J. Trop. Agric. and Fd. Sc. 38(2)(2010): 161–169

Classification of promising okra (*Abelmoschus esculentus)* genotypes based on principal component analysis

[Klassifikasi genotip okra (*Abelmoschus esculentus*) berasaskan analisis komponen utama]

R.K. Sharma* and K. Prasad**

Keywords: okra, *Abelmoschus esculentus,* correlation, variance, principal components

Abstract

Genetic diversity in 20 okra genotypes was determined using multivariate analysis. Various relationships were noticed among the attributes in the correlation studies. Positive significant correlations for plant height (PH) with number of branches (NB), days to 50% flowering (DF) with days to first harvest (DFH), pod yield per plant (PY) with pod yield per plot (PYP), number of pods per plant (NP) with PY and PYP ($p \leq 0.001$) and for NP with PH and NB ($p \leq 0.01$) were found, while pod weight (PW) was negatively correlated with NP ($p \leq 0.001$). Analysis of the extracted components, component patterns and Eigen values revealed that the first two principal components together accounted for 62.83% of the variance. The first component was found to be heavily loaded with PW, pod diameter (PD) and DFH in a positive direction, and NP, PY, PYP, PH and NB in a negative direction. Cluster analysis revealed that the selected genotypes could be grouped clearly into two groups accommodating 18 genotypes. The matrix obtained from this principal component analysis revealed that the genotypes Pb-57 and HRB-9-2 were in isolated positions in the third and fourth quadrants in the principal space.

Introduction

Abelmoschus esculentus (L.) Moench, commonly known as okra or bhindi in India is believed to be native to tropical Africa and belongs to the Malvaceae family. It is classified among the semi salt tolerant vegetable crops (Maas and Hoffman 1977). It is the only vegetable crop of economic importance in the Malvaceae family and cultivated throughout the tropics and subtropics (Kochhar 1986; Hammon and van Sloten 1989). It is mainly grown as a summer and rainy season (kharif) crop in India (Baloch 1994). The crop is highly nutritious and considered as a source of valuable nutrients. This vegetable provides protein, carbohydrate and significant amounts of vitamins and minerals, including calcium, which are often lacking in the diets of the people in developing countries.

The young tender pods of okra are used by different people in different ways as boiled vegetables (Magness et al. 1971), eaten fresh, canned, frozen or dried and used as a soup thickener (Tindall 1986) or cooked in curries, stews and soups. When ripe, the black or white-eyed seeds are sometimes roasted and used as a substitute for coffee

*Shivshankar Chemical Industry India Ltd., Bhagalpur – 812001, Bihar, India
**Department of Food Engineering and Technology, Sant Longowal Institute of Engineering and Technology, Longowal, 148 106 Punjab, India
Authors' full names: Ramesh Kumar Sharma and Kamlesh Prasad
E-mail: dr_k_prasad@rediffmail.com

(NRC 2006). It is considered as a rich source of dietary fibre. Nearly half of the okra pod is soluble fibre in the form of gums and pectin, which helps in lowering serum cholesterol (Jenkins et al. 2005) and thus reducing the risk of coronary heart disease (Jeff 2002). The other half is insoluble fibre, which helps to keep the intestinal tract healthy (Jeff 2002) and prevents the symptoms of irritable bowel syndrome. The mucilaginous extract of okra is often used to clarify sugarcane juice from which jaggary or brown sugar is manufactured (Prasad and Nath 2002). The stem of the plant provides non-digestible strong linear fibre, which finds uses in the paper, packaging and textile industries (Baloch 1994).

Stability with consistent performance under predictable and unpredictable environments is the main concern in the screening of phenotypically stable genotypes. Considerable effort is currently being made to improve different yield attributes of the okra such as the number of pods per plant, pod length and pod diameter. The weight of the pods and the number of pods per plant have been consistently identified as very important components of pod yield (Kaul et al. 1978). The number of pods per plant, days to flowering and plant height are some of the most variable quantitative characters of okra (Singh and Singh 1979). Genetic diversity is one of the important tools to qualify genetic variability in both cross and self-pollinated crops (Murty and Arunachalam 1966; Gaur et al. 1978; Hazara et al. 2002; Bendale et al. 2003; Sharma et al. 2008). Such a study also permits the selection of genetically divergent parent to obtain desirable recombinants in the segregating generations. Therefore, the present study was undertaken to analyse the genetic divergence in 20 okra genotypes and to classify them based on multivariate analysis.

