvement through direct selection. Heritability estimates ranged from moderate to high for all the traits studied. The highest heritability was recorded for oil content (99.42%), followed by plant height (97.57%), length of main raceme (95.21%), days to maturity (88.34%), days to 50% flowering (86.15%), number of secondary branches per plant (86.90%), and number of siliques on the main raceme (92.43%) (Figure 1). High heritability indicates that a large proportion of the observed variation is attributable to genetic factors and that phenotypic selection for these traits would be effective (Falconer and Mackay, 1996) [6]. Genetic advance provides an estimate of the expected

improvement under selection. The highest genetic advance was observed for the number of siliques per plant, followed by plant height, number of secondary branches per plant, and length of the main raceme. These traits also exhibited relatively high GAM values, suggesting the predominance of additive gene action and indicating that substantial genetic gain can be achieved through simple phenotypic selection. According to Johnson *et al.* (1955) [10], traits exhibiting high heritability coupled with high genetic advance are mainly governed by additive gene effects and are therefore more amenable to selection. High heritability accompanied by high GAM was observed for plant height, length of main raceme, number of secondary branches per plant, number of siliques per plant, average seeds per silique, and oil content. This combination suggests that these traits are primarily controlled by additive gene action and can be effectively improved through direct selection (Figure 1). Conversely, traits exhibiting high heritability but moderate genetic advance may be influenced by both additive and non-additive gene effects, requiring more careful breeding strategies. Overall, the results demonstrated the existence of substantial genetic variability among the mustard genotypes. Traits such as number of siliques per plant, plant height, length of main raceme, average seeds per silique, and oil content emerged as important selection criteria for the improvement of seed yield in Indian mustard. Similar findings have been reported by Meena *et al.* (2020) [16], Verma *et al.* (2022) [30], and Yadava *et al.* (2011) [32], Nageshwar *et al.* (2023) [21], who emphasized the significance of these traits in mustard breeding programmes.

**Fig 1:** Estimates of ECV, GCV, PCV, heritability, and genetic advance as a percentage of the mean for different quantitative traits in *Brassica juncea* genotypes

3.3. Correlation Analysis

Phenotypic correlation coefficients were estimated among *Brassica juncea* genotypes for 12 traits (Table 2); significant correlations were observed in either the positive or negative direction ($P \geq 0.5$). Among the morphological traits genotypic correlation attributed in seed yield per plant exhibits positive and highest significant genotypic correlation with test weight (0.4765) and oil content (0.4563), negative and significant correlation with the plant height (-0.4781), the correlation positive but non-significant association were observed with days to 50% flowering (0.3287), length of main raceme (0.0587), number of siliques on main raceme (0.0739), number of siliques per plant (0.1166), average seeds per silique (0.0606) and negative association with seed yield but non-significantly association were observed with days to maturity (-0.3240), number of secondary branches (-0.3281) and number of primary branches (0.057) with grain yield in Table 2. Similarly, the days to maturity show positive and significant relation with the plant height (0.4509), plant height with the length on main raceme (0.5546) number of siliques on main branches (0.5302) number of primary branches with number of secondary branches (0.4768), length on main raceme with the number of siliques on main branches (0.6141), number of siliques on main branches with the number of siliques per plant (0.6128) in Table 2. Traits showed a significant negative association: plant height with test weight (-0.4752) and number of secondary branches with the test weight (-0.6481). The correlation analysis revealed substantial associations among seed yield and its

component traits, indicating that several morphological characters can serve as effective selection criteria for yield improvement in Indian mustard. Seed yield per plant exhibited significant positive correlations with test weight and oil content, suggesting that improvement in these traits is likely to enhance grain yield simultaneously. Similar positive associations between seed yield and test weight have been reported by Singh *et al.* (2021) [27], Nageshwar *et al.* (2021) [22], and Yadava *et al.* (2011) [32], indicating the importance of seed weight as a major contributor to yield. Conversely, the significant negative association between seed yield and plant height suggests that taller genotypes may not necessarily be high yielding, a finding consistent with reports by Ameta *et al.* (2018) [2] and Shekhawat *et al.* (2019) [26]. Positive correlations among days to maturity, plant height, raceme length, number of siliques on the main raceme, and siliques per plant indicate that these traits are developmentally associated and may be improved simultaneously through indirect selection. Similar relationships among yield-contributing traits have been reported by Mahla *et al.* (2003) [15] and Kumar *et al.* (2017) [13]. The significant negative correlations of test weight with plant height and secondary branches suggest potential trade-offs between vegetative growth and seed development. Overall, the correlation analysis indicates that selection for higher test weight, increased oil content, and improved siliques-related traits would be effective strategies for enhancing seed yield in Indian mustard breeding programmes.

