PLANT SCIENCE TODAY
ISSN 2348-1900 (online)
Vol 10(2): 195–204
https://doi.org/10.14719/pst.2091

HORIZON
e-Publishing Group
HePG

RESEARCH ARTICLE

Genetic divergence studies for yield and quality traits in high protein landraces of rice (*Oryza sativa* L.)

B. Bhargavi[1*], Y. Suneetha[1] , J. Aravind Kumar[2] & T. Srinivas[1]

[1]Department of Genetics and Plant Breeding, Acharya N G Ranga Agricultural University, Andhra Pradesh, India
[2] Department of Plant Breeding, ICAR-Indian Institute of Rice Research, Hyderabad, Telangana, India

*Email: bhargavi.sjvb@gmail.com

OPEN ACCESS

ARTICLE HISTORY

Received: 26 August 2022
Accepted: 26 December 2022

Available online
Version 1.0 : 04 February 2023
Version 2.0 : 01 April 2023

Check for updates

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS etc.
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

CITE THIS ARTICLE

Bhargavi B, Suneetha Y, Kumar J A, Srinivas T . Genetic divergence studies for yield and quality traits in high protein landraces of rice (*Oryza sativa* L.). Plant Science Today. 2023; 10(2): 195–204. https://doi.org/10.14719/pst.2091

Abstract

The present study was undertaken to study the extent of genetic diversity in high protein rice landraces with respect to yield, yield components and quality characters. In this direction, 30 high protein rice landraces, collected from different parts of country by ICAR-Indian Institute of Rice Research (ICAR-IIRR), Hyderabad along with the high protein check, CR DHAN 310 were evaluated during *Kharif* 2021 at ICAR-IIRR farm located at International Crops Research Institute of Semi Arid Tropics (ICRISAT), Hyderabad. The study examined the genetic divergence of high protein rice cultures for yield, quality and nutritional parameters. Multivariate analysis techniques of Mahalanobis D^2 and Principal Component Analysis (PCA) were used to estimate the genetic diversity in the experimental material. In Mahalanobis D^2, the 31 high protein rice cultures were divided into six clusters. Cluster I had highest number of rice cultures (19), followed by Cluster III and V with five, four cultures, respectively. The clusters, II, IV, VI were mono-genotypic. It was discovered that grouping of these cultures into several clusters was random and was not related to geographical diversity. Inter-cluster distances between clusters V and VI were maximum. Cluster V had also exhibited higher intra-cluster distance. Further, Cluster VI had showed maximum yield plant^{-1}, grains per panicle^{-1}, zinc content and test weight, while, Cluster V had recorded high protein content. The greatest contribution to genetic divergence was recorded by yield plant^{-1} (21.60%), followed by iron (10.54%) and zinc content (9.54%). In Principal Component Analysis, the first five Principal Components (PCs) with eigen values >1 accounted for cumulative contribution of 67.69% to the total variability. The three traits, yield plant^{-1}, iron content, and amylose content contributed the most to variability. The 2D scatter diagram exhibited 18 different clusters, out of which 11 clusters were mono-genotypic. Mahalanobis D^2 Statistic and PCA concluded maximum genetic diversity between the landraces, JAK 248-3 and JAK 638 with JAK 611.

Keywords

Genetic diversity; grain yield; landrace; Mahalanobis D^2; principal component analysis; protein content quality characters; rice

Introduction

'Rice is Life' for billions of people, and rice production has influenced lifestyle, diet, and economic status of humans across the globe. It is being cultivated in over 100 countries worldwide. Asia produces and consumes 90% of the world's rice. It is referred to as 'Global Grain'' (1) and is the "pre-

