Unraveling Diversity and Character Association in Sesame (*Sesamum indicum* L.) using Different Agro-Morphological Traits

Manoj Kumar Patel, Deepti Tiwari, Versha Sharma, Dhirendra Singh, Shahil Kumar

ABSTRACT

Sesame is an important oilseed crop, carrying several beneficial properties to human kind. Given its wide-ranging applications, there is a constant need for enhancing the seed yield. Evaluation of genetic diversity and character association among various seed yield components is considered as the fundamental steps in any breeding program. Thus, the present study has been focused to investigate the diversity and character associations of twenty different sesame genotypes, using different agro-morphological traits. The diversity analysis conducted in this study, unveiled the presence of two distinct clusters. The first cluster encompassed three genotypes, while the remaining seventeen genotypes belonged to the second cluster. Furthermore, character associations were explored through correlation analysis, and subsequently partitioned into direct and indirect effects through path analysis. Correlation study revealed that seed yield per plant (SYP) was positively correlated with the capsules per plant (CPP), biological yield per plant (BYP) and harvest index (HI), while negative correlation was observed with days to 50% flowering (DFF) and days to maturity (DM). Through path analysis, we found that positive direct effects on SYP were observed for BYP, HI, DM, plant height (PH) and test weight (TW). Conversely, negative direct effects on SYP were noted for CPP, BPP and DFF. These findings highlighted the importance of focusing on BYP, HI, and CPP during the selection process, as these factors were identified as important components for seed yield. The insights gained from this study will not only aid in the selection of suitable parents for hybridization, but also assist in determining the key seed yield components for sesame improvement.

Keywords Cluster, Correlation, Diversity analysis, Path analysis, Sesame.

Manoj Kumar Patel[1]*, Deepti Tiwari[2], Versha Sharma[3], Dhirendra Singh[4], Shahil Kumar[5]

[1]Department of Plant Breeding and Genetics, SKNAU, Jobner 303329, Rajasthan, India

[2]Division of Plant Physiology, ICAR-IARI, New Delhi, India

[3]Department of Genetics and Plant Breeding, Agriculture University, Jodhpur, Rajasthan, India

[4]Ex-Professor and Head, Department of Plant Breeding and Genetics, SKNAU, Jobner, Rajasthan, India

[5]Division of Seed Science and Technology, ICAR-IARI, New Delhi, India

Email : patel.manoj39@gmail.com
*Corresponding author

INTRODUCTION

Sesame (*Sesamum indicum* L.) is an important oilseed crop, which contains important nutrients and carries several medicinal utilities (Patel *et al.* 2023a). Botanically, it is a diploid, self-pollinated species having chromosome number 2n=26, belongs to family Pedaliaceae. It is an oldest oilseed crop and grown in different parts of the globe including

Africa, America and Asia. Sesame is considered as third most important oilseed crop after groundnut and rapeseed-mustard. In the year 2022, the overall sesame production in India was 0.26 million tonnes (mt) in total area of 1.35 million hectares (mha) with productivity of 2.39 quintal/hectare (Anonymous 2022). In India, it is mainly grown in Uttar Pradesh, Madhya Pradesh, Rajasthan and Gujarat, which cumulatively contribute to 84% of total area and production (Anonymous 2022).

The success of any breeding program is determined by the amount of genetic diversity or variability available in any crop species. It is a pre-requisite component to initiate any crop improvement program (Ola *et al.* 2023, Patel *et al.* 2023b). Moreover, diversity analysis is an important method to study genetic similarity and dissimilarity among genotypes that eventually help in selecting parents for hybrid breeding programs (Harisha *et al.* 2021). Different kinds of diversity can be estimated either by using phenotypic data or genotypic data. In sesame Gebremichael and Parzies (2011) studied genetic divergence using SSR markers, whereas, Tanwar and Bisen (2018) and Swapna *et al.* (2023) carried out D^2 analysis using phenotypic data. Additionally, phenotypic data include many agro-morphological traits and these traits interact among themselves in several ways. Therefore, to study degree and direction of association among different traits correlation analysis is usually used. Furthermore, these traits contribute to seed yield either directly or indirectly, therefore such association can be estimated by using path analysis. In sesame, several researchers have studied different agro-morphological traits and performed path analysis to study the direct and indirect effects (Gnanasekaran *et al.* 2008, Patil and Lokesha 2018, Navaneetha *et al.* 2019). However, their results differ substantially, may be due to different genetic material and varying environmental conditions. Therefore, a breeder should know the interrelationship among the seed yield components and their association with seed yield for effective utilization of genetic material, that eventually leads to effective crop improvement.

