RESEARCH ARTICLE

SABRAO Journal of Breeding and Genetics
57 (3) 911-923, 2025
http://doi.org/10.54910/sabrao2025.57.3.4
http://sabraojournal.org/
pISSN 1029-7073; eISSN 2224-8978

GENETIC DIVERSITY ANALYSIS OF MUSTARD (*BRASSICA RAPA* L.) IN BANGLADESH

I. JAHAN[1], M.G. RASUL[1], M.M.I. RASHAD[2*], A. SARKAR[3], and S.M.M. FARHAN[4]

[1]Department of Genetics and Plant Breeding, Faculty of Agriculture, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur, Bangladesh
[2]Department of Genetics and Plant Breeding, Faculty of Agriculture, Patuakhali Science and Technology University, Patuakhali, Bangladesh
[3]Soil Resource Development Institute, Divisional Laboratory, Barishal, Bangladesh
[4]Faculty of Agriculture, Patuakhali Science and Technology University, Patuakhali, Bangladesh
*Corresponding author's email: mainul_gpb@pstu.ac.bd
Email addresses of co-authors: ijahan1959@gmail.com, mgrasul@bsmrau.edu.bd, avijitsrdi@gmail.com, skfrn13@gmail.com

SUMMARY

Knowledge about germplasm diversity is an invaluable aid in crop improvement strategies. The following research aimed to study the diversity of 53 mustard (*Brassica rapa*) genotypes. The study had a randomized complete block design with three replications, implemented at the Genetics and Plant Breeding Department experimental field of Bangabandhu Sheikh Mujibur Rahman Agricultural University. The collected germplasms came from the said university and the Bangladesh Agricultural Research Institute. Different multivariate analysis techniques used help classify genotypes across 21 studied characteristics, with all the genotypes grouped into seven clusters. Among the seven clusters, Cluster III had the most genotypes (12) and Cluster IV the least (1). Cluster VII showed the highest intra-cluster distance, and Cluster IV the lowest. The largest inter-cluster distance was between clusters IV and VII, and the smallest was between clusters I and III. Selecting genotypes from clusters with the greatest distances could enhance genetic diversity and heterosis. Genotypes from clusters with moderate to high inter-cluster distances and medium to high yields could be applicable for desirable segregants. Specifically, genotypes G_2, G_{12}, G_{17}, G_{33}, and G_{44} from Cluster I; G_1, G_{10}, G_{38}, and G_{40} from Cluster II; G_8, G_{13}, G_{27}, and G_{43} from Cluster III; G_{51} from Cluster IV; G_{21}, G_{34}, and G_{52} from Cluster V; G_{25} and G_{37} from Cluster VI; and G_3 and G_{41} from Cluster VII would be favorable selections as superior parents for hybridization programs.

Communicating Editor: Prof. P.I. Prasanthi Perera

Manuscript received: July 12, 2024; Accepted: January 25, 2025.

Citation: Jahan I, Rasul MG, Rashad MMI, Sarkar A, Farhan SMM (2025). Genetic diversity analysis of mustard (*Brassica rapa* L.) in Bangladesh. *SABRAO J. Breed. Genet.* 57(3): 911-923. http://doi.org/10.54910/sabrao2025.57.3.4.

Keywords: Oil seed, mustard, genetic diversity, morphological characteristics, cluster analysis, and D^2 statistic

Key findings: The highest genetic divergence emerged between clusters IV and VII, suggesting that selecting genotypes from these clusters could enhance heterotic F_1 generation and diversity in segregating generations.

INTRODUCTION

The Brassicaceae (syn. Cruciferae) family, commonly referred to as the mustard family, is one of the largest plant families, comprising around 340 genera and approximately 4,140 species (Walden *et al*., 2020; Francis *et al*., 2021; Dmitry *et al.,* 2023). This diverse family includes economically and agriculturally significant plants, ranging from vegetables and condiments to oilseed crops. Among its genera, *Brassica* stands out for its wide-ranging contributions, encompassing economically important vegetables (cabbage, broccoli), oilseed crops (field mustard), forage, condiments, and biodiesel (Pal *et al*., 2020). Within this genus, *Brassica rapa,* commonly known as field mustard or yellow sarson, is a widely cultivated oilseed crop (Rahman *et al*., 2022).

Brassica rapa, a diploid species in the Brassicaceae family, is predominantly outcrossing due to its floral morphology and mechanisms that promote cross-pollination (Rakow, 2004). It constitutes a vital component of the edible oil supply and is the leading oil crop concerning productivity area coverage of oilseed in Bangladesh (BBS, 2021). According to multiple records and papers, *B. campestris*, recently reclassified as *Brassica rapa,* has its origins in the Himalayan region, from which it underwent transfer to the European Mediterranean area and Asia (Ali *et al*., 2024). *Brassica rapa's* cultivation occurs on six different continents (Howlett *et al.*, 2018) for various goods and uses, such as oilseed, biofuel, vegetables and vegetable seed, and livestock feed. Their adaptability to germinate and thrive in low temperatures allows cultivation in regions with low temperatures, making them serve as winter crops in temperate zones. Major production areas include China, the Indian subcontinent, Canada, and Northern Europe (Rupa *et al.,* 2023).

Oilseed crops are high in vitamins A, D, E, and K and are mostly grown for edible oils. Their byproducts are useful for biofertilizers and animal feeds since they include calcium and phosphorus (Nehmeh *et al.*, 2022). While these oilseeds are crucial in human diets, Bangladesh's consumption falls well below the recommended 37 g per capita daily intake (Sharmin, 2016). Mustard seeds, rich in oil (40%) and protein (24%), offer a solution to this shortfall (Sharmin, 2016; Rao *et al.*, 2017). Oilseed contains high concentrations of magnesium and selenium; hence, its oil has anti-inflammatory qualities. Additionally, it aids in reducing body temperature and increasing sweat glands (Sultana, 2017).

