Biological Forum – An International Journal **14(4): 876-881(2022)**

ISSN No. (Print): 0975-1130
ISSN No. (Online): 2249-3239

Multivariate based Genetic Diversity Analysis among Linseed (*Linum usitatissimum* L.) Germplasm Accessions by Principle Component Analysis

Konreddy Sathish[1*], Deepika Parte[1], Khemlata Thakur[1], Deenanath Sahu[2], T. Lakshmi Pathy[3] and Nandan Mehta[4]

[1]Ph.D. Scholar, Department of Genetics and Plant Breeding, College of Agriculture, IGKV, Raipur (Chhattishgarh), India.
[2]M.Sc. Scholar, Department of Genetics and Plant Breeding, College of Agriculture, IGKV, Raipur (Chhattishgarh), India.
[3]Scientist, ICAR-SBI, Coimbatore (Tamil Nadu), India
[4]Principal Scientist (Linseed), Department of Genetics and Plant Breeding, College of Agriculture, IGKV, Raipur (Chhattishgarh), India.

(Corresponding author: Konreddy Sathish)*
(Received 19 September 2022, Accepted 29 October, 2022)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: The present study was conducted at the research cum Instructional farm, Indira Gandhi Krishi Vishwavidyalaya, (IGKV), Raipur, India. A total of 100 linseed genotypes and four standard checks were evaluated in Augmented RBD design during two rabi seasons *rabi* 2020-2021 and *rabi* 2021-2022. To estimate the genetic diversity the observations were recorded for 10 quantitative yield attributing traits *viz.,* days to 50% flowering, days to maturity, plant height (cm), number of primary branches plant^{-1}, number of secondary branches plant^{-1}, number of capsules plant^{-1}, number of seeds capsule^{-1}, 1000 seed weight (g), seed yield plant^{-1}, oil content (%) to determine multivariate based diversity analysis using Principle Component Analysis (PCA). Crop diversity studies provides scope for plant breeders for the development of novel and improved cultivars with most desirable traits by selecting suitable parents and also in studying the magnitude of genetic variability present in the germplasm accessions. Therefore, the present study on agro-morphological traits to identify the superior germplasm accessions that can contribute as potential donors for future exploitation in the selection and breeding of linseed.

Keywords: Augmented design, genetic diversity, Principle Component Analysis (PCA), germplasm.

INTRODUCTION

Linseed is botanically called *Linum usitatissimum* L. and locally called as a vise in Telugu and alsi in Hindi. Linseed is a highly self-pollinated crop since the beginning of the civilization and is considered one of the earliest crops to have been domesticated (Zohary and Hopf 2000). Apart from the cultivated *Linum usitatissimum*, in total five wild linseed species has been reported in India so far and they are, *"L. perenne*, *L. strictum*, *L. mysorense*, *L. augustifolium* and *L. grandiflorum"*. Linseed is also known as *alsi*, flax or flaxseed and belongs to the genus name *Linumis* a Latin word origioriginating *lin* or *thread* and *'Usitatissimum'* literally means "most useful".It is also an important oil seed crop with diploid chromosome number, 2n=30.The domestication of flax was also observed on the Indian subcontinent near the Mediterranean Sea and this region is known to have high biological diversity of the genus *Linum* (Fu, 2005; Kaur, 2017). *Linum bienne* Mill. (= *L. Angustifolium* Huds.) is perhaps the oldest flax wild form cultivated (Kumar *et al.,* 2021). Domestication of linseed by our ancestors for centuries and its cultivation by disruptive selection for years led linseed into two types, oil seeded and fibrous plants. Hence, today linseed is cultivated as a dual-purpose crop for oil and fiber extraction. Linseed seeds are traditionally used for the preparation of chutney powder, which is also rich in protein content, fat, dietary fiber and other micronutrients. At the Industrial level linseed oil is used for the manufacturing of paints, varnish, oilcloth, linoleum and inks (Walsh, 1965). Linseed contains the highest oil content among the crop plants grown with 36-40% which is also the richest source of PUFA (Poly Unsaturated Fatty Acids). Linseed is a rich source of ω-3 fatty acid: α-linolenic acid (ALA), short-chain polyunsaturated fatty acids (PUFA), soluble and insoluble fibers, Phyto estrogenic lignans, proteins and also an array of antioxidants (Ivanova *et al.,* 2011; Singh *et al.,* 2011; Alhassane and Xu 2010).