Thus, with this background, current study has been carried out with the aim to study genetic diversity among sesame genotypes using hierarchical clustering. In addition, association among different components traits of sesame have also been studied using correlation and path analysis.

MATERIALS AND METHODS

Plant materials

Twenty sesame genotypes were obtained from Agriculture University, Jodhpur, Rajasthan. The material consists of diverse sesame lines having substantial variation for various agro-morphological traits. The details of these genotypes including parentage and associated institute have been presented in the Table 1.

Experimentation

The present experiment was conducted at research farm of SKNAU, Jobner during *kharif* 2018. The

Table 1. Details of twenty sesame genotypes used for present investigation.

Sl. No.	Genotypes	Parentage	Releasing institute
1	RT - 46	T 12 × Punjab TIL no 1	Agricultural Research Station, Mandor, Jodhpur, Rajasthan
2	RT - 103	C-7 × A 6-5	Agricultural Research Station, Mandor, Jodhpur, Rajasthan
3	RT - 125	NA	Agricultural Research station, Mandor, Jodhpur, Rajasthan
4	RT - 127	SI-3500×Patan-64	Agricultural Research Station, Mandor, Jodhpur, Rajasthan
5	RT-346	RT 127×HY 24	Agricultural Research Station, Mandor, Jodhpur, Rajasthan
6	RT - 351	NIC 8409×RT 127	Agricultural Research Station, Mandor, Jodhpur, Rajasthan
7	RT - 372	RMT68×EC362397	Agricultural Research Station, Mandor, Jodhpur, Rajasthan
8	RT-378	NA	NA
9	RT - 383	NA	NA
10	RT - 384	NA	NA
11	RT - 385	NA	NA
12	RMT-425	NA	NA
13	RMT-447	NA	NA
14	RMT-450	NA	NA
15	RMT-479	NA	NA

Table 1. Continued.

Sl. No.	Genotypes	Parentage	Releasing institute
16	RMT-486	NA	NA
17	RMT- 505	NA	NA
18	Pragati	JLT-26 × RT-127	Crop Research Station, Mauranipur, Jhansi (CSAUA and T), Uttar Pradesh
19	TKG - 22	HT-6 ×JLT-3	Zonal Agricultural Research Station, Tikamgarh, Madhya Pradesh
20	GT - 10	Selection from TN AU-17	Agricultural Research Station, (JAU), Amreli, Gujarat

plant material was grown in three replicates using Randomized Complete Block Design (RCBD). Each genotype was grown in two rows of 4 m length with crop spacing 30×10 cm^2. All other package and practices of sesame was performed as per the recommended standards.

Observations recorded

The current research has been focused on ten different agro-morphological traits, namely days to 50% flowering (DFF), days to maturity (DM), plant height (PH, in centimeter), branches per plant (BPP), capsules per plant (CPP), seeds per capsule (SPC), test weight (TW in gram), biological yield per plant (BYP, in gram), harvest index (HI, in percentage) and seed yield per plant (SYP in gram). Observations on DFF and DM were recorded on plot basis, while remaining traits were recorded on randomly selected five individual plants.

