RESEARCH ARTICLE Indian Journal of Agricultural Research, Volume 60 Issue 2: 166-174 (February 2026)

Evaluation of Genetic Diversity in Linseed (*Linum usitatissimum* L.) Germplasm Through D^2 Cluster Analysis and PCA

Korra Shankar[1,2], Nalini Tiwari[1], Achila Singh[1], C.V. Sameer Kumar[2], Guglothu Suresh[2]

10.18805/IJARe.A-6385

ABSTRACT

Background: Linseed is used as an oilseed and fiber crop and is extremely rich in high Omga-3 fatty acid content. The main aim of breeders in crop improvement programs is to develop high-yielding crop varieties. For this, we need to select diverse high-yielding genotypes to produce the heterotic cross combinations. This experiment aims to evaluate the genetic diversity among linseed germplasm for yield-related traits.

Methods: In this experiment, a total of 75 diverse linseed genotypes (Indigenous and three exotic) along with ten checks were investigated during *Rabi*-2019-20 in randomized block design using three replications at Oilseeds Research Farm, Kalyanpur, Kanpur (Uttar Pradesh). Observations were recorded on ten yield-related traits. Clustering of genotypes by using the D^2 cluster of Tocher's method and Principal component analysis.

Result: Cluster analysis deciphered 75 genotypes into 16 distinct groups, with the highest number of individuals observed in cluster II. The greater average was recorded within clusters IV (14.34), II (13.20), XI (12.78) and I (6.31) order, whereas clusters between I and XI recorded the high between cluster distance. The results indicated that a total of 57.58% of variation was contributed by the first three principal components (PCA).

Key words: D^2 Cluster, Diversity, Linseed, PCA, Yield.

INTRODUCTION

A self-fertilizing winter season oilseed crop, linseed (*Linum usitatissimum* L.) is a member of the Linaceae family (Cloutier *et al*., 2012). It is believed that linseed originated in the Mediterranean region Darlington (1963) and Southwest Asia, specifically in India (Vavilov, 1935; Richharia, 1962). It has a somatic chromosome number of 2n=30, with other species exhibiting chromosomal variability (2n=16 to 60). Although primarily self-pollinating, insect activity can result in up to 2% outcrossing (Dilman, 1928). The term "linseed" is used when the crop is cultivated for oilseed, whereas "fibre flax" or simply "flax" is commonly used in Europe when it is grown for fibre (Vaisey-Genser and Diane, 2003). Flaxseed is rich in fats, proteins and dietary fibre and is available in two primary varieties: Brown and yellow or golden (also referred to as golden linseeds). Brown flaxseeds contain approximately 41% fat, 20% protein and a substantial 28% total dietary fibre. Also contains 7.7% moisture and 3.4% ash (left after combustion) (Gill, 1987). The oil content in linseed ranges from 33% to 45% (Arora *et al*., 2003) and it exhibits an inverse relationship with protein content. Flax fibres are widely used in the textile industry to make linen cloths, thread, rope and packaging materials, as well as in the production of currency notes (Mackiewicz-Talarczyk *et al*., 2008). Flax fiber is highly valued for its exceptional durability and strength. It is soft, flexible and non-lignified with 80-90% cellulose. Linseed is well known for its high content of omega-3 fatty acids Smy´kal *et al.* (2012) and it is responsible for its unique drying effect (Przybylski, 2005). These fatty acids contribute to cholesterol reduction and improve heart health (Westcott and Muir, 2003). Additionally, linseed oil is also used in the making of products like varnish, inks, paints and linoleum flooring due to its exceptional drying effect (Czemplik *et al*., 2011). Linseed is highly demanded for industrial applications due to its utilization in both edible and non-edible products. The lower productivity of linseed is due to poor fertility response and contains unsaturated fatty acid, linoleic acid causes the oxidation of oil thereby reducing the storage lifetime. To overcome these challenges, breeding efforts should focus

[1]Oilseeds Research Farm, Kalyanpur, Chandra Shekhar Azad University of Agriculture and Technology, Kanpur-208 002, Uttar Pradesh, India.
[2]Professor Jayashankar Telangana Agricultural University, Hyderabad-500 030, Telangana, India.

Corresponding Author: Korra Shankar, Oilseeds Research Farm, Kalyanpur, Chandra Shekhar Azad University of Agriculture and Technology, Kanpur-208 002, Uttar Pradesh, India.
Email: shankarchowhan5962@gmail.com
ORCID: 0009-0002-3208-1936

How to cite this article: Shankar, K., Tiwari, N., Singh, A., Kumar, C.V.S. and Suresh, G. (2026). Evaluation of Genetic Diversity in Linseed (*Linum usitatissimum* L.) Germplasm Through D^2 Cluster Analysis and PCA. *Indian Journal of Agricultural Research.* **60(2):** 166-174. doi: 10.18805/IJARe.A-6385.

