Journal of Spices and Aromatic Crops 6 (1) : 21-29 (1997)

Morphometrical studies on black pepper (*Piper nigrum* L.). II. Principal component analysis of black pepper cultivars

P N RAVINDRAN, R BALAKRISHNAN[1] & K NIRMAL BABU

Indian Institute of Spices Research
Marikunnu P.O., Calicut - 673 012, India.

Abstract

Forty four cultivars and seven wild accessions of black pepper (*Piper nigrum*) (51 Operational Taxonomic Units) were analysed adopting Principal Component Analysis to determine the nature and extent of divergence among them. Eight principal components emerged from the analysis and the dispersion pattern of Operational Taxonomic Units between the principal components (in principal component plots) showed the characters involved in the divergence of various cultivars. Dispersion of Operational Taxonomic Units between principal components showed that certain cultivars remain as independent entities thereby indicating their divergence from other cultivars. The majority of cultivars are distributed around the central point, and this group is comparable to the group D obtained by the centroid clustering. Computation of inter and intracluster - D^2 helped to establish further the extent of relationships among the clusters.

Key words : black pepper, morphometrics, *Piper nigrum*, Principal Component Analysis.

Abbreviations

OTU : Operational Taxonomic Unit

PC : Principal Component

PCA : Principal Component Analysis

Introduction

Cluster analysis of 51 black pepper (*Piper nigrum* L.) accessions (51 OTUs) led to 11 groups and 28 of them formed a single cluster (Ravindran *et al.*, this issue). Some of the cultivars were unique and formed independent groupings while a few others formed clusters of two or three. The study was then extended to find out the relative contribution of various characters in differentiating the cultivars by PCA. Forty four

[1]Present address : Sugarcane Breeding Institute, Coimbatore - 641 007, India.

cultivars (which included most of the common ones) and seven wild collections (51 OTUs) were made use of in this study. The philosophy of PCA has been outlined by Sneath & Sokal (1973) and Williams (1976) and it is one of the most useful numerical procedures used in taxonomic research since the degree of relatedness among individuals between or within species, or among genotypes within a breeding population can be approximated with principal components. This analysis can resolve groups of morphologically similar phenotypes when plotted against principal components. Adams (1977) used PCA to calculate Euclidean distances between cultivars of dry bean and these distances were found to be highly correlated with their genetic relationships. PCA was also found useful in establishing the relationship among character variation within a species to area of origin (Martin & Adams 1987; Hilling & Iezzoni 1988). Iezzoni & Pritts (1991) have outlined the usefulness of PCA in horticultural research.

Materials and methods

The PCA was carried out using 22 characters recorded from 51 OTUs (for details of OTUs and characters, see Ravindran *et al.,* this issue). The analysis was carried out using the BMDP programme package (BMDP 4M-PCA) developed at the Department of Biomathematics, University of California, USA, based on the principle outlined by Frane *et al.* (1981). The computer analysis was done at the Computer Centre of the Carneigie-Mellon University, Pittsburgh, USA.

Results and discussion

The PCs along with the variance explained by each PC and the cumulative proportion of variance explained by each PC are given in Table 1. The loadings for eight major PCs after rotation and rearrangement are given in Table 2. There is a steady increase in variance explained up to first eight PCs only; more then 75 per cent of the variation in data are accountable by the first eight PCs. The relative contributions of other characters are negligible, and hence can be ignored for all practical purposes. Tables 1 and 2 also show that the first eight PCs explain the whole variance in PC space and hence only these eight are important. As regards the first PC, the loadings are high for the set of characters leaf size index, leaf breadth and leaf length (characters 4, 2 and 1). In the average linkage cluster analysis (Ravindran *et al.,* this issue) also these characters are in one group, thereby confirming the results of cluster analysis. The second PC shows high loadings for characters leaf thickness, lower epidermal thickness and upper epidermal thickness (characters 5, 6 and 7). This also agrees with the results of the cluster analysis. The third PC shows high loadings for the set of characters 14, 12 and 13 (leaf length, spike length index, spike length and peduncle length) and hence these three characters are represented by PC-3, though for character 14 the loading is negative. In this way eight PCs could be recognised as given below:

