Ashish Patel
Department of Genetics and Plant Breeding, Collage of Agriculture, Raipur, Chhattisgarh, India

Abhinav Sao
Department of Genetics and Plant Breeding, Collage of Agriculture, Raipur, Chhattisgarh, India

Sunil Nair
Department of Genetics and Plant Breeding, Collage of Agriculture, Raipur, Chhattisgarh, India

Jageshwar Mandavi
Department of Genetics and Plant Breeding, Collage of Agriculture, Raipur, Chhattisgarh, India

Nishesh Tamrakar
Department of Genetics and Plant Breeding, Collage of Agriculture, Raipur, Chhattisgarh, India

Principal component analysis for eight quantitative traits in 55 indigenous rice germplasm (*Oryza sativa* L.)

Ashish Patel, Abhinav Sao, Sunil Nair, Jageshwar Mandavi and Nishesh Tamrakar

Abstract
Rice is the dietary staple for more than one third of the world's population and it is a model cereal species that can be used to compensate the world food hunger. In the present study, fifty five genotypes were studied for the eight quantitative traits. Principal Component Analysis was utilized to estimate the relative contribution of various traits for total variability. Principal component analysis revealed that 1st two component with Eigen value greater than 1 accounted 65.38% of total variation. Rotated component matrix showed that PC 2 accounted mainly for yield and yield related character, whereas the character accounted in PC1 which attained high PC score, could be effectively used in the hybridization program to create early duration genotypes.

Keywords: Rice, genotype, principal component

Introduction

Rice is the world's principal food crop, as it is dietary staple for more 67% of the world's population. Rice is the holder of two significant titles, i.e., the most significant food crop which assure the calorific needs of most of the worldwide population and a model cereal species filling in as the hereditary stage to study the novel qualities in rice. Because of independence in production, rice quality is vital to people involved in creating, handling and consuming rice and the most significant quality parts incorporates appearance, processing, cooking and handling quality and nutritional quality (Koutroubas *et al*., 2004) [2]. Rice has one of the largest germplasm collections in the world. This accessible collection of diverse cultivated and wild rice germplasm has made great contributions to rice breeding. The prolific idea about genetic variability present in the genepool is a prerequisite to adapt an efficient and valuable breeding approach. Characterization of this existing variability and realignment of characters in them through selective breeding might go a long way in meeting the existing and emerging challenges that threaten the world food security (Vanaja and Babu, 2006) [11]. Principal Component Analysis is one of the important tools used for identifying the plant characters that categorize the distinctiveness among the promising genotypes. PCA helps to eradicate redundancy in data sets due to regular variation occurring regularly in the crop species (Maji and Saibu, 2012) [4] and (Ramakrishnan *et al*., 2016) [6]. Hence, the importance of PCA is considered and the investigation was carried out in the rice germplasm.

Material and Methods

Fifty five genotypes of rice were used to reveal the genetic relationship between the different accessions. The experiment was conducted at Research Cum Instructional Farm, College of Agriculture, IGKV, Raipur. The experiment was laid out in augmented design with no replication (only check were replicated). The list of genotypes represented in Table 1. The Principal Component Analysis was carried out to identify plant traits that contribute most of the observed variations among the genotypes. Mean values of 55 genotypes for 8 quantitative traits were used. Principal Components are generally estimated from correlation matrix or covariance matrix. The analysis was conducted using XLSTAT software.

Results and Discussion

The main purpose of principal component analysis is reduce the dimension of a large dataset, increasing interpretability but at the same time minimizing information loss. Eigen vector values, percentage of variance and the cumulative percentage are presented in Table 2.

**Corresponding Author:
Ashish Patel**
Department of Genetics and Plant Breeding, Collage of Agriculture, Raipur, Chhattisgarh, India

In this case, 2 components had Eigen values greater than 1.0. PCA 1 and 2 had Eigen values of 3.141 and 2.049, respectively. Percentage of variance for the two factors was 39.26 and 26.12 per cent. Together they accounted for 65.38% of the variability of the genotypes used for the diversity analysis. Similar type of finding also reported by Saibu *et al.* (2017) and Salem *et al.* (2021) [10, 8].