dominant staple cereal food of Asia, Latin America and Africa (2)". However, protein malnutrition is a serious health threat in the children of many areas in these countries (3). In India, 80 percent of children under the age of five are malnourished with a daily protein intake of 13-19g only per child (4). Given that the recommended calorie intake for children is 1000-1400 calories per day *i.e.*, of 200-300g rice, with protein accounting for 150-450 calories (5), even a one per cent increase in grain protein content in rice would contribute significant protein to the diet. Rice protein is said to be the best among cereal proteins because it has improved balance of vital amino acids and is easier to digest. "Rice has a higher Protein Digestibility Corrected Amino Acid Score (PDCAAS) of 0.81 (6)" which shows the presence of essential amino acids and overall protein quality. However, the protein content of milled rice grain is typically 6-7 per cent, which is the lowest among cereals such as wheat (12-14 %) and maize (8-9 %). Therefore, efforts are needed to develop high protein rice genotypes with good yield. Besides yield, consideration of grain quality and nutritional characters, including protein content is increasing in the recent years (7). However, the evolution of highly resilient and good quality rice cultivars that are superior to existing varieties is significantly influenced by the level of genetic diversity present within the population. In order to develop the desired recombinants, it is essential to choose excellent and diverse parents for hybridization programmes. Information on genetic diversity with regard to the nature and degree of divergence for yield, yield components, quality, and nutritional characters would therefore be useful for the selection of the appropriate parents for hybridization programmes and the implementation of successful breeding strategies intended at the production of high yielding varieties with good grain and nutritional quality.

Multivariate analysis, like Mahalanobis D^2 statistic is the most commonly and widely used tool for measuring the genetic diversity within a population and identifying of superior and diverse parents. Principal Component Analysis (PCA), is a reduction technique for multivariate data and is used more frequently to assess the significance and contribution of each factor to the over-all variance in addition to providing details on the influence of a specific trait to the total variance. The scatter diagrams show the diversity of genotypes with respect to spatial distance, and each coefficient of eigenvector reveals the degree of contribution of each original variable with which each principal component is associated.

The success of any breeding programme depends on the amount of variability present in the germplasm. The knowledge of genetic divergence is very important in the identification of suitable parents for hybridization for varying conditions. Wider the diversity among the parents, the greater will be the chance of obtaining heterotic combinations and broader would be the spectrum of variability in the segregating generations (8). Though several studies on genetic diversity have been conducted in rice, genetic diversity studies in high protein landraces are lacking and this study would aid the breeders in their efforts and hasten the process of development of high yielding and protein rich rice varieties, which is the need of the hour to combat malnutrition, in a country like India with rice as the staple food. The present study was undertaken in this context to identify the amount of genetic diversity in the present experimental material comprising of protein rich rice landraces for yield, yield components, quality and nutritional traits along with identification of superior and diverse landraces through Mahalanobis D^2 analysis and principal component analysis for use in rice breeding programmes aimed at the development of high protein rice varieties with high yield, good grain and nutritional quality.

Materials and Methods

The experimental material contained 30 high protein rice landraces, collected from different parts of the country and the high protein check, CR Dhan 310 obtained from ICAR—IIRR, Rajendranagar, Hyderabad (Table 1). These cultures were sown at IIRR Farm at ICRISAT, Hyderabad

Table 1. Details of high protein rice genotypes studied in the present investigation

S. No.	Genotype	Origin
1	JAK 14	West Bengal
2	JAK 17	Jharkhand
3	JAK 25	Jharkhand
4	JAK 90	Uttar Pradesh
5	JAK 108	Jharkhand
6	JAK 120	Uttar Pradesh
7	JAK 124	Maharashtra
8	JAK 153	Jharkhand
9	JAK 163	Uttar Pradesh
10	JAK 247	West Bengal
11	JAK 248-3	West Bengal
12	JAK 287	West Bengal
13	JAK 341-2	Maharashtra
14	JAK 355	West Bengal
15	JAK 374	West Bengal
16	JAK 377-3	Maharashtra
17	JAK 390	Maharashtra
18	JAK 400	West Bengal
19	JAK 423	Maharashtra
20	JAK 424	Maharashtra
21	JAK 440	Uttar Pradesh
22	JAK 453	Uttar Pradesh
23	JAK 486	Uttar Pradesh
24	JAK 513-1	Jharkhand
25	JAK 519	Uttar Pradesh
26	JAK 552	West Bengal
27	JAK 595-1	Uttar Pradesh
28	JAK 611	Jharkhand

