Journal of Advances in Biology & Biotechnology

Volume 27, Issue 8, Page 1217-1227, 2024; Article no.JABB.121281
ISSN: 2394-1081

Variability, Correlation Patterns and Principal Component Analysis (PCA) for Seed Yield and Contributing Traits in Castor (*Ricinus communis* L.)

Kadam Abhishek Deepak [a], T. Manjunatha [b*], V. Hemalatha [a] and D. Srinivasa Chary [c]

[a] Department of Genetics and Plant Breeding, Professor Jayashankar Telangana State Agriculture University (PJTSAU), Rajendranagar, Hyderabad 500030, India.
[b] ICAR- Indian Institute of Oilseeds Research (ICAR-IIOR), Rajendranagar, Hyderabad 500030, India.
[c] Department of Statistics and Mathematics, Professor Jayashankar Telangana State Agriculture University (PJTSAU), Rajendranagar, Hyderabad 500030, India.

Authors' contributions

This work was carried out in collaboration among all authors. Author KAD wrote original draft, investigated the work, did software analysis, wrote, reviewed and edited the manuscript. Authors VH and DSC wrote and reviewed the manuscript. Authors KAD and DSC did data curation and formal analysis. Author TM conceptualized the study, did data curation, funding acquisition, investigated and supervised the study, performed methodology, administrated the project, searched for resources, did data validation, visualization, wrote, reviewed and edited the manuscript. All authors read and approved the final manuscript.

Article Information

DOI: https://doi.org/10.9734/jabb/2024/v27i81246

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/121281

Original Research Article

Received: 02/06/2024
Accepted: 07/08/2024
Published: 09/08/2024

**Corresponding author: E-mail: arsmanju@gmail.com;*

Cite as: *Deepak, Kadam Abhishek, T. Manjunatha, V. Hemalatha, and D. Srinivasa Chary. 2024. "Variability, Correlation Patterns and Principal Component Analysis (PCA) for Seed Yield and Contributing Traits in Castor (Ricinus Communis L.)". Journal of Advances in Biology & Biotechnology 27 (8):1217-27. https://doi.org/10.9734/jabb/2024/v27i81246.*

ABSTRACT

Castor (*Ricinus communis* L.) is a vital crop for industrial applications in more than 250 products including lubricants, paints, cosmetics, pharmaceuticals etc. This study is an attempt to understand the genetic diversity in 15 male (monoecious) and 15 female (pistillate) advanced breeding lines of castor. 11 quantitative traits were subjected to analysis of variance, correlation analysis, principal component analysis (PCA) and K-means clustering. Significant genetic variability and trait correlations were noticed, revealing opportunities for targeted improvement in castor. Clustering identified six distinct genetic groups, facilitating the identification of diverse parental lines. Principal component analysis elucidated key contributors of variation, enabling informed breeding decisions. This comprehensive study provides a foundation for further improvement in seed yield, oil content and environmental resilience in castor.

Keywords: Castor, pistillate; monoecious; K-means clustering; principal component analysis.

1. INTRODUCTION

Castor is a commercial, non-edible and industrial oilseed crop and member of the Euphorbiaceae family with a diploid chromosome count of 2n=20. It has extensive genetic variability, with Ethiopia considered as the most probable site of origin and India as an important centre of diversity [1,2]. Castor's reproductive mechanism is distinct, with cross-pollination facilitated by wind dispersal and monoecious nature [3]. Its seeds contain a unique oil composition, primarily ricinoleic acid, used in various industrial products and as a promising candidate for biodiesel production [4-6].

Castor is well-suited for both irrigated and dry land farming in India during the kharif / late kharif and rabi season, requiring minimal irrigation in the post-monsoon period [7]. Castor serves as a contingency crop, providing flexibility in agricultural planning to mitigate disruptions caused by monsoon irregularities. The major castor growing states in India include mainly Gujarat and Rajasthan followed by Andhra Pradesh, Tamil Nadu, Karnataka and to some extant in Telangana and maharashtra. During the period 2022-23, India produced 19.80 lakh tonnes of castor in an area of 10.19 lakh hectares [8]. Over the past forty years, collaborative research led by the Indian Council of Agricultural Research - Indian Institute of Oilseeds Research (ICAR-IIOR) and several State Agricultural Universities (SAUs) has resulted in the development of 15 high-yielding castor varieties and 23 castor hybrids. These efforts have significantly advanced castor breeding, boosting the productivity and sustainability in the industry.