PC-1 : Leaf size index, leaf length, leaf breadth

PC-2 : Leaf thickness, lower epidermal thickness, upper epidermal thickness

PC-3 : Leaf length-spike length index, spike length and peduncle length

Table 1. Variance explained and the cumulative proportion of variance* in black pepper cultivars

PC	Variance explained	Cumulative proportion of variance	
		Data space	PC space
1	4.0102	0.1823	0.2437
2	2.8760	0.3130	0.4184
3	2.4823	0.4258	0.5692
4	1.9442	0.5142	0.6874
5	1.4796	0.5815	0.7773
6	1.3124	0.6411	0.8570
7	1.2439	0.6979	0.9329
8	1.1038	0.7481	1.0000
9	0.9272	0.7902	
10	0.8559	0.8292	
11	0.6409	0.8583	
12	0.6075	0.8859	
13	0.5707	0.9118	
14	0.4830	0.9338	
15	0.4576	0.9546	
16	0.3768	0.9717	
17	0.2353	0.9824	
18	0.1875	0.9909	
19	0.0874	0.9949	
20	0.0657	0.9979	
21	0.0422	0.9999	
22	0.0031	0.0000	

PC = Principal Component

* The variance explained by each PC is the Eigen value for that factor. Total variance is defined as the sum of the positive Eigen values of the correlation matrix

PC-4 : Guard cell length and guard cell breadth

PC-5 : Fruit size and fruit shape

PC-6 : Leaf shape and leaf base

PC-7 : Stomatal frequency and mesophyll thickness

PC-8 : Leaf shape (orthotropic shoot) and colour of new shoot tip

Dispersion of cultivars between PC space

In order to study the relative positioning of various OTUs in relation to the PCs it is necessary to plot the OTUs against the PCs taking two each at a time (X and Y coordinates). Such a PC plot gives a visual idea on the contribution of each of the PC in differentiating the different OTUs. This is done by plotting the PC-scores of each of the OTU between the two coordinates, each one representing a PC. Figs. 1-3 represent such PC plots, showing the dispersion of OTUs. In these figures the OTUs with large values (+ve or -ve) on the Y-axis differ from others or the corresponding factor represented by the Y-axis.

Fig.1 gives the dispersion of 51 OTUs between first and second PCs (representing leaf size index, leaf length, leaf breadth and leaf thickness, lower epidermal thickness and upper epidermal thickness). This dispersion diagram highlights the following points. OTU 29 (Panniyur - 1) has large difference both with regard to X and Y coordinates indicating that both first and second PCs are important in differentiating this cultivar from others. OTU 4 (Balancotta) exhibits large difference from X coordinate thereby indicating that the first PC is important in differentiating this cultivar. OTUs 39 and 49 (Vadakkan and Wild Coll. 2060) have large difference with regard to X coordinate, thereby showing that these cultivars are differentiated from others mainly due to the first PC. OTUs 42

Table 2. Sorted and rotated loadings for the eight principal components in black pepper cultivars

Variable	PC-1	PC-2	PC-3	PC-4	PC-5	PC-6	PC-7	PC-8
1	0.970	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.876	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.823	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.875	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000	0.851	0.000	0.000	0.000	0.000	0.000	0.000
7	0.000	0.757	0.000	0.000	0.000	0.000	0.000	0.000
14	0.000	0.000	-0.906	0.000	0.000	0.000	0.000	0.000
12	0.000	0.000	0.891	0.000	0.000	0.000	0.000	0.000
13	0.341	0.000	0.528	0.000	0.000	0.000	-0.255	0.000
11	0.000	0.000	0.000	0.840	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.833	0.000	0.000	0.000	0.000
3	0.000	0.368	0.000	0.545	0.000	0.273	0.000	-0.343
19	0.000	0.000	0.000	0.000	0.816	0.000	0.000	0.000
17	0.000	0.000	0.000	0.000	0.641	-0.256	-0.332	0.000
20	-0.344	0.000	-0.253	0.000	0.577	0.000	0.000	0.000
16	0.000	0.000	0.000	0.000	0.000	0.885	0.000	0.000
15	0.000	0.000	0.000	0.000	0.000	0.795	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.700	0.000
8	0.000	0.358	0.000	0.000	0.000	0.000	0.000	0.000
18	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.779
22	0.000	0.000	-0.290	0.000	0.000	0.000	0.000	0.643
21	0.267	-0.297	0.000	0.000	0.000	0.000	0.388	0.416

