



Principal Component Analysis for Yield and Yield Related Traits in Sesame (*Sesamum indicum* L.)

Mukthambica K.^{1*}, Bisen R.² and Ramya K.T.³

¹Post graduate, Department of Genetics and Plant Breeding,

Jawaharlal Nehru Krishi Vishwavidyalaya, Jabalpur (Madhya Pradesh), India.

²Principal Scientist, AICRP, Project Coordinating Unit (Sesame and Niger),

Jawaharlal Nehru Krishi Vishwavidyalaya, Jabalpur (Madhya Pradesh), India.

³Scientist (Plant Breeding), Crop Improvement Division,

Indian Institute of Oilseeds Research, Rajendranagar, Hyderabad (Telangana), India.

(Corresponding author: Mukthambica K. *)

(Received: 14 January 2023; Revised: 18 February 2023; Accepted: 20 February 2023; Published: 22 March 2023)

(Published by Research Trend)

ABSTRACT: Sesame in India have been reported to have a wide variability for the characters but still no systemic efforts have been made to characterize and identify the genotypes having variable characters for the selection. The research was conducted using seventy genotypes of sesame based on yield and its contributing traits and also using Principal Component Analysis. The observations were recorded by selecting five random plants for fourteen quantitative characters. On the basis of Principal component analysis, out of fourteen components, only 4 principal components (PCs) exhibited more than 1.00 Eigen value and showed about 68.6 % variability among the traits studied. The PC1 had the highest variability (38.3%) followed by PC2 (12.2%), PC3 (10.2%) and PC4 (7.9%) for traits under study. Rotated component matrix revealed that the first principal component (PC1) was mostly related with traits such as days to 50% flowering, days to maturity, length to first capsule, plant height, secondary branches per plant and seed yield per plant. The second principal component (PC2) was related to the traits viz., capsule number per plant, number of leaf axils in main stem, capsule length, test weight and oil content while PC3 was consisting of traits viz., days to emergence, days to flower session. Fourth principal component was related to primary branches per plant. The genotypes like, Paiyur, VRI-3, TMV-6, TMV-3, DS-5, PKDS-8, Rajeshwari and JLT-408 were identified as putative genotypes and length to first capsule, plant height, primary branches per plant, secondary branches per plant, seed yield per plant are identified as main yield trait attributes. Thus, it can be utilized to select the more diverse germplasms for these traits and could be used as parents in heterosis breeding programmes.

Keywords: Sesame, Principal Component Analysis, variability, eigen values.

INTRODUCTION

Sesame (*Sesamum indicum* L.) is the one of ancient oilseed crop known to the mankind. Sesame is originated and domesticated in India and is widely grown in tropical and sub-tropical regions (Bedigian, 2003). In Hindu legend, the sesame seeds are the symbol of immortality. Sesame, commonly called as 'Til' is also known as 'benni seed' and 'gingelly' and falls in the order Tubiflorae and family Pedaliaceae. Sesame is diploid with chromosome number $2n=26$. Sesame seed has excellent nutritional properties and oil has antioxidant properties and long shelf life. So, sesame is considered as queen of oil seeds. Sesame ranks sixth in production among oilseeds i.e., after soybean, groundnut, cotton seed, sunflower, linseed and rapeseed globally. In India, sesame ranks fourth in area under cultivation after soybean, rapeseed-mustard and groundnut. International demand for sesame is increasing across the globe and increased to nearly 80%

since past 2000. During 2019, Myanmar is the leading country in production followed by India (FAOSTAT, 2020). In India, among oilseed crops, sesame ranks fourth in area (1,730,000 ha) with a production of 7,46,346 metric tons and productivity of 431 kg/ha. India has a large area under cultivation, but the total production and productivity is quite low (431 kg/ha) as compared to the world's average (512 kg/ha) (Myint *et al.*, 2020). This is because sesame is one of the most neglected oilseed crop grown on marginal lands under poor management. Further, poor yield is due to non-availability of improved cultivars for diverse agro-climatic condition and poor seed supply system. Sesame in India have been reported to have a wide variability for the characters but still no systemic efforts have been made to characterize and identify the genotypes having variable characters for the selection. Plant breeders frequently measure a vast number of traits, some of which may lack adequate discriminating power for germplasm assessment, characterization, and