The release of various mustard varieties has been progressing in Bangladesh from the Bangladesh Agricultural Research Institute (BARI), Bangladesh Institute of Nuclear Agriculture (BINA), Bangladesh Agricultural University (BAU), Sher-e-Bangla Agricultural University (SAU), and Bangladesh Agricultural Development Corporation (BADC). Despite the number of developed varieties, many released selections have issues, such as prolonged growth and stress sensitivity. Farmers are employing the short-duration Tori 7 variety, even if it is extremely poor-yielding, to fit into the transplanted Amon-Mustard-Boro cropping pattern, as the country has insufficient short-duration, high-yielding varieties. However, solving this problem can come from the production of short-duration, high-yielding hybrid cultivars by selecting the best parents for future hybridization programs.

Hybridization is essential for developing crops with specific desirable traits. Before hybridization, it is critical to understand

the genetic diversity among the current accessions. Genetic diversity study serves as the foundation for developing novel varieties and is crucial for the assessment, conservation, and use of germplasm resources (Delfina *et al.*, 2016; Saha *et al.,* 2024). The knowledge of relationships among superior breeding populations and genetic diversity aids in designing effective breeding programs, facilitating the selection of desirable parents for new breeding populations, and sustaining long-term selection gains (Sanwal *et al.*, 2015). Genetically diverse parents are likely to segregate and produce high heterotic crosses, with increased diversity enhancing the probability of yielding high heterotic progeny (Saha *et al.,* 2024).

Multivariate analysis serves as a valuable tool in quantifying genotypic divergence between biological populations. Principal component analysis and cluster analysis are useful techniques for genotypic population classification in biology and for determining the relative contributions of various components to overall divergence, both within and between clusters. Mahalanobis's (1936) D^2 statistic of multivariate analysis is becoming one of the most powerful tools for assessing the degree of genetic divergence among the populations (Indian *et al.*, 2019). Multivariate approaches help identify the best cross combinations by approximating genetic divergence while taking several factors into account at the same time in plant breeding projects. The variability among genotypes for distinct morphological parameters across different crops has already had successful assessments using similar methodologies (Ali *et al.,* 2017; Mecha *et al.,* 2017; Ilieva *et al.,* 2019).

Thus, the specific objectives of the presented study include assessing genetic divergence in mustard genotypes, classifying them based on divergence, identifying contributing characteristics, and selecting divergent parents for superior segregates in hybridization programs.

MATERIALS AND METHODS

Experimental site

The study commenced at the experimental field of the Genetics and Plant Breeding Field Laboratory of the Bangabandhu Sheikh Mujibur Rahman Agricultural University (BSMRAU) in November 2022 and continued until February 2023. The experimental site has a setting in a subtropical climatic zone. The soil at the BSMRAU farm belongs to the Salna series of shallow, red-brown terrace soils. It is silty clay loam, with a pH of approximately 6.5, characterized by medium-high elevation and moderate fertility.

Planting materials

The plant materials studied in the said experiment included 51 genotypes of the mustard (*Brassica rapa*) from the Bangabandhu Sheikh Mujibur Rahman Agricultural University and two genotypes from the Bangladesh Agricultural Research Institute (Table 1). These genotypes have their current maintenance in the breeding program of the Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur. The plant materials used in this study contain a collection of breeding lines.

Seed sowing, intercultural operations, and harvesting

Seeds' sowing at a depth of 1.5 cm began in the first week of November, 2022. Fertilizer application had the rates of 250 kg urea, 170 kg TSP, 85 kg muriate of potash, 150 kg gypsum, and 60 kg borax per hectare. Incorporating TSP, MP, gypsum, and borax ensued during the final land preparation, while applying urea was in three equal installments: at basal, before sowing, and at flowering. Three flood irrigations took place at the seedling, flowering, and siliqua-filling stages to ensure optimal growth and as necessary.

Table 1. Genotype number, code name, and origin/sources of collection of 53 *Brassica rapa* genotypes.

Genotype No.	Code Name*	Origin/Sources of collection**	Genotype No.	Code Name*	Origin/Sources of collection**
1	G_1 (BC-09-0001)	BSMRAU	28	G_{28} (BC-09-0068)	BSMRAU
2	G_2 (BC-09-0005)	BSMRAU	29	G_{29} (BC-09-0070)	BSMRAU
3	G_3 (BC-09-0007)	BSMRAU	30	G_{30} (BC-09-0071)	BSMRAU
4	G_4 (BC-09-0009)	BSMRAU	31	G_{31} (BC-09-0072)	BSMRAU
5	G_5 (BC-09-0012)	BSMRAU	32	G_{32} (BC-09-0073)	BSMRAU
6	G_6 (BC-09-0015)	BSMRAU	33	G_{33} (BC-09-0075)	BSMRAU
7	G_7 (BC-09-0017)	BSMRAU	34	G_{34} (BC-09-0078)	BSMRAU
8	G_8 (BC-09-0019)	BSMRAU	35	G_{35} (BC-09-0080)	BSMRAU
9	G_9 (BC-09-0021)	BSMRAU	36	G_{36} (BC-09-0084)	BSMRAU
10	G_{10} (BC-09-0027)	BSMRAU	37	G_{37} (BC-09-0086)	BSMRAU
11	G_{11} (BC-09-0029)	BSMRAU	38	G_{38} (BC-09-0088)	BSMRAU
12	G_{12} (BC-09-0033)	BSMRAU	39	G_{39} (BC-09-0092)	BSMRAU
13	G_{13} (BC-09-0035)	BSMRAU	40	G_{40} (BC-09-0094)	BSMRAU
14	G_{14} (BC-09-0037)	BSMRAU	41	G_{41} (BC-09-0099)	BSMRAU
15	G_{15} (BC-09-0040)	BSMRAU	42	G_{42} (BC-09-0104)	BSMRAU
16	G_{16} (BC-09-0042)	BSMRAU	43	G_{43} (BC-09-0105)	BSMRAU
17	G_{17} (BC-09-0043)	BSMRAU	44	G_{44} (BC-09-0107)	BSMRAU
18	G_{18} (BC-09-0044)	BSMRAU	45	G_{45} (BC-09-0110)	BSMRAU
19	G_{19} (BC-09-0047)	BSMRAU	46	G_{46} (BC-09-0114)	BSMRAU
20	G_{20} (BC-09-0049)	BSMRAU	47	G_{47} (BC-09-0119)	BSMRAU
21	G_{21} (BC-09-0052)	BSMRAU	48	G_{48} (BC-09-0120)	BSMRAU
22	G_{22} (BC-09-0055)	BSMRAU	49	G_{49} (BC-09-0122)	BSMRAU
23	G_{23} (BC-09-0058)	BSMRAU	50	G_{50} (BC-09-0123)	BSMRAU
24	G_{24} (BC-09-0061)	BSMRAU	51	G_{51} (BC-09-0124)	BSMRAU
25	G_{25} (BC-09-0063)	BSMRAU	52	G_{52} (BARI sarisha 12)	BARI
26	G_{26} (BC-09-0064)	BSMRAU	53	G_{53} (BARI sarisha 15)	BARI
27	G_{27} (BC-09-0066)	BSMRAU			