Materials and methods

A total of 20 okra genotypes were collected from different agro climatic regions of India (*Table 1*). Seeds of the selected genotypes were planted in a randomized complete block design at three different sowing times, staggered uniformly with sowing start from 18 January 2008 in 20-day intervals until

Table 1. Name and source of okra genotypes

No.	Genotype	Source	Symbol
1	AROH-1	Ankur Seeds Pvt. Ltd., Nagpur	G1
2	BO-1	Orissa University of Agriculture and Technology, Bhubneshwar	G2
3	Sel.-2	National Bureau of Plant Genetic Resources, New Delhi	G3
4	Vaishali Vadhu	Bihar Agricultural College, Sabour	G4
5	H-1-87-16	Bihar Agricultural College, Sabour	G5
6	Sel.-7	Indian Institute of Horticultural Research, Bengaluru	G6
7	HOE-301	Hoechst, Mumbai	G7
8	Sel.-4	Indian Institute of Horticultural Research, Bengaluru	G8
9	HRB-55	Chaudhary Charan Singh Haryana Agricultural University, Hissar	G9
10	BO-2	Orissa University of Agriculture and Technology, Bhubneshwar	G10
11	Pb-57	Marathwada Agricultural University Parbhani	G11
12	Sel.-10	Indian Institute of Horticultural Research, Bengaluru	G12
13	KS-312	Chandra Shekhar Azad University of Agriculture & Technology, Kanpur	G13
14	HOE-202	Hoechst, Mumbai	G14
15	71-14	Bihar Agricultural College, Sabour	G15
16	D-1-87-5	Bihar Agricultural College, Sabour	G16
17	NDO-25	Narendra Dev University of Agricultural and Technology, Faizabad	G17
18	HRB-9-2	Chaudhary Charan Singh Haryana Agricultural University, Hissar	G18
19	Sel.-8	Indian Institute of Horticultural Research, Bengaluru	G19
20	Pusa Sawani	Indian Agricultural Research Institute, Delhi	G20

the last sowing on 27 February 2008, during the spring-summer season. Each trial was performed in triplicate plots with 20 plants in each plot. The plot was 1.8 m long and 1.5 m wide with spacing of 45 cm between rows and 30 cm between plants. Paths between replications were 90 cm wide, while subplots were spaced 30 cm apart with border length of 75 cm for each trial. The net experimental area was 162 m^2 out of 323.64 m^2 of the total experimental area. The soil was sandy loam in texture, having good fertility and was properly levelled and well drained. Cultivation procedure adopted for the present study was as recommended for okra cultivation (Chadha 2002). Standard techniques were adopted in data collection from five competitive plants for each genotype on 10 characters, viz. plant height (PH), number of branches per plant (NB), days to 50% flowering (DF), days to first harvest (DFH), number of pods per plant (NP), pod length (PL), pod diameter (PD), pod weight (PW), pod yield per plant (PY) and pod yield per plot (PYP).

Statistical data analysis

In order to determine any statistically significant effects due to genotypes, analysis of variance was carried out for a randomized complete block design. Descriptive statistics, principal components and cluster analysis for multivariate data were statistically analysed using MINITAB v 13.2, SPSS v 11.0 and Microsoft Excel v 2000 software packages. The principal component and hierarchical cluster analyses provided the characteristic patterns to classify the selected genotypes. Clustering of the samples was done according to Ward (1963), which is based on minimizing the loss of information from joining two clusters.

Results and discussion

Data on comparative performance of 20 okra genotypes for diversity in comparison with the overall individual mean response are presented in *Figure 1*. The analysis of variance revealed that the differences among the selected genotypes were significant for all the characters, indicating the presence of variability among them. The results (*Figure 1*) indicated that genotype Sel.-7, 71-14 and KS-312 (*Table 1*) had the tallest plant height (PH) with means of 82.14, 83.06 and 84.89 cm respectively. The highest value for number of branches (NB) was observed on KS-312, days to 50% flowering (DF) and days to first harvest (DFH) on Pb-57, pod length (PL) on Sel.-8, pod diameter (PD) on Sel.-2, pod weight (PW) on H-1-87-16 while HRB-9-2 had the highest number of pods per plant (NP), pod yield per plant (PY) and pod yield per plot (PYP) (*Figure* 1). These results may support the varietal improvement programme as the selected germplasm had a cross-pollinating system.

