Biological Forum – An International Journal **13(3): 680-685(2021)**

ISSN No. (Print): 0975-1130
ISSN No. (Online): 2249-3239

Multivariate Based Diversity Analysis in Indian Mustard (*Brassica juncea* L.) Genotypes by Principle Component Analysis

***Shubham Chakraborty*[1], *Awnish Kumar*[1*], *Chandan Kishore*[2], *Ravi Shankar Singh*[2], *Anil Kumar*[2], *Anand Kumar*[2] *and Mankesh Kumar*[2]**
[1]M.Sc. Scholar (Ag) Department of Plant Breeding and Genetics,
Bihar Agricultural University, Sabour, Bhagalpur, (Bihar), India.
[2]Assistant Professor-Cum-Junior Scientist, Department of Plant Breeding and Genetics,
Bihar Agricultural University, Sabour, Bhagalpur, (Bihar), India.

(Corresponding author: Awnish Kumar)*

(Received 03 June 2021, Accepted 18 August, 2021)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Thirty three accessions including four checks of Indian mustard (*Brasssica juncea* L.) were evaluated in randomized block design having three replication for 10 quantitative characters to study the genetic diversity and correlation at Oilseed section, Bihar Agricultural University, Sabour, Bhagalpur. Multivariate based diversity analysis was determined in present research using Principle Component Analysis (PCA). To draw the inference for genetic diversity, data were recorded for ten quantitative traits *viz.* days to 50% flowering, days to maturity, plant height, Number of primary branches, Number of secondary branches per plant, Number of siliqua per plant, Siliqua length, Seed per siliqua, Test weight and Seed yield per plot. To estimate the relative contribution of each traits in total variability, principal component analysis was done. The analysis revealed the cumulative variability of 82.25 % expressed by the first five major axis. The eigen-value of the first principal component equaled 2.26 and represented 22.69 % of total variability. The second PC had eigen-value 1.92 giving 19.32 % variability. The remaining variability of 17.21 %, 12.92 % and 10.21 % was observed in component III, component IV and component V respectively by various traits such as seeds per siliqua, 1000 seed weight, number of secondary branches and days to maturity. Thus the findings of the experiment indicated the significant genetic variation among the genotypes and the traits contributing to the total variation were identified. However, the findings of the study could be more reliable if a multi-location or multi-seasonal polled data were available. Hence, genotypes used in this study can be utilized in further breeding programs for trait improvement contributing to major variation.

Keywords: Genetic diversity; Indian mustard; PCA; germplasm.

INTRODUCTION

Indian mustard (*Brassica juncea* L) is used as oilseed, vegetable and as a condiment on a global basis. It was among the earliest domesticated crop species during human civilization. Large number of wild as well as cultivated species are known over a large geographical region, across the continents like Asia, Africa, Europe, Australia, and America. (Hopkins *et al.*, 2007). It is an annual amphidiploid dicot crop (AABB) having chromosome number 2n = 4X = 36 and belongs to the family *Brasssicacae*. It is the third most important source of oil, after groundnut and soybean (Rabbani *et al.*, 1998; Abbas *et al.*, 2009). It is an important *Rabi* oil seed *brassica* cultivated mostly in North Indian under irrigated condition. With the ever increasing human population, introduction of high grain and oil yielding varieties are needed in order to fulfill consumer demand. Hence, to feed this huge population and for economic upliftment agrarian society, it is necessary to enhance the yield of this crop. To achieve this goal, it is essential to exploit the genetic diversity of these crops. Hence, an effort has been made to collect and evaluate diverse germplasm for their yield potential as well as other yield attributing traits.

Genetic variation refers to the heritable portion of the total variation present within and among the populations of a particular species. Diversity is estimate at three different level *viz.*, species diversity, ecosystem diversity and genetic diversity. Firstly, the Species diversity in which the abundance of different species at particular location is estimated. Secondly, the eco-system diversity, in which we study how a species interact with its environment and the third one is genetic diversity, in which variation at gene and germplasm level is observed. Genetic variation pool present within an inter-mating population forms the basis for selection and further crop improvement. In order to exploit a population for trait improvement, it is necessary to estimate the quantum of variability present in the population which is a must for genetic improvement in any crop species. For hybrid crop breeding, genetic distance estimations are used to choose parental combinations with appropriate genetic diversity and to classify germplasm into distinct heterotic groups. The distribution of population into various groups can be based on pedigree information, molecular marker data, agro-morphological traits or geographical origin. Since it's crucial to understand the population's exploitable variability, multiple methods can be used to determine genetic distance for population categorization. Multivariate analysis is one of the easiest ways. Multivariate methods have extensive use in summarizing and describing the inherent variation among crop genotypes estimates as it is the most widely used method for varietal differentiation. Multivariate statistical tools include Principal Component Analysis (PCA), Cluster analysis and discriminate analysis. Principal Component Analysis (PCA) is generally used to uncover similarities between variable and classify the genotypes into different groups. PCA examines a data table in which observations are characterized by a number of quantitative dependent variables that are all inter-correlated. Its major function is to extract significant information from the table and to depict the pattern of similarity between the observations and variables as points on maps by portraying it as a set of new orthogonal variables called principal components. The importance and contribution of each component to total variance is measured as 'Proper values', whereas the degree of contribution of every original variable associated with which each principal component is indicated as coefficient of proper vectors.

