Cluster and principal component analysis for the selection of maize (*Zea mays* L.) genotypes

Pratima Pahadi[1], Manoj Sapkota[2]*, Dhruba Bahadur Thapa[3] and Shreena Pradhan[4]

[1]Gokuleshowar Agriculture and Animal Science, Tribhuvan University, Baitadi, Nepal; [2]Institute of Agriculture and Animal Science, Tribhuvan University, Chitwan, Nepal; [3]Agriculture and Botany Division, Nepal Agricultural Research Council, Khumaltar, Lalitpur, Nepal; [4]Agriculture and Forestry University, Chitwan, Nepal.

*Corresponding Author

Abstract

Breeding for high yield crop requires information on the nature and magnitude of variation in the available materials, relationship of yield with other agronomic characters and the degree of environmental influence on the expression of these components characters.This study was conducted with the aim of identifying better performing maize genotypes and related traits with the help of principal component analysis and cluster analysis of major quantitative traits of the crop. Six genotypes of maize were tested and observed for days to tasseling, days to silking, days to pollen shed anthesis, ear height, silk length, plant height, ear length, ear circumference, number of kernel row per ear, number of kernel per row, five hundred kernel weight and grain yield.The first two components that explained 73.7% of the total variation were determined from Principal component analysis and were used for clustering genotypes. Second cluster comprising of four genotypes namely Rampur Yellow, CP808, Khumal Yellow and Rajkumar, had higher value of traits like number of kernel row per ear, number of kernel per row, and grain yield. The selection from the second cluster can be considered worthwhile as it has genotypes performing better in terms of yield and yield attributing characters and can be used for breeding purpose of hybrids.

Key words: Maize, PCA, Cluster analysis, selection, yield.

Introduction

Selection mainly based on phenotypic characters is the major technique used in breeding program. Response to selection depends on many factors including interrelationship of the characters. Plant breeders are interested in developing cultivars with improved yield along with other desirable agronomic and phenological characters. For this, breeders can choose between selecting

desirable genotypes in early generations and delaying any intense selection until advanced generations (Puri et al*., 1982).

The selection criteria may be yield or one or more of the yield attributing characters. However, breeding for high yield in crops require information of the nature and magnitude of variation in the available genotypes, relationship of yield with other agronomic characters and the intensity of environmental influence on the expression of these characters. Since grain yield in maize is quantitative in nature and polygenetically controlled, effective yield improvement and simultaneous improvement in yield components is crucial (Bellon and Olaye, 2009). Selection on the basis of grain yield character alone is usually not as effective and efficient as those based on its components characters (Muhammad et al., 2003). Knowledge regardingthe association between yield and its components traits and among the component parameters themselves will improve the efficiency of selection in plant breeding (Fakorede and Opeke, 1985).

This study was conducted with the aim to identify better performing maize genotypes with the help of principal component analysis and cluster analysis of major quantitative traits of the crop.The identification of such genotypes with superior traits could help in making the selection of good performer convenient and the planning of further breeding strategies effective.

Material and methods

The field experiment was conducted at the research field of Gokuleshowar Agriculture and Animal Science College (700 masl) from July 24, 2015 to October 29, 2015. The planting material was collected from National Maize Research

Program (NMRP), Rampur (Table 1).The experiment was laid out in randomized block design with three replications. The plot size was 2.25 m× 1.25 m = 2.8125 m^2 each. The row spacing of each treatment for maize sowing was 75 cm and there were four rows per plot.

Farm yard manure was applied at the rate of 10 ton/ha and chemical fertilizers were applied at the rate of 150:75:75 kg NPK/ha. The experiment was conducted under rainfed condition and one irrigation was given during grain filling stage. Harvesting of crop was done on the basis of physiological maturity of each genotype on October 29, 2015.

Observations were taken for days to tasseling (DT), days to silking (DS), days to pollen shed anthesis (PSAD), silk length (SL), plant height (PH), ear length (EL), ear height (EH), ear circumference (EC), number of kernel row per ear (NKRE), number of kernel per row (NKR), five hundred kernel weight (FKW) and grain yield (GY).

Data entry and processing was done in Microsoft excel. Principal component analysis was done from all the observed data through Minitab 17. Similarly, cluster analysis and dendogram were also prepared with the help of Minitab 17.

Result and Discussion
Principal component analysis (PCA)

Five components were extracted from 13 studied traits by PCA. The first two components that explained 73.7% of the total variation were used for clustering genotypes as shown in Table 2. Through this method 13 variables were reduced to two components with the help of the PCA and the number of component shown by the scree plot (Fig1). DT, DS, PSAD, SL, EL, EC, NKRE, NKR and GY had positive effect in the PC1 whereas TSI, PH and FKW had negative effect.