during *Kharif* 2021 on raised nursery. All suggested measures were followed to get a good crop. The 30 days old seedlings were transplanted in the field and laid out in Randomized Complete Block Design (RCBD) with three replications. Each culture was transplanted in five rows of 4.5m length with 20x15cm spacing. To ensure a healthy crop, recommended practices were followed throughout the crop period. The quantitative traits, namely, yield plant^{-1} and yield component traits, such as plant height, ear bearing tillers plant^{-1}, panicle length and grains panicle^{-1} were recorded from five randomly selected plants. The quality traits, namely, head rice recovery (%) and amylose content; and the nutritional traits, namely, protein, iron and zinc content in addition to test weight were obtained from a random grain sample taken from each plot in each replication. Days to 50% flowering and maturity were however, recorded on plot basis.

Standard statistical techniques were applied to the data collected (9) and diversity analysis was carried out through D^2 statistics (10) and Principal component analysis (11) as detailed below. Window Stat Version 8.5 was used to perform the statistical analysis.

Estimation and significance of Wilk's criterion

Analysis of covariance for the character pairs was estimated on the basis of mean values (9). The estimation of *Wilk's* criterion (12) was done using the following relationship.

$$\Lambda = \frac{(E)}{(E+V)}$$

Where,

Λ = Wilk's criterion

(E) = Dterminant of error matrix and

(E+V) = Determinant of error + variety matrix

Computation of D2 values

For the given combination of i and j genotype, the mean deviation *i.e.*, Y_{ti}-Y_{tj} for t=1, 2…p variables are computed and the D^2 values were calculated as

$$D^2_{ij} = \sum_{t=1}(Y_i^t - Y_j^t)_2$$

Where, Y_i^t is uncorrelated mean value of ith genotype for character 't' Y_j^t is uncorrelated mean value of jth genotype for character 't' D_{ij}^2 is D^2 value between ith and jth genotype.

Grouping of genotypes into various clusters

The grouping of genotypes into different clusters was done using the Tocher's method as described earlier (13). The criterion was that the two genotypes belonging to the same cluster at least on an average show a smaller D^2 value than those belonging to different clusters. For this purpose, D^2 values of all combinations of each genotype were arranged in ascending order of magnitude in a tabular form as described earlier (14).

29	JAK 625	West Bengal
30	JAK 638	Jharkhand
31	CR DHAN 310	Cuttack

Average intra and inter-cluster distance

For the measurement of intra-cluster distances, the formula used was $\Sigma D^2_i/n$ where, $\sum D^2_i$ was the sum of distances between all possible combinations (n) of the populations included in a cluster. The square root of the average D^2 value gave the genetic distance 'D' between the clusters. Based on D^2 values (inter-cluster distance), the cluster diagram was prepared adopting the scale given earlier (13) for rating of the distance.

Cluster means

The cluster mean for a particular trait is the summation of mean values of the genotypes included in a cluster divided by number of genotypes in the cluster.

Contribution of individual characters towards divergence

In all combinations, each character was ranked based on their contribution towards divergence between two entries.

Principal Component Analysis

Principal component analysis was carried according to procedure described earlier (11). In the present study, PCA was performed on the correlation matrix of traits, thereby removing the effects of scale (15).

Eigen values and eigen vectors

The eigen values and eigen vectors were computed from data matrix. The proportion of variation accounted for each principal component (PC) is expressed as the eigen value divided by the sum of the eigen values.

Per cent variance explained for PC1 = Eigen value PC1/ Sum of eigen values

PCA scores for each genotype under concerned PCs were computed and utilized to derive a 2D or 3D (dimensional) scatter plot of individuals (16).

