AATCC Review

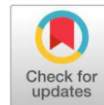**Research Article**

Open Access

Studies on genetic divergence and principal component analysis in okra [*Abelmoschus esculentus* L. (Moench)]

Balagoni Maruthi*[1], Sibsankar Das[1], Arup Chattopadhyay[1], Umesh Thapa[1], Anirban Maji[2], Adapa Kiran Kumar[3] and Pranab Hazra[1]

[1]Department of Vegetable Science, Faculty of Horticulture, BCKV, Mohanpur, West Bengal, 741252, India.
[2]Department of Genetics and Plant Breeding, Faculty of Agriculture, BCKV, Mohanpur, West Bengal, 741252, India.
[3]Director of Research, SKLTSHU, Mulugu, Telangana, India.

ABSTRACT

In the current investigation, 50 okra genotypes were grouped into 9 clusters by treating estimated D^2 values as the square of the generalized distance. Among the 9 clusters, cluster II had the highest intra-cluster value (74.77) followed by Cluster I (67.14), and the highest inter-cluster value was noticed between Cluster VII and IX (167.97) followed by Cluster II and IX (152.31). The maximum contribution towards divergence was shown by the characters like PDI of OELCV at 90 DAS (23.53 %), PDI of YVMV at 90 DAS (17.22), fruit length (9.27), days to 1st blooming (9.00 %) and the internodal length (7.69 %). The PCA was carried out to get a simplified view of the relationship between the attributes. From the PCA plot of PC_1 vs. PC_2 (Dim_1 vs Dim_2), selection may be refined considering all 8 principal components, with Punjab 8 being the best-performing cultivar having an optimum combination of all variables including OELCV and BYVMV disease tolerance. Seven diverse genotypes were selected based on yield potentiality, disease tolerance, multivariate analysis, and PCA.

Keywords: Okra, characters, cluster, divergence, genotypes, principal component analysis.

Introduction

Okra (*Abelmoschus esculentus* (L.) Moench) is a member of the Malvaceae family and earlier was referred to as *Hibiscus esculentus* (L.) [1]. It is also known as Lady's Finger, Okura, Okro, Bhindi, Gumbo Okra, and Quiabos in different countries [2]. As a vegetable, immature okra fruit can be utilized in soups, salads, and stews, dried or fresh, boiled or fried [3]. In India, the area under okra cultivation is around 5.28 million ha with a production of 61.46 million tons [4]. Okra has a high export potential among fresh vegetables, making it a valuable source of earnings [5].

The creation of new cultivars that fulfill the evolving demands for adjusting to varying growing conditions, resistance to abiotic and biotic stresses, yield of product, and specific quality criteria is unattainable for plant breeders in the absence of a diverse genetic pool consisting of heterogeneous plant material [6]. Hence, the presence of genetic diversity is an essential need for the efficacy of any improvement scheme [7]. Genetic diversity may appear employing natural processes within the gene pool or be intentionally generated through various artificial methods [8].

The successful creation of commercial hybrids emphasizes the significance of genetic divergence as an essential factor in choosing appropriate genotypes for hybridization programs in the context of genetic diversity in okra [9].

Materials and Methods

The current investigation was carried out at the Teaching Farm, College of Agriculture, Burdwan during the spring summer, and kharif seasons, in 2021. The experimental material consisted of 50 okra genotypes that were collected from various sources in India. All 50 genotypes were grown in Randomized Block Design with three replications. Observations on 17 quantitative traits were noticed on ten randomly selected plants in each treatment for each genotype. Standard cultural practices will be followed as per [10]. The average values of 17 quantitative traits (average of two seasons) were subjected to statistical analysis. Divergence was analyzed by using D^2 statistics [11], and the genotypes were grouped into clusters by Tocher's method given by [12].

Applying principal component analysis (PCA) to determine the factor dimension of the data, varietal information was simplified into a smaller number of factors to choose the genotypes that performed the best [13]. The current study sought to find distinct genotypes for use in breeding programs by examining the kind and extent of genetic diversity employing 17 important quantitative traits among 50 okra genotypes.

