Journal of Oilseed Brassica, 15 (2) : 279-283, July 2024

Exploring genetic variation and diversity patterns in Indian mustard (*Brassica juncea* L) germplasm for tocopherol content

Vijay Kamal Meena[1,2], Subhash Chand[1,3], Manoj Kumar Patel[1], Yashpal Taak[1], Navinder Saini[1],
Sujata Vasudev[1] and Devendra Kumar Yadava[1]*

[1]Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi 110012, India
[2]Agricultural Research Sub-Station, Agriculture University Jodhpur, Sumerpur, Pali 306902, Rajasthan, India
[3]ICAR-Indian Grassland and Fodder Research Institute, Jhansi 284003, Uttar Pradesh, India
**Corresponding author: dkygenet@gmail.com*

(Received: 22 January 2024; Revised: 07 June 2024; Accepted: 10 June 2024)

Abstract

Tocopherol (vitamin E) is considered an important vitamin carrying antioxidant properties. It plays a vital role in maintaining the quality and stability of the oil in *Brassica* species. Evaluation of genetic diversity and the choice of parents is the crucial step to augmenting the desired improvement of crops towards oil stability and quality improvement. In the present study, 96 Indian mustard genotypes were studied for genetic diversity employing cluster and principal component analysis (PCA) at ICAR-Indian Agricultural Research Institute, New Delhi during 2018-19. Mahalanobis D^2 statistics revealed six clusters where cluster II represented highest 23 genotypes, and cluster V with lowest 6 genotypes. The maximum inter-cluster distance was observed between clusters V and VI (403.4) followed by II and V (327.7) indicating more diversity between clusters. First two PCs accounted for 99% of total genetic variability using PCA. Furthermore, two diverse and complementary parents (Pusa Mahak and CN-101845) were identified that possessed complement traits, *viz.,* high GTC and TTC (Pusa Mahak), and high AGR (CN-101845). Therefore, Pusa Mahak and CN-101845 genotypes could be considered as donor parents for different traits to increase tocopherol content in Indian Mustard.

Keywords: *Diversity analysis, Indian mustard, principal component analysis, tocopherol content traits*

Introduction

Indian mustard [*Brassica juncea* (L.) Czern. & Coss.] is an important oilseed crop that is widely cultivated in many countries of the world including India, China, and Canada. The most prevalent forms of vitamin E in Indian mustard oil are α-and γ-tocopherol, whereas a trivial amount of δ-tocopherol is also reported (Yusuf and Sarin, 2007; Gupta *et al.,* 2015). Amongst the four isoforms (α, β, γ and δ), a-tocopherol has the utmost activity in humans for various purposes (Rizvi *et al.,* 2014). For instance, it acts as an antioxidant that protects the plant from the damage caused by harmful free radical molecules. In addition, a-tocopherol is considered the most effective antioxidant due to its high free radical scavenging activities (Rizvi *et al.,* 2014; Nimse and Pal, 2015). The cultivation of Indian mustard as a major oilseed crop is pivotal for various industrial and culinary applications, necessitating the enhancement of its stability and antioxidant properties. Achieving this improvement mandates the exploration of genetic diversity within the mustard genotypes for effective selection processes. Interestingly, prior to this study, there was a notable dearth of research focusing on the genetic diversity of Indian mustard.

Understanding and evaluating the entirety of genetic diversity within a crop's gene pool is a challenging task. Consequently, there is an increasing imperative for plant breeders to establish an optimal set of plant traits that can elucidate the variability within the population more effectively, spanning the entire crop growth cycle from sowing to harvest. This approach has been acknowledged in previous research (Saroj *et al.,* 2021; Meena *et al.,* 2022; Guguloth *et al.,* 2023; Singh *et al.,* 2024) as an essential strategy for facilitating the selection of superior genotypes. To harness the benefits of transgressive segregation, particularly in critical characteristics such as tocopherol traits, knowledge about the genetic distance between parental plants becomes crucial (Goffman and Becker, 2002). The genotypic improvement of plants in breeding programs is heavily reliant on the extent of genetic diversity present in the population (Meena *et al.,* 2021). The current investigation aims to evaluate the overall diversity among Indian mustard genotypes using multivariate analysis. The primary objective is to identify promising parent plants that can contribute to the augmentation of tocopherol content, thereby enhancing the self-life of the oil produced and fortifying its antioxidant properties. This research is expected to significantly contribute to the understanding of the genetic landscape of Indian mustard, paving the way for targeted and effective breeding programs focused on improving the crop's nutritional and storage-related attributes.