Table 2: Phenotypic correlation analysis for the 12 traits studied in *Brassica juncea* genotypes

Trait	DF 50	DM	PH	LMR	NPB	NSB	NSQMR	NSQ/P	AS/S	1000 SW	OC	SY/P
DF 50	1	0.2787 NS	-0.0282	0.1689	0.2123	-0.1048	-0.0296	-0.2397	0.1995	0.1439	0.1557	0.3287
DM		1	0.4509 **	0.3619 *	-0.0992	0.0563	0.143	-0.0375	-0.1088	-0.2049	0.0709	-0.324
PH			1	0.5546 **	0.0005	0.2221	0.5302 **	0.3895 *	0.0659	-0.4752 **	0.0301	-0.4781 **
LMR				1	-0.0954	0.1503	0.6141 **	0.1645	0.0831	-0.331 *	0.1438	0.0587
NPB					1	0.4768 **	0.0674	0.0773	0.0753	-0.2085	0.0742	-0.057
NSB						1	0.0095	-0.1323	-0.0473	-0.6481 **	-0.1448	-0.3281
NSQMR							1	0.6128 **	0.2138	-0.175	0.3558 *	0.0739
NSQ/P								1	0.249	-0.0069	0.1071	0.1166
AS/S									1	0.3705 *	0.1691	0.0606
1000 SW										1	0.0344	0.4765 **
OC											1	0.4563 **
SY/P												1

DF 50 - Days to 50% flowering; DM - Days to maturity; PH - Plant height (cm); LMR - Length of main raceme; NPB - Number of primary branches; NSB - Number of secondary branches; NSQMR - Number of silique on the main raceme; NSQ/P - Number of silique per plant; AS/S - Average seed per silique; 1000 SW - 1000 seed weight (g); OC - Oil content (%); SY/P - Seed yield per plant (g); *Indicates significant at 5%; **Indicates significant at 1%.

3.4. Cluster analysis

The genetic diversity among *Brassica juncea* genotypes was assessed using cluster analysis based on 12 morphological traits. The genotypes were grouped into six distinct clusters (Figure 2), indicating considerable genetic divergence. Cluster I was the largest, comprising 10 genotypes, followed by Cluster VI with 9 genotypes. Clusters II and IV each contained 7 genotypes, whereas Cluster III consisted of 2 genotypes and Cluster V contained only one genotype (IC-339597), indicating its unique genetic constitution.

The cluster mean analysis (Table 3) revealed substantial variation among clusters for the studied traits. Cluster IV recorded the highest mean values for seed yield per plant, oil content, days to 50% flowering (70 days), and days to maturity (129 days), indicating that the genotypes in this cluster possessed superior yield potential and oil productivity, although they were relatively late maturing. In contrast, the lowest mean value for days to 50% flowering (63 days) was observed in Cluster I, while the earliest maturity (122 days) was recorded in Cluster V, suggesting the presence of early-maturing genotypes.