The efficacy of varietal or hybrid development programmes fundamentally hinges on the magnitude and nature of genetic variability present in breeding lines.

The Indian germplasm repository boasts an impressive array of morphological diversity, mirroring the range of variants found in other centres of origin as well as tropical and subtropical regions. Twenty-four morphological descriptors have been crafted to characterize castor germplasm accessions. Notably, the native Indian collection predominantly features medium-tall plants with woody stems with red and green colour, accompanied by a varying number of nodes, divergent branching and a waxy stem coating. Furthermore, the leaves exhibit a semi-cup shape, while the racemes are characterized by their medium length, loose, conical shape, medium-sized, non-dehiscent, green and spiny capsules Anjani et al. [2].

The genetic diversity studies in castor breeding lines will aid plant breeders in selecting suitable parental lines for breeding programmes. Breeders can develop effective strategies for selection of genotypes by combining genetic variability, genotypic correlation analysis and principal component analysis. There is a need to understand correlation relationships between traits, highlighting potential synergies and trade-offs. Principal Component Analysis (PCA) is required to determine the primary factors contributing to the overall variation. Additionally, K- means clustering analysis can be employed to group the advanced breeding into distinct clusters. The present investigation aims to estimate genetic variability and character associations among yield-related traits in monoecious and pistillate breeding lines of castor.

Table 1. Means of 11 agro-morphological traits in 30 breeding lines

Entry	DFF	DM	PH	NN	TLP	ELP	TS	TC	OC	HSW	TY
K22-1	53.3	109.3	121.4	15.7	65.1	63.0	4.1	71.7	42.3	26.3	148.7
K22-5	53.7	105.0	86.6	15.3	58.4	54.2	1.9	78.1	38.3	25.4	102.1
R22-14	54.7	100.0	65.9	14.0	38.1	34.8	3.3	52.1	44.1	22.1	81.4
K22-19	53.7	97.7	86.6	15.1	44.9	32.3	3.7	49.7	41.1	22.7	105.2
K22-26	59.7	103.0	78.5	15.6	44.6	41.8	2.6	71.5	45.6	28.6	119.0
K22-35	54.3	101.3	72.9	14.5	55.7	45.4	4.0	54.9	41.7	21.4	114.9
K22-38	54.7	100.7	56.9	13.7	35.7	31.9	4.7	47.9	47.9	25.3	115.2
K22-39	55.3	90.7	60.9	13.1	63.3	53.4	4.3	67.9	38.6	18.3	102.2
K22-43	54.0	102.3	62.2	13.8	42.3	35.2	5.5	47.1	43.5	22.1	125.1
K22-45	55.7	103.0	60.2	13.9	35.9	31.0	5.7	51.2	46.1	18.3	123.0
K22-46	55.0	99.3	79.0	15.3	43.0	32.6	4.5	53.7	37.4	18.3	122.8
K22-47	56.3	106.3	102.3	15.5	40.6	33.8	4.6	30.2	43.0	21.8	77.3
K22-48	59.0	97.7	102.8	16.6	42.5	38.3	3.9	46.7	43.5	22.9	98.0
K22-49	56.0	101.0	80.6	13.5	41.5	37.7	5.3	38.7	39.0	23.1	103.4
ICS-164	49.3	100.3	104.6	17.7	47.0	42.9	3.7	81.7	41.1	22.1	178.7
IPC-41	38.0	103.0	63.6	11.8	50.5	45.4	3.9	40.5	47.7	17.9	73.7
DPC-25	59.7	104.7	60.5	19.0	64.7	60.1	2.2	67.5	39.2	21.6	62.2
IPC-46	48.0	110.3	86.5	18.3	54.6	40.5	3.2	50.4	48.1	26.5	79
DPC-22	62.3	108.7	70.3	16.6	40.0	39.9	2.0	70.0	38.3	21.3	48.2
JP-96	48.7	111.0	75.5	15.1	46.2	42.9	2.3	43.2	47.4	29.1	58.7
IPC-47	47.7	110.7	53.2	17.1	51.7	43.9	1.9	43.7	38.5	25.6	53.2
IPC-44	39.7	93.3	70.7	12.7	33.4	30.9	2.7	51.3	46.2	28.0	74.8
JP-86	62.3	108.0	50.7	16.8	36.9	35.4	2.7	40.7	46.4	28.3	47.5
M-619	62.7	105.7	56.2	21.1	49.3	49.2	2.7	46.1	35.6	17.4	39.5
IPC-48	39.3	94.7	69.0	8.2	50.6	48.2	2.0	79.9	44.8	16.8	70.6
IPC-31	61.7	108.3	43.9	17.4	31.4	29.8	2.6	40.0	38.3	22.1	46.7
IPC-49	45.0	98.0	45.7	13.9	40.7	39.1	1.9	37.2	38.3	34.8	59.1
IPC-53	42.3	98.0	61.4	12.5	36.3	34.7	2.9	43.1	44.3	20.1	58.9
IPC-52	38.0	93.0	49.3	12.0	36.9	33.5	2.9	55.4	37.1	18.7	69.3
IPC-51	38.7	100.7	54.9	13.1	40.6	37.5	4.1	53.6	46.2	25.7	113

