MULTIVARIATE ANALYSIS OF GENETIC DIVERSITY IN LINSEED GENOTYPES FOR YIELD AND ITS COMPONENT TRAITS

ABSTRACT

In the present investigation a total 30 linseed genotypes including released varieties were evaluated during *Rabi* 2021-22 at Agricultural Research Station, Ummedganj, Kota, Rajasthan (India) for understanding genetic diversity for grain yield and its component traits using principal component analysis. PCA was utilized to examine the variation and to estimate the relative contribution of various traits for total variability, first four principal components have more than one eigen values and accounted for 70.87% of total cumulative variance among 11 traits. PC1 had the contribution from traits *viz.*, days to maturity, plant height, capsule per plant, seed per capsule, plant stand, test weight, branches per plant, germination per cent and yield per plant, which accounted for 26.08% of the total variability whereas PC2, PC3 and PC4 exhibited 20.13%,13.43% and 11.23% to the total variability respectively. Thus, the results of principal component analysis revealed the wide genetic variation in linseed genotypes indicating that these accessions may be used as donors to improve the yield and quality traits in varietal development program. On the basis of Ward's linkage cluster analysis,five cluster were formed to identify relative closeness among 30 genotypes. Cluster V consisted of maximum 9 genotypes followed by cluster I and IV (8), cluster III (3) and cluster II (2). Maximum inter cluster distance was recorded between cluster I and IV indicating the possibilities of high heterosis if individual from these clusters were cross-bred. Cluster I had the highest mean values for plant height, seed per capsule, plant stand and yield per plant. Hence, suggesting that the genotypes constituted in these clusters may be used as a parent for future hybridization programs.

KEYWORDS: Cluster, Diversity,Eigen value, Linseed, Principal Component Analysis.

Comment [WU1]: Insert of

1. INTRODUCTION

Linseed (*Linumusitatissimum* L. 2n=30) also known as flax is a member of the genus *Linum* in the family Linaceae. It is one of the oldest crops being cultivated since the beginning of civilization. Every part of the linseed plant is utilized commercially, either directly or after processing (Qamar *et al*., 2019) Compared with other oil crops, linseed containing about 36-40

% oil is the richest source of polyunsaturated fatty acids (PUFA) like alpha-linolenic acid, the most abundant source of antioxidants (Andruszczak*et al*., 2015 and Goyal *et al*., 2014)and lignin (Kajla *et al*., 2015), which has a beneficial effect on the prevention of the diseases and human health.The oil has drying and hardening properties which emanates from the high linolenic acid (45-60%) content, and therefore it is mostly used for industrial purposes such as manufacturing of paints, varnishes, soaps and printing inks (Wakjira, 2007 and Biradar*et al*., 2016) while low linolenic acid content is necessary for its human consumption.Keeping in view, the increasing demand for linseed due to numerous health benefits, there is a consistent need to increase the genetic potential for seed yield.

In India it is grown in 181 thousand ha with a production of 41 thousand tones and productivity of 227kg/ha(Anonymous,2020-21). The area under linseed cultivation in Rajasthan is 5.7 thousand ha and production is 6.1 thousand tones with an average yield of 1066 kg/ha(Anonymous, 2021).There is consistent need to increase seed yield potential to increase the demand of linseed. Recombination of favorable genes is one of the ways to increase seed yield potential and related traits. One of the important approaches to linseed breeding is hybridization and subsequent selection. Appropriate selection of the parents is essential to be used in crossing to enhance the genetic recombination for potential yield increase (Islam *et al*., 2004). Genetic distance between parents is necessary to benefit transgressive segregation (Joshi *et al*., 2004). Higher the genetic distance between parents, the higher heterosis in progeny can be observed (Anand and Murrty 1968).