The greater the coefficient values, regardless of direction (positive or negative), the better they are in distinguishing different genotypes. In order to bridge the gap between global demand and supply, diversity must be present in different crops (Shiva, 1994). Identification and utilization of improved genotypes for crop improvement is accomplished by proper evaluation of important crop species (Jan *et al.*, 2016a; Jan *et al.*, 2016b; Khan *et al.*, 2016; Jan *et al.*, 2016c; Jan *et al.*, 2016d; González-Rubio *et al.*, 2016). Exploitation of the germplasm lines in breeding programs demands a wide variability for economic traits. Hence, determination of different traits contributing to diversity among the genotypes is a pre-requisite for making effective selection. As a result, the goal of this research was to describe and quantify genetic variation and connections across Indian mustard genotypes using phenotypic characters in order to identify and select promising genotypes for economically important traits.

MATERIALS AND METHODS

Thirty-three diverse genotypes of Indian mustard were procured from different agricultural institutes of India. The research trail was conducted at Oilseed Research Section, Bihar Agricultural University, Sabour (Bhagalpur) in randomized block design (RBD) with three replications in *Rabi* season (2019-20). Hand drill was used to sow the seeds of experimental materials at a depth of 3 to 4 cm. For a healthy crop, all suggested packages and practices were used. When about 75% of the plants had turned yellowish in colour, the crop was harvested. Distances between the two plots were maintained at 50 centimeter while distance between replications was one meter. Normal plant-plant spacing was followed i.e. 30 cm×10 cm. Each plot consisted of six rows out of which one border row each from left and right side were discarded before data recording. Data were recorded for ten quantitative traits *viz.* days to 50% flowering, days to maturity, plant height, Number of primary branches, Number of secondary branches per plant, Number of siliqua per plant, Siliqua length, Seed per siliqua, Test weight and Seed yield per plot.

Data for days to 50% flowering, days to maturity and seed yield per plot were taken on plot basis while rest data were taken from random sample of five individual plant basis from each replication. The recorded data on all ten quantitative traits were statistically analyzed in Windostat (Version 9.3) for multivariate based genetic diversity based on Principle component analysis (PCA) to cluster the accessions based on genetic similarity. By limiting the dimensions of multivariate data to a few major axes, the PCA analysis generates an Eigen vector for each axis and gives component scores for the characters (Sneath and Sokal 1973: Ariyo and Odulaja, 1991).

RESULTS

The contribution and significance of each component to overall variance are measured using Principal Component Analysis. The coefficient of appropriate vectors reflects the extent to which each original variable with which each principal component is associated. It also calculates a trait's independent contribution to overall variance. A screen plot depicting eigen-value variation for different components (Fig. 1) and percent of variation expressed by them (Fig. 2) has been shown.

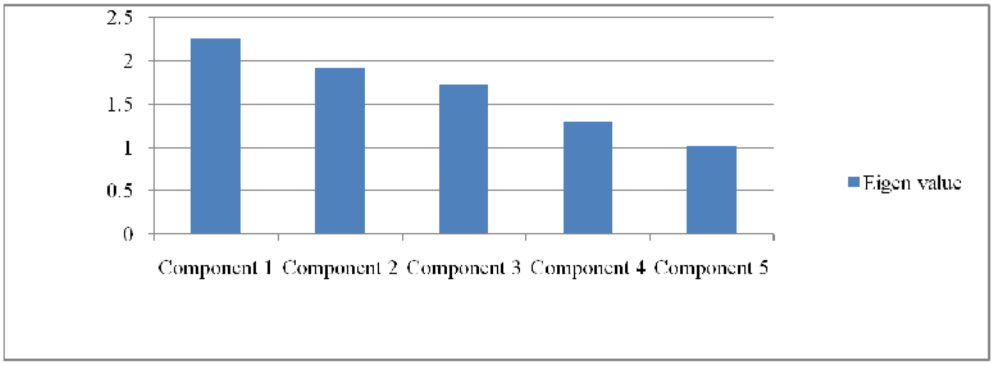

Fig. 1. Screen plot showing eigen value variation.