Although the area, production, and productivity of linseed in India during 2020-2021 is productivity of 637 kg/ha with an area of 174.87 thousand tons per hectare and in Chhattisgarh the productivity of 384 kg/ha with an area of 12.82 thousand tons per hectare (Indiastat, 2021). The introduction of new germplasm, gathering of local landraces, and the use of interspecific hybridization are all necessary for characterizing and evaluating germplasm accessions due to the availability and existence of low genetic variability. The most valuable source of the necessary traits for creating improved cultivars is germplasm. Estimating the variety that already exists among the community of people constitutes characterization (Ranjana *et al.,* 2019). Furthermore, research on morphological diversity characterization plays a significant part in managing crop diversity. Therefore, it is necessary to preserve the many genotypes and also to investigate linseed diversity research for future breeding purposes.

MATERIALS AND METHODS

Experimental site and materials. The experiment was conducted at the research cum Instructional farm, Indira Gandhi Krishi Vishwavidyalaya, (IGKV), Raipur, India. One Hundred linseed genotypes were taken for investigation in the present study. This study was conducted during two *Rabi* seasons, 2020-2021 and 2021-2022, for the estimation of genetic diversityby recording various quantitative characters by using the four standard checks namely Neelum, Kiran, RLC-148 (Varsha alsi) and RLC-153. Climatological data on rainfall, rainy days, temperature, relative humidity (RH) and sunshine hours were recorded at the Meteorological Observatory Unit, Department of Agro-meteorology, IGKV, Raipur during the cropping period. Favorable weather conditions were noticed during the crop growth and investigation of the current study.

Experimental design and methodology. To study the linseed germplasm accessions for characterization yield and yield attributing traits. The experimental field was adjusted and well-prepared with all recommended packages of practices. Proper fertilizer and irrigation were given to the crop at regular time intervals and proper weeding is done to control the weeds and to keep the crop free from weeds. Visual examination of the crop is done and observations were recorded randomly for five plants in each plot. Data were recorded on five randomly selected plants for all the agronomic traits viz., days to 50% flowering, days to maturity, Plant height (cm), number of primary branches plant^{-1}, number of secondary branches plant^{-1}, number of capsules plant^{-1}, number of seeds capsule^{-1}, oil content (%) and seed yield plant^{-1}. Oil content (%) was recorded for 100 lines by using the Soxhlet apparatus (Ozioko, 2012) at the department of genetics and plant breeding, college of agriculture, IGKV, Raipur

Statistical Analysis. The numerous agro-morphological traits data were analyzed using the accepted statistical methods. To cluster the accessions based on genetic similarity, the recorded data on all ten quantitative parameters were statistically analyzed in XLSTAT, and Add in soft for multivariate based genetic diversity based on Principle component analysis (PCA). The PCA approach creates an Eigen vector for each axis and provides component scores for the characters by condensing the dimensions of multivariate data to a few key axes.

RESULTS AND DISCUSSION

Principal Component Analysis. The contribution and significance of each component to overall variance are measured using Principal Component Analysis. It is a manifold variable statistical analysis to sort the data with a huge number of connected variables into a significantly lesser set of unique variables through the linear grouping of the variables that account greatest of the variation existing in the original variables. Principal components are normally projected either from the correlation matrix or covariance matrix. When the variables are measured in dissimilar components, scale effects can impact the structure of derived components. In such circumstances, it converts desirable to standardize the variables. The coefficient of appropriate vectors reflects the extent to which each original variable with which each principal component is associated. It also calculates a trait's independent contribution to overall variance. A scree plot depicting eigen-value variation and the percent of cumulative variation expressed by them for different components has been shown (Fig. 1), eigen-value variation expressed for different 5 components shown (Fig. 2) and the percent of variation expressed by them for five components has been shown (Fig. 3).

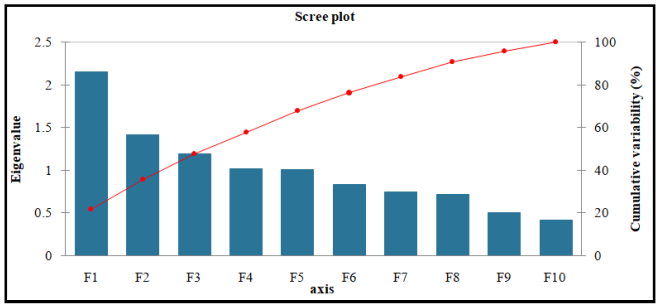

Fig. 1. Screen plot showing eigen value variation and % of expressed variation by different components.

