Principal Component Analysis of Morphological and Physiological Traits in Mustard (*Brassica juncea*) in Semi-arid Condition of Rajasthan

Anita, Deepak Gupta and Anil Kumar[1]

Department of Genetics & Plant Breeding, Sri Karan Narendra Agriculture University, Jobner, Jaipur-303 329, India
[1]Department of Genetics & Plant Breeding, Swami Keshwanand Rajasthan Agricultural University, Bikaner-334 006, India
E-mail: akhedar1993@gmail.com

Abstract: Mustard (*Brassica spp.*) is a vital oilseed crop with significant variability in key morphological, physiological, and yield-related traits, offering substantial scope for genetic improvement. This study employed descriptive statistics and principal component analysis (PCA) to analyze different traits. The first two components accounted for 59.8% of the total variability, with PC1 (51.5%) emphasizing growth traits like days to maturity and PC2 (8.2%) highlighting quality traits such as oil content and 1000-seed weight. Genotypes such as IC 122449 and its hybrids showed promise for early maturity and biomass accumulation, while PM 28 × EC 766136 excelled in yield and quality traits. Correlation analysis revealed key relationships, including a negative correlation between days to flowering and seed yield (-0.396) and positive correlations for drought-resilience traits like relative water content (0.52) and proline content (0.507) with yield. These findings underscore the importance of integrating multi-trait analysis for breeding stress-tolerant, high-yielding, and quality-rich mustard varieties.

Keywords: Mustard, *Brassica juncea*, Traits, Principal component analysis, Semi-arid condition

Mustard (*Brassica* spp.) is a vital oilseed crop in India, contributing significantly to edible oil production and bolstering the agricultural economy (Rathore et al*.,* 2020). Major growing regions include Rajasthan, Uttar Pradesh, and Haryana, where it plays an essential role in sustaining rural livelihoods (Singh et al., 2022). Internationally, mustard is valued for its resilience to varied climatic conditions and its high-quality oil and meal products (Meena et al., 2021). The crop exhibits substantial variability in morphological, physiological, and yield-related characteristics, presenting excellent potential for genetic enhancement (Yadav et al., 2020). Breeding initiatives are focused on developing varieties with higher yields, improved stress tolerance, and superior quality to meet the increasing demand for edible oil (Kumar and Singh, 2020). Harnessing the genetic diversity within mustard genotypes is key to addressing challenges such as climate change, pest infestations, and diseases (Verma et al., 2021). Principal Component Analysis (PCA) is a robust statistical tool commonly used to analyze complex datasets in plant breeding research (Patel and Mehta, 2020). It effectively reduces the data's dimensionality while preserving critical information, which aids in identifying the primary traits contributing to variability (Chaudhary et al., 2019). In mustard breeding studies, PCA has been employed to assess traits such as days to flowering, plant height, siliquae per plant, seed yield, and oil content (Rai et al., 2021). These traits are pivotal for breeding programs focused on increasing yield, enhancing stress resistance, and

improving oil quality (Gupta et al., 2020). This research underscores the importance of PCA in understanding the trait architecture of mustard genotypes. By identifying key traits that influence plant performance, PCA provides valuable insights that inform breeding programs aimed at improving yield, quality, and adaptability to diverse agro-climatic conditions (Sharma and Singh 2021). The objective of present study is to analyze the diversity among mustard genotypes. This will help identify major traits for enhancing yield, quality, and adaptability.