management (Upadhyay *et al.*, 2022). In this situation, Principal Component Analysis (PCA) can be used to uncover patterns in data sets and reduce redundancy. The present investigation was carried out to identify the most variable Principal components which may be given due importance in selecting the genotypes for the characters. Principle component analysis (PCA) is a multivariate analysis which helps in data transformation and reduction. This technique has been widely used for studying genetic variability in sesame (Teklu *et al.*, 2021; Durge *et al.*, 2022).

MATERIAL AND METHODS

The field experiment was conducted during Summer, 2021 at Research Farm, Indian Institute of Oilseeds Research (IIOR), Narkhoda, Hyderabad, Telangana. The plant material used in the experiment comprised of 70 sesame genotypes, all of which are released varieties of India except IIOS-1101 (pre-released variety), were collected from IIOR and AICRP, Sesame and Niger. The experimental design adopted was Randomized Complete Block design with two replications. Observations of fourteen quantitative characters were recorded on the basis of five random competitive plants for the evaluation of yield and yield contributing traits of proposed genotypes. The fourteen quantitative characters used for present study are days to emergence, days to 50% flowering, days to flower session, days to maturity, length to first capsule, plant height (cm), primary branches per plant, secondary

branches per plant, capsule number per plant, number of leaf axils in main stem, capsule length (cm), test weight (g), oil content (%) and seed yield per plant (g). Principal Component Analysis (PCA) is well-known method of dimension reduction that can be used to reduce a large set of variables to a small set which still contains most of the information present in large set of variables and calculated according to the method given by Massey (1965); Jolliffe (1986). PCA results in a smaller number of linear uncorrelated variables from a large number of correlated variables. These linear variables are known as principle components and are essentially used as criteria for plant selection.

RESULTS AND DISCUSSION

In the present investigation, Principal Component Analysis (PCA) was carried out and eigen values, % variance and cumulative eigen values are mentioned in Table 1. Out of fourteen components, only 4 principal components (PCs) exhibited more than 1.00 Eigen value and showed about 68.6 % variability among the traits studied. From the fifth component onwards, the values of each successive component have eigen values less than 1 which accounts smaller and smaller amounts of total variance so, they are of little use in selecting the characters. The PC1 had the highest variability (38.3%) followed by PC2 (12.2%), PC3 (10.2%) and PC4 (7.9%) for traits under study. Selection of lines for characters present in four principal components is desirable.

Table 1: Eigen values, percentage of total variation and cumulative percentage for corresponding fourteen traits in sesame genotypes.

Principal component (PC)	Eigen value	Variability (%)	Cumulative %
PC 1	5.36	38.3	38.3
PC 2	1.71	12.2	50.5
PC 3	1.43	10.2	60.7
PC 4	1.11	7.9	68.6
PC 5	0.87	6.2	74.8
PC 6	0.79	5.6	80.4
PC 7	0.64	4.6	85
PC 8	0.56	4	89
PC 9	0.46	3.3	92.3
PC 10	0.36	2.6	94.9
PC 11	0.34	2.4	97.3
PC 12	0.22	1.6	98.9
PC 13	0.14	1	99.9
PC 14	0.01	0.1	100

