

GENETIC DIVERGENCE IN PUMPKIN (*Cucurbita moschata* L.) GENOTYPES

S. SULTANA¹, M. A. KAWOCHAR², S. NAZNIN³
H. RAIHAN⁴ AND F. MAHMUD⁵

Abstract

Genetic diversity using Mahalanobis's D^2 technique was studied for yield and its components on twenty one genotypes of pumpkin (*Cucurbita moschata* L.). Quantification of variability for each character was done using the Shannon Weaver Diversity Index. High degree of variation was exhibited within the collection, as reflected by mean diversity index value of 0.80. Data were subjected to principal component analysis (PCA), principal coordinate analysis (PCO), canonical variate analysis (CVA) and non-hierarchical clustering to identify suitable parents having distant relationship for hybridization program. The genotypes were grouped into five different clusters. Cluster IV contained the maximum number of seven genotypes whereas cluster I contained least number having only one genotype. The lowest inter-genotypic distance (0.75) was found between BD-2174 and BD-9489 where the highest (47.46) was between BARI Mistikumra-1 and BD-2150. The maximum inter cluster distance was observed between cluster II and III (17.922) and the minimum inter cluster distance was observed between cluster II and IV (6.825). The maximum intra cluster distance was noticed for the cluster V (0.261) and the minimum intra cluster distance was found in cluster I (0.00). Cluster I contained the highest mean values for pedicel length of male flower, number of male flowers/plant, fruit length, fruit breadth, single fruit weight and fruits/plant. Cluster II contained the highest mean values for days to first male and female flowering. Cluster III contained the highest mean values for leaf breadth, pedicel length of female flower and number of female flowers/plant. Leaf breadth, pedicel length of male flower, number of male flowers/plant and fruits/plant were the important components of genetic divergence in the studied materials. Based on inter cluster distance, inter genotypic distance and consideration of desirable characters for high yield potential, the genotypes G19 (BARI mistikumra-1) and G20 (BARI mistikumra-2) from cluster II; G21 (BD-2150) from cluster I and G1 (BD-2151) and G13 (BD-266) from cluster III can be selected as better parents for future hybridization program.

Keywords: *Cucurbita moschata* L., principal component analysis, cluster analysis, genetic divergence.

¹⁻³Scientific Officer (Plant Breeding), Tuber Crops Research Centre, Bangladesh Agricultural Research Institute (BARI), Gazipur-1701, ⁴Scientific Officer, Agricultural Research Station, BARI, Gazipur-1701, ⁵Professor, Department of Genetics and Plant Breeding, Sher-e-Bangla Agricultural University, Dhaka-1207, Bangladesh.

Introduction

Pumpkin (*Cucurbita moschata* L.) is locally known as 'Misti kumra' or 'Misti lau' or 'Misti kadu' and is considered to have originated from Central and North America (Whitaker and Davis, 1962). It is an under exploited popular vegetable but has higher demand in Bangladesh. It is relatively high in energy and carbohydrates and a good source of vitamins, especially high carotenoid pigments and minerals (Bose and Som, 1986). It may contribute to improve the nutritional status of the people, particularly the vulnerable groups in respect of vitamin A requirement. It becomes available even in the lean period when other vegetables are scarce in Bangladesh. Among the non-traditional crops, Bangladesh has been earning a handsome amount of foreign currency by exporting pumpkin to the U.K., Pakistan and Middle East (Alamgir, 1998). The total production of pumpkin is 0.218 million tons in a year in this country (BBS, 2011). Lack of high yielding, disease and pest tolerant variety is the main constrains towards its production. For developing a high yielding variety with desired characters needs good parents. The selection of potential good parents in a breeding program is based on the knowledge of genetic diversity amongst them. Evaluation of genetic diversity is important to know the source of genes for a particular character within the available germplasm (Tomooka 1991). To realize heterosis, genetically divergent parents are generally considered to be useful. In such crosses more variability could be expected in the resulting segregating progenies. Genetic divergence can be estimated by D^2 Statistic suggested by Mahalanobis and in turn is based on multivariate analysis of quantitative characters. The present study has been undertaken to generate information on genetic divergence in pumpkin so that the useful parental material for the breeding programs could be selected.