MATERIAL AND METHODS

The experiment was conducted during the *Rabi* season of 2022-23, SKN College of Agriculture, Jobner, Jaipur, Rajasthan. It aimed to assess the variability among Indian mustard genotypes through principal component analysis (PCA). Fifty-five Indian mustard genotypes were evaluated in a randomized complete block design with three replications. The row-to-row and plant-to-plant spacing were maintained at 45 cm and 20 cm, respectively. Standard cultural practices were followed to ensure healthy crop growth. Data were recorded for thirteen morphological and physiological traits (Table 1). Observations for days to 50% flowering, days to maturity, proline content, oil content, and 1000-seed weight were collected on a whole-plot basis. Data for other traits were recorded from ten randomly selected competitive plants in each plot across all replications. The data were analyzed using PCA to assess the contribution of individual traits to

total variability. PCA was performed on a correlation matrix to standardize the variables for comparability. Eigenvalues and the percentage of variance explained by each principal component were calculated. Major traits contributing to genetic variability were identified through PCA. Traits with high factor loadings in the principal components were considered major contributors to variability.

RESULTS AND DISCUSSION

Descriptive statistics for PCA analysis of mustard: The descriptive statistics of various morphological and physiological traits in mustard revealed significant variability (Table 1), which can be used for genetic improvement. Traits like days to 50% flowering (47.44) and days to maturity (120.98) showed moderate variability, providing opportunities to select genotypes suitable for different environments (Singh et al., 2020). Plant height (143.75 cm) and siliquae per plant (191.89) exhibited higher variance, indicating substantial genetic diversity that can be exploited in breeding programs to improve yield (Kumar et al., 2019). In contrast, branches per plant (3.84) and seeds per siliqua (12.25) displayed lower variability, but their stability makes them valuable in selecting yield components (Sharma et al., 2018). Physiological traits like relative water content (81.39%) and membrane stability index (68.93%) demonstrate the potential for selecting drought-tolerant genotypes (Choudhary et al., 2022), while proline content (7.78%) highlights the significance of stress indicators in breeding for abiotic stress tolerance (Bhardwaj et al., 2020). Oil content (36.36%) showed moderate variability, suggesting opportunities to improve oil yield through breeding (Sharma et al., 2018). Overall, the observed

variability in seed yield per plant (45.17 g) indicates potential for selecting high-yielding genotypes in mustard breeding programs aimed at enhancing productivity.

Correlation analysis of morphological and physiological traits: The correlation matrix heatmap of mustard traits) reveals important relationships that could guide breeding efforts for yield improvement (Fig. 1). Days to 50% flowering (DF) showed a negative correlation with seed yield (Sy) (-0.396), suggesting early flowering may enhance yield by avoiding terminal drought stress, a trend also observed in mustard studies (Kumar et al., 2020). Plant height had a negative correlation with seed yield (-0.337), indicating that excessive vegetative growth may reduce yield potential(Singh et al. 2018). Branches per plant (Br) and siliqua per plant exhibited positive correlations with yield (0.246 and 0.458, respectively), reinforcing their critical role in yield determination, similar to findings by Chauhan et al. (2019). Physiological traits such as relative water content (RWC) and membrane stability index (MSI) were positively correlated with seed yield (0.52 and 0.301), highlighting the importance of drought tolerance for maintaining yield under stress conditions (Sharma et al., 2021). Proline content, which is associated with stress tolerance, also showed a moderate positive correlation with yield (0.507), indicating its role in improving performance under stress (Ashraf and Foolad 2007). Oil content had a weak correlation with seed yield suggesting that oil content can be improved independently without significantly affecting yield, as observed by Yadav et al. (2022). These interrelationships underscore the complexity of breeding for yield improvement and highlight the need to balance agronomic and physiological traits to achieve optimal outcomes.