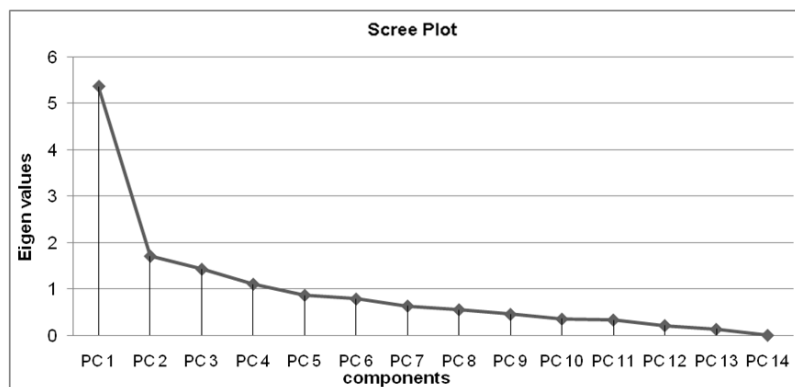


Fig. 1. Scree plot of sesame genotypes between Eigen values and principal components.

Scree plot was laid out between eigen value and principal components (Fig. 1). From the fifth component onwards, the line is almost flat, means that each successive component is accounting for smaller and smaller amounts of total variance and these are the

components having eigen values less than 1 so, they are of little use in selecting the characters. Selection of lines for characters present in four principal components is desirable.

Table 2: Principal Components for 14 yield contributing traits of sesame genotypes.

Traits	Principal Components			
	PC1	PC2	PC3	PC4
Days to emergence	-0.04	-0.03	0.57	0.5
Days to 50% flowering	0.35	-0.06	0.19	0.26
Days to flower session	0.32	-0.03	0.37	0.06
Days to maturity	0.29	-0.20	-0.18	-0.02
Length to first capsule	0.36	0.18	-0.13	0.11
Plant height	0.3	0.05	0.08	0.01
Primary branches per plant	0.29	-0.16	-0.12	0.37
Secondary branches per plant	0.28	-0.02	-0.37	0.07
Capsule number per plant	0.32	0.39	0.1	-0.21
Number of leaf axils in main stem	0.32	0.35	0.12	-0.2
Capsule length	-0.13	0.39	0.36	-0.33
Test weight	-0.14	0.47	-0.17	0.35
Oil content	-0.13	0.5	-0.29	0.36
Seed yield per plant	0.21	-0.04	-0.17	-0.28

Extraction method. Rotated component matrix revealed that the first principal component (PC1) which accounted for the highest variation (38.3%) was mostly related with traits such as days to 50% flowering, days to maturity, length to first capsule, plant height, secondary branches per plant and seed yield per plant (Table 2).

The second principal component (PC2) was related to the traits viz., capsule number per plant, number of leaf axils in main stem, capsule length, test weight and oil

content while PC3 was consisting of traits viz., days to emergence, days to flower session. Fourth principal component was related to primary branches per plant. Rotated component matrix revealed that first four PCs are representing maximum variability (68.6%) hence, the traits falling in these four PCs may be given due importance in sesame breeding. The top principal component scores (PC score) for all the traits were estimated in four principal components and are presented in Table 3.

Table 3: Interpretation of rotated component matrix for the traits having maximum values in each PCs.

Characters	PC 1	PC 2	PC 3	PC 4
	Days to 50% flowering	Capsule number per plant	Days to emergence	Primary branches per plant
	Days to maturity	Number of leaf axils in main stem	Days to flower session	
	Length to first capsule	Capsule length		
	Plant height	Test weight		
	Secondary branches per plant	Oil content		
	Seed yield per plant			

PC scores of genotypes. The PC scores of the each component (PC1, PC2, PC3, PC4) had positive and negative values (Table 4). These scores can be utilized to propose precise selection indices whose intensity can be decided by variability explained by each of principal component. High PC score for a particular genotype in a particular component denotes high values for the variables in that particular genotype.

VRI-3 had the highest PC score followed by Paiyur, TMV-6, TMV-3, Usha, Thilothama, N-32, Nirmala, TMV-7 and TMV-4 in PC1 indicating that these genotypes possess high variability for traits viz., days to 50% flowering, days to maturity, length to first capsule, plant height, secondary branches per plant and seed yield per plant which are the main yield trait attributes.