Results and Discussion

The results on genetic divergence of 31 rice landraces for yield, yield components, quality and nutritional traits are presented and discussed hereunder.

Test with Wilk's criterion 'Λ'

Analysis of variance (ANOVA) for the different traits studied is presented in Table 2. The results revealed significant differences between the genotypes for all traits studied. The Wilk's statistic 'Λ' was highly significant with value of 1340.221, justifying the calculation of D^2 values. The results are in conformity with the findings of earlier workers (17) who also reported 'V' statistic to be highly significant with value of 2736.30 in their studies on 107 rice genotypes for yield and yield component traits, justifying the need to calculate D^2 values. Similar results were also reported ear-

Table 2. Analysis of variance for grain yield, yield components, quality and nutritional traits in high protein rice genotypes

Source of variation	d. f.	Yield component traits						
		Days to 50 per cent flowering	Days to maturity	Plant height (cm)	Ear bearing tillers plant^{-1}	Paniclelength (cm)	Grains panicle^{-1}	Test weight (g)
		Mean sum of squares						
Replications	2	0.899	4.00	5.140	0.075	0.916	369.220	10.454
Genotypes	30	237.846**	309.25**	1098.79**	11.387**	10.872**	3221.7**	18.875**
Error	60	26.062	105.18	18.890	0.741	4.262	603.700	7.449

Table 2 Contd...

Source of variations	d. f	Grain yield plant^{-1} (g)	Quality traits		Nutritional traits		
			HRR (%)	Amylose content (%)	Iron content (ppm)	Zinc content (ppm)	Protein Content (%)
		Mean sum of squares					
Replications	2	1.431	19.02	0.468	0.0001	0.062	0.005
Genotypes	30	18.481**	127.017**	39.307**	0.090**	2.222**	1.367**
Error	60	1.119	9.904	10.569	0.001	0.410	0.166

lier (18) in his studies on 33 rice genotypes including red, black and white colour rice for yield and quality characters.

Grouping of rice cultures into different clusters

Using Tocher's method, the 31 cultures were divided into six clusters based on the D^2 value and the results are presented in Table 3 and Fig.1. An analysis of these findings showed Cluster I with 19 genotypes, (JAK 108, JAK 247, JAK 17, JAK 486, JAK 377-3, JAK 595-1, JAK 120, JAK 453, CR DHAN 310, JAK 90, JAK 519, JAK 374, JAK 513-1, JAK 287, JAK 355, JAK 124, JAK 625, JAK 153 and JAK 440), representing rice cultures collected from Jharkhand, Uttar Pradesh, West Bengal, Maharashtra and Cuttack. Cluster III comprised of five cultures (JAK 400, JAK 423, JAK 424, JAK 25 and JAK 341-2) obtained from West Bengal, Maharashtra and Jharkhand, while, Cluster V consisted of four genotypes (JAK 14, JAK 611, JAK 638 and JAK 552) collected from Jharkhand and West Bengal states. Cluster II (JAK 390), Cluster IV (JAK 163) and Cluster VI (JAK 248-3) were observed to be mono-genotypic clusters. No association was observed for geographic location and genetic diversity in the distribution of rice cultures into different clusters. The cultures from various geographical regions were grouped into the same cluster, while, cultures from identical eco-geographical location were grouped in different clusters. This unorganised and random arrangement can be attributed to the movement of genetic material between breeding locations, in addition, to natural and human selection for large, adaptable gene complexes, resulting in genetic drift leading to higher genetic diversity. The results in conformity with the reports of earlier workers (19), who also reported random clustering of 82 rice genotypes into 10 clusters, with no relation to geographic diversity.