* G=Genotype, **BSMRAU= Bangabandhu Sheikh Mujibur Rahaman Agricultural University, BARI= Bangladesh Agricultural Research Institute.

Additional agronomic practices included drainage, earthing-up, thinning at 10 and 17 days after sowing (DAS), and gap-filling as needed. The treatment of Malathion 57 EC transpired one month after sowing at 12-day intervals (three times) at 1 ml per 2.5 l of water to manage aphid infestations. Rovral-50 WP spraying was at 20 g per 10 l of water during siliqua setting and 15 days later to control the *Alternaria* leaf spot. Although no significant disease outbreaks manifested. Harvesting occurred in mid-February when 80% of the plants reached maturity.

Experimental design and layout

The study implemented a randomized complete block design (RCBD) with three replications. Each plot consisted of two rows, each measuring 4 m in length. The distances between block-to-block, row-to-row, plot-to-plot, and plant-to-plant are 50, 30, 50, and 10 cm, respectively. The assigning of genotypes and treatments was random to experimental plots, with intercultural operations carried out to promote optimal growth and development of the *Brassica* seedlings.

Data recorded

The study involved collecting data on 21 different traits. For instance, recording the days to first flowering (DFF), days to 50% flowering (DFPF), days to 100% flowering (DHPF), and days to last flowering (DLF) progressed by noting the intervals between sowing and specific flowering stages. Measuring the days to maturity (DM) was the

time length from sowing to 80% siliquae maturity. Plant height (PH) as determined at harvest was by averaging measurements from 10 randomly selected plants per plot. Branching patterns' assessment comprised counting primary branches per plant (PBPP) and secondary branches per plant (SBPP). Leaf characteristics, including leaf length (LL), leaf width (LW), and petiole length (PL), underwent measurement from the largest leaves of 10 selected plants per plot. Determining the root length (RL) and the length of the main racemose (LMR) consisted of averaging measurements from 10 randomly selected plants per plot. Additionally, recording the number of pods on the main racemose (NPMR), length of siliquae (LS), beak length (BL), number of siliquae per plant (NSPP), number of seeds per silique (NSPS), thousand seed weight (TSW), and seed yield per plant (SYPP) succeeded.

The Harvest Index calculation was according to Radford (1967), as follows: Harvest Index =

$$\frac{\text{Seed weight}}{\text{Total plant weight}} \times 100$$

Total plant weight's computation comprised adding the seed weight and straw weight, i.e., Total plant weight = seed weight + straw weight

Straw weight had the selected plants sun-dried for a few days.

Statistical analysis

The statistical analysis of data for various characteristics revealed significant differences among Brassica genotypes, requiring data from all the characteristics subjected to multivariate analysis.

Multivariate analysis proceeded using the GENSTAT program. Using the least significant difference (LSD) test helps determine the significance of the difference between the treatment means at a 5% probability level. Genetic diversity assessment used Mahalanobis's (1936) generalized distance (D^2), as further elaborated by Rao (1952).

RESULTS AND DISCUSSION

Univariate analysis

The ANOVA results in Table 2 reveal highly significant variations in nine of the 21 studied characteristics. These indicate a wide scope of selection for these characteristics, i.e., the data expressed substantial variability and thus a high possibility of improvement in most traits.

Multivariate analysis

Principal component analysis

The principal component analysis (PCA) identified 21 principal component axes for coordinating genotypes, with the first axis accounting for 43.73% of the total variation. The first two axes together explained 58.27% of the variation, while five eigenvalues above unity captured 83.05% (Table 3).

A two-dimensional chart (Z1-Z2) attained construction based on PCA scores I and II. Scatter distribution of 53 *Brassica rapa* genotypes based on principal component scores superimposed with clustering and the superimposed scatter diagram revealed seven apparent clusters, with genotypes distantly located from each other (Figure 1).

Principal coordinate analysis

The principal coordinate analysis (PCoA), conducted as a complement to the principal component analysis, revealed inter-genotypic distances. The greatest distance (2.2) was notable between genotypes 19 and 42, followed by the genotypes 1 and 42 (2.1), 10 and 42 (2.1), and 21 and 42 (2.0). The smallest distance was evident between genotypes 22 and 29 (0.2), followed by 13 and 24 (0.2), 20 and 39 (0.2), and 11 and 36 (0.2) (Table 4). The difference between the highest and the lowest inter-genotypic distances indicated the prevalence of variability among the 53 genotypes of *Brassica rapa* studied.