The highest PC score of Paiyur followed by VRI-3, TMV-6, Shubra, TMV-3, TSS-6, Tarun and GT-3 in PC2 was mainly related with capsule number per plant, number of leaf axils in main stem, capsule length, test weight and oil content

The highest PC score was obtained by Prachi followed by TKG-306, IIOS-1101, Nirmala, TMV-7, Usha, RT-351, Uma, T-78 and RT-125 in PC3 for characters namely days to emergence and days to flower session.

In PC4, high PC scores were recorded for characters viz., primary branches per plant by the genotypes Thilothama, VRI-1, E-8, TKG-306, Swetha til, Thilak, TKG-22, TKG-55 and VRI-2.

Table 4: PC scores of sesame genotypes.

Sr. No.	Genotypes	PC 1	PC 2	PC 3	PC 4
1.	AKT-101	-0.25	0.28	-1.7	-1.6
2.	AKT-64	-0.79	0.63	-0.44	-1.84
3.	Amrit	-0.73	0.09	0.57	-2.45
4.	B-67	0.96	-1.15	0.04	-1.07
5.	CUMS-17	0.58	-0.85	-1.46	-1.08
6.	Chandana	-0.21	-0.5	-0.09	-1.11
7.	DS-5	-0.27	0.02	-1.25	0.93
8.	DSS-9	0.1	0.76	0.44	0.67
9.	E-8	-0.17	-1.09	-0.35	1.69
10.	GJT-5	-1.07	0.8	-0.21	-0.25
11.	GT-1	-1.07	0.53	0.24	0.39
12.	GT-10	0.77	-0.12	-2.31	-1.04
13.	GT-2	-0.65	-0.07	-2.03	0.57
14.	GT-3	-0.67	1.02	-1.02	-1.28
15.	GT-4	-0.56	0.51	-1.16	-0.45
16.	HT-1	-1.39	-0.09	-0.55	-0.87
17.	HT-2	-1.22	0.54	0.29	0.14
18.	Hima	0.01	0.08	-1.08	0.13
19.	IIOS-1101	0.97	-0.39	1.74	0.11
20.	JLT-408	-0.65	0.69	-1.11	0.81
21.	JLT-7	-0.27	0.28	0.45	-0.23
22.	JTS-8	-0.85	0.36	-0.87	-0.39
23.	Kanak	-0.27	-1.34	-1.85	-1.09
24.	Krishna	0.43	-2.8	0.37	-1.6
25.	N-32	1.34	-0.15	0.68	-2.2
26.	Nirmala	1.33	-0.92	1.68	-1.25
27.	PKDS-11	0.75	-0.13	-1.07	0.1
28.	PKDS-8	-0.12	0.06	-0.35	0.38
29.	Phule Til-1	-1.44	0.02	-0.66	-0.82
30.	Paiyur	2.69	2.45	-1.73	-0.91
31.	Prachi	0.49	-2.29	2.49	-0.34
32.	Pragati	-1.04	-0.18	0.32	0.02
33.	Punjab til-1	-0.9	-0.15	-0.05	0.67
34.	Punjab til-2	-0.76	0.78	0.74	0.13
35.	RT-103	-1.16	-0.36	0.77	0.94
36.	RT-125	-1.19	0.23	1.01	0.66
37.	RT-127	-1.09	0.002	0.42	0.43
38.	RT-346	-0.98	0.22	0.79	0.87
39.	RT-351	-0.76	0.96	1.2	0.21
40.	RT-372	-0.87	0.47	0.21	0.3
41.	RT-46	-1.02	-0.62	0.31	-0.24
42.	Rajeshwari	-0.12	0.48	-0.42	0.18
43.	Rama	0.26	-0.45	0.57	-0.96
44.	Savitri	0.37	-1.21	-0.85	0.6
45.	Shubra	-0.8	1.6	0.69	-0.17
46.	Smarak	-0.59	-0.19	-0.16	-0.1
47.	Swetha til	-0.57	0.75	0.75	1.29
48.	T-78	-0.78	0.25	1.02	-0.26
49.	TKG-22	-0.08	0.54	0.66	1.27
50.	TKG-306	-0.4	0.38	1.95	1.37
51.	TKG-308	-0.95	-0.99	0.11	-0.79
52.	TKG-55	-0.85	0.68	0.05	1.08
53.	TMV-3	1.98	1.15	-0.31	0.38
54.	TMV-4	1.09	-0.77	-0.11	0.85
55.	TMV-6	2.08	2.36	0.98	0.2
56.	TMV-7	1.16	0.02	1.56	-0.48
57.	TSS-6	-0.11	1.14	0.81	0.33
58.	Tarun	-1.03	1.14	0.42	-0.89
59.	Thilak	0.59	-1.12	-0.4	1.29
60.	Thilothama	1.45	-1.65	-1.32	2.61
61.	Thilarani	0.67	-0.93	0.61	-0.36
62.	Uma	0.83	-2.35	1.09	0.71
63.	Usha	1.75	0.61	1.51	-1.70
64.	VRI-1	0.73	-0.56	-1.28	2.06
65.	VRI-2	0.81	-0.15	0.0004	1.02
66.	VRI-3	2.9	2.43	0.16	0.76
67.	Vinayak	0.5	-1.47	-0.98	0.001
68.	YLM-11	0.16	0.54	-0.37	0.52
69.	YLM-17	0.03	-0.52	-0.22	0.93
70.	YLM-66	0.91	-0.3	0.09	0.19