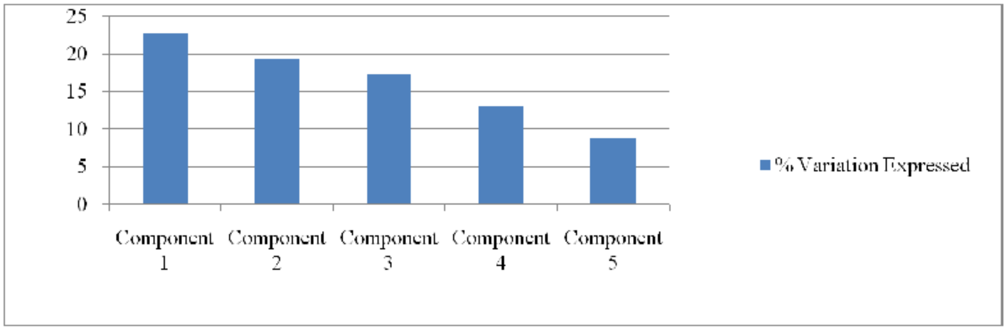

Fig. 2. Screen plot showing % of expressed variation by five components.

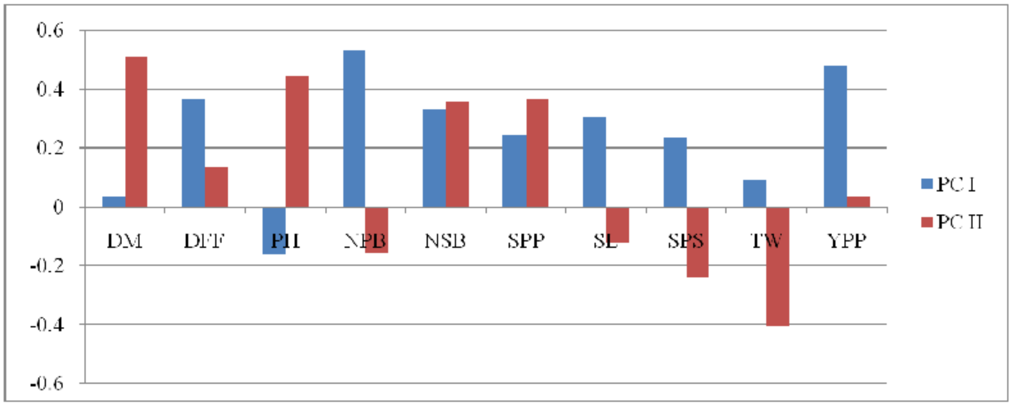

Fig. 3. Contribution of agro-morphological traits in PC I and PC II.

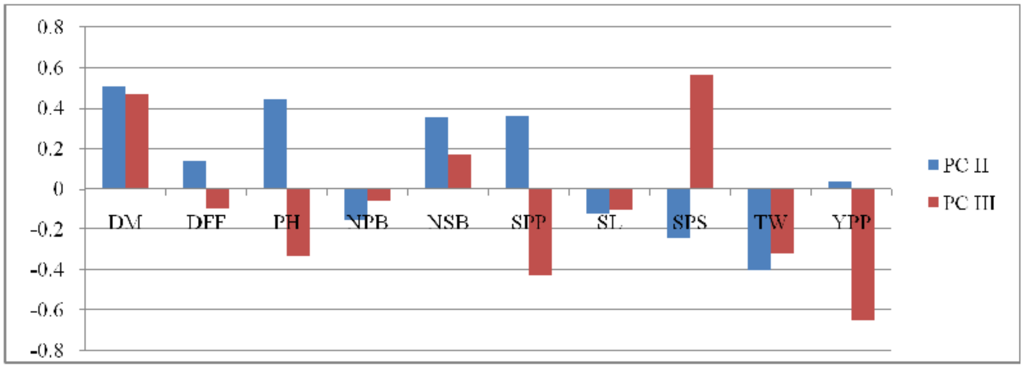

Fig. 4. Contribution of agro-morphological traits in PC II and PC III.