Data analysis

Best Linear Unbiased Predictions (BLUPs) values were determined using software metaR (Alvarado *et al.* 2016) by taking genotypes as random effects. These BLUPs values were further used in different analysis such as diversity, correlation and path analysis. Diversity analysis was carried out with the help of R package "stats" as per Ward's method (Ward 1963) using Euclidean distance. To estimate optimum number of clusters, different indices were used and their index value was calculated. The optimum number

of clusters were determined based on the majority rule. Further, based on the output a dendrogram was created using iTOL (Letunic and Bork 2021). Correlation analysis among different agro-morphological traits was performed as per Pearson's correlations and correlation matrix was pictographed using R based package "ggcorrplot". Thereafter, correlation coefficients partitioned into direct and indirect effects using path analysis as per Dewey and Lu (1959).

RESULTS AND DISCUSSION

Optimization of number of clusters and diversity analysis

In the present study, diversity analysis of 20 sesame genotypes has been carried out on ten agro-morphological traits using hierarchical clustering. To optimize number of clusters, 24 different test indices were calculated, out of which 11 indices viz., Krzanowski-Lai (KL), Calinski-Harabasz (CH),

Table 2. Estimation of optimum number of clusters employing different indices.

| Sl. No.	Indices	Number of clusters	Index value
1	Krzanowski-Lai (KL)	2	1.53
2	Calinski-Harabasz (CH)	2	5.82
3	Hartigan	3	1.25
4	Cubic clustering criterion (CCC)	2	-1.63
5	Scott	10	112.05
6	Marriot	3	35182258
7	Trace Cov W	3	109.73
8	Trace W	3	9.48
9	Friedman	10	19867.45
10	Rubin	3	-0.04
11	C-index	4	0.47
12	Davies–Bouldin (DB)	10	0.9
13	Silhouette	2	0.28
14	Duda	2	0.8
15	Pseudo T2	2	3.67
16	Beale	2	1.54
17	Ratkowsky	3	0.35
18	Ball	3	32.04
19	Point Biserial	2	0.63
20	Frey	2	3.24
21	McClain	2	0.26
22	Dunn	2	0.53
23	SD-index	9	0.96
24	SDbw	10	0.28
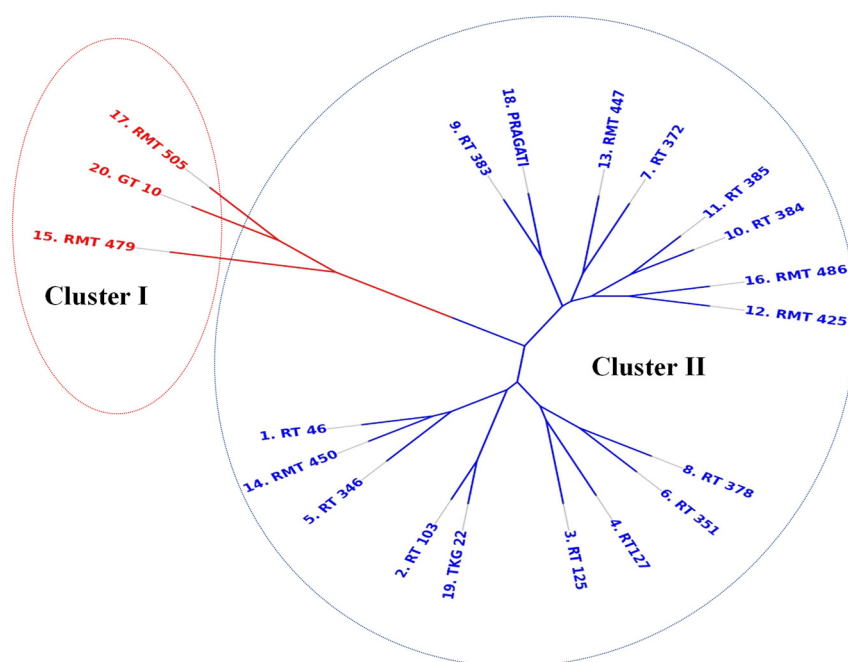
Fig. 1. Tree dendrogram showing two major clusters (cluster I with red color and cluster II with blue color) grouping twenty sesame genotypes into two groups.