The correlation matrix of the 10 attributes is shown in *Table 2*. Nearly 70% of the correlation coefficients in the matrix had values over 0.20. The highest positive significant correlations could be observed for plant height (PH) with number of branches (NB), days to 50% flowering (DF) with days to first harvest (DFH), pod yield per plant (PY) with pod yield per plot (PYP), number of pod per plant (NP) with PY and PYP ($p \leq 0.001$) and NP with PH and NB ($p \leq 0.01$), while a negative correlation was observed for pod weight (PW) with NP ($p \leq 0.001$). The strong correlations are indicative of the genetic control, and feeble correlations are influenced by environmental factors and are yield-limiting (Subramanyan et al. 1995). The results are in agreement with earlier studies on *Hibiscus sabdariffa* L. (Banarjee et al. 1998).

The data set of the measurements was subjected to principal component analysis (PCA), which removed the highly inter-correlated nature of the prevalent variations. The initial statistics of Eigen analysis are given in *Table 3*. It can be seen that three principal components (PCs) appeared to account for 75.49% of the total variance in the data. According to the Kaiser Criterion (Kaiser 1960), only the first three PCs were

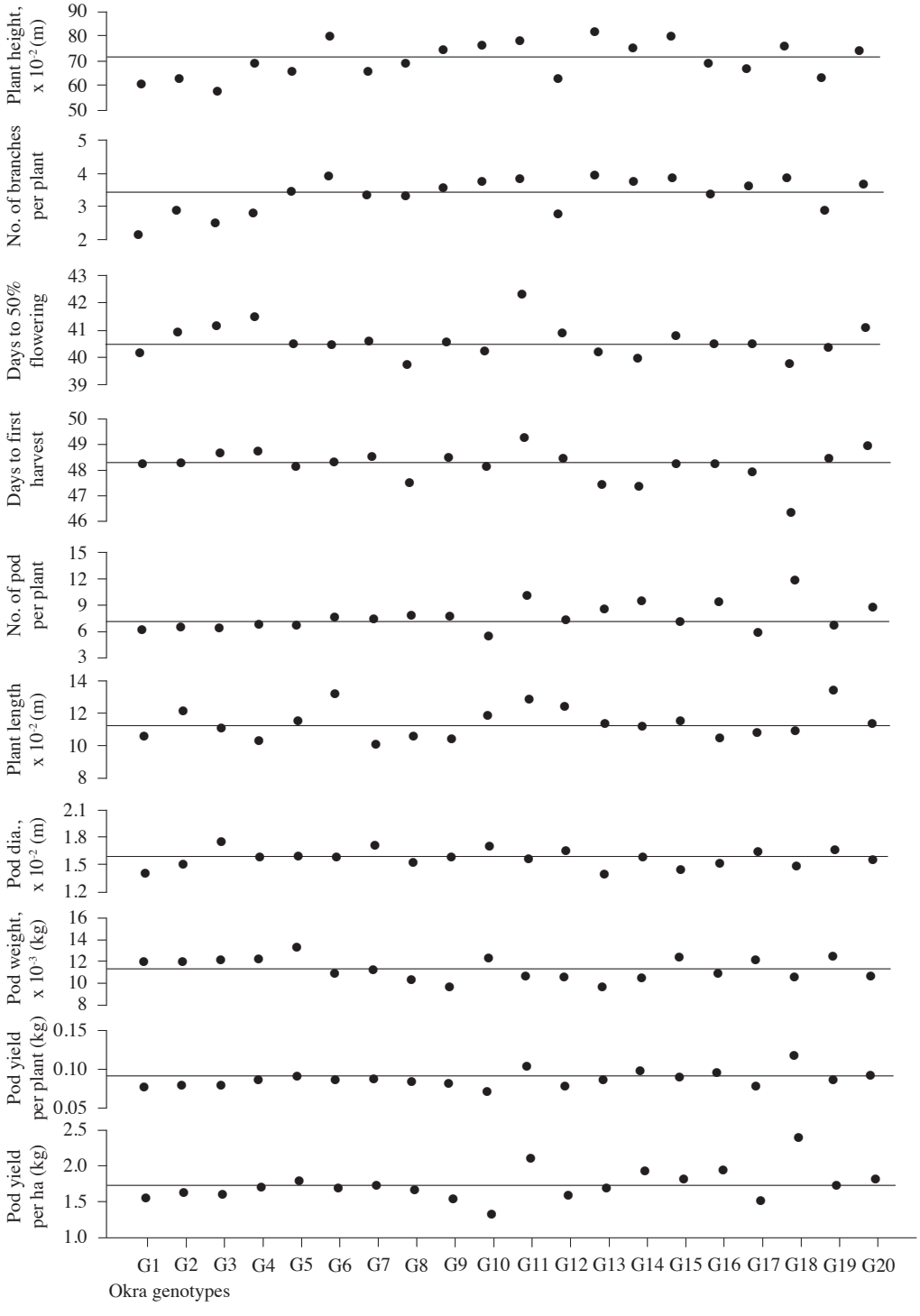

Figure 1. Performance of 10 characteristics of 20 okra genotypes

Table 2. Correlation matrix of performance variables of the okra genotypes

Variables	Plant height (PH)	No. of branches (NB)	Days to 50% flowering (DF)	Days to first harvest (DFH)	No. of pod per plant (NP)	Pod length (PL)	Pod diameter (PD)	Pod weight (PW)	Pod yield per plant (PY)	Pod yield per plot (PYP)
Plant height (I)	1.000									
No. of branches (II)	0.879***	1.000								
Days to 50% flowering (III)	0.060	0.074	1.000							
Days to first harvest (IV)	0.271	0.301	0.805***	1.000						
No. of pod per plant (V)	0.559**	0.515*	0.052	0.362	1.000					
Pod length (VI)	0.116	0.115	0.324	0.237	0.095	1.000				
