Plant Res. 2(2) 75~80(1999)

Varietal Classification by Multivariate Analysis on Quantitative Traits in Pecan

Dong-Young Shin[1], Ill-Sup Nou[2]
[1]Department of Resource Plant, [2]Department of Horticulture, College of Agriculture, Sunchon National University, 315 Maegokdong, Sunchon 540-742.

ABSTRACT

Twenty two varieties of pecan including wild types were classified based on 6 characters measured by principal component analysis score distance. The results are summarized as fellow. Twenty two varieties were classified into 5 groups based in PCA score distance. Five groups were distinctly characterized by many morphological characters. Total variation could be explained by 51%, 95%, 99% with first, third and fifth principal components respectively. Varimax rotation of the factor loading of the first factors indicated that the first component was highly loaded with leaf characters, the second component with fruit characters, but fruit length was negative loaded. The second, the third and the fourths groups of cultivars had very close genetic parentage similarity.

Key words : Principal component analysis, Characters, Pecan cultivars, Classification

INTRODUCTION

Pecan is a deciduous tree and belongs to the Juglandaceae family. Pecan is economically important as nut and timber crop. The major producer of pecan is the United States, where it is the third most important nut crop after almond and walnut. Pecan is cultivated in the southern United States for use in various types of bread, ice cream and candy. Pecan is typically cross-pollinated, and its heterozygosity is expected to be high(Madden and Malmalstorm 1975; Romberg 1946). Pecan remains as an unimproved crop because of its long generation time and the relatively recent emphasis on nut production. Selection for different pecan cultivars has been in progress for only about 100 years, and thus these pecan cultivars possess a very recent history compared with most other crops(Sparks 1992). According to Sparks(1992), all pecan cultivars have been derived from the three sources: 1)chance seedling, 2) selections from seedling orchards or seeds planted by homeowners, and 3) breeding programs. Existing cultivars are thought to have a very narrow germplasm base due to the ways they were generated. Sparks(1985) indicated that controled pollination of pecan can result in some contamination. Pedigree information gives an incomplete indication of genetic diversity of different pecan cultivars. The sources of some cultivars are unknown or questionable. The breeding process generally consists of a large number of crosses. However the long generation time and the quantitative nature of most of the important traits of pecan have

Corresponding author: **Dong-Young Shin**, Sunchon National University, 315 Maegokdong, Sunchon 540-742; E-mail: sdy@sunchon.ac.kr

hindered classical genetic studies in this group. Pecan cultivars are classified into 6 group with pollen by Seneter(1992). Goff(1991, 1992)etc classified 46 cultivars by thickness of nut which is an important consideration because consumers often prefer thin shelled nuts that are easy to crack. Darrelld explained various types of cultivated and wild pecan in southern areas, but there are few reports for the classification by quantitative characters of pecan. The most important thing is to select good parents to cross for breeding a new cultivar. The variation of next generation of cross depends on difference of combining ability.

Generally, if it is no problem of physiological combining ability, various genetic variations is induced by the more different genetic background and genetic distance. If we know the genetic distance between the cultivars, it is very convenient to select parents for cross breeding. The best method of good parentage selecting is by crossing each other, however it is impossible and not convenient to cross each cultivar. The method of estimate genetic distance as similarity of total genetic distance without crossing has been used recently and approved by breeders(Bhatt 1970, Adams1978).

The methods of genetic distance and classification by multivariate analysis are Uclid distance(Adams, 1978), Mahalanobis distance,(Mahanobis, 1936) Generized Distance(D2), Principal component analysis(Adams, 1978) and Factor analysis(Denis ,1978). The purpose of this study is to carry out to classified into pecan cultivar by Principal component analysis and factor sanalysis.

MATERIALS AND METHODS

Pecan cultivar and wild types from 22 cultivars were sampled in October 1995 from the farm horticulture and campus of UGA. Investigating items were 6 quantitative characters which were leaf width(mm), leaf length(mm), leaf weight(g), fruit length(mm), fruit width(mm), fruit weight(g). Data analysis is by SAS program(version 6.1).