Comment [WU2]: Small r

Principal Component Analysis (PCA) data consists of numerous intercorrelated quantitative dependent variables. It collects data from a table and displays it as a collection of new orthogonal variables. It extracts the information from a table and represents it as a set of new orthogonal variables called principal components also it may be used to reveal patterns and eliminates redundancy in data sets (Maji and Shaibu, 2012).It reduces the dimensions of a multivariate data to a few principal axes, generates an Eigen vector for each axis and produces component scores for the characters (Leonard and Peter, 2009). It analyses data consisting of several inter correlated quantitative dependent variables as observations (Kaur *et al*., 2018 and Mahendran *et al*., 2015). By using a few components, each sample can be represented by relatively few numbers instead of by values for thousands of variables (Ringer, 2008). Thus, the

primary benefit of PCA arises from quantifying the importance of each dimension for describing the variability of a data set in more interpretable and more visualized dimensions through linear combinations of variables that accounts for most of the variation present in the original set of variables. The higher the coefficients, regardless of the direction (positive or negative), the more effective they will be in discriminating between accessions, whereas Cluster analysis, which is based on generalized distance, is a useful tool for determining the level of genetic divergence at the genotypic level (Mahalanobis, 2015). Multivariate analysis, which includes cluster analysis and PCA, has already been shown to be an effective biometrical tool for evaluating the degree of variability in a germplasm collection of different crop plants (Acquaah, 2012). Hence, the main objective of this study was to assess the potential genetic diversity by using PCA method for selection of parents in hybridization programme to obtain desirable segregants in advanced generation. Multivariate statistical approaches are commonly used to summarize and characterize the intrinsic diversity among genotypes and to measure the degree of difference between the genotypes.

2. MATERIAL AND METHODS

The germplasm used in this study consisted of 30linseed accessions(Table 1). The trial was conducted during *Rabi* 2021-2022 at Agricultural Research Station, Ummedganj, Kota, in Randomized Block Design (RBD) using three replications. Each entry had four rows of 3 m length, having a spacing of 30 cm between rows and 10 cm between plants. For each genotype, plant height (cm), capsule per plant, seed per capsule, plant stand, test weight (g), branches per plant and yield per plant (g) were determined on randomly selected five plants. On a full plot basis, days to 50% flowering, days to maturity and germination per cent were recorded. To uncover the best connections among features, R software was used to perform principal component analysis (PCA) and a hierarchical clustering.

3. RESULTS AND DISCUSSION

3.1 Principal component analysis

Principle Component Analysis revealed that out of eleven characteristics studied, only four principal components (PCs) exhibited more than 1.00 eigen value. ‘Eigen values’ measure the importance and contribution of each component to total variance, whereas each coefficient of

eigen vectors indicates the degree of contribution of every original variable with which each principal component is associated. There are no standard tests to prove significance of Eigen values and the coefficients (Jollife, 2002),for the selection of various parents, the main components with more than one eigenvalue showed greater variability among the linseedgenotypes. According to PCA, the first four PCs explained 70.87 per cent of the variation in yield and yield component attributes among 30 genotypes(Table 2). Scree plot elucidated the variation percentage between Eigenvalues and the Principal components (Fig. 1). Among the four principal components, PC1 shared high proportion of total variation 26.08% with an Eigenvalue of 2.86. In comparison to the other PCs, the graph clearly shows that PC1 showed the most variability.The rest of the three principal components *viz*., PC2, PC3 and PC4 contributed 20.13%, 13.43% and 11.23% of the total variance respectively. The Eigen values are gradually declined from PC1 to PC4. The Eigen values are 2.21, 1.47 and 1.23 for PC2, PC3 and PC4, respectively. Elbow type line is obtained after PC4 tended to straight with minute difference observed in each PC.Similar results were earlier reported by (Dabalo*et al*., 2020).

Results of rotated component matrix showed that the first principal component PC1 which accounted for the maximum variabilitywas positively contributed by characters such as days to maturity (0.30), plant height (0.40), capsule per plant (0.48), seed per capsule (0.11), plant stand (0.00), test weight (0.15), branches per plant (0.38), germination per cent (0.14) and yield per plant (0.14), clearly indicated that the variation in PC1 is mainly due to these characters. Hence, the genotypes selected from PC1 would be useful in future breeding programmes for the improvement of the traits contributing to maximum variability.