Table 3. Clustering pattern of 31 high protein rice genotypes for grain yield, yield components, quality and nutritional characters

Cluster	No. of Genotypes	Genotypes
1 Cluster	19	JAK 108, JAK 247, JAK 17, JAK 486, JAK 377-3, JAK 595-1, JAK 120, JAK 453, CR DHAN 310, JAK 90, JAK 519, JAK 374, JAK 513-1, JAK 287, JAK 355, JAK 124, JAK 625, JAK 153 and JAK 440
2 Cluster	1	JAK 390
3 Cluster	5	JAK 400, JAK 423, JAK 424, JAK 25 & JAK 341-2
4 Cluster	1	JAK 163
5 Cluster	4	JAK 14, JAK 611, JAK 638 & JAK 552
6 Cluster	1	JAK 248-3

Average intra and inter-cluster D2 value

Intra-cluster distances (Table 4 and Fig. 2) are a measure of diversity within the genotypes grouped into a cluster. The intra-cluster D^2 values ranged from 0.000 for the mono-genotypic clusters (Clusters II, IV, and V) to 50.13 (Cluster V). Therefore, genotypes of Cluster V are inferred to be diverse or dissimilar, compared to the genotypes grouped in Cluster I with intra-cluster D^2 value of 36.53 and Cluster III with D^2 value of 34.22. The cultures of Cluster V (JAK 14, JAK 611, JAK 638 & JAK 552) had recorded maximum inter-cluster D^2 value with JAK 248-3, culture of Cluster VI. Hence, hybridization of cultures from these clusters is estimated to result in wide variability for yield, yield components, quality and nutritional parameters aiding in the improvement of high yielding cultures with good grain and nutritional quality. Minimum inter-cluster distance was recorded between the Clusters II and IV (44.67) indicating their closer association and resemblance with respect to the traits studied. The results are in conformity with findings of earlier workers (1) who studied 37 rice genotypes for their genetic diversity for various yield and yield component traits. They reported maximum inter-cluster distance between Cluster XVI