Table 5: List of selected genotypes in each principal component.

Sr. No.	PC 1	PC 2	PC 3	PC 4
1.	VRI-3	Paiyur	Prachi	Thilothama
2.	Paiyur	VRI-3	TKG-306	VRI-1
3.	TMV-6	TMV-6	IIOS-1101	E-8
4.	TMV-3	Shubra	Nirmala	TKG-306
5.	Usha	TMV-3	TMV-7	Swetha til
6.	Thilothama	TSS-6	Usha	Thilak
7.	N-32	Tarun	RT-351	TKG-22
8.	Nirmala	GT-3	Uma	TKG-55
9.	TMV-7		T-78	VRI-2
10.	TMV-4		RT-125	

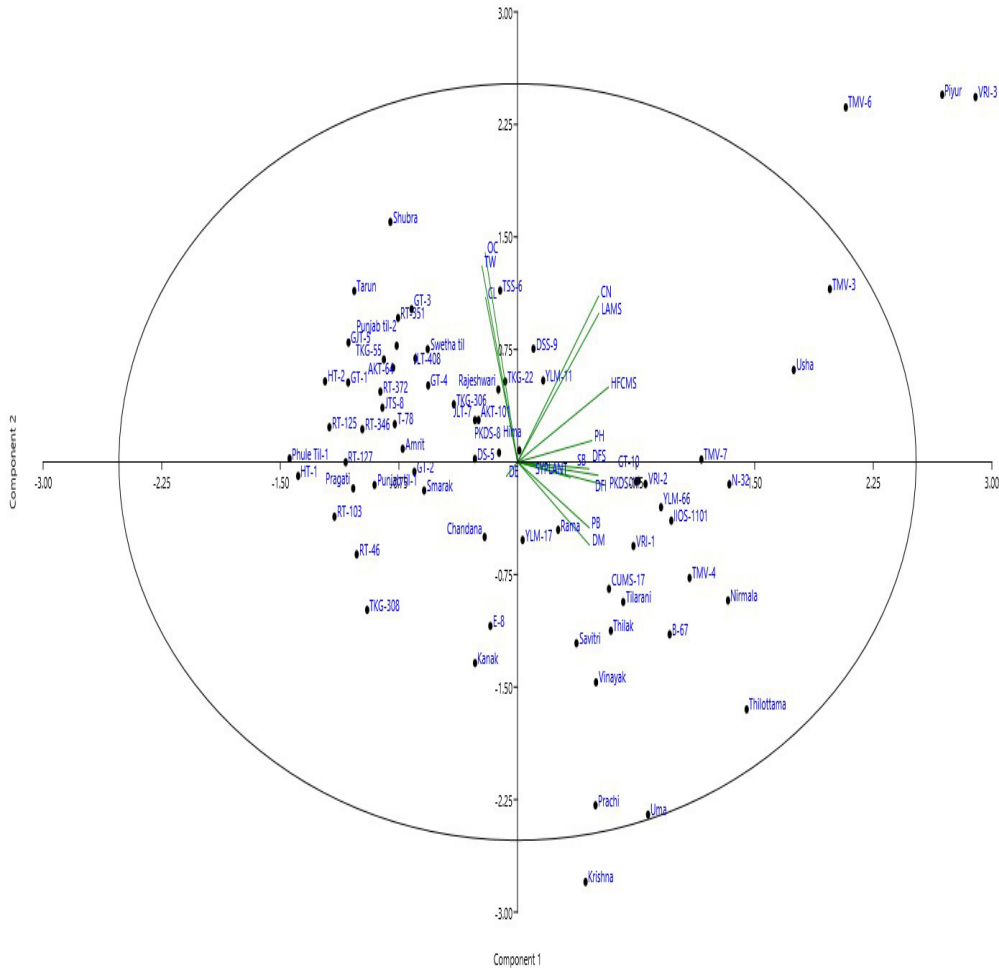


Fig. 2. Two dimensional biplot of Principal Component Analysis.

Germplasm lines having maximum positive PC scores and common in PC1, PC2 for days to 50% flowering, days to maturity, length to first capsule, plant height, secondary branches per plant, seed yield per plant, capsule number per plant, number of leaf axils in main stem, capsule length, test weight, oil content were Paiyur, VRI-3, TMV-6, TMV-3. Selection of these lines can help in development of high yielding cultivars as these lines have high variability. Genotypes showing positive PC scores in PC2, PC4 for capsule number per plant, number of leaf axils in main stem, capsule length, test weight, oil content and primary branches per plant and low PC scores in PC1 and PC3 for phenological traits days to 50% flowering, days to maturity, days to emergence and days to flower session are DS-5, PKDS-

Mukthambica et al., Biological Forum – An International Journal 15(3): 227-232(2023)

8, Rajeshwari and JLT-408 and selection of these lines are used for developing early maturing varieties. On the basis of Principal component analysis, Paiyur, VRI-3, TMV-6, TMV-3, DS-5, PKDS-8, Rajeshwari and JLT-408 were identified as putative genotypes