Rotated component matrix showed that PC1 exhibited highest variability 39.16% with highly loaded character *viz*, days to maturity (0.950) days to 50% flowering (0.948) plant height (0.830) panicle length (0.610). This result showed that PC1 is mainly depended on days to maturity and days to 50% flowering. PC2 contained 27.26% variability with high loaded character such as grain yield per plant (0.827), 100 grain wt. (0.786), number of effective tillers (0.729) and panicle length (0.458). Similar type of finding also reported by Sanni *et al.* (2012), Radhamani *et al.* (2015) and Lakshmi *et al.* (2019) [9, 5, 3]. Rotated component matrix explained each principal component separately with various yield character presented in Table 2.

The distribution of the scores for the 8 quantitative traits in the Principal Component Analysis were broad showing the large diversity. The analysis explained the variance structure through a few linear combination of the variables, proportion of variability. Thus on the basis of factor loadings, an exact picture of the component traits that are contributing maximum variability is obtained. The information on genetic diversity are essential for shaping breeding strategies, heterotic grouping and to predict future hybrid performance (Acquaah, 2012) [1].

Scree plot is a graph which order the Eigen valve from largest to smallest and also showed the % of variability in terms of Eigen value and principal component. PC 1 showed 39.26% variability with Eigen value 3.14 and then the graph gradually decline for other PC. Steep curve followed by bend and then straight found for different PCs. This graph (fig: 2) showed clear cut explanation of variability which was maximum for PC1. The biplot diagram (Fig. 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].

Table 1: List of the germplasm used in the study

| No	Accessions no	Accessions name	No	Accessions no	Accessions name
1	CGR: 252	Lohandi	28	CGR: 366	Sathaka
2	CGR: 254	Beda Luchai	29	CGR: 370	Samsarlu
3	CGR: 255	Machhari Kata	30	CGR: 377	Sariya
4	CGR: 260	Malchi	31	CGR: 5683	Bans Chare
5	CGR: 264	Malwi	32	CGR: 5687	Bans Kot
6	CGR: 267	Mathuli(1)	33	CGR: 5688	Bans Kothiya
7	CGR: 270	Mekra Ghol	34	CGR: 5691	Bans Pati
8	CGR: 271	Minda Wadalu	35	CGR: 5702	Bans Patri
9	CGR: 272	Modki	36	CGR: 5706	Bans Phool
10	CGR:274	Bhainsa Mindariya	37	CGR: 5707	Bans Por
11	CGR: 276	Nagina -22	38	CGR: 5708	Bansmati
12	CGR: 279	Newa Dhan	39	CGR: 5716	Baragi
13	CGR: 280	Mewsega	40	CGR: 5772	Banreg
14	CGR: 281	Newa Singo	41	CGR: 5782	Barhi
15	CGR: 285	Nela Vensi	42	CGR: 5793	Barah Kisma
16	CGR: 287	Nella Wadalu	43	CGR: 5795	Barik
17	CGR: 289	Newari Deshi	44	CGR: 5799	Bar Khutiya
18	CGR: 294	Pande	45	CGR: 5800	Barro
19	CGR: 301	Parewa	46	CGR: 5801	Barwi
20	CGR: 315	Karan Phool	47	CGR: 5802	Barungi
21	CGR: 324	Pora	48	CGR: 5803	Basant
22	CGR: 329	Pora Dhan	49	CGR: 5807	Basengi
23	CGR: 332	Kariya Pora	50	CGR: 5808	Basin
24	CGR: 334	Sukal Pora	51	Indira Bhrani Dhan 1	54 Chattisgarh Zinc Rice
25	CGR: 349	Potasu	52	Indira Aerobic 1	55 Samleshwari
26	CGR: 352	Rakhi	53	Danteshwari	

Table 2: Eigen value, contribution of variability and factor loadings for the Principal Component axes

	Principal component	
	PC1	PC2
Eigen value	3.141	2.093
Variability	39.26	26.12
Cumulative %	39.26	65.38
Traits	**Factor loading after varimax rotation**	
Days to 50% flowering	0.948	-0.076
No. of effective tiller	-0.318	0.729
Plant height	0.830	-0.261
Panicle length	0.610	0.458
Days to maturity	0.950	-0.085
Grain length width ratio	0.211	-0.240
100 grain weight	-0.097	0.786
Yield per plant	0.192	0.827

