



Plant Archives

Journal homepage: <http://www.plantarchives.org>

DOI Url : <https://doi.org/10.51470/PLANTARCHIVES.2026.v26.supplement-2.257>

MULTIVARIATE ANALYSIS OF GENERAL COMBINING ABILITY EFFECTS IN LINSEED (*Linum usitatissimum* L.) USING PRINCIPAL COMPONENT ANALYSIS AND HIERARCHICAL CLUSTERING

Helee Patel¹, Pranay C. Patel², Himani P. Vadodariya², N. V. Soni², H. N. Zala¹, A. M. Patel² and S. K. Shah¹

¹C. P. College of Agriculture, Sardarkrushinagar Dantiwada Agricultural University (SDAU), Sardarkrushinagar – 385 506, Gujarat, India.

²Anand Agricultural University (AAU), Anand – 388 110, Gujarat, India.

*Corresponding author E--mail: pranaypatel0407@aaui.in

(Date of Receiving : 17-04-2026; Date of Revision : 09-06-2026; Date of Acceptance : 01-07-2026)

ABSTRACT

Principal Component Analysis (PCA) and Ward's hierarchical clustering were applied to the General Combining Ability (GCA) matrix of 10 linseed (*Linum usitatissimum* L.) parents (4 lines $\times$ 6 testers) across 15 agronomic and oil-quality traits (rabi 2024–25). Three principal components explained 81.86% of total variance: PC1 (47.60%) represented a yield–agronomic axis, PC2 (21.55%) captured a fatty-acid axis reflecting Δ -15 desaturase antagonism and PC3 (12.71%) separated palmitic acid from seeds per capsule. SHA-7 exhibited the highest yield-GCA (PC1 score +4.93), while Varsa Alsi showed unique high- ω -3 GCA (PC2 score –4.19). Four parent clusters and four trait clusters were identified. The MGIDI ranked Varsa Alsi first (1.295) and SHA-7 second (4.633), confirming SHA-7 $\times$ Varsa Alsi as the optimal cross for simultaneous yield and ω -3 improvement.

Keywords : Principal component analysis, general combining ability, linseed, line $\times$ tester, MGIDI

Introduction

Linseed (*Linum usitatissimum* L.) is valued simultaneously as an industrial fibre source, edible oil and nutraceutical, owing to its high alpha-linolenic acid (ALA, ω -3) content of total 40–55% oil (Goyal *et al.*, 2014). Breeding strategies maximising seed yield may not simultaneously maximise ω -3 content, making the identification of parents with complementary genetic background a critical priority (Bhateria *et al.*, 2006).

Line $\times$ tester design (Kempthorne, 1957) remains a standard framework for estimating general combining ability (GCA). GCA effects, attributable to additive gene action, are essential for identifying superior recurrent parents. However, when GCA is estimated across many traits simultaneously, univariate interpretation provides limited insight into parental genetic relationships or trait co-inheritance structure (Sharma *et al.*, 2023).

Principal Component Analysis (PCA) reduces dimensionality of complex datasets while retaining maximum variance (Pearson, 1901; Greenacre *et al.*, 2022). When applied to GCA matrices, PCA reveals parental clusters with similar additive constitutions, identifies orthogonal genetic dimensions corresponding to distinct trait complexes and visualises parent–trait relationships through biplots (Gabriel, 1971). Such approaches have been applied in cotton (Ullah *et al.*, 2022), field pea (Sharma *et al.*, 2023) and *Cucurbita pepo* (Chikh-Rouhou *et al.*, 2024), but not to a linseed line $\times$ tester GCA matrix along with a agronomic and fatty-acid quality traits simultaneously. Hierarchical clustering using Ward's minimum variance method (Ward, 1963) complements PCA through dendrogram-based classification.

The present study applies PCA and Ward's hierarchical clustering to the GCA matrix from a 4-line $\times$ 6-tester programme in linseed evaluated for 15 traits, with the objectives of: (i) identifying principal genetic

dimensions (ii) classifying parents (iii) characterising trait co-inheritance and (iv) identifying complementary parental combinations for hybrid breeding.

Materials and Methods

Experimental material comprised of different 4 lines (SHA-4, SHA-5, SHA-6, SHA-7) used as females and 6 different testers (Azad Alsi-1, Varsa Alsi, Kota barani Alsi-3, Kota barani Alsi-4, JLS-66, T-397) used as males. A line × tester crossing programme was conducted at the Botanical Garden, Department of Genetics and Plant Breeding, SDAU, during rabi 2023–24. The resulting 24 F₁ hybrids, all parental sets and check variety JLS-95 were evaluated in a randomised complete block design with three replications at the Centre for Vegetable Research, Chimanbhai Patel College of Agriculture, SDAU, during rabi 2024–25. Data were recorded for 15 different traits, which includes, days to 50% flowering (DF), days to maturity (DM), plant height (PH), primary branches per plant (PB), capsules per plant (NCP), capsule size (CS), seeds per capsule (NSC), seed yield per plant (SYP), 1000-seed weight (TSW), oil content (OC), palmitic acid (PA), stearic acid (SA), oleic acid (OA), linoleic acid (LA) and linolenic acid (LNA).

GCA effects were estimated using the line × tester method (Kempthorne, 1957; Panse and Sukhatme, 1978). The 10 × 15 GCA matrix was standardised column-wise (z-score) to ensure equal trait contribution. PCA was performed via singular value decomposition; components were retained by the Kaiser criterion (Kaiser, 1960; eigenvalue > mean;

>6.67% individual variance for a 15-trait matrix). Biplots for PC1 × PC2 and PC1 × PC3 were constructed with trait-loading vectors scaled by 3.3 (Gabriel, 1971). Double-clustered heatmaps were produced by Ward's minimum variance clustering (Ward, 1963) using Euclidean distance. The MGIDI (Olivoto and Lúcio, 2020; Olivoto, 2022) was applied to rank parents by proximity to a simultaneous multi-trait ideotype. All analyses were performed in Python 3 using scikit-learn (Pedregosa *et al.*, 2011), SciPy (Virtanen *et al.*, 2020) and Matplotlib.