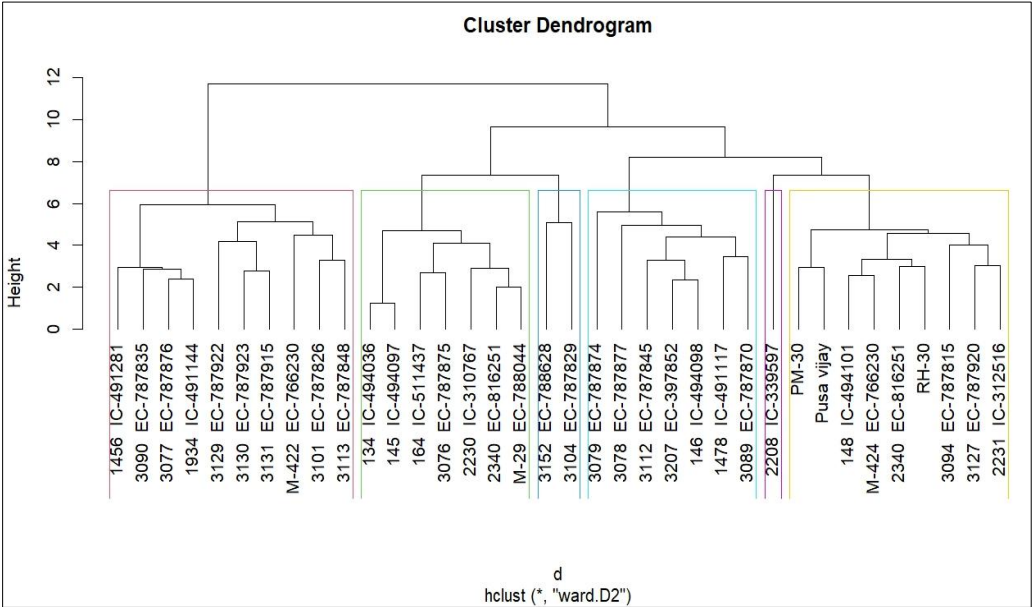


Fig 2: Cluster dendrogram of 36 genotypes distributed into various clusters

Table 3: Cluster mean for different characters in *Brassica juncea* germplasm

Cluster	I	II	III	IV	V	VI
DF 50	63.9	67	66.79	70.86	66.93	66.19
DM	126.81	126.33	125.21	129.48	122.67	128.67
PH	141.6	144.82	142.43	154.73	126.35	169.31
LMR	53.1	41.82	56.94	64.27	42.78	70.34
NPB	5.48	6.83	7.42	5.95	6.67	6.57
NSB	14.43	18.67	22.13	15.1	14.13	18.81
NSQMR	66.1	57	65.33	71.05	62.33	80.05
NSQ/P	461.43	397.33	411.29	453.05	469.47	514.86
S/SQ	14.52	12.5	14.83	15.48	16.27	15.33
1000SW	4.84	4.59	4.54	4.97	5.26	4.59
OC	38.29	27.83	41.25	41.57	43.03	41.88
SY/P	23.43	20.9	23.86	24.78	24.41	22.93

Table 4: Intra and inter-cluster distances among eight clusters in thirty-six *Brassica juncea* genotypes

Cluster	I	II	III	IV	V	VI
I	3.884					
II	4.574	4.072				
III	3.995	4.249	3.095			
IV	4.236	3.741	3.591	3.562		
V	6.129	5.644	4.683	4.846	0	
VI	6.734	6.728	6.662	5.868	8.104	5.097

Bold figures indicate that the intra-cluster distance among different genotypes is divided into seven clusters

The cluster analysis revealed substantial genetic diversity among the 36 *Brassica juncea* genotypes, indicating considerable scope for genetic improvement through hybridization. Similar clustering patterns have been reported by Thakral *et al.* (2015) [28], Mohan *et al.* (2017) [19] and Ponnithyath *et al.* (2025) [24], demonstrating the effectiveness of morphological traits in assessing genetic divergence. The superior cluster mean values for seed yield and oil content observed in Cluster IV suggest that its genotypes are promising parents for yield improvement. Furthermore, the high inter-cluster distance between Clusters VI and V indicates greater genetic divergence, and according to Mahalanobis (1936) [14] and Rao (1952) [25], hybridization between genotypes from genetically distant clusters is expected to generate greater variability and enhance the likelihood of obtaining superior transgressive segregants in breeding programmes.

3.5. Principal Component Analysis (PCA)

Principal component analysis (PCA) was conducted on *Brassica juncea* genotypes using 12 quantitative traits to assess genetic variability. The first five principal components (PCs) had eigenvalues greater than one and together explained 74.23% of the total variation (Table X). Among these, PC1 explained the highest proportion of variation (23.68%; eigenvalue 2.841), followed by PC2 (17.61%; 2.113), PC3 (12.29%; 1.433), PC4 (11.94%; 1.423), and PC5 (8.72%; 1.046) (Table 5). In PC1, plant height (25.17%), number of siliques on the main raceme (20.66%), length of the main raceme (19.69%), and test weight (9.73%) contributed most to the total variation, whereas days to 50% flowering (0.03%) and average seeds per silique (0.47%) contributed the least (Table 5).

