International Journal of Chemical Studies

Atul Singh
Project Coordinating Unit, Sesame and Niger, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur, Madhya Pradesh, India

Rajani Bisen
Project Coordinating Unit, Sesame and Niger, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur, Madhya Pradesh, India

Akanksha Tiwari
Project Coordinating Unit, Sesame and Niger, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur, Madhya Pradesh, India

Assessing inter: Relationship of sesame genotypes and their traits using cluster analysis and principal component analysis

Atul Singh, Rajani Bisen and Akanksha Tiwari

Abstract
Seventy five sesame genotypes were evaluated to assess the genetic divergence available in the sesame germplasm based on the Mahalanobis distance and cluster analysis for the identification of genetically diverse and agronomically superior accessions which may be further used in the hybridization programme. Clustering pattern indicated that majority of genotypes, i.e. 64 (85%) were genetically close to each other and grouped in 3 clusters, while apparent diversity was mainly noticed due to 11 genotypes (15%) distributed over 10 clusters. Maximum inter cluster divergence was observed between clusters XII and IV (108.84) followed by clusters XIII and IX (107.11) indicating that good recombinants can be realized by mating between the genotypes KMS-5-371, EC-303304-A, KMR-48-A, KMS-5-587, IS-491-A, NIC-16393 and IC-204063 in a definite fashion. The trait days to flower initiation (39.82) contributed maximum to genetic divergence followed by oil content (28.58%) and number of capsules/plant (13.12%). These three characters contributed more than 81% to the total genetic divergence in the genotypes studied. PCA yielded twelve PCs from the twelve agronomic traits of sesame out of which seven PCs showing about 83.193% variability. Results revealed that the genotypes *viz*., S-0644, KMS-5-361, T-1-A, NIC-16214, SI-1125, NIC-8080-A have highest PC values for characters number of capsules/plant, number of seeds/capsule, number of primary branches/plant, number of secondary branches/plant, 1000 seed weight and seed yield/plant. Thus, these genotypes might be used for development of new varieties of sesame.

Keywords: Sesame, Genetic Divergence, Principal Component Analysis, Genotypes, Hybridization

1. Introduction
Sesame is one of the most ancient oilseed crop cultivated and domesticated in the Indian subcontinent during Harappan and Anatolian eras over 4,000 yrs ago. The importance of sesame lies in the quality of the oil, the presence of antioxidants sesamin and sesmolin, its antiquity and use in religious rituals. The climatic and edaphic conditions of India are quite suitable for cultivation of sesame crop. Sesame has important agricultural attributes as it is adapted to tropical and temperate conditions, grows well on stored soil moisture with minimal irrigation or rainfall, can produce good yields under high temperatures, and its seed has a high value. However, the average productivity of sesame is low as compared to other oilseed crops due to the lack of high yielding cultivars. Since sesame has been treated as less input intensive crop, the role of breeding improved varieties has been considered as a promising approach.

Before exploiting a population for trait improvement, it is necessary to understand the magnitude of variability in the population which is fundamental for genetic improvement in all crop species. To develop segregating population, genetic distance estimates form the basis for selecting parental combinations with sufficient genetic diversity. Variability in terms of genetic divergence for agronomic traits is the key component of breeding programmes for broadening the gene pool of sesame. Genetic distance estimates can be estimated by different methods as it is crucial to understand the usable variability existing in the population. One of the approaches is to apply multivariate analysis. Multivariate statistical tools include Principal Component Analysis (PCA), Cluster analysis and discriminate analysis. Principal Component Analysis (PCA) can be used to uncover similarities between variable and classify the cases (genotypes), while cluster analysis on the other hand is concerned with classifying previously unclassified materials. It is particularly useful in characterizing individual accessions, to detect duplications of genetic materials in germplasm collections, and as a general guide in selecting parents for crossing in breeding programs and in developing informative mapping populations

Correspondence
Rajani Bisen
Project Coordinating Unit, Sesame and Niger, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur, Madhya Pradesh, India

for genome mapping (Harlan 1975; Jain *et al*. 1975; Jain, 1977; Arunachalam 1981; Souza and Sorrells, 1989; Smith and Smith, 1989; Mohammadi, 2003) [3, 6, 5, 1, 18, 17, 10].