Results and Discussion

Combining ability ANOVA and GCA effects

The analysis of variance for combining ability (Table 1) revealed significant mean squares due to lines for days to flowering, days to maturity, plant height, number of primary branches per plant, capsule size, number of seeds per capsule, seed yield per plant, 1000-seed weight, oil content, palmitic acid, stearic acid, oleic acid and linoleic acid. Mean squares due to testers were significant for plant height, number of seeds per capsule, seed yield per plant, oleic acid, linoleic acid and linolenic acid. The lines × testers interaction was significant for all traits except 1000-seed weight and oil content, indicating the involvement of both additive and non-additive gene action. The $\sigma^2_{gca}/\sigma^2_{sca}$ ratio exceeded unity for days to flowering, plant height, capsule size, 1000-seed weight and stearic acid, confirming additive gene action for these traits. For the remaining traits the ratio was less than unity, indicating predominance of non-additive gene action (Table 1).

Table 1 : Analysis of variance (mean squares) and variance components for combining ability across 15 traits in linseed.

Source	df	DF	DM	PH	PB	NCP	CS	NSC	SYP	TSW	OC	PA	SA	OA	LA	LNA
Replications	2	4.13	1.09	7.53	0.02	0.98	0.01	0.03	0.46*	0.48	30.93**	0.01	0.34*	0.18	0.17	0.06
Lines	3	166.46**	23.38*	714.45**	4.05**	342.34	4.43**	3.49**	14.60**	5.01**	9.96*	0.56*	3.49**	31.75**	2.52**	16.09
Testers	5	13.61	11.96	195.13**	0.63	55.44	0.15	1.04**	0.54**	0.68	3.97	0.30	1.23	71.89**	3.28**	76.97**
Lines × Testers	15	6.35*	6.48*	27.17**	0.71**	287.35**	0.32**	1.31**	1.68**	0.70	5.52	0.15**	0.45**	12.31**	1.14**	13.70**
Error	46	2.93	2.86	7.73	0.14	5.26	0.01	0.08	0.10	0.42	3.13	0.04	0.08	0.22	0.14	0.39
σ^2_{gca}		5.58**	0.75	28.51**	0.10	–	0.13**	0.06*	0.39	0.14	0.11	0.02*	0.13**	2.63**	0.12**	2.19**
σ^2_{sca}		1.14*	1.20*	6.48**	0.19**	94.04**	0.10**	0.41**	0.52**	0.09	0.71	0.04**	0.12**	4.03**	0.33**	4.44**
$\sigma^2_{gca}/\sigma^2_{sca}$		4.90	0.62	4.40	0.53	–	1.30	0.15	0.75	1.56	0.15	0.50	1.08	0.65	0.36	0.49
Gene action		Add.	Non-add.	Add.	Non-add.	Non-add.	Add.	Non-add.	Non-add.	Add.	Non-add.	Non-add.	Add.	Non-add.	Non-add.	Non-add.

DF = days to 50% flowering, DM = days to maturity, PH = plant height, PB = primary branches/plant, NCP = capsules/plant, CS = capsule size, NSC = seeds/capsule, SYP = seed yield/plant, TSW = 1000-seed weight, OC = oil content, PA = palmitic acid, SA = stearic acid, OA = oleic acid, LA = linoleic acid, LNA = linolenic acid. * and ** significant at 5% and 1%, respectively. – indicates negative estimate. Add. = additive; Non-add. = non-additive.

GCA effects of 10 parents across all 15 traits are presented in Table 2. Among the lines, SHA-7 showed significant positive GCA for seed yield per plant (SYP = +1.23**), capsules per plant (NCP = +3.27**), 1000-seed weight (TSW = +0.76**), capsule size (CS = +0.71**), seeds per capsule (NSC = +0.28**) and primary branches per plant (PB = +0.58**), identifying

it as the best general combiner for yield. Among testers, Varsa Alsi exhibited the highest positive GCA for linolenic acid (LNA = +4.76**) and negative GCA for oleic acid (OA = –4.51**), reflecting Δ -15 desaturase pathway antagonism. These raw GCA values formed the 10 × 15 input matrix subjected to PCA and MGIDI.

Table 2 : General combining ability (GCA) effects of 10 linseed parents (4 lines and 6 testers) for 15 traits.

	DF	DM	PH	PB	NCP	CS	NSC	SYP	TSW	OC	PA	SA	OA	LA	LNA
SHA-4	-2.88**	-0.71*	-5.13**	-0.36**	-1.89**	-0.41**	-0.23**	-0.64**	-0.29	0.64	0.07	-0.19**	1.36**	0.25**	-1.36**
SHA-5	-2.38**	1.18**	-5.57**	-0.42**	-5.26**	-0.26**	-0.50**	-0.71**	-0.40*	0.13	0.06	-0.43**	0.80**	-0.51**	0.14
SHA-6	2.46**	-1.21**	3.82**	0.20**	3.88**	-0.04**	0.45**	0.12	-0.07	-1.07**	-0.26**	0.02	-0.58**	-0.03	0.80**
SHA-7	2.79**	0.74*	6.88**	0.58**	3.27**	0.71**	0.28**	1.23**	0.76**	0.30	0.13**	0.60**	-1.57**	0.30**	0.42**
SE± (lines)	0.4031	0.3988	0.6552	0.0880	0.5406	0.0135	0.0671	0.0759	0.1534	0.4171	0.0487	0.0678	0.1126	0.0891	0.1477
Azad Alsi-1	-0.74	0.79	-3.55**	-0.02	1.71*	0.01	0.23**	0.07	0.10	-0.16	-0.14*	0.11	1.92**	-0.44**	-1.37**
Varsa Alsi	0.85	0.46	7.15**	0.09	1.91**	-0.12**	0.35**	0.35**	0.43*	0.99	-0.21**	-0.55**	-4.51**	0.90**	4.76**
Kb Alsi-3	-1.57**	0.13	-1.32	0.39**	-0.19	-0.11**	-0.17*	-0.18	-0.24	-0.15	0.06	0.39**	2.14**	-0.45*	-2.25**
Kb Alsi-4	0.76	-0.96*	-3.84**	-0.02	-0.38	0.17**	-0.12	-0.05	-0.07	0.06	-0.03	-0.14	-0.62*	-0.17	0.79**
JLS-66	-0.40	-1.46**	0.66	0.22*	-3.94**	0.08**	-0.44**	-0.23*	-0.07	0.05	0.23**	0.19*	0.08	-0.15	-0.83**
T-397	1.10	1.04*	0.89	-0.23*	0.91	-0.02	0.15	0.03	-0.15	-0.79	0.09	0.01	0.98**	0.31*	-1.09**
SE± (testers)	0.494	0.489	0.803	0.108	0.662	0.017	0.082	0.093	0.188	0.511	0.060	0.083	0.138	0.109	0.181