DF 50 - Days to 50% flowering; DM - Days to maturity; PH - Plant height (cm); LMR - Length of main raceme; NPB - Number of primary branches; NSB - Number of secondary branches; NSQMR - Number of silique on the main raceme; NSQ/P - Number of silique per plant; AS/S - Average seed per silique; 1000 SW - 1000 seed weight (g); OC - Oil content (%); SY/P - Seed yield per plant (g).
The estimates of intra- and inter-cluster distances (Table 4) further confirmed the extent of genetic divergence among the genotypes. The highest intra-cluster distance was recorded in Cluster VI (5.097), followed by Cluster II (4.072) and Cluster I (3.884), indicating greater variability among genotypes within these clusters. As expected, Cluster V exhibited zero intra-cluster distance (0.000) because it contained only a single genotype. The maximum inter-cluster distance was observed between Clusters VI and V (8.104), followed by Clusters I and VI (6.734), II and VI (6.728), and III and VI (6.662), indicating high genetic divergence between these groups. In contrast, the minimum inter-cluster distance was recorded between Clusters III and IV (3.591), suggesting close genetic similarity. The wide inter-cluster distances observed among several cluster pairs indicate that hybridization between genotypes from these clusters may generate greater genetic variability and increase the probability of obtaining superior recombinants in subsequent breeding generations.

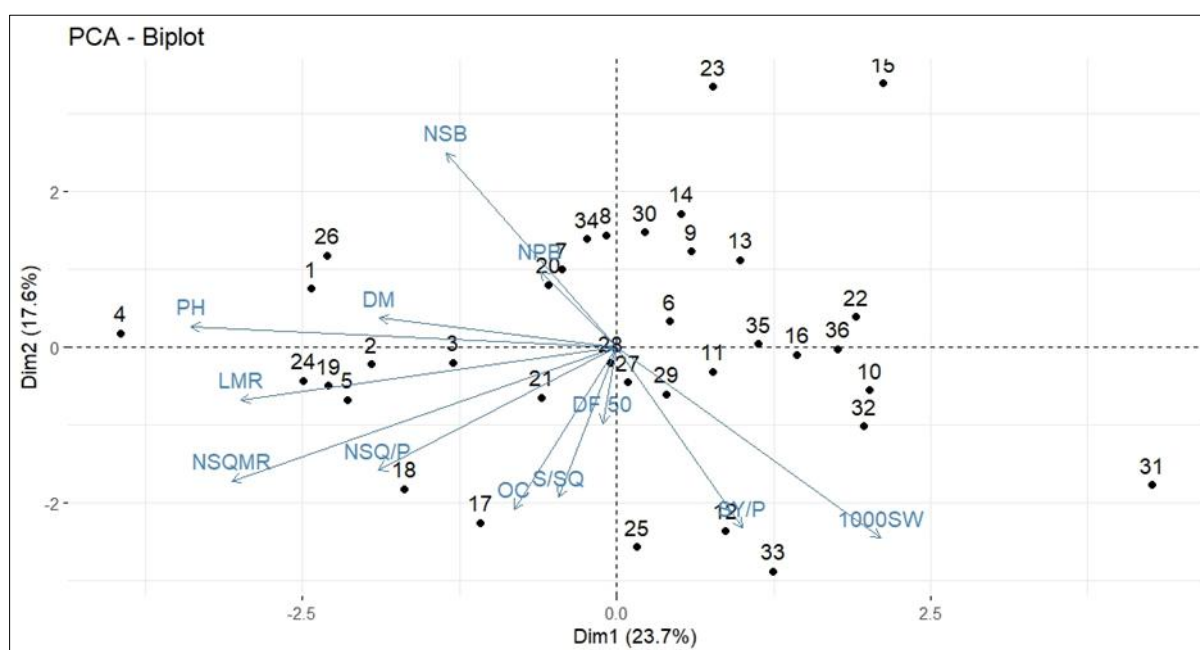
Table 5: Principal component analysis for 13 quantitative traits in pigeon pea germplasm

Trait	PC1	PC2	PC3	PC4	PC5
DF 50	0.03	2.87	47.47	0.01	5.29
DM	7.84	0.4	11.92	18.81	2.99
PH	25.17	0.2	0.59	2.29	2.23
LMR	19.69	1.38	2.55	3.28	1.86
NPB	0.81	2.78	5.66	43.21	0.52
NSB	4.04	18.39	1.87	14.46	1.52
NSQMR	20.66	8.79	2.13	1.04	1.36
NSQ/P	7.91	7.47	20.69	3.92	1.21
S/SQ	0.47	11.04	0.11	9.84	32.32
1000SW	9.73	17.85	0	0.2	8.04
OC	1.46	12.86	4.39	2.14	17.63
SY/P	2.21	15.97	2.63	0.8	25.01
Eigenvalue	2.841	2.1127	1.4747	1.4329	1.0464
Variance Percentage	23.68%	17.61%	12.29%	11.94%	8.72%
Cumulative% Variance	23.68%	41.28%	53.57%	65.51%	74.23%