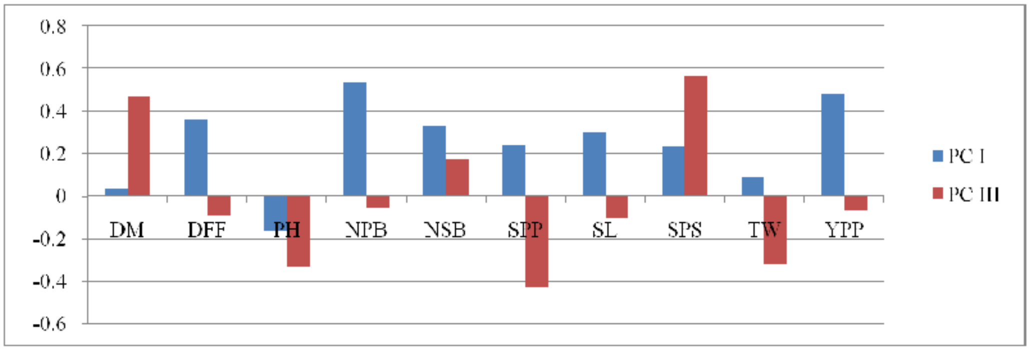

Fig. 5. Contribution of agro-morphological traits in PC I and PC III.

The 5 principal components accounted for 80.75 % of the total variability among the studied *B. juncea* L. genotypes (Table 1, Figs. 2-5). Among the five principal components (PCs), the contribution of **PC-I** was found to have 22.69 % out of the total variability. Number of primary branches (0.531) followed by seed yield per plot (0.481), days to maturity (0.36), number of secondary branches (0.332), siliqua length (0.322), number of siliqua per plant (0.24), seeds per siliqua (0.231), 1000 seed weight (0.091), days to 50% flowering (0.034) contributed positively to first principal component. In contrast, plant height (-0.164) contributed negatively.

The contribution of **PC-II** in total variability was 19.2% which was positively associated with days to 50% flowering (0.512) followed by plant height (0.445), number of siliqua per plant (0.364), number of secondary branches (0.356), days to maturity (0.133) and yield per plot (0.03) whereas siliqua length (-0.125) followed by number of primary branches (-0.152), seeds per siliqua (-0.246) and 1000 seed weight were negatively associated with PC-II.

PC-III showed 17.21% of the total agro-morphological variation and was positively associated with seeds per siliqua (0.562) followed by days to 50% flowering (0.478), number of secondary branches (0.174) while negatively associated with number of primary branches (-0.052), days to maturity (-0.096), yield per plot (-0.062), siliqua length (-0.012), 1000 seed weight (-0.325), plant height (-0.331) and siliqua per plant (-0.422).

The **PC-IV** contributed 12.92% in the total variability and were positively correlated to 1000 seed weight (0.552) followed by number of secondary branches (0.342), siliqua per plant (0.192), seeds per siliqua (0.174), number of primary branches (0.155), days to 50% flowering (0.055) and plant height (0.012). Similarly the traits i.e. days to maturity (-0.133), yield per plot (-0.175) and siliqua length (-0.652) contributed negatively.

In the overall variability, the contribution of **PC-V** was observed 10.21% and the parameter i.e. days to maturity (0.801) followed by number of primary branches (0.166), days to 50% flowering (0.074) contributed positively while the characters i.e. 1000 seed weight (-0.056), plant height (-0.074), seeds per siliqua (-0.091), siliqua per plant (-0.133), number of secondary branches (-0.294), siliqua length (-0.291) and yield per plot (-0.321) contributed negatively.

Table 1: Principal components (PCs) of agro-morphological traits among *B. juncea* L. accessions.

	PC I	**PC II**	**PC III**	**PC IV**	**PC V**
Eigen value	2.26	1.92	1.72	1.29	1.02
Cumulative eigen value	2.26	4.18	5.90	7.19	8.21
Percent variance	22.69	19.20	17.21	12.92	10.21
Cumulative variance	22.69	41.89	59.10	72.03	82.25
Traits	Eigen vectors				
DFF	0.034	0.513	0.471	0.058	0.078
DM	0.365	0.131	-0.094	-0.130	0.808
PH	-0.163	0.444	-0.334	0.019	-0.072
NPB	0.533	-0.158	-0.055	0.153	0.169
NSB	0.330	0.357	0.174	0.342	-0.0294
NSP	0.241	0.363	-0.428	0.196	-0.137
SL	0.304	-0.120	-0.104	-0.659	-0.291
SPS	0.233	-0.241	0.566	0.172	-0.092
TW	0.092	-0.403	-0.323	0.551	-0.053
YPP	0.480	0.034	-0.065	-0.175	-0.323