DF = days to 50% flowering, DM = days to maturity, PH = plant height (cm), PB = primary branches/plant, NCP = capsules/plant, CS = capsule size (mm), NSC = seeds/capsule, SYP = seed yield/plant (g), TSW = 1000-seed weight (g), OC = oil content (%), PA = palmitic acid (%), SA = stearic acid (%), OA = oleic acid (%), LA = linoleic acid (%), LNA = linolenic acid (%). * and ** significant at 5% and 1%, respectively.

Variance decomposition

The first three principal components explained 81.86% of total variance in the standardised GCA matrix (Table 3; Fig.1), PC1 47.60%, PC2 21.55% and PC3 12.71%. All three exceeded the Kaiser threshold

(6.67%) and satisfied the 80% cumulative variance criterion (Greenacre *et al.*, 2022). A clear scree elbow between PC3 and PC4 confirmed the three-component solution.

Table 3 : Eigenvalue decomposition summary for PCA of the GCA effect matrix of 10 linseed parents across 15 traits. Bold rows indicate PCs retained per the Kaiser criterion.

PC	Eigenvalue	Variance (%)	Cumulative (%)	Kaiser criterion
1	7.14	47.60	47.60	Retained (>6.67%)
2	3.23	21.55	69.15	Retained (>6.67%)
3	1.91	12.71	81.86	Retained (>6.67%)
4	1.17	7.81	89.67	Excluded by scree criterion
5	0.62	4.12	93.79	Not retained

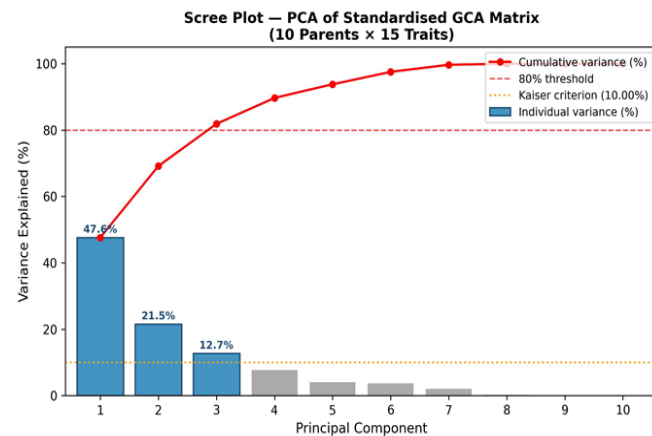


Fig. 1 : Scree plot of variance explained by each principal component from PCA of the standardised GCA matrix (10 parents × 15 traits). Blue bars = individual variance (%); red line with markers = cumulative variance (%); red dashed line marks the 80% threshold; orange dotted line marks the Kaiser criterion (6.67%). Three PCs cumulatively explain 81.86% of total variance.

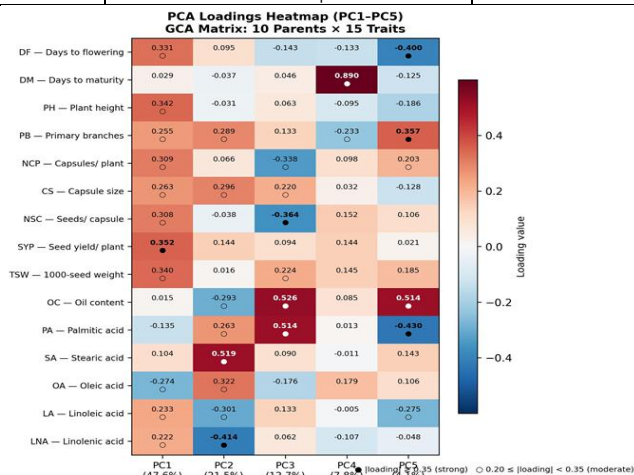
PC loadings and biological interpretation

PC1 loaded strongly and positively on the agronomic-yield complex: SYP (+0.352), PH (+0.342), TSW (+0.340), DF (+0.331), NCP (+0.309), NSC (+0.308) and CS (+0.263) (Table 4; Fig.2), indicating that alleles favouring yield, stature and seed number are co-inherited. Oleic (−0.274) and palmitic (−0.135) acids loaded negatively, revealing an inverse relationship between yield alleles and oleic-acid GCA.

PC2 represented an independent fatty-acid quality axis: stearic (+0.519), oleic (+0.322) and palmitic (+0.263) loaded positively, while linolenic (−0.414), linoleic (−0.301) and oil content (−0.293) loaded negatively. The opposing SA/OA vs LNA/LA loadings directly reflect Δ -15 desaturase pathway antagonism: flux towards linolenic acid via FAD3 reduces the oleic pool (Hall *et al.*, 2006). PC3 contrasted palmitic acid (+0.479) with NSC (−0.316) and DM (+0.360), revealing a third orthogonal dimension linking saturated-fat accumulation and phenology.

