J. Agric. Res. Technol., 49 (3) : 371-382 (2024)

Comparative Study of Genetic Diversity and Principal Component Analysis in Linseed (*Linum Usitatissimum* L.) Germplasm for Agro-Climatic Conditions of Prayagraj

A. Kirankumar[1], N. Thomas[2], G. M. Lal[3], A. Dayal[4], and M. Revanth[5]
Department of Genetics and Plant Breeding, SHUATS, Naini Prayagraj - 211 007 (India)
Corresponding author Email : neha.thomas@shuats.edu.in
(Received : 14.05.2024 Accepted : 05.06.2024)

Abstract

In the present investigation, fifty linseed genotypes, including two checks, were evaluated for genetic diversity and principal component analysis at Sam Higginbottom University of Agriculture Technology and Sciences during the 2023-24 Rabi season. Data was collected on fourteen quantitative traits, with significant differences observed for most traits, except the harvest index. The minimal variations in the coefficients of variation between genotypic and phenotypic variables indicated strong additive gene action with minimal environmental influence a substantial degree of genetic advancement and heritability was discovered for number of primary branches per plant. Hierarchical The genotypes were categorized by Euclidean cluster analysis into four clusters, with clusters III and IV being largest, comprising twenty-three and sixteen genotypes, correspondingly.Clusters II and I were found to have the largest inter-cluster distance (D^2=9.79), with the highest contribution to genetic difference coming from plant seed yield. With eigenvalues greater than one, four components (PC1 through PC4) were found using principal component analysis (PCA), which accounted for 82.74% of the variance. The number of primary branches per plant and plant height were the key drivers of PC1's 53.47% variability. PC2 explained 12.26%, influenced by traits such as days to maturity and number of seeds per capsule. PC3 (9.72%) was associated with days to 50% flowering and seed yield traits, while PC4 (7.28%) was linked to the number of seeds per capsule and biological yield. In order to create superior linseed varieties with increased yield and quality, breeders can benefit greatly from these studies' insightful selection of a variety of genotypes and essential features.

Key words : Genetic variability, Cluster analysis, Principal component analysis, Heritability, *Linum usitatissimum* L.

The flax plant seed (*Linum usitatissimum*), commonly referred to as linseed or flaxseed, is a multifunctional crop farmed for both its fiber and oil. Linseed has long been used to make linen and as a source of linseed oil, which is prized for its ability to dry paint and varnish. The self-pollinating annual crop linseed (Linum usitatissimum L.2n=2x=30) is native to either the Middle East or IndiaIt is a member of the Linaceae family and genus Linum. Linnaeus gave the plant *Linum usitatissimum* its scientific name in his "Species Plantarum" (Linnaeus, 1857). Based on morphological, genetic, and molecular data, pale flax (*L. augustifolium* Huds.), the wild parent of farmed flax, is where it all began (FU *et al*. 2002).

Canada, Argentina, the United States, China, India, and several European countries are important producers of linseed (Kaur *et al*. 2023). With an average productivity of 1094 kg ha^{-1} worldwide, 3.5 million metric tons of linseed are produced on an area of 3.2 million ha. (FAO Stat 2023). It would occupy 2 lakh hectares in India in 2022–2023, producing and using 600 kg ha^{-1} and 1.2 lakh metric tons per hectare, respectively (Ministry of Agriculture and Farmers Welfare, Government of India, 2023).

Genetic divergence between parents is required to achieve maximum heterosis and to acquire transgressive segregants in the

1. M. Sc. Scholar, 2. Assistant Professor, 3. Associate Professor and 4. M. Sc. Scholar.

segregating generations since the crossing program consists of genetically dissimilar parents. Two statistical techniques that are used to explain the relative contribution of traits and find patterns of genotype-to-genotype interactions are Principal Component Analysis (PCA) and cluster analysis. Plant characteristics that identify prospective genotypes can be found using PCA, a very effective and useful technique (Chakravorty *et al*. 2013). The PCA provides information on the independent influence of a particular feature on the total variance, with each coefficient of eigenvalue reflecting the degree of contribution of each original variable to which each principal component is associated (Baraki *et al*. 2020; Teklu *et al*. 2021).

Materials and Methods

In the present investigation, 50 linseed genotype accessions were used (Table 1). A Randomized Block Design (RBD) was utilized to sow three replications of this material during rabi, 2023–2024. One row with a length of two meters was used to cultivate each genotype. Plant to plant and row to row spacing was kept at 10 and 30 cm, respectively,in the agricultural study field at Sam Higginbottom University of Agriculture Technology and Sciences, Prayagraj, Uttar Pradesh, in the Department of Genetics and Plant Breeding. From each genotype, five randomly selected plants were taken in order to document observations on twelve characteristics such as plant height (cm), days to 50% flowering, days to maturity, technical height (cm), Number of primary branches per plant, Number of secondary branches per plant, Number of capsules per plant, Number of seeds per capsule, test weight (g), Biological yield (g), Seed yield per plant (g), Harvest index (%).Using R software, the gathered data wereanalysed for genetic divergence in fifty genotypes using PCA and Wards hierarchical clustering.

