www.ThePharmaJournal.com

The Pharma Innovation

Swapnil
Department of Plant Breeding
and Genetics, Bihar Agricultural
University, Sabour, Bhagalpur,
Bihar, India

Jaishreepriyanka R
Department of Plant Breeding
and Genetics, Bihar Agricultural
University, Sabour, Bhagalpur,
Bihar, India

Digvijay Singh
Department of Plant Breeding
and Genetics, Bihar Agricultural
University, Sabour, Bhagalpur,
Bihar, India

SS Mandal
Department of Plant Breeding
and Genetics, Bihar Agricultural
University, Sabour, Bhagalpur,
Bihar, India

Principal component analysis in maize (*Zea mays* L.) under normal sown condition of Bihar

Swapnil, Jaishreepriyanka R, Digvijay Singh and SS Mandal

Abstract
In the present study, 32 genotypes of maize were grown under normal sown condition of Bihar at Maize Research Farm, Bihar Agricultural University, Sabour (Bhagalpur) during *Kharif,* 2019-20. The experiment was laid in randomized complete block design with three replications to estimate the contribution of ten quantitative traits to the total variability in maize through Principal component analysis. The traits used for analysis were days to 50% tasseling, days to 50% silking, days to 75% brown husk, plant height, ear height, cob length, cob girth, grain weight, shelling percent and test weight. Out of 10, 3 principal components (PCs) exhibited greater than 1.00 eigen value, and explained 71.76% cumulative variability among the traits studied. The PC1 displayed 42.25% while, PC2 and PC3 exhibited 15.20% and 14.31% variability respectively among the inbreds for the traits under study. This study will be helpful in identifying the variability contributing parameters and selection of suitable genotypes for breeding and utilization in maize improvement for yield attributing traits.

Keywords: Component, analysis, maize, normal, *Zea mays* L.

Introduction
Maize (*Zea mays* L.) is one of the popular crops grown in the world, ranking second to wheat and followed by rice (Vasal (2000)) [8]. It occupies an important position in the world economy as food, feed, and industrial grain crop. It is a staple food for several million people in the developing world where they derive their protein and calorie requirements from it. The right choice of maize genotypes for a given region is a crucially important step to obtain high grain yield of different maize inbreds (Ali *et al*. 2013) [4]. Principal component analysis is a sort of multivariate analysis where canonical vectors or roots representing different axes of differentiation and amount of variation accounted for each of such axes, respectively, are derived (Rao, 1952) [6]. It is called principal component analysis as it reflects the importance of the largest contributor to the total variation at each axis of differentiation. It was first coined by Pearson (1901) [3], and developed independently by Hotelling (1933) [1]. The central idea of principal component analysis is to reduce the dimensionality of a data set in which there are a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This reduction is achieved by transforming to a new set of variables, the principal components, which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables. Computation of the principal components reduces to the solution of an eigenvalue-eigenvector problem for a positive-semi-definite symmetric matrix.
This study was conducted with the aim to identify better performing maize genotypes with the help of principal component analysis of the 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 experimental material comprised of thirty-two genotypes of maize in *Kharif,* 2019-20 at Maize Research Farm, Bihar Agricultural University, Sabour (Bhagalpur). The experiment was laid in randomized complete block design with three replications. Individual plot size measured 9 m × 1.5 m with 4 rows planted at a spacing of 60*25 cm. Recommended doses of fertilizers were applied. Data for ten quantitative traits were recorded *viz.* days to 50% tasseling, days to 50% silking, days to 75% brown husk, plant height, ear height, cob length, cob girth, grain weight, shelling percent and test weight. The days to 50% tasseling, days to 50% silking and days to 75% brown husk were recorded on plot basis and rest of the traits

Corresponding Author:
SS Mandal
Department of Plant Breeding
and Genetics, Bihar Agricultural
University, Sabour, Bhagalpur,
Bihar, India

were recorded for five randomly selected plants in each replication. Principal Components are generally estimated from correlation matrix or covariance matrix. Here, correlation matrix was used to generate the principal components. The Principal Component Analysis was conducted using STAR software.