Results and Discussion

Based on divergence determination, all 50 genotypes were grouped into 9 clusters (Table 1). Cluster I was the largest having 26 genotypes followed by Cluster II with 10 genotypes and Cluster III having 5 genotypes. Cluster IV with 3 genotypes, Cluster V with 2 genotypes, and Cluster VI, VII, VIII, and IX have monotypic genotypes each. The genotype grouping pattern was found to be random, demonstrating that there is no direct association between geographical distribution and genetic distance.

The lack of relationship between genetic diversity and

geographical distance implies that forces other than geographical origin such as genetic stock exchange, genetic drift, spontaneous mutation, and natural and artificial selection are responsible for genetic diversity [14]. As a result, genotypes for hybridization should be chosen based on genetic divergence rather than geographic diversity. Many earlier investigations also revealed that various sets of bhindi genotypes were grouped under 4-10 clusters [15; 16; 17; 18].

The intra and inter-cluster distances among 50 bhindi genotypes are displayed in Table 2. Among the 9 clusters, cluster II had the highest intra-cluster value (74.77) followed by cluster I (67.14) indicating a prevalence of wide genetic divergence among the genotypes as compared to the remaining clusters. Clusters VI, VII, VIII, and IX showed the least intra-cluster value (0.00). At the inter-cluster level, the highest inter-cluster value was noticed between Cluster VII and IX (167.97) followed by between Cluster II and IX (152.31) indicating the genotypes included in these clusters had a high divergence. The minimum inter-cluster value was noticed between Cluster II and VII (82.39) which indicated a close relationship among the genotypes included in these clusters. Kalloo *et al.* [19] recommended that the crosses between selected varieties from widely separated clusters were most likely to give ideal recombinants.

The top five characters (fig. 2) that contributed most towards the genetic divergence (Table-4) were PDI of OELCV at 90 DAS (23.53 %) followed by PDI of YVMV at 90 DAS (17.22 %), fruit length (9.27), days to 1st blooming (9.00 %) and inter nodal length (7.69 %). These traits may be utilized in selecting genetically diverse parents for hybridization programs to exploit either maximize heterosis or to execute efficient selection in the segregating generation.

The cluster means of 50 genotypes (Table 3) displayed that the mean values of the clusters varied in magnitude for all 17 attributes. Cluster VI had the highest fruit yield/plant followed by clusters IV and Cluster V. Cluster VI also displayed the highest plant height, taking a minimum of days to 1st flowering and days to 50 % flowering. The maximum cluster mean was in cluster VI for the number of primary branches/plant, number of fruits/plant, number of nodes on main stem/plant, and seed number /fruit. The highest cluster mean for fruit diameter was recorded by Cluster III followed by Clusters I and II. Cluster II showed the highest cluster mean for PDI of BYVMV at 90 DAS and Cluster IX for PDI of OELCV at 90 DAS. Genotypes belonging to Clusters VI could be regarded as useful sources of genes for improving productivity with early maturity that can be fitted well in multiple cropping systems.

For crop improvement in Bhindi, intercrossing among genotypes with outstanding mean performance was suggested by earlier workers [20; 21; 22].

Principal component analysis: The PCA was carried out to get a simplified view of the relationship between the attributes, PDI of OELCV at 90 DAS, PDI of BYVMV at 90 DAS, days to 1st flowering, internodal length, fruit length, fruit yield per plant, no. of primary branches per plant and days to 50 % flowering

which explained almost 100% contribution towards divergence, and variable loadings for components PC_1 (PDI of OELCV at 90 DAS), PC_2 (PDI of BYVMV at 90 DAS), PC_3 (days to 1st flowering), PC_4 (internodal length), PC_5 (fruit length), PC_6 (fruit yield per plant), PC_7 (no. of primary branches per plant) and PC_8 (days to 50 % flowering) were estimated (Table 5). These components were chosen because they explained almost 100.00 % of the total variance. The eigenvalues were more than 1.0 for the first 3 components.