PC2 was positively contributed by characters *viz*., days to 50% flowering (0.42), days to flowering (0.45), days to maturity (0.00), plant height (0.20), capsule per plant (0.15), plant stand (0.47), branches per plant (0.27) and germination per cent (0.47), while characters *viz.,* yield per plant (-0.03), test weight (0.06) and seed per capsule (0.09)showed negative loadings. PC3 had the positive contribution from the characters *viz.,* days to flowering (0.09), days to maturity (0.49), capsule per plant (0.06) and branches per plant (0.23), while remaining factors showed negative loadings. PC4 had the positive contribution from the characters *viz.*, days to 50% flowering (0.13), days to flowering (0.20), capsule per plant (0.19), seed per capsule (0.05), test weight (0.79), branches per plant (0.19) and germination per cent (0.03). Consequently, the most

diverse accessions could be selected on the basis of these four PCs. (worku*et al*., 2015; Kaur *et al*., 2018 and Patial *et al*., 2019)reported similar results and thus confirmed the independence of some traits.

The biplot diagram gives the picture of interaction among the characters and also the genotypes performing better for the traits. The vector length of each trait depicts its contribution to total divergence, longer the vector length, more is the contribution of concerned traits. In this study, the distribution and nature of diversity for genotypes and quantitative traits are described in the biplot diagram (Fig. 2) between PC1 and PC2.The trait yield per plant (g) and capsule per plant showed maximum vector length indicating its contribution to the total divergence followed by branches per plant, plant height and days to maturity.

3.2 Diversity analysis

Cluster analysis groups a large number of accessions into a few numbers of homogeneous clusters which in turn facilitates the selection of diverse accessions. The genotypes were grouped into five clusters based on non-hierarchical euclideancluster analysis statistics as shown in Table 3 and Fig 3. The number of accessions fallen under cluster V was the highest (9) followed by cluster I (8), cluster IV (8), cluster III (3) and cluster II (2). Distribution pattern of all the genotypes into five clusters showed the presence of considerable genetic diversity among the genotypes for most of the traits under consideration. Different clustering patterns were also reported in linseed by some earlier workers(Kandil *et al*., 2011; Pali and Mehta, 2017 and Patial *et al*., 2019). The cluster mean of different quantitative characters for different clusters have been presented in Table 4. The maximum value for grain yield per plant was recorded in cluster I followed by cluster V, cluster III and cluster II.Cluster I had genotypes having higher mean values for characters like days to maturity, plant height, seeds per capsule, plant stand, yield per plant and also desirable for early flowering trait. Whereas, clusters V had high mean values for yield contributing traits *viz*., capsule per plant, branches per plant and germination per cent. In order to generate transgressive segregates with higher yield, the genotypes in the above cluster may be targeted for multiple crossing programmes.

Table 1: List of 30 linseed genotypes along with source used in the study.

1.	RL-15597	9.	RL-18113	17.	RL-18121	25.	RL-18102
2.	RL-18104	10.	RL-18114	18.	RL-18122	26.	RL-18103
3.	RL-18105	11.	RL-18115	19.	RL-18123	27.	PA-2
4.	RL-18106	12.	RL-18116	20.	RL-18124	28.	KBA-5
5.	RL-1007	13.	RL-18117	21.	RL-13165	29.	KBA-6
6.	RL-18110	14.	RL-18118	22.	RL-18109	30.	T-397
7.	RL-18111	15.	RL-18119	23.	RL-18107	Source- AICRP on Linseed and ARS, Kota	
8.	RL-18112	16.	RL-18120	24.	RL-18101		

Comment [WU3]: Note : Describe RL for understanding easily

Table 2:-Principal Component analysis (PCA) based on standard data for 11 quantitative traits of linseed.

Characters	PC1	PC2	PC3	PC4
Days to 50% flowering	-0.38	0.42	-0.03	0.13
Days to flowering	-0.31	0.45	0.09	0.20
Days to maturity	0.30	0.00	0.49	-0.22
Plant height	0.40	0.20	-0.03	-0.22
Capsule per plant	0.48	0.15	0.06	0.19
Seed per capsule	0.11	-0.09	-0.59	0.05
Plant stand	0.00	0.47	-0.12	-0.32
Test weight	0.15	-0.06	-0.04	0.79
Branches per plant	0.38	0.27	0.23	0.19
Germination per cent	0.14	0.47	-0.23	0.03
Yield per plant	0.25	-0.03	-0.49	-0.13
Eigen values	**2.86**	**2.21**	**1.47**	**1.23**
Percent variance	**26.08**	**20.13**	**13.43**	**11.23**
Cumulative Proportion	**26.08**	**46.21**	**59.64**	**70.87**

Table 3: Grouping of thirty genotypes of linseed into five clusters.