PC = Principal Component

(Vellanamban), 26 (Nedumchola), 44 (Vokkalu) and 3 (Arimulaku) show large negative differences from the X coordinate thereby indicating that these characters forming the first PC are important in differentiating them from other OTUs. OTU 5 (Bilimalligesara) has large negative difference in the Y coordinate representing the second PC, while OTU 37 (Uddakere) has a large positive difference from the Y coordinate. This PC is therefore important in differentiating the OTUs 5 and 37 from all the others.

Fig. 2 gives the dispersion of OTUs between the first the third principal components. The largest variation with regard to these PCs was given by OTU 44 (Vokkalu), thereby indicating that both leaf and spike characters led to the divergence of this cultivar. PC-3 (leaf length - spike length index, spike length

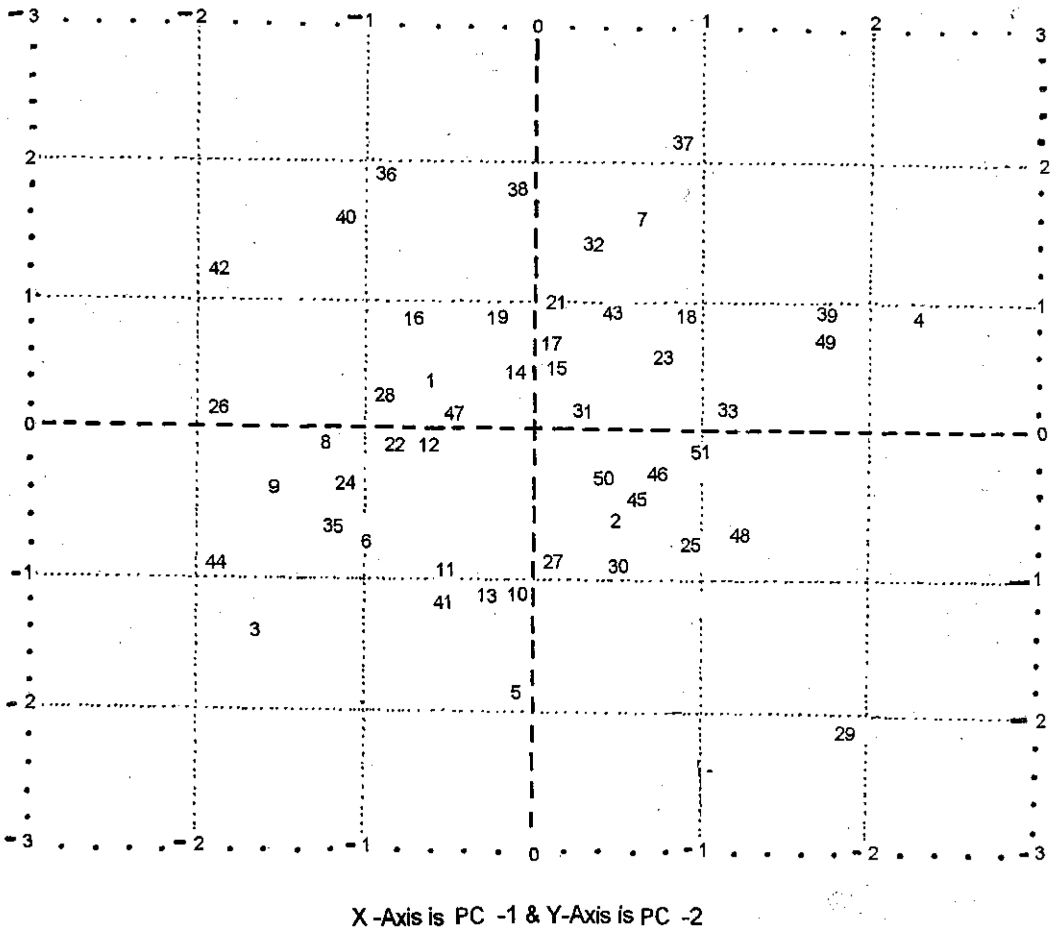

Fig. 1. Dispersion of OTUs between first and second principal components

and peduncle length) is important in differentiating OTU 19 (Kuriyalmundi), 47 (Wild Coll. 2009) and 49 (Wild Coll. 2060), all the three are having large negative differences from the Y coordinate, indicating that these cultivars get differentiated from others by the third PC. OTUs 13 (Karimkotta) and 20 (Kuthiravally) have large positive difference from the Y coordinate indicating that these cultivars are differentiated from others mainly by long spikes, peduncles and small LL/SL values. OTUs 35 (Thommankodi), 32 (Poonjaranmunda) and 29 (Panniyur-1) are the other cultivars having large positive differences from the Y coordinate indicating that the third PC is important in differentiating them from others.

Similarly, distribution of OTUs between first and fourth PCs (Fig. 3) showed that the fourth PC was responsible mainly for differentiating OTUs 39 (Vadakkan), 20 (Kuthiravally), 17 (Kottanadan), 19 (Kuriyalmundi), 10 (Karimunda), 43