The first component (PC_1) explained 36.40 % of the total accounted for variance in which a decrease of OELCV disease severity leads to an increase in fruit yield/plant, fruit length, primary branches number, and a decrease of BYVMV disease severity and days to first flowering. The second component (PC_2) explained an additional 23.16 % of the variance in which a decrease in BYVMV disease severity associated with increased days to 1st blooming, days to 50 % blooming, OELCV disease severity, fruit yield per plant, internodal length, primary branches number, and fruit length. The third component (PC_3) explained an additional 17.38 % of the variance in which an increase in days to 1st flowering leads to a decrease of OELCV disease severity, fruit yield per plant, and fruit length is also associated with increased YVMV disease severity, days to 50 % flowering, primary branches count and internodal length. The fourth component (PC_4) explained an additional 8.85 % of the variance in which a decrease in internodal length leads to an increase in fruit length, primary branches count, and fruit yield per plant.

The PCA was also utilized to detect correlations among Indian okra genotypes and to find diverse parents for a hybridization scheme (22; 23). There are no definitive rules for determining the significance of a trait coefficient for each principal component. A coefficient larger than half of the coefficient divided by the square root of the standard deviation of the eigenvalue of the relevant principal component is considered significant by Johnson and Wichern [24]. In PCA, genotypes that are near together are regarded as similar; genotypes that are farther away are perceived as more divergent.

The differences noticed in the data, and summarized in the PCA (fig. 3), indicated genotypes, Punjab 8, AKO 107, Ajeet 121, Greengold, Anima, Hissar unnat, and Hoshiarpur local were quantitatively dissimilar from others. The remainder of the genotypes had similar features making a separate cluster. From the plot of PC1 vs. PC2 (Dim_1 vs Dim_2), selection may be refined considering all 8 principal components, with 'Punjab 8' being the best-performing cultivar having an optimum combination of all variables including OELCV and BYVMV disease tolerance, followed by genotypes, 'Hissar Unnat', and 'Hoshiarpur local', and can be used as improved genetic material for disease resistant breeding against OELCV and BYVMV disease.

Based on D^2 statistics, principal component analysis, and average performance for fruit yield/plant, OELCV and BYVMV disease severity traits, seven diverse genotypes, Punjab 8, AKO 107, Ajeet 121, Greengold, Anima, Hissar Unnat and Hoshiarpur local were identified as good candidates for utilization in disease resistant breeding against OELCV & BYVMV diseases.

Table 1: Cluster classification of the 50 genotypes of okra

Clusters Genotypes
Cluster 1 Deepti, Kiran, NRB 208, Annika, Ankur 41, Raja, Ruchi, Dhanvi 66, Hisar Unnat, Bhindi Selection 51, Mahima Super, Hari Pari, Mahijha No. 25, Punjab 7, Danteshwari, Thakath (TS-102), Hoshairpur Local, Rajrani, Anima, Gold 207, Super Anamika, Madurai Local, Panchwati, Hina RCH, Suguna A-51, and MH 310
Cluster 2 Hari Kranti, Summer Beauty, GFS GOLD (V-4), Pusa Sawani, Maharani, Rani 792, Harika, Research Soniya, AKO 107 and Ajeet-121
Cluster 3 Pusa Makhmali, NOL 303+ (JULIE), Bambeshwari, Chidambaram Local and Super Champion 55
Cluster 4 NOL 1307 (SILKY), Green Gold and Chandra IMP
Cluster 5 Arka Anamika and Arka Abhay
Cluster 6 Punjab 8
Cluster 7 Super Lady luck
Cluster 8 Sarala (LS-11)
Cluster 9 Mahyco 777

Table 2: Intra and inter-cluster distances of 50 genotypes of okra

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9
Cluster 1	**67.14**	90.73	90.37	87.70	89.61	85.35	109.26	93.53	111.67
Cluster 2		**74.77**	99.14	100.19	102.37	124.44	82.39	113.85	152.31
Cluster 3			**60.19**	95.75	124.91	106.36	113.19	90.92	99.96
Cluster 4				**66.78**	114.57	103.23	111.71	108.99	118.78
Cluster 5					**56.63**	114.25	121.41	100.92	148.40
Cluster 6						**0.00**	133.49	122.83	91.90
Cluster 7							**0.00**	137.97	167.97
Cluster 8								**0.00**	130.95
Cluster 9									**0.00**

Bold diagonal values indicate intra-cluster distance; the remainder of values indicate the inter-cluster distance.