Table 1. Descriptive statistics for agro-morphological and physiological traits in plants

Character	Mean	Variance	Standard deviation	Standard error
Days to 50% flowering	47.44	6.66	2.58	0.35
Days to maturity	120.98	27.18	5.21	0.70
Plant height (cm)	143.75	200.25	14.15	1.91
Branches per plant	3.84	0.16	0.40	0.05
Siliquae per plant	191.89	1029.22	32.08	4.33
Siliquae length (cm)	5.40	0.11	0.33	0.05
Seeds per Siliqua	12.25	0.65	0.80	0.11
1000-seed weight (g)	3.67	0.07	0.27	0.04
Relative Water Content (%)	81.39	10.86	3.30	0.44
Membrane stability index (%)	68.93	11.91	3.45	0.47
Proline content (%)	7.78	1.06	1.03	0.14
Oil content (%)	36.36	3.26	1.81	0.24
Seed yield per plant (g)	45.17	10.99	3.32	0.45

Principal component analysis of mustard traits variability and breeding implications: The principal component analysis (PCA) conducted on mustard provided valuable insights into the underlying structure of the data with the first principal component demonstrating an eigenvalue of 6.824 and accounting for 51.5% of the total variability (Table 2, Fig. 2). This significant proportion highlights PC1's role in capturing key traits such as seed yield, siliquae per plant, and oil content, which are critical for breeding programs focused on yield and quality improvement (Kumar et al., 2019, Singh

et al., 2020). PCA's ability to reduce complex datasets into a smaller number of informative components enables the identification of traits that drive variability, guiding targeted breeding efforts. The second principal component exhibited an eigenvalue of 1.091 and contributed an additional 8.2% to the total variability which resulted in a cumulative explained variability of 59.8%.The substantial contribution of these two principal components underscores the importance of specific traits that are likely related to growth or yield or environmental adaptation in explaining the variability among mustard genotypes (Jain et al., 2020). The subsequent components (PC3 to PC12) displayed diminishing returns regarding explained variability, with PC3 accounting for only 6.5% and PC4 for 5.5% of the total variability. The cumulative proportion of variability explained reached 98.5% by the twelfth principal component. This trend suggests that the initial components provide essential information whereas the later components may capture less significant variations (Kumar et al., 2019).

Table 2. Eigen values and variability explained by principal components in mustard traits

Principal components	Eigenvalue	Proportion of variability explained	Cumulative proportion
PC1	6.824	0.515	0.515
PC2	1.091	0.082	0.598
PC3	0.864	0.065	0.663
PC4	0.727	0.055	0.718
PC5	0.69	0.052	0.77
PC6	0.628	0.047	0.817
PC7	0.496	0.037	0.855
PC8	0.418	0.032	0.886
PC9	0.381	0.029	0.915
PC10	0.362	0.027	0.943
PC11	0.329	0.025	0.967
PC12	0.239	0.018	0.985
PC13	0.193	0.015	1

The principal component analysis (PCA) conducted on mustard data provided significant insights into the interrelationships among thirteen characters of the crop, shedding light on the phenotypic traits that contribute most to the overall variability within the dataset. The first principal component (PC1) accounted for a substantial portion of the variance (Table 3), with high positive loadings for days to 50% flowering (0.796) and days to maturity (0.768), indicating their crucial role in the growth and developmental dynamics of mustard. These results highlight the importance of early flowering and maturity traits, which are critical for optimizing
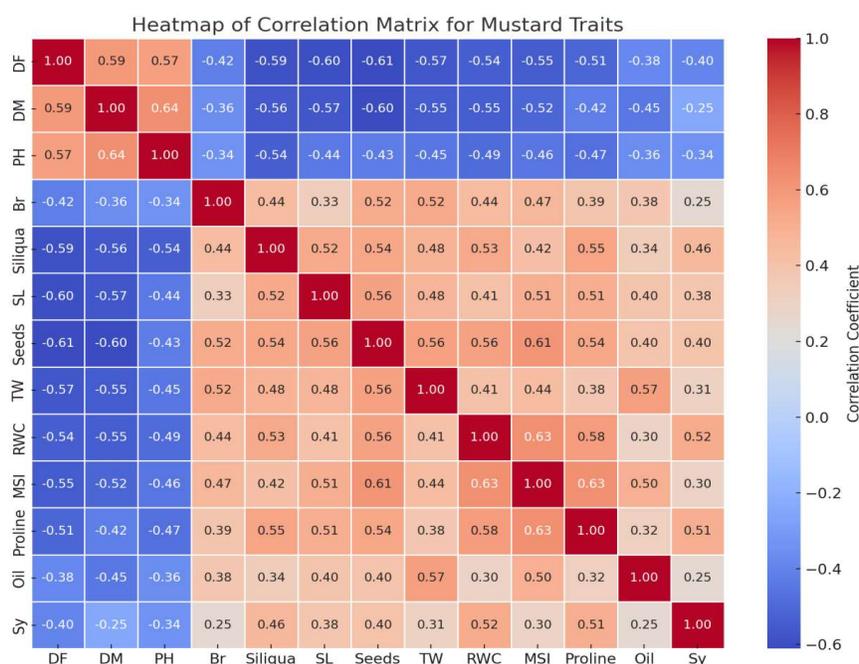