The most effective trait in the first component were days to tasseling, days to silking, PSAD, EH, EC, NKRE and NKR. DT, DS, PSAD, EH, PH, EC, NKRE and FKW had positive whereas TSI, SL, EL, NKR and GY had negative effect in the PC2. TSI, PH and FKW were the most effective traits in the second component (Table 3).

Cluster analysis

Cluster analysis was done based on PCA. From the dendogram (Fig 2), we can find the genotypes categorized into two clusters. First cluster comprises of two genotypes, Khumal Rato and Local which are found to be about 40% similar. This cluster is characterized by high value of TSI, PH and FKW (Table 4). Second cluster comprises of four genotypes, Rampur Yellow, CP808, Khumal Yellow and Rajkumar, in which Rampur Yellow and CP808 were the most similar genotypes. This cluster has higher value of PSAD, NKRE, NKR and GY (Table 4).

Biplot analysis

The biplot (Fig 3) provides similar analysis of the data as that stated from the cluster analysis and the cluster characteristics.The biplot shows that the traits GY, SL, EL, NKR, NKRE, DT, DS, PSAD and EHlie together whereas FKW, PH, EH/SH and TSI are further and they form a different grouping. We can thus observe the traits in such group behave similarly.

The selection from the second cluster is worthwhile as it has genotypes performing better in terms of yield and yield attributing characters.

The selection of genotypes from second cluster means the selection of genotypes having higher value of EC, EL, NKRE and NKR which leads to selection of high yielding genotypes considering yield was found positively related with ear girth, ear height and number of kernels

row and it was in correspondence with the finding of Wali et al., 2006. The second cluster also includes lower plant height. Thus, the genotypes from Cluster II can be used for breeding program with hybridization for a dwarf plant and high yield.

Biplot also madeit convenient for selection of the traits that are closer and behave similar in order to enhance the yield. The same can be concluded from the two components determined from the PCA.

Rahim et al., (2010) showed that the hybrids of genotypes with maximum dissimilarity result in high yield and so the cross between the most dissimilar genotypes (Rajkumar with Khumal Rato or the local genotype) shown from the cluster analysis can be done in breeding program to achieve higher heterosis. Similarly, the use of different measurement techniques can be appropriately used for genotypes grouping (Bauer et al., 2007; Kraic et al., 2009).

Similar to the findings by Ali et al., (2008) who reported that cluster analysis can be helpful for finding high yielding genotypes and Singh and Dwivedi (2002), the results of this study showed the presence of a genetic divergence among maize genotypes which led to distinction of better performing genotype of maize.
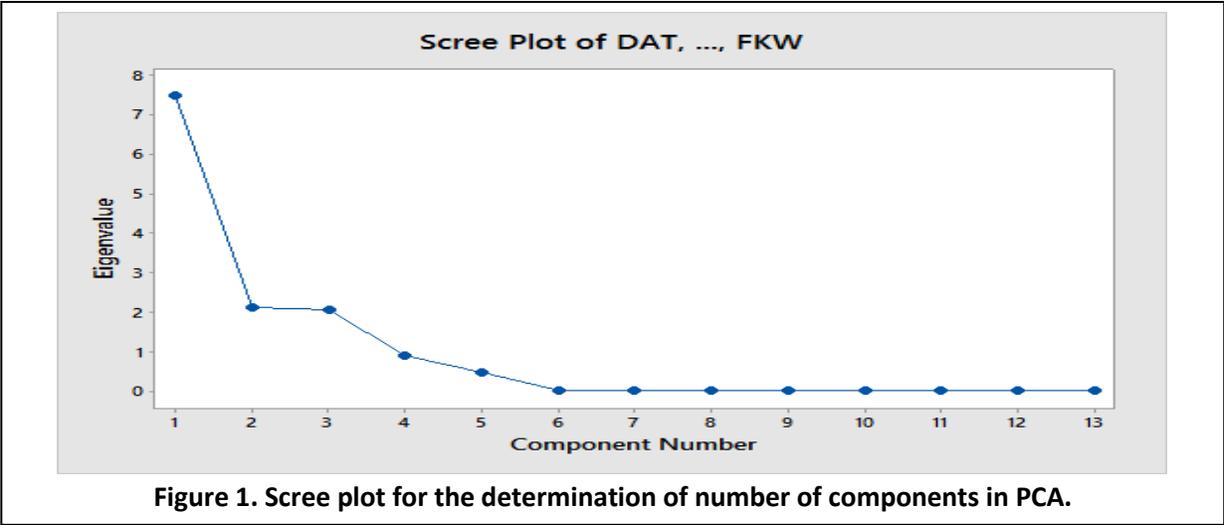
Figure 1. Scree plot for the determination of number of components in PCA.
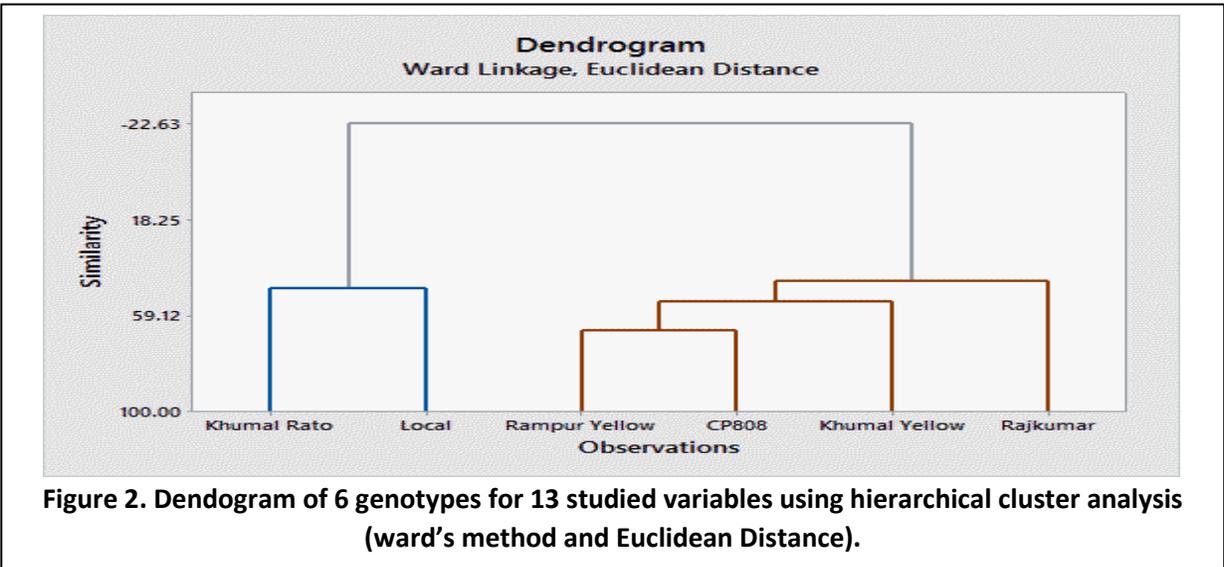