and length to first capsule, plant height, primary branches per plant, secondary branches per plant, capsule number per plant, number of leaf axils in main stem, seed yield per plant are identified as main yield trait attributes. In sesame breeding programmes, selection for these genotypes and traits may be utilized in hybridisation programme to develop improved genotypes in future. Similarly, Furat and Uzun (2010) evaluated sesame genotypes using 21 morphological and agronomical descriptors and observed that seven PC explained

69.9% of total multivariate variation. The characters such as days to emergence, flowering and capsule initiation and seed yield contributed significantly towards genetic diversity; Akbar *et al.* (2011) revealed that four PC explains 63.64% and revealed that plant height, days to maturity, seed yield per plant, capsule number per plant are the major determinants of genetic diversity; Ahadu (2012) evaluated sixty four sesame genotypes of Ethiopia and reported that biomass yield, capsule filling period, days to maturity, seed yield per plot had contributed to variability and also present in first principle component (PC1); Baraki *et al.* (2015) evaluated 13 sesame genotypes and reported that out of eight, first three PCs accounted for 88.49% of the total variance and these three PCs were considered as significant; Tanwar and Bisen (2018) have reported that nine PCs exhibited more than 1 eigen value and showed 95.19 % total variability among the characters; Baraki and Berhe (2019) reported for seed yield, primary branches per plant and secondary branches per plant and plant height.