To boost the yield potential and genetic gain under low or no input condition, it is necessary to explore the existing genetic diversity. A focused research on genetic diversity and breeding of sesame is needed to improve its worldwide productivity. The present study was carried out to ascertain the nature and magnitude of genetic divergence among 75 sesame genotypes and to select better accessions for sesame improvement programme.

Materials and methods
A total of 75 sesame genotypes were grown in a Randomized Complete Block Design during *Kharif* 2016 at Project Coordinating Unit (Sesame and Niger) Research Farm, JNKVV, Jabalpur (M.P). The soil of the experiment is medium black with uniform topography and free from water logged conditions. Each genotype was grown in 4 row plots of 3 m long, with a spacing of 30 cm between rows and 15 cm between plants. Standard agronomic package and practices were followed to raise a healthy crop. The analysis is based on the determination of the 12 quantitative traits viz., days to flower initiation, days to 50% flowering, plant height (cm), number of primary branches/plant, number of secondary branches/plant, number of capsules/plant, capsule length(cm), days to maturity, number of seeds/capsule, 1000 seed weight (g), seed yield/plant (g) and oil content (%). Genetic divergence among the sesame genotypes was estimated using the Mahalanobis D^2 statistics. Character means were transformed into sets of uncorrelated variables using the pivotal condensation of common dispersion matrix according to Rao (1952) [14]. Although the D^2 statistics can handle a multidimensional situation, higher order interactions do not contribute very much in any experiment. Grouping of genotypes into different clusters was done according to Tocher's method (Rao 1952) [14]. The data were also subjected to Principal Component Analysis (PCA) using method of dimension reduction (Massy, 1965; Jolliffe, 1986) [9, 7].

Results and Discussion
Based on D^2 values, all the genotypes could be grouped into different clusters using non-hierarchical Euclidean cluster analysis. The genotypes within each cluster were closer to each other than the genotypes in different clusters. Genetic divergence among 75 genotypes was determined using seed yield and its attributing traits. In the present study, 75 genotypes were grouped into thirteen clusters based on divergence analysis (**Table 1**). Clustering of genotypes was not associated with the geographical distribution and were mainly grouped due to their morphological differences. Clustering pattern indicated that majority of genotypes, i.e. 64 (85%) were genetically close to each other and grouped in 3 clusters, while apparent diversity was mainly noticed due to 11 genotypes (15%) distributed over 10 clusters. Cluster I was the largest among all clusters comprising 35 genotypes, Cluster II had twenty five genotypes and Cluster IX had four genotypes. Cluster VI had two genotypes, clusters IV, V, VII, VIII, X, XI, XII and III were solitary containing single genotype each. Such grouping of the genotypes is indispensably important for breeding program since, it clearly indicates the significance of a given trait in any group (cluster).
Contribution of each character towards genetic divergence has been estimated from the number of times that each character

appeared in the first rank (Table 2). The trait days to flower initiation (39.82) contributed maximum to genetic divergence by taking 1105 times first rank followed by oil content (28.58%) by 793 times and number of capsules/plant (13.12%) by 364 times, while lowest contribution was given by trait days to maturity (0.04%) by taking one time rank first. Contradictory results were reported by Parameshwarappa *et al*. (2012) [12] for seed yield; Tripathi *et al*. (2013) [19] and Narayanan and Murugan (2013) [11] for seed yield and days to 50% flowering.

The inter-cluster distances were higher than the intra-cluster distances suggesting wider genetic diversity among the genotypes of different groups. By using these inter-genotypic distances intra-cluster genotypic distances were calculated (Table 3). Maximum inter cluster divergence was observed between clusters XII and IV (108.84) followed by clusters XIII and IX (107.11) indicating that good recombinants can be realized by mating between the genotypes KMS-5-371, EC-303304-A, KMR-48-A, KMS-5-587, IS-491-A, NIC-16393 and IC-204063 in a definite fashion. Whereas, the lowest inter cluster distance was observed between cluster IV and cluster III (4.60) indicating close relationship between the genotypes of these clusters and hence, may not be emphasized upon to be used in hybridization programme. Accessions falling under distant clusters could be hybridized to get the higher heterotic responses. Crosses involving parents belonging to the most divergent clusters would be expected to manifest maximum heterosis and wide variability of genetic architecture (Souroush *et al*., 2004). Similarly, Begum *et al*. (2011) [2] also studied the genetic diversity in sesame genotypes through Mahalanobis's (D2) distance and reported high genetic diversity among the genotypes. The greater the distance between two clusters, the wider the genetic diversity among the parents to be included in hybridization program.