Table 3: Cluster means of 50 genotypes of okra

Cluster	PH	NPB	DFF	D50F	N1STF	D1H	NN	IL	FL	FD	NFS	FW	FYPP	PF	1000SW	OELCV90	BYVMV90
1	84.00	2.29	37.36	42.05	4.83	47.79	15.91	5.26	11.18	1.82	18.55	12.68	238.72	38.46	47.24	22.28	14.64
2	84.84	2.30	37.36	41.78	4.78	47.77	15.67	5.87	12.62	1.81	16.41	14.00	223.90	39.59	47.15	15.66	23.49
3	84.47	2.22	38.82	43.23	4.89	48.75	15.99	5.25	10.91	1.83	19.28	12.32	238.24	38.60	51.93	22.29	9.04
4	87.62	2.00	35.01	39.54	4.97	45.40	16.92	5.73	12.18	1.79	21.16	14.51	307.86	48.69	56.28	15.34	12.15
5	91.31	2.52	35.51	39.63	3.88	45.38	17.20	4.19	15.67	1.65	20.41	14.76	299.64	45.88	59.95	22.27	8.94
6	98.80	3.20	33.31	38.30	4.00	44.61	21.75	4.93	14.43	1.75	27.99	15.88	444.82	70.74	63.96	17.96	17.44
7	73.76	2.35	38.48	42.95	4.20	49.26	15.30	4.84	12.91	1.71	20.00	13.78	275.73	42.45	53.04	0.00	19.03
8	86.20	2.20	35.51	40.53	4.50	46.75	16.10	4.23	12.34	1.70	19.99	12.62	252.44	40.11	53.64	25.80	0.00
9	92.22	3.15	33.99	40.13	4.80	45.70	15.70	7.05	10.88	1.67	16.70	12.88	213.33	33.98	43.52	36.47	0.00

PH: Plant height, NPB: Number of primary branches, DFF: Days to 1st flowering, D50F: Days to 50 % flowering, N1stF: Node to 1st flowering, D1H: Days to 1st harvest
NN: Number of nodes/main stem, IL: Internodal length, FL: Fruit length, FD: Fruit diameter, NFs: Number of fruits/plant, FW: Fruit weight, FYPP: Fruit yield /plant
NRPF: Number of ridges/fruit, NSPF: Number of seeds/fruit, 1000SW: 1000 seed weight, OELCV90: PDI of OELCV at 90 DAS, BYVMV90: PDI of BYVMVV at 90 DAS

Table 4: Percentage contribution of each character towards divergence

Character	Singh statistic	Proportion percentage	Cumulative percentage
OELCV90	22940224.51	23.53	23.53
BYVMV90	16784671.89	17.22	40.75
FL	9032834.61	9.27	50.01
DFF	8771029.19	9.00	59.01
IL	7496128.33	7.69	66.70
FYPP	6249896.47	6.41	73.11
NPB	4224944.02	4.33	77.44
D50F	3957852.03	4.06	81.50
N1STF	3748597.22	3.85	85.35
NN	3635648.30	3.73	89.08
NSPF	2821843.30	2.89	91.97
1000SW	2433313.34	2.50	94.47
FD	2134182.86	2.19	96.66
FW	1399339.38	1.44	98.09
NFS	1034427.54	1.06	99.15
PH	458876.35	0.47	99.62
D1H	367430.00	0.38	100.00

PH: Plant height, NPB: Number of primary branches, DFF: Days to 1st flowering, D50F: Days to 50 % flowering, N1stF: Node to 1st flowering, D1H: Days to 1st harvest, NN: Number of nodes/main stem, IL: Internodal length, FL: Fruit length, FD: Fruit diameter, NFs: Number of fruits/plant, FW: Fruit weight, FYPP: Fruit yield /plant, NRPF: Number of ridges/fruit, NSPF: Number of seeds/fruit, 1000SW: 1000 seed weight, OELCV90: PDI of OELCV at 90 DAS, BYVMV90: PDI of BYVMVV at 90 DAS

Table 5: Results of principal component analysis (PCA) for 8 quantitative traits contributing to the divergence of okra

Principal component (PC)	Eigenvalue	percentage of variance	cumulative percentage of variance
PC$_1$	**2.91**	36.40	36.40
PC$_2$	**1.85**	23.16	59.56
PC$_3$	**1.39**	17.38	76.94
PC$_4$	0.71	8.85	85.79
PC$_5$	0.62	7.75	93.54
PC$_6$	0.32	3.95	97.48
PC$_7$	0.18	2.30	99.78
PC$_8$	0.02	0.22	100.00

