20. CLASSIFICATION OF LOCAL STRAINS OF MAIZE IN JAPAN AND SELECTION OF BREEDING MATERIALS BY APPLICATION OF PRINCIPAL COMPONENT ANALYSIS

Noboru MOCHIZUKI

Researcher, Division of Genetics, Department of Physiology and Genetics,
National Institute of Agricultural Sciences
Hiratsuka, Kanagawa

Collection of adapted local strains and introduction of exotic germ plasm for breeding use has been frequently carried out in maize breeding. These strains were, in most cases, classified into groups on the basic of their origin and characteristics observed in introduction fields, tested on their combining ability, and breeding materials were selected from the strains.

The objective of this paper is to report on the application of principal component analysis to classification of local strains (open-pollinated varieties) of maize and to selection of breeding materials from them for use in hybrids and synthetic varieties.

The source materials in the present study were a series of the reports on the characteristics and the combining ability of Caribbean flint local strains collected from Fuji, Shikoku, and Kyushu in Japan.

Maize strains or collections were generally classified into strain group, varieties, or races based on observations on variations of many characters of a plant. So principal component analysis, a method of multivariate statistical analysis, was applied to the classification of the maize strains.

Classification of the Local Strains from Shikoku and Meaning of Principal Component in Relation to Maize Breeding

Out of 65 agronomic and botanical characters in 57 strains collected from Shikoku observed at Hiratsuka, Kanagawa, in 1967, 10 characters were selected, which were of importance in maize breeding, in which there were distinct differences between strains, and among which correlation coefficients were not very high. These characters were silking date, stalk and leaf length, number of leaves, tassel and ear length, ear diameter, ear weight, number of husks, and 100 kernel weight.

The correlation matrix of the 10 characters were calculated, following principal component analysis. The results are presented in Table 1. Twenty-eight 21, 18 and 13

Table 1. Percent of eigen value (λ_k) in total variation (p) obtained from principal component analysis of the 10×10 correlation matrix in strains from Shikoku.

Principal component	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
λ_k/p (%)	28.3	21.4	18.2	12.5	6.3	5.3	3.4	2.2	1.8	0.6
$\left(\sum_{l=1}^{k}\lambda_l\right)/p$ (%)	28.3	49.6	67.8	80.3	86.7	92.0	95.4	97.6	99.4	100.0

percent of the total variation of the 10 characters were accounted for by the first four principal components, respectively. Hence about 80 percent of the total variation could be explained by the principal components in total.

So as to classify the strains into strain groups or varieties having similar characteristics, the squared distances between the 57 strains in the four dimensional space were calculated from the first four principal component's scores in the strains. The smaller the squared distance was between strains, the more similar the characteristics of the strains were expected to be. So the strains among which the squared distances were very small were grouped as a variety. The criterion of the grouping was that the average distances within a variety were always smaller than the ones among varieties. In consequence, the 57 strains were classified into 14 varieties. Furthermore, by the same way the varieties were classified into four major varietal groups such as A, B, C, and D. The average squared distances between varietal groups and within groups are presented in Table 2.

Table 2. Average squared distance between varietal groups and within one in strains from Shikoku.

Varietal group	A	E	C	D
A	6.0	22.2	37.4	31.5
B		11.4	18.9	38.1
C			10.6	40.5
D				0

For the purpose of understanding the relation between characters and principal components, the characters were assorted into three classes such as plus, minus, and zero in Table 3. Class zero was omitted in the Table. This assortment was based on the element of eigen vector associated with eigen values. In the case of the first principal component, the characters concerned with size of vegetative characters and earliness, such as stalk and leaf length, number of leaves, and silking date, contributed greatly to the principal component. So the meaning of the first principal component appeared to correspond to the general size of vegetative characters in relation to the duration of growing period. By the similar consideration the meaning of the second principal component appeared to correspond to yield potentiality, especially to the efficiency of photosynthesis and translocation in a plant, and the third to the degree of differentiation in organs.

Table 3. Assortment of characters by degree of contribution to principal components in strains from Shikoku.