Note: Figure in bold letter showed highly loaded traits/character in respective PC.
Extraction method: principal component analysis
Rotation method: varimax with Keiser normalization
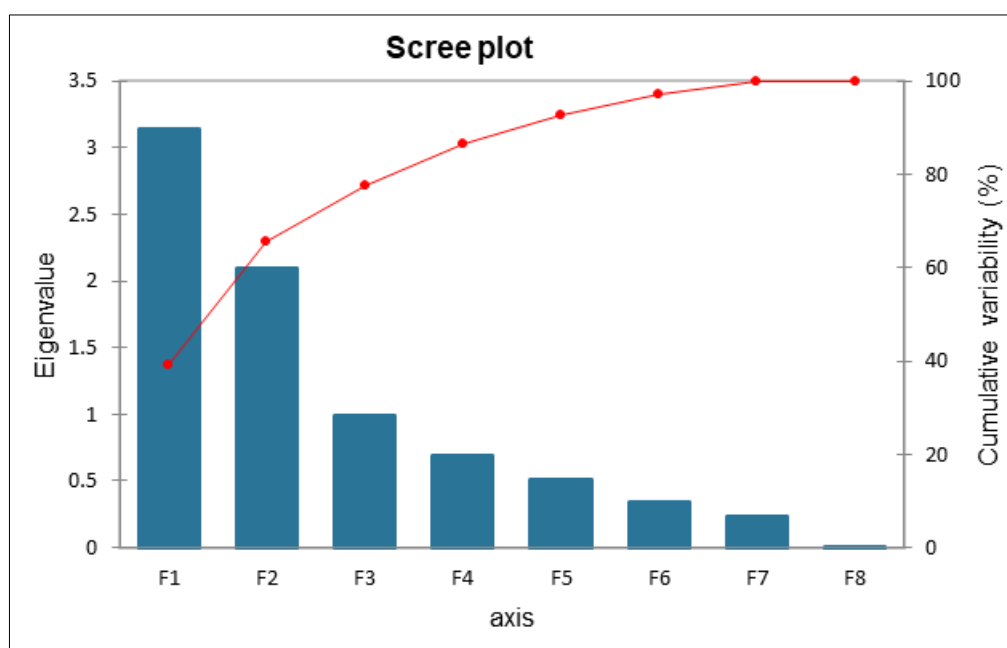
Fig 1: Scree plot showing Eigen value and cumulative variability with their respective PCs

The Biplot diagram (fig:3) 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 related traits.

The result of Biplot analysis presented in fig 3. The result showed that the 2 PCs divided into 2 cluster. The result of box plot revealed that cluster 1st was mainly characterized by days to 50% flowering, plant height, Panicle length, days to maturity, The 1st cluster consisted Lohandi, Beda Luchai, Macchari Kata, Malchi, Malwi, Mathuli (1), Mekra Ghol, Minda Wadalu, Modki, Bhainsa Mindariya, Nagina -22, Newa Dhan, Mewsega, Newa Singo, Nela Vensi, Nella Wadalu, Newari Deshi, Pande, Parewa, Karan Phool, Pora, Pora Dhan, Sukul Pora, Potasu, Rakhi, Ranikajal, Sathaka, Samsarlu, Sariya, Indira Bhrani Dhan 1, Indira Aerobic 1, Danteshwari, Chhattisgarh Zinc Rice 1 and Samleshwari while 2nd cluster was mainly characterized by number of effective tillers 100 grain weight and grain yield per plant. 2nd cluster consisted Banreg, Basin, Bans Chare, Bansmati, Bans Por, Barhi, Barah Kism, Baragi, Bans Phool, Bar Khutiya,

Basngi, Barro, Barwi, Basant, Bans Patri, Banspati, Barik, Bans Kothiya, Bans Kot and Barungi. So this result suggest that for early duration rice variety we must select the germplasm from the 1st cluster, while constructing a breeding programme for yield improvement 2nd cluster landraces should be select as a parent. The result of Dendrogram also supported by Biplot analysis.

Principal component score for all the accessions were estimated and presented in Table 3. The high score of accessions can be utilized for precise selection of parent, whose intensity is determined by variability explained by each PCs. Indira Aerobic 1, Chhattisgarh Zinc Rice 1, Indira Bharni Dhan 1, Danteshwari and Samleshwari had high PCs score which is present in PC2. It mean these accession are high PC score for number of effective tillers, 100 grain weight and grain yield per plant. Bans Chare, Basant, Bans Kot, Basengi and Bans Pati had high PC score for PC1 and highly loaded for Days to 50% flowering, Panicle length, Plant height and days to maturity. On the basis of top 5 PC score, accessions are selected and depicted in Table 4