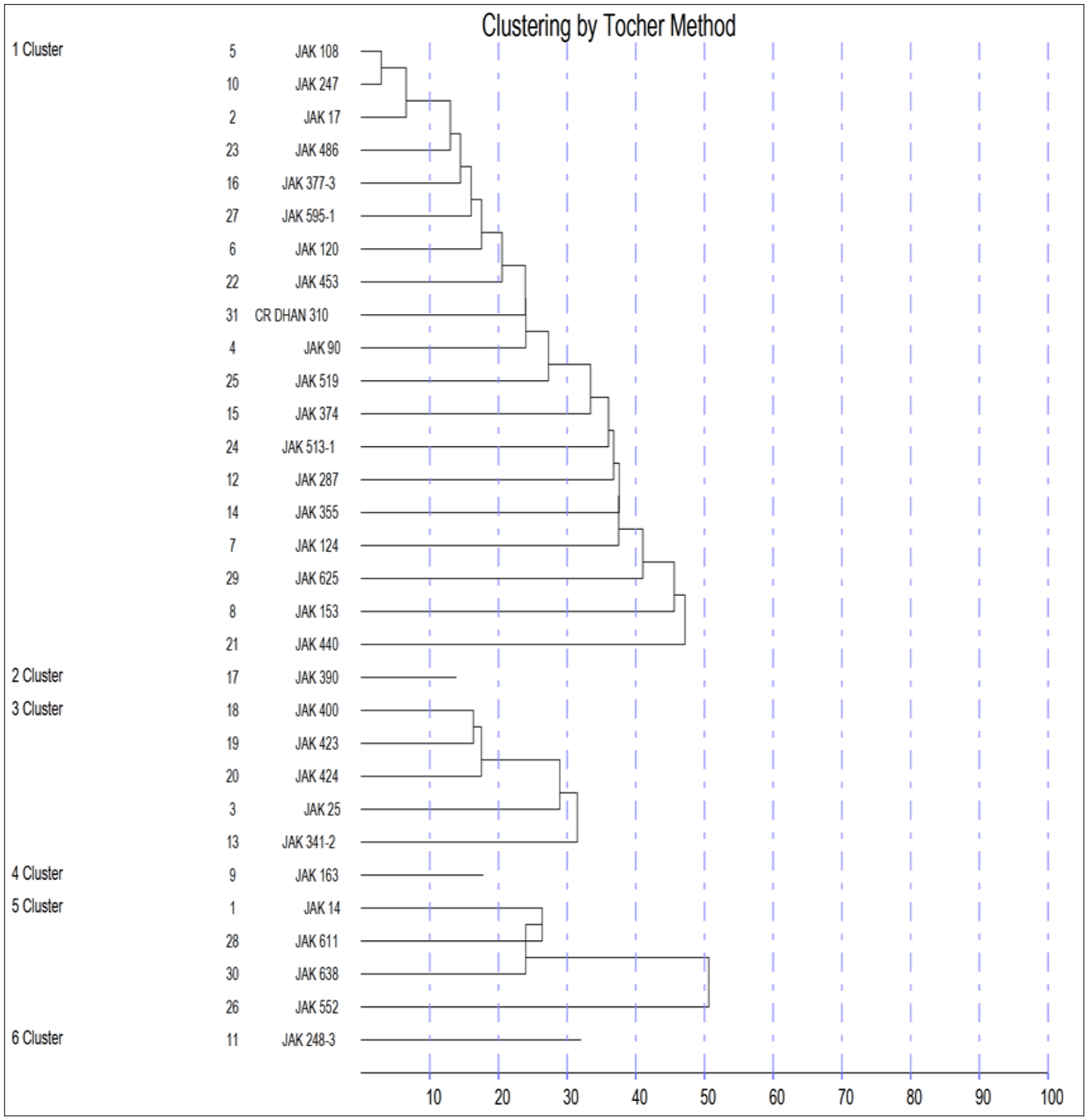

Fig. 1. Dendrogram showing relationship among 31 high protein rice genotypes in six clusters based on Mahalanobis D^2 valus.

Table 4. Average intra-and inter-cluster D^2 values among six clusters of high protein rice genotypes for grain yield, yield components, quality and nutritional characters

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Cluster 1	**36.53**	66.08	75.15	51.75	75.89	103.62
Cluster 2		**0.00**	53.98	44.67	85.09	162.96
Cluster 3			**34.22**	113.37	140.97	69.04
Cluster 4				**0.00**	63.62	191.70
Cluster 5					**50.13**	228.83
Cluster 6						**0.00**

and XVII and minimum inter-cluster distance between Cluster II and III.

Cluster means

Results describe the performance of all cultures consisting

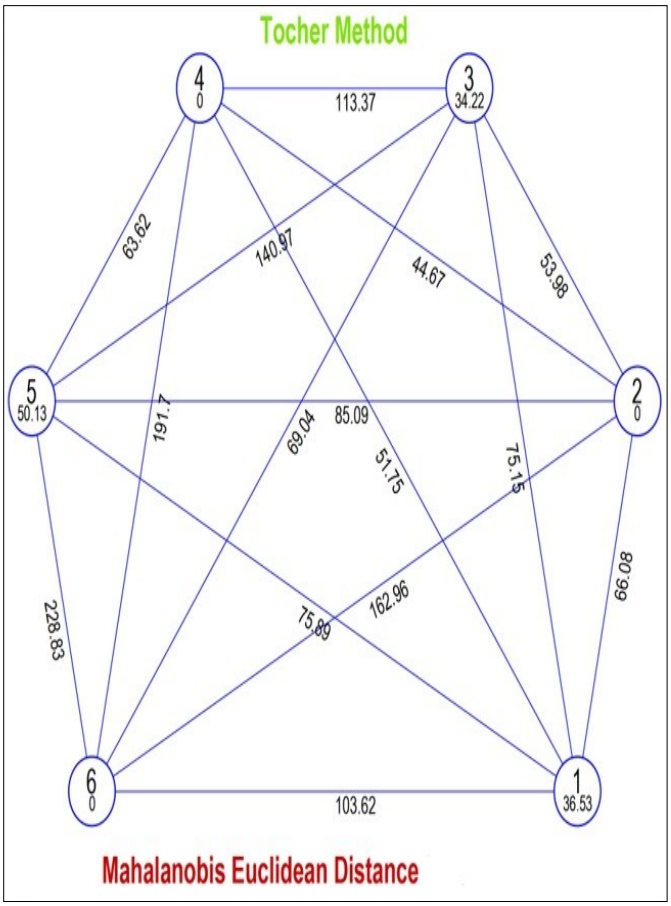

Fig. 2. Intra and inter-cluster distance of 31 high protein rice genotypes in six clusters