Principal component	Class	Corresponding character
X_1	+	Stalk length, leaf length, silking date, number of leaves
	−	None
X_2	+	Ear weight, ear diameter, 100 kernel weight, ear length
	−	Silking date, number of leaves
X_3	+	Tassel length, ear length, 100 kernel weight
	−	Ear diameter, number of husks, number of leaves

Classification of Representative Caribbean Flint Local Strains in Japan and Relation Between Classification and Maize Breeding

In order to ascertain the universality of this statistical classification method mentioned above, principal component analysis was applied to the data on the characteristics of representative 72 Caribbean flint local strains in Japan. The strains including 24 from each of Fuji, Shikoku, and Kyushu were observed at Hiratsuka, Kanagawa in 1958.

The correlation matrix of 12 characters selected from the 65 characters were calculated, following the analysis. Fourty-eight, 16, 10, and 8 percent of the total variation was accounted for by the first four principal components respectively, and hence more than 80 percent could be explained in total.

The 72 strains were grouped into 14 varieties which were further classified into four major varietal groups such as A, B, C, and D. The interrelationship of the varietal groups in the four dimentional space are shown graphically in Figure 1. The figures in the diagram indicate the average distances between varietal groups and within one.

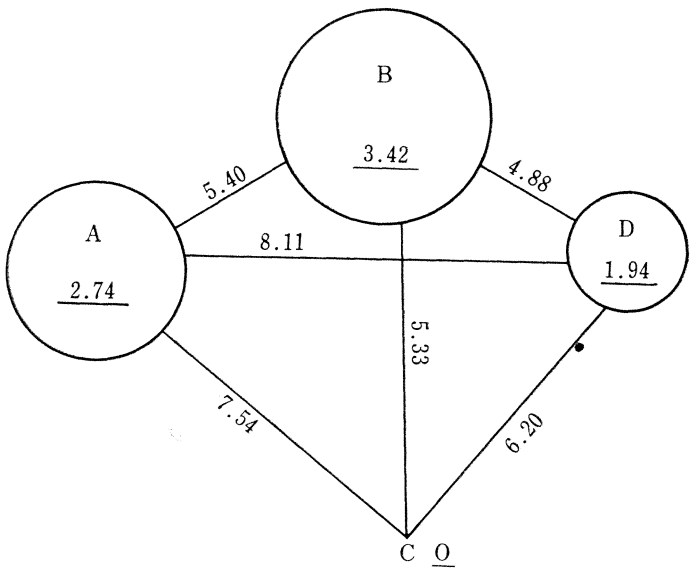

Fig. 1. Clusters representing varietal groups and their interrelations in representative Caribbean flint local strains.

The meanings of the principal components are similar to that obtained in the strains from Shikoku. In the case of the first principal component which could explain nearly half of the total variation, the characters concerned with earliness, plant size, and grain yield contributed greatly to the principal component. This indicates that the compound variation of the characters concerned with earliness, plant size, and grain yield represented the main variation of the Caribbean flint local strains in Japan.

Furthermore most of the representative local varieties, which had been evaluated as superior breeding stocks in use for hybrids and actually had been used as the parental sources of the recommended hybrid varieties in Japan, belonged to the varietal group B in this classification. This fact suggested that the classification based on principal com-

ponent analysis was of significance for selection of breeding material from the classification.

Relation between Combining Ability and Principal Component

For the purpose of verifying the applicability of principal component analysis not only to classification of strains but also to preliminary selection of breeding materials from the strains, the relation between combining ability and principal components was studied.

The materials in this particular study consisted of two sets of data. One was on the adaptability trial of the 72 local strains including all the ones mentioned before observed at Kuma, Ehime in 1958. The other was on the combining ability trials of 24 local strains from Fuji in 1960 and 24 from Shikoku in 1961 at the same location. Combining ability of the strains (open-pollinated varieties) was tested in top-cross trials using four U.S. dent testers, three inbred lines and a single cross in both years.

Ten agronomic characters were selected following principal component analysis and classification on the adaptability trial data. Also combining ability for grain yield and silking date were analysed on the combining ability trials by the procedures suggested by Federer *et al* (1947) and Plasted *et al* (1962).

The results may be summerized as follows. The 72 strains were grouped into 13 varieties which were further classified into six major varietal groups. High correlation coefficients between the estimates of general combining ability and the first principal component's scores were observed, *i.e.* r=.919** and .645** in the strains from Fuji and Shikoku, respectively. The relationship for strains from Fuji is shown in Figure 2. Zero or low correlation coefficients between principal components scores and the estimates of the standard deviation of specific combining ability within a strain were found. Most of the strains having superior general combining ability were included in a particular varietal group, the B varietal group in the present study.