Cubic clustering criterion (CCC), Silhouette, Duda, Pseudo T2, Beale, Point Biserial, Frey, McClain and Dunn index suggested 2 optimum clusters, whereas 7 indices namely, Hartigan, Marriot, Trace Cov W, Trace W, Rubin, Ratkowsky and Ball index proposed 3 clusters. Moreover, C and SD-index optimized 4 and 9 clusters, respectively; while remaining 4 indices i.e., Scott, Friedman, Davies–Bouldin (DB) and SDbw index proposed 10 optimum number of clusters (Table 2). Therefore, majority rule was applied in all indices and two clusters were optimized for the present study as maximum number of indices suggested 2 optimum number of clusters. Result indicated that three genotypes viz., GT-10, RMT-505 and RMT-479 into one cluster, while remaining 17 genotypes belonged to another cluster (Fig. 1). Genotypes such as RT - 346, RT - 351 and Pragati have a common parent RT-127 (Table 1), thus they have shown similarity with RT-127 and belonged to same cluster. Moreover, genotypes belonging to the same clusters have shown less genetic divergence, while greater divergence was observed among genotypes belonging to different clusters. The diversity observed in the present study can be attributed to a combination of factors such as the interchange of breeding material, both natural and artificial selection, genetic drift, the flow of genes between populations, and variations due to environmental conditions (Singh and Chaudhary 1977, Bharadwaj *et al.* 2001, Sood *et al.* 2005). Therefore, genotypes between clusters can be used as parents in hybrid breeding programs to exploit maximum heterosis (Harisha *et al.* 2021, Swapna *et al.* 2023).

Correlation studies

Correlation analysis indicates the degree and direction of association between different variables. In the present study, to study association among different agro-morphological traits, correlation study was carried out, in which SYP has shown significant positive correlation with CPP (r=0.61), BYP (r=0.7) and HI (r=0.61), while negative correlation observed with DFF (r= -0.49) and DM (r=-0.63). For remaining traits, viz., PH, BPP, SPC and TW have shown non-significant correlation (Fig. 2). Traits which have shown positive correlation with SYP are considered as important component traits, have play an import-
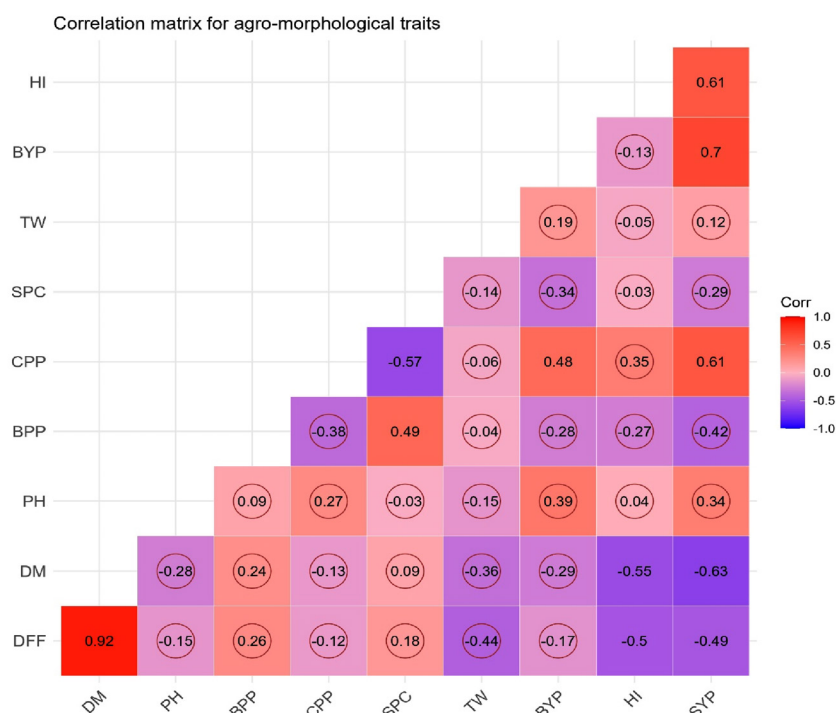
Fig. 2. Correlation matrix showing correlation among different agro-morphological traits. The boxes highlighted with red showing positive correlation, while boxes with blue color exhibiting negative correlation. Moreover, values without a circle showing significant correlation at 5% level of significance, whereas encircled value is showing non-significant correlation.