Pod diameter (VII)	0.454*	0.229	0.235	0.422	0.432	0.087	1.000			
Pod weight (VIII)	0.489*	0.396	0.123	0.185	0.751***	0.171	0.313	1.000		
Pod yield per plant (IX)	0.360	0.359	0.031	0.376	0.862***	0.005	0.361	0.344	1.000	
Pod yield per plot (X)	0.421	0.423	0.052	0.336	0.880***	0.009	0.382	0.352	0.971***	1.000

Table 3. Eigen analysis of principal components

No.	Eigen value	Individual percentage
1	4.411	44.11
2	1.871	18.71
3	1.267	12.67
4	0.969	9.69
5	0.713	7.13
6	0.562	5.62
7	0.112	1.12
8	0.073	0.73
9	0.021	0.21
10	0.001	0.01

retained because subsequent Eigen values were all less than one. Hence, the reduced dimensionality of descriptor space was three.

The loading of components on the principal axes indicated a high degree of genetic diversity among the selected genotypes (*Figure 2*). To study the prevailing patterns for the attributes and to identify factors that were substantively meaningful, a multivariate approach was used. The loading plots in *Figure 2* are for the three principal components. It can be seen that the first component with 44.11% of the variance was highly correlated both positively (PD and PW) and negatively (PH, NB, NP, PY and PYP). Thus, it classified and distinguished the scores of the okra genotypes on the basis of these components. The second principal component with 18.71% of total variance was highly correlated with DF and DFH, while the third component with only 12.67% of total variance was strongly and positively correlated with PY and PYP, and negatively with PH and NB (*Figure 2*).

Figure 3 illustrates the biplots of scores and loadings of the three principal components of observed parameters on the okra genotypes, which further affected their positions according to their characteristics with respect to the extracted principal components. The biplot of PC1 and PC2 shows the separation of the okra genotypes according to their respective scores. The first quadrant of the plot contains samples having positive PC1 and PC2 scores. The

