

GENETIC DIVERSITY OF EXOTIC WHITE MAIZE GENOTYPES THROUGH PRINCIPAL COMPONENT AND CLUSTER ANALYSIS FOR YIELD-RELATED TRAITS

GOUTOM ROY, BODRUN NESSA SHOMPA, IRIN AKTER EMU,
NILOY GAIN AND JAMILUR RAHMAN*

*Department of Genetics and Plant Breeding, Faculty of Agriculture,
Sher-e-Bangla Agricultural University, Sher-e-Bangla Nagar, Dhaka-1207, Bangladesh*

Keywords: White maize, Genetic diversity, D^2 statistics, Cluster analysis, Principal component analysis

Abstract

Genetic variation is a fundamental requirement in crop improvement programs. The study evaluated 17 white maize genotypes to analyze genetic diversity across 13 yield-related parameters at the farm of Sher-e-Bangla Agricultural University, Dhaka, Bangladesh. Substantial genetic variation was observed among the traits of interest, and the first three principal components had eigenvalues greater than 1, accounting for 87.64% of the total variation. The cluster phylogeny classified the genotypes into five clusters, with genotypes between clusters I and V, followed by clusters I and II, showing the highest inter-cluster distance, indicating that genotypes in these clusters were more diverse than others. Comparatively, genotypes in cluster V showed short plant height, whereas clusters I and II showed the highest yield-contributing performance. Collectively, genotype CLTHW15008 from cluster I and genotypes CLTHW15004, CLTHW15005, CLTHW15006, CLTHW15011, CLTHW15001, and CLTHW15015 from cluster II could be selected as potential parents to develop vigorous maize hybrids.

Introduction

Maize (*Zea mays* L.) is one of the most widely cultivated cereal crops, playing a crucial role in food security, nutrition, and industrial applications. It exhibits substantial genetic variation due to its domestication history and adaptation to diverse agroecological conditions, particularly variations in endosperm characteristics, such as hardness and color, pericarp, starch type, and kernel structure (García-Lara *et al.* 2019). Among its diverse types, yellow and white maize cultivars are widely grown worldwide. White maize is particularly significant in Africa, Latin America, and parts of Asia, serving as a staple food and key ingredient in traditional diets (Arora *et al.* 2024). Exploring exotic white maize inbred lines can introduce novel alleles into breeding programs, thereby improving yield potential, stress tolerance, and disease resistance (Menkir *et al.* 2018). Hence, understanding genetic diversity through various approaches is crucial for maintaining genetic resources, introducing promising varieties, and optimizing heterosis breeding of hybrids.

Multivariate and cluster analyses are widely employed to assess genetic diversity and relationships among germplasm collections. These methods enable researchers to analyze multiple traits simultaneously, reducing data complexity while preserving essential variability information (Adebayo *et al.* 2021). Among these techniques, PCA identifies the most influential traits contributing to diversity, describes the pattern of variation among characters in an individual, and maps multi-dimensional relationships (Nwangburuka *et al.* 2011). Additionally, D^2 -based clustering classifies genotypes into distinct groups and provides insights into inter- and intra-cluster distances, aiding in the selection of genetically divergent parental lines (Antony *et al.* 2021).

*Author for correspondence: <jamilur@sau.edu.bd>.

Maize is a short-duration crop cultivated in the spring and summer seasons to achieve desirable yields and boost the economy (Dogar *et al.* 2023). In Bangladesh, maize cultivation covered 1,227 thousand acres, yielding 4.563 million tons during the 2022-2023 season (BBS 2023). Both white and yellow maize are cultivated in Bangladesh, whereas white maize is particularly favored in human diets because carotenoid degradation during cooking processes such as baking or frying imparts a distinct aroma and flavor (Malvar *et al.* 2008). A remarkable surge in the demand for white maize flour has been noticed in recent years in India and Bangladesh (Roy *et al.* 2025). The rising demand for white maize, owing to its nutritional benefits, contrasts with Bangladesh's limited domestic production and reliance on substantial imports, highlighting the urgent need to introduce high-yielding and regionally adaptable white maize cultivars. Considering this, the present study evaluated 17 exotic white maize inbred lines using multivariate and cluster analyses, which help identify promising inbred lines with desirable traits and simultaneously broaden the genetic pool of white maize for future hybridization programs.