ant role in improvement of seed yield. Although, usually maturity traits (DFF and DM) show positive correlation with seed yield, but in the present investigation contradictory result was observed as most of the genotypes in the present were early maturing and high yielding varieties. The detailed correlation among different agro-morphological traits have been presented in the Fig. 2. In support to present findings, positive correlation of CPP with SYP was reported by Sankar and Kumar (2003) and Navaneetha *et al.* (2019), whereas negative correlation of DFF with SYP was observed by Patil and Lokesha (2018).

Path analysis

In a breeding program, our primary concern often revolves around enhancing crop yield, which is the cumulative result of various agro-morphological traits. These traits are interrelated and they mutually influence each other's effect on seed yield. Path coefficient analysis partitions the correlation coefficients of

dependent variable with other independent variables into direct and indirect effects (Dewey and Lu 1959). In the present investigation, seed yield per plant (SYP) used as the dependent variable and considered the other yield components as independent variables. Path analysis results revealed that correlation of SYP with other traits is mainly due to direct effects, while the contribution of indirect effects was relatively less (Table 3). We observed positive direct effects on SYP for BYP (0.8096), followed by HI (0.7466), DM (0.0299), PH (0.0284), and TW (0.0087). Conversely, negative direct effects on SYP were noted for CPP (-0.055), BPP (-0.0248), and DFF (-0.0012). Similarly, positive direct effect for TW and negative direct effect for CPP was observed by Navaneetha *et al.* (2019). Traits that demonstrate positive direct effects should be improved through selection to enhance the seed yield. Conversely, traits showing negative direct effects should be minimized to increase seed yield (Patil and Lokesha 2018).

Table 3. Path analysis showing direct (diagonal values in bold) and indirect effects of various agro-morphological traits on seed yield. '*' Significant at 5% level of significance.

Traits	DFF	DM	PH	BPP	CPP	SPC	TW	BYP	HI
DFF	**-0.0012**	-0.0011	0.0002	-0.0003	0.0001	-0.0002	0.0005	0.0002	0.0006
DM	0.0276	**0.0299**	-0.0083	0.0073	-0.004	0.0027	-0.0108	-0.0087	-0.0164
PH	-0.0042	-0.0079	**0.0284**	0.0026	0.0076	-0.0009	-0.0043	0.011	0.0011
BPP	-0.0064	-0.0061	-0.0023	**-0.0248**	0.0094	-0.0123	0.001	0.0069	0.0066
CPP	0.0068	0.0073	-0.0148	0.0208	**-0.055**	0.0315	0.0033	-0.0267	-0.0191
SPC	-0.0014	-0.0007	0.0002	-0.0038	0.0044	**-0.0077**	0.0011	0.0026	0.0003
TW	-0.0038	-0.0031	-0.0013	-0.0004	-0.0005	-0.0012	**0.0087**	0.0017	-0.0004
BYP	-0.1339	-0.2368	0.3133	-0.2264	0.3924	-0.2729	0.1538	**0.8096**	-0.106
HI	-0.3738	-0.4103	0.0279	-0.1981	0.2588	-0.0254	-0.037	-0.0978	**0.7466**
Correlation coefficients	**-0.49***	**-0.63***	**0.34**	**-0.42**	**0.61***	**-0.29**	**0.12**	**0.7***	**0.61***
Residual effect	**0.07**								

