

GENETIC DIVERGENCE ANALYSIS OF LINSEED GENOTYPES

SAURAV ROY, SHAMIMA AKTER*, MD MAZBHA UDDIN,
MK HOSSAIN AND NAZMUL ALAM

Department of Botany, Jahangirnagar University, Savar, Dhaka-1342, Bangladesh

Keywords: Linseed, Genetic divergence, Principal Component Analysis,
Mahalanobis's D^2 statistics

Abstract

An experiment was conducted to assess the genetic divergence among ten genotypes of linseed (*Linum usitatissimum* L.) for 16 quantitative traits using a Randomized Complete Block Design with three replications at the Botanical Garden, Jahangirnagar University during the Rabi season 2024-2025. Genotypic coefficient of variation (GCV) and phenotypic coefficient of variation (PCV) attained their highest values for seed yield per plant (37.16 and 38.58, respectively). Heritability was high for all traits except the number of seeds per capsule and oil content (%). Based on Mahalanobis's D^2 statistics, the 10 genotypes were grouped into five clusters; cluster I contained two genotypes, clusters II and III contained three genotypes each, while clusters IV and V were solitary. The greatest inter-cluster distance was recorded between clusters II and III while the lowest between IV and V. Cluster III had the greatest intra-cluster distance, indicating the highest level of heterogeneity among the clusters. Significant contributors to variability included days to first flowering (21.2%), number of primary branches per plant (12.1%) and 1000 seed weight (11.4%).

Introduction

Linseed (*Linum usitatissimum* L.) which is locally known as 'tishi' and 'moishna' in Bangladesh is a multipurpose Rabi crop. It is an annual diploid crop ($2n = 30$) of the Linaceae family, comprising 14 genera and more than 200 species (Patra 2018). Linseed is gaining renewed interest due to its adaptability to marginal lands, low input requirements and the diverse utility of its seed oil, which is rich in α -linolenic acid- a crucial omega-3 fatty acid beneficial for human health (Diederichsen and Ulrich 2009).

Linseed oil reduces the risk of heart disease, inflammation, arthritis, bowel disorders, and several other health issues. Additionally, it contains lignin- a group of compounds that contribute significantly to cancer prevention (Dabalo *et al.* 2020). The fibre obtained from linseed stems is processed into linen yarn, which is widely used in high-quality textile and apparel manufacturing. Moreover, high-quality writing papers are also manufactured from linen fibre (Sahu 2015).

Linseed offers numerous benefits, yet its production in Bangladesh has been steadily declining. The primary cause of limited linseed cultivation is its low yield potential, which occurs due to a lack of improved cultivars. Most linseed genotypes currently cultivated are traditional landraces or older strains, which have limited productivity and adaptability. As a result, developing cultivars with high yield potential has become a primary objective to address low production levels (You *et al.* 2019). According to Arunachalam (1981), genetic diversity can be crucial in selecting parental materials to achieve maximum recombination in hybridization programmes. Mahalanobis's D^2 statistics is a crucial tool that illustrates the nature and magnitude of genetic divergence, aiding in the selection of sustainable genotypes for improved yield. Therefore, the present investigation was carried out on ten linseed genotypes using Mahalanobis's D^2 statistics to assess the magnitude of genetic diversity.

*Author for correspondence: <sanupurbotju@juniv.edu>.

Materials and Methods

This investigation was carried out in the experimental field of the Botanical Garden of Jahangirnagar University, Savar, Dhaka, Bangladesh, during the Rabi season from November 2024 to March 2025. The experiment was arranged in a Randomized Complete Block Design (RCBD) with three replications. The materials for the present investigation consisted of 10 linseed genotypes collected from different sources (Table 1). The dosages of fertilizers and manures were applied to the experimental field according to BARC (2018) and cultural practices were done when required. Data were collected from ten randomly chosen plants in each replication at the field level. For this study, 16 quantitative traits were examined, including seed germination percentage, plant height (cm), number of primary and secondary branches per plant, days to 1st and 50% flowering, days to final harvesting, total biomass per plant (gm), harvest index (%), number of capsules per plant, capsule diameter (mm), number of seeds per capsule, 1000-seed weight (gm), oil content (%), free fatty acids content (%), and seed yield per plant (gm).