Results and Discussion

With the exception of the harvest index, the analysis of variance showed that every feature in this study is displaying a notable degree of variability, at 0.01% level of significance (Table 2). These findings are consistent with those of previous studies, includingsatishpaul*et al*. (2019), Toor *et al*. (2023), Choudhary *et al*. (2023), Satish *et al*. (2022), and Prajapati *et al*. (2022). These authors also reported significant variability across a range of traits in their respective studies, highlighting the genetic diversity present in their plant populations. However, similar to the current study, they also noted instances where certain traits, such as the harvest index, exhibited limited variability. This recurrent observation across different studies

Table 1. Experimental Material used in this study

01. RL-10103	11. IC-4621	21. IC-564630	31. NEELAM	41. RST(1)20-RC-1019
02. RL-29210	12. SGB	22. RST (1) Brl-130543	32. PRATAP ALSI-2	42. RST(1)/RL-10008-7
03. RL-13519	13. IC-564622	23. NP-RR-10	33. RL-130063	43. RST(1)/RL-10008-7-1
04. IC 564609	14. IC-564623	24. IC-564632	34. NP-RR-328	44. RL-14509
05. IC-564612	15. IC-564624	25. NP-22	35. SGB-1	45. RL-10135-1
06. IC-564612-1	16. IC-564625	26. PADMINI	36. RL-10189	46. A-203(185)(SGB-19)
07. IC-564614	17. IC-564626	27. NP-23k	37. RL-10135	47. RL-10205
08. RLSP-13505	18. IC-564627	28. NP-RR-191 IDS-W-RL-26015	38. GPR-13-2/	48. RL-10203
09. RL3502	19. IC-564628	29. PARVATI	39. RST CR11/5-RL-10121	49. SHEKHAR(NC)
10. RL-13191	20. IC-564629	30. A-202B(184) (SGB-18)	40. RST(1)A/R11/MEERA	50. T-397 (CHECK)

Source: Directorate of Research, SHUATS, Prayagraj.

Table 2. Analysis of variance for fourteen characters in linseed genotypes

Source of variance	Mean sum of squares (MSS)		
Degrees of freedom	**Replication**	**Genotype**	**Error**
	2	**49**	**98**
Days to 50% flowering	38.847	32.956**	1.085
Plant Height(cm)	4.0757	12.7501**	5.2877
Technical height(cm)	8.2829	15.2381**	5.2763
Number of primary branches per plant	0.5171	3.1663**	0.3543
Number of secondary branches per plant	7.7355	16.3407**	3.1376
Capsules per plant	5.492	108.619**	15.695
Seeds per capsule	0.27227	0.79752**	0.29458
Days to maturity	61.627	173.347**	5.042
Seed yield per plant (g)	0.00048	0.35654**	0.01949
Biological yield (g)	0.04131	1.57596**	0.23234
Test weight (g)	0.87147	1.02503**	0.22902
Harvest index (%)	10.558	17.253*	10.485
Seed yield per plot (g)	4.07	1128.33**	738.68
Seed yield per hectare (kg)	373	227453**	60837

**Significant at 0.01% level of significance.
* Significant at 0.05% level of significance

suggests a potential commonality in the genetic architecture or environmental influence affecting these traits.

Genetic variability : There is evidence that environmental factors as well as genotypes contribute to the apparent variance, as seen by the greater Estimates of the variation in phenotype coefficient relative to genotypic coefficient. Greater estimates of variation in phenotypic coefficient was observed only for the number of primary branches per plant and it was recorded as moderate estimate for genotypic coefficient of variation, and moderate estimates (10-20%) of both phenotypic coefficient of variation and genotypic coefficient of variation was observed for the characters *viz*., number of secondary branches per plant and test weight. Low estimates of Phenotypic coefficient of variation and Genotypic coefficient of variation were observed among the characters days to 50% flowering, days to maturity, plant height, technical height, number of capsules per plant, number of seeds per capsule, seed yield per plant, biological yield, test weight, seed yield per plot and seed yield per hectare. It indicates that improving these characters is difficult by simple selection. These estimates are shown in Table 3.Similar, findings were reported byAnkit *et al*. (2019), Khemalata*et al*. (2019), Ruchika *et al*. (2022), Garima *et al*. (2022), Patil *et al*. (2022), Choudhary *et al*. (2023), Toor *et al*. (2023), Paliwal *et al*. (2024).

For the parameters days to 50% flowering (91%), days to maturity (92%), seed yield per plant (85%), number of primary branches per plant (73%), biological yield (66%), and seed yield per hectare (68.35%), high heritability estimates were found.This indicates that additive genes controlled these traits, and substantial estimates of genetic advancement were noted for test weight (20.82%) and the number of major branches per plant (30.29%) as a percentage of mean. For the number of major branches per plant, high heritability and strong genetic advance as a percentage of mean were found. This suggests that selection may be effective in segregating generations that improve certain traits, and that the heritability may most likely be attributable to additive gene effect. These estimates are shown in table 3.These results are consistent with the findings of previous researchers, including Chandrawati *et al*. (2016), Fekadu *et al*. (2020), Satish *et al*. (2022), Ashok *et al*. (2017), Leelavati & Mogali *et al*. (2018), Vipin *et al*. (2019), Ankit *et al*. (2019), and Ronika *et al*. (2020). These studies similarly reported high heritability and genetic advance for key agronomic traits, reinforcing the importance of these parameters in predicting the effectiveness of selection in plant breeding.

Table 3. Estimation of genetic parameters among fourteen traits for fifty linseed genotypes

Genetic Parameters	DFF	PH (cm)	TH (cm)	PB	SB	CPP	SPC	DM	SYP (g)	BY (g)	TW (g)	HI (%)	SYPP (g)	SYPH (kg)
Genotypic Variance (2g)	10.6	2.49	3.32	0.94	4.40	30.9	0.17	56.1	0.11	0.45	0.27	2.26	5.16	1340.3
Phenotypic Variance (2p)	11.7	7.78	8.60	1.29	7.54	46.6	0.46	61.1	0.13	0.68	0.49	12.74	12.70	1961.1
Genotypic Coefficient of Variation (%)	4.27	2.37	4.77	17.2	10.2	7.00	4.86	6.65	8.66	7.33	13.80	3.51	2.99	5.72
Phenotypic Coefficient of Variation (%)	4.48	4.19	7.67	20.2	13.4	8.59	8.07	6.94	9.38	9.03	18.83	8.34	4.69	6.92
Heritability (h2 %)	91	32	39	73	58	66	36	92	85	66	54	18	40.65	68.35
Genetic Advance	6.40	1.84	2.33	1.70	3.30	9.34	0.51	14.7	0.64	1.12	0.78	1.30	2.98	62.35
Genetic Advance as % of mean	8.38	2.76	6.11	30.2	16.1	11.7	6.03	13.1	16.4	12.25	20.82	3.04	3.93	9.75

DFF-Days to 50% flowering, DTM- Days to Maturity, PH- Plant height, TH- Technical height, PB-Number of primary branches per plant, SB- Number of secondary branches per plant, CPP-Number of capsules per plant, SPC- Number of seeds per capsule, TW- Test weight, BY- Biological Yield, SY/P- Seed yield per plant, HI- Harvest index.