On the other hand, we have noticed a network of positive indirect effects on SYP through various components (Table 3). For instance, through DFF, we observed positive indirect effects for DM (0.0276) and CPP (0.0068), while studying DM, only CPP (0.0073) exhibited a positive indirect effect on SYP. Through PH, we noted positive indirect effects for DFF (0.0002), SPC (0.0002), BYP (0.3133) and HI (0.0279), while DM (0.0073), PH (0.0026) and CPP (0.0208) have shown positive indirect effects on SYP through BPP. Additionally, positive indirect effects on seed yield were observed through CPP for DFF (0.0001), PH (0.0076), BPP (0.0094), SPC (0.0044), BYP (0.3924), and HI (0.2588). Moreover, positive indirect effects for DM (0.0027) and CPP (0.0315) was observed while studying SPC, while DFF (0.0005), BPP (0.001), CPP (0.0033), SPC (0.0011), and BYP (0.1538) exhibited the same through TW. Additionally, positive indirect effects on seed yield were noted for DFF (0.0002), PH (0.011), BPP (0.0069), SPC (0.0026), and TW (0.0017) *via* BYP. while, DFF (0.0006), PH (0.0011), BPP (0.0066), and SPC (0.0003) provided the same through HI. For the remaining traits, negative indirect effects on seed yield per plant (SYP) were evident when examined through different components (Table 3). These interconnected observations shed light on the complex dynamics affecting seed yield. Interestingly, in the present study we observed that CPP had shown highest positive indirect effects on seed yield through various components (Table 3). Therefore, CPP should be prioritized in selection to improve seed yield. Similar results were also obtained by Navaneetha

et al. (2019), where, high positive indirect effect was observed for number of capsules on main stem through other components.

In this study, we observed that traits such as BYP and HI have shown high positive direct effects as well as significant positive correlation with the SYP, while DFF have shown negative direct effect and significant negative correlations (Table 3). These results indicated the true association of these traits with seed yield either positive or negative direction (Akintunde 2012). However, CPP have shown significant positive correlation but negative direct effect, this unusual result suggested that the significant correlation with the seed yield is not true, but due to indirect effect of other component traits and these components could be considered in selection (Akintunde 2012). On the other hand, positive direct effect and significant negative correlation as observed in DM, suggested the undesirable indirect effects of other components, which should be excluded (Singh and Chaudhary 1977). The residual effect in the present study was almost negligible with magnitude of 0.07, suggested that components included in the present investigation sufficiently defines the variability of seed yield per plant.

CONCLUSION

In any crop improvement program, the initial steps involve assessing the genetic diversity and identifying important seed yield components. The present study highlighted the importance of these foundational

studies in twenty sesame genotypes using agro-morphological traits. Diversity analysis revealed two optimum clusters and genotypes between clusters can be used as parents in hybrid breeding programs, which will facilitate to develop superior sesame hybrids. Furthermore, correlation and path analyses revealed biological yield per plant (BYP), harvest index (HI), and capsules per plant (CPP) as pivotal components of seed yield. Thus, these findings can be effectively integrated into selection programs. Ultimately, present study provides valuable insights for breeders by assisting them in selecting suitable parent plants for hybrid breeding and determining the key components that contribute to superior seed yield in sesame.

REFERENCES

Akintunde A (2012) Path analysis step by step using excel. *J Technical Sci Technol* 1 (1) : 09—15. https://doi.org/10.31578/. v1i1.29.

Alvarado G, López M, Vargas M, Pacheco Á, Rodríguez F, Burgueño J, Crossa J (2016) META-R (Multi Environment Trial Analysis with R for Windows) Version 6.01. CIMMYT Research Data and Software Repository Network. http://hdl. handle.net/11529/10201

Anonymous (2022) Sesame *kharif* Crop Survey. IOPEPC https://iopepc.org/misc/2021_22/Final%20_Report_Sesame_2022.pdf.

