Numerical Analysis of Variation Among Nigerian Accessions of 'Egusi' Melon (*Citrullus lanatus* (Thunb.) Matsum & Nakai)

E. O. Idehen*[1], O. B. Kehinde[1] and O. J. Ariyo[1]
[1] Department of Plant Breeding and Seed Technology, University of Agriculture
P. M. B. 2240, Abeokuta, Nigeria
emmaidehen@yahoo.com

Abstract: Fourteen quantitative characters were measured in 20 accessions of 'egusi' melon. Principal Component Analysis (PCA) and Single Linkage Cluster Analysis (SLCA) were employed to analyse the variation pattern in these accessions. The first three principal components accounted for 76.33% and 78.70% of the total variation in the early and late seasons respectively. SLCA summarized the relationship among the accessions at various levels of similarity into a dendogram while the accessions were sorted into six distinct groups. The implication of these results for 'egusi' melon breeding is discussed. [The Journal of American Science. 2007;3(2):7-15]. (ISSN: 1545-1003).

Keywords: *Citrullus lanatus;* 'egusi' melon; variation; Principal Component Analysis and Single Linkage Cluster Analysis

INTRODUCTION

In cross pollinated crops such as 'egusi' melon, germplasm often exist in the form of heterozygous genotypes which could be improved through intraspecific hybridization and selection. Development of improved varieties of crop plant necessarily involves the incorporation of specific gene complexes governing desired traits. Plant breeders firstly identify traits in plant species that makes them suitable for utilization by man, then search out new genes for the desired traits in new cultivars (or strains) using traditional techniques or biotechnology. They then finally access the performance of the new cultivars before releasing them as registered cultivars to be grown widely by farmers and enjoyed by consumers. It is important to classify the range of variability among accessions to facilitate the maintenance and further acquisition of germplasm resources. Thus, the wealth of any germplasm collection is measured in terms of the genetic variability of the crop species it contains.

Many techniques have been employed by researchers to determine the extent of variability in a germplasm collection. The Principal component analysis (PCA) has been used to measure genetic divergence among genotypes. This method which is a common ordination technique reduces the dimension of multivariate data by removing inter-correlations among attribute-variables (characters on which units are to be compared), and enables multidimensional relationships to be plotted on two or three principal axes (Hayman, 1967). PCA chooses independent or orthogonal axes, which are minimally correlated and represents linear combinations of the original characters (Clifford and Stephenson, 1975). The relative discriminating power of axes and their associated characters are measured by eigen-values and factor scores, respectively. The first principal component axis accounts for as much of the remaining variability as possible. Single Linkage Cluster Analysis (SLCA) is an agglomerative technique which shows the pattern of relationship between individuals of a population. The aim of this study was to identify the major characters responsible for the variation and to classify the variation pattern within the germplasm collection.

MATERIALS AND METHOD

The twenty accessions used in this study comprised fourteen accessions from National Institute for Horticultural Research (NIHORT) Ibadan, and six from different towns in Nigeria (Table 1). The field evaluation of these accessions were carried out at the Teaching and Research Farms, University of Agriculture, Abeokuta (Latitude 7.35^0N, 3.88^0E 450m asl) during the early (March) and late (August) growing seasons in 2005. Two-row plot was adopted for the while a Randomized Complete Block Design (RCBD) with three replications was used. A block consisted of 40 rows and planting was done in 6- meter-long rows. The rows were 1 meter apart while the plant – to – plant distance was also 1 meter. Two seeds of each accession was planted per hole and later thinned to 1 plant per stand. Each row therefore contained seven plants and five competitive plants within the row were observed. Manual weeding was carried out when necessary.

Data were collected on the following characters: Days to germination, days to flowering, fruit circumference, fruit weight, number of branches per plant, vine length, number of fruits per plant, number of seeds per fruit, seed weight per fruit, 100- seed weight, days to maturity and seed yield.

Cluster analysis was performed by the Genstat computer program (Genstat, 1993) at the International Institute of Tropical Agriculture (IITA) Ibadan, Nigeria. A similarity matrix was developed from the data matrix by comparing each accession with every other one. Principal component analysis produced an eigen vector for each principal axis. The character loadings were used to determine the accession component scores. Using the character variation, SLCA summarized the position of accessions into dendogram at intervals of 5 % level of similarity.