Cluster No.	Name of genotypes	Number of genotypes
I	RL-18101, RL-18107, RL-18124, PA-2, RL-1007, RL-18102, RL-13165, KBA-5	8
II	RL-18104, T-397	2
III	RL-18119, RL-18120, RL-18121	3
IV	RL-18103, RL-18110, RL-18116, RL-18117, RL-18113, RL-18105, RL-18114, RL-18115	8
V	RL-18118, RL-18123, RL-18106, KBA-6, RL-18109, RL-18112, RL-18122, RL-15597, RL-18111	9

Table 4: Mean values of different characters for thirty genotypes of linseed grouped in different clusters.

Cluster No	Character's Name										
	DFF	DF	DM	PH	CPP	SPC	PS	TW	BPP	GP	YPP
I	61.000	65.000	121.000	79.270	111.170	8.530	353.000	7.830	4.330	70.000	2.183
II	62.000	66.000	120.800	60.014	93.122	7.816	311.800	7.668	4.174	73.000	1.743
III	61.077	65.385	120.769	62.201	87.425	8.161	269.154	7.720	3.918	66.409	1.916
IV	60.750	65.000	120.750	60.785	92.325	8.165	205.250	7.975	4.035	60.000	1.619
V	60.143	65.000	122.143	70.280	113.014	8.303	271.571	7.870	4.409	73.333	2.151