Materials and Methods

In the current study, a comprehensive panel comprising 96 genotypes of *Brassica juncea* was assembled, encompassing varieties, cultivars, introgression lines, derived lines, as well as exotic and indigenous collections. This panel was raised at the experimental farm of the Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi (28°380 N, 77°090 E, and 228.61 MSL), India during the 2018-19 growing season. The experimental trials were organized in a randomized alpha lattice design with two replications, where each plot was composed of four rows measuring 2 m in length, resulting in a plot size of 2.4 m^2. The assessment of tocopherol traits, namely ATC (α-tocopherol content), DTC (β-tocopherol content), and GTC (γ-tocopherol content) in harvested seeds was conducted through high-performance liquid chromatography (HPLC) (Fritsche *et al.,* 2012). In this process, 50 mg of freshly harvested seeds were ground in 1500 µL of n-heptane, and the resulting tocopherol solution was incubated at -20°C for 2 hours. Subsequently, 20 µL of this solution was subjected to HPLC analysis. A silica gel column (5 µM LiChrospher® Si 60, Merck, Darmstadt, Germany) was utilized for tocopherol extraction, employing a mobile phase consisting of an n-heptane/isopropanol mixture (99 + 1; v + v). Tocopherols were quantified using fluorescence detection with excitation at $\lambda = 290$ nm and emission at $\lambda = 328$ nm. Retention times were compared with Merck's tocopherol kit (Merck, Darmstadt, Germany) to identify specific tocopherol forms (GTC: 8.9 min and ATC: 10.6 min), and the concentration of each tocopherol form was calibrated based on signal output. Linear calibration ranges were established for the concentration of the analyzed samples. The tocopherol composition was expressed as the ratio of ATC to GTC, referred to as AGR (α- by γ- tocopherol content ratio) and the TTC (total tocopherol content) was the sum of ATC, DTC and GTC. To analyze the field data and identify promising parents, D^2 statistics (Mahalanobis, 1928; Harman, 1976), principal component analysis (PCA), and agglomerative hierarchical clustering (Ward's method) were employed. The STAR version 2.0.1 (STAR, 2014) software was used for both principal component analysis and agglomerative hierarchical clustering.

Results and Discussion

Genetic diversity employing Mahalanobis D^2 statistics

Total 96 genotypes, accessed from different origins, were clustered into six groups (Table 1). The constellation of grouping exhibited a maximum number (23) of genotypes in cluster II followed by cluster I with 21 genotypes, and lowest number of genotypes (6) were observed under cluster V.

The advantage of genetic diversity analysis based on Mahalanobis D^2 distance over the euclidian distance is that it can take account of the correlation between a highly correlated variable and can scale the contribution of individual variables to the distance value according to the variability of each variable (Ghorbani, 2019). The Mahalanobis D^2 distance among genotypes in the diversity panel ranged from 30.8 to 403.4, which was huge and higher than previous reports by (Bind *et al.,* 2015; Chandra *et al.,* 2018). The huge extent of genetic diversity in the present study was due to the involvement of genetic material from different origins. The intra-cluster heterogeneity of cluster number I (91.2) was high indicating considerable diversity existed among genotypes belonging to this group and might be served as a guideline to select desirable parents for recombination breeding programs within the same cluster (Table 1). was much higher than earlier reports related to D^2 clustering studies by (Bind *et al.,* 2015; Chandra *et al.,* 2018), indicating a wider genetic base of materials within cluster I in the present study. Cluster V showed the highest mean values for ATC, GTC and TTC; however, Cluster III had the highest mean value for AGR (2.1) (Table 2). Which suggest that cluster V is superior in Total tocopherol content but it has lower conversion of g- tocopherol to a- tocopherol whereas Cluster III have more AGR which suggest that it has high conversion of ATC from GTC (Fritsche *et al.,* 2012; Meena *et al.,* 2022). The D^2 analysis for the tocopherol content characters among the 96 Indian Mustard genotypes revealed that all the studied four traits have almost equally contributed towards total divergence (Table 2). The genotypes belonging to cluster V (Pusa Mahak); cluster II (CN-101845) and cluster IV (NPJ-203) were found more diverse and could be utilized as parents for the hybridization programs to obtain maximum heterosis in tocopherol content i.e., ATC and AGR.