Submitted: 06-03-2025 **Accepted:** 24-10-2025 **Online:** 17-11-2025

on developing linseed lines with enhanced yield potential and reduced linoleic acid content, thereby improving productivity and oil quality. To accelerate crop improvement programs, selecting genetically diverse parents is essential for the creation of high-yielding varieties with improved climate resilience for diversified agroecological recovery. For this purpose, it is essential to quantify the diversity among parents (Govindaraj *et al*., 2015). The more diverse the parents, the greater the chances of achieving high heterotic lines for seed yield and its component characters.

MATERIALS AND METHODS

A total of 75 genotypes were used in this experiment to assess the genetic diversity (Table 1) including three exotic (Canada) and other diverse regions of Indian including checks (T-397, Shekhar, Parvathi, Rajan, Gaurav, Surya, Ruchi, Meera, Rashmi and Shikha) of linseed (*Linum usitatissimum* L.) during *Rabi* season 2019-20. The location of this experiment is at Oilseeds Research Farm, Kalyanpur with the collaboration of C.S. Azad University of Agriculture and Technology, Kanpur, Uttar Pradesh, (India). The experiment was carried out in a randomized block design with a spacing of 30 × 7 cm using three replications. The equally competitive plants within a row were chosen for taking observation on ten yield component traits like days for 50 per cent blooming, primary branches, secondary branch number, height of plant (cm), capsule number, seeds per capsule, days needed for maturity, 1000 seed weight, oil content of each line (%) and yield of single plant. A total of 25 grams of seeds from each genotype was used for oil content estimation using the NMR technique. The multivariate methods based on Tocher's cluster analysis, which Rao (1952) and Mahalanobis (1936) described and PCA were used to evaluate the phenotypic divergence among the accessions and calculate the genetic distance using D^2 statistics.

RESULTS AND DISCUSSION

Clustering of genotypes

Members within the same cluster are expected to be in a high close relationship in terms of the characteristics being studied than members from different clusters. D^2 clustering pattern deciphered that a total of 75 genotypes were grouped into 16 distinct clusters, displayed in Table 2 and Fig 1. Cluster II exhibited the greatest number of genotypes, consisting of 38, while cluster IV recorded the second maximum number of genotypes with 17. Cluster XI composed of five genotypes follows Cluster I, which contains three genotypes. The lowest number of genotypes was found in clusters III, V, VI, VII, VIII, IX, X, XII, XIII, XIV, XV and XVI, which had a single genotype each. The exotic germplasms, Hermes, Redwood 65 and AR-2 were classified into clusters, II, XVI and X, respectively. Similar patterns of clustering were corroborated by Kant *et al*. (2011); Meena *et al*. (2021) and Kumar and Kumar (2021). The intra-cluster and inter-cluster distance estimations of D^2 values were presented in Table 3. Cluster IV (14.34) has recorded the highest intra-cluster distance, followed by cluster II (13.20), cluster XI (12.78) and cluster I (6.31) in descending order. Only one genotype was present in clusters X, XI, VIII, VII, IV and III (0.00), which exhibited the lowest intra-cluster D value. The highest intra-cluster distance was recorded in descending order between cluster XI and cluster I (40.62), cluster V and XI (38.13), cluster V and cluster XIV (35.57), cluster I and cluster XIV (34.73) and cluster VI and cluster XV (33.98). Due to the greater inter-cluster distances among these clusters, crossing between these clusters can result in hybrids with better heterosis (Acquaah, 2012). Similar types of results in diversity studies in linseed have been previously confirmed by (Tewari *et al*., 2020; Nizar and Mulani, 2015; Pali and Mehta, 2015). The lowest inter-cluster distance was recorded among clusters VI and X (8.79), then between VII and XII (9.99), III and VII (10.88), III and VIII (11.72) and cluster III and cluster XIV (11.73).

Cluster means (Average)

There was a considerable variation between 16 clusters concerning cluster average for distinct traits, as revealed by intra-cluster means ten characters shown in Table 4. Clusters XIV (89.00 days), cluster XIII (83.00 days), cluster X (82.00 days) and cluster VII (81.00 days) were recorded the highest average and clusters IX (64.00 days) and XV (65.00) had the lowest average for days for 50 per cent blooming. The greater cluster mean for the primary branch was obtained for cluster X (7.67) followed by cluster IX (7.33), cluster VI (5.33) and cluster XIV (5.00). The lowest cluster mean was observed for cluster V (2.67). Clusters X (31), IX (30.33), XVI (29.00) and VI (20.67) showed the highest cluster average value in descending order for the secondary branch and cluster I (8.67) recorded the lowest average. For plant height, clusters I (104.00), XVI (102.67), V (79.00) and VI (75.33) were observed the highest cluster average and cluster XV (49.00) had the lowest average. The highest cluster average was observed in clusters VI (172.00), XI (165.80), X (156.33) and XVI (132.33) and the lowest cluster mean was in cluster XV (36.00) for capsules per plant. Cluster V (7.67) exhibited the lowest average and clusters VIII (9.33), III (9.00), XIV (8.59) and XI (8.93) for trait, seeds per capsule. When it comes to days taken for maturity, clusters XIV (146.33), XII (145.00), X (141.00) and I (140.67) demonstrated the highest cluster average and cluster VIII (126.00) indicated the lowest mean. Clusters IX (45.50), V (42.00), IV (39.24) and I (38.67) recorded the highest average oil content. In contrast, Cluster XIV (24.86) recorded the lowest mean. The maximum average of 1000-seed weight was observed in clusters VI (8.84), X (8.71), XII (8.52) and XIV (8.26). On the other hand, cluster XIV (4.18) had the lowest average weight. The character seed yield recorded the maximum average in cluster XIII (12.47) then in cluster X (10.12), cluster VI (10.09), cluster IX (9.37)

Table 1: List of linseed accession utilized in this experiment.