Table 1. List and sources of the accessions used for the study

SERIAL NUMBER	ACCESSION NAME	SOURCE
1	V2	NIHORT, Nigeria
2	131DA	NIHORT, Nigeria
3	DL99/46	NIHORT, Nigeria
4	DL99/71	NIHORT, Nigeria
5	DL99/75	NIHORT, Nigeria
6	DL99/76	NIHORT, Nigeria
7	DD95/549	NIHORT, Nigeria
8	DD98/3	NIHORT, Nigeria
9	DD98/4	NIHORT, Nigeria
10	DD98/7	NIHORT, Nigeria
11	DD98/506	NIHORT, Nigeria
12	DD98/511	NIHORT, Nigeria
13	DD98/533	NIHORT, Nigeria
14	DD98/550	NIHORT, Nigeria
15	L1	OKENE, Nigeria
16	L2	MINNA, Nigeria
17	L3	BENIN, Nigeria
18	L4(PAPA SAKI)	SAKI, Nigeria
19	L5(SOFIN 'BIG')	ABEOKUTA, Nigeria
20	L6(SOFIN 2)	SAKI, Nigeria

*NIHORT: National Institute of Horticultural Research, Ibadan, Nigeria

RESULTS

Principal Component Analysis (PCA) produced an eigen vector for each principal component axis. In the early season, only two of the fourteen principal component axes had eigen values greater than 2.0 and altogether accounted for 64.84% of the total variation and PCA 1, PCA 2 and PCA 3 accounted for 34.56%, 30.28% and 11.50% of the variation respectively (Table 2). While in the late season, three of the fourteen principal axes had eigen values greater than 2.0 and altogether accounted for 78.70% of the total variation while PCA 1, PCA 2 and PCA 3 accounted for 37.48%, 23.66% and 17.56% respectively (Table 2). The scores of the major characters describing the first three principal axes in the early and late seasons are presented in Table 3 and Table 4 respectively. The arithmetic sign of the coefficients is irrelevant since a common rule of thumb for determining the significance of a character coefficient is to treat coefficients greater than 0.3 as having a large effect to be considered important (Raji, 2002). Characters having less than 0.2 coefficient value were considered to be of no effect to the over-all variation observed in this study.

In the early season, the first principal component which accounted for the highest proportion (34.56%) was mostly correlated with characters such as fruit circumference, seed weight per fruit, vine length per plant, number of fruits per plant and fruit weight. Second principal axis was dominated by days to germination, days to flowering and days to maturity. The third principal component was dominated by fruit circumference, fruit weight per plant, number of branches per plant and 100- seed weight (Table 3).

In the late season, the first principal component which accounted for the highest proportion (37.48%) was mostly correlated with characters such as number of fruits per plant, fruit circumference per plant, vine length per plant, days to germination and days to maturity, while the second principal component was dominated by fruit weight, fruit circumference and seed weight per fruit. The third principal component was dominated by fruit weight per plant, days to flowering, number of seeds per fruit and 100- seed weight (Table 4).

A plot of accessions on axes 1 and 2 (Fig. 1A) showed that L5, L2, L4 and DD98/3 were the most distant from all other accessions. Figure 1B, graphing axes 1 and 3, illustrated the distinction between L5, L2, L4 and DD98/3 and their relationship with other accessions. L5 was also described by characters associated with principal component one. Figure 1C, showed another configuration of the accessions (axes 2 and 3) with L5 still distant from all other accessions.