Principle component analysis

The utilization of principal component analysis (PCA) serves to fortify the conclusions derived from the D^2 analysis, revealing the primary contributors to variability among genotypes (Sheela *et al.,* 2020; Rahangdale *et al.,* 2021). PCA achieves this by condensing extensive datasets and identifying a concise set of essential independent variables while preserving the original variability. In this study, two principal components (PCs) with eigenvalues surpassing one (Table 2) were derived from the initial data, explaining 99.2 % of the total variation across the 96 finger millet genotypes assessed for tocopherol content traits. The

Table 1: Grouping of 96 Indian genotypes into six clusters, average inter (non-diagonal) and intra (diagonal) cluster distances based on Mahalanobis D^2 statistics (D^2 values)

Cluster	I	II	III	IV	V	VI
Genotype	21[a]	23[b]	19[c]	16[d]	6[e]	11[f]
I	91.2	232.4	176.9	141.4	186.5	308.1
II		51.1	86.1	123.3	327.7	126.8
III			30.8	67.9	272.1	161.8
IV				32.3	236.7	199.0
V					64.6	403.4
VI						64.2

a: AJ-11, BASANTI, CN-101813, CN-105309, DRMRIJ 17-38, DTM-50, EC-564949-1, IC-597869, IC-597910, IC-766097, JM-1, KDM-1049, NDRE-4, PBR-97, Pusa jai kisan, RB-60, RC-371-1, RC-571, RC-891-1, RE-15, RH749

b: AJ-3, CN-101845, CN-105312, CN-113780, CN-34005, IC-597875, IC-597876, IC-597881, IC-597949, JAGANNATH, KDM-49-1, NC-533726, NC-660, PCR-7, PCR-9403, PTJ-3-64, RC-1270, RC-132, RC-371, RGN-34, RGN-73, RJ-10, VASUNDHRA

c: ASHAVATI, CN-105233, CN-105305, CN-105379, I-79 (M), IC-597878, IC-597880, IC-597890, IM-76, JM-06010-1, NPJ-113, PBR-357, PM29, PM-67, PTJ-3-89, RCQR-9901, RE-11, RLC3, YSRL-9

d: CN-105257, DJ-12, DJ-65, IC-597882, IC-597904, IM-108, IM-152, IM-59, NC-37362, NPJ203, PR-2001-42, PUSA BOLD, RC-1271, RE-44, RNN-505, TN-3

e: CN-105306, EM-1, IM-170, KRISHNA, PUSA MAHAK, RE-7-1

f: DJ-5, EJ-22, GIRIRAJ, GR-325, IC-597873, IM-39, NPJ-161, PUSA BAHAR, RH-4193, RHUR-2-1, SITARA SAGAR

Table 2: Estimates of principal component analysis and cluster mean values of tocopherol traits in Indian Mustard genotypes

Character	Principal component analysis				D^2 Statistic								
					Contribution (%)	Cluster numbers						Mean	CV
	PC1	PC2	PC3	PC4		I	II	III	IV	V	VI		
ATC	0.7	0.6	-0.1	0.0	24.7	3	2.5	3.1	3.4	5.8	1.2	3.2	1.9
GTC	0.8	-0.4	0.0	0.0	25.1	219.4	116.7	147.9	171.2	299.5	70.8	170.9	1.9
TTC	0.9	-0.4	0.0	-0.0	25.2	223.4	119.7	152.1	175.6	307.4	72.3	175.1	1.9
AGR	0.4	0.9	0.1	0.0	24.8	1.4	2.1	2.1	2	1.9	1.7	1.9	1.3
Eigenvalue	2.3	1.6	0.0	0.0									
Proportion of variance (%)	58.1	40.9	0.8	0.0									
Cumulative proportion (%)	58.1	99.1	99.9	100									

noteworthy influence of the initial PCs on overall variability aligns with previous findings in the literature, in relation to various plant traits in Indian mustard (Yadav *et al.,* 2022).