Genotypes	Parents/Pedigree	Source
BRLS 111-3	(RL 26018 × Shekhar) × LCK 7035	Bikramganj, Bihar
LCK 1807	Shekhar × T-397	Kanpur, U.P.
RL 15582	RL-2204 × Triveni	Kota, Rajasthan.
NL 315	Shekhar × JRF-5	Nagpur, Maharastra.
T-397	T- 491 × T 1193-1	Kanpur, U.P.
RL 15580	Pratap Alsi-l × NL-18	Kota, Rajasthan.
BAU 13-01	Shweta × PKDL 41	Kanke, CG.
LMS 2016-I-2	LMS -9-2K × Cl-15888	Mauranipur, U.P.
OL 10-15	KL 248 × Padmini	Keonjhar, Odisha
BRLS 119	(LCK 7035 × Shekhar) × PKDL 71	Bikramganj, Bihar
RLC 174	Neelum × R 7	Raipur, CG.
LMS 2016-I-5	LMS-9-2K × JRF-5	Mauranipur, U.P.
RCRL 14	NL-115 × KL-168	Raichur, Karnataka
SLS 123	EC 541222 × JLS 73	Sagar, M.P.
Shekhar	BAU 1008 × Kiran	Kanke, CG.
LCK 1812	Laxmi-27 × EC-1387	Kanpur, U.P.
SLS 124	Jeevan × LCK 9209	Raipur, CG.
NL 335	Shubhra × Shikha	Kanpur, U.P.
RLC 173	-	Raipur, CG.
SLS 122	JLS 73 × JLS 66	Sagar, M.P.
SL 2	EC 1398 × EC 23595	Shillongani, Assam.
RL 15581	Pratap Alsi-l X RL-2204	Kota, Rajasthan.
LCK 1829	Shweta × Sheela	Kanpur, U.P.
SL 1	T 491 × T 1193-1	Kanpur, U.P.
RCRL 11	EC 1398 × FRW I	Shillongani, Assam.
RL 15583	NL-115 × FRW-12	Raichur, Karnataka.
RLC 171	PKDL-41 × Meera	Kota, Rajasthan.
BRLS 104	POLF-22 × JRF-5	Raipur, CG.
SLS 121	LCK 7035 × Shekhar	Bikramganj, Bihar
OL 10-2	JLS-73 × JLT-215	Sagar. M.P.
LMS 2016-R-16	OL 18-4 × OLC 10	Keonjhar, Odisha
LCK 1827	LMS-9-2K × PKDL-62	Mauranipur, U.P.
RL 172	CJS-234 × R-552	Raipur, CG.
LMS 2016-R-7	LMS-9-2K × Surbhi	Mauranipur, U.P.
BAU 14-09	Shekhar × Garima	Kanke, CG.
RFC 2018-1	GS234 × Ayogi	Raipur, CG.
LCK 1836	Shekhar × JRF-4	Kanpur, U.P.
BAU 14-07	Jeeven × Shekhar	Kanke, CG.
Ruchi	Garima × LCK 88062	Kanpur, U.P.
RFC 2018-2	Ayogi × GS-234	Raipur, CG.
LCK 1820	Neelam × JRF-1	Kanpur, U.P.
Tiara	-	Barrackpore, W.B.
KL 284	Giza7 × KLS-1	Palampur, H.P.
Meera	(RL 75-6-2 × RL 29-8) × LCK 8528	Kota, Rajasthan.
Parvati	EC41628/EC77959//DPL20 × Neelum/// EC216 × Hira//BR1 × NP440	Mauranipur, U.P.
RMLS 11	Sel. From Malghatti local	Raichur, Karnataka.
RL 13165	RLC 25127 × RL 26016	Raipur, CG.
TL 99	RLC 6 × Solin	BARC, Mumbai
Gaurav	Selection-3/EC-1552	Kanpur, U.P.

Table 1: Continue...

Table 1: Continue...

