Eco. Env. & Cons. 30 (May Suppl. Issue) : 2024; pp. (S344-S348)
Copyright@ EM International
ISSN 0971–765X

DOI No.: http://doi.org/10.53550/EEC.2024.v30i03s.059

Assessment of genetic diversity in Indian sesame (*Sesamum indicum* L.) germplasm

T. Ezhilarasi[1]*, A. Mahalingam and N. Manivannan[1]

Regional Research Station, Vridhachalam,
[1]Centre for Plant Breeding & Genetics,
Tamil Nadu Agricultural University, Coimbatore 641 003, T.N., India

(Received 6 September, 2023; Accepted 6 November, 2023)

ABSTRACT

Principal component analysis and hierarchical clustering were carried out for 151 Indian sesame germplasm for five quantitative traits, *viz*., days to fifty percent flowering, days to maturity, plant height (cm), number of capsules per plant and seed yield per plant (g). Among five principal components, PC I and PC II alone recorded above 1 eigenvalue. The first two principal components accounted for 62.45% of cumulative total variations. PC I contributed the maximum variability of 40%, followed by PC II (21.68%) of the total variation. Days to fifty percent flowering and days to maturity had high and negative loading values in PC I. In the case of PC II, plant height had the highest and negative loading value. Among characters, days to flowering and maturity had a positive association with seed yield per plant. Similarly, plant height had a positive association with seed yield per plant. Hence these traits may be considered as selection indices for the improvement of seed yield. Sesame germplasm accessions were grouped into eight distinct clusters using the agglomerative clustering method. The hierarchical cluster analysis indicates that cluster I had the highest genotypic members (58) followed by cluster II (50 genotypes). Cluster IV holds 15 genotypes and Cluster V has 11 genotypes. Clusters III and VI had seven genotypes for each cluster. However, cluster VII holds only one genotype and cluster VIII has two genotypes. In the present study, the level of variability in sesame germplasm was found to be similar in both principal component analysis and hierarchical cluster method.

Key words: *Cluster analysis, Genetic diversity, Germpalsm, Principal component analysis, Sesame*

Introduction

Sesame (*Sesamum indicum* L.) is one of the most important oldest oilseed crops cultivated in semi-arid tropics and subtropical regions of Asia and Africa. It belongs to the family Pedaliaceae and is native to Africa. Sesame is considered the queen of the oilseed crop because of its stabilized keeping quality of oil. Sesame seed contains about 50% oil content, 18% protein content, 20% carbohydrates and 12% dietary fiber. Sesame oil is highly rich in antioxidants such as sesamin, sesmolin and sesaminol. The sesame seeds are used for baking and confectionaries. In 2017 globally, 99.83 lakh ha of the area is under sesame cultivation with a production of 55.31 metric tons and productivity of 550 kg/ha. In India, sesame is cultivated in an area of 17.30 lakh ha with a productivity of 431 kg/ha. India shares 12.4 percent of world sesame production (Daisy Myint *et al*., 2020).

Seed yield is a complex trait in sesame and it is reportedly associated with several component traits and interrelated. Selection of genotypes for seed yield and yield contributing traits will be only effective if the desired genetic variability is present in the gene pool. Genetic diversity studies are an efficient tool to select parents for hybridization for a success-

ful breeding programme.

Principal Component Analysis (PCA), Cluster analysis and discriminate analysis are important multivariate analysis methods (Oyelola, 2004). Principal component analysis transforms a set of correlated variables into a new set of uncorrelated variables that are principal components that are derived in decreasing order. The first principal component accounts for as much of the variance of the original data as possible, followed by the second, *etc*. (Richman, 1986). It helps in reducing the data dimensionality sets and it increases reliability. This technique has been widely used for studying genetic variability in sesame (Teklu *et al*., 2021; Durge *et al*., 2022 and Sasipriya *et al*., 2023). Crossa, (1990) stated that cluster analysis is a numerical arrangement technique that explains groups of individuals and hierarchical classification, which is important for categorizing individuals into groups having the objective of studying associations in the data.