RESULTS

The simple statistics of the 6 quantitative characters recorded in Table 1. Six characters appeared highly significant by analysis of variance at 1% level. Leaf width is from 115mm to 20mm, leaf weight is from 26 to 3g, fruit weight is from 21 to 2g. There are a lot of

Table 1. The simple statistics of the quantitative characters of pecan by average linkage cluster analysis

Content / Character	Mean	Std. dev.	Max.	Min.	Range	C.V.(%)	Skewness	Kurtosis	Biomodality	F. value
Leaf Width (mm)	43.91	2.97	115.67	20.33	95.34	46.31	2.25	6.82	0.59	24.4**
Leaf Length (mm)	121.33	3.90	262.67	64.00	198.67	31.50	2.36	8.38	0.55	68.2**
Leaf Weight (g)	6.49	0.49	25.97	2.80	23.17	75.00	3.32	13.03	0.73	15.4**
Fruit Width (mm)	6.37	2.90	45.66	24.03	21.63	60.31	1.25	1.07	0.56	9.7**
Fruit Length (mm)	33.31	5.93	34.63	18.40	16.23	17.39	0.13	-0.85	0.38	21.4**
Fruit Weight (g)	22.21	3.09	21.20	2.37	18.83	56.72	0.72	2.17	0.27	19.4**

**significant at 1%

Table 2. The eigen values of the correlation matrix of quantitative character of pecan

	Eigen Value	Difference	Proportion	Cummulative
1	3.06441	1.09192	0.510735	0.51074
2	1.97249	1.29829	0.328748	0.83948
3	0.67421	0.45651	0.112369	0.95185
4	0.21770	0.16311	0.036283	0.98814
5	0.05459	0.03800	0.009099	0.99723
6	0.01660	.	0.002766	1.00000

* The data have been standarized to mean 0 and varience 1
* Root-mean-square total sample standard deviation=1
* Root-mean-square distance between observation=3.464102

Table 3. A matrix of simple correlation coefficients for 6 characteristics of 22 pecan genotypes

Characters	X1	X2	X3	X4	X5	X6
Leaf Width(X1)	100	97*	93*	-2	-38	1
Leaf Length(X2)		100	94*	-3	-37	-4
Leaf Weight(X3)			100	2	-28	10
Fruit Length(X4)				100	35	87*
Fruit Width(X5)					100	22
Fruit Weight(X6)						100

* : Printed values are multiplied by 100 and rounded to the nearest integer values greater than 0.4 have bee flaged by an '*'

genetic variance in this material of pecan varieties. The highest of coefficient variance is leaf weight, the lowest is fruit length and that of leaf width, leaf length were 30-46%, the other characters were 50-60%. A principal factor matrix after orthogonal rotation for the 6 variables is given in Table 2. Factors were constructed using the principal factor analysis technique to establish the dependence relationship between morphological characteristics and yield components from data recorded from 22 genotype of pecan. They shows eigen value and its contribution to total variation obtained from principal component analysis. The proportion of the first component is 51%, the second is 33% , the third is 11%. The total three of cumulative components are 95% and 5 component are 99%. A matrix of simple correlation coefficient of six characters of 22 cultivars is presented in Table 3. Leaf width is positive correlated with leaf length and leaf weight, the same result of leaf length with leaf weight and fruit length with fruit weight. However fruit width is negative correlated with leaf width, leaf length, and leaf weight.

Table 4. Principal factor matrix after orthogonal rotation for 6 characteristics of 22 genotypes of pecan

Variables	Factor 1	Factor2
Leaf Length (X2)	98*	9
Leaf Width (X1)	97*	12
Leaf Weight (X3)	94*	20
Fruit Length (X4)	-51*	40
Fruit Width (X5)	-13	95*
Fruit Weight (X6)	-7	93*

Note : Printed values are multiplied by 100 and rounded to the nearest integer values greater than 0.4 have been flaged by an '*'

A summary of the composition of variables of the two factors with loading are given Table 4. The coefficient of correlation is same as factor loading , so to clear contribution between each factor and character, Varimax method is used. It appeared that characters affect factor exactly, first factor was strongly associated with leaf width, leaf length, leaf weight(Table 4). This

Table 5. Varietal compositions of each groups classified by the principal component score computed in 6 characters

Group	Accession Number	Cultivar Name	Group	Accession Number	Cultivar Name
I	1	Farley	II	18	Schley
	5	KOPF		21	Stuart
	11	Cheyenne	III	2	Pawnee
	12	Chickascum		16	Wichta
	14	Kiowa		17	Desirable
	15	Osage	IV	6	Candy
	19	Elliot		7	Red hickory
	20	Clarke		22	Carya corpiformis
II	3	Sioux	V	8	C. tomentosa
	4	USDA139		9	C. glabra
	13	Shawnee		10	Black walnut