DFF- Days to 50% flowering, DM - Days to maturity of primary spike, PH - Plant height (cm), NN —Number of nodes up to primary spike, TLP – Total length of primary spike (cm), ELP – Effective Length of primary spike (cm), NS - Number of spikes, NC-Number of capsules in Primary spike, HSW- Hundred seed weight (g.) , OC- Oil content (%), TY-Total seed yield (g.)

2. MATERIALS AND METHODS

Thirty advanced and stabilized breeding lines of castor, comprising 15 monoecious and 15 pistillate lines, developed by pedigree method at the ICAR-Indian Institute of Oilseeds Research, Hyderabad and SAUs were utilized for genetic diversity study. Prior to evaluation, all the breeding lines underwent six (monecious) to eight (pistillate) generations of self-pollination at research farm of ICAR-IIOR, Hyderabad under controlled pollination conditions.

An experiment was conducted at the ICAR-Indian Institute of Oilseeds Research at Narkhoda research farm near Hyderabad, India (17.366°N and 78.478°E). The crop was grown under irrigated conditions in red sandy soil and sowing was carried out on 24th July 2023. The study was conducted in randomized block design (RBD) with three replications and two rows per genotype with 10 plants per row in each replication with the spacing of 90 x 60 cm. Eleven characters *viz.,* days to 50% flowering, days to maturity , plant height (cm), number of nodes to primary spike, total length of primary spike (cm), effective length of primary spike (cm), number of spikes/plant, number of capsules on primary spike,100-seed weight (g), oil content (%) and total seed yield per plant were recorded in each replication from five randomly selected plants per breeding line (Table 1). All statistical analyses were performed using INDOSTAT statistical software, INDOSTAT Services, India [8].

3. RESULTS AND DISCUSSION

3.1 Analysis of Variation

The analysis of variance showed significant differences among the genotypes for all characters. The estimates of genetic variability, heritability and genetic advance are presented in Table 2. The phenotypic coefficient of variation (PCV) values, which quantify the total variability present in the population, including both genetic and environmental components, were calculated. The number of capsules on primary spike (29.19%), plant height up to primary spike (27.42%), total seed yield per plant (44.18%) and total number of spikes per plant (36.71%) showed high phenotypic coefficient of variation. Whereas, days to 50% flowering (15.28), number of nodes (17.54) and oil content (10.47) showed moderate phenotypic coefficient of variation. These PCV values indicate that total seed yield per plant, total and effective length of primary spike, number of capsules on primary spike and total number of spikes per plant have high variability, while days to 50% flowering, number of nodes and oil content have moderate phenotypic variability.

In contrast, the genotypic coefficient of variation (GCV) values, which quantify the genetic component of variability. High GCV values were recorded for, total seed yield per plant (34.04%), number of capsules on primary spike (31.13%), plant height up to primary spike (26.32%) and total number of spikes per plant (23.79%). However, days to 50% flowering (14.89), number