The present study aims to ascertain the level of germplasm variation in sesame to identify diverse phenotypes and to group the diverse phenotypes based on the similarity of agronomic traits by taking into account of several characteristics and relationships between them. This study will pave the way for the selection of diverse parents for the improvement of sesame.

Materials and Methods

A total of 151 Indian sesame germplasm available at the gene pool of Regional Research Station, Tamil Nadu Agricultural University, Vridhachalam was utilized for the present study. The experimental area is located at about a latitude of 11.30°N and a longitude of 79.26°E with red laterite/sandy soil topography at an altitude of 46.7 meters above mean sea level. The average annual rainfall is around 1050 mm. Sesame germplasm was raised during the summer 2019 at Regional Research Station, Tamil Nadu Agricultural University, Vridhachalam, Cuddalore District, Tamil Nadu, India to study genetic diversity and seed multiplication. Each germplasm was sown in 4m of row length with a spacing of 30 cm between rows and 30 cm between plants. All the recommended agronomic practices were strictly followed to raise a healthy crop. Morphological traits *viz*., days to fifty percent flowering, plant height (cm), number of capsules per plant, and seed yield per plant (g) were recorded on five randomly selected plants of each genotype. Biometrical data were statistically analysed to identify phenotypically similar groups based on the five biometrical traits studied. Analysis of variance using descriptive statistics such as mean and standard deviation for all the traits were calculated. Grouping of genotypes into similar phenotypic groups was carried out using Ward's hierarchical algorithm based on squared Euclidean distances. Principal Component Analysis (PCA) was used to study the variations present in the gene pool. The observations recorded were statistically analysed using the Statistical Tool for Agricultural Research (STAR 2.0) package for analysis. The principal component analysis was computed using the following equation: PC1= P Ó1ajXj where ajXj = Linear coefficient – Eigen vectors. Cluster analysis was performed using the agglomerative clustering method adopting Euclidean distance measure.

Results and Discussion

First-order statistics, namely the mean, standard deviation, minimum and maximum values for the quantitative traits studied are shown in Table 1. Days to fifty percent flowering ranged from 30 to 42 days, with a mean value of 35 days. The mean value for days to maturity was 87 days, ranging between 82 and 95 days. Accordingly, the genotype Si 1246 has recorded an early flowering of 30 days with an early maturity of 82 days. Though the genotypes Si

Table 1. Characteristic mean values of sesame germplasm

Variable	Mean	SD	Minimum value	Accessions	Maximum value	Accessions
Days to Fifty percent flowering	35	2.19	30.00	Si 1246	42.00	Si 983 PSR 2987 IC 204541
Days to Maturity	87	2.51	82.00	Si 1246	95.00	IC 204541
Plant height (cm)	98.75	7.70	80.00	Si 1171	147.00	IS 564
No. of capsules per plant	55.62	11.50	32.00	Si 37	88.00	Si 864
Seed yield per plant (g)	8.80	2.29	5.40	IS 62-1	15.60	TN-8452- 2

983, PSR 2987 and IC 204541 expressed late flowering of 42 days, the genotype IC 204541 alone expressed late maturity of 95 days duration. The germplasm recorded a mean plant height of 98.75 cm, with the shortest stature of 80 cm (Si 37) and the tallest plant of 147 cm (Si 864). The genotypes Si 37 and Si 864 have registered the lowest number of capsules per plant (32) and the highest number of capsules per plant (88) respectively. The germplasm recorded a mean seed yield of 8.8g per plant. Among genotypes, IS 62-1 was found to be the low yielder with 5.4g of seed yield and the genotype TN-8452-2 was found to be the high yielder with 15.60g seed yield per plant.