The primary contributors to the PCs were identified based on loading factor values. PC1 represented approximately 58.20% of the overall variation, with TTC (0.9) and GTC (0.8) playing a major role in driving this variation positively. PC1 was identified as a key component for tocopherol traits, encompassing TTC and GTC traits associated with higher total tocopherol content. PC2 accounted for 41% of the total variation, with ATC (0.6) and AGR (0.9) displaying highly positive scores due to their strong correlation – an increase in ATC resulted in an increase in AGR. Consistent positive scores for eigenvectors were observed in PC1 and PC2 for all tocopherol traits, collectively explaining 99% of the total variability (Table 2). Within the PC1 genotypes, Pusa Mahak significantly contributed positively, while exhibiting an inverse contribution in PC2. Genotype CN-101845 made significant positive contributions in both PC1 and PC2 directions (Fig. 1A).

The loading plot (Fig. 1B) illustrates the variables' directions based on the sign of score coefficients. ATC and AGR show opposite directions, indicating a negative correlation with GTC and TTC. This implies that GTC primarily contributes to the increase in TTC and ATC. The AGR ratio is a crucial trait for selecting genotypes with high ATC, as it is the only form usable by humans (Meena *et al.,* 2022). The PCA score plot (Fig. 1) and D^2 statistic (Table 2) indicate that the genotype PUSA MAHAK is distantly related to CN-101845. Selecting these two diverse genotypes could effectively enhance grain tocopherol content traits in Indian mustard, as they exhibit complementary traits. Pusa Mahak has high GTC and TTC, while CN-101845 demonstrates high ATC and AGR (Table 2). The clustering pattern of genotypes, as shown in the dendrogram (Fig. 2), further supports the

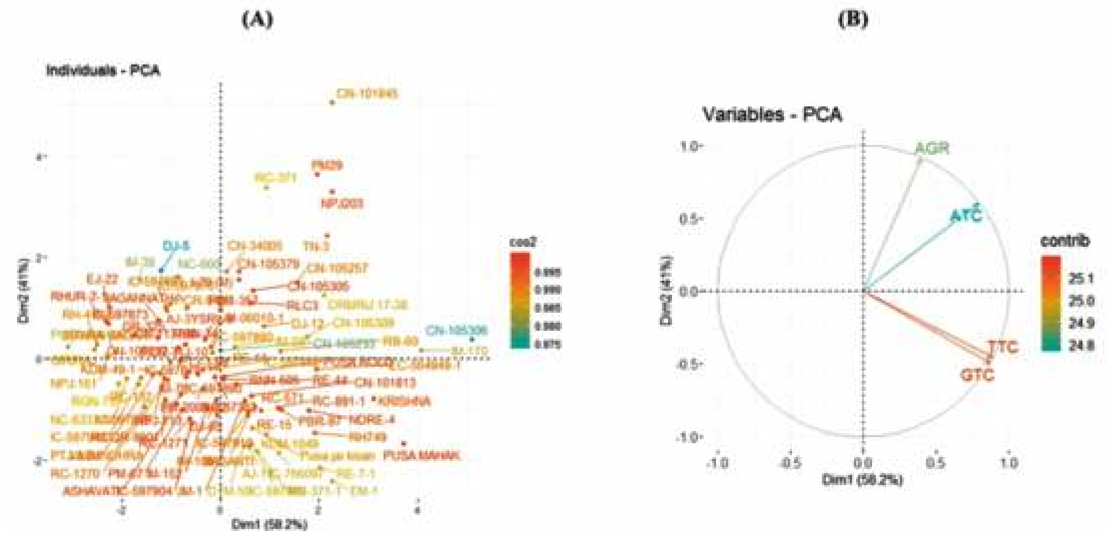

Fig. 1: The principal component analysis showing the distribution pattern of 96 Indian Mustard genotypes (A) and 4 tocopherol content traits (B); Dim1 and Dim2 are the first and second principal components, respectively