Table 2. Analysis of variance for 21 characteristics of 53 genotypes of *Brassica rapa.*

Characteristics	Block	Genotype	Error	CV%	LSD (5%)
Days to 1st flowering (DFF)	0.15	4.03	3.36	5.97	3.68
Days to 50% flowering (DFPF)	1.36	14.99**	5.36	6.22	4.65
Days to 100% flowering (DHPF)	0.24	12.71**	6.45	5.35	5.1
Days to last flowering (DLF)	0.09	10.76**	0.09	3.12	4.63
Days to maturity (DM)	0.95	3.76	2.75	1.96	4.63
Plant height (PH) (cm)	0.29	79.25*	49.43	6.96	14.11
Primary branches per plant (PBPP)	0.0001	1.76	1.82	20.74	2.7
Secondary branches per plant (SBPP)	0.39	0.23	8.85	35.83	5.97
Leaf length (LL) (cm)	0.014	5.26**	1.94	12.35	2.8
Leaf width (LW) (cm)	1.22	3.23**	1.04	12.14	2.05
Petiole length (PL) (cm)	4.43	4.22*	2.34	18.52	3.07
Root length (RL) (cm)	1.56	2.51	1.77	13.64	2.67
Length of main racemose (LMR) (cm)	2.79	20.09	15.59	9.45	7.92
Number of pods on main racemose (NPMR)	7.23	42.06	36.30	14.96	12.09
Length of silique (LS) (cm)	0.012	0.23	0.19	11.22	0.89
Length of beak (LB) (mm)	0.55	5.15**	2.39	14.97	3.11
Number of silique per plant (NSPP)	164.93	1516.47	1420.13	22.94	75.62
Number of seeds per silique (NSPS)	73.19	103.96	111.33	59.81	21.17
Thousand seeds weight (TSW) (g)	0.038	0.15	0.11	10.53	0.66
Harvest index (HI) (%)	14.94	0.49	12.39	11.92	7.06
Seed yield per plant (SYPP) (g)	0.017	12.368**	1.44	22.17	2.41

* indicates significant at 5% level, ** indicates significant at 1% level, CV = Coefficient of variation, LSD = Least significant difference

Table 3. Eigen-values and percentage of variation concerning 21 characteristics of 53 *Brassica rapa* genotypes.

Principal component axis	Eigen-values	Total variation accounted (%)	Cumulative percentage of variation
1	10.309	43.73	43.73
2	3.427	14.54	58.27
3	2.350	9.97	68.24
4	2.234	9.48	77.72
5	1.256	5.33	83.05
6	0.794	3.37	86.42
7	0.696	2.95	89.37
8	0.564	2.39	91.76
9	0.385	1.63	93.39
10	0.341	1.45	94.84
11	0.281	1.19	96.03
12	0.258	1.09	97.12
13	0.203	0.86	97.98
14	0.138	0.59	98.57
15	0.134	0.57	99.14
16	0.073	0.31	99.45
17	0.050	0.21	99.66
18	0.042	0.17	99.84
19	0.026	0.11	99.94
20	0.011	0.05	99.99
21	0.003	0.01	100.00

Table 4. The 10 of each lower and higher intercluster distances (D^2 values) between pairs of *Brassica rapa* genotypes.

Genotype pair	10 lower D^2 values	Genotype pair	10 higher D^2 values
22-29	0.2	19-42	2.2
13-24	0.2	1-42	2.1
20-39	0.2	10-42	2.1
11-36	0.2	21-42	2.0
20-29	0.2	26-42	2.0
20-22	0.2	8-42	2.0
30-32	0.2	3-19	2.0
24-31	0.3	19-41	1.9
20-24	0.3	10-42	1.9
20-27	0.3	4-42	1.9

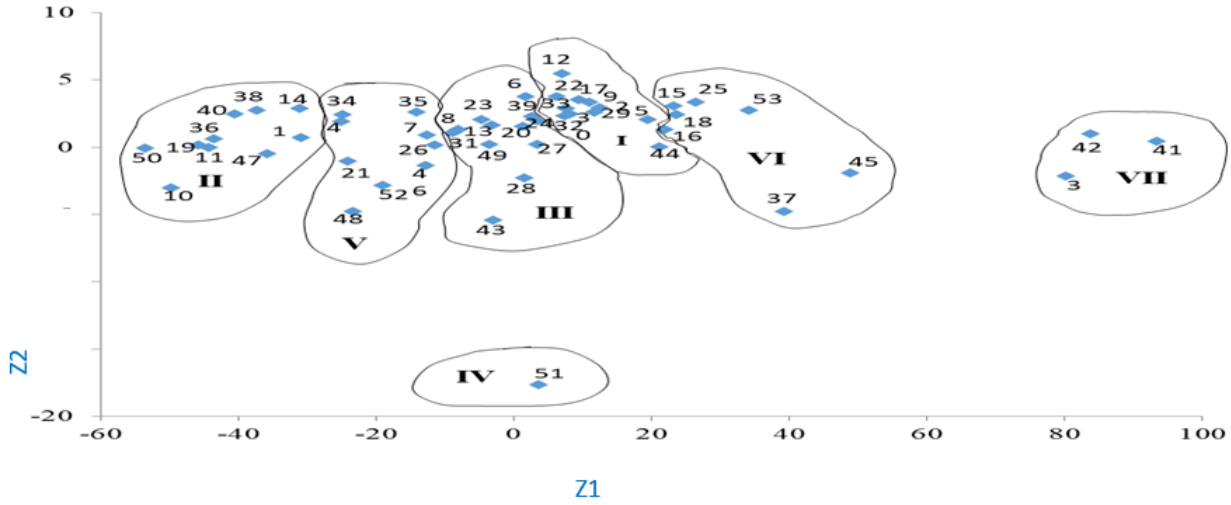

Figure 1. Scatter distribution of 53 *Brassica rapa* genotypes based on principal component scores 1 (Z1) and 2 (Z2) super-imposed with clustering.

Non-hierarchical clustering

A non-hierarchical clustering method, utilizing the covariance matrix, as employed, led to the classification of 53 genotypes into seven distinct clusters. Jahan *et al.* (2013) and Rahman *et al.* (2022) observed four clusters in 18 genotypes and five clusters in 14 genotypes in *Brassica rapa*, respectively. Rout *et al.* (2019) reported seven clusters in 71 genotypes of mustard. These findings corroborate the clustering pattern of the genotypes derived from the PCA, further supporting the validity of the results.