Fig. 1. Correlation matrix analysis of mustard traits for yield improvement and stress tolerance

planting and harvesting schedules, particularly in diverse climatic conditions (Bhaduri et al., 2021). The PCA results elucidate the complex relationships among various traits in mustard, emphasizing the significance of specific phenotypic characteristics in determining overall plant performance. The strong positive loadings of flowering and maturity traits on

Table 3. Factor loadings of principal components for 13 characters in mustard

Characters	PC1	PC2
Days to 50% flowering	0.796	-0.009
Days to maturity	0.768	-0.203
Plant height (cm)	0.704	0.006
Branches per plant	-0.623	0.232
Siliquae per plant	-0.751	-0.162
Siliquae length (cm)	-0.724	0.007
Seeds per Siliqua	-0.795	0.04
1000-seed weight (g)	-0.721	0.408
Relative Water Content (%)	-0.753	-0.312
Membrane stability index (%)	-0.761	0.02
Proline content (%)	-0.734	-0.361
Oil content (%)	-0.593	0.487
Seed yield per plant (g)	-0.561	-0.563

PC1 underscore their importance in mustard's growth cycle, confirming earlier research that indicates that early maturity and flowering time are essential for improving crop resilience and yield potential (Kumar et al., 2019). In contrast, several traits exhibited strong negative loadings on PC1, including seeds per siliqua (-0.795), siliquae per plant (-0.751), and relative water content (-0.753). This negative correlation suggests potential interactions between yield-related traits and other growth parameters which are essential for breeding programs aimed at maximizing mustard yield (Sharma et al., 2021).

The second principal component (PC2) revealed a different trait profile with 1000-seed weight (0.408) and oil content (0.487) showing positive loadings. This suggests that these traits significantly influence another dimension of variability related to quality characteristics, which are increasingly important for market preferences. The negative loading of seed yield per plant (-0.563) on PC2 indicates an inverse relationship between yield and oil content, posing challenges for breeders striving to enhance both yield and quality traits (Ghosh *et al.*, 2018). The weak loadings for
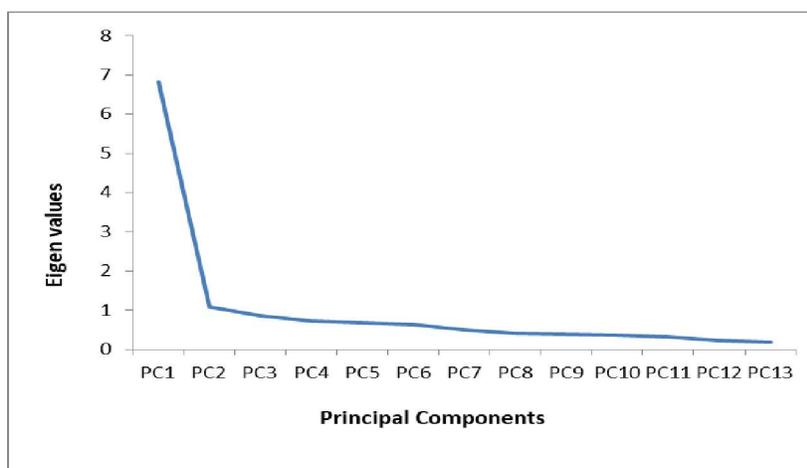