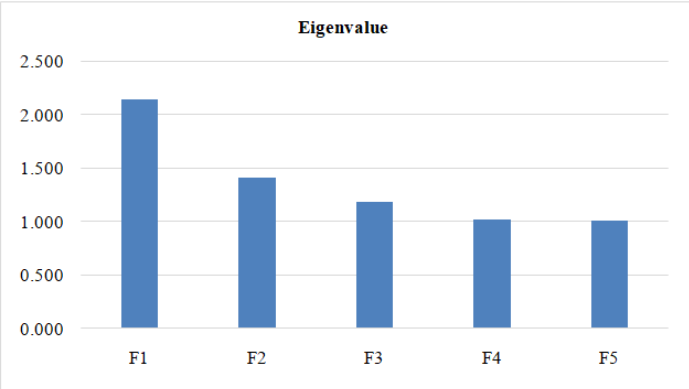

Fig. 2. Screen plot showing eigenvalue variation.

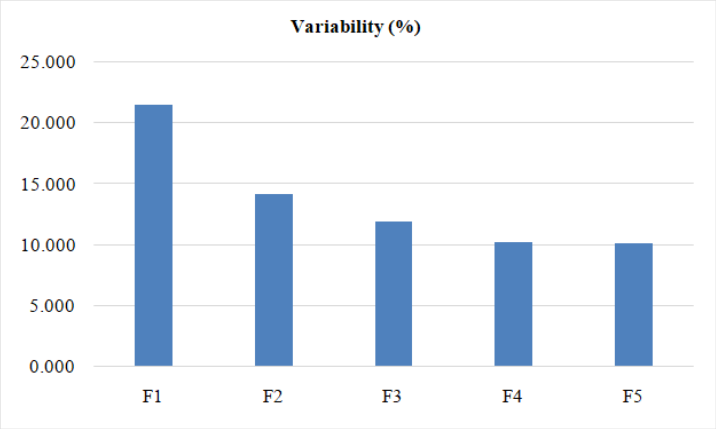

Fig. 3. Screen plot % of expressed variation by different components.

In the current exploration, PCA was accomplished for ten agro-morphological characters of Linseed and presented in Table 1. The PC with Eigenvalues of more than 1 and which explained at least 5% of the variation in the data were considered in the present study. The PC with higher Eigenvalues and variables which had high factor loading was considered the best representative of system attributes. Out of 10, only five principal components (PCs) exhibited more than 1.00 Eigenvalue and showed about 10% variability among the traits studied. So, these 5 PCs were given due importance for further explanation. The PC1 showed 21.54 % whereas, PC2, PC3, PC4 and PC5 revealed 14.20%, 12.00%, 10.20% and 10.10% variability correspondingly between the genotypes for the characters under study. The first PC accounts for as much of the variability in the data as possible, and each following component accounts for as much of the remaining variability as possible. A similar kind of work was done by Song *et al.* (2015); Kaur *et al.* (2018); Chakraborty *et al.* (2021); Kumar *et al.,* (2021).

Table 1: Principal components (PCs) of agro-morphological traits among *Linum usitatissimum* L. accessions.

	F1	F2	F3	F4	F5
Eigenvalue	2.16	1.42	1.20	1.10	1.00
Variability (%)	21.54	14.20	12.00	10.20	10.10
Cumulative %	21.54	35.74	47.66	57.85	67.94
Traits	Eigen vectors				
50% flowering	-0.114	-0.403	0.435	0.025	-0.474
Days to maturity	-0.192	-0.142	0.260	0.763	0.211
Plant height(cm)	0.235	0.058	0.334	-0.036	0.758
Primary branches plant-1	0.468	-0.221	-0.242	0.208	-0.149
Secondary branches plant-1	0.554	-0.191	-0.060	0.059	-0.057
Number of capsules plant-1	0.529	-0.112	0.075	-0.078	-0.024
Number of seeds plant-1	0.076	0.148	0.595	-0.460	-0.074
1000 seed weight(g)	-0.223	-0.292	-0.446	-0.348	0.259
Seed yield plant-1	0.181	0.530	-0.088	0.138	-0.058

Oil content(%)	0.016	0.573	-0.062	0.109	-0.233

By creating a graph between Eigen values and principle component numbers, the Scree plot was able to show the percentage of variation related to each principal component. With an Eigen value of 2.16, the PCI had 21.54% variability, which then rapidly decreased. After the third PC, the elbow form line tended to be straight with little variation seen in each PC. It is evident from the graph that PCI experienced the greatest amount of volatility (Fig. 1). Similar kind of analysis was also presented by Paul *et al.* (2017); Kumar *et al.* (2016); Thakur *et al.* (2021); You *et al.* (2017).