Factor loadings due to Pcs

	PC$_1$	PC$_2$	PC$_3$	PC$_4$	PC$_5$	PC$_6$	PC$_7$	PC$_8$
PDI of OELCV at 90 DAS	-0.46636	0.01314	-0.80625	0.09339	0.19059	0.19835	0.21879	0.00622
PDI of YVMV at 90 DAS	-0.35406	-0.25958	0.83595	0.11684	0.10999	0.17581	0.22757	-0.00115
DFF	-0.56273	0.79791	0.13213	0.04286	-0.11255	-0.07623	0.02058	0.09212
IL	0.57326	0.39189	0.05813	-0.69489	0.02591	0.14772	0.09515	0.00156
FL	0.71250	0.36702	-0.00127	0.35115	-0.32938	0.35174	-0.04595	-0.00210
FYPP	0.88192	0.09878	-0.06516	0.18346	-0.13429	-0.29897	0.25912	-0.00048
NPB	0.53594	0.47911	0.11078	0.20732	0.65067	-0.00006	-0.06789	-0.00002
D50F	-0.59235	0.78807	0.06753	0.02652	-0.09122	-0.06798	0.03335	-0.09347

DFF: Days to 1st flowering, IL: Internodal length, FL: Fruit length, FYPP: Fruit yield /plant, NPB: Number of primary branches, D50F: Days to 50 % flowering
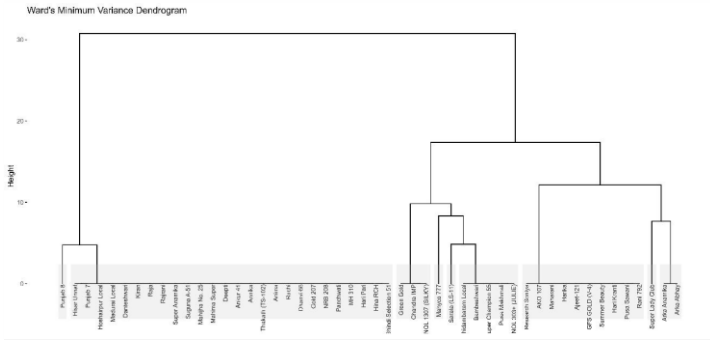
Figure 1: Wards minimum variance dendrogram
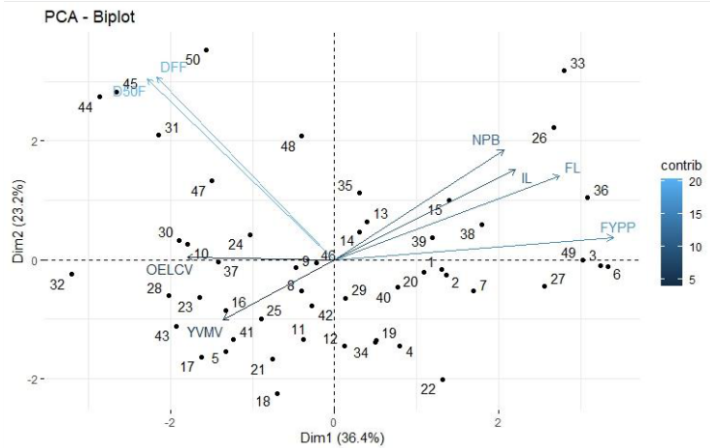
Figure 3. *Biplot diagram factor scores for the first and second components as determined by principal component analysis. Points in the diagram closest to the intersection of 0 on the X- and Y-axes indicate similarity. Outliers on the X-axis, that is, 33= AKO 107, 26 = Ajeet 121, 36 = Green gold, 6 = Punjab 8, 27 = Anima, 49= Hoshiarpur local and 3 = Hissar Unnat, indicate diversity. Numbers correspond to the name of the genotype.*

Figure 2. Percentage contribution of each trait on total divergence

Figure 4. PCA variables plot

Future Scope

In this study we have identified genotypes for YVMV and OELCV disease resistance noticed with low incidence values can be studied further in hot spot regions to verify their long-term viability.