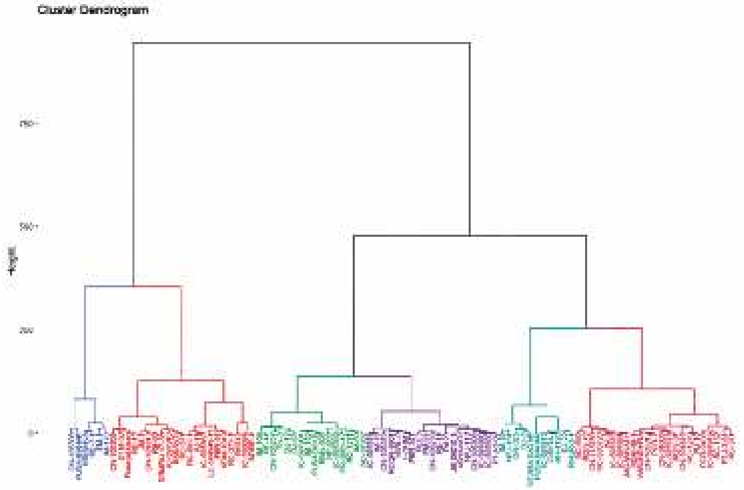

Fig. 2: Dendrogram and agglomerative hierarchical clustering (Ward's method) of 96 genotypes of Indian Mustard. Genotypes were grouped into Six main clusters based on similarity euclidean distances and represented by different colors

selection of Pusa Mahak and CN-101845 as parents in hybridization programs. This approach addresses the limitations associated with developing promising tocopherol content trait cultivars in Indian mustard.

Conclusion

The findings from this research offer valuable insights and lay the foundation for designing upcoming breeding initiatives aimed at enhancing Tocopherol content in Indian mustard. However, a molecular investigation is essential to enhance the parent selection process, introduce new gene combinations into genotypes, and create cultivars of Indian Mustard with elevated tocopherol content.

References

Bind D, Singh SK and Dwivedi VK. 2015. Assessment of genetic diversity and other genetic parameters in Indian mustard (*B. Juncea*). *Indian J Agric Res* **49**: 554–557.

Chandra K, Pandey A and Mishra SB. 2018. Genetic diversity analysis among Indian mustard (*B. juncea*) genotypes under rainfed condition. *Int J Curr Microbiol Appl Sci* **7**: 256–268.

Chaudhary R, Chand S, Alam BK, Yadav P, Meena VK, Patel MK, Pardeshi P, Rathore SS, Taak Y, Saini N, *et al.* 2022. Codon usage bias for fatty acid genes FAE1 and FAD2 in oilseed *Brassica species*. *Sustain* **14**: 11035.

Fritsche S, Wang X, Li J, Stich B, Kopisch-Obuch FJ, Endrigkeit J, Leckband G, Dreyer F, Friedt W, Meng J and Jung C. 2012. A candidate gene-based association study of tocopherol content and composition in rapeseed (*B. napus*). *Front Plant Sci* **3**: 1–24.

Ghorbani H. 2019. Mahalanobis distance and its application for detecting multivariate outliers. Facta Universitatis, Series: Mathematics and Informatics 5:583. https://doi.org/10.22190/FUMI1903583G

Goffman FD and Becker HC. 2002. Genetic variation of tocopherol content in a germplasm collection of *Brassica napus* L. *Euphytica* **125**:189–196.

Guguloth N, Saha S, Verma RK, Tirkey AE, Patel A, Prakash D and Srivasta K. 2024. Genetic variability and character association studies for yield and yield traits in Indian mustard (*B. juncea*). *J Oilseed Brassica* **15**: 81-87.