Materials and Method

The investigation was carried out at the experimental field of Sher-e-Bangla Agricultural University, Bangladesh during the period from March 2010 to August 2010 to study the genetic diversity in pumpkin. Twenty one genotypes of pumpkin were used in Randomized Complete Block Design (RCBD) with three replications. Those genotypes were assigned at random into pits of each replication. Pits of 55 cm x 55 cm x 50 cm were prepared in each plot with spacing of 3 m x 3 m. Number of pits/plot were 3. Standard package of cultural practices was followed for raising healthy crops. For studying different genetic parameters and inter-relationships, thirteen characters were taken into consideration like leaf length (cm), leaf breadth (cm), internodes distance (cm), days to first male flowering, days to first female flowering, pedicel length of male flower (cm), pedicel length of female flower (cm), number of male flowers, number of female flowers, fruit length (cm), fruit breadth (cm), fruit weight (Kg) and fruit yield/ plant (Kg).

Quantification of variability for each character was done using the Shannon-Weaver Diversity Index. Estimate of variability for each character was computed using the standardized Shannon-Weaver Diversity Index, designated as H' and has the formula:

$$H' = -\sum P_i(\log_2 P_i) / \log_2 n$$

Where, P_i is the proportion of the total number of genotypes belonging to the i^{th} class.

For each quantitative characters, the overall genotype means (\bar{x}) and standard deviation (σ) were used to subdivide the population values (x_i) into 10 frequency classes, ranging from class 1 (if $x_i \leq -2\sigma$) to class 10 (if $x_i \leq X+2\sigma$), the class interval being 0.5σ . The lowest and highest values were considered to determine the number of classes construct. The diversity considered high when $H' > 0.75$, moderate when $H' = 0.50 - 0.75$ and low when $H' < 0.50$. The Shannon-Weaver Diversity Index has a value ranging from 0 to 1, where 0 indicates absence of diversity and 1 indicates maximum diversity.

For both univariate and multivariate analysis, mean data for each character was used. In case of univariate analysis, analysis of variance was done individually by F test (Panse and Shukhatme, 1978) and MSTATC software was used for this purposes. Multivariate analysis viz. Principal Component Analysis (PCA), Principal Coordinate Analysis (PCO), Cluster Analysis, and Canonical Vector Analysis (CVA) were done by using Genestat 5.13 software program.

Result and Discussion

The analysis of variance showed significant variations among the genotypes for all the characters studied (Table 1).

Estimation of variation Using the Shannon- Weaver Diversity Index

Only two characters like days to first male flowering (0.73) and leaf breadth (0.74) exhibited medium variation while all the rest gave high diversity values. The computed diversity ranged from 0.73 (days to first male flowering) to 0.85 (leaf length) with a mean diversity value of 0.80 indicated existence of high variation within the collection (Table 2).

Cluster analysis

Based on cluster analysis, the twenty one genotypes were grouped into five clusters (Table 3). Cluster IV contained the maximum number of seven genotypes followed by cluster III, V and II having six, five and two, respectively, while cluster I contained least number of one genotype. In many cases, the same cluster included genotypes from different eco-geographic region indicating that

the geographic distribution and genetic divergence did not follow the same trend. The mean performances of thirteen characters in five clusters are shown in Table 4. Most of the characters showed distinct differences among the clusters. Cluster I contained the highest mean values for pedicel length of male flower (22.60), number of male flowers/plant (9.33), fruit length (74.40), fruit breadth (35.00), single fruit weight (3.58) and fruits/plant (7.15), whereas the lowest mean values for internodes distance (11.33). Cluster II contained the highest mean values for days to first male flowering (85.83), days to first female flowering (86.83) and the lowest mean values for leaf length (12.83), leaf breadth (18.33), pedicel length of male flower (8.23), pedicel length of female flower (2.09), number of male flowers/plant (1.00), number of female flowers/plant (1.00), fruit length (40.06), fruit breadth (18.03), single fruit weight (0.85) and fruits/plant (0.85). Cluster III got the highest mean values for pedicel length of female flower (5.52) and number of female flowers/ plant (4.78) while the lowest values for days to first male flowering (63.06) and days to first female flowering (67.78). Cluster IV contained the highest mean values for leaf breadth (25.77) and internodes distance (15.19).