Table 2. Eigen values and percentage of total variation accounted by the first four principal component axes of the ordination of 'egusi' melon accessions in both seasons

Principal axis	Season	Eigen value	Total variation accounted for (%)	Cummulative Percentage
I	ES	4.83	34.56	34.56
	LS	5.25	37.48	37.48
II	ES	4.23	30.28	64.84
	LS	3.31	23.66	61.14
III	ES	1.60	11.50	76.33
	LS	2.46	17.56	78.70
IV	ES	0.97	6.91	83.24
	LS	1.10	0.09	86.56

ES-Early Season
LS-Late Season

Table 3. Scores of the major characters of the first principal components used in ordination in the early season

Axis I	Axis II	Axis III
Fruit circumference per plant (0.40)	Days to flowering (0.42)	Number of branches per plant (0.48)
Seed weight per fruit (0.38)	Days to maturity (0.41)	100- seed weight (0.45)
Vine length per plant (0.33)	Days to germination (0.34)	Fruit circumference (0.43)
Fruit weight (0.32)		Fruit weight per plant (0.41)
Number of fruits per plant (0.28)		

Table 4. Scores of the major characters of the first principal components used in ordination in the late season

Axis I	Axis II	Axis III
Fruit circumference per plant (0.39)	Fruit weight (0.47)	Fruit weight per plant (0.44)
Number of fruits per plant (0.39)	Fruit circumference (0.40)	Days to flowering (0.43)
Vine length per plant (0.36)	Seed weight per fruit (0.36)	Number of seeds per fruit (-0.31)
Days to germination (-0.32)	Number of branches per plant (-0.27)	100- seed weight (0.25)
Days to maturity (-0.30)		