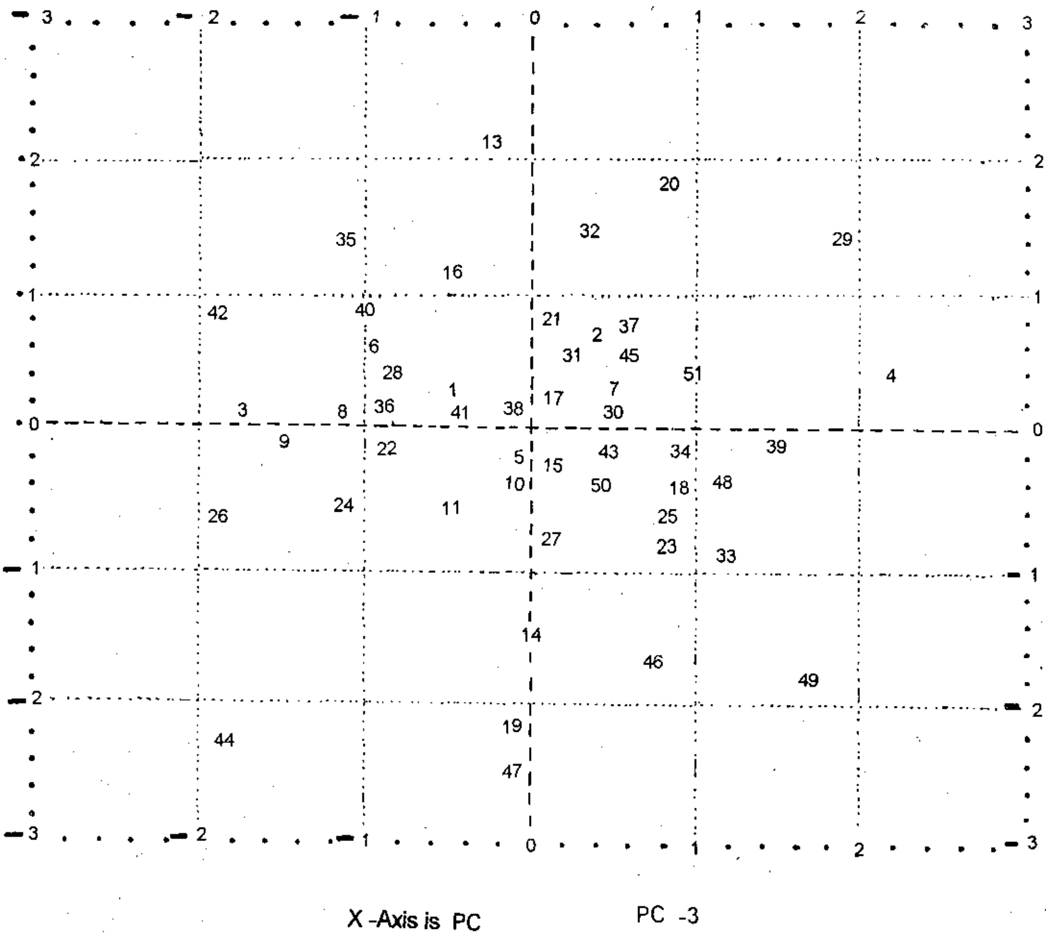

Fig. 2. Dispersion of OTUs between first and third principal components

(Vellayaranmunda) and 36 (Thulamundi).

The dispersion patterns also reveal the variability between the cultivars included in the same group based on centroid clustering. In other words, the PC analysis helped to bring out the finer differences among the related cultivars, which are grouped by centroid linkage. The study has brought out the following general conclusions.

1) Certain cultivars remain as independent entities thereby indicating their divergence from each other and from other cultivars. Such cultivars are Panniyur - 1, Vokkalu, Nedumchola, Kuthiravally, Vadakkan and Karimunda.

2) Panniyur - 1, a hybrid between Uthirancotta and Cheriyakaniakkadan, did not show any similarity with its parents. This absence of resemblance was seen in all the scatter plots. In the centroid linkage also, Panniyur-1 was in an independent group as seen in the