Table 1. Range, Mean, Mean sum of square (MS_G), Percent coefficient of variation (CV) of 21 pumpkin genotypes.

Parameters	Range	Mean	MS _G	CV (%)
Leaf length without petiole (cm)	11.33-21.60	17.20	24.51**	5.23
Leaf breadth (cm)	15.20-30.67	23.39	38.32**	4.70
Internodes distance	10.00-20.00	14.17	29.12**	14.73
Days to first male flowering	60.00-86.33	70.38	150.44**	6.07
Days to first female flowering	65.67-87.00	75.17	135.82**	5.31
Pedicel length of male flower (cm)	8.07-26.47	16.67	105.05**	6.73
Pedicel length of female flower (cm)	2.07-8.63	4.45	9.94**	8.35
Number of male flowers/plant	1.00-15.33	6.77	40.94**	34.57
Number of female flowers/plant	1.00-7.00	3.53	11.34**	52.68
Fruit length (cm)	39.60-74.40	53.34	201.36**	9.87
Fruit breadth (cm)	17.33-35.00	27.59	65.34**	10.04
Single fruit weight (kg)	0.80-3.58	2.09	1.77**	17.02
Fruit yield/plant (kg)	0.80-9.92	4.01	21.00**	40.09

** Variation is significant at 1% level of probability.

Table 2. Computed diversity indices (H') for 13 different characters of 21 pumpkin genotypes.

Characters	H'
Leaf length without petiole (cm)	0.85
Leaf breadth (cm)	0.74
Internodes distance	0.81
Days to first male flowering	0.73
Days to first female flowering	0.82
Pedicle length of male flower (cm)	0.81
Pedicle length of female flower (cm)	0.83
Number of male flowers/plant	0.78
Number of female flowers/plant	0.81
Fruit length (cm)	0.84
Fruit breadth (cm)	0.75
Single fruit weight (kg)	0.84
Fruit yield/plant (kg)	0.82
Average	0.80

Table 3. Distribution of genotypes in different clusters.

Cluster no.	No. of Genotypes	No. of population	Name of genotypes
I	G21	1	BD-2150
II	G19, G20	2	BARI mistikumra-1, BARI mistikumra-2
III	G1, G2, G9, G10, G11, G13	6	BD-266, BD-2214, BD-2151, BD-2153, BD-2229, BD-2222
IV	G3, G4, G6, G14, G15, G16, G18	7	BD-2174, BD-2177, BD-2196, BD-9489, BD-9494, BD-9491, BD-9490
V	G5, G7, G8, G12, G17	5	BD-264, BD-2203, BD-2212, BD-9493, BD-4590

Principal Coordinate Analysis (PCO)

Inter-genotypic distances (D^2) were obtained from PCO for all possible combinations between pair of genotypes. Intergenotypic distances among the pumpkin genotypes ranged from 0.75 (between BD-2174 and BD-9489) to 47.46 (between BARI mistikumra-2 and BD-2150) (Table 5). The difference between the highest and lowest inter genotypic distance indicated the presence of variability among the twenty one pumpkin genotypes.

Table 4. Cluster mean values of 13 different characters of 21 pumpkin genotypes.