Table 4: List of selected accessions in each PCs on the basis of top 5 PC score

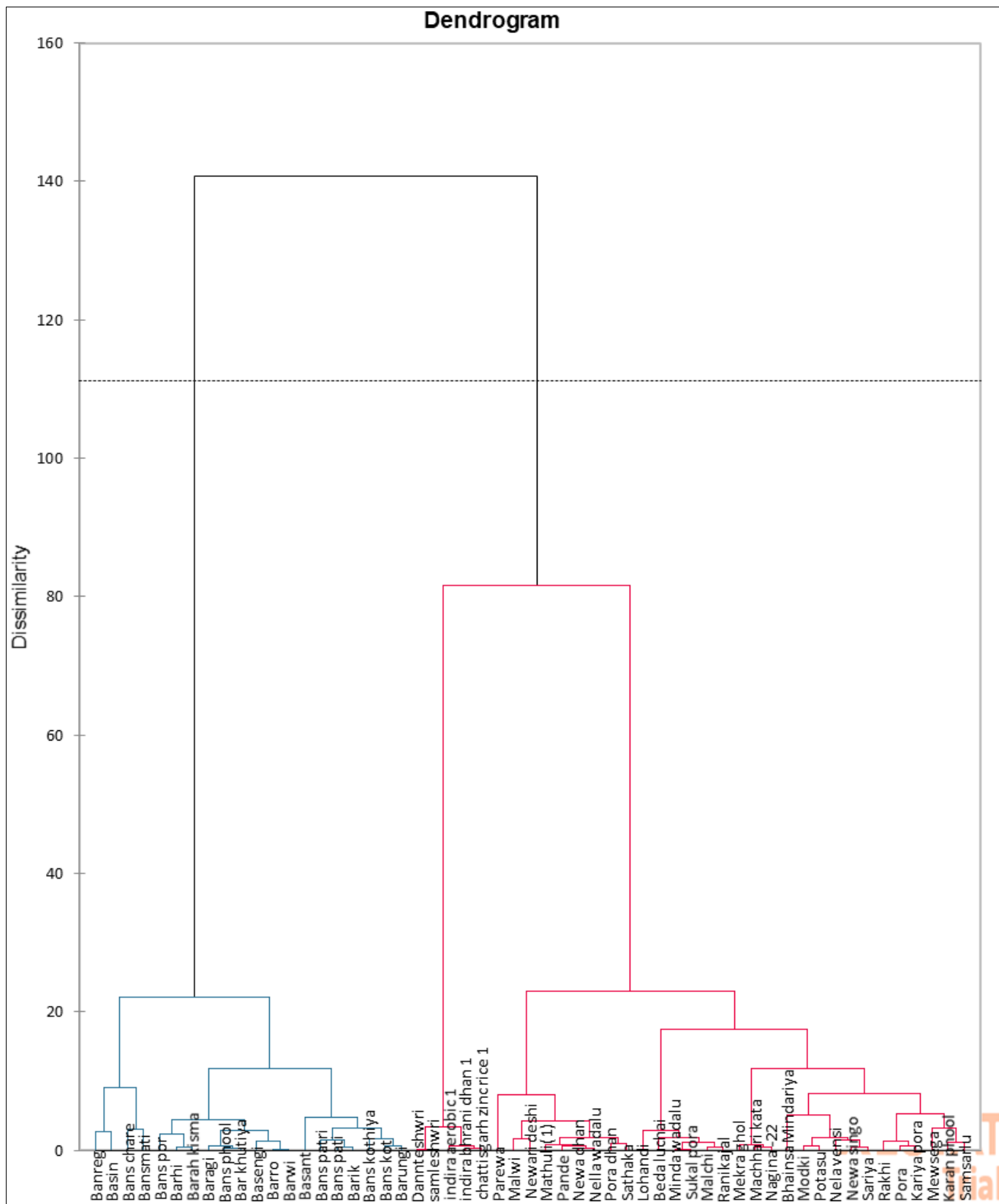| S. No.	PC1	PC2
1	Bans Chare (3.652)	Indira Aerobic 1 (3.519)
2	Basant (3.115)	Chattisgarh Zinc Rice 1 (3.519)
3	Basengi 9(2.671)	Indira Bharni Dhan 1 (3.146)
4	Bans Kot (2.595)	Danteshwari (2.356)
5	Bans Pati (2.588)	Samleshwari (2.292)