Authors Contributions

Conceptualization of research work and designing of an experiment: Dr. Balagoni Maruthi, Dr. Sibsankar Das, Arup Chattopadhyay, Umesh Thapa, and Anirban Maji
Investigation: Dr. Balagoni Maruthi
Data Curation and Analysis: Dr. Balagoni Maruthi and Arup Chattopadhyay
Writing: Dr. Balagoni Maruthi
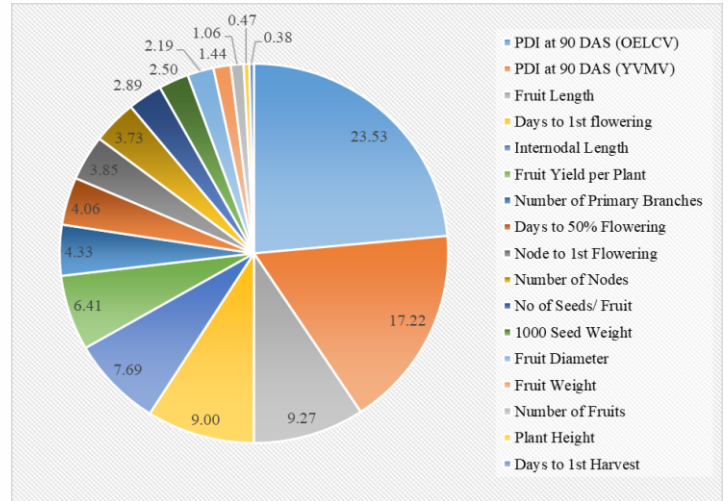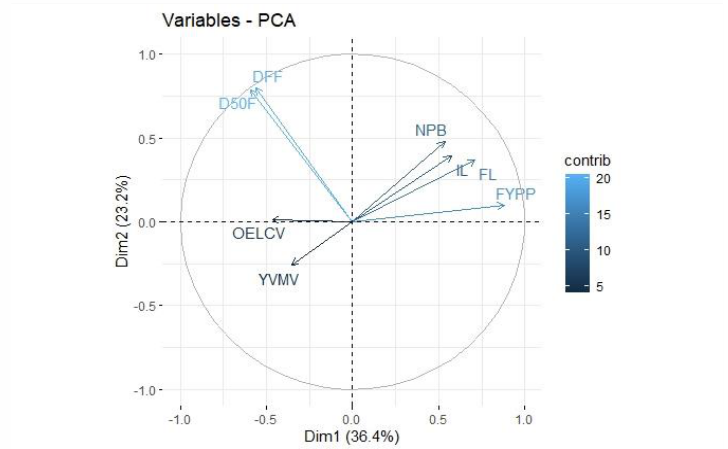Writing review and editing: Dr. Balagoni Maruthi and Adapa Kiran Kumar

Conflict of Interest: Nil

References

1. Kendre AH, Gholve VM, Navale MD (2023) Studies on the host range of okra enation leaf curl virus. *J Pharm Innov* 12(3):1950-1953.

2. Kumar D, Singh J, Pathania R, Dogra BS, Chandel VGS (2023) Revealing genetic diversity for the improvement of pod yield in okra (*Abelmoschus esculentus* L. Moench). *Electron J Plant Breed* 14(4):1497-1504.

3. Gemede HF, Ratta N, Haki G.D, Woldegiorgis A.Z, Beyene F (2015) Nutritional quality and health benefits of okra (*Abelmoschus esculentus*) A review. *J Food Process Technol* 6(458):p.2.

4. Anonymous, (2022) Horticulture statistics at a glance. Ministry of agriculture and farmers welfare (Department of Agriculture, Cooperation and Farmers Welfare) Krishi Bhawan, New Delhi.

5. Srivastava R, Meena K, Rai AK, Rai A, Singh S, Tiwari A, Singh N (2023) Appraisals of technical interventions to boost okra [*Abelmoschus esculentus* (L.) Moench] productivity in eastern Uttar Pradesh. *J Pharm Innov* 12(8):2387-2390.