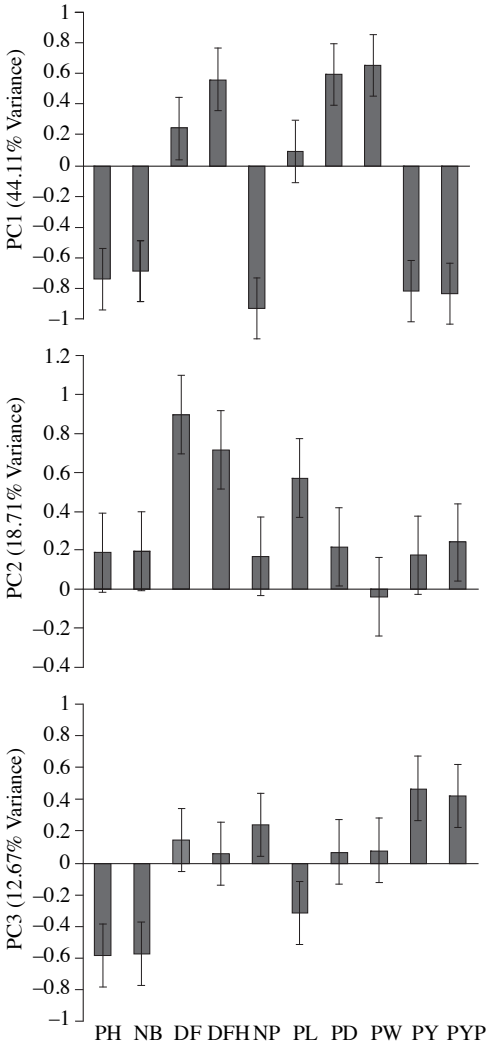

Figure 2. Loadings of 10 characteristics component on principal axes

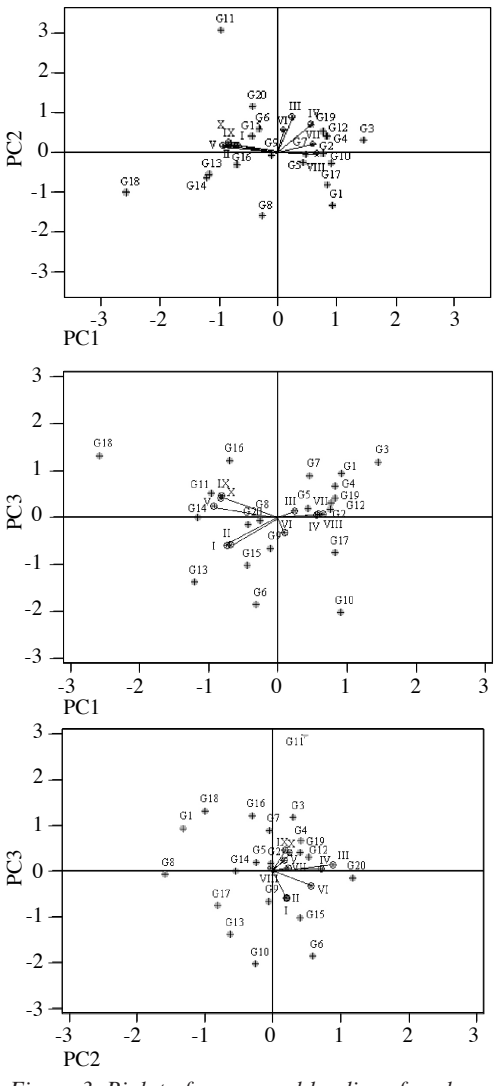

Figure 3. Biplot of scores and loadings for okra genotypes on principal axes 1, 2 and 3

genotypes of this quadrant were heavily loaded with days to 50% flowering (DF), days to first harvest (DFH), pod length (PL), pod diameter (PD) and pod weight (PW), and comprised eight genotypes in cluster I (*Table 4*). The second quadrant is composed of samples with negative PC1 and positive PC2 scores, and is heavily loaded mainly with the phenotypic and yield related characters. Similarly, the third quadrant contains the high-yielding promising genotype HRB-9–2 (G18) having negative PC1 and PC2 scores. Positive PC1 and negative PC2 scores are placed in fourth quadrant made up of less promising okra genotypes.