The variation observed in cluster mean also points to the degree of variability (Veni *et al.* 2008 and Peluzio *et al.* 2012) [20, 13]. The genotypes viz., NIC-8080-A and IS-347-1-A in cluster VI had maximum cluster mean values for the traits seed yield (13.16), number of secondary branches/plant (5), number of capsules/plant (56.51) and 1000 seed weight (3.51), and genotypes viz., KMS-5-587, IS-491-A, NIC-16393 and IC-204063 in cluster IX had maximum mean for oil content (51.44%) (Table 4). These genotypes may be important for further breeding program to increase grain yield and oil content meanwhile cluster XI could serve as a parent material in the improvement of capsule number as it possess highest capsule number than any other cluster.

PCA is a well-known method of dimension reduction that can be used to reduce a large set of variables to a small set that still contains most of the information in the large set (Massay, 1965; Jolliffe, 1986) [9, 7]. Twelve Principal Components (PCs) were extracted from the twelve quantitative traits of sesame and only 9 principal components (PCs) exhibited more than 0.5 Eigen value and showed about 94.86% variability among the traits studied. According to eigen value and cumulative percent criterion, first seven principal components represent the data efficiently. The first principal component i.e. PC1 accounted for maximum proportion of total variability in the set of all variables and remaining components accounted for progressively lesser amount of variation. The first principal component accounted for maximum variability i.e., 21.51% which reduced gradually to 5.61% in ninth principal component.

Rotated component matrix revealed that each principal component was separately loaded with various yield and

quality attributing traits under study. It revealed that the first principal component (PC1) which accounted for the highest variation (21.51%) was mostly related with phenological traits *viz.*, days to flower initiation, days to 50% flowering and days to maturity. The second principal component (PC2) was dominated by traits *viz.*, number of capsules/plant, number of seeds/capsule while, PC3 consisted mainly of plant height. Fourth and fifth principal components were dominated by yield related traits *viz.*, number of primary branches/plant, number of secondary branches/plant, 1000 seed weight and seed yield/plant. Sixth principal component mainly consisted of oil content and seventh principal component for capsule length (Table 7). Similar results were reported by Saha *et al.* (2012) for days to 50% flowering, days to maturity and number of capsules/plant, Ismaila and Usman (2014) [4] for number of capsules/plant, Shim *et al.* (2016) [16] for days to 50% flowering, days to maturity and plant height. Contradictory reports have been given by Kangbo *et al.* (2009) [8] for days to 50% flowering and days to maturity, Ismaila and Usman (2014) [4] for number of branches/plant.

From this study, it was clear that most of the important yield attributing and quality traits were present in PC2, PC4, PC5 and PC6. Genotypes showing maximum positive PC scores and common in PC2, PC4 and PC5 for the yield and yield attributing traits viz., number of capsules/plant, number of seeds/plant, number of secondary branches/plant, number of primary branches/plant, 1000 seed weight and seed yield/plant are S-0644-A and KMS-5-361; genotype viz., T-1-A is common in PC4, PC5 and PC6 having positive PC scores for yield and quality traits and having negative value for phonological traits viz., days to flower initiation, days to 50% flowering and days to maturity. Thus, selection of these can help in further development of new high yielding quality varieties. Maximum negative values were recorded in genotypes SI-1003, SI-269, NIC-16409, NIC-16204, IC-56162, IC-43102-A and IS-347-1-A for traits days to 50% flowering and days to maturity.