CONCLUSIONS

PCA is an important technique for enhancing the breeding programme as it extracts all the key components and highlights their contribution to total variability. PCA biplot revealed the high performing genotypes viz., Paiyur, VRI-3, TMV-6, TMV-3, DS-5, PKDS-8, Rajeshwari and JLT-408 can be effectively utilized for crop improvement programmes.

FUTURE SCOPE

Principal Component Analysis (PCA) helps to speed up crop enhancement by sorting the key components and their contribution to variation. Selection of genotypes which showing high contribution to yield can be made which could be utilized as parents in future breeding programmes.

Acknowledgments. The authors are thankful to all members involved in the entire field operations. We are also thankful to the ICAR-IIOR crop improvement laboratory, Rajendranagar, Hyderabad; JNKVV; AICRP-Sesame and Niger, PC Unit, Jabalpur for providing experimental material.

Conflicts of Interest. None.

REFERENCES

Ahadu, M. (2012). Phenotypic variability, divergence analysis and heritability of characters in sesame (*Sesamum indicum* L.) genotypes. *Nature and science*, 10(10), 117-126.

- Akbar, F., Rabbani, M. A., Shinwari, Z. K. and Khan, S. J. (2011). Genetic divergence in sesame (*Sesamum indicum* L.) landraces based on qualitative and quantitative traits. *Pakistan Journal of Botany*, 43(6), 2737-2744.
- Baraki, F. and Berhe, M. (2019). Evaluating performance of sesame (*Sesamum indicum* L.) genotypes in different growing seasons in northern Ethiopia. *International Journal of Agronomy*, 7, 1-7.
- Baraki, F., Tsehaye, Y. and Abay, F. (2015). Grain yield based cluster analysis and correlation of agronomic traits of sesame (*Sesamum indicum* L.) genotypes in Ethiopia. *Journal of Natural Sciences Research*, 5(9), 11-17.
- Bedigian, D. (2003). Evolution of sesame revisited: domestication, diversity and prospects. *Genetic Resources and Crop Evolution*, 50, 779-787.
- 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.
- FAOSTAT (2020). Food and Agricultural Organisation. QCL Crops and livestock products 2020. Available from: <https://www.fao.org/faostat/en/#data/QCL>
- Furat, S. and Uzun, B. (2010). The use of agro-morphological characters for the assessment of genetic diversity in sesame (*Sesamum indicum* L.). *Plant Omics Journal*, 3(3), 85-91.
- Jolliffe, I. T. (1986). *Principal Component Analysis (second edition)*. Springer-Verlag, New York Inc. pp.487.
- Massey, W. F. (1965). Principal components regression in exploratory statistical research. *Journal of the American Statistical Association*, 60, 234-246.
- Myint, D., Gilani, S. A., Kawase, M. and Watanabe, K. N. (2020). Sustainable sesame (*Sesamum indicum* L.) production through improved technology: An overview of production, challenges and opportunities in Myanmar. *Sustainability (MDPI)*, 12, 3515.
- Tanwar, A. and Bisen, R. (2018). Genetic diversity analysis in sesame (*Sesamum indicum* L.) germplasm based on morphological and quality traits. *Electronic Journal of Plant Breeding*, 9(1), 9-17.
- 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.
- Upadhyay, S., Rathi, S., Choudhary, M., Snehi, S., Singh, V., Singh, P. K. and Singh, R. K. (2022). Principal Component Analysis of Yield and its Attributing Traits in Advanced Inbred Lines of Rice under Sodicity condition (*Oryza sativa* L.). *Biological Forum – An International Journal*, 14(2), 1273-1276.

How to cite this article: Mukhthambica K., Bisen R. and Ramya K.T. (2023). Principal Component Analysis for Yield and Yield Related Traits in Sesame (*Sesamum indicum* L.). *Biological Forum – An International Journal*, 15(3): 227-232.