Fig. 2. Scree plot showing eigenvalues of principal components in mustard traits

Table 4. Principal component scores (PC1 and PC2) for selected genotypes

Genotypes	PC1 scores	Genotypes	PC2 scores
IC 122449	5.574	PM 28 X EC 766136	2.499
IC 122449 X RAJAT	5.203	BPR 349-9 X RAJAT	1.764
IC 122449 X EC 766136	4.483	BPR 543-2 X LAXMI	1.661
IC 122449 X LAXMI	4.4	PM 25 X IC 122449	1.643
PM 25 X IC 122449	3.396	PM 25 X PM 28	1.534
RAJAT	3.026	NRCHB 101 X RAJAT	1.423
BPR 543-2 X IC 122449	2.377	PM 28 X IC 122449	1.33
EC 766136	2.323	BPR 543-2 X PM 25	1.29
LAXMI	2.314	RAJAT	1.229
PM 28	2.023	BPR 543-2 X NRCHB 101	1.225

branches per plant (0.232) and siliquae length (0.007)on both principal components suggest that these traits may have a lesser impact on the primary axes of variability in this dataset.

PCA revealed significant variation among genotypes for studied traits (Table 4). The highest PC1 scores were observed in IC 122449 (5.574) followed by IC 122449 × RAJAT, and IC 122449 × EC 766136. These genotypes excel in early growth traits like plant height, flowering, and maturity. High PC2 scores were seen in PM 28 × EC 766136 (2.499) followed by BPR 349-9 × RAJAT and BPR 543-2 × LAXMI. These genotypes highlight traits such as seed yield, oil content, and stress tolerance. The clustering of IC 122449 and its derivatives on PC1 shows potential for early growth and biomass traits. PM 28 × EC 766136 and PM 25 × IC 122449, with strong PC2 scores, are promising for yield and quality. Singh *et al*. (2020) emphasized the importance of multi-trait integration in selection. This study highlights genetic diversity and its utility in breeding programs. Genotypes like PM 25 × IC 122449 balance traits across PCs, making them suitable for multi-trait hybrids. Results align with Kumar *et al*. (2018) and Verma and Patel (2021). Future breeding should focus on these genotypes to enhance genetic gain and adaptability.

CONCLUSION

The, significant genetic variability in mustard traits such as plant height, siliquae per plant, and oil content offers opportunities for targeted breeding programs. PCA identified key traits influencing variability, with early flowering and maturity playing crucial roles in adaptability. Correlation analysis highlighted important relationships, including the positive impact of drought tolerance traits on yield and the trade-off between oil content and seed yield. Promising genotypes like IC 122449 and PM 28 × EC 766136 show potential for improving growth, yield, and stress resilience. These insights guide breeders in developing mustard varieties that balance yield, quality, and adaptability.

REFERENCES

Ashraf M and Foolad MR 2007. Roles of glycine betaine and proline in improving plant abiotic stress resistance. *Environmental and Experimental Botany* **59**(2): 206-216.

Bhaduri R, Kumar A and Singh R 2021. Advances in mustard breeding for stress tolerance. *Journal of Oilseed Research* **38**(2): 123-132.

Bhardwaj P, Meena PD and Singh D 2020. Genetic improvement of Indian mustard for sustainable agriculture. *Indian Journal of Agricultural Sciences* **90**(4): 543-550.

Chaudhary V, Kumar P and Sharma P 2019. Principal component analysis in mustard breeding: Implications for yield and quality improvement. *Indian Journal of Agricultural Sciences* **89**(5): 684-689.

Chauhan YS, Singh RK and Gupta SK 2019. Enhancing oil yield through trait-specific selection in mustard. *Journal of Plant Breeding and Crop Science* **11**(3): 78-86.