A graphical depiction of various characteristics of different okra genotype groups was obtained by means of the hierarchical cluster analysis (HCA) of standardized compositions according to Ward (1963) as an amalgamation rule, and squared Euclidean distances as the measure of proximity between samples. A dendogram

Table 4. Distribution of 20 okra genotypes into five clusters

Group/cluster No.	No. of genotypes	Genotypes in different cluster
1	10	AROH-1 (G1), BO-1 (G2), Sel.-2 (G3), Vaishali Vadhu (G4), H-1-87-16 (G5), HOE-3019 (G7), BO-2 (G10), Sel.-10 (G12), NDO-25 (G17), Sel.-8 (G19)
2	4	Sel.-7 (G6), HRB-55 (G9), 71-14 (G15), Pusa Sawani (G20)
3	1	Pb-57 (G11)
4	4	Sel.-4 (G8), KS-3 (G13), HOE-202 (G14), D-1-87-5 (G16)
5	1	HRB-9-2 (G18)

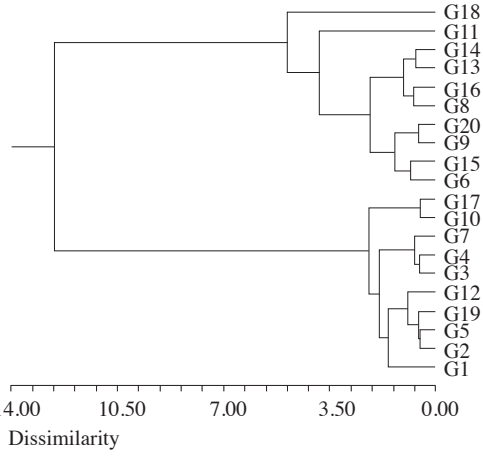

Figure 4. Dendogram of cluster analysis for selected okra genotypes

is shown in *Figure 4*. As a result of applying HCA to the principal component score matrix, the okra genotypes were grouped into five different clusters (*Table 4*). It was revealed that clusters 1, 2 and 4 have 10, 4 and 4 genotypes respectively. Cluster 3 and 5 had only one genotype each, i.e. the lowest. The genotype Pb-57, a high yielding and promising genotype also showed complete resistance to yellow vein mosaic virus (Bora et al. 1992). The clustering pattern of the okra genotypes revealed considerable genetic diversity among them by occupying five clusters. Thus, the above characterization of the okra genotypes on the basis of dissimilarity in scores with respect to the extracted principal components justifies the conclusion that variability existed.

Conclusion

It was concluded that genetic variation existed among the genotypes in all the characters studied. Genotypes Pb-57 and HRB-9-2 showed greater potential in terms of yield attributes as they outperformed the other cultivars, indicating their usefulness as promising genotypes. The crosses among the other clusters would probably exhibit high heterosis and may also produce new okra recombinants with the desired characters.

Acknowledgement

The authors would like to acknowledge the coordinators of the 'All India Coordinated Research Projects on Vegetables' for their timely support in completing the present work.

References

Baloch, M.A. (1994). Factors influencing the growth of okra. *Pakistan J. Sci. Res.* 82: 363–367

Banarjee, R., Sinha, M.K. and Roy, M.K.G. (1998). Correlation and path coefficient analysis of yield components in *Hibiscus sabdariffa* L. *Jute Development Journal* 8: 643

Bendale, V.W., Atanur, S.S., Bhave, S.C., Mehta, J.L. and Pethe, U.B. (2003). Genetic divergence for yield and yield components in okra. *Orissa J. Hort.* 3: 30–33

Bora, G.C., Saikia, A.K. and Shadeque, A. (1992). Screening of okra genotypes for resistance to yellow vein mosaic virus disease. *Indian J. Virol.* 8: 55–57

Chadha, K.L. (2002). *Hand Book of Horticulture*. New Delhi: ICAR Publication

Gaur, P.C., Gupta, P.K. and Kishore, H. (1978). Studies on genetic divergence of potato. *Euphytica* 27: 361–368

Hammon, S. and van Sloten, D.H. (1989). Characterization and evaluation of okra. The use of plant genetic resources, p. 173–174. (Town: Publisher??)