Genetic Diversity : In the current study, wards hierarchical clustering was used to divide 50 genotypes of linseed into four clusters based on D^2 values. The highest cluster of all, cluster III, had the greatest number of genotypes (23), followed by cluster IV (16 genotypes), cluster II (6 genotypes), and cluster I (5 genotypes) (Table 4). A dendrogram is created, which illustrates how the genotypes are grouped into distinct clusters (Fig. 1).

Table 4. Distribution of fifty genotypes of linseed in different clusters on the basis of D^2 statistics

Cluster group	No. of genotypes	Name of Genotype
1	5	RL-3502, IC-564630, IC-564622, RL-29210, RL-13191
2	6	A-203(185)(SGB-19), A-202B(184)(SGB-18), RLSP-13505, RST(1)RL-10008-7, SGB, RST(1)RL-10008-7-1.
3	23	Rl-14509, IC-564623, NP-RR-328, RL-10205, NP-23K, RL-10135-1, NP-22, RL-10203, NP-RR-10,NP-RR-191, PADMINI, RST (1) Brl-130543, PARVATI, PRATAP ALSI-2, IC-564625, IC-564627, SHEKHAR(NC), IC-564614, IC-4621, SGB-1, IC-564624, IC-564628, IC-564629
4	16	IC-564626. RL-10103, RL-13159, IC-564612, IC-564612-1, RL-130063, RST(1)MEERA, RST(20)-RC-1019, T-397(CHECK), NEELAM, RL-10189, RL-10135, IDS-W-RL-26015

According to the analysis of estimates of intra- and inter-cluster heterogeneity provided by intra- and inter-cluster values, there was minimal variation across genotypes within a cluster with respect to the total influence of the 14 parameters that were being studied. The projected genetic diversity was larger between two groups the farther apart they were. Table 4 and Figure 2 showing the intra and inter-cluster distances between the four clusters. Cluster II had an intra-cluster distance of 4.38 and Cluster IV had a distance of 2.71. With a 4.17 intra-cluster distance, Cluster II was the largest, followed by Cluster III (3.30), Cluster I (3.22), and Cluster IV (2.71). Cluster II and cluster I had the highest inter-cluster distance (9.79), followed by cluster IV and cluster II (8.78), cluster III and cluster II (8.20), and cluster IV and cluster I (4.43). The lowest inter-cluster distance (3.89)

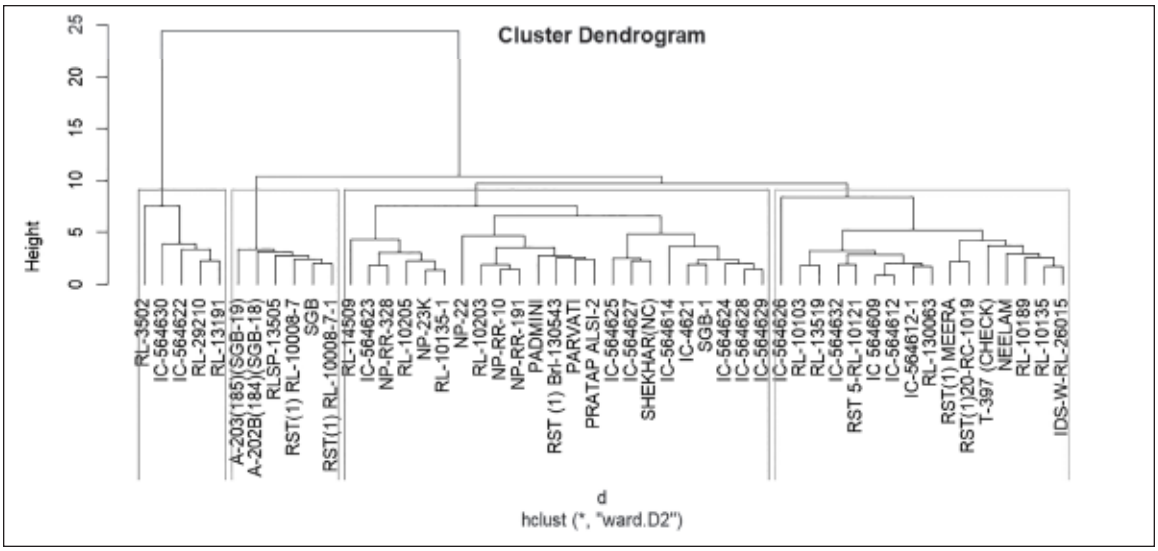

Fig. 1. Cluster Dendrogram of Linseed Genotypes Based on Ward's Method

was found between cluster III and cluster I, followed by cluster IV and cluster III (4.42), and cluster IV and cluster I (4.43).

The mean performance of the fourteen quantitative traits across the four clusters was analysed. Cluster II, which consists of five genotypes, exhibited the highest mean values for days to 50% flowering (78.12) and plant height (67.24). However, this cluster had the lowest mean values for several traits, including the number of primary branches per plant (5.22), number of secondary branches per plant (19.01), number of capsules per plant (75.49), number of seeds per capsule (8.10), biological yield (8.80), seed yield per plot (3.68), and seed yield per hectare (614.56).