DF 50 - Days to 50% flowering; DM - Days to maturity; PH - Plant height (cm); LMR - Length of main raceme; NPB - Number of primary branches; NSB - Number of secondary branches; NSQMR - Number of silique on the main raceme; NSQ/P - Number of silique per plant; AS/S - Average seed per silique; 1000 SW - 1000 seed weight (g); OC - Oil content (%); SY/P - Seed yield per plant (g).

In PC2, the highest contributions were from the number of secondary branches (18.39%), test weight (17.85%), seed yield per plant (15.97%), and oil content (12.85%), while plant height (0.20%) and days to maturity (0.40%) contributed minimally. PC3 was primarily influenced by days to 50% flowering (47.47%) and number of siliques per plant (20.69%), whereas PC4 was dominated by number of primary branches (43.21%) and days to

maturity (18.81%). In PC5, the number of siliques per plant (32.32%), seed yield per plant (25.01%), and oil content (17.63%) accounted for the greatest variation. The PCA biplot based on PC1 and PC2 explained 41.29% of the total variability and illustrated the distribution of genotypes and the contribution of individual traits to genetic diversity (Figure 3).

**Fig 3:** PCA Biplot of 36 genotypes, including three checks of *Brassica juncea* genotypes for yield and yield-related traits.

The first five principal components explained a substantial proportion of the total variability, indicating considerable genetic diversity among the evaluated mustard genotypes. Traits such as plant height, siliques number, test weight, seed yield, and oil content contributed most to genetic divergence, suggesting that these characters are important for genotype discrimination and selection. Similar findings were reported by Chakraborty *et al.* (2021) [4], who observed that the first few principal components captured most of the genetic variation in Indian mustard. Likewise, Dewey and Lu (1959) [5] emphasized that principal component scores are effective for identifying traits contributing to genetic diversity and for selecting superior genotypes in

breeding programmes.

4. Conclusion

The present investigation revealed substantial genetic variability among the 36 *Brassica juncea* genotypes for all the studied traits, demonstrating considerable potential for genetic improvement through selection and hybridization. High heritability coupled with high genetic advance for important yield-related traits indicated the predominance of additive gene action and the effectiveness of direct selection. Correlation analysis identified test weight and oil content as the major traits positively associated with seed yield, suggesting that these characters can be used as

reliable selection criteria in mustard breeding programmes. Cluster analysis grouped the genotypes into six distinct clusters, with the greatest genetic divergence observed between Clusters VI and V, indicating that hybridization between genotypes from these clusters may generate superior recombinants and greater heterosis. Principal component analysis further confirmed the existence of broad genetic diversity, with the first five principal components explaining 74.23% of the total variation and identifying plant height, siliques-related traits, test weight, seed yield, and oil content as the major contributors to genetic divergence. Overall, the integration of variability, correlation, cluster, and PCA analyses provides a comprehensive understanding of the genetic architecture of the evaluated germplasm and offers valuable information for parent selection and the development of high-yielding, genetically diverse Indian mustard cultivars.