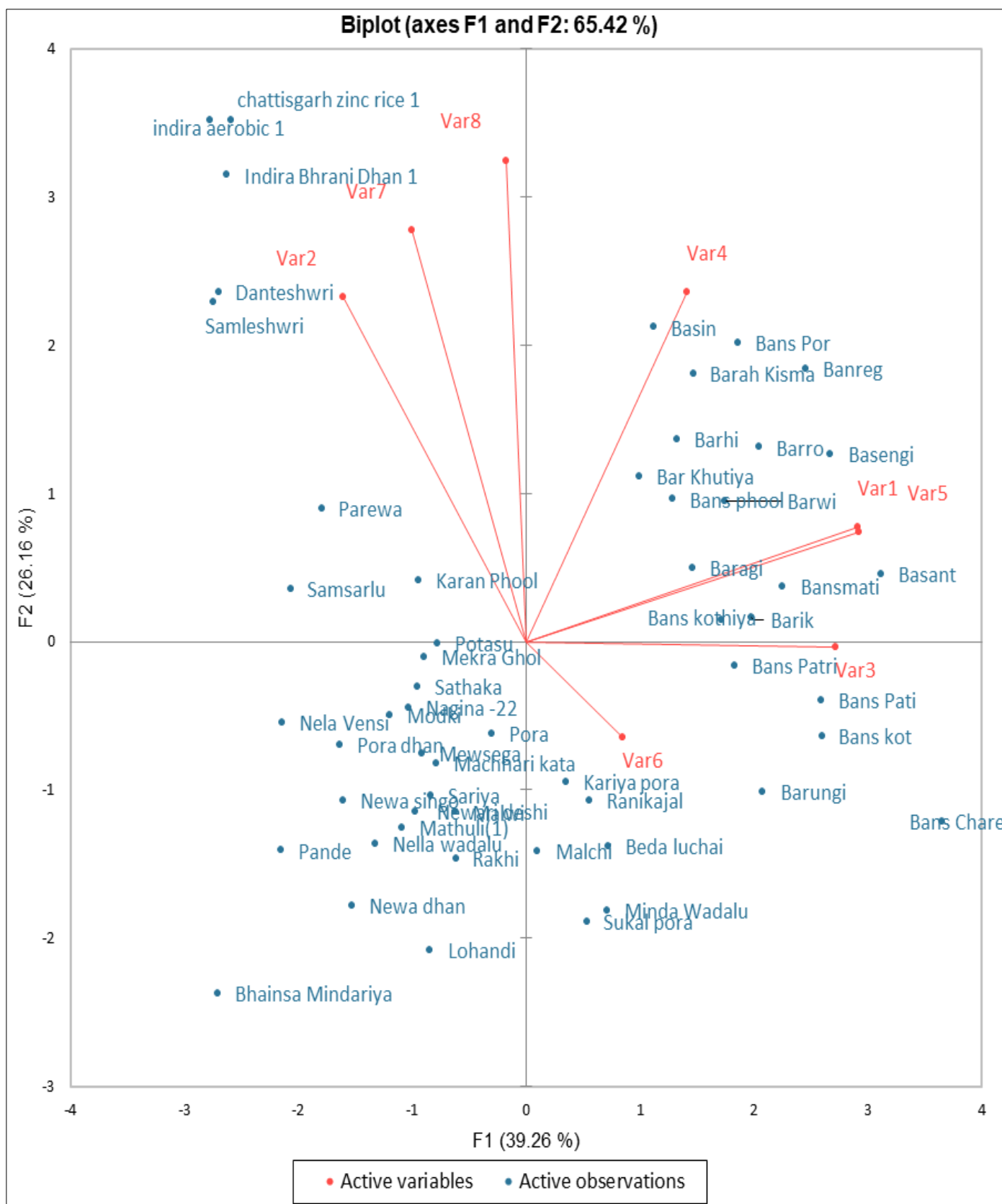**Fig 2:** Dendrogram showing PCs with different classes