Table 1: Genotypes used in the study

Code	Genotypes
1	VL18510
2	VL162579
3	VL162212
4	VL18523
5	VL18552
6	VL18526
7	VL18688
8	VL18604
9	VL18610
10	VL175847
11	VL172395
12	VL175824
13	VL162291
14	VL18665
15	VL162542
16	VL18670
Code	Genotypes
17	VL18668
18	VL162552
19	VL18673
20	VL143996
21	VL153209
22	VL144011
23	VL176197
24	VL176261
25	VL176005
26	VL176346
27	VL176185
28	VL144287
29	VL175873
30	VL18733
31	KL155958
32	BML-7

Source **of genotypes:** CIMMYT, Hyderabad (India)

Result and Discussion

Principal component analysis (PCA) is a multivariate technique that analyzes a data table in which observations are described by several inter-correlated quantitative dependent variables. It helps to eradicate redundancy in data sets due to regular variation occurring regularly in the crop species (Majiand Saibu, 2012) [2] and (Ramakrishnan *et al*., 2016). PCA is the oldest and best-known technique of multivariate data analysis.

Minimum, maximum and average values were calculated for the analyzed genotypes, as well as standard deviation as main indicators of variability (Table- 2).

Table 2: Basic statistics for various characters of maize genotypes.

	Minimum	Maximum	Mean	Standard deviation
Days to 50% tasselling	106.5	120.0	114.5	3.5
Days to 50% silking	110.0	124.5	118.6	3.8
Days to 75% brown husk	146.0	167.5	158.9	6.0
Plant height	71.8	133.8	96.2	13.3
Ear height	14.8	66.0	35.7	13.1
Cob length	85.0	161.0	122.3	20.8
Cob girth	22.0	44.0	32.7	5.3
Grain weight	8.5	122.5	47.0	28.9
Shelling percent	23.0	85.5	68.4	15.9
Test weight	129.7	314.7	229.5	51.9

The first PC accounts for as much of the variability in the data as possible and each succeeding component accounts for as much of the remaining variability as possible. Out of 10, 3 principal components (PCs) exhibited more than 1.00 eigen value, and showed about 71.76% cumulative variability among the traits studied. Hence, these 3 PCs were given due importance for further explanation. Eigen values of three principal component axes and percentage of variation accounting for them obtained from the principal component analysis are presented in Table 3. The PC1 showed 42.25%

while, PC2 and PC3 exhibited 15.20% and 14.31% variability respectively among the inbreds for the traits under study (Table 3).

Further, the scree plot is also shown in Fig. 1 This exhibit plots of 10 Principal Components used in this study. Starting with the first PC, the plot slopes steeply downward initially and then slowly become an approximately horizontal line. The point at which the curve first begins to straighten out is considered to indicate the maximum number of components to extract. As seen in this Fig. 1, 3 Principal Components were considered. Therefore, these results illustrate the need for multiple decision criteria in deciding the number of components to be retained.

Table 3: Eigen value, contribution of variability and eigen vectors for the principal component axis of maize.

Parameters	PC1	PC2	PC3
Eigen Values	4.225	1.520	1.431
% Variance explained	42.250	15.200	14.310
Cumulative% Variance	42.250	57.450	71.760
Days to 50% tasselling	0.403	0.165	-0.365
Days to 50% silking	0.393	0.183	-0.349
Days to 75% brown husk	0.186	0.045	-0.354
Plant height	-0.154	0.619	-0.083
Ear height	-0.284	0.420	-0.138
Cob length	-0.205	0.391	0.426
Cob girth	-0.349	-0.046	-0.215
Grain weight	-0.428	0.116	-0.280
Shelling percent	-0.280	-0.092	-0.521
Test weight	-0.346	-0.449	-0.129
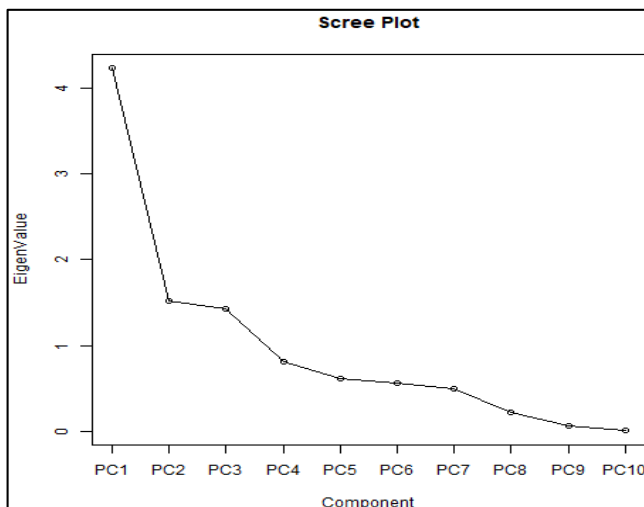
Fig 1: Scree Test for Component Analysis