Overall results revealed that the genotypes S-0644, KMS-5-361, T-1-A, NIC-16214, SI-1125, NIC-8080-A have highest PC values for characters number of capsules/plant, number of seeds/capsule, number of primary branches/plant, number of secondary branches/plant, 1000 seed weight and seed yield/plant and negative value for PC1 *viz.*, phenological traits is T-1-A. Thus, these genotypes might be used for development of new varieties of sesame.

References:
1. Arunachalam V. Genetic distance in plant breeding. Indian J. Genet. 1981; 41:226-236.
2. Begum S, Islam MA, Husna A, Hafiz TB, Ratna M. Genetic diversity analysis in sesame (*Sesamum indicum* L.). Journal of Agriculture. 2011; 9(2):65-71.
3. Harlan JR. The living fields: Our agricultural heritage. Cambridge University Press Cambridge. 1975.
4. Ismaila A, Usman A. Genetic variability for yield and yield components in sesame (*Sesamum indicum* L.). International Journal of Science and Research. 2014; 3(9):63-66.
5. Jain SK. Genetic diversity of weedy rye populations in California. Crop Sci. 1977; 17:480-482.
6. Jain SK, Qualset CO, Bhatt GM, Wu KK. Geographical patterns of phenotypic diversity in a world collection of durum wheats. Crop Sci. 1975; 15:700-704.
7. Jolliffie IT. Principal Component Analysis. Springer, New York. 1986.
8. Kangbo S, Dong HC, Bok PS, Whan PJ, Cheon BJ, Yong PK. Interpretation of agronomic traits variation of sesame cultivar using principal component analysis. Korean Journal of Crop Science. 2009; 54(1):24-28.
9. Massay WF. Principal components regression in exploratory statistical research. Journal of the American Statistical Association. 1965; 60:234-246.
10. Mohammadi SA, Prassana BM, Singh NN. Sequential path model for determining interrelationship among grain yield and related characters in Maize. Crop Sci. 2003; 43:1690-1697.
11. Narayan R, Murgan S. Studies on variability in sesame (*Sesamum indicum* L.). International Journal of Agricultural Research. 2013; 2(11):52-55.
12. Parameshwarappa SG, Palakshappa MG, Shinde, Gayatree G. Studies on genetic divergence for yield and yield attributing traits in sesame (*Sesamum indicum* L.). germplasm. International Journal of Agricultural Sciences. 2012; 8(2):441-444.
13. Peluzio JM, Pires LPM, Cancellier LL, Afferri FS, Colombo GA, Teixeira Junior T *et al.* Genetic divergence among soybean cultivars in irrigated lowland in the state of Tocantins. Ciencia Rural. 2012; 42(3):395-400.
14. Rao CR. Advanced Statistical Methods in BiometericaJ Research, John Wiley and Sons Inc., New York. 1952.
15. Saha Sruba, Begumb T, Dasgupta T. Analysis of genotypic diversity in sesame based on morphological and agronomic traits. Conference on International Research on Food Security, Natural Resource Management and Rural Development organised by Georg-August Universität Göttingen and University of Kassel-Witzenhausen University of Calcutta, Department of Genetics and Plant Breeding, 35 Ballygaunge Circular Road, 700019 Calcutta, India. 2012.
16. Shim KB, Shin SH, Shon JY, Kang SG, Yang WH, Hew SG. Classification of a collection of sesame germplasm using multivariate analysis. Journal of Crop Science and Biotechnology. 2016; 19(2):151-155.
17. Smith JSC, Smith OS. The Description and assessment of distances between inbred lines of maize: The utility of morphological, biochemical and genetic descriptors and a scheme for the testing of distinctiveness between inbred lines. Maydica, 1989; 34:151-161.
18. Souza E, Sorrells ME. Pedigree Analysis of North American oat cultivars released from 1952 to 1985. Crop Sci. 1989; 29:595-601.
19. Tripathi A, Bisen R, Ahirwal RP, Paroha S, Sahu R, Ranganatha ARG. Study on genetic divergence in sesame (*Sesamum indicum* L.) germplasm based on morphological and quality traits. Bioscan. 2013; 8(4):1387-1391.
20. Veni BK, Rani BP, Ramana MV. Genetic divergence studies in soybean [*Glycine max* (L.) Merrill]. Soybean Research. 2008; 6:77-80.