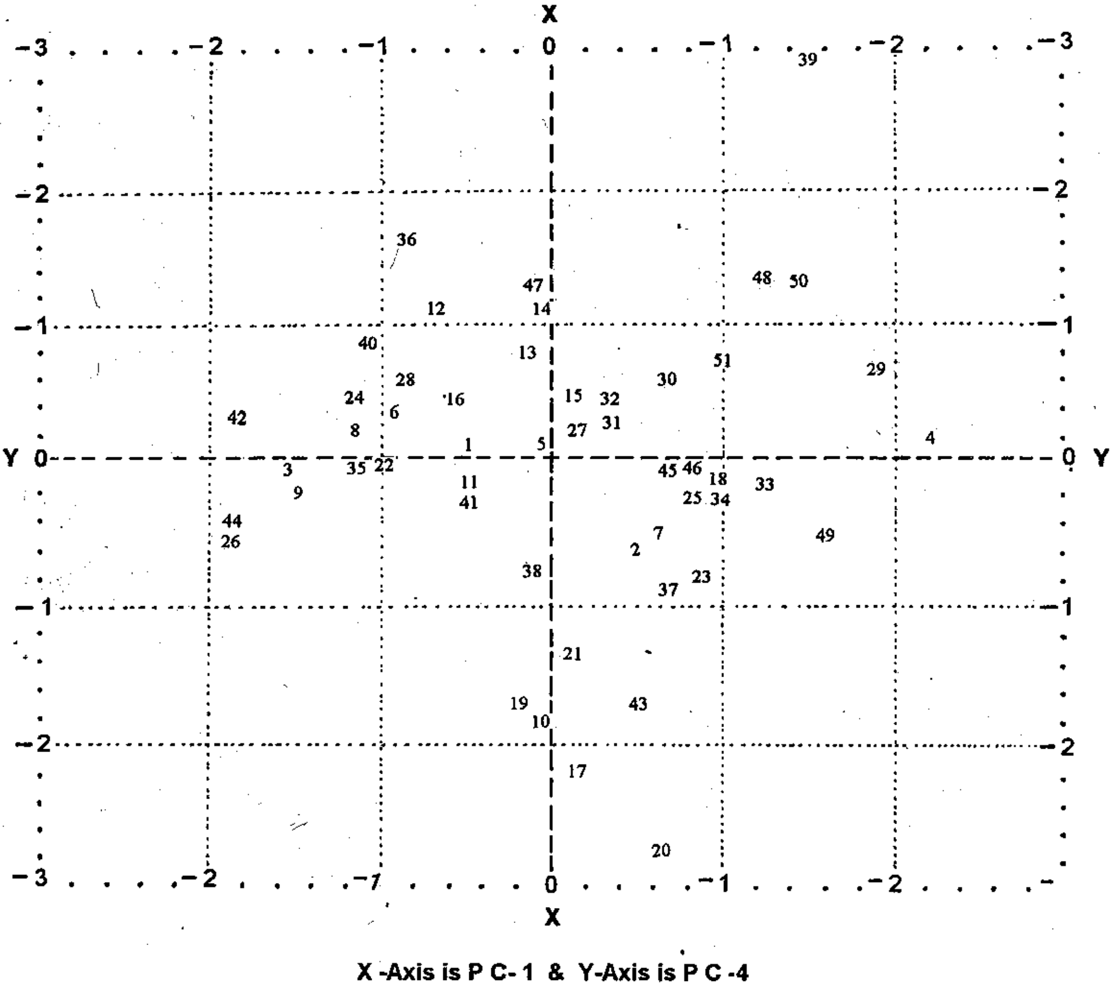

Fig. 3. Dispersion of OTUs between first and fourth principal components

dendrogram and correlation diagram.

3) The majority of cultivars are distributed around the central point indicating their relative similarity, and thereby can be included in a single group. This is comparable to the group D obtained in the centroid linkage clustering. The more distantly related ones occupied positions outside the central group. Cultivars such as Panniyur - 1, Vadakkan and Kuthiravally are so distinct that they stand out in all the scatter plots.

The above numerical taxonomic analysis was useful in relating the extent of divergence among black pepper and in pointing out the divergence of characters that led to the differentiation of cultivars. In order to understand further the role of various PCs in differentiating the OTUs, intra and inter cluster D^2s were computed for each group formed as a result of the centroid linkage (Ravindran *et al.*, this issue).

From this analysis the following conclusions were drawn:

1) PC - 1 delineates cluster H from the group of clusters E, I, J and K and also from the group of clusters A, B, C, D, F and G. The group of clusters E, I, J, and K can be considered as distinct from the group of clusters A, B, C, D, F and G though the distinction is only marginal.

2) PC-2 joined clusters B and C and separated them from the rest (A, D, E, F, G, H, I, J and K). In other words, the nine cultivars represented in cluster B and C are distinct from the remaining OTUs as far as PC-2 is concerned.

3) PC-3 joined clusters F and H and delineated them from the remaining OTUs.

4) PC-4 delineated the original group of 11 clusters into four groups, where cluster F and J are quite distinct between themselves as well as from others. Cluster G and I could be joined with respect to PC-4 and the rest (A, B, C, D, E and K) could be joined as another group of clusters.

5) PC-5 could delineate cluster H and F and also these two clusters from the rest.

6) PC-6 could join cluster E and G and separate them from the rest.

7) PC-7 could show cluster K as a separate group from the rest.

8) PC-8 could form three major groups of clusters, A & K, B, I & F and C, D, E, H & J.