Table 4. Average Intra and Inter cluster distance D^2 valuesin linseed

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	3.227064			
Cluster 2	9.795351	4.386323		
Cluster 3	3.898989	8.201341	3.305555	
Cluster 4	4.430408	8.785527	4.429041	2.718986

*Diagonal values represent Intra-cluster values and off-diagonal values represent Inter-cluster values

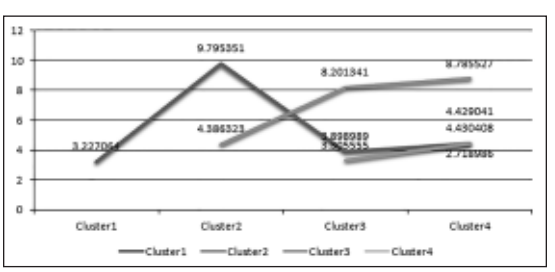

Fig. 2. Graph showing Intra and Inter cluster distance D^2 values

On the other hand, Cluster III, which includes 23 genotypes, recorded the highest mean values for days to maturity (118.36) and technical height (39.40). Despite this, it had the lowest mean values for plant height (65.21), test weight (3.41), and harvest index (42.06).

Cluster IV, comprising 16 genotypes, showed the highest mean values for a variety of traits, including the number of primary branches per plant (8.42), number of secondary branches per plant (25.94), number of capsules per plant (93.84), number of seeds per capsule (9.50), test weight (4.76), biological yield (10.63), seed yield

per plant (4.81), harvest index (45.87), seed yield per plot (81.13), and seed yield per hectare (742.74). The detailed statistics are presented in Table 5.

Table 5. Cluster mean values of different characters for 50 linseed genotypes

Clusters	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Characters				
Days to 50% flowering	78.12281	74.02381	75.91667	77.33333
Days to maturity	116.75439	107.28571	118.36111	98.53333
Plant height (cm)	67.24035	66.75714	65.21722	66.09333
Technical height (cm)	38.30526	36.98571	39.40833	38.44667
Primary branches	5.224561	5.352381	5.344444	8.426667
Secondary branches	19.01404	19.80952	21.08333	25.94667
Number of capsules per plant	75.49474	77.94762	81.90000	93.84000
Number of seeds per capsule	8.101754	8.471429	8.416667	9.506667
Test weight (g)	3.549123	3.885714	3.419444	4.760000
Biological yield (g)	8.803509	8.954762	9.250556	10.63200
Seed yield per plant (g)	3.687368	3.794762	3.868889	4.810667
Harvest index (%)	42.10676	43.29120	42.06159	45.87647
Seed yield per plot (g)	73.74737	75.89524	77.37778	81.13200
Seed yield per hectare (kg)	614.5614	632.4603	644.8148	742.7407

The previous observation, which indicated that genotypes with distinct mean performance for different features were classified into different clusters, has proven a wide range in cluster mean from one cluster to another.

Seed yield per plant exhibited the highest contribution towards genetic divergence, accounting for 23.5% of the total variability. This trait was closely followed by days to maturity (23.1%) and days to 50% flowering (21.1%), indicating their critical roles in influencing genetic diversity which was depicted in Table 6 and fig 3.Scatter plot matrix (Fig. 4) visualizes bivariate relationships between combinations of variables used in this study for cluster analysis. These findings are consistent with earlier reports by Nizar *et al*. (2015), Singh *et al*. (2015), Kumar *et al*. (2017), Amit *et al*. (2017), and Ankit *et al*. (2019), where traits such as seed yield and flowering time were found to contribute significantly to genetic divergence in linseed and related crops.

Table 6. Per cent contribution of fourteen characters towards divergence

Characters	Percent contribution
Days to 50% flowering	21.1%
Days to maturity	23.1%
Plant height (cm)	2.4%
Technical height (cm)	2.8%
Primary branches	5.8%
Secondary branches	3.3%
Number of capsules per plant	4.4%
Number of seeds per capsule	2.2%
Test weight (g)	3.2%
Biological yield (g)	3.6%
Seed yield per plant (g)	23.5%
Harvest index (%)	1.5%
Seed yield per plot (g)	3%
Seed yield per hectare (kg)	0%

Principal component analysis (PCA) :

The four components PC1, PC2, PC3, and PC4 accounted for 53.472%, 12.263%, 9.722%, and 7.278% of the variations among

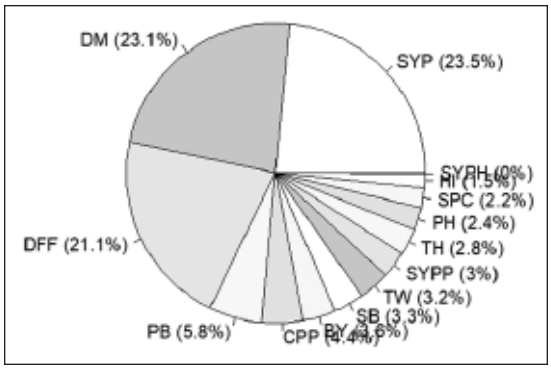

Fig. 3. Pie chart showing percentage contribution of traits towards total divergence

the characters, respectively, and the PCA result explained the genetic variability of the linseed germplasm (Table 7). Out of the complex fourteen components, the first four main PCAs (PC1, PC2, PC3, and PC4) accounted for 82.74% of the total cumulative variance. The eigen vectors decreased significantly from PC1 (53.47%) to PC4 (7.28%). This shows that beyond PC4, additional principal components did not explain significant variation. Thus, only four PCs were considered. The results of the component matrix revealed that PC 1 with an eigen value 7.49 which accounted for highest variation 53.47% showed high positive loadings for, plant height (0.238), technical height (0.019), number of primary branches per plant (0.023), it showed high negative loadings for days to 50% flowering (-0.020), number of secondary branches per plant (-0.331), number of capsules per plant (-0.329), number of seeds per capsule (-0.337), seed yield per plant (-0.251), biological yield (-0.320), test weight (-0.359), harvest index (-0.119), seed yield per plot (-0.309), seed yield per hectare (-0.349). Principal component II with an eigen value 1.72 which accounted for 12.26% variation showed positive loadings for days to maturity (0.111), number of primary branches per plant (0.385), number of secondary branches per plant (0.017), number of capsules per plant (capsule (0.065), harvest index (0.458) and it showed negative loadings for days to 50% flowering (-0.522), plant height (-0.1440, technical height (-0.505), seed yield per plant (-0.130), biological yield (-0.232), test weight (-0.019), seed yield