Principal Component Analysis (PCA) is used to distinguish the similarities between variables and to classify the genotypes, while cluster analysis on the other hand is concerned with classifying previously unclassified materials. PCA plays a pivotal role in selecting the most significant data from a large number of data. PCA is analysed by a sequence of steps, by the number of components extracted is equal to the number of variables being analysed. The first component can be expected for a fairly large amount of total variance. A total of 5 PC components were generated in this study. PC I contributed the maximum variability of 40% followed by PC II (21.68%) contributed 62.45% of cumulative total variations from the sesame gene pool studied. However, the remaining three components *viz*., PC III, PC IV and PC V contributed 37.55 % of diversity in the gene pool. The biplot diagram explains the distribution and nature of diversity for variables and genotypes among the sesame germplasm were distributed in different groups, which clearly showed genetic diversity among the germplasm accessions (Fig. 1 to 3). Among traits, days to fifty percent flowering and days to maturity had high and negative loadings in PC I. The positive influence of the number of capsules per plant on the first principle component was also recorded by Iqbal *et al*. (2016) and Sasipriya *et al*. (2023). In the case of PC II, traits plant height and number of capsules per plant had high negative and positive loadings. Hence these traits are important in contributing more variation in the respective PCs (Table 2). Based on the vector length of characters, it can be inferred that days to fifty percent flowering, days to maturity and plant height had more variation than other traits (Fig. 1). The trait seed yield per plant had relatively less variation among genotypes. Among characters, days to flowering and maturity had a positive association with seed yield per plant. Plant height had a positive association with seed yield. Hence these traits may be considered as selection indices for the improvement of seed yield.

Using PCA, a scree plot diagram was depicted to determine the optimum number of components by

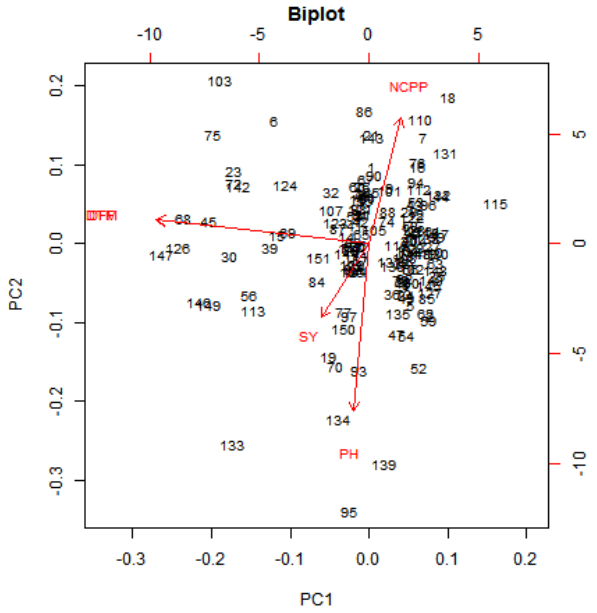

Fig. 1. Distribution of sesame germplasm for first two components

Table 2. Eigen value and percent of total variation and component matrix for the principal component axes

Principal Components	1	2	3	4	5
Days to Fifty percent flowering	-0.6935	**0.1071**	-0.0454	-0.0750	0.7070
Days to maturity	-0.6935	**0.1058**	-0.0467	-0.0737	-0.7070
Plant height (cm)	-0.0505	-0.7451	**0.1004**	-0.6574	0.0000
No. of capsules per plant	**0.1053**	**0.5623**	**0.6060**	-0.5528	-0.0016
Seed yield per plant (g)	-0.1560	-0.3257	**0.7864**	**0.5012**	0.0001
Percent of variance	0.4077	0.2168	0.1982	0.1772	0.0001
Cumulative %	0.4077	0.6245	0.8227	0.9999	1.0000
Eigen values	2.0383	1.0842	0.9910	0.8859	0.0005

drawing a graph between eigenvalues and principal components which explains the variability available in the gene pool (Fig. 2). The cumulative percent of the variance of 62.45% was contributed by PC I and PCII by means of expressing the eigenvalues of greater than 1.0 indicating that these two factors were optimum to assess the variability present in the agronomic traits. Data were considered in each component with Eigen values > 1 as per the suggestions given by Brejda *et al*. (2000), which determines a minimum of 10 % of the variation. Superior eigenvalues are considered as best attributes in principal components.