The first principal component vector (PC I) observed 42.25% of the total variability. It could be seen from the first vector that the characters like days to 50% tasseling (0.403), days to 50% silking (0.393) and days to 75% brown husk (0.186) displayed maximum positive contribution towards genetic divergence, while, the remaining characters like plant height (-0.154), ear height (-0.284), cob length (-0.205), cob girth (-0.349), grain weight (-0.428), shelling percent (-0.280) and test weight (-0.346) displayed negative contribution. However, the second vector (PC II) accounted for 15.20% of total variation present in the population. The major characters that contributed to the second component include traits like days to 50% tasseling (0.165), days to 50% silking (0.183), days to 75% brown husk (0.045), plant height (0.619), ear height (0.420), cob length (0.391) and grain weight (0.116)

had showed maximum positive contribution however, traits like cob girth (-0.046), shelling percent (-0.092) and test weight (-0.449) had maximum negative contribution towards genetic diversity.

Table 4: Principal component scores of different genotypes

| Genotypes	PC1	PC2	PC3
1	1.389	-1.099	1.848
2	-0.980	-0.773	2.742
3	-0.227	0.008	1.343
4	-0.643	0.463	-0.329
5	1.856	0.876	2.041
6	0.038	-1.295	-0.718
7	0.373	-0.567	-0.587
8	-1.387	-1.438	0.549
9	-3.137	-0.656	0.371
10	-1.790	-0.125	-1.036
11	2.080	-0.139	1.919
12	-0.511	-0.521	-0.899
13	3.579	0.249	0.399
14	2.345	-0.280	-1.096
15	0.896	-1.211	-1.202
16	-5.560	0.680	0.600
Genotypes	PC1	PC2	PC3
17	-0.659	-0.491	-0.635
18	2.206	-0.017	-1.815
19	-2.010	4.094	-0.987
20	3.528	1.119	-0.133
21	-0.223	-1.000	-2.278
22	-0.380	0.010	-0.227
23	0.394	-0.639	0.781
24	-1.256	-0.518	0.520
25	1.350	-1.258	-0.793
26	-0.309	1.133	-0.256
27	2.582	1.141	0.383
28	0.052	-0.208	-1.893
29	-1.723	-1.202	-0.443
30	0.938	-0.461	1.263
31	1.272	3.321	0.026
32	-4.084	0.805	0.542

PC III accounted 14.31% of the total variability. The traits like days to 50% tasseling (-0.365), days to 50% silking (-0.349), days to 75% brown husk (-0.354), plant height (-0.083), ear height (-0.138), cob girth (-0.215), grain weight (-0.280), shelling percent (-0.521) and test weight (-0.129) had maximum negative contribution towards genetic diversity. However, single character, cob length (0.426) displayed positive contribution towards genetic diversity.

High principal factor score for a particular genotype in a particular factor denotes high values for the variables in that particular genotype. Table 4 revealed that genotypes 1, 5, 11, 13, 14, 18, 20, 25, 27 and 31 had high principal component scores in factor 1 indicating that they had high plant height and ear height which, this principal factor is representing. Similarly, principal component scores of 19, 20, 26, 27 and 31 genotypes were high in factor 2 indicating high cob length and grain weight. Principal component scores of 1, 2, 3, 5, 11 and 30 were high in factor 3 indicating high days to 50% tasseling, days to 50% silking, days to 75% brown husk, plant height, ear height and cob length. Figure 2 displays the distribution of maize genotypes for the first two principal components. Genotypes that were distant from the other genotypes indicates their usefulness in breeding programmes. However, the biplot diagram (Figure-3) between PC 1 and PC2 explained the distribution and the nature of diversity for