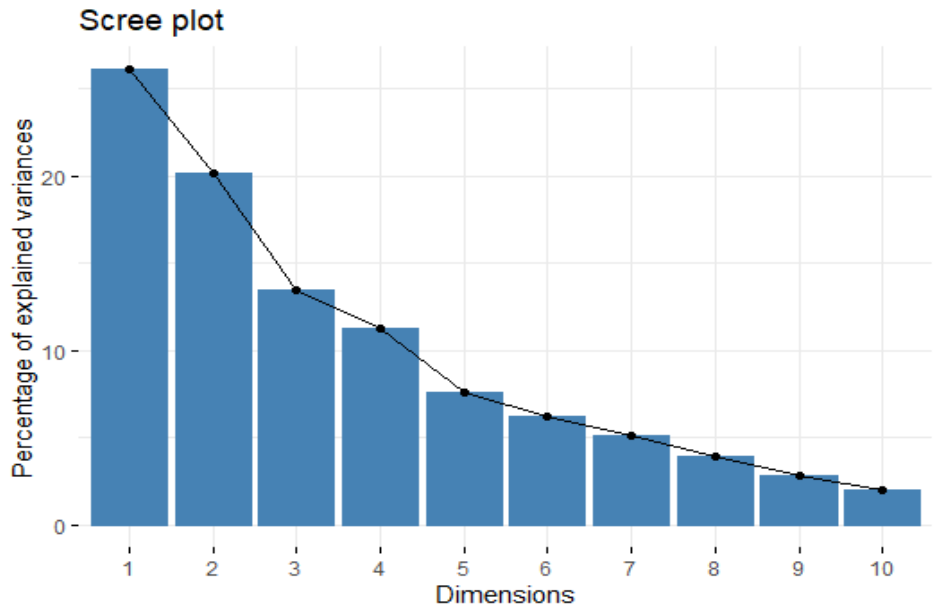

Figure 1: Scree plot for contribution of different principal components

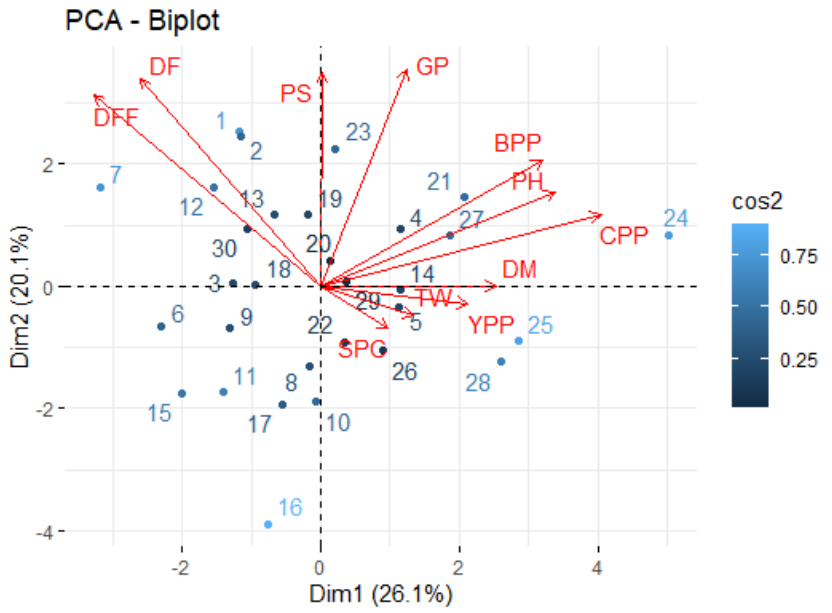

Figure 2: **Biplot of 30 genotypes of linseed on principal component axis 1 and 2.**

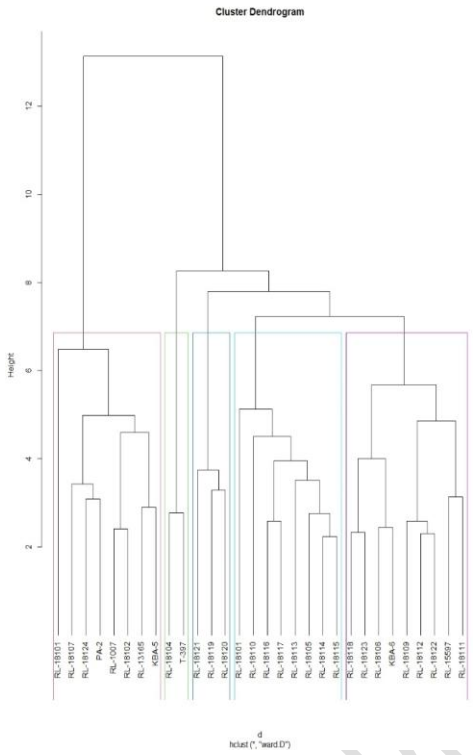

Figure 3: Dendrogram depicting relationships among the linseed genotype for different traits

4. CONCLUSION

Both multivariate statistical approaches revealed that the germplasm used in the study had tremendous of genetic diversity. Four of the eleven principal components were significant (eigenvalue >1) and accounted for 70.87 per cent of the variance, according to PCA. The yield and its contributing characteristics dominated in PC1 and PC2. As a result, selecting germplasm with a high PC1 and PC2 score may result in increased yield and yield characteristics, by harnessing heterosis. High-performing germplasm from each cluster can be employed in a hybrid breeding programme to generate superior high yielding lines. As a result, breeders will benefit greatly from the research in terms of selecting promising parental lines from germplasm in the study.

6. REFERENCES

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

Anand, I. J. and Murrty, B. R.1968. Genetic divergence and hybrid performance in linseed. *Indian Journal of Genetics and Plant Breeding*.**28**(2):178-85.

Andruszczak, S., Gawlik-Dziki, U., Kraska, P.,Kwiecinska-Poppe, E.,Rozyło, K.and Pałys, E. 2015. Yield and quality traits of two linseed (*Linum usitatissimum* L.) cultivars as affected by some agronomic factors. *Plant Soil and Environment*,**61**: 247–252.

Anonymous, 2021. Directorate of Economics and Statistics, Department of Agriculture Cooperation and Farmers Welfare. Government of India.

Anonymous, 2020- 2021. Annual Project Report. ICAR-All India Coordinated Research Project on Linseed.