These results indicate that significant classification in relation to maize breeding

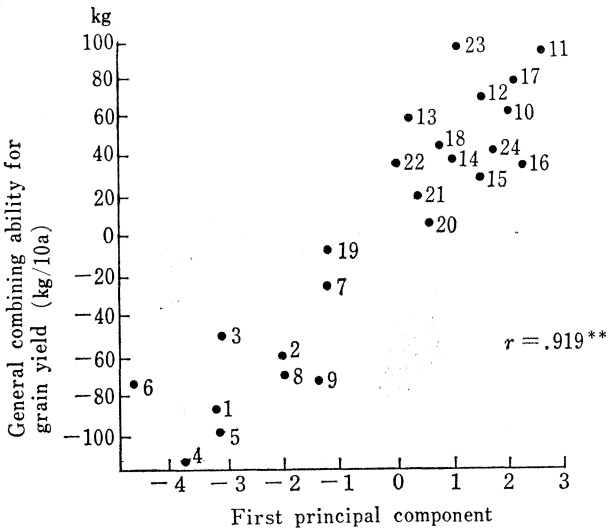

Fig. 2. Correlation between general combining ability for grain yield and first principal component in strains from Fuji. Figure in diagram indicates strain number.

and preliminary selection of breeding materials without testing procedures of combining ability were possible by application of principal component analysis to the data on the characteristics of the strains.

Thus, it was concluded from the present study that classification of many strains collected, introduced, and maintained and preliminary selection of superior breeding materials for use in hybrids and synthetic varieties could by achieved by application of principal component analysis.

References

1. Federer, W. T. and Sprague, G. F. (1947): A comparison of variance components in corn yield trials: I. Error, tester × line, and line components in top cross experiments. J. Amer. Soc. Agron. 39: 453–463.
2. Mochizuki, N. (1968): Classification of maize lines and selection of breeding materials by the application of principal component analysis. Bull. Nat. Inst. Agric. Sci. series D 19: 85–149.
3. Plasted, R. L., Sanford, L., Federer, W. F., Kehr, A. E., and Peterson, L. C. (1962): Specific and general combining ability for yield in potetoes. Amer. Poteto J. 39: 185–197.
4. Seal, H. (1964): Multivariate Statistical analysis for Biologists. Methuen & Co. Ltd. London. pp 209.

Discussion

A. Anós, Spain: Doesn't the use of high correlated characters affect the successful classification of strains when the principal component analysis is used? If so, how could you classify the strains using high correlated characters?

Answer: Highly correlated characters can be also used for classification of strains by principal component analysis. I have a result on this point. In the text I reported only the case using ten characters for the classification of the Shikoku strains. However, I did principal component analysis using 24 characters from which the ten characters were selected, and among which correlation coefficients were in wide range from high to low. Even in thic case the four principal components which could be explained nearly 80 percent of total variation were obtained.

A. Senanarong, Thailand: What is the genetic background of local Caribbean flint of Japan? How long it has been grown in Japan?

Answer: The Caribbean flint distributes in mountaneous area of central and southwestern Japan. It was introduced by Portuguese about four hundred years ago. Since then it has been cultivated and selected by farmers, so it is highly adapted to Japanese climate and soil, and also has high productivity.

D. Sharma, India: Were general combining ability estimate and principal component analysis also related with the (*per se*) performance of a particular strain?

Answer: As I mentioned before, the correlation coefficients between the estimate of general combining ability and the first principal component were .919** and .645** in Fuji and Shikoku strains, respectively. Using the same data I found that the correlation coefficients between the estimate of general combining ability and the grain yield of strain *per se* were .762** and .502* in Fuji and Shikoku strains, respectively. Therefore, the application of principal component analysis seems to be more effective to select high combining ability strains than the performance of a strain.

D. Sharma, India (comment): I have considerable data which shows that general combining ability of strains estimated even by use of seven testers gives more or less

the same ranking as are gotten from the yield performance of the strains. If groups based on principal component analysis are also highly correlated with the general combining ability estimate, then it would be more practical to select on the basis of yield itself.

V. R. Carangal, Philippines: Do you have any idea of the origin of these strains? How is this related with you varietal grouping?

Answer: The strains were originally introduced by Portuguese about four hundred years ago. The varietal grouping in the text is within the Caribbean flint of Japan.