Table 1. Linseed genotypes with their sources used in the experiment.

Codes	Genotypes name	Sources
G ₁	BARI Tishi-1(Neela)	ORC, BARI
G ₂	Narail	Local
G ₃	Faridpur	Siddiq Bazar, Dhaka
G ₄	Thakurgaon	ACI Ltd.
G ₅	Gopalganj	Local
G ₆	Rajshahi	Local
G ₇	Natore	Local
G ₈	Shariatpur	Namabazar, Savar
G ₉	Sirajganj	Local
G ₁₀	Habiganj	Local

*G: Genotype, ORC: Oil Research Centre, BARI: Bangladesh Agricultural Research Centre.

The collected data were assembled and formulated in proper form of statistical analysis. The genetic diversity among the genotypes was evaluated using Mahalanobis's (1936) distance (D^2) and general statistics with Indostat 9.1 software. Principal Component Analysis (PCA) was utilized to analyze the inter-relationships among various characters using R software. Average intra-cluster and inter-cluster distances were determined using the formula, recommended by Singh and Chaudhary (1985) using Indostat 9.1 software.

Results and Discussion

To assess the extent of variability caused by genetic factors, the phenotypic coefficient of variation (PCV), genotypic coefficient of variation (GCV), heritability (h^2), genetic advance (GA), and genetic advance as a percentage of the mean (GAM) were calculated for sixteen traits and are shown in Table 2. The study revealed that all characters showed higher PCV than GCV, suggesting that environmental factors influence these traits. The highest GCV and PCV were noted for seed yield per plant with a value of 37.16 and 38.58, respectively, followed by total biomass per plant (37.03 and 38.17, respectively), free fatty acid (36.42 and 36.66, respectively), and number of capsules per plant (33.34 and 33.67, respectively). The heritability estimates in percentage and genetic advance have been presented in Table 2. Most of the characters showed high heritability in broad sense (94 to 98) except oil content (29.10). The number of capsules per plant (46.01) showed the highest value of genetic advancement at 5% level, followed by plant height (41.58), Seed germination (21.91%) and harvest index (17.36). The highest genetic advance as a

percentage of the mean was observed for the character free fatty acid (74.51), followed by total biomass per plant (74.02), seed yield per plant (73.73), number of capsules per plant (67.99), and harvest index (59.28). The lowest percentage was recorded in oil content at 2.91%.

Table 2. Estimation of genotypic coefficient of variation (GCV), phenotypic coefficient of variation (PCV), heritability (h^2) and genetic advancement.

Characters	GCV	PCV	h^2 (Broad sense as %)	Genetic advancement (5%)	Genetic advancement as % of mean
Seed germination (%)	13.25	13.66	94.10	21.91	26.49
Plant height (cm)	27.12	27.61	96.50	41.58	54.90
No. of primary branches/ plant	15.87	16.24	95.50	2.46	31.97
No. secondary branches / plant	21.10	21.45	96.80	5.07	42.75
Days to first flowering	9.80	9.92	97.50	11.30	19.93
Days to 50% flowering	9.99	10.22	95.40	12.31	20.10
Days to final harvesting	5.30	5.31	99.60	12.48	10.90
Total biomass per plant(gm)	37.03	38.17	94.10	7.30	74.02
Harvest index (%)	29.90	31.08	92.60	17.36	59.28
No. of capsules / plant	33.34	33.67	98.00	46.01	67.99
Capsules diameter (mm)	8.86	9.01	96.60	1.07	17.94
No. of seed / capsule	13.15	15.42	72.70	1.58	23.10
1000 seed weight (gm)	12.46	12.68	96.50	1.13	25.21
Oil content (%)	2.62	4.86	29.10	0.95	2.91
Free fatty acid (%)	36.42	36.66	98.70	0.91	74.51
Seed yield / plant (gm)	37.16	38.58	92.80	1.99	73.73

Using non-hierarchical clustering based on 16 quantitative characters, 10 genotypes were divided into 5 clusters: clusters II and III were the largest, each containing 3 genotypes; cluster I included 2 genotypes; and clusters IV and V consisted of one genotype each (Table 3). The genotype of the solitary cluster is highly distinctive and useful for breeding. Based on 7 characters, Begum *et al.* (2007) found 5 distinct clusters for 36 linseed genotypes, whereas Khan *et al.* (2013) found 13 distinct clusters based on 10 characters using 55 linseed genotypes.