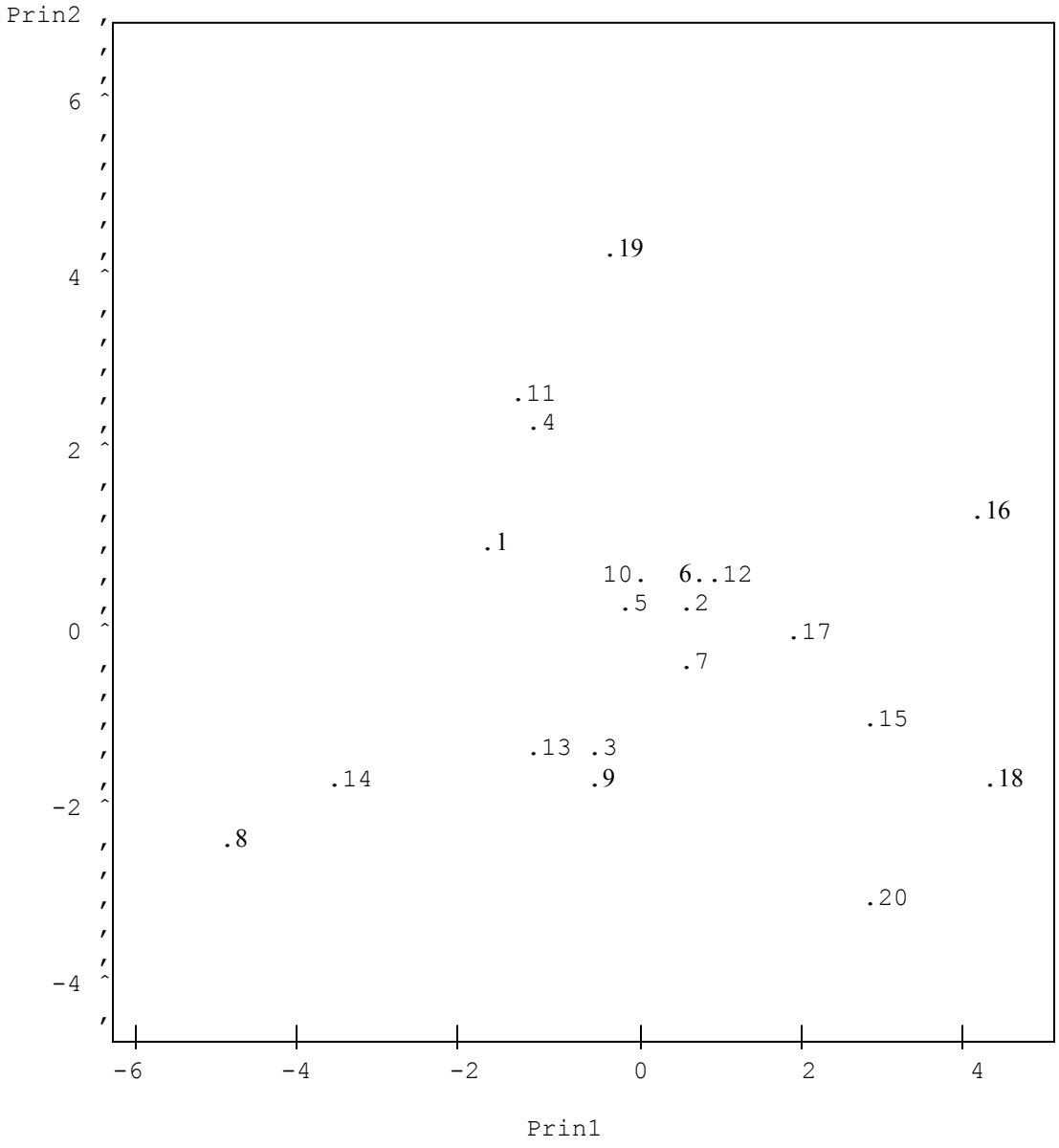

Figure 1A. Configuration of the twenty accessions under axis 1 and 2

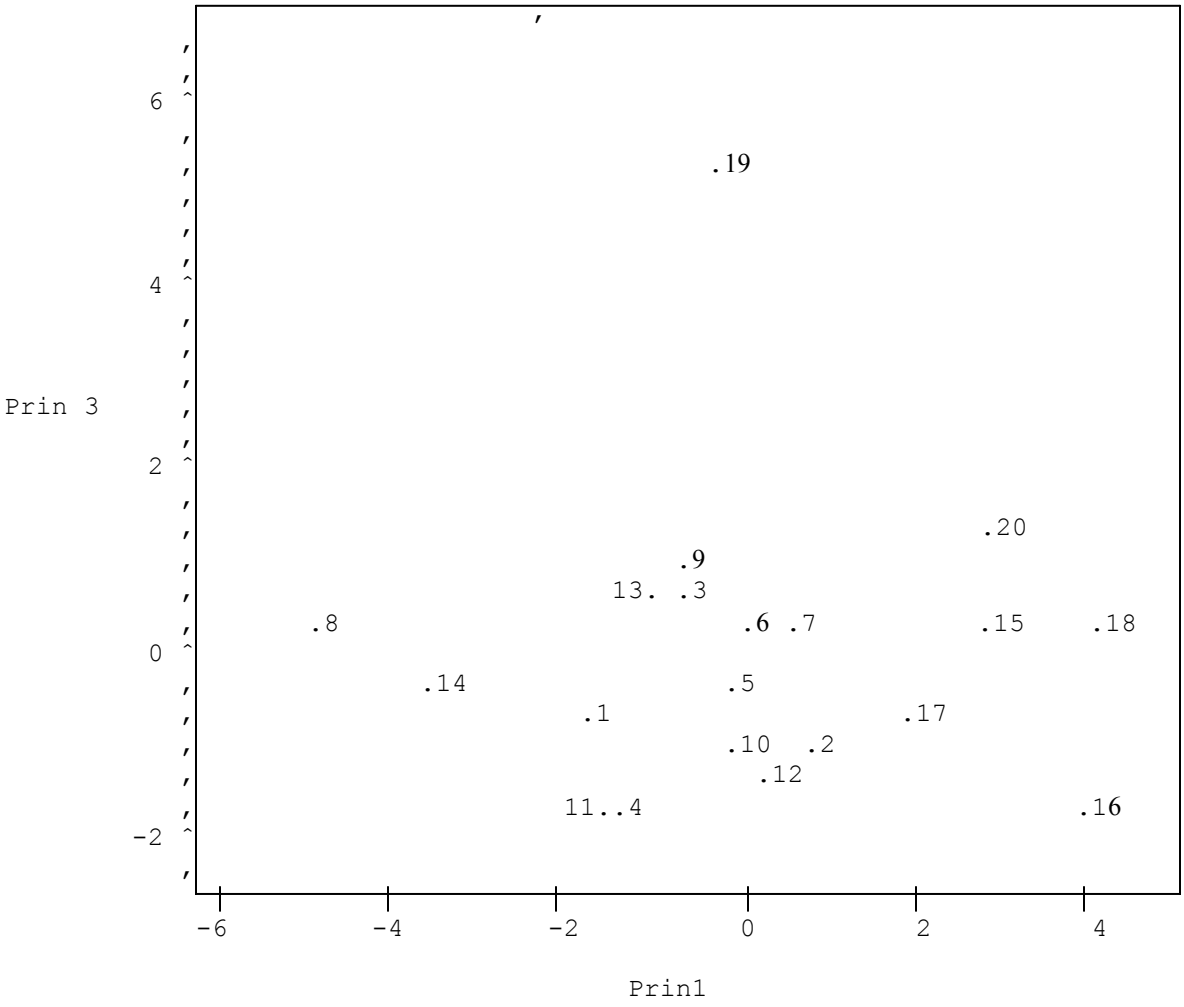

Figure 1B. Configuration of the twenty accessions under axis 1 and 3

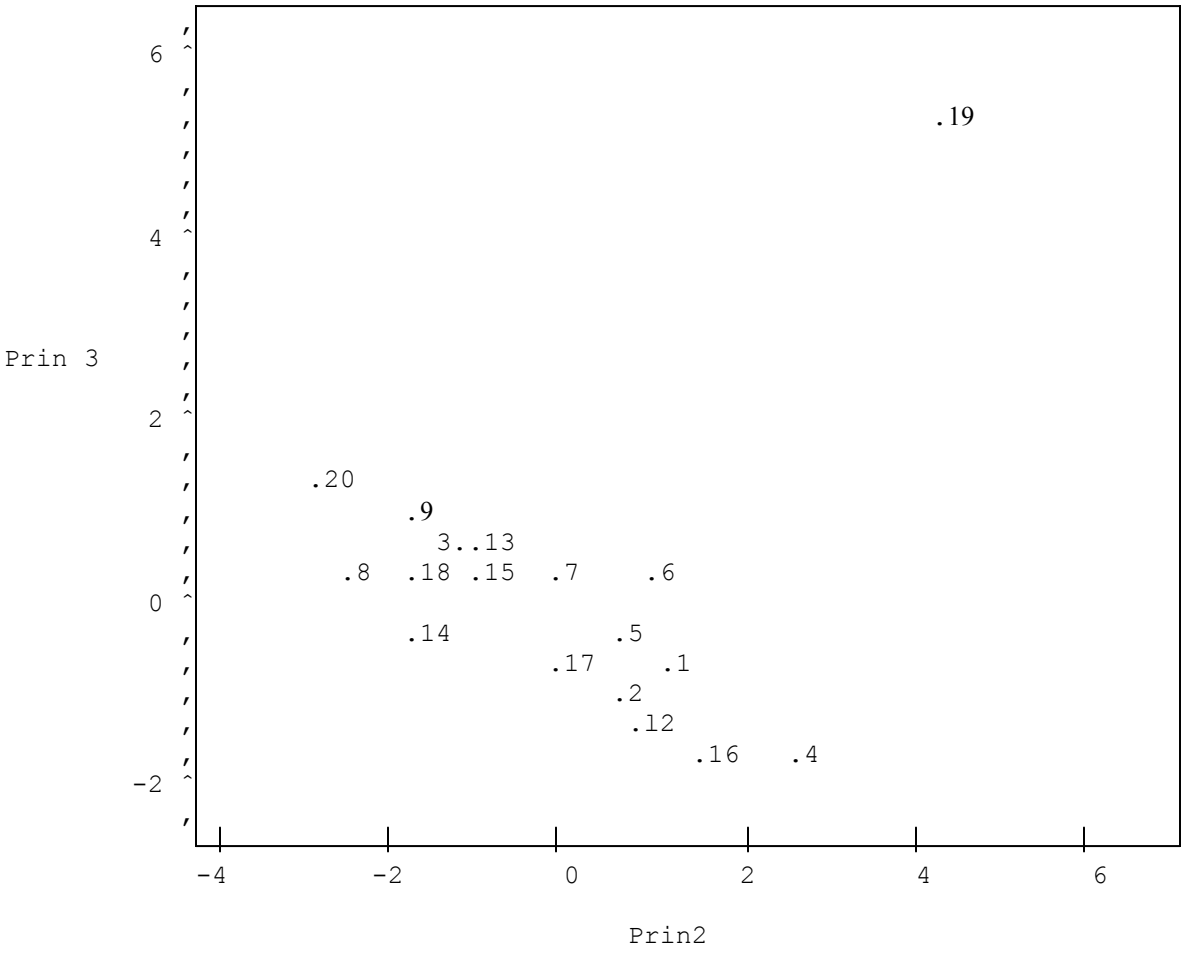

Figure 1C.Configuration of the twenty accessions under axis 2 and 3
Note: 2 observations hidden (10 and 11)

Figure 2, shows the dendogram drawn from Single Linkage Cluster Analysis (SLCA) to illustrate the relationship between the twenty accessions. At a minimum distance of 0.00 level of similarity, all twenty accessions were distinct from each other, while at a distance of 2.0 all had formed a single cluster indicating that accessions had at least one neighbor with more than 2.0 similarity level. At a distance of 0.3 accessions131DA, DL99/75, L3 and DL99/76 had been joined to accessions DL99/71, DD98/7, DD98/506 and DD98/511. At a distance of 0.5 level of similarity, 18 of the accessions had formed a single cluster but at a distance of 2.0, only accession L5 could be distinguished from the rest of the population.