Figure 1 illustrates that genotypes from the same location may be in different clusters, while those from different locations may cluster together, suggesting no correlation between geographic and genetic diversity. Moreover, Cluster III had the highest number of genotypes (12), followed by Clusters I (11), II (10), and V (9). Cluster IV contained a single genotype, and Cluster VII included three genotypes.

Cluster I exhibited the second-highest mean for LB (10.73 mm) and the third-highest means for SBPP (8.50), LL (11.50 cm), LW (8.59 cm), and harvest index (30.11%). It displayed the lowest mean for DFF (30.00) and DHPF (46.23) (Table 5).

Table 5. Cluster means for 21 characteristics of *Brassica rapa* genotypes.

Characteristics	Cluster Means						
	I	II	III	IV	V	VI	VII
Days to 1st flowering (DFF)	30.00	30.05	30.83	30.50	30.89	31.29	33.17
Days to 50% flowering (DFPF)	36.14	36.05	37.54	39.50	38.22	37.57	38.83
Days to 100% flowering (DHPF)	46.23	47.10	49.17	46.50	47.83	47.14	46.83
Days to last flowering (DLF)	73.36	74.05	74.54	73.00	75.17	74.71	68.33
Days to maturity (DM)	83.77	84.35	85.08	85.00	84.50	83.93	83.67
Plant height (PH) (cm)	96.65	104.95	103.68	102.00	103.43	99.83	88.65
Primary branches per plant (PBPP)	6.27	6.26	6.60	5.25	6.52	6.79	6.77
Secondary branches per plant (SBPP)	8.50	9.93	7.56	6.70	9.16	7.00	5.28
Leaf length (LL) (cm)	11.50	11.61	11.47	11.18	11.65	11.05	7.88
Leaf width (LW) (cm)	8.59	8.72	8.99	8.05	8.52	7.75	6.09
Petiole length (PL) (cm)	7.80	8.26	8.89	9.54	8.86	7.85	5.46
Root length (RL) (cm)	9.58	10.51	9.85	8.34	9.77	9.08	9.00
Length of main racemose (LMR) (cm)	41.55	43.80	42.97	40.33	42.16	40.42	33.41
Number of pods on main racemose (NPMR)	39.90	43.99	42.35	40.95	40.49	38.68	23.75
Length of silique (LS) (cm)	4.01	4.07	3.88	4.47	3.84	4.28	3.37
Length of beak (LB) (mm)	10.73	10.57	10.80	8.60	10.14	10.11	7.22
Number of silique per plant (NSPP)	153.99	205.62	165.48	163.05	183.37	133.57	82.12
Number of seeds per silique (NSPS)	15.76	15.42	16.69	27.76	16.74	17.99	20.79
Thousand seeds weight (TSW) (g)	3.14	3.16	3.07	3.65	3.18	3.03	3.23
Harvest index (HI%)	30.11	30.95	29.98	31.00	29.37	27.85	24.83
Seed yield per plant (SYPP) (g)	5.13	6.18	5.68	5.30	5.73	4.68	3.83

Cluster II had the highest mean values for PH (104.95 cm), SBPP (9.93), RL (10.51 cm), LMR (43.80 cm), NPMR (43.99), NSPP (205.62), and SYPP (6.18 g). It ranked second for LL (11.61 cm), LW (8.72 cm), and harvest index (30.95%), and third for LB (10.57 mm), LS (4.07), and DM (84.35). Conversely, it showed the lowest mean values for DFPF (36.05) and NSPS (15.42) (Table 5).

Cluster III had the supreme mean for DHPF (49.17), DM (85.08), LW (8.99 cm), and LB (10.80 mm). It ranked second for PH (103.68 cm), PL (8.89 cm), RL (9.85 cm), LMR (42.97 cm), and NPMR (42.35). Additionally, it had the third-highest mean for DLF (74.54), PBPP (6.60), NSPP (165.48), and SYPP (5.68 g) (Table 5).

Cluster IV has a sole representation of genotype G51. This cluster recorded the ultimate mean for DFPF (39.50), PL (9.54 cm), LS (4.47 cm), NSPS (27.76), TSW (3.65 g), and harvest index (31%). It ranked second for DM (85.00) and third for NPMR (40.95). However, it had the lowest mean values for PBPP (5.25) and RL (8.34 cm) (Table 5).

Cluster V displayed the highest range for DLF (75.17). It ranked second for DHPF (47.83), SBPP (9.16), LL (11.65 cm), NSPP (183.37), and SYPP (5.73 g). Additionally, it had the third-highest mean for DFF (30.89), DFPF (38.22), PH (103.43 cm), PL (8.86 cm), RL (9.77 cm), LMR (42.16 cm), and TSW (3.18 g) (Table 5).

Cluster VI revealed this group contributed genotypes with the topmost mean for the characteristics PBPP (6.79); the second-highest mean for DFF (31.29), DLF (74.71), and LS (4.28 cm); and the third-highest mean for DHPF (47.14) and NSPP (17.99). The lowest mean value appeared for the character TSW (3.03).

Cluster VII exhibited the maximum mean for DFF (33.17) and the second-highest mean for DFPF (38.83), PBPP (6.77), NSPS (20.79), and TSW (3.23 g). Conversely, it had the lowest mean values for DLF (68.33), DM (83.67), PH (88.65 cm), SBPP (5.28), leaf length (7.88 cm), LW (6.09 cm), PL (5.46 cm), LMR (33.41 cm), NPMR (23.75), LS (3.37 cm), LB (7.22 mm), NSPP (82.12), harvest index (24.83%), and SYPP (3.83 g) (Table 5).