Table 3. Distribution of different linseed genotypes into clusters based on D^2 statistics.

Cluster	Number of the genotypes	Name of genotypes
I	2	G ₃ , G ₈
II	3	G ₆ , G ₉ , G ₁₀
III	3	G ₂ , G ₄ , G ₅
IV	1	G ₇
V	1	G ₁

The PCA provided eigenvalues and percentage contributions for each principal component axis, capturing the variation among genotypes. A two-dimensional scatter plot was created using the first two principal components, PC₁ (44.1%) and PC₂ (37.3%), which together explained 81.45% of the total genetic divergence (Table 4). This scatter diagram illustrates the distribution pattern of genotypes, showing significant variability largely due to the high eigenvalues and contribution percentages of these axes (Fig. 1). Based on the principal component scores, distribution pattern among 25 genotypes was observed by Kumar and Paul (2016) and 16 genotypes by Paul *et al.* (2017).

Table 4. Lateral roots (eigenvalues) and percentage of variation for corresponding 9 components characters in 10 linseed genotypes.

Principal components	Eigen value	Percentage of variation	Cumulative percentage of variation
PC ₁	7.057	44.106	44.106
PC ₂	5.974	37.341	81.448
PC ₃	1.359	8.499	89.947
PC ₄	0.942	5.889	95.837
PC ₅	0.317	1.981	97.818
PC ₆	0.240	1.502	99.321
PC ₇	0.071	0.448	99.769
PC ₈	0.029	0.182	99.951
PC ₉	0.007	0.048	100.000

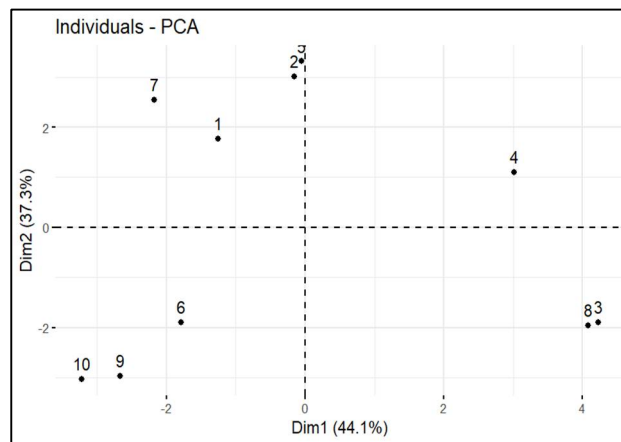


Fig.1. Distribution of various genotypes represented in a scatter plot based on PCA scores.

In the present investigation, 16 characters contributed towards divergence. Among the 16 characters analyzed, days to first flowering contributed the most to divergence at 21.2%, followed by the number of primary branches per plant at 12.1%, 1000 seed weight at 11.4%, plant height at 9.2%, capsule diameter at 8%, total biomass per plant at 6.4%, days to 50% flowering at 5.9%, number of secondary branches per plant at 5%, seed yield per plant at 3.4%, with the remaining characters contributing less to genetic divergence (Fig. 2). Nizar and Mulani (2015) observed highest contribution for number of capsules per plant followed by days to maturity, days to 50% flowering, plant height, seed yield per plant and gave emphasis on selecting genotypes based on high contributing characters. Therefore, parent selection should focus on days to first flowering, the number of primary branches per plant, 1000 seed weight, plant height, capsule diameter, number of capsules per plant, total biomass per plant, days to 50% flowering, seed germination percentage, and seed yield per plant due to their significant role in genetic divergence in linseed. Using PCA, it was found that intra-cluster distances ranged from 37.50 (cluster I) to 717.56 (cluster III), indicating the presence of divergent genotypes within the clusters. Clusters IV and V showed 0 intra-cluster distance due to a solitary cluster (Table 5). The inter-cluster D^2 values also ranged widely, with a minimum value (1153.33) between clusters IV and V, indicating a homogenous nature among the genotypes of these clusters, and a maximum value (7786.62) between clusters II and III, indicating high diversity among the genotypes of these clusters. Therefore, genotypes within these clusters are more genetically diverse. Hybridizing genotypes