Table 2. Mean, range, variability parameters, heritability and genetic advance

Characters	Mean	Range		PCV (%)	GCV (%)	h^2	GA %
		Min.	max				
DFF	51.96	38	62.67	15.28	14.89	95.0	29.86
DM	102.19	90.67	111	5.56	5.29	90.7	10.39
PH	71.09	43.96	121.40	27.45	26.32	92.0	52.00
NN	14.97	8.20	21.13	17.54	16.50	88.5	31.99
TLP	45.41	31.43	65.13	21.54	19.84	84.80	37.62
ELP	40.78	29.83	63.00	23.09	20.92	82.10	39.05
NS	53.52	30.20	81.73	29.19	31.13	66.40	39.95
NC	3.39	1.87	5.67	36.71	23.79	71.90	54.39
HSW	23.08	16.80	34.80	22.17	8.19	50.30	23.01
OC	42.31	35.57	48.13	10.47	15.74	61.10	13.18
TY	89.03	39.47	178.70	44.18	34.06	59.5	54.11

DFF- Days to 50% flowering, DM - Days to maturity of primary spike, PH - Plant height (cm), NN – Number of nodes up to primary spike, TLP – Total length of primary spike (cm), ELP – Effective Length of primary spike (cm), NS - Number of spikes, NC-Number of capsules in Primary spike, HSW- Hundred seed weight (g.) , OC- Oil content (%), TY-Total seed yield (g.). PCV- Phenotypic coefficient of variation, GCV- Genotypic coefficient of variation, h^2- Heritability at broad sense, GAM- Genetic advance as apercent mean

of nodes (16.50), total length (19.84) and effective length (20.92) of primary spike and oil content (15.74) showed moderate range of genotypic variability, whereas, days to 50% maturity and hundred seed weight recorded low genotypic variability. These findings are consistent with previous studies [9-12]. Genotypic coefficient of variation (GCV) recorded lower than phenotypic coefficient of variation (PCV) for all characters, indicating that environmental factors were minimum and phenotypic selection can be effectively used for improvement.

High heritability estimates were recorded for days to 50% flowering (95.0%), days to 50% maturity (90.70%), number of nodes (88.50%), total length of primary spikes (84.80%) and effective length of primary spikes (82.10%). High the heritability value indicated the presence of additive gene action and further improvement in these traits could be effective through direct selection. These results are in accordance with the findings of Dapke *et al*. [13], Movalia *et al*. [14] and Rukhsar *et al.* [10].

In contrast, moderate heritability estimates were recorded for hundred seed weight (50.30%) and total seed yield per plant (59.50%), indicating the influence of both genetic and environmental factors. There are contradictory reports for the heritability of total seed yield per plant Patel et al.[15], Ranjitha et al.[16],Rukhsar et al.[10] reported high heritability (>60%). Whereas, Alhaji *et al.* [11] and Mullualem et al. [17] reported low (30%) to moderate (30-60%) heritability in broad sense.

3.2 Correlation Analysis

The correlation analysis revealed various relationships between traits (Table 3, Fig. 1). Days to 50% flowering showed a significant positive correlation with number of nodes to primary spike (0.6820) and days to 50% maturity (0.4150). Days to 50% maturity had a significant positive correlation with number of nodes to primary spike (0.6328) and 100-seed weight while showing a negative non-significant correlation with total number of spike number of capsules on primary spike, and total yield.

Plant height up to primary spike showed a significant positive correlation with total seed yield per plant (0.6053), followed by total length of primary spike, number of capsules on primary spike, effective length of primary spike and number of spikes. Total length of primary spike had a high significant positive correlation with effective length of primary spike (0.9355) and number of capsules on primary spike (0.5348). Effective length of primary spike showed a high significant positive correlation with number of capsules on primary spike (0.6267) and a significant negative correlation with total number of spikes per plant (-0.3417). Additionally, number of spikes per plant was significantly correlated with total seed yield per plant (0.6160) and number of capsules on primary spike had a significant positive correlation with total seed yield per plant (0.4047) similar reports were given by Tewari *et al.* [18], Manjunatha *et al.* [19], Goodzari *et al.* [20], Patel et al. [21]. However, hundred seed weight exhibited a significant positive correlation with days to 50% flowering. Total seed yield per plant was significantly correlated with number of capsules per spike (0.6160), plant height to primary spike (0.6053) and number of spikes per plant (0.4047), with a significant negative association with days to 50% maturity.

3.3 Diversity Analysis

The study employed K-means clustering to categorize genotypes into distinct groups, as presented in Table 4. This method identified patterns and relationships among the characteristics, grouped them into six clusters with high degrees of similarity. The cluster size ranged from 2 (cluster V) to 11 (cluster II) breeding lines, indicating varying levels of diversity within each group. The cluster mean differentiated members of each cluster from other clusters, highlighting the distinct characteristics of each group (Table 5).