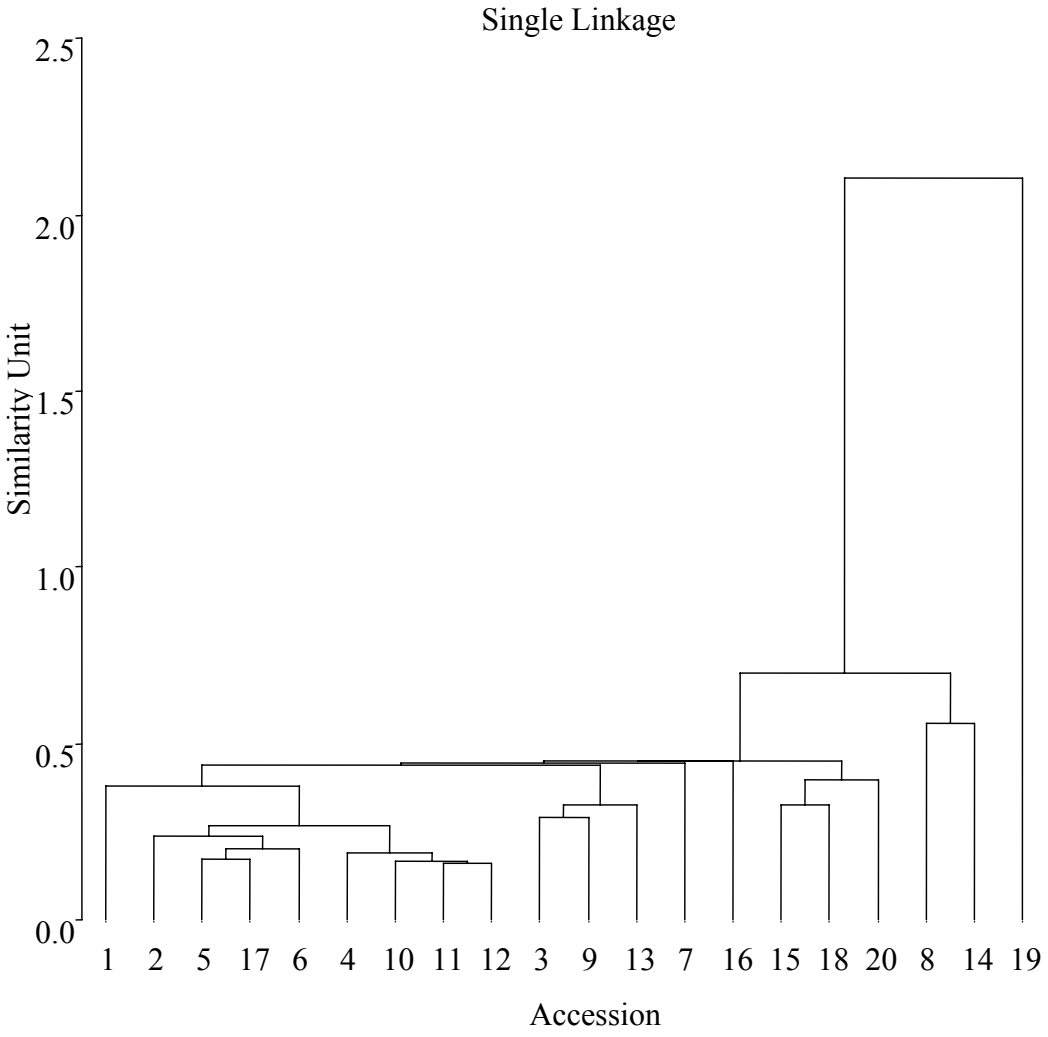

Figure 2. Dendogram resulting from Single Linkage Cluster Analysis (SLCA) of twenty accessions of 'egusi' melon.

DISCUSSION

The result from the Principal Component Analysis (PCA) identified three axes to have accounted for 76.33% and 78.70% of the total variation in the early and late seasons respectively for the twenty accessions of 'egusi' melon (*Citrullus lanatus* (Thunb.) Matsum. & Nakai) evaluated. In the early season, fruit circumference per plant, seed weight per fruit, vine length per plant and number of fruits per plant were identified as major factors that contributed 34.56% of variation in PCA 1. While in the late season, number of fruits per plant, fruit circumference per plant, vine length per plant, days to germination and days to maturity accounted for 37.48% of variation in PCA 1. Two dimension ordination of the 20 accessions drawn based on combinations of the first, second and third principal axes, revealed accession L5 distant from all others. One basic assumption in the use of numerical analysis is that numerical evaluation

of overall similarity is a measure of genetic similarity (Sneath and Sokal, 1973). The SLCA revealed a range of genetic diversity among the accessions. It is significant that accession L5 formed a distinct cluster.

It would be suitable to use quantitative characters mostly especially number of days to flowering as criterion for grouping these accessions since number of days is always correlated with life span (Ariyo, 1993). The importance of flowering behaviour in numerical taxonomic studies has already been highlighted by Sneath and Sokal (1973). Accession L4 is early maturing and high yielding and becomes an obvious parent for hybridization whenever breeding for early maturity is the objective.

Correspondence to:
E. O. Idehen*
Department of Plant Breeding and Seed Technology, University of Agriculture
P. M. B. 2240, Abeokuta, Nigeria
emmaidehen@yahoo.com

Received: 5/21/2007

REFERENCES

1. Akoroda, M.O. 1983. Principal component analysis and metroglyph of variations among Nigerian yellow yams. *Euphytica* 32:565 – 573
2. Ariyo, O. J. 1993. Genetic diversity in West African okra (A. caillei). Multivariate analysis of morphological and agronomic characteristics. *Genetic Res. and Crop Evolution* 40: 25-32
3. Clifford, H. T. and Stephenson, W. 1975. An Introduction to Numerical Classification. Academy Press, London, New York, xii, 229pp.
4. Genstat, 1993. Genstat Five Release Three. Reference Manual Oxford University Press, Walton Street, Oxford OX2 6DP
5. Hayman, H.H. 1967. Modern Factor Analysis. 2nd edition. University of Chicago Press, Chicago pp 474.
6. Raji, A. A. 2002. Assessment of genetic diversity and heterotic relationship in African improved and local cassava (*Manihot esculentus* Crantz) germplasm. PhD Thesis. University of Ibadan, Nigeria.
7. Sneath, P. M. and R. P. Sokal, 1973. Numerical Taxonomy. The principle and practice of numerical classification. San Francisco WH Freeman 573pp.
8. Ward, J. H. 1963. Hierarchical grouping to optimize an objective function. *Journal of American Statistical Association* 58: 236-244