Hazara, P., Basu, D. and Sahu, F.K. (2002). Genetic divergence in okra. *Indian J. Hort*. 59: 406–410

Jeff, S. (2002). Growing okra, *Backyard gardener*. Retrieved in 2 June 2010 from http://ag.arizona.edu/yavapai/anr/hort/byg/archive/growingokra.html

Jenkins, D.J.A., Kendall, C.W.C., Marchie, A., Faulkner, D.A., Wong, J.M.W., de Souza, R., Emam, A., Parker, T.L., Vidgen, E., Trautwein, E.A., Lapsley, K.G., Josse, R.G., Leiter, L.A., Singer, W. and Connelly, P.W. (2005). Direct comparison of a dietary portfolio of cholesterol-lowering foods with a statin in hypercholesterolemic participants. *Am J Clin Nutrition* 81(2): 380–387

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement* 20: 141–151

Kaul, T., Lal, G. and Peter, K.V. (1978). Correlation and path coefficient analysis of components of earliness, pod yield and seed yield in okra. *Indian J. Agric. Sc*. 48(8): 459–63

Kochhar, S.L. (1986). *Tropical Crops: A text book of economic botany*. India: Macmillan Publisher

Magness, J.R., Marke, G.M and Compton, C.C. (1971). Food and feed crops of the United States. Inter regional Research project IR-4, IR-Bul. 1. *New Jersey Agric. Epxt. Station Bulletin* 1971: 828

Mass, E.V. and Hoffman, G.J. (1977). Crop salt tolerance – current assessment. *J. Irrig. Drain. Div. Am. Soc. Civil Eng*. 103: 115–134

Murty, B.R. and Arunachalam, V. (1966). The nature of genetic divergence in relation to breeding system in crop plant. *Indian J. Genet*. 26A: 188–198

NRC (2006). *Lost Crops of Africa: Volume II: Vegetables*, National Research Council. Washington D.C.: National Academy Press

Prasad, K. and Nath, N. (2002). Effect of pre-treatments and clarificants on sugarcane juice characteristics. *Asian Journal of Chemistry* 14(2): 723–731

Sharma, J.P., Singh, A.K., Kumar, S. and Sharma, N. (2008). Genetic divergence studies in okra *[Abelmoschus esculentus* (L.) Moench]. *Journal of Research, SKUAST-J* 7(1): 99–104

Singh, S.P. and Singh, H.N. (1979). Genetic divergence in okra *(Abelmoschus esculentus* L. Moench). *Indian J. Hort*. 36(2): 166–170

Subramanyan, D., Kumar, P.V. R., Krishnamurthy, B. and Islam, S. (1995). Heritability and correlation studies in Kenaf (*Hibiscus cannabinus* L.). *Indian J. Genet. Plant Breed*. 55: 279–282

Tindall, H.D. (1986). *Vegetable in the tropics,* 1st Edition. London: Macmillan Publishers

Ward, J.H. (1963). Hierachical grouping to optimize an objective function. *J. Am. Statist. Assoc*. 58: 236–244

Abstrak

Kepelbagaian genetik 20 genotip okra telah dikenal pasti menggunakan analisis multivariat. Kajian korelasi yang dijalankan telah menunjukkan beberapa perhubungan antara atribut yang dikaji. Korelasi positif yang signifikan telah diperoleh antara tinggi pokok (PH) dengan bilangan cabang (NB), hari untuk 50% berbunga (DF) dengan hari pertama menuai (DFH), hasil buah setiap pokok (PY) dengan hasil buah setiap plot (PYP), jumlah buah setiap pokok (NP) dengan PY dan PYP ($p \leq 0.001$) serta hubungan antara NP dengan PH dan NB ($p \leq 0.01$). Walau bagaimanapun, berat buah (PW) dengan NP mempunyai hubungan korelasi yang negatif ($p \leq 0.001$). Analisis komponen terpilih, pola komponen dan nilai Eigen menunjukkan dua komponen utama yang menyebabkan 62.83% keberlainan. Komponen pertama yang merangkumi PW, garis pusat buah (PD) dan DFH menjurus ke arah hubungan positif, sementara NY, PY, PYP, PH dan NB menjurus ke arah hubungan negatif. Analisis kluster mendedahkan genotip terpilih boleh dikumpulkan dalam dua kumpulan yang jelas merangkumi 18 genotip. Matriks yang diperoleh daripada analisis komponen utama ini menunjukkan kedudukan genotip Pb-57 dan HRB-9-2 terpencil dalam kuadran ketiga dan keempat ruang utama.

Accepted for publication on 1 December 2010