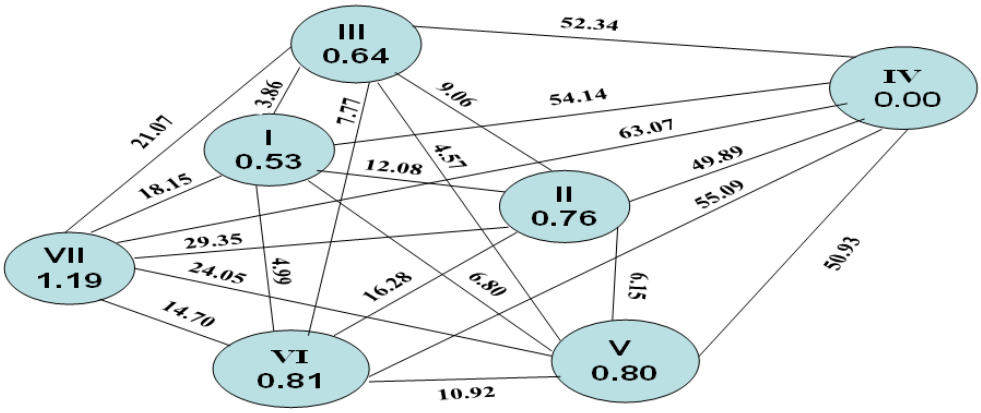

Figure 2. Diagram showing intracluster (inside the circle) and intercluster (outside the circle) distances of 53 *Brassica rapa* genotypes**.**

Canonical variate analysis

The inter-cluster Mahalanobis D^2 values, derived via the Canonical Variate Analysis (CVA), are available in Figure 2 alongside intracluster distances. The greater intercluster distances compared with intracluster distances indicate higher genetic diversity among genotypes across different groups, consistent with findings by Singh *et al.* (1987) in a multivariate study of mustard.

The results revealed that the highest intercluster distance was remarkable between clusters IV and VII (63.07), followed by IV and VI (55.09), I and IV (54.14), and III and IV (52.34) (Figure 2). These higher intercluster distances suggest a broader spectrum of segregating populations when parents from these distant clusters are selected for hybridization programs. Similar findings came from Singh *et al*. (1991), who noted that the greater genetic distances correlate with higher heterosis than clusters with similar genetic distances. In contrast, the lowest intercluster distances were evident between clusters I and III (3.86), followed by III and V (4.57), II and V (6.16), I and V (6.80), III and VI (7.77), and II and III (9.06). It reflects a closer genetic relationship among the genotypes within these clusters (Figure 2).

However, for a practical plant breeder, the objective is not only to obtain high heterosis but also to achieve high-level production. Considering duration and yield, crosses involving cluster IV and cluster VII may exhibit high heterosis for yield. Mian and Bahl (1989) reported that the parents separated by D^2 values of moderate magnitude generally showed higher heterosis. Keeping this in view, it appeared that the crosses between the genotype belonging to cluster IV with cluster I and genotypes in cluster VI with cluster VII could produce high heterosis for yield, as well as for earliness. Likewise, the crosses between the genotype from cluster IV with clusters VI and III can produce a high level of segregating population. Thus, the genotypes belonging to Cluster I and Cluster IV, Cluster III and Cluster IV, and Cluster VI and Cluster VII can be the best selections for future hybridization programs.

Intracluster distances' calculation occurred from the inter-genotypic distances in the distance matrix, as suggested by Singh *et al.* (1987) (Figure 2). The highest intracluster distance, observed in cluster VII (1.19), comprised three genotypes, followed by cluster VI (0.81), which included seven genotypes. The lowest recorded intracluster distance surfaced in cluster IV (0.00), consisting of a single genotype, followed by cluster I (0.53), which included 11 genotypes. These findings indicate that the genotypes within cluster VII were the most genetically diverse, whereas those in cluster IV were the least diverse. This highlights maximum intracluster diversity in cluster VII and minimum diversity in cluster IV.

Contribution of characteristics toward divergence of the genotypes

Table 6 highlights the contribution of various traits to the genetic divergence among genotypes. In Vector I, key traits contributing to genetic divergence along the first axis include DFF (0.23), DFPF (0.17), DHPF (0.05), DM (0.01), NSPS (0.12), and TSW (0.10), all with positive contributions. The remaining 15 traits out of 21 played a minor role, indicated by negative values.

In Vector II, traits contributing significantly along the second axis include SBPP (0.15), LS (0.05), LB (0.12), TSW (0.12), and HI% (0.12), all showing positive values. The other traits had minor roles with negative values.

SYPP exhibited a minor role in genetic divergence in both vectors, as indicated by negative values. In contrast, TSW had positive values in both vectors, signifying its importance as a key trait contributing to the genetic divergence among the *Brassica rapa* genotypes studied.

Comparison of different multivariate techniques

The cluster pattern of D^2 analysis through non-hierarchical clustering has taken care of simultaneous variations in all characteristics under study.

The study found that the D^2 and principal component analysis can be alternative methods to give information regarding the characteristics' contribution toward the mustard genotypes' divergence. The results of different multivariate analyses were superimposed in Figure 2. One can conclude from this figure that all the techniques gave, more or less, similar results, and the different multivariate techniques supplemented and confirmed the results of one another.

Selection of parents for future hybridization

Selecting genetically diverse parents is essential for successful hybridization. East (1936) observed the degree of heterosis in

Table 6. Latent vectors for 21 characteristics in *Brassica rapa* genotypes.

Characteristics	Vector-I	Vector-II
Days to 1st flowering (DFF)	0.23	-0.27
Days to 50% flowering (DFPF)	0.17	-0.40
Days to 100% flowering (DHPF)	0.05	-0.42
Days to last flowering (DLF)	-0.12	-0.33
Days to maturity (DM)	0.01	-0.44
Plant height (PH) (cm)	-0.22	-0.24
Primary branches per plant (PBPP)	-0.00	-0.14
Secondary branches per plant (SBPP)	-0.27	0.15
Leaf length (LL) (cm)	-0.32	-0.08
Leaf width (LW) (cm)	-0.30	-0.08
Petiole length (PL) (cm)	-0.22	-0.32
Root length (RL) (cm)	-0.23	-0.06
Length of main racemose (LMR) (cm)	-0.28	0.01
Number of pods on main racemose (NPMR)	-0.27	0.00
Length of silique (LS) (cm)	-0.05	0.05
Length of beak (LB) (mm)	-0.27	0.12
Number of silique per plant (NSPP)	-0.32	-0.08
Number of seeds per silique (NSPS)	0.12	-0.02
Thousand seeds weight (TSW) (g)	0.10	0.12
Harvest index (HI%)	-0.20	0.12
Seed yield per plant (SYPP) (g)	-0.24	-0.00

hybrids is directly proportional to the genetic differences between parental strains. Crosses between genetically distant parents tend to exhibit higher heterosis (Kaeppler, 2012; Marcón *et al*., 2019). Considering the magnitude of genetic distance, the contribution of different characteristics toward divergence, the magnitude of cluster means for different characteristics, and the agronomic performance of genotypes, the following genotypes are the best options for optimal performance in hybridization programs.