Materials and Methods

The experiment was conducted at the farm of Sher-e-Bangla Agricultural University, Dhaka, to evaluate seventeen exotic white maize inbred lines viz., CLTHW15004, CLTHW15005, CLTHW15006, CLTHW15007, CLTHW15008, CLTHW15009, CLTHW15003, CLTHW15010, CLTHW15011, CLTHW15012, CLTHW15001, CLTHW15014, CLTHW15015, CLTHW14001, CLTHW14003 and CLTHW13001, collected from CIMMYT, Mexico and a local genotype Suvra from BARI, Gazipur. The seeds of the genotypes were planted in lines in the well-prepared soil on 16th November 2017 under a Randomized Complete Block Design (RCBD) with three replications, having the line-to-line and hill-to-hill distances of 75 cm and 25 cm, respectively. The field was grouped into three blocks, maintaining an individual block size of 33.15 m × 2.5 m. The plants were spaced 25 cm apart within the same row and 65 cm between rows. All agronomic operations were performed as required. Moreover, a summary of weather conditions is presented in Table 1. Fifteen plants from each line were randomly selected for recording data on plant height (cm), base to cob distance (cm), base diameter of plant (cm), leaf length (cm), leaf breadth (cm), days to 50% flowering (DAS), days to maturity, cob length (cm), cob diameter (cm), number of rows cob⁻¹, number of grains row⁻¹, 100-grain weight (g), and yield plant⁻¹ (g). Analysis of variance (ANOVA) and multivariate analysis were done following Mahalanobis' D² statistics (1936), and cluster analysis, was performed using GENSTAT 5.13 software.

Table 1. Monthly weather conditions at the experimental site during the study period.

Month	Maximum temp. (°C)	Minimum temp. (°C)	Relative humidity (%)	Total rainfall (mm)	Sunshine (h)
November 2017	32.40	18.00	77.00	0.00	5.80
December 2017	29.10	16.30	69.00	227.00	7.90
January 2018	28.10	13.00	79.00	0.00	3.90
February 2018	28.10	11.10	72.00	10.00	5.70

Bangladesh Meteorological Department (Climate Division), Agargaon, Dhaka-1212.

Results and Discussion

Thirteen yield-attributing traits, exhibited substantial genetic variation through variance analysis, suggesting that genetic improvement utilizing these breeding materials is achievable (Table 2). Evaluation of mean performance revealed that plant height among the inbred lines ranged from 172.67 cm (CLTHW14001) to 223.93 cm (CLTHW15010). Conversely, the maturity duration extended from 129.00 to 151.67 DAS, whereas the lowest duration was noted in

CLTHW15012 (129.00 DAS), CLTHW15003 (129.33 DAS), and CLTHW14001 (129.33 DAS), highlighting that these lines could be utilized as early-matured varieties (Fig. 1). Focusing on the yield-attributing traits, the highest cob length (16.47 cm), cob diameter (4.85 cm), number of grains row⁻¹ (30.14), number of rows cob⁻¹ (16.57), 100-grain weight (34.10 gm), and yield plant⁻¹ (170.01 gm) were observed in the CLTHW15008 genotype. Ahmed *et al.* (2020), Olawamide and Fayeun (2020), and Shompa *et al.* (2020) also noted substantial variations in maize genotypes. Based on the agronomic traits' performance, it can be concluded that CLTHW15008, CLTHW15004, CLTHW15005, CLTHW15006, CLTHW15001, and CLTHW15015 are the most promising high-yielding inbreds of white maize.

Table 2. Analysis of variance (ANOVA) of different characters in white maize genotypes.

Sources of variation	df	Mean sum of squares												
		PH	DBC	BD	LL	LB	DFF	DM	CL	CD	RPC	GPR	HGW	YPP
Genotypes	16	604.78**	369.84**	0.82	63.47**	1.44	24.04**	96.46**	2.97**	0.13	3.93**	29.42**	34.57**	15505.68**
Replications	2	1079.99	181.02	1.31	177.36	0.74	12.51	55.90	0.09	0.01	0.03	2.01	115.01	300.23
Error	32	385.93	6.12	0.01	59.02	0.02	0.09	17.86	0.10	0.01	0.09	0.45	0.01	77.48

PH: Plant height (cm), DBC: Distance between base and cob (cm), BD: Base diameter of plant, LL: Leaf length, LB: Breadth of leaf (cm), DFF: Days to 50% flowering, DM: Days to maturity, CL: Cob length (cm), CD: Cob diameter (cm), RPC: No. of rows cob⁻¹, GPR: No. of grains row⁻¹, HGW: 100-grains weight (gm), and YPP: Yield plant⁻¹ (gm).