Cluster II and III stood out for their notable characteristics, comprising high-yielding and tall lines with late maturity and produced higher number of spikes per plant. In contrast, Cluster IV was characterized by genotypes with short height, long primary spikes with higher number of capsules on primary spike. Clusters V and VI were the pistillate lines with early flowering and maturity, short stature, fewer nodes, high oil content and average seed yield. In contrast, Cluster I comprised late-maturing, short pistillate lines with a more number of nodes but low total seed yield per plant.

DFF	DM	PH	NN	TLP	ELP	NS	NC	OC	HSW	TY	
1	0.415		0.682					-0.322			DFF
	1		0.632						0.309	-0.240	DM
				0.356	0.268	0.248	0.318			0.605	PH
			1	0.207				-0.375			NN
				1	0.935		0.5348				TLP
					1	-0.341	0.626				ELP
						1			-0.406	0.616	NS
							1			0.404	NC
								1			OC
									1		HSW
										1	TY

Fig. 1. Shaded correlation matrix of different traits under study

Table 3. Genotypic correlation coefficients of yield attributing traits in the advanced breeding lines of castor studied

	DFF	DM	PH	NN	TLP	ELP	NS	NC	OC	HSW	TY
DFF	1	0.4150* *	0.1351	0.6820 **	0.0636	0.0491	0.1343	0.0056	-0.3226**	-0.0323	0.0111
DM		1	0.1315	0.6328 **	0.1719	0.1669	-0.2066	-0.1715	0.0894	0.3094**	-0.2406*
PH			1	0.1601	0.3565**	0.2689**	0.2483*	0.3181**	0.1000	0.0329	0.6053 **
NN				1	0.2079*	0.1575	-0.2008	-0.0672	-0.3753 *	0.1906	-0.1366
TLP					1	0.9355 **	-0.1553	0.5348 **	-0.2547**	-0.1188	0.1773
ELP						1	-0.3417**	0.6267 **	-0.2829**	-0.1053	0.0352
NS							1	-0.2970**	0.2277*	-0.4062 **	0.6160**
NC								1	-0.2048*	-0.1924	0.4047 **
OC									1	0.1743	0.0912
HSW										1	-0.1406
TY											1

*** Significant at 0.01: *Significant at 0.05*

DFF- Days to 50% flowering, DM - Days to maturity of primary spike, PH - Plant height (cm), NN —Number of nodes up to primary spike, TLP – Total length of primary spike (cm), ELP – Effective Length of primary spike (cm), NS - Number of spikes, NC-Number of capsules in Primary spike, HSW- Hundred seed weight (g.) , OC- Oil content (%), TY-Total seed yield (g.).

3.4 Principal Component Analysis

The determination of the optimal number of principal components (PCs) that explains the maximum variability is crucial in Principal Component Analysis (PCA). According to Rencher (2002), a threshold of 70% variance explained by PCs is recommended. In our analysis, out of 11 the first four PCs accounted for approximately 75.80% of the total variability in the data set, with PC1 explaining 26.50%, PC2 explaining 20.20%, PC3 explaining 15.90%, and PC4 explaining 13.20% (Table 6). This exceeds the recommended threshold, indicating that the first four PCs are sufficient to capture the majority of the variation in the data.

Alternatively, the Kaiser criterion (1958) suggests retaining PCs with eigenvalues greater than 1 ($\lambda i > 1$), which also supports the retention of the first four PCs. The scree plot further confirms this decision, showing a clear break in the slope after the fourth component (Fig. 2). Therefore, the first four PCs were retained for further analysis. Simplifying the data while capturing a substantial amount of the variation.