Rashmi	Gaurav/Janki	Kanpur, U.P.
Shikha	Hira/CRISTA	Kanpur, U.P.
Rajan	Gaurav × GS 234	Kanpur, U.P.
Surya	Shubhra × H-25	Kanpur, U.P.
Indu	LMS 1-2K × LMS 4-27	
JRF-1	-	Barrackpore, W.B.
JRF-3	-	Barrackpore, W.B.
JRF-4	-	Barrackpore, W.B.
Hermes	-	Canada
Redwood 65	-	Canada
AR-2	-	Canada
NL 356	-	Nagpur, MH.
LCK-9319-1	-	Kanpur, U.P.
CI-2007-3	-	Kanpur, U.P.
EC-1413-4	Exotic line	Kanpur, U.P.
BAU-09-7	-	Ranchi, Jharkhand
SLS-77-8	-	Sagar, M.P.
EC-9902-9	Exotic line	Kanpur.
NDL-204-10	-	Ayodhya, U.P.
LCP-11-102-11	-	Ludhiana, Punjab
LMS-2011-19-12	-	Mauranipur, U.P.
RL-10135-13	-	Kota, Rajasthan
EC-384154-14	Exotic line	Kanpur.
RL-29005-16	-	Kota, Rajasthan
LSL-100-21	-	Kanpur, U.P.
NDL-2011-36-22	-	Ayodhya, U.P.

Table 2: Composition of seventy-three linseed genotypes into different clusters.

Cluster number	No. of genotypes	Genotypes included in the cluster
I	3	JRF-3, JRF-4, JRF-1
II	38	RCRL 11, NL 356, SLS 121, NDL-204-10, BAU 14-07, BRLS 111-3, BAU-09-7, BRLS 104, LMS 2016-R-7, LMS 2016-1-5, LSL-100-21, EC-9902-9, RL 13165, SL2, LCK 1827, LMS 2016-R-16, LCP-11-102-11, BRLS 119, BAU 13-01, RL 15582, RLC 171, Rashmi, OL 10-2, RL-10135-13, SLS-77-8, EC-1413-4, CI-2007-3, T-397, Tiara, RLC 174, SL 1, Hermes, RL 15580, OL 10-15, LCK-9319-1, Shekhar, SLS 124, SLS 123
III	1	LMS-2011-19-12
IV	17	LCK 1807, Meera, Rajan, RCRL 14, RLC 173, TL 99, LCK 1812, NL 335, Indu, Gaurav, RMLS 11, LMS 2016-1-2, SLS 122, NL 315, RL 15583, Surya, LCK 1836
V	1	Shikha
VI	1	LCK 1829
VII	1	RFC 2018-1
VIII	1	Ruchi
IX	1	Parvat
X	1	AR-2
XI	5	RL 15581, BAU 14-09, RFC 2018-2, RL-2905-16, NDL-2011-36-22
XII	1	EC-384154-14
XIII	1	LCK 1820
XIV	1	KL 284
XV	1	RL 172
XVI	1	Redwood 65

and cluster XI (9.35). The clusters V and XIV (2.82) were noticed the lowest average. Ranjana *et al*. (2019) and Meena *et al*. (2021) have obtained similar results.

The contribution of different traits towards the divergence

According to average D^2, the highest contribution (presented in Table 5) was from the capsule number (35.75%) followed by plant height (19.96%), oil content (18.70%), seed yield (9.48%), test weight (8.72%). The other characteristics *viz.*, days needed for 50 per cent blooming, maturity duration, primary branch, secondary branch and seeds per capsule recorded negligible contribution. These characteristics, like the capsules per plant, plant height, grain yield and oil percentage should be emphasized to choose the appropriate parents for crossing. Tewari *et al.* (2013) also followed the high contribution of capsule number and seed yield toward total genetic divergence.

Principal component analysis (PCA)

A multivariate approach (Crossa, 1990) like Principal component analysis (PCA) is used to determine how the different traits contribute to overall variability and to offer a basis for choosing characteristics. The primary goal of PCA is to compress the total variation from studied variables into a smaller set of factors (Sharma, 1998; Brejda *et al*., 2000). In this study, a total of ten principal components (PCs) were extracted, equivalent to the number of traits studied and it revealed the four most informative PCs with eigenvalues of more than one which accounted for 68.38 per cent cumulative variance (Table 6). However, more than 50 per cent of the variance in the population was explained by three major PCs (PC1 25.88%; PC2-18.58% and PC3-13.12%). Dabalo *et al*. (2020); Guei *et al*. (2005) confirmed these findings. As a result, parameters showed positive weight in the first three PCs considered to be more crucial (Patial *et al*., 2019). The secondary branch number of each plant (0.54), primary branch number (0.49), capsule number (0.47), seed yield (0.45), days needed for 50 per cent blooming (0.13) and maturity duration (0.07) recorded positive weightage in PC1 while other traits were negative weight (Fig 2B). In PC2, the parameters *viz*., duration of 50 per cent blooming (0.32), seeds per fruit (0.27) and capsules (0.16) showed positive loading while other traits showed negative loading. In PC3, plant height (0.71), days needed for 50 per cent blooming (0.61) and maturity duration (0.31) showed highest positive loading and the remaining parameters showed negative loading. Generally, only a single component was chosen from these determined classes. The secondary branch of each plant was the better choice, as it exhibited the highest loading from PC1. Similarly, the time taken for 50 per cent blooming, plant length and seed number were best choices for the second, third and fourth principal components, respectively. PCA clearly demonstrated that the secondary branch of a plant, days to 50 per cent flowering, plant height and seeds per capsule were the most important traits showing a strong effect on total variation.