from these groups is likely to produce more advantageous segregants, attain higher hybrid vigor, and potentially develop high-yielding linseed varieties. Genotypically distant parents can transfer genes effectively and produce heterosis Begum *et al.* (2007) observed a minimum intra-cluster distance in cluster I and a maximum in cluster III; the highest inter-cluster distance was between clusters II and V.

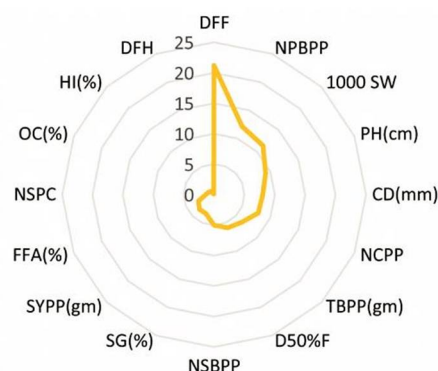


Fig. 2. Contribution of 16 characters toward the genetic divergence.

SG% = Seed germination percentage, PH = Plant height, NPBPP = Number of primary branches per plant, NSBPP = Number of secondary branches per plant, DFF = Days to first flowering, D50%F = Days to 50% flowering, DFH = Days to final harvesting, TBPP = Total biomass per plant, HI = Harvest index, NCPP = Number of capsules per plant, CD = Capsule diameter, NSPC = Number of seed per capsule, 1000 SW = 1000 seed weight, OC = Oil content, FFA = Free fatty acid, SYPP = Seed yield per plant.

Table 5. Average intra (bold) and inter-cluster distance in linseed genotypes.

Cluster	I	II	III	IV	V
I	37.5	5496.27	2195.51	3536.96	4378.07
II		344.61	7786.62	3792.05	2374.88
III			717.56	1821.39	3969.63
IV				0.00	1153.33
V					0.00

Table 6. Cluster mean for 16 characters in linseed genotypes.

Characters	I	II	III	IV	V
Seed germination (%)	68.67	85.89	82.44	96.67*	88.33
Plant height (cm)	111.1*	58.73	76.91	67.73	60.53
No. of primary branches /plants	8.43	6.91	8.8*	7.27	5.93
No. of secondary branches /plants	15.1*	10.4	13.16	9.27	8.47
Days to first flowering	62.67*	49.44	58.67	55.67	61.67
Days to 50% flowering	67.83*	53.22	63.78	59.33	66.33
Days to final harvesting	123*	110.22	114.67	107	117
Total biomass per plant (gm)	13.56*	5.08	12.38	9.38	9.76
Harvest index (%)	15.06	33.66	29.67	39.06*	33.81
No. of capsules / plant	53.78	44.87	83.66	94.47*	89.07
Capsule diameter (mm)	5.46	5.61	6.62*	6.03	6.05
No. of seeds / capsule	5.67	6.64	7.47	7.8*	7.8*
1000 Seed weight (gm)	4.66	3.94	5.05*	4.86	3.97
Oil content (%)	31.8	31.49	33.62	34.07*	33.32
Free fatty acid (%)	1.74*	1.47	0.96	0.88	0.66
Seed yield / plant (gm)	2.04	1.69	3.63	3.67*	3.3