Biradar, S.A.,Kumar, K.A., Rajanna, B.,Savitha, A.S.,Shubha, G.V.,Shankergoud, I.,Chittapur, B.M. and Singh, P.K. 2016. Prospects and challenges in linseed (*Linum usitatissimum* L.) production: A review. *Journal of Oilseeds Research*,**33**(1): 1-13.

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

Goyal, A., Sharma, V., Upadhyay, N., Gill, S. and Sihag, M. 2014. Flax and flaxseed oil: an ancient medicine & modern functional food. *Journal of Food Science and Technology*,**51**(9):1633-53.

Islam, M.R., Faruquei, M.A.B., Bhuivan, M.A.R., Biswas, P.S. and Salam, M.A. 2004. Genetic diversity in irrigated rice. *Pakistan Journal of Biological Sciences*,**7**(2):226-29.

Jolliffe IT. Principal component analysis (2nd Edn.). Springer-Verlag, New York, USA. 2002.

Joshi, B.K., Mudwari, A., Bhatta, M.R. and Ferrara, G.O. 2004. Genetic diversity in Nepalese wheat cultivars based on agro-morphological traits and coefficients of parentage. *Nepal Agriculture Research Journal*,**5**(1):7-17.

Kajla, P., Sharma, A. and Sood, D.R.2015. Flaxseed—a potential functional food source. *Journal of Food Science and Technology-Mysore*, **52**: 1857–1871.

Kandil, A.A., Sheraif, A.E., Abo-Zaied, T.A. and Ganil, A. 2011. Genetic divergence and heterosis in linseed (*Linum usitatissimum* L.).*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. *Journal of Environmental Biology*, **39**(5):702-709.

Leonard, K. and Peter, R.J. 2009. Finding Groups in Data: An Introduction to Cluster Analysis. 344.

Mahalanobis, P.C. 1936. On the generalized distance in statistics.In Proceedings of the National Institute of Science of India. Calcutta.**2**:49–55.

Mahendran, R., Veerabadhira, P., Robin, S. and Raveendran, M.2015. Principal component analysis of rice germplasm accessions under high temperature stress. *International Journal of Agricultural Science and Research*, **5**(3): 355–360.

Maji, A.T. and Shaibu, A.A. 2012. Application of principal component analysis for rice germplasm characterization and evaluation. *Journal of Plant Breeding and Crop Science*,**4**(6): 87–93.

Nachimuthu, V.V., Robin, S., Sudhakar, D., Raveendran, M., Rajeswari, S. and Manonmani, S. 2014. Evaluation of rice genetic diversity and variability in a population pannel by principal component analysis. *Indian Journal of Science and Technology*, **7**(10):1555–1562.

Pali, V. and Mehta, N. 2017. Genetic diversity assessment of flax (*Linum usitatissimum* L.) germplasm using molecular and morphological markers. *Electronic Journal of Plant Breeding*,**7**(4): 986-995.

Patial, R., Paul, S., Sharma, D., Sood, V.K. and Kumar, N. 2019. Morphological characterization and genetic diversity of linseed (*Linum usitatissimum* L.). *Journal of Oilseed Research*,**36**(1):8-16.

Qamar, H., Ilyas, M., Shabbir, G., Irshad, G., Nisar, F., Abbas, S.M., Ghias, M. and Arshad, A. 2019. Flax: ancient to modern food. *Pure and Applied Biology*,**8**: 2269–2276.

Ringer, M. 2008. What is principal component analysis?.*Nature Biotechnology*,**26**(3): 303–304.

Wakjira, A. 2007. Linseed (*Linum usitatissimum* L.) In: Vegetable oils and fats, Plant Resources of Tropical Africa (PROTA), Vandervosson, HAM and Kamilo GSM (ed) PROTA Foundation, Wageningen, Netherland.108-115.

Worku, N., Heslop-Harrison, J.S. and Adugna, W. 2015. Diversity in 198 Ethiopian linseed (*Linum usitatissimum* L.) accessions based on morphological characterization and seed oil characters. *Genetic Resources and Crop Evolution*,**62**:1037-1053.