Table 4. Clustering of 30 castor accessions using K-clustering method based on agro-morphological data set

Group K	Number of members	Cluster member
I	5	DPC-25, DPC-22, JP-86, M-619, IPC-31
II	11	K22-1, R22-14, K22-26, K22-35, K22-38, K22-43, K22-45, K22-46, K22-49, IPC-46, JP-96
III	4	K22-19, K22-47, K22-48, ICS-164
IV	3	K22-5, K22-39, IPC-47
V	2	IPC-44, IPC-53
VI	5	IPC-41, IPC-48, IPC-49, IPC-52, IPC-51

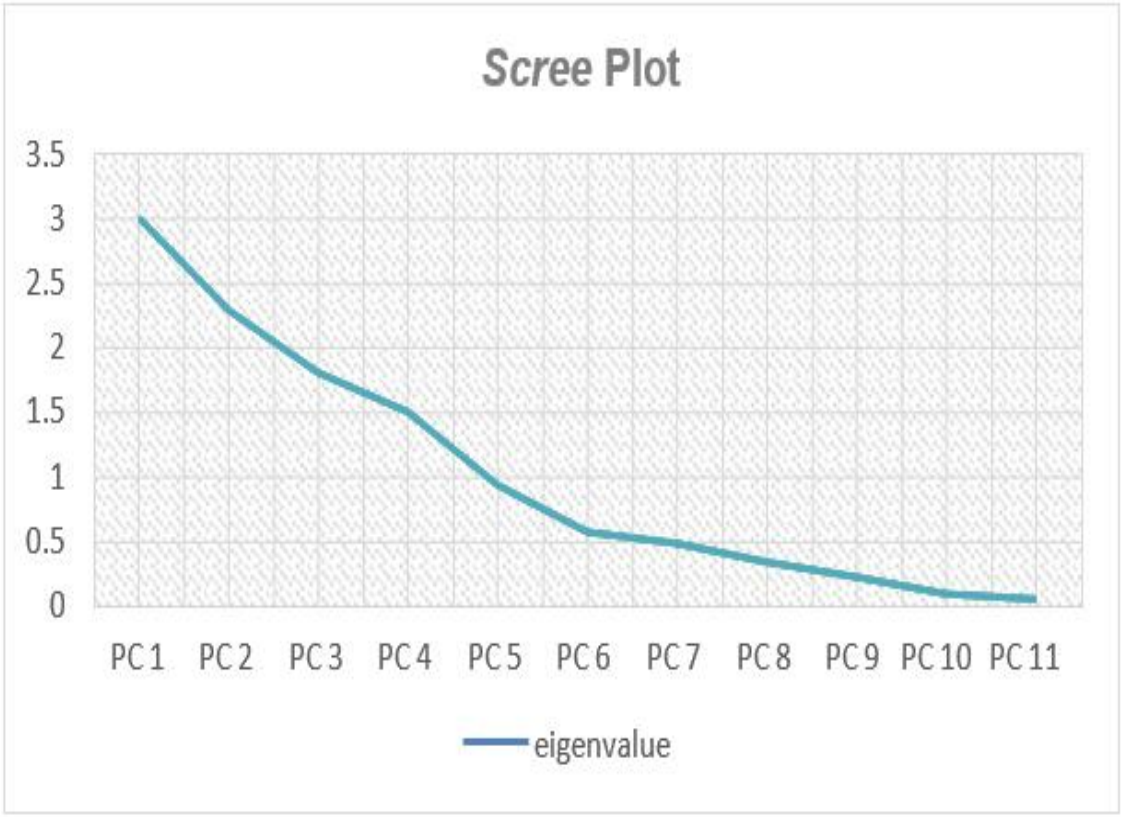

Fig. 2. *Scree* plot of eigenvalues (Variances of principal components)

Table 5. K-Cluster means For 11 agro-morphological character studied

	DFF	DM	PH	NN	TLP	ELP	NS	NC	OC	HSW	TY
1 Cluster	61.73	107.07	56.33	18.19	44.46	42.88	2.44	52.87	39.55	22.13	48.81
2 Cluster	54.00	103.76	76.33	14.86	45.70	39.71	4.10	52.96	43.91	23.72	108.33
3 Cluster	54.58	100.50	99.06	16.22	43.73	36.82	3.97	52.08	42.18	22.39	114.78
4 Cluster	52.22	102.11	66.90	15.18	57.82	50.51	2.71	63.24	38.46	23.11	85.83
5 Cluster	41.00	95.67	66.02	12.57	34.88	32.83	2.80	47.20	45.25	24.03	66.87
6 Cluster	39.80	97.87	56.52	11.79	43.85	41.55	2.95	53.31	42.83	22.79	77.13