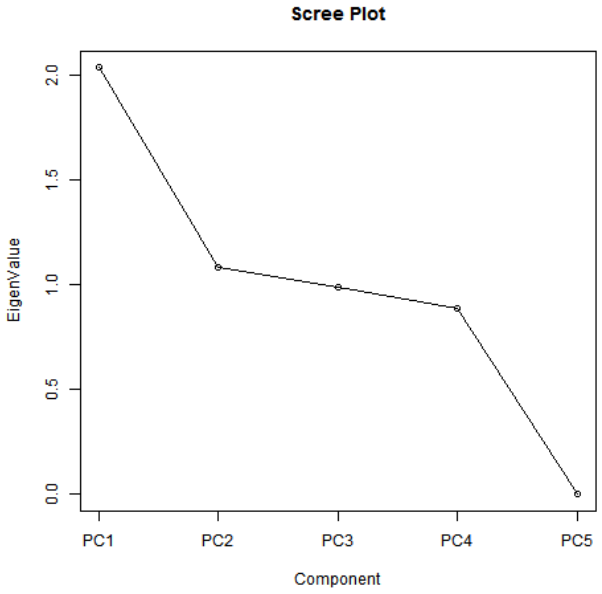

Fig. 2. Scree plot diagram depicted the five components of sesame germplasm

Sesame germplasm accessions were grouped into eight distinct clusters using the agglomerative clustering method. The distance between the clusters occurs between 0.390 and 8.79, as shown in Fig. 3. Among the clusters, cluster I had the highest number of members (58) followed by cluster II (50 genotypes), cluster IV (15 genotypes) and cluster V (11 genotypes). Similar results were obtained in PCA also exhibiting 62.45% of total variability in the first two principal components. Kavithamani *et al*., (2019) also observed grouping of genotypes by a combination of both PCA and clustering provides similar results of variability that existed in the available gene pool. Cluster III and VI accommodated each seven genotypes. However, Cluster VII holds only one genotype and cluster VIII had two genotypes namely TNAU 12 and VS 486-4. Cluster I had the desirable mean values for days to fifty percent flowering and cluster III for the number of capsules per plant and seed yield per plant. Grouping of genotypes into different clusters indicates the presence of diversity among the germplasm studied. The genotypes belonging to these clusters can be utilized for sesame crop improvement.

Knowledge of the extent of genetic diversity present in the available germplasm will pave the