References

1. Agnihotri A, Sarkar G, Kaushik N, Prem D, Gupta K. Introgression of high oleic acid in Indian mustard through inter-specific hybridization. *Genet Breed*. 2007;188-190.
2. Ameta KD, *et al*. Correlation and path coefficient analysis for yield and its contributing traits in Indian mustard. 2018.
3. Burton GW, DeVane EH. Estimating heritability in tall fescue (*Festuca arundinacea*) from replicated clonal material. *Agron J*. 1953;45:478-481.
4. Chakraborty S, Kumar A, Kishore C, Singh RS, Kumar A, Kumar A, Kumar M. Multivariate based diversity analysis in Indian mustard (*Brassica juncea* L.) genotypes by principal component analysis. *Biol Forum Int J*. 2021;13:680-685.
5. Dewey DR, Lu K. A correlation and path-coefficient analysis of components of crested wheatgrass seed production. *Agron J*. 1959;51:515-518.
6. Falconer DS, Mackay TFC. Introduction to quantitative genetics. 4th Ed. Essex, UK: Longman Group Ltd.; 1996.
7. Fu J, Zhang MF, Qi XH. Genetic diversity of traditional Chinese mustard crops *Brassica juncea* as revealed by phenotypic differences and RAPD markers. *Genet Resour Crop Evol*. 2006;53:1513-1519.
8. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate data analysis. 8th Ed. Cengage Learning; 2019.
9. Jat RS, Singh VV, Sharma P, Rai PK. Oilseed Brassica in India: demand, supply, policy perspective and future potential. *OCL*. 2019;26:8. doi:10.1051/ocl/2019005.
10. Johnson HW, Robinson HF, Comstock RE. Estimates of genetic and environmental variability in soybeans. *Agron J*. 1955;47:314-318.
11. Jolliffe IT. Principal component analysis and factor analysis. In: *Principal component analysis*. New York, NY: Springer; 1986. p. 115-128. doi:10.1007/978-3-642-04898-2_455.
12. Kassambara A, Mundt F. factoextra: Extract and visualize the results of multivariate data analyses. R package version 1.0.7.999; 2020. <https://github.com/kassambara/factoextra>
13. Kumar S, *et al*. Character association and path coefficient analysis for seed yield and its component traits in Indian mustard. 2017.
14. Mahalanobis PC. On the generalized distance in statistics. *Proc Natl Inst Sci India*. 1936;2:49-55.
15. Mahla HR, *et al*. Correlation and path coefficient analysis in Indian mustard. 2003.
16. Meena HS, Ram B, Singh VV. Assessment of genetic variability and correlation studies in Indian mustard. *Indian J Agric Res*. 2020;54(4):471-476.
17. Mehta H, Neha N, Kumar R. Genetic variability of quantitative traits in Indian mustard (*Brassica juncea* L.) germplasm lines. *Int J Environ Agric Biotechnol*. 2025;10(3):618600.
18. Mitiku T, Yali W. Breeding for quality traits improvement of mustard (*Brassica juncea*). *Int Res J Biotechnol*. 2022;12:1-6.
19. Mohan S, Yadav RK, Tomar A, Singh M. Genetic divergence analysis in Indian mustard (*Brassica juncea* (L.) Czern & Coss). *J Pharmacogn Phytochem*. 2017;6(1):350-351.
20. Nagaharu U. Genome analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *J Jpn Bot*. 1935;7:389-452.
21. Nageshwar SS, Yadav VK, Kumar J, Kumar K, Tiwari U. Study of heritability and genetic advance for the selection of improved genotype based on quantitative and quality characters in bread wheat (*Triticum aestivum* L.). *AMA*. 2023;54(8):15077-15085.
22. Nageshwar SVS, Singh M, Singh L, Kumar S, Kumar N, Singh AK. Correlation studies among yield and its components in bread wheat (*Triticum aestivum* L.). *Pharma Innov J*. 2021;10(12):978-982.
23. Pant SC, Singh P. Genetic variability in Indian mustard. *Agric Sci Dig*. 2001;21(1):28-30.
24. Ponniyath A, Singh SK, Upadhyay P, Yadav PK, Krishna B. Genetic characterization for economic yield and its related characters in Indian mustard (*Brassica juncea* L. Czern and Coss). *Biopestic Int*. 2025;21(1).
25. Rao CR. Advanced statistical methods in biometric research. New York: John Wiley & Sons; 1952.
26. Shekhawat N, *et al*. Genetic variability, correlation and path analysis in Indian mustard. 2019.
27. Singh V, *et al*. Correlation and path coefficient analysis in Indian mustard (*Brassica juncea* L.). 2021.
28. Thakral NK, Avtar R, Singh A, Neeru. Evaluation and classification of Indian mustard (*Brassica juncea* L.) genotypes using principal component analysis. *J Oilseed Brassica*. 2015;6(1):167-174.
29. Unified Portal for Agricultural Statistics. 3rd estimate of 2024-25. <https://upag.gov.in/>
30. Verma OP, Kumar K, Singh M. Genetic divergence and variability studies in Indian mustard. *Pharma Innov J*. 2022;11(6):1250-1255.
31. Wright S. Correlation and causation. *J Agric Res*. 1921;20:557-585.
32. Yadava DK, Giri SC, Vignesh M, *et al*. Genetic variability and trait association studies in Indian mustard. *Indian J Agric Sci*. 2011.