Gupta S, Sangha MK, Kaur G, Banga S, Gupta M, Kumar H and Banga SS. 2015. QTL analysis for phytonutrient compounds and the antioxidant molecule in mustard (*B. juncea*). *Euphytica* **201**: 345–356.

Harman HH. 1976. Modern Factor Analysis. 3rd edition Chicago, University of Chicago Press, pp. 376.

Kumari A, Kumari V. 2018. Studies on genetic diversity in Indian mustard (*B. Juncea*) for morphological characters under changed climate in the mid-hills of Himalayas. *Pharm Inno J* **7**: 290–296.

Mahalanobis P. 1928. Statistical study of the Chinese head. *Man India*. **8**: 107–108.

Meena VK, Sharma R, Kumar N and Kumar M. 2021. Genetic diversity analysis of bread wheat varieties and pre-release lines evaluated under timely sown irrigated conditions. *Indian J Plant Genet Resour* **34**: 86–89.

Meena VK, Sharma RK, Chand S, Kumar M, Kumar N, Jain N and Singh A. 2022. Elucidating molecular diversity in spring wheat (*T. aestivum*) under terminal heat stress environment using morpho-physiological traits and SSR markers. *Indian J Genet Plant Breed* **82**: 47–55.

Meena VK, Taak Y, Chaudhary R, Chand S, Patel MK, Muthusamy V, Yadav S, Saini N, Vasudev S and Yadava DK. 2022. Deciphering the genetic inheritance of tocopherols in Indian mustard (*B. juncea*). *Plant* **11**: 1779.

Nimse SB and Pal D. 2015. Free radicals, natural antioxidants, and their reaction mechanisms. *RSC Adv* **5**: 27986–28006.

Rahangdale S, Singh Y, Upadhyay PK and Koutu GK. 2021. Principal component analysis of input lines of rice for the important traits responsible for yield and quality. *Indian J Genet Plant Breed* **81**: 127–131.

Rizvi S, Raza ST, Ahmed F, Ahmad A, Abbas S and Mahdi F. 2014. The role of Vitamin E in human health and some diseases. *Sultan Qaboos Univ Med J* **14**: 157–165.

Saroj R, Soumya SL, Singh S, Sankar SM, Chaudhary R, Yashpal R, Saini N, Vasudev S and Yadava DK. 2021. Unraveling the relationship between seed yield and yield-related traits in a diversity panel of *Brassica juncea* using multi-traits mixed model. *Front Plant Sci* **12**: 651936

Singh T, Gupta SK, Sharma M and Sheera A. 2024. Morphological characterization and correlation analysis of F cross of Indian 2:3 mustard × East-European mustard for yield traits in *Brassica juncea*. *J Oilseed Brassica* **15**: 96-100.

Sheela KRVS, Robin S and Manonmani S. 2020. Principal component analysis for grain quality characters in rice germplasm. *Electron J Plant Breed* **11**: 127–131.

Yadav D, Singh L, Kulsoom S, Jafri F, Singh A and Singh V. 2022. Study on genetic variability and principal component analysis in Indian mustard (*B. juncea*). *Pharm Inno J* **11**: 380–383.

Yadava DK, Giri SC, Vignesh M, Vasudev S, Kumar Yadav A, Dass B, Singh R, Singh N, Mohapatra T and Prabhu KV. 2011. Genetic variability and trait association studies in Indian mustard. *Indian J Agric Sci* **81**: 712–716.

Yadava DK, Taak Y, Saini N, Nanjundan J and Vasudev S. 2022. Brassica Breeding. In: Yadava DK, Dikshit HK, Mishra GP, Tripathi S (Eds.), Fundamentals of Field Crop Breeding. 1st ed., Springer Singapore, pp. 779–835.

Yusuf MA and Sarin NB. 2007. Antioxidant value addition in human diets: Genetic transformation of *Brassica juncea* with γ-TMT gene for increased α-tocopherol content. *Transgenic Res* **16**: 109–113.