in a particular cluster. The estimation of cluster means gives details on suitable donors for a specific trait. Mean performance of clusters for yield, yield parameters, quality and nutritional traits for the 31 cultures evaluated in the experiment are presentedin Table 5. Means of clusters varied from 94.00 days (Cluster IV) to 108.33 (Cluster III) for days to 50 per cent flowering; 120.33 days (Cluster VI) to 134.07 days (Cluster III) for days to maturity; 71.00 cm (Cluster VI) to 141.33 cm (Cluster II) for plant height; 9.00 (Cluster V) to 14.00 (Cluster IV) for ear bearing tillers plant^{-1}; 22.95 cm (Cluster I) to 27.12 cm (Cluster VI) for panicle length; 94.25 (Cluster V) to 207.00 (Cluster VI) for grains panicle^{-1}; 18.00 g (Cluster V) to 23.81 g (Cluster VI) for test weight; 15.78g (Cluster V) to 24.17g (Cluster VI) for grain yield plant^{-1}; 56.70% (Cluster V) to 70.30% for head rice recovery (Cluster II); 18.92% (Cluster V) to 23.73% (Cluster II) for amylose content; 9.16% (Cluster IV) to 9.82% (Cluster V) for protein content; 1.4 ppm (Cluster V) to 2.7ppm (Cluster III) for iron content; and 18.5ppm (Cluster V) to 23.8 ppm (Cluster VI) for zinc content. Cluster VI recorded maximum grain yield plant^{-1} (24.17g), in addition to test weight (23. 81g) along with zinc content (23.8ppm) and the Cluster had also recorded maximum mean for panicle length (27.12 cm) and grains panicle^{-1} (207.00), in addition to low and desirable mean values for maturity (105.33) and plant height (71 cm). Cluster V showed maximum protein content (9.82%). Maximum head rice recovery was observed in Cluster II (70.30%). while Iron content was maximum in Cluster III (2.7 ppm). However, no single cluster had recorded desirable values for all the traits studied. Similar findings were reported earlier (11 and 12). In the context of hybridization programmes aimed at improvement of the desired traits, the selection of cultures from clusters with high mean values is suggested. Therefore, it is recommended to combine between the selected cultures from the divergent clusters in order to carefully integrate all the desired traits. In this path, selection of cultures from clusters V and VI need to be utilised in hybridization programs to obtain high yielding rice cultures with desirable quality and nutritional parameters. The results are in broad agreement with the reports of earlier workers (1). They had also reported high cluster mean values of grain yield per plant, grains per panicle and panicle length for Cluster VII.

Relative contribution of individual characters towards divergence

"Information on the relative contribution of various characters towards divergence was reported to aid the breeder in the choice of parents for hybridization and effective selection (1)". Table 6 and Fig. 3 reveal the results for the percentage contribution of yield, yield components, quality, and nutritional attributes to genetic divergence. These findings showed that yield plant^{-1} (21.60%), followed by iron content (10.54%) had contributed the maximum towards genetic diversity of the high protein rice cultures. However, in the present research, the minimum contribution to the genetic divergence of high protein cultures was found for days to 50% flowering. Grouping of 27 genotypes into four clusters and maximum contribution of yield per

Table 5. Cluster means of 31 high protein rice genotypes for grain yield, yield components, quality and nutritional characters