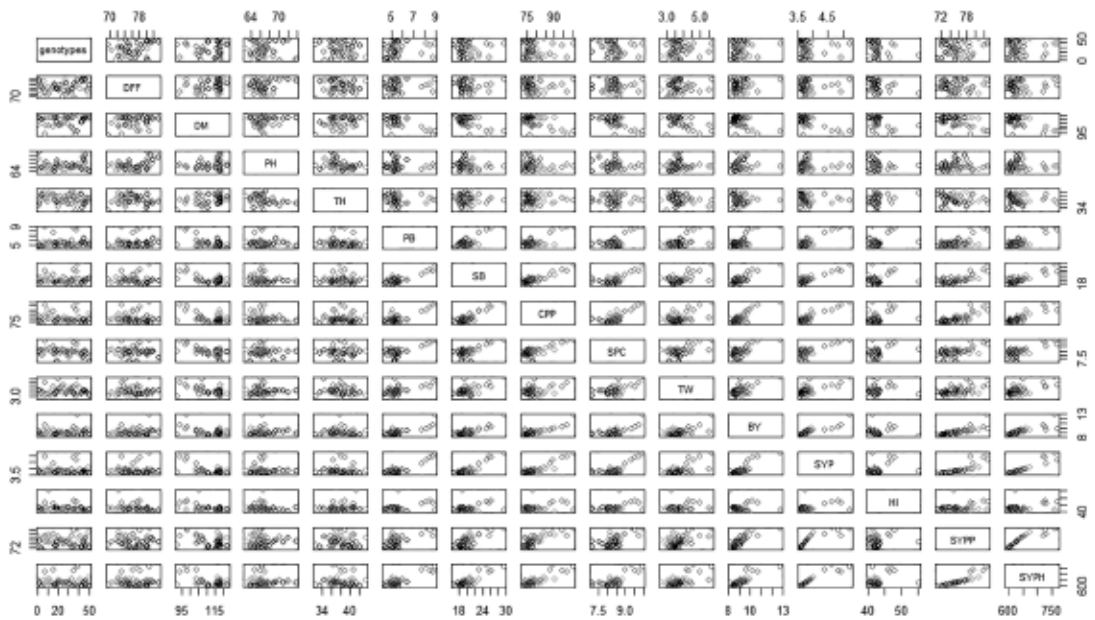

Fig. 4. Scatterplot Matrix of different Traits Across Linseed Genotypes

per plot (-0.075), seed yield per hectare (-0.015). Principal component III with an eigen value 1.36 which accounted for 9.72% variation showed positive loadings for days to 50% flowering (0.324), plant height (0.514), number of primary branches per plant (0.599), number of capsules per plant (0.144), number of seeds per capsule (0.157), biological yield (0.168), test weight (0.041), seed yield per plot (0.083), and seed yield per hectare (0.045). It showed negative loadings for technical height (-0.309), number of secondary branches per plant (-0.001), days to maturity (-0.068), seed yield per plant (-0.142), and harvest index (-0.260). Principal component IV with an eigen value 1.02 which accounted for 7.28% variation showed positive loadings for number of capsules per plant (0.145), number of seeds per capsule (0.098), biological yield (0.188), seed yield per plot (0.097), it showed negative loadings for days to 50%v flowering (-0.449), plant height (-0.022), technical height (-0.442), number of primary branches per plant (-0.357), number of secondary branches per plant (0.198), days to maturity (-0.070), seed yield per plant (-0.012), test weight (-0.041), harvest index (-0.595), and seed yield per hectare (-0.026).Projection of 14 quantitative characters on the first two components of PCAwas depicted in fig 5 using major PCA vectors showing distribution of different traits.

The percentage of variation associated with each PC was derived by constructing a graph between eigen values and principal component numbers shown in Fig 6. With an eigen value of 7.49 PC1 explains the 53.47% of genetic variability. From, PC1 to PC4 eigen values decrease drastically from 7.49 to 1.02 respectively. After PC4 bar graph began to straighten, an elbow-shaped line was formed. It clearly demonstrates that PC1 had the most variation. Similarly, Chatfield, C. and Collins, A. J. (1980suggested that in order to deal with fewer components, components with an eigen value of less than one should be removed. Sharma, J.R. (1998) reported that the significance of the biggest contributor to the overall variation at each differentiation axis is reflected by PCA. Fenty, J. 2004stated that a big set of variables is reduced to a smaller set of components by PCA. Individual and biplot loadings of the different traits and the genotypes distributed throughout the first two main componentsare given in fig 7 and fig 8. The similar results were reported earlier by Adugna *et al*. (2003), Rahimi *et al*. (2011), Kumar *et al*. (2015), Kumar *et al*. (2016), Kaur *et al*. (2018), Fekadu *et al*. (2020), Ronika *et al*. (2020), Sandhya *et al*. (2022), Sandhya *et al*. (2022), Sandhya *et al*. (2023).