Table 4 : PCA loadings (eigenvectors) of 15 linseed traits on PC1, PC2 and PC3. Values $\geq |0.35|$ are in bold. Trait direction arrows: $\uparrow$ higher GCA desirable; $\downarrow$ lower GCA desirable.

Trait	Full name	PC1 (47.6%)	PC2 (21.55%)	PC3 (12.71%)	Trait cluster
DF	Days to 50% flowering $\uparrow$	0.331	0.095	0.039	Yield
DM	Days to maturity $\downarrow$	0.029	-0.037	0.360	Phenology
PH	Plant height $\uparrow$	0.342	-0.031	-0.299	Yield
PB	Primary branches/plant $\uparrow$	0.255	0.289	-0.161	Yield
NCP	Capsules/plant $\uparrow$	0.309	0.066	-0.115	Yield
CS	Capsule size $\uparrow$	0.263	0.296	0.128	Yield
NSC	Seeds/capsule $\uparrow$	0.308	-0.038	-0.316	Yield
SYP	Seed yield/plant $\uparrow$	0.352	0.144	-0.219	Yield
TSW	1000-seed weight $\uparrow$	0.340	0.016	-0.127	Yield
OC	Oil content $\uparrow$	0.015	-0.293	-0.163	Oil
PA	Palmitic acid $\downarrow$	-0.135	0.263	0.479	Saturated FA
SA	Stearic acid $\downarrow$	0.104	0.519	-0.184	Saturated FA
OA	Oleic acid $\uparrow$	-0.274	0.322	-0.264	MUFA
LA	Linoleic acid ω -6 $\uparrow$	0.233	-0.301	0.362	PUFA
LNA	Linolenic acid ω -3 $\uparrow$	0.222	-0.414	-0.100	PUFA

**Fig. 2 :** PCA loadings heatmap of 15 traits on the first three PCs. Red = strong positive loading; blue = strong negative loading; white = near zero. Filled circle (●) indicates $|loading| \geq 0.35$ (strong); open circle (○) indicates $0.20 \leq |loading| < 0.35$ (moderate). PC1 is dominated by the yield complex; PC2 contrasts ω -3 fatty acids against saturated/monounsaturated fatty acids; PC3 reveals a palmitic acid–phenology axis.**PC1 $\times$ PC2 biplot: parental classification**

The PC1 $\times$ PC2 biplot (Fig.3) captures 69.15% of variance. SHA-7 occupied the most extreme positive PC1 position (+4.93), over two-fold further from the

origin than any other parent (Table 5), confirming its exceptional yield GCA: SYP (+1.23**), NCP (+3.27**), TSW (+0.76**), CS (+0.71**), NSC (+0.28**) and PB (+0.58**) (Table 5). SHA-6 showed a moderate positive PC1 score (+2.04). SHA-4 (−3.14) and SHA-5 (−3.97) clustered tightly in the negative PC1 region with overlapping convex hulls, indicating functional genetic redundancy both carry early-flowering alleles but limited yield-improving resources. Similar redundancy diagnostics have been reported in field pea (Sharma *et al.*, 2023) and cotton (Ullah *et al.*, 2022).

On PC2, Varsa Alsí occupied the extreme negative position (−4.19), reflecting the highest linolenic-acid GCA in the tester group (+4.76**) combined with strongly negative oleic GCA (−4.51**), the pattern predicted by Δ -15 desaturation antagonism (Hall *et al.*, 2006). The large angular separation between SHA-7 (+4.93, +2.18) and Varsa Alsí (+3.63, −4.19) provides multivariate confirmation of their 49.50% standard heterosis for SYP (Panse and Sukhatme, 1978), consistent with the principle that heterosis increases with parental genetic divergence (Sprague and Tatum, 1942).

Table 5 : PC scores (biplot coordinates) for 10 linseed parents and concise breeding interpretation.

Parent	Type	PC1	PC2	PC3	Breeding interpretation
SHA-4	L	−3.14	−1.25	0.48	Early-flowering alleles; limited yield GCA
SHA-5	L	−3.97	−1.29	0.54	Genetic redundancy with SHA-4
SHA-6	L	2.04	0.18	−2.96	Favourable low-palmitic GCA
SHA-7	L	4.93	2.18	1.86	Best general combiner for yield
Azad Alsí-1	T	−0.40	0.90	−1.32	Moderate oleic GCA; versatile partner
Varsa Alsí	T	3.63	−4.19	0.60	Unique high- ω -3/high-oil parent
Kb Alsí-3	T	−1.51	2.07	−0.04	Elevated stearic GCA
Kb Alsí-4	T	−0.27	−0.26	0.01	Balanced moderate GCA
JLS-66	T	−1.27	1.16	1.80	Early maturity contribution
T-397	T	−0.04	0.50	−0.97	Balanced oleic/stearic GCA

L = line; T = tester.