Table 6. Group means of each characters classified by the PCA score distance

Character / Group(mm)	Leaf Width (mm)	Leaf Length (g)	Leaf Weight (mm)	Fruit Width (mm)	Fruit Length (g)	Friut Weight (g)
I	32.14	102.53	0.45	5.86	30.81	22.86
II	38.60	110.60	0.68	4.39	37.34	20.19
III	37.00	112.90	0.51	10.84	41.72	24.31
IV	57.67	146.00	0.68	3.96	27.65	18.70
V	72.23	171.67	0.39	13.53	31.47	29.21

factor was regarded as a leaf characteristics. Second factor was fruit length and fruit weight This means that the first factor was leaf characters and the second factor was fruit character. Leaf characteristics had positive loading in factor 1, but fruit characters had negative loading in factor 1. These two characters had positive loading in factor 2. The sign of the loading indicates the direction of the relationship between the factor and the variable. Thus, three variables with high loading indicates in the same factor with the same sign would be expected to exhibit a positive correlation. Fruit length was negative to the first factor. Factor analysis provides more information than a simple correlation matrix because it indicates groups of variables and indicates groups of variables and indicates percentage contribution of variables to each factor. As a previous result of principal analysis, we know that first and second factor contributed classification. Fig 1 shows dendrogram of 22 cultivars based on distance from average cluster analysis in 6 characters. Five group are classified into by cluster analysis depend on 1 average distance. The first group was composed 8 cultivar (Farley, KOPF, Cheyenne, Chickasaw, Kiowa, Osage, Elliot, Clarke), the second group was composed 5 cultivar(Sioux, USDA139, Shawnee, Schley, Stuart), third group composed Pawnee, Wichta, Dessirable and the other groups were also 3 cultivar. Forth and fifth group were composed wild cultivar(*Carya corrdiformis*) and campus trees of UGA(C.ttomentosa, C.glabra, Black walnut). Table 6 shows that group means of each characters were classified by the PCA score distance. 5 th group of leaf width was the largest of any other group ; the first, the second and third group means range was from 32mm to 37mm. The leaf length of the 4th and the 5th groups was 145-171mm; but first, second and third groups of range were 100-113mm. The fifth group mean of leaf weight was highest among all the group, but the other groups had similar means that were very