Principal component analysis (PCA) : The four components PC1, PC2, PC3, and PC4 accounted for 53.472%, 12.263%, 9.722%, and 7.278% of the variations among the characters, respectively, and the PCA result explained the genetic variability of the linseed germplasm(Table 7).Out of the complex fourteen components, the first four main PCAs (PC1, PC2, PC3, and PC4) accounted for 82.74% of the total cumulative variance. The eigen vectors decreased significantly from PC1 (53.47%) to PC4 (7.28%). This shows that beyond PC4, additional principal components did not explain significant variation. Thus, only four PCs were considered. The results of the component matrix revealed that PC 1 with an eigen value 7.49 which accounted for highest variation 53.47% showed high positive loadings for, plant height (0.238), technical height (0.019), number of primary branches per plant (0.023), it showed high negative loadings for days to 50% flowering (-0.020), number of secondary branches per plant (-0.331), number of capsules per plant (-0.329), number of seeds per capsule (-0.337), seed yield per plant (-0.251), biological yield (-0.320), test weight (-0.359), harvest index (-0.119), seed yield per plot (-0.309), seed yield per hectare (-0.349).

Principal component II with an eigen value 1.72 which accounted for 12.26% variation showed positive loadings for days to maturity (0.111), number of primary branches per plant (0.385), number of secondary branches per plant (0.017), number of capsules per plant (capsule (0.065), harvest index (0.458) and it showed negative loadings for days to 50% flowering -0.522), plant height (-0.1440, technical height (-0.505), seed yield per plant (-0.130), biological yield (-0.232), test weight (-0.019), seed yield per plot (-0.075), seed yield per hectare (-0.015). Principal component III with an eigen value 1.36 which accounted for 9.72% variation showed positive loadings for days to 50% flowering (0.324), plant height (0.514), number of primary branches per plant (0.599), number of capsules per plant (0.144), number of seeds per capsule (0.157), biological yield (0.168), test weight (0.041), seed yield per plot (0.083), and seed yield per hectare (0.045). It showed negative loadings for technical height (-0.309), number of secondary branches per plant (-0.001), days to maturity (-0.068), seed yield per plant (-0.142), and harvest index (-0.260). Principal component IV with an eigen value 1.02 which accounted for 7.28% variation showed positive loadings for number of capsules per plant (0.145), number of seeds per capsule (0.098), biological yield (0.188), seed yield per plot (0.097), it showed negative loadings for days to 50%v flowering (-0.449), plant height (-0.022), technical height (-0.442), number of primary branches per plant (-0.357), number of secondary branches per plant (0.198), days to maturity (-0.070), seed yield per plant (-0.012), test weight (-0.041), harvest index (-0.595), and seed yield per hectare (-0.026).Projection of 14 quantitative characters on the first two components of PCAwas depicted in fig 5 using major PCA vectors showing distribution of different traits.

The percentage of variation associated with each PC was derived by constructing a graph between eigen values and principal component numbers shown in Fig 6. With an eigen value of 7.49 PC1 explains the 53.47% of genetic variability. From, PC1 to PC4 eigen values

Table 7. Principal Component Analysis (PCA) Loadings, Eigenvalues, and Variance Explained for different Traits in Linseed

Characters	Principal components			
	PC1	PC2	PC3	PC4
Days to 50% flowering	-0.020	-0.522	0.324	-0.449
Days to maturity	0.238	-0.144	0.514	-0.022
Plant height (cm)	0.019	-0.505	-0.309	-0.442
Technical height (cm)	0.023	0.385	0.599	-0.357
Primary branches	-0.331	0.017	-0.001	-0.198
Secondary branches	-0.329	0.012	0.144	0.145
No. of capsules per plant	-0.337	0.065	0.157	0.098
No. of seeds per capsule	-0.290	0.111	-0.068	-0.070
Test weight (g)	-0.251	-0.130	-0.142	-0.012
Biological yield (g)	-0.320	-0.232	0.168	0.188
Seed yield per plant (g)	-0.359	-0.019	0.041	-0.041
Harvest index (%)	-0.119	0.458	-0.260	-0.595
Seed yield per plot (g)	-0.309	-0.075	0.083	0.097
Seed yield per hectare (kg)	-0.349	-0.015	0.045	-0.026
Eigen values	7.49	1.72	1.36	1.02
% variance	53.47	12.26	9.72	7.28
Cumulative proportion	53.47	65.74	75.46	82.74

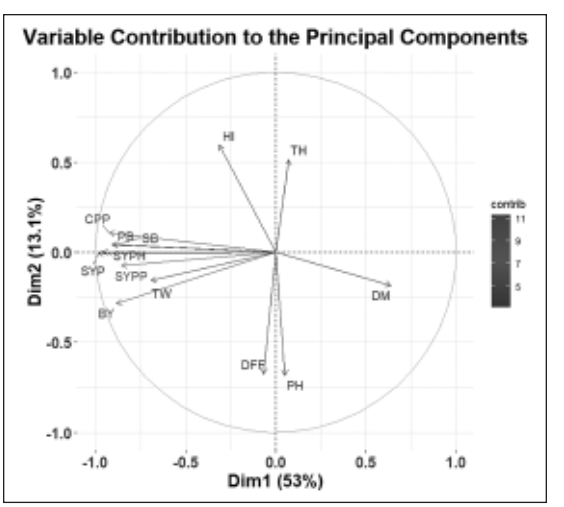

Fig. 5. Projection of 14 quantitative characters on the first two components of PCA

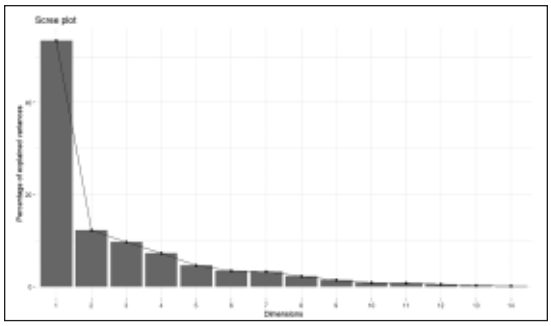

Fig. 6. Scree plot showing eigen value and percentage of variation explained by different components