Bharadwaj C, Satyavathi T, Subramanyam D (2001) Evaluation of different classificatory analysis methods in some rice (*Oryza sativa* L.) collections. *Ind J Agric Sci* 71(2):123—125.

Dewey DR, Lu K (1959) A correlation and path coefficient analysis of components of crested wheatgrass seed production. *Agron J* 51 (9) : 515—518. https://doi.org/10.2134/agronj1959.00021962005100090002x.

Gebremichael DE, Parzies HK (2011) Genetic variability among landraces of sesame in Ethiopia. *Afr Crop Sci J* 19(1):1—13. https://doi.org/10.4314/acsj.v19i1.68656.

Gnanasekaran M, Jebaraj S, Muthuramu S (2008) Correlation and path co-efficient analysis in sesame (*Sesamum indicum* L.). *Pl Arch* 8(1):167—169.

Harisha R, Bhadru D, Bhargava K, Vanisri S, Shankar VG, Balakrishnan AP, Gowda MM, Singhal S (2021) Assessment of Variability and Genetic Diversity for Elite Rice (*Oryza sativa* L.) Genotypes of Telangana and Andhra Pradesh. *Int J Environ Clim* 12(11): 3612–3622. https://doi.org/10.9734/ijecc/2022/v12i111409.

Letunic I, Bork P (2021) Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49(W1):W293-296. https://doi.org/10.1093/nar/gkab301.

Navaneetha JS, Murugan E, Parameswari C (2019) Correlation and path analysis for seed yield and its components in sesame (*Sesamum indicum* L.). *Electron J Pl Breed* 10(3):1262—1268. https://doi.org/10.5958/0975-928x.2019.00161.3.

Ola MP, Patel MK, Kumar A, Ram M, Pant NC (2023) Genetic variability for seed yield and its component characters in taramira (*Eruca sativa* Mill.). *Pharma Innov J* 12(3): 3102—3105.

Patel MK, Singh D, Sharma V, Tiwari D, Ola MP, Kumar S, Saroj R (2023a) Stability analysis in sesame (*Sesamum indicum* L.) under different Doses of Fertilizers. *Biol Forum Int J* 15 (2) : 365—370.

Patel MK, Tiwari D, Sharma V, Ola MP, Saroj R, Kumar S, Harisha R, Singh D (2023b) Deciphering the genetic variability for seed yield components in sesame (*Sesamum indicum* L.). *Int J Environ Clim* 13 (10) : 1118—1126. https://doi.org/10.9734/ijecc/2023/v13i102761.

Patil MK, Lokesha R (2018) Estimation of genetic variability, heritability, genetic advance, correlations and path analysis in advanced mutant breeding lines of sesame (*Sesamum indicum* L.). *J Pharmacog Nat Prod* 4:1-5. https://doi.org/10.4172/2472-0992.1000151.

Sankar PD, Kumar CR (2003) Character association and path coefficient analysis in sesame (*Sesamum indicum* L.). *Agricult Sci Digest* 23 (1) : 17—19.

Singh BS, Chaudhary RK (1977) Biometrical methods in quantitative genetic analysis. Kalyani Publishers, New Delhi, pp 20—35.

Sood S, Sood KC, Kumar S (2005) Genetic diversity in rice. *Res Crop* 6 (2) : 290—292.

Swapna B, Thippani S, Padmaja D, Nayak SS (2023) Genetic diversity for yield and yield contributing characters in sesame (*Sesamum indicum* L.). *Int J Environ Clim* 13(8):1717—1725. https://doi.org/10.9734/ijecc/2023/v13i82124.

Tanwar A, Bisen R (2018) Genetic diversity analysis in sesame (*Sesamum indicum* L.) germplasm based on morphological and quality traits. *Electron J Pl Breed* 9(1) : 9—17. https://doi.org/10.5958/0975-928X.2018.00002.9.

Ward Jr JH (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58(301) : 236—244. https://doi.org/10.1080/01621459.1963.10500845.