Figure 2. Dendogram of 6 genotypes for 13 studied variables using hierarchical cluster analysis (ward's method and Euclidean Distance).
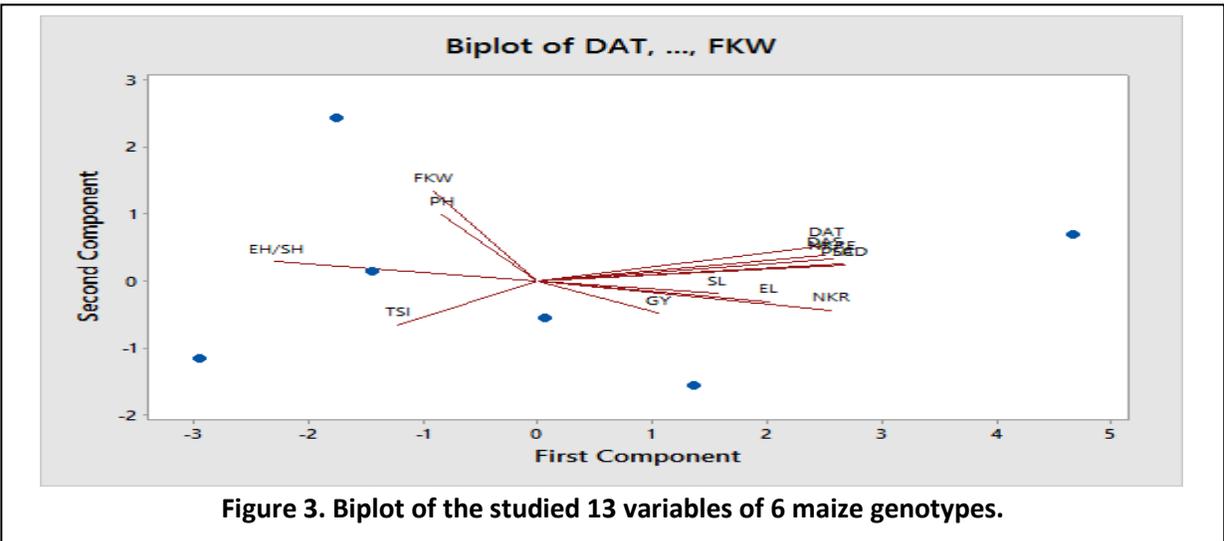

Figure 3. Biplot of the studied 13 variables of 6 maize genotypes.