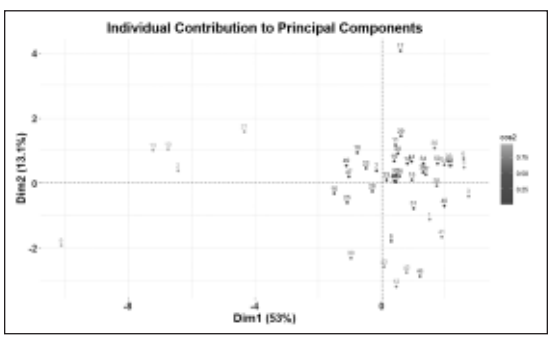

Fig. 7. Individual contribution of 50 linseed genotypes

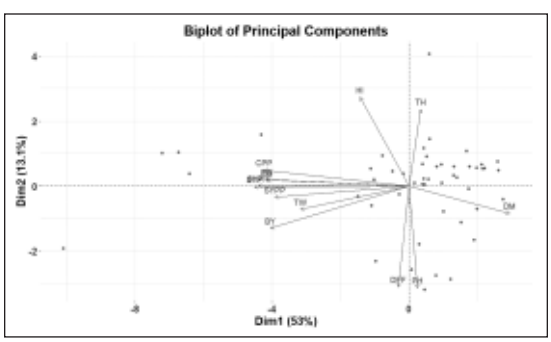

Fig. 8. Biplot of 50 linseed genotypes on principal component axis 1 and 2

decrease drastically from 7.49 to 1.02 respectively. After PC4 bar graph began to straighten, an elbow-shaped line was formed. It clearly demonstrates that PC1 had the most variation. Similarly, Chatfield, C. and Collins, A. J. (1980suggested that in order to deal with fewer components, components with an eigen value of less than one should be removed. Sharma, J.R. (1998) reported that the significance of the biggest contributor to the overall variation at each differentiation axis is reflected by PCA. Fenty, J. 2004stated that a big set of variables is reduced to a smaller set of components by PCA. Individual and biplot loadings of the different traits and the genotypes distributed throughout the first two main componentsare given in fig 7 and fig 8. The similar results were reported earlier by Adugna *et al*. (2003), Rahimi *et al*. (2011), Kumar *et al*. (2015), Kumar *et al*. (2016), Kaur *et al*. (2018), Fekadu *et al*. (2020), Ronika *et al*. (2020), Sandhya *et al*. (2022), Sandhya *et al*. (2022), Sandhya *et al*. (2023).

Conclusion

The current study reveals considerable genetic diversity among the 50 linseed genotypes. It is concluded that linseed genotypes RL-3502, IC-564622, Rl-13191, IC-564630, Rl-29210, were found best as depicted highest according to PCA biplot and mean performance for each character.These results demonstrated that, relative to the mean for the number of primary branches per plant, high heritability and high genetic advancement are notably, promising for selection in breeding programs due to their substantial genetic control and lesser environmental influence and traits showing high heritability and moderate genetic advance in days to maturity, seed yield per plant, biological yield these were indicating a greater portion of their inheritance coming from non-additive gene action. These features could be enhanced by heterosis breeding. Clusters I and II were found to have the greatest inter-cluster distances based on genetic diversity research. This suggests that the genotypes within this cluster are highly diverse and may be employed as parents in upcoming breeding initiatives.

References

Adugna, W. and Labuschagne, M. T. 2004. Diversity analysis in Ethiopian and some exotic collections of linseed. S. Afr. J. Plant Soil. 21(1): 53-58.

Amit Kumar, S. A., Kerkhi and Rohit Kumar. 2017 Genetic diversity (D2) analysis in Linseed (*Linum usitatissimum* L.). J. Pharmacogn. Phytochem. 6(4): 886-888.

Anam Rizvi, Shailesh Marker and Ashok Choudhary. 2018. Multivariate analysis in linseed (*Linum usitatissimum* L.). J. Pharmacogn. Phytochem.7(2):528-531.

Ankit, S. P. S., Singh, V. K., Singh, A., Singh, A. K., Tiwari, A. and Singh, P. 2019. Character association and genetic divergence analysis in linseed (*Linum usitatissimum* L.). J. Pharmacogn. Phytochem. 8(3): 348-351.

Ashok, K. M., Sandhya and Manoj, K. 202.) Assessment of genetic diversity in linseed (*Linum usitatissimum* L.) genotypes. Electron. J. Plant Breed.12(2): 597-601.

Baraki, F., Gebregergis, Z., Belay, Y., Berhe, M, Teame G., Hassen, M. and Araya, G. 202.) Multivariate analysis for yield and yield-related traits of sesame. (*Sesamum indicum* L.) genotypes. Heliyon. 6(10).

Chandrawati, N., Singh, R., Kumar, S., Kumar and Yadav, H. K. 2016. Genetic variability and interrelationship among morphological and yield traits in linseed (*Linum usitatissimum* L.). Genetika.48(3): 881-892.

Chakravorty, A., Ghosh, P. D. and Sahu, P. K. 2012. Multivariate analysis of phenotypic diversity of landraces of rice of West Bengal. Am. J. Exp. Agric. 3(1): 110-123.

Choudhary, S. K., Nirala, R. B. P., Anand, S. and Kumar, A. 2023 Character association and path coefficient analysis for yield and yield components in linseed (*Linum usitatissimum* L.). Biol. Forum-Int. J. 15(3): 40-44.

Food and Agriculture Organization of the United Nations (2023) FAOSTAT Statistical Database.

Fekadu Amsalu. 2020. Estimates of heritability, genetic, and principal components analysis for yield and its traits in linseed genotypes (*Linum usitatissimum* L.) in Central Highlands of Ethiopia. Int. J. Res. Agric. For. 7(9):1-6.