6. Friedt W, Snowdon R, Ordon F, Ahlemeyer J (2007) Plant breeding: assessment of genetic diversity in crop plants and its exploitation in breeding. In *Progress in botany* (pp. 151-178). Berlin, Heidelberg: Springer Berlin Heidelberg.

7. Begna T, Begna T (2021) Role and economic importance of crop genetic diversity in food security. *Int J Agricultural Sci Food Technology* 7(1):pp.164-169.

8. Ikram-ul-Haq AAK, Azmat MA (2013) Assessment of genetic diversity in okra (*Abelmoschus esculentus* L.) using RAPD markers. *Pak J Agric Sci* 50(4):pp.655-662.

9. Pattan F, Rajan REB, Kumar CPS, Ruban JS (2023) Multivariate analysis for assessing genetic diversity in different genotypes of okra (*Abelmoschus esculentus* L. Moench) for varietal improvement. *J Appl Nat Sci* 15(3):1006-1011.

10. Chattopadhyay A, Dutta S, Bhattacharya I, Karmakar K, Hazra P (2007) Technology For Vegetable Crop Production. All India coordinated research project on vegetable crops. Directorate of Research, West Bengal, India.

11. Mahalanobis P (1936) On the generalized distance in statistics. Proceedings of the National Institute of Science, India. 12:49-55.

12. Rao CR (1952) Advance Statistical Methods in Biometrics. John Willey and Sons Pvt. Ltd., New York.

13. Seth T, Chattopadhyay A, Chatterjee S, Dutta S, Singh B (2016) Selecting parental lines among cultivated and wild species of okra for hybridization aiming at YVMV disease resistance. *J Agric Sci Technol* 18(3):751-762.

14. Das S, Mandal AB, Hazra P (2010) Genetic diversity in brinjal genotypes under eastern Indian conditions. *Indian J Hortic* 67:166-169.

15. Ranpise PS, Joshi VR, Patil BT (2018) Diversity studies in okra (*Abelmoschus esculentus* L. Moench.). *J Pharmacogn Phytochem* **7**(1):pp.1412-1414.

16. Mohammed J, Mohammed W, Shiferaw E, (2022) Performance and genetic variability of okra (*Abelmoschus esculentus* (L.) Moench) genotypes in Ethiopia for agromorphology and biochemical traits. *Adv Agric* pp.1-8.

17. Reddy JP, Anbanandan V, Kumar BS (2022) Genotypic, phenotypic variability and evaluation of okra [*Abelmoschus esculentus* (L.) Moench] genotypes for yield components. *J Appl Nat Sci* 14(1):pp.180-187.

18. Munshi R, Acharyya P, Chattopadhyay A, Mishra P, Sahoo B, Chakraborty L (2023) Diversity Analysis in Okra [*Abelmoschus esculentus* (L.) Moench] Genotypes under Saline Situation. *Biological Forum-An International Journal* 15(6):385-394.

19. Kalloo G, Singh VP, Dudi BS, Pratap PS (1980) Analysis of variation and genetic diversity in garden peas. *Haryana Agricultural University Journal of Research* 10:540-546.

20. Priyanka DV, Reddy MT, Begum H, Sunil N, Jayaprada M (2017) Genetic divergence analysis of inbred lines of okra [*Abelmoschus esculentus* (L.) Moench]. *Int J Curr Microbiol Appl Sci* 6(11):379-388.

21. Ramgiry M, Singh S, (2017) Genetic divergence analysis in okra (*Abelmoschus esculentus* (L.) Moench). *Int J Pure Appl Biosci* 5(2):981-986.

22. Badiger M, Rao PG, Pitchaimuthu M, Indiresh KM, Lingaiah HB, (2018) Multivariate assessment of yield and its component in okra [*Abelmoschus esculentus* (L.) Moench] genotypes. *J Agric Res* 55(1):140-143.

23. Abdou R, Zango O, Toudou AK, So TKA, Bakasso Y (2022) Agro morphological characterization and evaluation of three okra [*Abelmoschus esculentus* (L.) Moench] Varieties from Zinder (Niger) for yield and other Variability Components. *Agric Sci* 13(3):321-329.

24. Johnson RA, Wichern DW, (1988) Applied multivariate statistical analysis (2nd edn) Prcnticc-Hall: Enlglewood Cliffs, New Jersey, pp. 340-377.