Characters	I	II	III	IV	V
Leaf length without petiole (cm)	17.67	12.83	18.40	17.83	16.53
Leaf breadth (cm)	21.87	18.33	23.47	25.77	22.29
Internodes distance	11.33	13.00	14.33	15.19	13.60
Days to first male flowering	70.67	85.83	63.06	69.91	73.60
Days to first female flowering	74.33	86.83	67.78	76.19	78.13
Pediceal length of male flower (cm)	22.60	8.23	20.86	12.47	19.71
Pediceal length of female flower (cm)	5.10	2.09	5.52	4.65	3.72
Number of male flowers/plant	9.33	1.00	7.72	7.05	7.07
Number of female flowers/plant	4.67	1.00	4.78	3.67	2.67
Fruit length (cm)	74.40	40.06	54.18	57.27	47.95
Fruit breadth (cm)	35.00	18.03	27.65	30.11	26.35
Single fruit weight (kg)	3.58	0.85	2.05	2.54	1.71
Fruit yield/plant (kg)	7.15	0.85	6.10	3.63	2.70

Table 5. Ten of each higher and lower inter- genotypic distance (D^2) among the 21 pumpkin genotypes.

Sl No.	Genotypic Combination		Ten Maximum (D^2) Values	Sl No.	Genotypic Combination		Ten Minimum (D^2) Values
01	BD-2150-	BARI mistikumra-2	47.46	01	BD-2174	- BD-9489	0.75
02	BD-2151 -	BARI mistikumra-2	46.39	02	BD-2177	- BD-2214	1.42
03	BD-2150-	BARI mistikumra-1	46.04	03	BARI mistikumra-1	- BARI mistikumra-2	1.57
04	BD-2151 -	BARI mistikumra-1	44.82	04	BD-9489	- BD-9490	2.13
05	BD-266 -	BARI mistikumra-2	43.86	05	BD-2174	- BD-9490	2.41
06	BD-2229 -	BARI mistikumra-2	43.04	06	BD-264	- BD-9491	2.92
07	BD-266 -	BARI mistikumra-1	42.30	07	BD-9490	- BD-9491	3.45
08	BD-2229 -	BARI mistikumra-1	41.46	08	BD-2196	- BD-2214	3.60
09	BD-2196 -	BARI mistikumra-2	38.98	09	BD-2151	- BD-2229	3.71
10	BD-2196 -	BARI mistikumra-1	37.42	10	BD-2174	- BD-9491	3.74

The intra-cluster distances were computed by using the values of inter-genotypic distances from distant matrix according to Sing and Chaudhury (1985). The

magnitude of the intra cluster distances were not always proportional to the number of genotypes in the clusters (Huque *et al.*, 2012), as the maximum intra cluster distance was noticed for the cluster V (0.261) followed by cluster III (0.177) and cluster IV (0.109). The minimum intra cluster distance was found in cluster I (0.00) followed by cluster II (0.01) (Table 6). This result supported by Gaffar (2008) and he reported that the genotypes were grouped into five clusters in which the highest intra cluster distance was noticed for the cluster II (0.999) and the lowest for the cluster IV (0.439).

Table 6. Intra (Bold) and inter cluster distances (D^2) for 21 pumpkin genotypes.

Cluster	I	II	III	IV	V
I	00.00	15.472	11.858	6.825	10.326
II		00.01	17.922	14.444	10.447
III			0.177	7.284	9.274
IV				0.109	6.450
V					0.261

Table 7. Eigenvalues and percentage of variation in respect of 13 principal components in 21 genotypes of pumpkin

Principal component axis	Eigen values	Percent variation	Cumulative % of Percent variation
1	5.735	44.12	44.12
2	2.066	15.89	60.12
3	1.684	12.95	72.96
4	1.176	9.05	82.01
5	0.812	6.24	88.25
6	0.558	4.29	92.54
7	0.328	2.53	95.07
8	0.233	1.80	96.87
9	0.185	1.42	98.29
10	0.137	1.06	99.35
11	0.035	0.27	99.62
12	0.033	0.25	99.88
13	0.018	0.13	100.00

Canonical Variate Analysis (CVA)

The inter-cluster distances were obtained from CVA. The inter cluster distances were larger than intra cluster distances which indicated that wider genetic diversity among the genotypes of different groups. The maximum inter cluster distance was observed between cluster II and III (17.922) followed by cluster I and II (15.472), cluster II and IV (14.444). The maximum inter cluster distance indicated that the genotypes belonging to cluster II were far away from those of cluster III. The minimum inter cluster distance was observed between cluster IV

and V (6.45) followed by I and IV (6.825) (Table 6). The results revealed that the crosses between the genotypes of cluster II and III would exhibit high heterosis and produce new recombinant with desirable traits.