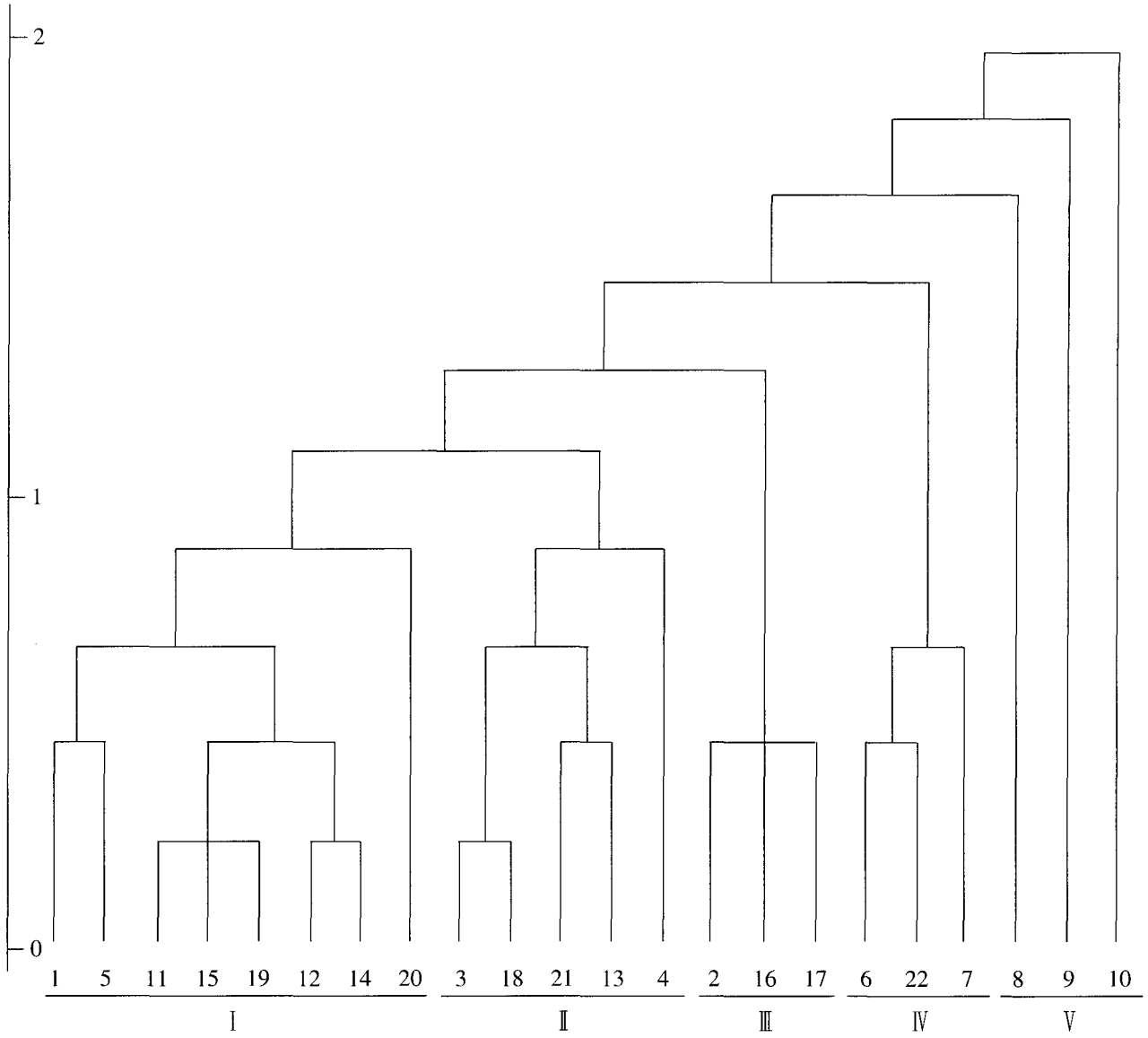

Fig. 1. Dendrogram of twenty two taxa based on PCA score distance in six characters. The sample numbers are same to shown as in Table 5

low compared to the fifth group. The 3th and 5th group means of fruit width were bigger than the first, second group means.

DISCUSSION

The classification of multi-variate analysis has been confirmed in rice, red pepper, cotton, corn *etc*. Comparison method of similarity between cultivar are Uclid distance(Sneath, 1973), Mahalanobis D2 (Mahalnobis, 1936), Factor distance of Principal Component Analysis(Adams, 1978). It is very difficult to know the best method to classify cultivars, but depending on variation of characters among crops, Principal Component Analysis have been a good method to adjust crops that have a lot of variations since this method is calculated by means data. There were some reports(Adams 1978, Denis 1978) which are used

by this method. The overall result of classification seems promising as a tool for pecan cultivar. All cultivars were distinguishable using quantitative characters by Principal Component Analysis. The factor analysis approach is one that can be used successfully for analysis of large amounts of multi-variate data, and should be applied more frequently in the field of genotype classification. This research describes one application of factor analysis can be a subjective procedure. The greatest benefit of factor analysis can be delineating areas of further research designed to test the validity of the suggested factors. Use of factor analysis by plant breeders has the potential of increasing the comprehension of the causal relationships of variables, and help determine the nature and sequence of traits to be selected in a breeding program. Dendrogram was constructed based on quantitative characters that classified these 22 cultivars into 5 major groups(Fig. 1). Each group shows their special characters in relation to each other. Pecan cultivars in second group are related to each other more similar than other group. The parentage of Sioux and Shawnee is Schley which was composed same second group. Sioux is offspring of Schley cross with Carmich. Shawnee is next generation of crossing Schley and Barton. The parentage of third group is one half of Mahan, Succes, Jewett . There seems to has a lot of genetic similarity within each group. Fourth group of pecan was very small nut characters. The 5th group of pecan cultivars was wild types.

ACKNOWLEDGMENT

This paper was supported by NON DIRECTED RESEARCH FUND, Sunchon National University.

REFERENCE

Adams, M. W. and J. V. Wiersma. 1978. An adaptation of principal component analysis to an assessment of genetic distance. Michigan State University. Agricultural Experiment Station Research Report.347.

Bhat, G. M. 1970. Multivariate analysis approach to selection of parents for hybridization aiming at yield improvement in self-pollinated crops. Aust. J. Agri. Res.21:1-7

Darrell Sparks. 1992. Pecan Cultivars the orchard' s foundation Pecan production innovations Watkinville Georgia

Denis, J. C. and M. W. Adams.1978. A factor analysis of plant variables related to yield in dry beans. I. Morphological traits. Crop.Sci. 18:74-78.

G. J. Seiler and R. E. Stafford. Factor Analysis of Components of Yield in Guar. Crop Science, vol.25, November- December.

Goff W. D. ,. Braster L. R. and West M. S.1992. Shell thickness classification for 46 pecan cultivars grown in the Southeastern United States. Hort. Sci. 26 :11-1427.

Goff. W. D., Brasher, L. R., West M. S. 1991. Shell thickness classifications for 46 pecan cultivars grown in the southern United States. Hortscienc Alexandria, Va.: American Society for Horticultural Science. 226:11-1427.

Mahaanobis, P. C. 1936. On the generized distance in statistic. Proc. Nat. Inst. Sci. India 2:49-55.

Seneter,S. D., Thompson, T. E. 1992. Lipid profiles of pollen from six pecan cultivars. Annu. Rep. North nut brow assoc. hamden, conn.: The association. 63-67.

Sneath, P. H. a., and R. P. Sokal. 1973. Numerical taxonomy. W. H. Freeman. San Franciso. 573.

Received September 20, 1999
Accepted November 20, 1999