Cluster Analysis. Genetic divergence analysis was established to be a useful tool to measure the comparative involvement of divergent traits in the total divergence of both inter and intra-group intensities. Cluster analysis or clustering is the task of configuring a set of entities in such a manner that entities in a similar set (called a cluster) are more related to each other than to those in other sets (clusters). Cluster analysis has no tool for differentiating between related and unrelated variables. Hence, the choice of variables involved in cluster analysis is necessary to be strengthened by intangible kindnesses. This is very significant because the clusters made can be very reliant on the variables involved. In the present study, Euclidian distance between genotypes was calculated from the standardized data matrix by Un-weighted Pair Group Method using the Arithmetic Averages (UPGMA) method, and clustering was done by Agglomerative Hierarchical method using XLSTAT 2020 software.

Analysis was performed by the Unweighted variable Pair Group Method of the Average Linkage Cluster Analysis (UPGMA) using Euclidean distance as a dissimilarity measure divided the 100 entries and 4 checks of linseed into Five clusters (Table 2). The entries were not uniformly distributed between the clusters. Cluster II constituted 30 accessions, forming the largest cluster followed by cluster I and cluster III (24 accessions), cluster V and IV (16 and 10 accessions respectively). The pattern of group arrangement proved the reality of substantial aggregate variability.

Calculated Intra and Inter-cluster distances between ten clusters and were have been given in Table 3. The cluster distance ranged from 8.38 intra (cluster II) to 1397.02 (cluster V). The cluster distance was extreme between inter-cluster I and V (798.23) and inter-cluster distance was least detected between cluster I and cluster IV (755). To recognize much variability and high heterotic effect, parents should be particular from two clusters taking broader inter-cluster distances. The mean value for each cluster (Table 4) revealed that genotypes in cluster I had the highest values for traits, 1000 seed weight (g), and seed yield plant^{-1}. Cluster II had the highest values for oil content (%). Cluster III had the highest values for primary branches plant^{-1}, secondary branches plant,andnumber of capsules plant^{-1}. Cluster IV had the highest values for plant height (cm) and number of seeds capsule^{-1}. ClusterV had the highest values for 50% Flowering followed by days to maturity. Cluster analysis revealed a wide range of genetic divergence, which is useful for future hybridization breeding programs for getting desirable transgressive segregants. These results are supported with the verdicts of previous workers (Fulkar *et al.,* 2007; Srivastava *et al.,* 2009; Kandil *et al.,* 2011; Rajanna *et al.,* 2019).

Table 2: Distribution of 100 linseed genotypes into five cluster groups.

Clusters	No. of Genotypes	Genotypes
I	24	Neelum (C1), EC-41495, EC-99056, IC-AR-6, ILS-169, L-106, KANPUR-40/2, LCK-8605, L.S.-1, NP-10, NP-30, K A-153, NP-39, NP-65, NP-103, NP-138, NPRR-61, CI-1375, BAULK, KLS-A-3, KLS-B-2, R.S.-2, EC-1393-1, EC-1391
II	30	Kiran(C2), RLC-148(varsha alsi) (C3), RLC-153(C4), EX-304-1, Fzox Natural, AYOGI, L-103, NPRR-272, NPRR-449, Punjab T-4, JLT-90, KL-1, LC-2014, LCK-9303, LCK-9406, Sumerpur Local, EC-397752, KFS-11 65/12, A-9-2-1, A-469B, EC-41770, RSJ-31, KL-230, RKY-17, R.S.-6, BR-26, NL-97, L-470-Eng, CI-1971
III	24	EC-22529, EC-41595, A-58, EI-47-21-42, GS-128, A-97, A-236, A-195, OR-8-38, A-301, A-389, BNL-126, RLC-40, EC-4752, SJKO-8, SJKO-13, L-88-LHCK-7, B-81-81, RJK-34, OL-98-13-9, CANADA, CP-43, EC-1456, LCK-9319
IV	10	A-93, NPRR-68, RS-203, RJK-32, BAU-9906, BAU-06-8, BAU-08-07, SJKO-6, LCK-9119, CI-1413
V	16	RL-910, RLC-25, RLC-49, RR-464, S-91-43, LC-2057, NP-HYB-8, PKDL-55, BS-18, EC-718840, Raisa, NDL-8809, Dingoahi, EC-1453, LCK-9324, LCK-9325

Table 3: Estimates of intra (diagonal and bold) and inter cluster distances among five clusters.