Thus the original 11 clusters were grouped into eight PCs as summarised below:

PC-1 : (A, B, C, D, F, G) (E, I, J, K) H

PC-2 : (B, C) (A, B, C, D, E, G, I, J, K)

PC-3 : (F, G) (A, B, C, D, E, G, I, J, K)

PC-4 : (A, B, C, D, E, H, K) (G, I) F and J

PC-5 : (A, B, C, D, E, G, I, J, K) F and H

PC-6 : (A, B, C, D, F, H, I, J, K) (E, G)

PC-7 : (A, B, C, D, F, G, H, I, J) (K)

PC-8 : (A, K) (B, F, I) (C, D, E, G, H, J)

This study helps to recognise the underlying similarities among the cultivars. Though a large group (cluster D) showed fair degree of resemblances, the presence of distinct groups also pointed out the fact that all the cultivars could not have originated from a common stock but that their origins were separated in space and time. Panniyur - 1 was for example, formed to be distincly differentiated on account of more than one PCs as seen in the PC plots. Similar situations exist in the case of Karimunda, Kuthiravally, Vadakkan, etc. Probably they also might have orginated as natural hybrids in the past. This is plausible as natural crossing followed by reproductive isolation as a result of the absence of any active pollen transfer mechanism in the otherwise dioecious type, coupled with successful vegetative propagation might have been responsible for the divergence noticed in the cultivars (Ravindran *et al.,* 1990). The cultivated black pepper are all selections or straight adoption from forest grown plants. Such selection could have been carried out many times at many locations separated both temporally and spatially. It is for the first time that such a study has been carried out in black pepper. In a similar study with *Piper* species ocurring in Western Ghats, it was shown that PCA was quite useful in establishing the nature of divergence among the various species (Ravindran *et al.*, in press). The usefulness of PCA in the delimitation, both at specific and intraspecific levels, has also been shown in *Cinnamomum* spp. of Western Ghats (Ravindran *et al.* 1995). Absence of random mating and free gene flow in *Piper* leads to isolation of small populations (Ravindran 1990) and these populations might have undergone divergence through segregation, chance natural crossing and accumulation of mutations leading to the origin of variability in black pepper.

Acknowledgments

We are grateful to the authorities of Carneigie - Mellon University Computer Centre, Pittsburgh, USA, for the computer analysis. We are thankful to our colleagues Mr. K Samsudeen, Ms. Minoo Diwakaran and Ms. Geetha S Pillai for their help in the preparation of this paper. We are also thankful to Dr. K V Peter, Director, Indian Institute of Spices Research, Calicut for a critical reading of the manuscript.

References

Adams M W 1977 An estimation of homogeneity in crop plants, with special reference to genetic vulnerability in the dry bean, *Phaseolus vulgaris*. Euphytica 26 : 665-679.

Frane J, Jenrich R & Sampson P 1981 Factor analysis. In : Dixon W J (Ed.) BMDP - 81 Manual (pp. 480-491). University of California Press, USA.

Hilling K W & Iezzoni A F 1988 Multivariate analysis of a sour cherry germplasm collection. J. Amer. Soc. Hort. Sci. 113 : 928-934.

Iezzoni D F & Pritts M P 1991 Application of principal component analysis to horticultural research. HortScience 26 : 334-338.

Martin G B & Adams M W 1987 Land races of *Phaseolus vulgaris* (Fabaceae) in northern Malawi D. 1. Regional variation. Econ. Bot. 41 : 190-203.

Ravindran P N 1990 Studies on black pepper (*Piper nigrum*) and some its wild relatives. Ph D Thesis, University of Calicut, Calicut.

Ravindran P N, Asokan Nair R, Nirmal Babu K, Chandran K & Nair M K 1990 Ecological and morphological notes on *Piper* species from Silent Valley forests, Kerala. J. Bombay Nat. Hist. Soc. 87 : 421-426.

Ravindran S, Manilal K S & Ravindran P N 1995 Numerical taxonomy of *Cinnamomum*. II. Principal component analysis of major species from Kerala, India. In : Manilal K S & Pandy A K (Eds.) Taxonomy and Plant Conservation (pp. 153-164). CBS Publishers, New Delhi.

Sneath P H A & Sokal R R 1973 Numerical Taxonomy. Freeman, San Francisco.

Williams W T 1976 Pattern Analysis in Agricultural Science. Elsevier Scientific, New York.