Note: DFF = Days to 50% maturity, **DM** = Days to maturity, **PH** = Plant Height (cm), **NPB**= Number of primary branches, **NSB**=Number of secondary branches, **NSP** = Number of siliqua per plant, **SL** = Siliqua length (cm), **SPS** = Number of seeds per siliqua, **TW** = Test weight (gm), **YPP**= Seed yield per plot(g)

DISCUSSIONS

Maximum variability observed among the genotypes from different locations was accounted by the first three Principal components (PC I), (PC II) and (PC III). It accounted for 59 per cent of the morphological variation in the Indian mustard genotypes and was loaded on number of primary branches, seeds per siliqua, siliqua per plant and yield per plot. The first five components in the PCA analysis with eigen values >1 contributed 82.25 per cent of the variability among genotypes evaluated for different phenotypical traits. Hence, the important traits coming together in a particular PC by contributing towards variability has the tendency to remain together which in turn offers opportunity for their effective utilization in crop breeding. Results of the present investigation were supported by the findings of previous literature. Using PCA, a number of plant scientists and breeders in the past were able to obtain outstanding results in terms of diversity in agro-morphological traits for diverse Brassica individuals. These approaches have also been effectively used in Indian mustard (*Brassica juncea* L.) germplasm (Gupta *et al.*, 1991), Ethiopian mustard (Alemayehu and Becker, 2002), and white head cabbage germplasm (Balkaya *et al.*, 2005). Similar results on variability in Indian mustard contributed by the major traits mentioned in the present study have been also reported. (Awasthi *et al.*, 2020*;* Lakra *et al.* 2020; Rout *et al.*, 2019)

Findings of the above mentioned literature amply supported our present investigation that PCA is a good statistical tool in estimating associations among individuals having originated from different environments The classification of *Brassica juncea* L. genotypes into various groups was not only due to their origin from various ecological zones throughout the country and the world (Amurrio *et al.*, 1995), who investigated qualitative and quantitative traits of Iberian pea genotypes, which were superimposed with the current findings. During the current study, significant genetic differences were observed for many agro-morphological parameters across various *Brassica juncea* L. genotypes. In the future, these traits investigated might play a significant role in Brassica varietal development programme.

CONCLUSION

Principal component analysis has found a few main characteristics that play a key role in categorizing the existing variation of the germplasm set. The number of primary branches, seeds per siliqua, siliqua per plant, and seed yield per plot in different principle components were shown to be the most essential for accounting the variance. Hence, these traits must be given due importance while selecting genotypes for varietal development. Since, variability in any population is the main component of any breeding programme and presence of genetic variability is a must for crop breeding. As a result of the findings, the high degree of genetic diversity among genotypes was observed, as well as the traits that contribute to this diversity. Hence, the results of this study will be more useful in selecting promising parents for enhancing different morphological characteristics that affect yield in *Brassica juncea*.

Acknowledgement. Authors are highly thankful to the Chairman of Department of Plant Breeding and Genetics, BAU, Sabour for providing logistic support necessary to conduct the research work.

Conflict of Interest. None.

REFERENCES

Abbas, S. J., Farhatullah, M. K. B., Khan, I. A., & Munir, I. (2009). Molecular Analysis of Genetic Diversity in Brassica Species. *Pakistan Journal of Botany, 41*: 167-176.

Alemayehu, N., & Becker, H. (2002). Genotypic Diversity and Patterns of Variation in a Germplasm Material of Ethopian mustard (*Brassica carinata* A. Braun). *Genetic Resources and Crop Evolution, 49,* 573-582.

Amurrio, J. M., De Ron, A. A., & Zeven, A. C. (1995). Numerical Taxonomy of Iberian Pea Landraces Based on Quantitative and Qualitative Characters. *Euphytica, 82:* 195-205.

Ariyo, Q. J., & Odulaja, A. (1991). Numerical analysis of variation among accessions of Okra [*Abelmoschus esculentus* (L.) Moench], Malvaceae. *Annals of Botany; 67:* 527–531.

Awasthi, D., Vimlesh, D. K., & Kandalkar, V. S. (2020). Evaluation of Heritability and Genetic Advance for Morphological Traits of Indian mustard Germplasms. *Current Journal of Applied Science and Technology, 39*(21): 39-47.