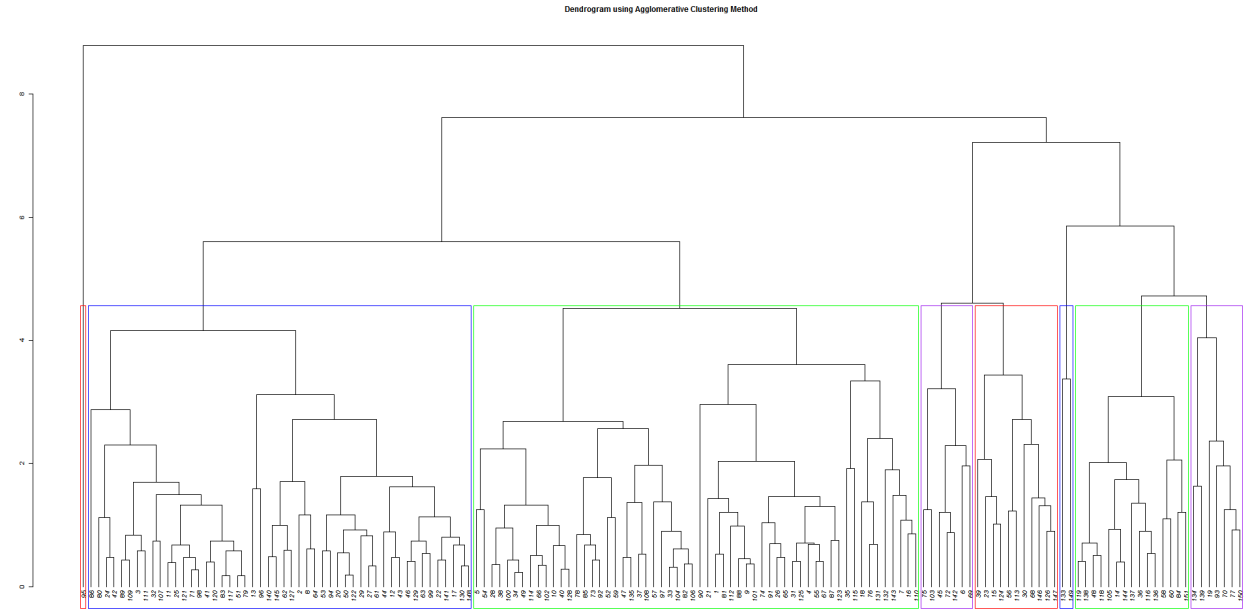

Fig. 3. Dendrogram using Agglomerative Clustering Method

way for the selection of parents and the selection of breeding methodology for the sesame crop improvement programme. The information on genetic diversity is essential for shaping breeding strategies, heterotic grouping and predicting future hybrid performance (Acquaah, 2012). The genotypes *viz*., SO 182, TN 8454, TN 8452, TNAU 12, VS 486-4 and TNAU 64 were found to be superior performance for seed yield trait. These germplasm accessions can be utilized in the development of high-yielding sesame varieties through hybridization followed by selection. From the cluster analysis, Cluster I had desirable mean values for days to fifty flowering. Similarly, cluster III for the number of capsules per plant and seed yield per plant. The genotypes belonging to these clusters can be utilized for the sesame crop improvement programs.

References

Acquaah, G. 2012. *Principles of Plant Genetics and Breeding*. Second edition. Wiley Blackwell publications.; pp. 127-129.

Brejda, J.J., Moorman, T.B., Karlen, D.L. and Dao, T.H. 2000. Identification of regional soil quality factors and indicators. I. Central and Southern High- Plains. *Soil Sci. Soc. Am. J.* 64: 2115-2124.

Crossa, J. 1990. Statistical analyses of Multilocation trials. *Advances in Agronomy, Elsevier.* 44: 55-85

Daisy Myint., Syed A. Gilani, Makoto Kawase and Kazuo N. Watanabe, 2020. Sustainable Sesame (*Sesamum indicum* L.) Production through Improved Technology: An Overview of Production, Challenges, and Opportunities in Myanmar. *Sustainabili*ty. 12: 3515.

Durge, B.D., Geethanjali, S. and Sasikala, R. 2022. Assessment of genetic variability for seed yield and its components in sesame *Sesamum indicum* L. based on multivariate analysis. *Electronic Journal of Plant Breeding*. 13 (3): 974-982. https://www.ejplantbreeding.org/index.php/EJPB/article/view/4404

Iqbal, A., Ahktar, R., Begum, T. and Dasgupta, T. 2016. Genetic estimates and diversity study in Sesame (*Sesamum indicum* L.). *IOSR Journal of Agriculture and Veterinary Science.* 9(8): 1-5.

Kavithamani, D., Yuvaraja, A. and Selvi, B. 2019. Principal component analysis and grouping of sorghum (*Sorghum bicolor* L. Moench) gene pool for genetic diversity. *Electronic Journal of Plant Breeding*. 10(4): 1426-1434 DOI:10.5958/0975-928X.2019.00182.0

Oyelola, B.A. 2004. *The Nigerian Statistical Association Preconference Workshop*; 2004 Sep 20–21; University of Ibadan.

Richman, M. 1986. Rotation of principal components. *J. Climatol*. 6: 293-335.

Sasipriya, S., Parimala,K., Eswari, K.B. and Balram, M. 2023. Principal component analysis and metroglyph of variation among sesame *Sesamum indicum* L. genotypes. *Electronic Journal of Plant Breeding.* 13 (4): 1234-1242. https://www.ejplantbreeding.org/index.php/EJPB/article/view/4433

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