Table 3: Average intra and inter cluster distance in linseed.

Clusters	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI
I	**6.31**	29.06	28.30	23.37	15.20	33.30	30.66	26.81	32.61	31.41	40.62	27.98	27.21	34.73	26.92	20.96
II		**13.20**	15.85	17.46	24.60	17.70	15.59	17.13	16.21	17.00	19.45	15.57	18.17	23.05	22.14	23.21
III			**0.00**	21.29	26.78	23.68	10.88	11.72	23.13	21.65	19.13	14.32	19.22	11.73	21.52	25.45
IV				**14.34**	17.41	23.78	21.13	19.99	18.28	21.74	27.79	18.43	17.71	29.90	17.74	22.79
V					**0.00**	32.78	28.01	25.13	26.62	31.19	38.13	25.77	26.30	35.57	16.41	26.34
VI						**0.00**	22.36	24.69	17.01	8.79	18.45	21.42	20.16	27.54	33.98	19.21
VII							**0.00**	10.57	24.59	19.37	19.05	9.99	16.76	14.47	22.71	25.24
VIII								**0.00**	25.75	22.29	22.02	15.42	19.26	17.88	22.30	22.00
IX									**0.00**	16.58	21.30	22.40	20.37	30.77	24.66	25.65
X										**0.00**	18.40	17.05	16.27	24.41	31.58	16.95
XI											**12.78**	21.65	22.89	21.60	32.86	29.27
XII												**0.00**	13.13	19.08	21.41	23.34
XIII													**0.00**	25.32	22.87	21.59
XIV														**0.00**	30.98	29.13
XV															**0.00**	33.58
XVI																**0.00**

Bold figure denotes intra-cluster distance.

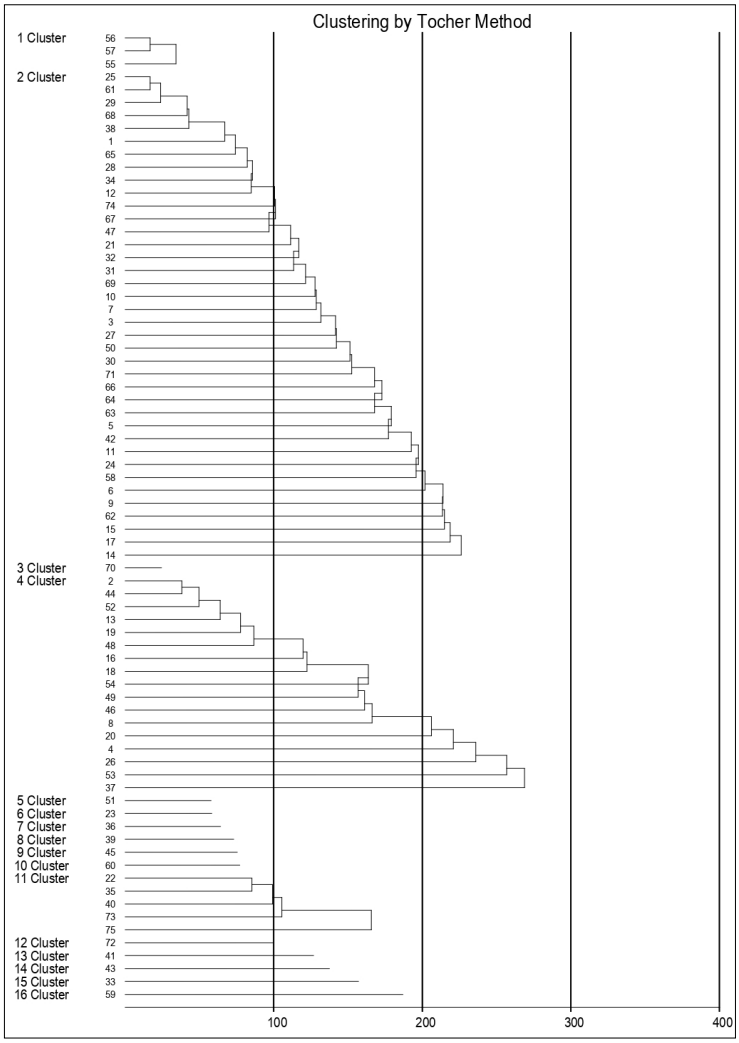


Fig 1: Dendrograms showing grouping of 75 linseed genotypes generated using D^2 cluster analysis.

Table 4: Cluster means of different characters to genetic diversity in linseed.