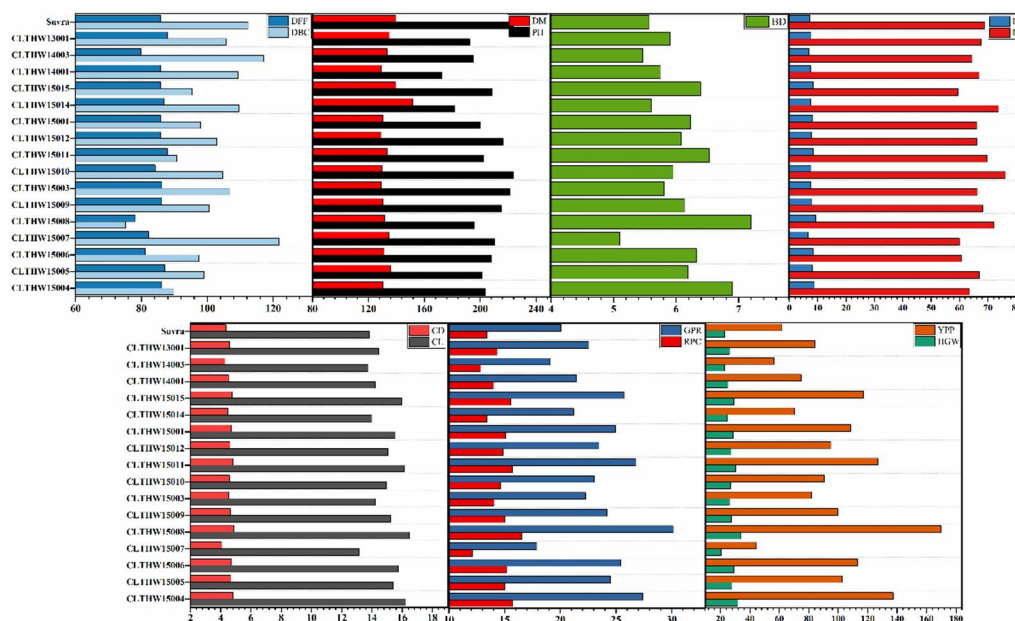


Fig. 1. Mean performance of various growth parameters and yield components of white maize genotypes.

DFF: Days to 50% flowering, DBC: Distance between base and cob (cm), PH: Plant height (cm), DM: Days to maturity, BD: Base diameter of plant, LL: Leaf length, LB: Breadth of leaf (cm), CL: Cob length (cm), CD: Cob diameter (cm), RPC: No. of rows cob⁻¹, GPR: No. of grains row⁻¹, HGW: 100-grain weight (gm) and YPP: Yield plant⁻¹ (gm).

Knowledge of genetic diversity among the genotypes is essential for selecting suitable parents for the hybridization program, particularly in a cross-pollinated crop like maize (Doggalli *et al.* 2024). In the present study, the eigenvalues, per cent of total variation, and cumulative variation

accounted for thirteen traits of seventeen white maize genotypes obtained from principal component analysis are presented in Table 3. The first principal component (PC1), which accounts for 68.21% of the total variation, is predominantly associated with plant height. This suggests that plant height is a major determinant of genetic diversity among the studied genotypes. Similarly, the second principal component (PC2) contributes 11.20% of the variation and is strongly linked to the distance between base to cob. The first three components exhibited eigenvalues greater than one, suggesting that they had the largest contribution to diversifying the genotypes. The cumulative variance explained by the first three PCs reached 87.64%, indicating that these components capture most of the total variability, a trend commonly observed in crop trait evaluations (Ghafoor *et al.* 2022, Singh *et al.* 2023). Among the measured traits, leaf length and leaf breadth also contributed to the observed variation, explaining an additional 6.79% and 4.65% of the variance, respectively. This suggests that leaf morphology plays a secondary yet notable role in genotype differentiation. Therefore, these four traits could be considered for selecting superior genotypes in the breeding program.

Table 3. Eigenvalues, variance percentage, and the relative contribution of 13 characters towards divergence.