Class	I	II	III	IV	V
I	**10.81**	10.99	10.48	7.55	798.23
II		**8.38**	18.36	10.62	797.40
III			**12.39**	11.94	797.63
IV				**10.73**	797.56
V					**1397.02**

Table 4: Estimates of quantitative trait means grouped under different clusters.

Cluster	I	II	III	IV	V
50% Flowering	66.23	66.97	66.97	66.90	**864.34**
Days to Maturity	107.71	114.65	110.34	110.95	**115.03**
Plant Height(cm)	38.13	38.94	37.05	**42.26**	33.50
Primary branches plant^{-1}	1.61	1.63	**2.33**	1.88	1.83
Secondary branches plant^{-1}	4.93	5.39	**9.12**	6.61	4.69
No. of Capsules plant^{-1}	24.21	15.91	**33.13**	23.39	14.54
No. of Seeds Capsule^{-1}	7.52	7.48	7.17	**8.31**	7.73
1000 seed weight (g)	**6.67**	5.98	5.75	6.27	6.61
Seed yield plant^{-1}	**6.61**	6.38	4.96	2.13	3.67
Oil Content (%)	36.72	**38.02**	36.61	34.60	34.16

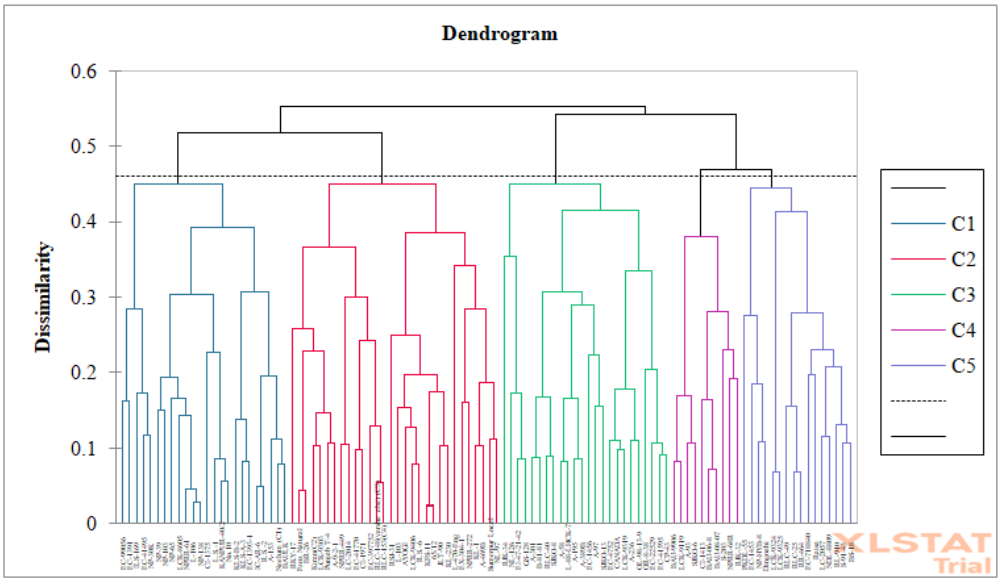

Fig. 4. Dendrogram showing the distribution of genotypes.

CONCLUSION AND FUTURE SCOPE

By Principal component analysis (PCA) few important traits were identified which have a primary role role in categorizing the existing variation in the germplasm accessions studied. The Primary branches plant-1, Secondary branches plant-1, No. of capsules plant^{-1}, and seed yieldplant^{-1} in different principlecomponents were shown to be the most essential foraccounting the variance. Hence, these characters must be given prior importance while selecting genotypes forvarietal development. Since, within a population variability is the important component plant breeding and also availability of genetic variability is must. From the findings, the high degree of genetic diversity among genotypes was identified, also the traits that are contributing towards thediversity of population. Therefore, the results found from this study will be assisting in selecting suitable parental lines for enhancing different morphological traits.

Acknowledgement. Authors of the present study are highly thankful to chairperson the program, Department of Genetics and Plant Breeding, CoA, IGKV, Raipur for providing logistic support necessary to conduct the research work.
Conflict of Interest. None.