From cluster I, G_5 could be a selection for the lowest DFF and DM; G_2 for the highest RL and NPMR; G_{17} for the lowest DFPF and highest SBPP; G_{32} for the lowest DM; G_{29} and G_{33} for the lowest DHPF; G_{33} for the highest LW; and G_{44} for the highest PBPP. In cluster II, G_1 could be a choice for the lowest DFPF and the highest HI% and SYPP; G_{40} for the highest LL; and G_{47} for the lowest DHPF. From cluster III, G_8 could be an option for the lowest DHPF; only G_{51} of cluster IV could be chosen for the highest LS. In the case of cluster V, G_{34} could be a selection for the lowest DFPF and the highest LMR; G_{46} for the lowest DLF; G_{52} for the highest PH, NSPS, and TSW; and G_{21} for the lowest PL and the lowest LB; G_{25} could be a choice from cluster VI for the earliest DFF and DM. Finally, G_{41} could be an option from cluster VII for the lowest PL. These genotypes may be the best suggestions for future hybridization programs.

In obtaining superior segregants, the suggested targeted crossings can be as follows: crossing G_2 (Cluster I) with G_{51} (Cluster IV) is a good proposal for improved RL and greater NPMR, and for early DM, crossing G_5 and G_{32} (Cluster I) with G_3 and G_{41} (Cluster VII). Crosses between G_{33} and G_{44} (Cluster I) with G_{41} and G_{42} (Cluster VII), respectively, may yield better segregants for higher LW and the maximum PBPP. Additionally, crossing G_{17} (Cluster I) with G_{38} (Cluster II) could enhance segregants for a higher SBPP and the lowest DFPF. In attaining a higher HI% and better SYPP, crossing of G_1 (Cluster II) with G_3 (Cluster VII) would result in the best segregants. G_{40} and G_{47} (Cluster II) could be crossed with G_{15} and G_{16} (Cluster VI) for higher LL and lower DHPF. A higher LS could result in crossing G_{51} (Cluster IV) with G_{42} (Cluster VII). Crossing G_8 (Cluster III) and G_{34} (Cluster V) with G_3 (Cluster VII) could yield superior segregants with shorter DHPF and higher LMR for future hybridization programs. Similarly, crossing G_{13} (Cluster III) with G_{22} (Cluster I) may produce better results for dwarf PH. For traits such as higher PH, NSPS, and TSW, crossing G_{52} (Cluster V) with G_{42} (Cluster VII) would result in the best segregants. G_{21} (Cluster V) with G_{41} (Cluster VII) for achieving lower PL and LB, and G_{25} (Cluster VI) with G_{51} (Cluster IV) for early DFF and DM. Crossing G_{27} (Cluster III) with G_{41} (Cluster VII) could improve DLF, while G_{29} and G_{33} (Cluster I) crossed with G_{10} (Cluster II) may effectively shorten DHPF.

CONCLUSIONS

In the potential study, we assessed the genetic diversity among 53 *Brassica rapa* genotypes. Significant genetic variability was prominent between clusters, particularly between clusters IV and VII, but less so within the same clusters. Genotypes within the same clusters shared closer genetic relationships, while genotypes from different clusters showed potential as diverse parents for hybridization. These diverse genotypes can be applicable to broaden the genetic base of mustard breeding programs, enabling the development of superior, high-yielding, short-duration hybrid varieties tailored to Bangladesh's agricultural needs.

ACKNOWLEDGMENTS

The authors are grateful to BSMRAU and BARI authorities for providing the seed materials.

REFERENCES

Ali N, Farhatullah NU, Rabbani MA, Hussain I, Ali S, Khan SA, Kakar MQ (2017). Genetic diversity of *Brassica rapa* L. indigenous landraces based on cluster and principal component analyses. *Pak. J. Bot*. 49(5): 1891-901.

Ali N, Qayyum A, Sher A, Nadeem F, Fiaz S, Kumar S, Alharthi B, Ercisli S (2024). Analyzing phenotypic diversity in *Brassica rapa* genotypes based on morpho-physiological characteristics using multivariate analyses. *Genet. Resour. Crop Evol*. 71(5): 2053-64. https://doi.org/10.1007/s10722-023-01757-8.

BBS (2021). Statistical Yearbook. 41th ed. Statistics Division, Ministry of Planning. Government of Peoples Republic of Bangladesh.

Delfina B, Alessandro T, Francesca D, Andrea V, Patrizia V, Giampiero V, Luigi C (2016). Next generation breeding. *Plant Sci*. 242: 3–13. https://doi.org/10.1016/j.plantsci.2015.07.010.

Dmitry A, German K, Hendriks P, Marcus A, Koch F, Lens F, Lysak MA, Donovan C, Bailey A, Mummenhoff K, Ihsan A, Al-Shehbaz (2023). An updated classification of the Brassicaceae (Cruciferae). *PhytoKeys* 8: 220. doi: 10.3897/phytokeys.220.97724.

East EM (1936). Heterosis. *Genetics.* 21(4): 375–397.

Francis A, Lujan-Toro BE, Warwick SI, Macklin JA, Martin SL (2021). Update on the Brassicaceae species checklist. *Biodivers.* Data J. 9. https://doi.org/10.3897/BDJ.9.e65310.