Table 1. List of plant material used in the experiment.

Genotype Entry	Cross Name	Origin
1	Khumal Rato	Agriculture Botany Division Nepal Agriculture Research Council (NARC)
2	Rampur Yellow	National Maize Research Program, Rampur
3	Khumal Yellow	Agriculture Botany Division NARC
4	CP808	Foreign Germplasm
5	Local	Baitadi
6	Rajkumar	Foreign Germplasm

Table 2. Eigen analysis of the Correlation Matrix.

Eigen value	7.4701	2.1167	2.0404	0.9062	0.4666
Proportion	0.575	0.163	0.157	0.07	0.036
Cumulative	0.575	0.737	0.894	0.964	1

Table 3. PCA analysis of 13 studied traits in maize genotypes.

Variable	PC1	PC2
Days to Tasseling	0.338	0.252
Days to silking	0.335	0.179
Tasseling silking Interval	-0.162	-0.311
Days to pollen shed anthesis	0.358	0.114
Ear height	-0.304	0.133
Silk length	0.21	-0.092
Plant height	-0.111	0.466
Ear length	0.27	-0.15
Ear circumference	0.357	0.115
Number of kernel row per ear	0.344	0.157
Number of kernel per row	0.342	-0.208
Grain yield per plant (g)	0.141	-0.232
Five hundred kernel weight (g)	-0.12	0.629

Table 4. Average value of the studied traits of two clusters formed from cluster analysis.

Variable	Cluster1	Cluster2	Grand Centroid
Days to Tasseling	54.833	55.833	55.500
Days to silking	59.000	59.583	59.389
Tasseling silking Interval	4.167	3.750	3.889
Days to pollen shed anthesis	74.000	76.250	75.500
Ear height	116.302	111.700	113.234
Silk length	7.280	7.083	7.149
Plant height	265.667	244.458	251.528
Ear length	24.917	25.858	25.544
Ear circumference	16.900	17.642	17.394
Number of kernel row per ear	12.967	13.950	13.622
Number of kernel per row	27.967	32.000	30.656
Grain yield per plant (g)	126.786	151.984	143.585
Five hundred kernel weight (g)	171.698	145.106	153.970

References

Ali, Y., Atta, B. M., Akhter, J., Monneveux, P. and Lateef, Z. (2008). Genetic variability, association and diversity studies in wheat (Triticumaestivum L.) germplasm. *Pak J Bot.* 40: 2087-2097.

Atanaw, A., Wali, M. C., Salimath, P. M. and Jagadeesha, R. C. (2010). Combining ability and heterosis for grain yield and ear characters in maize. *Karnataka Journal of Agricultural Sciences.* 19(1).

Bauer, I., MladenovicDrinic, S., Drinić, G., and IgnjatovićMicić, D. (2007). Assessing temporal changes in genetic diversity of maize hybrids using RAPD markers. *Cereal Research Communications.* 35(4): 1563-1571.

Bello, O. B. and Olaoye, G. (2009). Combining ability for maize grain yield and other agronomic characters in a typical southern guinea savanna ecology of Nigeria. *Afr. J. Biotechnol.* 8 (11): 2518-2522.

Fakorede, M. A. B. and Opeke, B. O. (1985). Weather factors affecting the response of maize to planting dates in a tropical rainforest location. *Experimental Agriculture*. 21(01): 31-40.

Kraic, F., Mocák, J., Roháčik, T. and Sokolovičová, J. (2009). Chemometric characterization and classification of new wheat genotypes. *Nova Biotechnol.* 9: 101-106.

Muhammad, B. A., Muhammad, R., Muhammad, S. T., Amer, H., Tariq, M. and Muhammad, S. A.(2003). Character association and path coefficient analysis of grain yield and yield Components in maize. *Pak. J. Biological Sci. 6* (2): 136-138.

Puri, Y. P., Qualset, C. O. and Williams, W. A. (1982). Evaluation of yield components as selection criteria in barley breeding. *Crop Sci.* 22: 927-931.

Rahim, M. A., Mia, A. A., Mahmud, F., Zeba, N. and Afrin, K. (2010). Genetic variability, character association and genetic divergence in Mungbean (*Vignaradiata* L. Wilczek). *Plant Omic.* 3: 1-6.

Singh, S. P. and Diwivedi, V. K. (2005). Genetic divergence in wheat (*Triticumaestivum* L.). *Agricultural Science Digest.* 25(3): 201-203.

Wali, M. C., Salimath, P. M., Prashanth, M., and Harlapur, S. I. (2006). Studies on character association as influenced by yield, starch and oil in maize (*Zea mays* L.). *Karnataka Journal of Agricultural Science.* 19(4): 932-935.