REFERENCES

Alhassane, T. and Xu, X. M. (2010). Flaxseed lignans: source, biosynthesis, metabolism, antioxidant activity, bio-active components and health benefits. *Compr. Rev. Food Sci. Food Saf., 9*, 261–269.

Anonymous (Indiastat 2021).

Chakraborty, S., Kumar, A., Kishore, C., Singh, R. S., Kumar, A., Kumar, A. and Kumar, M. (2021). Multivariate Based Diversity Analysis in Indian Mustard (*Brassica juncea* L.) Genotypes by Principle Component Analysis. In *Biological Forum–An International Journal*, *13*(3), 680-685).

Fu, Y. B. (2005). Geographic Patterns of RAPD Variation in Cultivated Flax. *Crop Sci., 45*, 1084–1091.

Fulkar, P. L., Ghorpade, P. B., Maheshwari, J. J., Patil, S. R., Reddy, M. N. and Pavithran, C. (2007). Evaluation of linseed germplasm forgenetic divergence and choice of parents. *Journal of Soils and Crops*, *17*, 333-338.

Ivanova, S., Rashevskaya, T. and Makhonina, M. (2011). Flaxseed additive application in dairy products production. *Procedia Food Sci., 1*, 275–280.

Kandil, A. A., Sheraif, A. E., Abo-Zaied, T. A. and Ganil, A. (2011). Genetic divergence and heterosis in linseed (*Linum usitatissimum* L.). *Mansoura Journal of Plant Production*, *2*, 335-349.

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.

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

Kumar, N. and Paul, S. (2016). "Selection criteria of linseed genotypes for seed yield traits through correlation, path coefficient and principal component analysis." *JAPS: Journal of Animal & Plant Sciences, 26*(6).

Ozioko, Francis Uchenna (2012). "Extraction and characterization of soybean oil based bio-lubricant." *AU Journal of Technology, 15*(4).

Paul, S., Kumar, N. and Chopra, P. P. (2017). Correlation and genetic diversity of linseed (*Linum usitatissimum* L.) genotypes based on principal component analysis in Mid-Hills of North-West Himalayas. *Journal of Pharmacognosy and Phytochemistry, 6,* 287-290.

Rajanna, P., Paul, S. A., Sharma, D. E., Sood, V. K, Kumar, N. I. (2019). Morphological characterization and genetic diversity of linseed (*Linum usitatissimum* L.). *Journal of Oilseeds Research, 36*(1), 8-16.

Singh, K. K., Mridula, D., Rehal, J. and Barnwal, P. (2011). Flaxseed- a potential source of food, feed and fiber. *Crit. Rev. Food Sci. Nutr., 51*, 210–222.

Song, J., Dang, Z., Zhang, J., Wang, L., Li, W., Zhang, Y. and Zhang, Q. (2015). Principal component analysis and cluster analysis of oil flax germplasm based on agronomic traits. *Southwest China Journal of Agricultural Sciences*, *28*(2), 492-497.

Srivastava, R. L., Singh, H. C., Karam, H., Malik, Y. P. and Prakash, O. (2009). Genetic divergence in linseed, *Linum usitatissimum* L. under salt stress condition. *Journal of Oilseeds Research*, *26*, 159-161.

Thakur, R., Paul, S., Thakur, G., Kumar, A. and Dogra, R. (2020). Assessment of genetic diversity among linseed (*Linum usitatissimum* L.) germplasm based on morphological traits. *Journal of Pharmacognosy and Phytochemistry*, *9*(6), 1785-1790.

Walsh, R. J. (1965). Linseed oil protection for New York State Thruway bridges. *Civil engineering*, *15*(2), 39-41.

You, F. M., Jia, G., Xiao, J., Duguid, S. D., Rashid, K. Y., Booker, H. M., & Cloutier, S. (2017). Genetic variability of 27 traits in a core collection of flax (*Linum usitatissimum* L.). *Frontiers in plant science*, *8*, 1636.

Zohary, D. and Hopf, M. (2000) Domestication of plants in the old World, 3rd edn. Oxford University Press, Oxford, 125–132.

How to cite this article: Konreddy Sathish, Deepika Parte, Khemlata Thakur, Deenanath Sahu, T. Lakshmi Pathy and Nandan Mehta (2022). Multivariate based Genetic Diversity Analysis among linseed (*Linum usitatissimum* L.) Germplasm Accessions by Principle Component Analysis. *Biological Forum – An International Journal, 14*(4): 876-881.