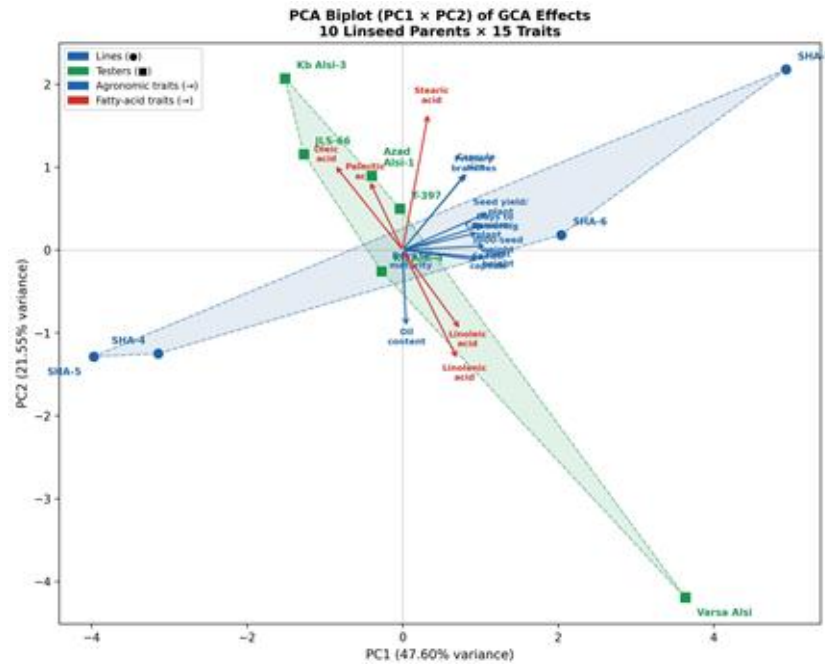


Fig. 3 : PCA biplot (PC1 × PC2) of GCA effects for 10 linseed parents across 15 traits. Lines = filled circles (●, blue); testers = filled squares (■, green). Trait-loading vectors: blue = agronomic traits; red = fatty-acid quality traits. Convex hulls delineate line and tester genetic spaces. PC1 (47.60%) = yield–agronomic axis; PC2 (21.55%) = fatty-acid quality axis.

PC1 × PC3 biplot

The PC1 × PC3 biplot (Fig. 4) captures 60.31% of variance and reveals structure invisible in the primary biplot. SHA-6 occupied the most extreme negative PC3 score (−2.96), driven by strongly negative

palmitic-acid GCA a nutritionally desirable profile. JLS-66 (+1.80) and SHA-7 (+1.86) showed the most positive PC3 scores, adding discrimination among parents clustering near the origin in the primary biplot.

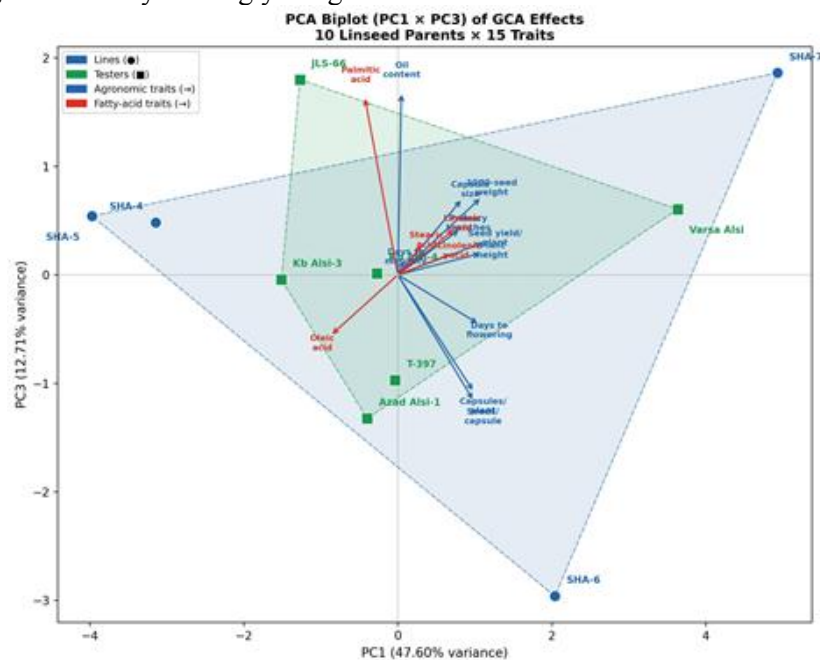


Fig. 4 : PCA biplot (PC1 × PC3) of GCA effects. PC3 (12.71%) primarily contrasts palmitic acid (+0.479) against seeds per capsule (−0.316) and oleic acid (−0.264), adding orthogonal discrimination. Combined with the PC1 × PC2 biplot, three-dimensional interpretation accounts for 81.86% of total GCA variation.

Hierarchically clustered heatmap

The double-clustered heatmap (Fig.5) identified four parent and four trait clusters. Parent clusters: (i) SHA-4 and SHA-5 showed strongly negative yield GCA with positive early-flowering GCA (ii) SHA-6 and SHA-7 reflects positive yield GCA but opposing fatty-acid profiles (iii) Varsa Alsí alone distinctly separated by its unique positive-LNA/negative-OA profile (iv) remaining testers forming a heterogeneous

moderate GCA group. Trait clusters involves, (i) yield–agronomic complex (SWP, TSW, NCP, PH) (ii) ω -3 PUFA cluster (LNA, LA) (iii) saturated-monounsaturated FA cluster (SA, OA) (iv) phenological cluster (DF, DM). These confirm that multi-trait GCA improvement requires targeting four independent genetic dimensions via appropriate parental combinations