Character-wise means were computed for all genotypes across five clusters which is shown in Table 6. Cluster I exhibited the highest average for plant height (111.10 cm), number of secondary branches per plant (15.10), days to first flowering (62.67), days to 50% flowering (67.83), days to final harvest (123.00), total biomass per plant (13.56 g), and free fatty acid content (1.74%). Cluster III showed highest mean value for number of primary branches per plant (8.80), capsule diameter (6.62 mm), number of seeds per capsule (7.47) and 1000 seed weight (5.05 gm) (Table 6). Cluster IV showed highest mean value for seed germination (%) (96.67 %), harvest index (39.06), number of capsules per plant (94.47), number of seeds per capsule (7.8), oil content (34.07 %) and seed yield per plant (3.67 gm). Cluster V showed the highest mean value only for the number of seeds per capsule (7.8). Cluster II does not show the highest mean value for any character, indicating that selection of parental lines from this cluster for future hybrid linseed breeding programme will be ineffective. These findings support the observations reported by Khan *et al.* (2013) and Begum *et al.* (2007).

Based on the most divergent contributing characters, greater inter-cluster distance, and overall mean performance, it can be concluded that G₇ (Natore local) from cluster IV and G₅ (Gopalganj local) from cluster III are promising candidates for future breeding efforts.

References

- Arunachalam V 1981. Genetic distance in plant breeding. *Indian J. Genet. Plant Breed.* **41**: 226-236.
- Bangladesh Agricultural Research Council (BARC) 2018. Fertilizer recommendation guide. Dhaka, Bangladesh. pp. 92.
- Begum H, Alam AKMM, Chowdhury MJA and Hossain MI 2007. Genetic divergence in linseed (*Linum usitatissimum* L.). *Int. J. Sustain. Crop Prod.* **2**(1): 4-6.
- Dabalo DY, Singh BCS and Weyessa B 2020. Genetic variability and association of characters in linseed (*Linum usitatissimum* L.) plant grown in central Ethiopia region. *Saudi J. Biol. Sci.* **27**(8): 2192-2206.
- Diederichsen A and Ulrich A 2009. Variability in stem fibre content and its association with other characteristics in 1177 flax (*Linum usitatissimum* L.) gene bank accessions. *Ind. Crops Prod.* **30**(1): 33-39.
- Khan MA, Mirza MY, Amjad M, Nawaz N, Nawa MS and Baig D 2013. Assessment of genetic diversity in germplasm of linseed (*Linum usitatissimum* L.). *Pakistan J. Agric. Res.* **26**(3): 178-184.
- Kumar N and Paul S 2016. Selection criteria of linseed genotypes for seed yield traits through correlation, path coefficient and principal component analysis. *JAPS: J. Anim. Plant Sci.* **26**(6): 1688-1695.
- Mahalanobis PC 1936. On the generalized distance in statistics. *Proc. Natl. Inst. Sci. India.* **2**: 49-55.
- Nizar MA and Mulani RM 2015. Genetic diversity in indigenous and exotic linseed germplasm (*Linum usitatissimum* L.). *Electron. J. Plant Breed.* **6**(3): 848-854.
- Patra D 2018. Morphological and Biochemical Characterization of Elite Genotypes of Linseed. (Doctoral dissertation), Orissa University of Agriculture and Technology, Bhubaneswar, India.
- Paul S, Kumar N and Chopra P 2017. Correlation and genetic diversity of linseed (*Linum usitatissimum* L.) genotypes based on principal component analysis in Mid-Hills of North-West Himalayas. *J. Pharmacogn. Phytochem.* **6**(1): 287-290.
- Sahu D 2015. Characterization and genetic analysis in germplasm accessions of linseed (*Linum usitatissimum* L.) (M.Sc. Ag. Thesis). Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh, India.
- Singh RK and Chaudhary BD 1985. *Biometrical Methods in Quantitative Genetic Analysis* (rev. Ed.). Kalyani Publishers. New Delhi, India. pp. 252-269.
- You FM, Cloutier S, Rashid KY and Duguid SD 2019. Flax (*Linum usitatissimum* L.) Genomics and Breeding. In: *Advances in Plant Breeding Strategies: Industrial and Food Crops*. Cham: Springer **6**: 277-317.