Fu, Y. B., Diederichsen, A., Richards, K. W. and Peterson, G. W. 2002. Genetic diversity within a range of germplasm accessions of cultivated flax and its implications for conservation. Theor. Appl. Genet. 104(4): 563-569.

Garima, T., Satish, P. and Ronika, T. 2022. Genetic variability for seed yield and its component traits in linseed (*Linum usitatissimum* L.). Himachal J. Agric. Res. 48(2): 151-156.

Kaur, V., Kumar, S., Yadav, R., Wankhede, D. P., Aravind, J, Radhamani, J., Rana, J. C. and Kumar, A. 2018. Analysis of genetic diversity in Indian and exotic linseed germplasm and identification of trait-specific superior accessions. J. Environ. Biol. 39: 702-709.

Khemlata, T,, Biswas, K. and Mehta, N. 2019. Variability analysis for seed yield and its contributing traits among flax type linseed (*Linum usitatissimum* L.) genotypes in Chhattisgarh. Int. J. Chem. Stud. 7(5): 1633-1636.

Kumar, N. and Paul, S. 2016. Selection criteria of linseed genotypes for seed yield traits through correlation, path coefficient, and principal component analysis. J. Anim. Plant Sci. 26(6): 1688-1695.

Kumar, A., Kerkhi, S. A. and Kumar, R. 2017. Genetic diversity (D^2) analysis in linseed (*Linum usitatissimum* L.). J. Pharmacogn. Phytochem. 6(4): 886-888.

Leelavathi, T. M. and Mogali, S. C. 201.) Genetic variability, character association, and path analysis for yield and yield components in mutant population of linseed (*Linum usitatissimum* L.). J. Farm Sci. 31(1): 17-20.

Linnaeus, C. 1857. Species Plantarum. Printed for Royal Society of London, UK.

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

Paliwal, S., Tripathi, M. K., Sikarwar, R. S. and Sharma, S. 2024. Study of genetic variability and path analysis for yield attributing traits in linseed (*Linum usitatissimum* L.). Int. J. Plant Soil Sci. 36(7): 20-29.

Patil, A. R, Wadikar, P. B., Rao, M. V., Mundhe, S. S. and Patil, S. S. 2022. Genetic variability for quantitative traits in linseed (*Linum usitatissimum* L.). Pharma Innov. J. 12(1): 488-490.

Prajapati, P. K., Ratan. S., Singh, Y., Pyasi, D. K., Rakesh, K. and Thakur, S. 2022. Genetic variability and path analysis for yield and associated traits in linseed genotypes. Int. J. Plant Soil Sci. 34(24): 685-691.

Rahimi, M. M., Zarei, M. A. and Arminian, A. 2011. Selection criteria of flax (*Linum usitatissimum* L.) for seed yield, yield components, and biochemical compositions under various planting dates and nitrogen. Afr. J. Agric. Res. 6(13): 3167-3175.

Ruchika, D., Satish, P. and Satasiya, P. 2020. Assessment of genetic variability, correlation, and path analysis for yield and its components in linseed (*Linum usitatissimum* L.). J. Pharmacogn. Phytochem. 9(5): 2384-2389.

Ronika, T., Satish, P. and Satasiya, P. 2020. Genetic variability and path analysis for yield and its related traits

in linseed (*Linum usitatissimum* L.). Int. J. Curr. Microbiol. Appl. Sci. 9(10): 2579-2586.

Sandhya, Manoj K. and Pawan, K. 202.) Principal component and correlation studies in linseed under subtropical region of Rajasthan, India. Environ. Ecol. 40(1): 93-98.

Sandhya, Mukul, Manoj K., Pawan, K., Akash, G., Deepak, M. and Ravi Kishan S. 2022. Assessment of principal component analysis and genetic diversity among linseed (*Linum usitatissimum* L.) germplasm under humid South Eastern Plain of Rajasthan. Agric. Mech. Asia. 53(2).

Sandhya, Manoj K., Mukul, Pawan, K. and Akash, G. 2023. Multivariate analysis of genetic diversity in linseed genotypes for yield and its component traits. Int. J. Environ. Clim. Change. 13(12): 537-544.

Satish paul and Nimit kumar. 2019 Selection Strategy and Estimation of Interrelationships for Improvement of Seed Yield and Other Related Traits in Linseed (*Linum usitatissimum* L.). Indian Journal of Plant Genetic Resources. 32(01): 28-35.

Sathish, K., Parte, D., Rohit, M. V. S. K., Roy, S., Sahu, J. and Mehta, N. 2022. Genetic-morphological characterization and diversity analysis of linseed (*Linum usitatissimum* L.) germplasm. Int. J. Plant Soil Sci. 34(22): 1655-1663.

Singh, P., Mehta, N., Nair, S. and Sao, A. 2015. Genetic divergence analysis for seed classes of four popular linseed (*Linum usitatissimum* L.) varieties of Chhattisgarh. Int. J. Farm Sci. 5(2): 42-46.

Teklu, D. H., Shimelis, H., Tesfaye, A., Mashilo, J., Zhang X., Zhang, Y. and Shayanowako, A. I. T. 2021. Genetic variability and population structure of Ethiopian sesame (*Sesamum indicum* L.) germplasm assessed through phenotypic traits and simple sequence repeats markers. Plants. 10(6): 1129.

Toor, A. K., Alka and Kumar, A. 2023. Genetic variability and character association in linseed (*Linum usitatissimum* L.) genotypes for yield. Environ. Ecol. 41(4C): 2928-2938.

Wang, Y., Zhang, X., Li, J. and Smith, B. 2007. Genetic diversity and geographic differentiation in cultivated flax (*Linum usitatissimum* L.). Theor. Appl. Genet. 114(8): 1517-1527.

Ministry of Agriculture & Farmers Welfare, Government of India (2023) Agriculture Statistics at a Glance.