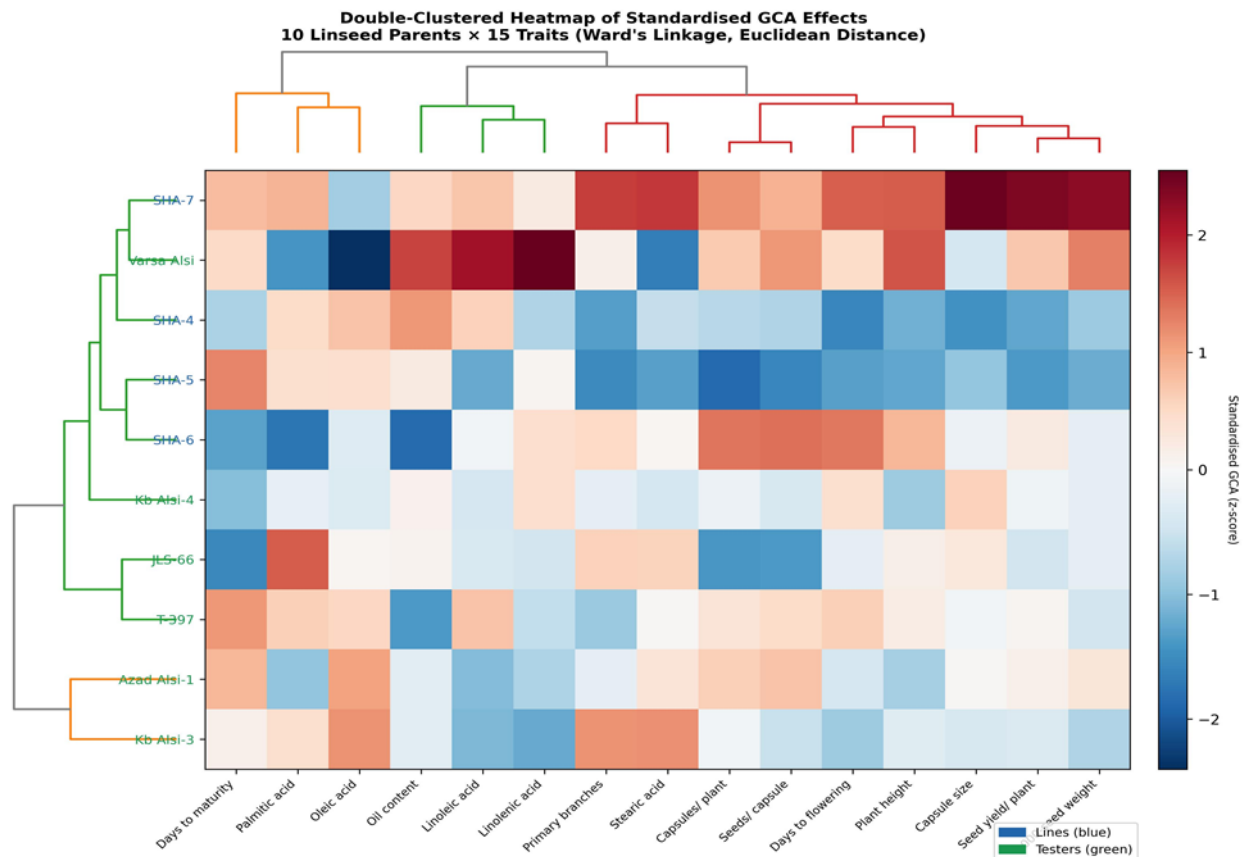


Fig. 5 : Double-clustered heatmap of standardised GCA effects (z-scores) for 10 parents (rows) $\times$ 15 traits (columns) using Ward's minimum-variance hierarchical clustering with Euclidean distance. Red = positive GCA; blue = negative GCA. Row-label colours: blue = lines; green = testers. Four parent and four trait clusters with distinct biological identities are visible.

Multi-trait Genotype–Ideotype Distance Index (MGIDI)

The MGIDI (Olivoto and Lúcio, 2020; Olivoto, 2022) was applied to the standardised GCA matrix to rank parents by proximity to a simultaneous multi-trait ideotype. A lower MGIDI score indicates a superior multi-trait combiner.

Two factors were retained by the Kaiser criterion: F1 (52.88% variance; yield–agronomic complex) and F2 (19.43%; fatty-acid quality axis), jointly explaining 72.31% of GCA variation. The ideotype was set at maximum observed scores: F1 = 4.910 (SHA-7) and

F2 = 3.472 (Varsa Alsí). The two-factor solution for MGIDI, compared with the three-component standalone PCA (81.86%), reflects sign-reorientation applied in the MGIDI algorithm to ensure all factor scores are directionally aligned with the ideotype (Olivoto, 2022).

Varsa Alsí ranked first (MGIDI = 1.295), owing to its maximum F2 score (3.472) and strong F1 (3.615). SHA-7 ranked second (MGIDI = 4.633) having its high F1 is offset by a negative F2, reflecting limited ω -3 GCA. SHA-5 ranked last (MGIDI = 8.702), confirming its redundancy with SHA-4 (Table

6; Figs.6a–6d). Factor contributions demonstrate that Varsa Alsí's gap from the ideotype is entirely on the F1 (yield) axis, while SHA-7's deficit is entirely on F2 (ω -3 quality), which confirming that the SHA-7 $\times$ Varsa Alsí cross simultaneously captures both axes.

Table 6 : MGIDI scores, factor scores and parent rankings for 10 linseed parents. F1 and F2 = factor scores from PCA of the sign-oriented, standardised GCA matrix. Lower MGIDI = closer to multi-trait ideotype.

Parent	Type	F1 (52.88%)	F2 (19.43%)	MGIDI	Rank
Varsa Alsí	T	3.615	3.472	1.295	1
SHA-7	L	4.910	−1.160	4.633	2
SHA-6	L	1.348	−1.394	6.030	3
Kb Alsí-4	T	−0.131	−0.032	6.140	4
T-397	T	−0.185	−0.702	6.586	5
JLS-66	T	−0.695	−0.248	6.727	6
Azad Alsí-1	T	−0.893	−1.200	7.450	7
SHA-4	L	−2.941	1.595	8.073	8
Kb Alsí-3	T	−1.556	−1.465	8.136	9
SHA-5	L	−3.472	1.136	8.702	10
<i>Ideotype</i>	—	4.910	3.472	0.000	—

L = line; T = tester. Ideotype: F1 = 4.910, F2 = 3.472. Lower MGIDI = closer to multi-trait ideotype.