Character	DF	PB	SB	PH (cm)	CP	SC	DM	OC (%)	TSW (g)	SY (g)
I	77.33	3.22	8.67	104.00	54.67	8.78	140.67	38.67	6.83	4.32
II	69.17	4.50	16.93	60.90	118.89	8.78	130.82	36.36	7.37	6.12
III	74.00	4.00	11.00	60.67	105.33	9.00	139.00	30.02	4.49	4.45
IV	66.35	4.67	18.84	68.25	84.76	8.59	130.69	39.24	7.68	7.28
V	66.33	2.67	12.33	79.00	41.67	7.67	127.33	42.00	7.40	2.54
VI	77.67	5.33	20.67	75.33	172.00	9.67	133.00	39.25	8.84	10.09
VII	81.00	4.33	14.33	56.00	104.33	8.00	130.00	27.21	6.96	5.16
VIII	69.33	4.67	20.00	69.33	98.00	9.33	126.00	25.82	5.94	5.39
IX	64.00	7.33	30.33	57.00	135.33	8.00	134.00	45.50	6.96	9.37
X	82.00	7.67	31.00	71.33	156.33	8.00	141.00	37.00	8.71	10.12
XI	71.07	4.60	23.47	52.13	165.80	8.93	132.20	31.76	5.81	9.35
XII	78.00	3.00	11.00	59.00	97.67	8.33	145.00	31.01	8.52	6.27
XIII	83.00	4.00	16.00	67.00	101.00	8.33	137.00	34.83	7.01	12.47
XIV	89.00	5.00	19.00	59.00	122.00	9.00	146.33	24.86	4.18	2.82
XV	65.00	3.67	11.67	49.00	36.00	8.67	129.00	38.06	5.66	4.12
XVI	78.00	4.67	29.00	102.67	132.33	8.33	138.00	35.00	8.26	8.72

Table 5: Per cent contribution of ten traits towards total genetic divergence.

Characters	Number of times ranked first	Per cent contribution (%)
Days to 50 per cent flowering (DF)	107	3.86
Number of primary branches per plant (PB)	0	0.00
Number of secondary branches per plant (SB)	5	0.18
Plant height (PH) (cm)	554	19.96
Number of capsules per plant (CP)	992	35.75
Number of seeds per capsule (SC)	0	0.00
Days to maturity (DM)	93	3.35
Oil content (OC) (%)	519	18.70
1000-seed weight (TSW) (g)	242	8.72
Seed yield per plant (SY) (g)	263	9.48

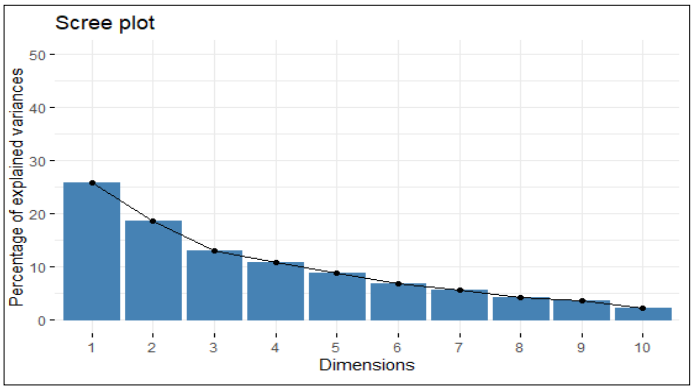


Fig 2A: Scree plot showing the proportion of total variance explained by each principal component.

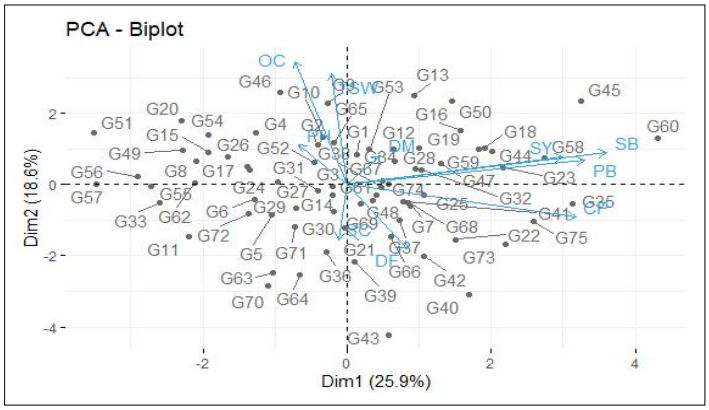


Fig 2B: PCA biplot illustrating the spatial distribution of genotypes across the first two principal components and the corresponding trait loadings.

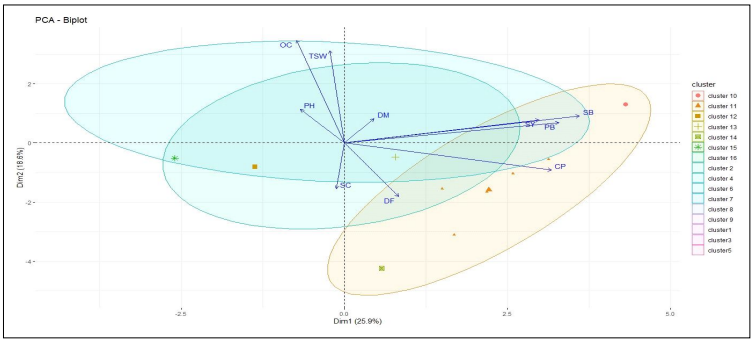


Fig 3: 2-dimensional PCA biplot, represents the relationships among clusters of data, variables and the contribution of variables to the principal components.