Principal Component Analysis (PCA)

PCA is a statistical method which attempts to describe the total variation in multivariate sample using fewer variables than in the original data set (Bartolome *et al.*, 1999). The analysis results in the identification of the major attributes that are responsible for the observed variation within a given collection. From principal component analysis the values were found as 72.96% in the first three components and it was 82.01% in the four components of the total variance (Table 7). The two dimensional scatter diagram was prepared by using score component 1 in X axis and 2 in Y axis, showing the groups into five clusters among the genotypes which supported the result of cluster analysis (Fig 1).

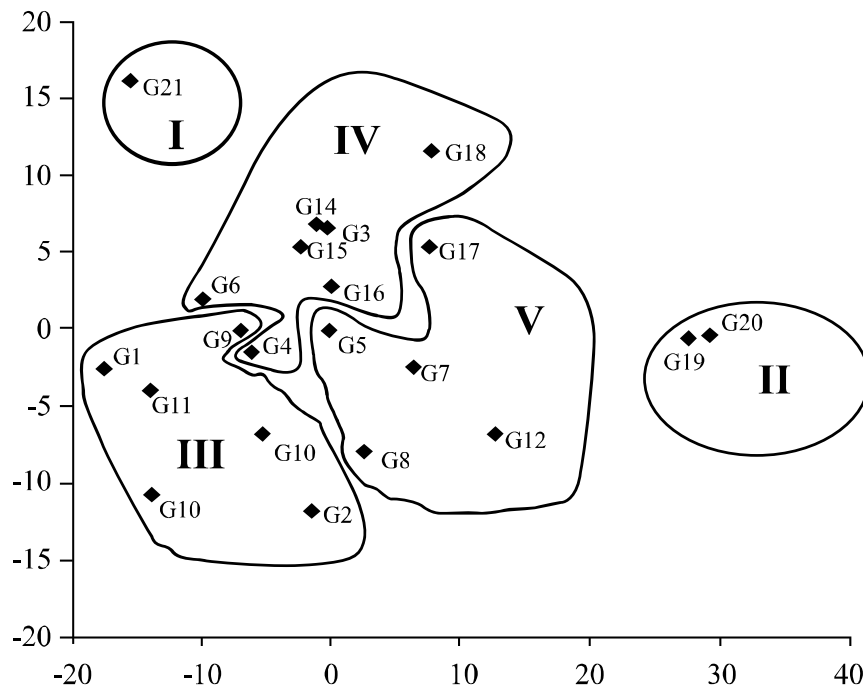


Fig. 1. Scattered diagram of 21 pumpkin genotypes of superimposed clusters

Contribution of characters towards divergence of the genotypes

Contribution of characters towards divergence of the genotypes is presented in table 8. The result of principal component analysis revealed that in vector I (Z_1), the important characters responsible for genetic divergence in the major axis of differentiation were leaf length, leaf breadth, days to first male flowering, pedicel

length of male flower, number of male flowers/plant, number of female flowers/plant, single fruit weight and fruits/plant which accounted for 44.12% of the total variation. In vector II (Z_2), which was the second axis of differentiation, leaf breadth, internodes distance, days to first female flowering, pedicel length of male flower, pedicel length of female flower, number of male flowers/plant and fruits/plant were important. Those characters contribute 15.89% of the total variation. The role of leaf breadth, pedicel length of male flower, number of male flowers/plant and fruit yield/plant in both the vectors was positive, which indicate the important components of genetic divergence in the collected materials. Divergence in the collected materials due to these four characters will offer a good scope for improvement of yield through rational selection of parents for producing hybrids. Banik (2003) reported that main vine length, first female flower node number, nodes on main vine, fruit length and number of seeds/fruit in snake gourd had the highest contribution towards the divergence.