Principal components	Eigenvalues	Percent variation	Cumulative Percent variation	Vector-1	Vector-2
Plant height (cm)	8.87	68.21	68.21	0.01	-0.46
Base to cob distance (cm)	1.46	11.20	79.41	0.33	-0.03
Base diameter of plant (cm)	1.07	8.23	87.64	-0.33	-0.02
Leaf length (cm)	0.88	6.79	94.43	-0.04	0.47
Leaf breadth (cm)	0.61	4.65	99.08	-0.33	0.03
Days to 50% flowering (DAS)	0.07	0.50	99.58	0.01	0.44
Days to maturity (DAS)	0.02	0.19	99.77	0.09	0.59
Cob length (cm)	0.02	0.13	99.90	-0.33	-0.03
Cob diameter (cm)	0.01	0.07	99.97	-0.32	0.12
No. of rows cob ⁻¹	0.00	0.02	99.99	-0.33	0.01
No. of grains row ⁻¹	0.00	0.01	100.00	-0.34	0.01
100-grain weight (gm)	0.00	0.00	100.00	-0.33	0.01
Yield plant ⁻¹ (gm)	0.00	0.00	100.00	-0.33	-0.02

The contributions of characters to the divergence obtained from canonical variates analysis are presented in Table 3. In this study, the loadings of agronomic traits on Vector-1 and Vector-2 suggested their relative importance in explaining genetic variability. In Vector-1, positive values were observed for plant height (0.01), distance between base to cob (0.33), days to 50% flowering (0.01), and days to maturity (0.09). In vector 2, positive values were noted for leaf length (0.47), leaf breadth (0.03), days to 50% flowering (0.44), days to maturity (0.59), cob diameter (0.12), number of rows per cob (0.01), number of grains per row (0.01), and 100-grain weight (0.01). Days to maturity and days to 50% flowering showed positive values for both the vectors, revealing the highest contribution towards divergence among the genotypes. On the other hand, negative values were observed in the base diameter of plant and cob length in both vectors, which had lower contributions toward the divergence. The present findings are consistent with Fatema *et al.* (2019) and Islam *et al.* (2019). For crop improvement, the selection of parental lines should be based on the agronomic performance of genotypes and on traits that contribute most to genetic divergence, as these factors increase the likelihood of developing superior recombinants.

Based on the variation in mean performance, cluster analysis following their relative D² values separated the seventeen maize inbred lines into five clusters. Within the five clusters, the

maximum number of inbred lines were formed in cluster II, which included six genotypes. On the other hand, clusters III, IV, and V each contained four, three, and three genotypes, respectively. However, cluster I had only one genotype (Fig. 2). The genetic distance obtained from the cluster distance analysis provides insightful information on the genetic diversity and potential heterotic patterns among the inbred lines. In the current study, the magnitude of divergence varied substantially among cluster pairs (Table 4). The highest inter-cluster D^2 values were noted between clusters I and V (58.22), followed by I and II (44.91) as well as I and III (44.35), suggesting that the genotypes in these clusters possessed the maximum phenotypic differentiation. The wide divergence between these clusters indicated their potential as complementary parental groups for hybridization, where crossing genetically divergent inbreds may generate broader genetic variation and increase the possibility of recovering desirable recombinants (Khan *et al.* 2022, Doggalli *et al.* 2024). In contrast, the minimum inter-cluster distance was observed between clusters II and III ($D^2 = 9.18$), followed by clusters V and III ($D^2 = 11.35$), indicating comparatively greater similarity in their trait profiles. Therefore, the observed D^2 pattern can be used as an initial criterion for strategic parent selection. Genotypes included in relatively narrow groups may be exploited through selection breeding to accumulate and fix desirable traits. In contrast, highly divergent and agronomically complementary inbreds can be prioritized for hybridization, followed by evaluation of progeny performance and heterosis.

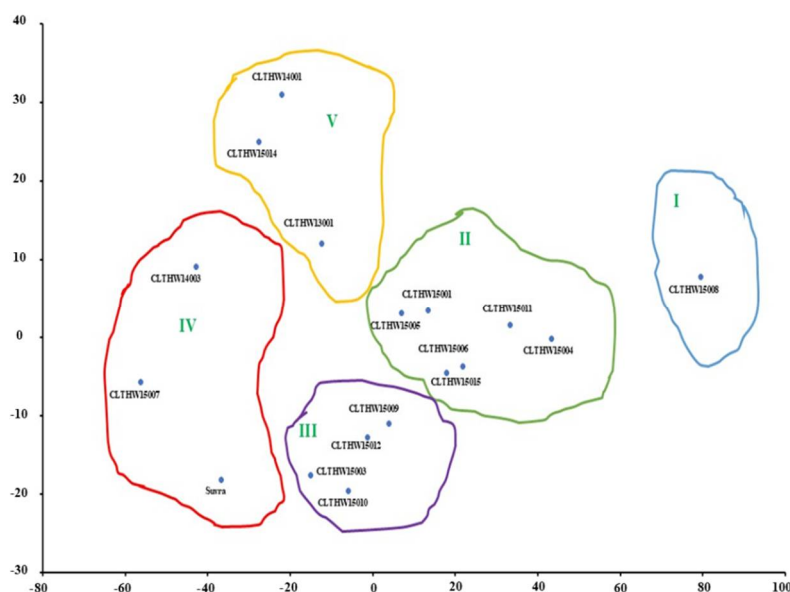


Fig. 2. Grouping of seventeen white maize genotypes into five clusters based on thirteen characters.