Table 6: Principal components extracted with eigen values, percentage of variance explained and factor loading of different traits.

Parameters	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
DF	0.13	0.32	0.61	-0.096	0.05	-0.416	0.319	-0.43	-0.025	0.17
PB	0.49	-0.12	0.09	-0.17	-0.114	0.317	0.477	0.022	0.255	-0.54
SB	0.54	-0.16	-0.02	-0.127	0.017	0.116	0.113	0.31	0.0033	0.737
PH (cm)	-0.099	-0.199	0.71	-0.099	0.27	0.19	-0.27	0.439	-0.20	-0.15
CP	0.47	0.16	-0.15	0.18	-0.047	-0.401	-0.108	0.22	-0.618	-0.29
SC	-0.017	0.27	-0.01	0.71	0.48	0.225	0.329	0.129	0.053	0.04
DM	0.07	-0.14	0.31	0.58	-0.703	0.074	-0.15	-0.029	0.102	0.06
OC (%)	-0.109	-0.61	-0.03	0.097	0.081	0.167	0.29	-0.412	-0.55	0.06
TSW (g)	-0.032	-0.55	-0.02	0.168	0.18	-0.649	0.13	0.208	0.38	-0.07
SY (g)	0.45	-0.136	0.01	0.13	0.36	0.088	-0.579	-0.49	0.213	-0.059
Cumulative variance (%)	25.88	44.46	57.58	68.38	77.17	84.05	89.67	93.99	97.69	100.00
Percentage variance (%)	25.88	18.58	13.12	10.80	8.78	6.88	5.62	4.313	3.704	2.305
Eigen value	2.588	1.858	1.312	1.080	0.879	0.688	0.562	0.432	0.371	0.231

DF: Days to 50% flowering; PB: Primary branches per plant; SB: Secondary branches per plant; PH: Plant height; CP: Number of capsules per plant; SC: Number of seeds per capsule; DM: Days to maturity; OC: Oil content; TSW: 1000-seed weight and SY: Seed yield per plant.

The scree plot illustrated the percentage difference between principal components and eigenvalues (Fig 2A). PC1 showed 25.88 per cent variability with an Eigenvalue of 2.588 in this study. The first component (PC1) recorded the highest variance compared to other PCs (Fig 2A). Due to greater variance explained by the first component, genotypes selected from this group might be helpful in breeding programs for trait improvement. The selection of clusters from PC1 is highly beneficial in trait enhancement approaches, as it exhibits the greater variability (Fig 3). The biplot diagrams represent how the trait interacts and which genotypes perform better towards traits (Fig 2B).

CONCLUSION

The outcomes of this experiment demonstrated the substantial genetic diversity among all the accessions evaluated. The clustering analysis revealed 16 distinct groups, highlighting the potential for heterosis through strategic hybridization among more distantly related clusters. The capsule number appeared as the most critical trait influencing divergence, later on the height of plant and oil content. Principal component analysis indicated that the first three PC contributed a total of 57.58% of the variance. Overall, this experiment revealed that the selection of diverse genotypes may lead to improvement in linseed yield and quality by heterotic cross combinations.

ACKNOWLEDGEMENT

The authors immensely thank the Oilseeds Research Farm of CSAUAT, Kanpur, Uttar Pradesh from the Plant Breeding and Genetics department, for rendering necessary facilities and for their moral support during the time of investigation.

Informed consent

No animals are involved during research.

Conflict of interest

The authors confirm that no competing exist.

REFERENCES

Arora, S., Modgil, R., Sood, S., Bhataria, S. (2003). Physio-chemical and nutritional quality of different cultivars of linseed (*Linum usitatissimum* L.). *Journal of Food Science and Technology-Mysore.* **40(3):** 324-327.

Cloutier, S., Ragupathy, R., Miranda, E., Radovanovic, N., Reimer, E., Walichnowski, A., Ward, K., Rowland, G., Duguid, S. and Banik, M. (2012). Integrated consensus genetic and physical maps of flax (*Linum usitatissimum* L.). *Theoretical Applied Genetics*. **125:** 1783-1795.

Crossa, J. (1990). Statistical Analysis of Multilocation Trials. *Advances in Agronomy.* **44:** 55-86.

Czemplik, M., Boba, A., Kostyn, K., Kulma, A., Mitula, A., Sztajnert, M., Wróbel Kwiatkowska, Zuk, M., Jan J., Szopa J., Telichowska, K.S. (2011). Flax Engineering for Biomedical Application. In: Biomedical Engineering, Trends, Research and Technologies, [(Eds.) Komorowska, M.A. and Olsztynska-Janus, S.]. InTech, Rijeka, Croatia. pp. 407-434.

Dabalo, D.Y., Singh, B.C.S., Weyessa, B. (2020). Genetic variability and association of characters in linseed (*Linum usitatissimum* L.) plant grown in central Ethiopia region. *Saudi Journal of Biological Sciences*. **27(8):** 2192-2206.