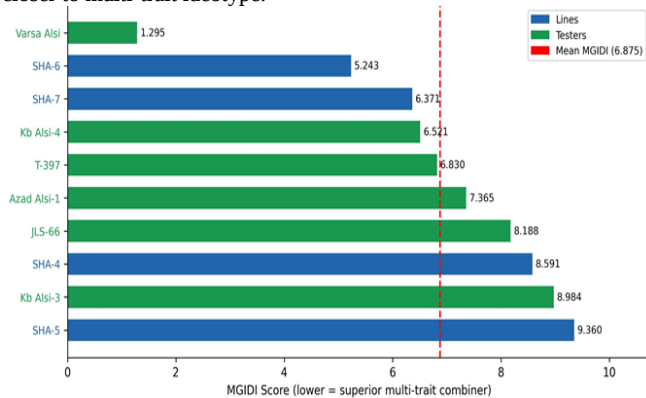


Fig.6a : MGIDI bar chart ranking 10 linseed parents by multi-trait genotype-ideotype distance. Lower MGIDI = superior multi-trait combining ability. Blue bars = lines; green bars = testers. Red dashed line = mean MGIDI. Varsa Alsí (MGIDI = 1.295) ranks first, followed by SHA-7 (4.633).

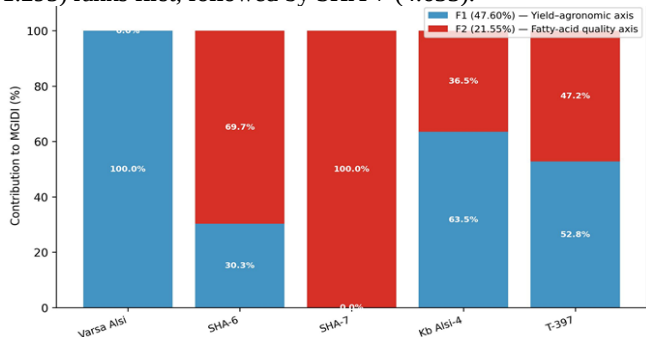


Fig.6b : Stacked bar chart of factor contributions (%) to MGIDI score for the five best-ranked parents. F1 (52.88%) = yield-agronomic axis; F2 (19.43%) = fatty-acid quality axis. Varsa Alsí's MGIDI distance is almost entirely F1-driven; SHA-7's deficit is entirely F2-driven.

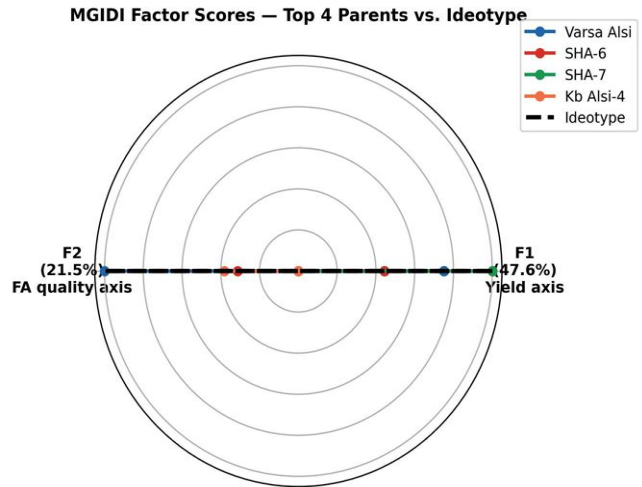


Fig.6c : Spider (radar) chart of normalised MGIDI factor scores for the top four parents and the ideotype (dashed black line). F1 = yield axis; F2 = fatty-acid quality axis. No single parent matches the ideotype on both axes; the complementarity between Varsa Alsí (dominant F2) and SHA-7 (dominant F1) is graphically evident.

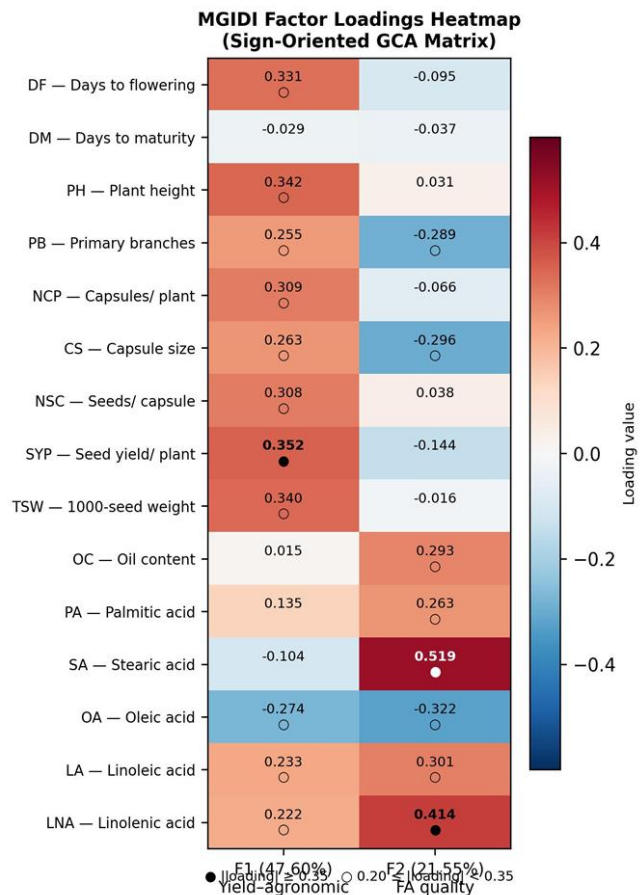


Fig.6d : MGIDI factor loadings heatmap for two retained factors (F1: 52.88%; F2: 19.43%) across 15 traits using the sign-oriented GCA matrix. Red = strong positive loading; blue = strong negative loading. ● indicates $|\text{loading}| \geq 0.35$ (strong); ○ indicates $0.20 \leq |\text{loading}| < 0.35$ (moderate). F1 encompasses the yield-agronomic complex; F2 represents the fatty-acid quality axis.