Table 8. Relative contributions of the thirteen characters of 21 pumpkin genotypes to the total divergence.

Characters	Vector-1	Vector-2
Leaf length without petiole (cm)	0.5349	-0.1190
Leaf breadth (cm)	0.1333	0.0045
Internodes distance	-0.1342	0.1140
Days to first male flowering	0.9228	-0.3450
Days to first female flowering	-0.0185	0.1123
Pedicel length of male flower (cm)	0.1329	0.0509
Pedicel length of female flower (cm)	-0.5019	0.2415
Number of male flowers/plant	0.1079	0.2099
Number of female flowers/plant	0.3500	-1.9159
Fruit length (cm)	-0.0214	-0.2980
Fruit breadth (cm)	-0.4121	-0.1813
Single fruit weight (kg)	1.3447	-1.4068
Fruit yield/plant (kg)	0.0004	1.4252

Selection of genotypes

It is generally assumed that maximum amount of heterosis would be manifested in cross combinations involving the genotypes belonging to the most divergent clusters. Genotypes in cluster II if crossed with cluster I and cluster III might exhibit high heterosis as well as wide spectrum of genetic variation in F_2 generation. Based on inter cluster distance, inter genotypic distance and consideration of desirable characters for high yield potential, the genotypes G19

(BARI mistikumra-1) and G20 (BARI mistikumra-2) from cluster II; G21 (BD-2150) from cluster I and G1 (BD-2151) and G13 (BD-266) from cluster III may be considered better parents for future hybridization program.

References

- Alamgir, M. 1998. Export potential of non-traditional agricultural products (Bangladesh). *Bangladesh Quarterly*, **1**(3): 12-13.
- Bartolome, V. I., L. C. Quintana, A. B. Olea, L. C. Paunlagui, M. A. Ynalvez and C. G. Maclaren. 1999. Experimental design and data analysis for agricultural research. Volume 2. Training Documents. Biometrics Unit. International Rice Research Institute.
- Banik, B. R. 2003. Variability, gene action and heterosis in snake gourd (*Trichosanthes anguina* L.). Ph.D. thesis, Department of Genetics and Plant Breeding. Bangabandhu Sheikh Mujibur Rahman Agricultural University, Salna, Gazipur, Pp. 1-60.
- BBS (Bangladesh Bureau of Statistics). 2011. Statistical Year Book of Bangladesh 2010-2011. Bangladesh Bureau of Statistics Division, Ministry of Planning, Govt. of the People 's Republic of Bangladesh, Dhaka.
- Bose, T. K. and M. G. Som. 1986. Vegetables Crops in India. Naya Prokash, Calcutta, India, Pp. 92-95.
- Gaffar, A. 2008. Characterization and genetic diversity of sponge gourd (*Luffa cylindrica* L.). MS Thesis. Department of Genetics and Plant Breeding. Sher-e-Bangla Agricultural University, Dhaka.
- Huque, A. K. M. H., M. K. Hossain, N. Alam, M. Hasanuzzaman and B. K. Biswas. 2012. Genetic divergence in yardlong bean (*Vigna unguiculata* (L.) walp. Ssp. *Sesquipedalis verdc*). *Bangladesh J. Bot.* **41**(1):61-69.
- Panse, V.G., and Sukhatme, P.V. 1978. Statistical methods for agricultural workers. I.C.A.R., New Delhi.
- Singh, R.K. and B.D. Chaudhury. 1985. Biometrical methods in quantitative genetic analysis. Kalyani Publishers, New Delhi, India, P. 56.
- Tomooka, N. 1991. Genetic diversity and landrace differentiation of mungbean, *Vigna radiata* (L) Wilczek, and evaluation of its wild relatives (The subgenus *Ceratotropic*) as breeding materials. Tech. Bull., Trop. Res. Center, Japan. No.28. Ministry of Agriculture, Forestry and Fisheries, Japan, Pp. 1.
- Whitaker, T. W. and G. N. Davis. 1962. Cucurbits. Interscience Pub. INC. New York, Pp. 13.