Darlington, C.D. and Heinisch, O. (1963). Chromosome Botany and the Origins of Cultivated Plants. 2. Aufl. George Allen and Unwin Ltd., London 1963. XVI + 231 S., Preis 35,- s." Biometrische Zeitschrift **8(1966):** 283-284.

Dillman, A.C. (1928). Natural crossing in flax. *Journal of the American Society of Agronomy.* pp 279-286.

Acquaah, G. (2012). Principles of Plant Genetics and Breeding, John Wiley and Sons, Hoboken, NJ, USA.

Gill, K.S. (1987). Introduction of Linseed. ICAR Publication, Krishi Anusandhan Bhawan, Pusa, New Delhi. pp 1-11.

Govindaraj, M., Vetriventhan, M., Srinivasan, M. (2015). Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. *Genetics Research International.* pp 1-14.

Guei, R.G., Sanni, K.A. and Fawole, A.F.J. (2005). Genetic diversity of rice (*Oryza sativa* L.). *Agronomie Africaine.* **5:** 17-28.

Brejda, J.J., Moorman, T.B., Moorman, D.L. and Dao, T.H. (2000). Identification of regional soil quality factors and indicators I. Central and southern high plains. *Soil Science Society of America Journal.* **64(6):** 2115-2124.

Kant, R., Chauhan, M.P., Srivastava, R.K. and Yadav, R. (2011). Genetic divergence analysis in linseed (*Linum usitatissimum* L.). *Indian Journal of Agricultural Research*. **45:** 59-64.

Kumar, N. and Kumar, V. (2021). Assessment of genetic diversity in linseed germplasm using morphological traits. *Electronic Journal of Plant Breeding.* **12(1):** 66-73.

Mackiewicz-Talarczyk, M., Barriga-Bedoya, J., Mankowski, J. and Pniewska, I. (2008). Global flax market situation. *ID 97 International Conferences on Flax and Other Bast Plants*. pp 408-412.

Mahalanobis, P.C. (1936). On the Generalized Distance in Statistics. In: Proceedings of the National Academy of Science (India). **2:** 49-55.

Meena, A.K., Kumar, M. (2021). Assessment of genetic diversity in linseed (*Linum usitatissimum* L.) genotypes. *Electronic Journal of Plant Breeding*. **12(2):** 597-601.

Nizar, M.A. and Mulani, R.M. (2015). Genetic diversity in indigenous and exotic linseed germplasm (*Linum usitatissimum* L.). *Electronic Journal of Plant Breeding*. **6(3):** 848-854.

Pali, V. and Mehta, N. (2015). Character association analysis for seeds yield and its components in linseed (*Linum usitatissimum* L.). *Trends in Biosciences*. **8(17):** 4573-4576.

Patial, R., Paul, S., Sharma, D., Sood, V.K., Kumar, N. (2019). Morphological characterization and genetic diversity of linseed (*Linum usitatissimum* L.). *Journal of Oilseeds and Research*. **36:** 8-16.

Przybylski, R. (2005). Flax oil and high linolenic oils. *Bailey's Industrial Oil and Fats Products*. **2:** 281-301.

Ranjana, P., Satish, P., Devender, S., Sood, V.K., Nimit, K. (2019). Morphological characterization and genetic diversity of linseed (*Linum usitatissimum* L.). *Journal of Oilseeds Research.* **36(1):** 8-16.

Rao, C.R. (1952). Advanced Statistical Methods in Biometric Research. John Wiley and Sons, Inc., New York.

Richharia, R.H. (1962). The Indian Central Oilseed Committee. Hyderabad. pp.155.

Sharma, J.R. (1998). Statistical and Biometrical Techniques in Plant Breeding. New Age international publishers, New Delhi.

Smy´kal, P., Bacˇova´-Kerteszova´, N., Kalendar, R., Corander, J., Schulman, A.H. and Pavelek, M. (2012). Genetic diversity of cultivated flax (*Linum usitatissimum* L.) germplasm assessed by retrotransposon-based markers. *Theoretical Applied Genetics*. **122:** 1385-1397.

Tewari, N., singh, M. and Singh, H.C. (2013). Genetic divergence in linseed (*Linum usitatissimum* L.). *Progressive Research.* **8:** 335-387.

Tewari, N., Singh, A. and Husain, M.F. (2020). Assessment of genetic and geographic divergence in linseed (*Linum usitatissimum* L.) genotypes. *SSRG International Journal of Agriculture and Environmental Science*. **7(4):** 48-53. ISSN: 2394-2568.

Vaisey-Genser, M., Morris, D.H. (2003) Introduction: History of the Cultivation and Uses of Flaxseed. In: Flax: The genus Linum. [Muir, A.D., Westcott, N.D. (eds)], Taylor and Francis, London. pp 1-21.

Vavilov, N.I. (1935). Studies on the origin of cultivated plants. *Bull. Bot. Pl. Progressive Agriculture*. **1(1):** 11-15.

Westcott, N.D., Muir, A.D. (2003). Flax seed lignan in disease prevention and health promotion. *Phytochemistry Reviews*. **2(3):** 401-417.