Table 6. Principal component analysis for 11 agro-morphological traits

Variables	PC1	PC2	PC3	PC4
DFF	0.222	-0.391	-0.231	-0.374
DM	0.21	-0.46	-0.19	0.218
PH	0.297	0.191	-0.445	0.013
NN	0.300	-0.49	-0.148	-0.108
TLP	0.498	0.161	0.115	0.048
ELP	0.500	0.147	0.217	0.116
NS	-0.142	0.207	-0.489	-0.424
NC	0.375	0.336	0.126	0.018
OC	-0.208	0.136	-0.306	0.435
HSW	0.000	-0.185	-0.216	0.643
TY	0.162	0.320	-0.485	0.047
Eigenvalue	3.01	2.296	1.813	1.507
Variance %	26.50	20.20	15.90	13.20
Cumulative variance %	26.50	46.60	62.60	75.80

DFF- Days to 50% flowering, DM - Days to maturity of primary spike, PH - Plant height (cm), NN – Number of nodes up to primary spike, TLP – Total length of primary spike (cm), ELP – Effective Length of primary spike (cm), NS - Number of spikes, NC-Number of capsules in Primary spike, HSW- Hundred seed weight (g.) , OC- Oil content (%), TY-Total seed yield (g.)

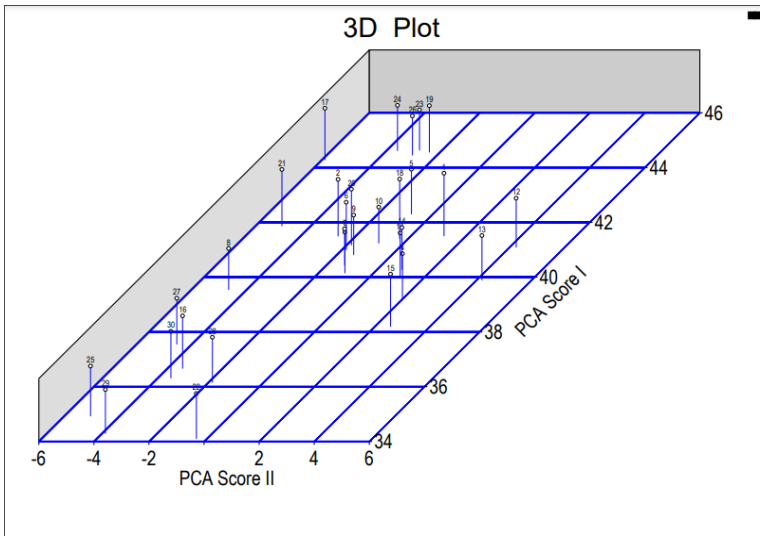

Fig. 3. PCA analysis for the castor breeding lines

The Principal Component Analysis (PCA) identified four significant components that captured the underlying variation pattern in the data. The first component primarily defined by the effective length of primary spike as the main contributor followed by total length of primary spike, number of capsules on primary spike, number of nodes and plant height to primary spike, which were highly correlated and represented the main source of variation among the breeding lines. The second component was characterized by number of capsules on primary spike, total seed yield, number of spikes per plant and plant height. The third component highlighted effective length of primary spike, number of capsules on primary spike and number of spikes per plant as the main contributors of variation, while the fourth component emphasized hundred seed weight, oil content and days to 50% maturity, which were less correlated with the first three components and captured a distinct aspect of the variation. The resulting principal component plot (Fig. 3) displaying a broad dispersion of the 30 parental lines, indicating substantial diversity among them.

4. CONCLUSION

This study revealed significant genetic variability among 30 advanced castor breeding lines for 11

quantitative characters. Based on mean performance for 11 traits, K22-1, K22-5, K22-26, K22-39, K22-43, K22-46 and ICS-164 were identified as best monoecious lines, whereas IPC-41, IPC-46, IPC-44, IPC-48, IPC-52 and IPC-58 were identified as best pistillate lines.

Correlation analysis revealed significant positive associations between traits, including plant height, spikes per plant and capsules on primary spike with total seed yield per plant. Principal component analysis (PCA) identified primary contributors to variation, such as primary spike length, capsules on primary spike and plant height. K-means clustering divided the 30 breeding lines into 6 clusters based on K-cluster mean. Using superior lines from divergent clusters as parents could exploit maximum heterosis in breeding programmes.

DISCLAIMER (ARTIFICIAL INTELLIGENCE)

Authors hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of manuscripts.