Table 4. Nearest and farthest clusters from each cluster between D^2 values of white maize genotypes.

Cluster	Nearest with D^2 values	Farthest with D^2 values
I	IV (35.73)	V (58.22)
II	III (9.18)	I (44.91)
III	II (9.18)	I (44.35)
IV	II (14.87)	I (35.73)
V	III (11.35)	I (58.22)

The cluster means of thirteen traits showed considerable differences between clusters for all the characters studied (Table 5). The genotypes in cluster III recorded the highest mean plant height (219.29), while those in cluster V showed the lowest (182.36). Maximum base to cob distance (7.2), leaf length (72.21), leaf breadth (9.39), cob length (16.47), cob diameter (4.85), number of rows cob⁻¹ (16.57), number of grains row⁻¹ (30.14), 100-gram weight (34.1), and yield plant⁻¹ (170.01) were observed in cluster I. On the contrary, cluster IV showed the lowest mean value for cob length (13.58), cob diameter (4.23), number of grains row⁻¹ (19.02), 100-gram weight (22.16), and yield plant⁻¹ (54.28). Cluster V showed the maximum days to 50% flowering (86.83) and days to maturity (138.67). However, the magnitude of the mean value for each cluster varied, suggesting greater genetic diversity across the thirteen traits that could be employed in hybridization programs.

Table 5. Cluster mean values of 17 white maize genotypes across 13 characters.

Characters	I	II	III	IV	V
Plant height (cm)	195.81	204.17	219.29	209.93	182.36
Distance between base to cob (cm)	75.20	95.04	103.72	117.11	108.27
Base diameter of plant (cm)	7.20	6.43	5.99	5.38	5.75
Leaf length (cm)	72.21	64.40	69.20	64.47	69.44
Leaf breadth (cm)	9.39	8.41	7.74	6.95	7.54
Days to 50% flowering (DAS)	78.17	85.67	85.58	82.61	86.83
Days to maturity (DAS)	131.67	133.50	129.58	135.78	138.67
Cob length (cm)	16.47	15.84	14.87	13.58	14.21
Cob diameter (cm)	4.85	4.75	4.59	4.23	4.52
No. of rows cob ⁻¹	16.57	15.37	14.64	12.78	13.91
No. of grains row ⁻¹	30.14	25.80	23.25	19.02	21.74
100-grain weight (gm)	34.10	29.57	26.94	22.16	25.27
Yield plant ⁻¹ (gm)	170.01	117.66	91.89	54.28	76.55

Collectively, we noted considerable genetic divergence among the inbred lines, which were classified into five distinct clusters. The significant inter-cluster distances indicated that crossing inbred lines from distant clusters could maximize heterotic effects of the genetic recombinants, potentially leading to novel combinations of desirable traits. Considering the diversity pattern and other agronomic performance of these seventeen genotypes, CLTHW15008 from cluster I and CLTHW15004, CLTHW15005, CLTHW15006, CLTHW15011, CLTHW15001 and CLTHW15015 from cluster II might serve as a pool of potential parents for future hybridization programs. However, the present assessment was based primarily on yield-related traits under the evaluated environment. Therefore, the selected inbred lines should undergo multilocation and multi-season evaluation, together with systematic screening for major biotic and abiotic stresses, before being advanced as stable parental lines or potential white maize varieties.

Acknowledgment

The authors acknowledge the University Grants Commission of Bangladesh (UGC) for providing the research grant.