both variables and the genotypes. The loading plot depicted that almost all the genotypes and variables showed high degree of variation. Similar reports were observed by Ravi *et al.,* 2018 [7].
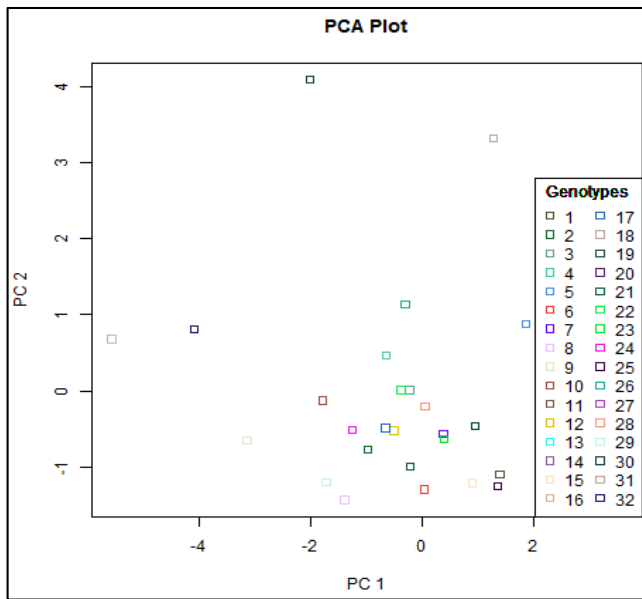
Fig 2: Distribution of maize genotypes for the first two principal components
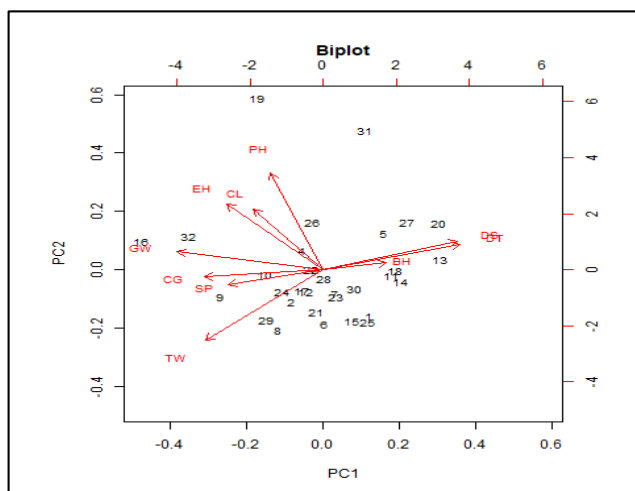
Fig 3: Biplot diagram of principal components

Conclusion

Three principal components (PCs) exhibited more than 1.00 eigen value, and showed about 71.76% cumulative variability among the traits studied. The PC1 showed 42.25% while, PC2 and PC3 exhibited 15.20% and 14.31% variability respectively among the inbreds for the traits under study. The biplot diagram exhibited high variability among the genotypes and between the parameters. In this regard, the study will be helpful in identifying the variability contributing parameters and selection of suitable genotypes for breeding and utilization in crop improvement for yield attributing traits.

References

1. Hotelling H. Analysis of a complex of statistical variables into principal components. J Educ Psychol 1933;25:417-441.
2. Maji AT, Shaibu AA. Application of principal component analysis for rice germplasm characterization and evaluation. Journal of Plant Breeding and Crop Science 2012;4(6):87-93.
3. Pearson K. On lines and planes of closest fit to systems of points in space. Philos Mag A 1901;6:559-572.
4. Ali Q, Ahsan M, Ali F *et al*. Heritability, heterosis and heterobeltiosis studies for morphological traits of maize (*Zea mays* L.) seedlings, Advanced Life Sciences, 2013;1(1):52-63.
5. Ramakrishnan P, Babu C, Iyanar K, Manivannan N. Principal component analysis for evaluation of Guinea grass (*Panicum maximum* Jacq.) germplasm accessions. The Bioscan 2014;11(4):2859-2862.
6. Rao CR. Advanced statistical methods in biometrical research. John Wiley and Sons Inc., New York 1952, 236-272.
7. Ravi Yugandhar P, Suneetha Kota Usha Kiran B, Sridhar M. Principal component analysis for Agromorphological and quality characters in germplasm of rice (*Oryza sativa* L.). International Journal of Advanced Biological Research 2018;8(2):268-273.
8. Vasal SK. Quality maize story. In: Improving Human Nutrition through Agriculture. The Roll of International Agricultural Research. A workshop hosted by IRRI, Philippines, Organized by International Food Policy Institute, Los Banos, USA 2000.