ACKNOWLEDGEMENT

This work is supported by ICAR-Indian institute of oilseeds research (ICAR-IIOR) Rajendranagar, Hyderabad.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Vavilov NI. The origin, variation, immunity and breeding of cultivated plants. LWW; 1951.
2. Anjani K. Castor genetic resources: A primary gene pool for exploitation. Industrial Crops and Products. 2012;35: 1–4.
3. Moshkin VA. History and origin of castor. Castor. Oxonian Press Pvt. Ltd., New Delhi. 1986;6-10.
4. Ogunniyi DS. Castor oil: A vital industrial raw material. Bioresource Technology. 2006;97(9):1086-91.
5. Mutlu H, Meier MA. Castor oil as a renewable resource for the chemical industry. European Journal of Lipid Science and Technology. 2010;112(1):10-30.
6. Shrirame HY, Panwar NL, Bamniya BR. Bio diesel from castor oil-a green energy option. LowCarbon Economy. 2011;2(1):1.
7. DOR, 2006. Strategies for enhancing castor production in India. Directorate of Oilseeds Research,Hyderabad.
8. Indiastat. 2022-2023. Available:https://www.indiastat.com/default.aspx
9. Chaudhari D, Parmar V, Dube D, Bhakta R, Patel A, Lodam V. Genetic variability and divergence study in castor (*Ricinus communis* L.). Advances in Life Science. 2016;5(16):6418-6422.
10. Rukhsar Patel MP, Parmar DJ, Kumar S. Genetic variability, character association, and genetic divergence studies in castor (*Ricinus communis* L.). Annals of Agrarian Science. 2018; 16(2):143- 148.
11. Alhaji KA, Oliseh A, Salihu BZ, Kabaraini MA. Studies on genetic variability and trait relationship in castor (*Ricinus communis* L.). Journal of Plant Development. 2019; 26:3 11.
12. Kumar RM, Madival Y, Paramesh V, Narayanappa N, Ranganath G, Jinger D, Chillarge S. Optimizing crop nutrition in rainfed castor cultivation for higher productivity, energy efficiency, and reduced carbon footprint. Environmental and Sustainability Indicators. 2024;22:100409.
13. Dapke JS, Naik MR, Singh AP, Baraskar VV, Narwade AV, Indivar P. Genetic association yield with different agronomic traits in castor (*Ricinus communis L.).* International Journal of Biology Research. 2016;1(1):47-49.
14. Movaliya HM, Chovatia VP, Madariya RB, Mungala RA, Pipaliya HR, Bhuva SK. Study of variability and correlation for seed yield and its attributes in castor (*Ricinus communis* L.). Journal of Pharmacognosy and Phytochemistry. 2018;7(2):1474-7.
15. Patel JR, Saiyed MP, Patel CG, Bhatt RK, Bhatt VK. Genetic divergence in castor (Ricinus communis L.).
16. Ranjitha V, Venkatachalam SR, Arutchenthil P, Kathirvelan P. Genetic divergence in castor (Ricinus communis L.). Electronic Journal of Plant Breeding. 2019;10(2):766-71.
17. Destaw Mullualem DM, Getinet Alemaw GA, Yohannes Petros YP, Shiferaw Alemu SA. Phenotypic variability and association of traits among yield and yield-related traits

in castor (Ricinus communis L.) accessions at Melkassa, central rift valley of Ethiopia.
18. Tewari N, Kumar Y. Genetic variability and character association in castor (Ricinus communis L.). Current Advances in Agricultural Sciences (An International Journal). 2013;5(2):256-8.
19. Manjunatha T, Lavanya C, Senthilvel S, Patel MP, Ramya KT, Sarada C. Relationship of node number with days to flowering and yield traits in castor (Ricinus communis L.) hybrids under rainfed and irrigated conditions. J. Oilseeds Res. 2020;37(4):245-52.
20. Goodarzi F, Darvishzadeh R, Hassani A, Hassanzaeh A. Study on genetic variation in Iranian castor bean (Ricinus communis L.) accessions using multivariate statistical techniques. Journal of medicinal plants Research. 2012;6(7):1160-7.
21. Patel MP, Parmar DJ, Kumar S. Genetic variability, character association and genetic divergence studies in castor (Ricinus communis L.). Annals of Agrarian Science. 2018;16(2):

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/121281