Balkaya, A., Yanmaz. R., & Kar, A. A. (2005). Morphological Characterization of White Head Cabbage (*Brassica oleracea* var. capita tasubvar. alba) Genotypes in Turkey. *New Zealand Journal of Crop and Horticultural Science, 33,* 333-341.

González-Rubio, C., Rodríguez-Estrella, R., Lozano-Garza, O. A., & García-DeLeón, F. J. (2016). Genetic Diversity of the Endemic Xantus' Hummingbird Using 16 Novel Polymorphic Microsatellite Loci, and their Cross Amplification between Six Related Species. *Open Journal of Genetics, 6*: 19-27.

Gupta, V. P., Sekhon, M. S., & Satija, D. R. (1991). Studies on Genetic Diversity, Heterosis and Combining Ability in Indian Mustard (*Brassica juncea* (L.) Czern. & Coss). *Indian Journal of Genetics, 51*: 448-453.

Hopkins, C. J., Cogan, N. O. I., Hand, M., Jewell, E., Kaur, J., Li, X., Lim, G. A. C., Ling, A. E., Love, C., Mountford, H., Todorovic, M., Vardy, M., Spangenberg, G. C., Edwards, D., & Batley, J. (2007). Sixteen New Simple Sequence Repeat Markers from *Brassica juncea* Expressed in Sequences and Their Cross-Species Amplification. *Molecular Ecology Notes, 7*: 697-700.

Jan, S. A., Shinwari, Z. K., & Rabbani, M. A. (2016a). Morpho-Biochemical Evaluation of *Brassica rapa* Sub-Species for Salt Tolerance. *Genetika, 8,* 323-338.

Jan, S. A., Shinwari, Z. K., & Rabbani, M. A. (2016b). Determining Genetic Divergence among *Brassica rapa* Ecotypes through Electrophoretic Mobility of Total Seed Proteins. *Journal of Animal and Plant Sciences, 26,* 1758-1764.

Jan, S. A., Shinwari, Z. K., & Rabbani, M. A. (2016d). Agro-Morphological and Physiological Responses of *Brassica rapa* Ecotypes to Salt Stress. *Pakistan Journal of Botany*, *48*: 1379-1384.

Jan, S. A., Shinwari, Z. K., Rabbani, M. A., Shah, S. H., Ibrahim, M. I., & Ilyas, M. (2016c). Optimization of an Efficient SDS-PAGE Protocol for Rapid Protein Analysis of *Brassica rapa*. *Journal of Biodiversity and Environmental Sciences*, *9*: 17-24.

Khan, Q., Khan, A. S., Khurshid, H., Jan, S. A., Ahmad, N., Khan, S. A., Shah, S. H., Saleem, N., Ibrahim M. I., Ilyas, M., & Arif, M. (2016). Exploring Durable Genetic Resistance against Leaf Rust through Phenotypic Characterization and LR34 Linked STS Marker in Wheat Germplasm. *Bioscience Journal*, *32*: 986-998.

Lakra, A., Tantuway, G., Tirkey, A. E., & Srivastava, K. (2020). Genetic variability and Trait association studies in Indian mustard (*Brassica juncea* L. Czern & Coss). *International Journal of Current Microbiology and Applied Sciences*, *9*(1): 2556-2563.

Rabbani, M. A., Iwabuchi, A., Murakami, Y., Suzuki, T., & Takayanagi, K. (1998). Phenotypic Variation and the Relationships among Mustard (*Brassica juncea*) *Germplasm from Pakistan. Euphytica*, *101*: 357-366.

Rout, S., Kerkhi, S. A., & Gupta, A. (2019). Estimation of genetic variability, heritability and genetic advance in relation to seed yield and its attributing traits in Indian mustard [*Brassica juncea* (L.) Czern and Coss.]. *Journal of Pharmacognosy and Phytochemistry*, *8*(3): 4119-4123.

Shiva, V. (1994). Agriculture and Food Production. UNESCO/Environment Education Dossiers, *9*, 2-3

Sneath, P. H., & Sokal, R. R. (1973). Numerical Taxonomy. *The Principles and Practice of Numerical Classification*. W.H. Freeman & Co Ltd.

How to cite this article: Chakraborty, S., Kumar, A., Kishore, C., Singh, R.S., Kumar, A., Kumar, A. and Kumar, M. (2021). Multivariate Based Diversity Analysis in Indian Mustard (*Brassica juncea* L.) Genotypes by Principle Component Analysis. *Biological Forum – An International Journal*, *13*(3): 680-685.