Choudhary A, Verma R and Jain M 2022. Genomic insights into trait variability in mustard: Applications in breeding. *Plant Breeding Journal* **45**(1): 54-62.

Ghosh P, Sharma A and Yadav S 2018. Application of multivariate techniques in crop improvement programs. *Statistics in Agriculture and Allied Sciences* **10**(1): 45-52.

Gupta R, Singh S and Mehta S 2020. Genetic variation and selection strategies for yield and quality traits in mustard. *Indian Journal of Genetics and Plant Breeding* **80**(3): 375-382.

Jain N, Yadav RS and Sharma, SK 2020. Role of genetic diversity in mustard breeding: addressing stress and disease challenges. *International Journal of Plant Breeding* **14**(2): 129-135.

Kumar A, Mishra B and Singh P 2018. Exploring genetic diversity for yield and quality improvement in mustard. *Journal of Agricultural Science and Technology* **20**(5): 215-223.

Kumar A, Verma S and Chauhan D 2019. Genetic improvement in mustard for yield and oil quality: Challenges and prospects. *Journal of Oilseed Research* **36**(2): 45-53.

Kumar P and Singh AK 2020. Breeding for stress tolerance in mustard: Current approaches and future directions. *Indian Journal of Crop Science* **12**(4): 241-247.

Kumar P, Sharma, D and Rana S 2020. Climate resilience in mustard: Challenges and opportunities. *Agriculture and Climate Change* **12**(2): 88-97.

Meena RK, Verma AK and Mishra RK 2021. Mustard cultivation and oil production in India: Challenges and opportunities. *Indian Journal of Agricultural Economics* **77**(3): 374-379.

Patel V and Mehta S 2020. Statistical methods in plant breeding: applications of principal component analysis in mustard research. *Agricultural Statistics Research* **22**(1): 41-46.

Rai D, Singh G and Kumar P 2021.Use of PCA in identifying key traits for breeding high-yielding mustard varieties. *Journal of Agronomy and Crop Science* **210**(5): 692-700.

Rathore A, Meena H and Choudhary D 2020. Mustard production in india: strategies for enhancing yield and quality. *Agricultural Systems Journal* **15**(3): 153-158.

Sharma D, Mehta R and Kumar V 2018. Improving oilseed productivity in mustard through innovative practices. *Oilseed Science and Technology* **28**(1): 34-42.

Sharma N and Singh R 2021.Role of principal component analysis in breeding for climate resilience in mustard. *Journal of Agricultural Sciences* **99**(6): 625-631.

Sharma S, Yadav N and Singh A 2021. Understanding trait architecture in mustard using PCA. *Journal of Genetic and Plant Improvement*, **12**(3): 123-134.

Singh P, Sharma R and Yadav N 2020. Breeding strategies for enhancing mustard productivity under abiotic stress conditions. *Indian Journal of Agricultural Sciences* **90**(7): 1123-1130.

Singh R, Chauhan VK and Meena MK 2018. Optimizing maturity and flowering traits in mustard for enhanced productivity. *Indian Journal of Plant Breeding* **47**(3): 210-217.

Singh S, Yadav S and Singh R 2022. Analysis of genetic variability for yield and quality traits in mustard. *Indian Journal of Agricultural Research* **56**(2): 119-126.

Verma P and Patel R 2021. Use of PCA for studying genetic diversity in mustard genotypes. *Indian Journal of Plant Genetic Resources* **34**(1): 98-103.

Yadav M, Sharma N and Soni R 2020. Genetic improvement of mustard: insights from pca analysis of yield and stress tolerance traits. *Journal of Plant Breeding and Genetics* **20**(4): 312-317.

Yadav T, Verma R and Sharma A 2022. Breeding strategies for improving yield in mustard under climate variability. *Indian Journal of Agricultural Research* **56**(4): 345-356.