Fig 3: Biplot analysis for 55 germplasm of rice with 2 PCs

Table 3: Principal component score of different accessions of rice

No.	Accessions name	PC1	PC2	No.	Accessions name	PC1	PC2
1	Lohandi	-2.079	-0.844	29	Samsarlu	0.358	-2.067
2	Beda Luchai	-1.380	0.725	30	Sariya	-1.037	-0.837
3	Machhari Kata	-0.823	-0.784	31	Bans Chare	-1.217	3.652
4	Malchi	-1.414	0.099	32	Bans Kot	-0.637	2.595
5	Malwi	-1.151	-0.625	33	Bans Kothiya	0.146	1.716
6	Mathuli (1)	-1.256	-1.093	34	Bans Pati	-0.395	2.588
7	Mekra Ghol	-0.101	-0.897	35	Bans Patri	-0.165	1.833
8	Minda Wadalu	-1.814	0.709	36	Bans Phool	0.961	1.280
9	Modki	-0.497	-1.200	37	Bans Por	2.019	1.863
10	Bhainsa Mindariya	-2.376	-2.703	38	Bansmati	0.371	2.253
11	Nagina -22	-0.449	-1.035	39	Baragi	0.496	1.461
12	Newa Dhan	-1.780	-1.528	40	Banreg	1.843	2.456
13	Mewsega	-0.755	-0.912	41	Barhi	1.363	1.325
14	Newa Singo	-1.073	-1.612	42	Barah Kisma	1.806	1.465
15	Nela Vensi	-0.546	-2.146	43	Barik	0.161	1.974
16	Nella Wadalu	-1.360	-1.326	44	Bar Khutiya	1.117	0.997
17	Newari Deshi	-1.143	-0.969	45	Barro	1.311	2.045
18	Pande	-1.410	-2.152	46	Barwi	0.950	1.743
19	Parewa	0.897	-1.792	47	Barungi	-1.012	2.077
20	Karan Phool	0.412	-0.948	48	Basant	0.454	3.115
21	Pora	-0.623	-0.299	49	Basengi	1.263	2.671
22	Pora Dhan	-0.699	-1.635	50	Basin	2.120	1.122
23	Kariya Pora	-0.947	0.346	51	Indira Bhrani Dhan 1	3.146	-2.631
24	Sukal Pora	-1.890	0.532	52	Indira Aerobic 1	3.519	-2.778
25	Potasu	-0.012	-0.776	53	Danteshwri	2.356	-2.700
26	Rakhi	-1.462	-0.617	54	Chattisgarh Zinc Rice1	3.519	-2.588
27	Ranikajal	-1.072	0.555	55	Samleshwri	2.292	-2.747
28	Sathaka	-0.304	-0.952				

Conclusion
PCA offer a large breeding population in a small data set in order to creating a group based on similarity in one or more characters. It will help in choice of parent in hybridization program. In the present study PCA revealed high level of genetic variation among the accession and the variable contributing the diversity. So the result will be useful in choosing the parent in hybridization program for early duration rice variety with higher yield performance.

References
1. Acquaah G. Principles of plant genetics and breeding. Second edition. Wiley Blackwell publications, 2012, 127-129.
2. Koutroubas SD, Mazzini F, Pons B, Ntanos. Grain quality variation and relationships with morphological traits in rice (*Oryza sativa* L.) genetic resources in Europe. Field Crops Research. 2004;86:115-130.
3. Lakshmi VI, Sreedhar M, Vanisri S, Anantha MS, Rao LS, Gireesh C. Multivariate analysis and selection criteria for identification of African rice (*Oryza glaberrima*) for genetic improvement of *indica* rice cultivars. Plant Genetic Resources. 2019;17(6):499-505.
4. 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.
5. Radhamani T, Sassikumar D, Packiaraj D, Saraswathi R, Ravi V. Principal component analysis of variation among the induced mutants of rice (*Oryza sativa* L.). Research on Crops. 2015;16(4):698-703.
6. 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.
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. Salem KF, Saleh MM, Aldahak L, Elabd AB. Assessment phenotypic diversity of rice (*Oryza sativa* L.) genotypes by multivariate analysis. Journal of Aridl and Agriculture. 2021;7:52-59.
9. Sanni KA, Fawole I, Ogunbayo SA, Tia DD, Somado EA, Futakuchi K, *et al*. Multivariate analysis of diversity of landrace rice germplasm. Crop Science. 2012;52(2):494-504.
10. Shaibu AA, Uguru MI. Application of Multivariate Analysis to Access Selected Rice Germplasm Phenotypic Diversity. Int. J Exp. Agri. 2017;25:1-8.
11. Vanaja T, Babu LC. Variability in grain quality attributes of high yielding rice varieties (*Oryza sativa* L.) of diverse origin. Journal of Tropical Agriculture. 2006;44(1-2):61-63.