Howlett BG, Butler RC, Walker MK, Teulon DA (2018). Are insect flower visitor assemblages distinguishable between *Brassica napus* and *B. rapa*?. *N. Z. Plant Prot*. 71: 189-97. https://doi.org/10.30843/nzpp.2018.71.162.

Ilieva V, Markova Ruzdik N, Mihajlov L, Ilievski M (2019). Assessment of agro-morphological variability in rice using multivariate analysis. *J. Agric. Plant Sci*. 17(1): 79–85.

Indian G, Sankaranarayanan R, Murugesan S, Rajangam J (2019). Assessment of genetic divergence using Mahalanobis D2 analysis in mango. *J. Agric. Ecol.* 7: 38–46. http://doi.org/10.53911/JAE.

Jahan N, Bhuiyan SR, Talukder MZA, Alam MA, Parvin M (2013). Genetic diversity analysis in *Brassica rapa* using morphological characters. *Bangladesh J. Agric. Res.* 38(1): 11–18.

Kaeppler S (2012). Heterosis: Many genes, many mechanisms—end the search for an undiscovered unifying theory. *ISRN Bot.* 2012(1): 682824. https://doi.org/10.5402/2012/682824.

Mahalanobis PC (1936). On the generalized distance in statistics. *Proc. Natl. Sc. India*. 2: 49–55.

Marcón F, Martínez EJ, Rodríguez GR, Zilli AL, Brugnoli EA, Acuña CA (2019). Genetic distance and the relationship with heterosis and reproductive behavior in tetraploid bahiagrass hybrids. *Mol. Breed.* 39: 1–13. http://doi.org/10.1007/s11032-019-0994-3.

Mecha B, Alamerew S, Assefa A, Assefa E, Dutamo D (2017). Genetic diversity based on multivariate analysis for yield and its contributing characters in bread wheat (*Triticum aestivum* L.) genotypes. *Agric. Res. Technol.* 8(5): 1–10.

Mian MAK, Bahl PN (1989). Genetic divergence and hybrid performance in chickpea. *Indian J. Genet.* 49: 119–124.

Nehmeh M, Rodriguez-Donis I, Cavaco-Soares A, Evon P, Gerbaud V, Thiebaud-Roux S (2022). Bio-refinery of oilseeds: Oil extraction, secondary metabolites separation towards protein meal valorisation—A review. *Processes*. 10(5): 841. https://doi.org/10.3390/pr10050841.

Pal D, Singh D, Kumar A, Singh S (2020). Seasonal incidence of insect pests on rapeseed-Mustard. *Int. J. Curr. Microbiol*. *Appl. Sci.* 9(4): 2525–2531.

Radford PJ (1967). Growth analysis formulae-their use and abuse 1. *Crop Sci.* 7(3): 171–175.

Rahman J, Sultana F, Fatima K, Hasan MM, Gain N, Hossain MS, Chowdhury AK, Rahman A (2022). Genetic diversity of field mustard (*Brassica rapa* L.) and their saturated and unsaturated fatty acids association. *SABRAO J. Breed. Genet.* 54 (2): 249–266. http://doi.org/10.54910/sabrao2022.54.2.4.

Rakow G (2004). Species origin and economic importance of *Brassica*. In: E.C. Pua and C.J. Douglas (Eds) Brassica. Biotechnology in Agriculture and Forestry, Vol 54. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-06164-0_1.

Rao CR (1952). Advanced Statistical Method in Biometrical Research. John Wiley and Sons. New York.

Rao P, Avtar R, Kumari N, Jattan M, Rani B, Sheoran RK (2017). Multivariate analysis in Indian mustard genotypes for morphological and quality traits. *Electron. J. Plant Breed*. 8(2): 450–458.

Rout S, Das A, Sur B, Umamaheswar N, Sadhu S, Ghimiray TS, Mondal HA, Hijam L, Chakraborty M, Roy SK (2019). Genetic divergence in Indian mustard [*Brassica juncea* (L.) Czern and Coss] under sub-Himalayan region. *J. Crop and Weed.* 15(2): 18–24.

Rupa NI, Bhuiyan MS, Rashid MH, Parveen S, Arif MA, Akter S (2023). Evaluation of the potentiality of eight advanced populations of *Brassica rapa* L. based on yield and yield-contributing characters. *SAARC J. Agric*. 21(2): 109–118. https://doi.org/10.3329/sja.v21i2.70320.

Saha SJ, Chowdhury AK, Rashad MM, Alam MM, Farhan SMM (2024). Genetic variability in exotic cowpea (*Vigna unguiculata* L.) genotypes. *Asian J. Med. Biol. Res.* 10(4):148-56. https://doi.org/10.3329/ajmbr.v10i4.74668.

Sanwal SK, Singh B, Singh V, Mann A (2015). Multivariate analysis and its implication in breeding of desired plant type in garden pea (*Pisum sativum*). *Indian J. Agric. Sci.* 85(10): 1298–1302.

Sharmin S (2016). Comparative and genetic study of the advanced populations of *Brassica rapa* L. MS Thesis. Sher-E-Bangla Agricultural University. Dhaka. Bangladesh.

Singh P, Khera MK, Gupta VP (1991). Variability and correlation studies for oil and seed yield in gobhi sarson. *Crop Improv.* 18(2): 99–102.

Singh R, Singh D, Chaudhury B (1987). Morphological variation in Indian mustard. *Annal. Biol.* 3(1): 26–31.

Sultana F (2017). Genetic diversity and fatty acid composition analysis of mustard (*Brassica rapa* L.). MS Thesis. Sher-E-Bangla Agricultural University. Dhaka. Bangladesh.

Walden N, German DA, Wolf EM, Kiefer M, Rigault P, Huang XC, Kiefer C, Schmickl R, Franzke A, Neuffer B, Mummenhoff K (2020). Nested whole-genome duplications coincide with diversification and high morphological disparity in Brassicaceae. *Nat. Commun.* 11(1): 3795. https://doi.org/10.1038/s41467-020-17605-7.