References

- Adebayo MA, Menkir A, Gedil M, Blay E, Gracen V, Danquah E and Funmilayo L 2021. Diversity assessment of drought-tolerant exotic and adapted maize (*Zea mays* L.) inbred lines with microsatellite markers. *J. Crop Sci. Biotech.* **24**(2): 147-154.
- Ahmed N, Chowdhury AK, Uddin MS and Rashad MMI 2020. Genetic variability, correlation and path analysis of exotic and local hybrid maize (*Zea mays* L.) genotypes. *Asian J. Med. Biol. Res.* **6**(1): 8-15.
- Antony BJ, Kachapur RM, Naidu GK and Harlapur SI 2021. Genetic diversity study among maize (*Zea mays* L.) inbred lines. *J. Farm Sci.* **34**(4): 352-356.
- Arora A, Das AK, Kumar R, Sharma S, Kaur N, Dixit S, Kaur Y, Saxena DC and Rakshit S 2024. Development of high-yielding white maize hybrids with better chapatti-making quality compared to traditionally used local landraces. *Front. Nutr.* **11**: 1330662.
- BBS (Bangladesh Bureau of Statistics) 2023. Statistical Yearbook Bangladesh. Statistics and Informatics Division, Ministry of Planning, The Government of the People's Republic of Bangladesh, Dhaka. p.151.
- Dogar MW, Hussain Z, Ashraf MI and Naveed M 2023. Evaluation of genetic variability in maize (*Zea mays* L.) based on yield and its attributed traits. *Asian J. Biotech.* **6**(2): 200-210.
- Doggalli G, Mahadevu P, Lohithaswa HC and Shekar BG 2024. Assessment of genetic diversity in the inbred lines for forage traits in maize. *Plant Arch.* **24**: 867-874.
- Fatema R, Rahman J, Shozib HB, Nazrul MI and Fatima K 2019. Genetic diversity and nutritional components evaluation of Bangladeshi germplasms of kidney bean (*Phaseolus vulgaris* L.). *J. Genet. Res.* **5**(2): 83-96.
- Garcia-Lara S, Chuck-Hernandez C and Serna-Saldivar SO 2019. Development and structure of the corn kernel. In: Serna-Saldivar SO (Ed.), *Corn* (Third Edition). AACC International Press, Oxford. pp. 147-163.
- Ghafoor A, Ahmad Z and Khan M 2022. Principal component analysis in rapeseed: Identifying key traits for yield improvement. *Crop Sci. Rev.* **18**(1): 23-34.
- Islam NU, Ali G, Dar ZA, Maqbool S, Kumar B and Bhat A 2020. Genetic divergence in maize (*Zea mays* L.) inbred lines. *Int. J. Chem. Stud.* **8**(1): 425-428.
- Khan S, Mahmud F and Ahmed T 2022. Genetic diversity with cluster analysis of maize genotypes (*Zea mays* L.). *Adv. Biosci. Biotech.* **13**(7): 273-283.
- Malvar RA, Revilla P, Moreno-González J, Butrón A, Sotelo J and Ordás A 2008. White maize: genetics of quality and agronomic performance. *Crop Sci.* **48**(4): 1373-1381.
- Menkir A, Gedil M and Danquah E 2018. Genetic diversity and population structure of early-maturing tropical maize inbred lines. *Theor. Appl. Genet.* **131**(12): 2665-2678.
- Nwangburuka CC, Kehinde OB, Ojo DK, Denton OA and Popoola AR 2011. Morphological classification of genetic diversity in cultivated okra (*Abelmoschus esculentus* (L) Moench using principal component analysis (PCA) and single linkage cluster analysis (SLCA). *Afr. J. Biotech.* **10**(54): 11165-11172.
- Olawamide DO and Fayeun LS 2020. Correlation and path coefficient analysis for yield and yield components in late-maturing provitamin a synthetic maize (*Zea mays* L.) breeding lines. *J. Exp. Agric. Int.* **42**: 64-72.
- Pranay G, Shashibhushan D, Bhadraru D, Rani KJ and Kumar CS 2024. Genetic diversity studies in elite maize lines. *Plant Arch.* **24**(2): 1689-1694.
- Roy G, Shompa BN, Emu IA, Fatima K and Rahman J 2025. Genetic variability, correlation and path coefficient analysis of exotic white maize (*Zea mays* L.) genotypes. *Bangladesh J. Agri.* **50**(2): 132-146.
- Shompa BN, Fatima K, Jony M, Sarkar S, Ullah MJ, Chowdhury AK and Rahman J 2020. Selection of dwarf stature yield potential lines from F₃ populations of white maize (*Zea mays* L.) *J. Genet. Res.* **6**: 95-105.
- Singh B, Gupta R and Verma N 2023. Multivariate statistical approaches for trait selection in oilseed crops. *Plant Genet. Res.* **9**(4): 201-217.

(Manuscript received on 20 April, 2025; revised on 03 September, 2026)