Discussion

The three-component PCA solution (81.86% variance) provides a sound multivariate framework for interpreting combining-ability structure in linseed. The three axes—yield-agronomic (PC1), fatty-acid quality (PC2) and palmitic-phenology (PC3) map directly onto three independent breeding objectives, confirming PCA's utility for disentangling multi-dimensional GCA in oilseeds (Sharma *et al.*, 2023; Ullah *et al.*, 2022).

PC1 dominance (47.60%) demonstrates that additive variation for yield components is the primary source of combining-ability variation in this population, with SHA-7 driving most of this variation. The orthogonality of PC2 to PC1 is the key practical finding: yield improvement and ω -3 improvement are independent genetic objectives. A parent optimised for PC1 (SHA-7) is not optimal for PC2, and vice versa (Varsa Alsi). This supports a targeted strategy where SHA-7 $\times$ Varsa Alsi simultaneously captures positive PC1 (yield, from SHA-7) and negative PC2 (high LNA and oil content, from Varsa Alsi) — a complementarity no other pairing in this population achieves.

The tight biplot clustering of SHA-4 and SHA-5 reveals practical parental redundancy. Deploying both provides negligible additional genetic diversity; future panels should replace one with a genotype occupying a different biplot position, particularly one with positive PC2 scores. Similar redundancy diagnostics have been reported in field pea (Sharma *et al.*, 2023) and cotton (Ullah *et al.*, 2022).

Ward's clustering independently confirmed the biplot classifications, Varsa Alsi's singleton status and the SHA-4/SHA-5 grouping. The four trait clusters provide a genetic pathway map for multi-trait improvement: traits within a cluster can be improved through a single parent, while traits across clusters require pairing parents from different biplot quadrants.

This study is among the first to apply PCA directly to a GCA matrix, rather than phenotypic means, in linseed, yielding more precise parental classification. The findings support a three-dimensional parent-selection framework: maximise PC1 divergence for yield heterosis (SHA-7 as line); exploit PC2 complementarity for ω -3 improvement (Varsa Alsi as tester); and consider PC3 for palmitic-acid reduction (SHA-6). The framework is recommended for routine use in multi-trait oilseed breeding programmes.

MGIDI formalised the PCA complementarity finding into a single index: Varsa Alsi ranked first (MGIDI = 1.295) by virtue of its maximum F2 score,

while SHA-7's F1 dominance places it second. No single parent reaches the ideotype on both axes, reinforcing SHA-7 $\times$ Varsa Alsi as the optimal cross for simultaneous yield and ω -3 improvement (Olivoto and Lúcio, 2020).

Acknowledgement

The authors acknowledge the Department of Genetics and Plant Breeding, Chimanbhai Patel College of Agriculture, SDAU, Sardarkrushinagar for providing the all basic facilities for completion of the research work. We would like to thank to ICAR-Indian Institute of Pulse Research (ICAR-IIPR), Kanpur; Directorate of Oilseeds Research (DOR), SHUATS, Prayagraj and Regional Agricultural Research Station, Jawaharlal Nehru Krishi Vishwavidyalaya (JNKVV), Sagar, Madhya Pradesh for sharing genotype with us to conducting the present research work.

References

- Bhateria, S., Sood, S. P. and Pathania, A. (2006). Genetic analysis of quantitative traits across environments in linseed (*Linum usitatissimum* L.). *Euphytica*, **150**, 185–194.
- Chikh-Rouhou, H., Abdedayem, W., Solmaz, I., Sari, N. and Garcés-Claver, A. (2024). Diversity analysis and assessment of morphological variation for main agronomic traits in *Cucurbita pepo* germplasm. *Horticulturae*, **10**, 724.
- Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, **58**, 453–467.
- Goyal, A., Sharma, V., Upadhyay, N., Gill, S. and Sihag, M. (2014). Flax and flaxseed oil: An ancient medicine and modern functional food. *Journal of Food Science and Technology*, **51**, 1633–1653.
- Greenacre, M., Groenen, P. J. F., Hastie, T., Iodice D'Enza, A., Markos, A. and Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, **2**, 100.
- Hall, C., Tulbek, M. C. and Xu, Y. (2006). Flaxseed. *Advances in Food and Nutrition Research*, **51**, 1–97.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, **20**, 141–151.
- Kempthorne, O. (1957). *An introduction to genetic statistics*. John Wiley & Sons.
- Olivoto, T. (2022). Multivariate genotype-ideotype distance index for plant breeding. *Agronomy Journal*, **114**, 2556–2566.
- Olivoto, T. and Lúcio, A. D. (2020). metan: An R package for multi-environment trial analysis. *Methods in Ecology and Evolution*, **11**, 783–789.
- Panase, V. G. and Sukhatme, P. V. (1978). *Statistical methods for agricultural workers* (3rd ed.). Indian Council of Agricultural Research.
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **2**, 559–572.

- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830.
- Sharma, A., Yadav, R., Sheoran, R., Kaushik, D., Mohanta, T. K., Sharma, K., Yadav, A., Dhanda, P. S. and Kaushik, P. (2023). Estimation of heterosis and the combining ability effect for yield and its attributes in field pea (*Pisum sativum* L.) using PCA and GGE biplots. *Horticulturae*, **9**, 256.
- Sprague, G. F. and Tatum, L. A. (1942). General vs. specific combining ability in single crosses of corn. *Journal of the American Society of Agronomy*, **34**, 923–932.
- Ullah, A., Shakeel, A., Ahmed, H. G. M.-D., Naeem, M., Ali, M., Shah, A. N., Wang, L., Jaremko, M., Abdelsalam, N. R., Ghareeb, R. Y. and Hasan, M. E. (2022). Genetic basis and principal component analysis in cotton (*Gossypium hirsutum* L.) grown under water deficit condition. *Frontiers in Plant Science*, **13**, 981369.
- Virtanen, P., Gommers, R., Oliphant, T. E. and SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, **17**, 261–272.
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, **58**, 236–244.