	DFF	DM	PH	EBT	PL	GPP	TW	HRR	PC	Fe	Zn	AC	GYPP
Cluster 1	102.00	129.07	102.04	11.56	22.95	124.09	19.60	64.07	9.53	0.16	1.87	23.07	20.65
Cluster 2	98.00	123.00	141.33	11.67	25.40	78.33	18.40	70.30	9.77	0.23	2.15	23.73	20.60
Cluster 3	108.33	134.07	111.47	11.53	23.52	101.53	19.30	64.80	9.55	0.27	2.15	22.72	20.58
Cluster 4	94.00	124.00	127.33	14.00	26.72	130.67	22.38	57.90	9.16	0.13	2.32	21.32	23.40
Cluster 5	101.00	124.75	129.21	9.00	23.12	94.25	18.00	56.70	9.82	0.14	1.85	18.92	15.78
Cluster 6	105.33	120.33	71.00	10.33	27.12	207.00	23.81	60.80	9.50	0.24	2.38	21.25	24.17

Table 6. Contribution of different characters towards genetic divergence in high protein rice genotypes

S. No.	Character	Contribution (%)
1.	Days to 50 per cent flowering	4.15
2.	Days to maturity	5.10
3.	Plant height	4.54
4.	Ear bearing tillers plant^{-1}	6.24
5.	Panicle length	5.62
6.	Grains panicle^{-1}	8.94
7.	Test weight	6.33
8.	Head Rice Recovery percentage	5.18
9.	Protein content	4.68
10.	Iron content	10.54
11.	Zinc content	9.54
12.	Amylose content	7.54
13.	Grain yield plant^{-1}	21.60

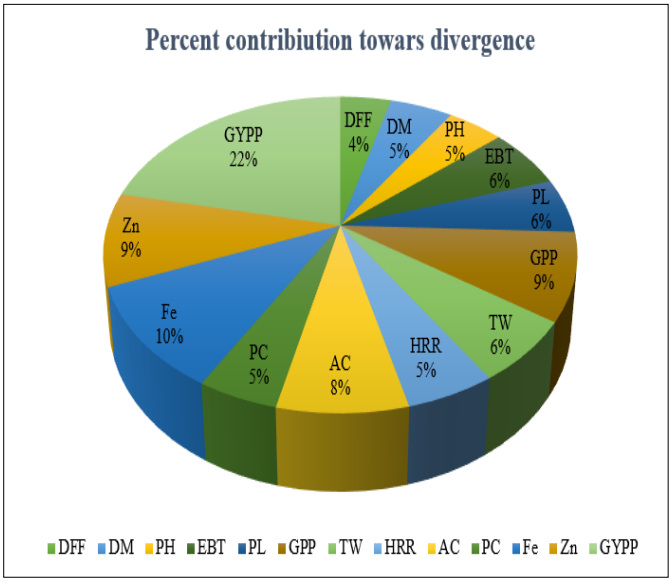

Fig. 3. Contribution of grain yield, yield components, quality and quality characters towards divergence.

DFF= Days to 50 per cent flowering, **DM**= Days to maturity, **PH**= Plant height, **EBT**=Ear Bearing tillers plant^{-1}, **PL**= Panicle length, **GPP**= Grains panicle^{-1}, **TW**= Test weight, **HRR**= Head rice recovery%, **AC**= Amylose content, **PC**=Protein content, **Fe C**= Iron content, **Zn C**= Zinc content, **GYPP**= Grain yield plant^{-1}

plant (20.80%) towards genetic divergence followed by plant height (16.81%), kernel breadth (16.52%) and kernel length (15.10%) was also reported earlier (20). In order to improve the probability of obtaining desired recombinants for breeding of high yielding rice cultures with good grain and nutritional quality, it would be useful to choose divergent parents based on yield plant^{-1} and iron content.

Principal Component Analysis (PCA)

Tables 7-8 and Figure 4 reveal results of PCA. Table 7 contains the eigen values, amount of total variance denoted by principal components of significance, and component loading of diverse traits for the principal components. The first five PCs had eigen values >1. These principal components were noticed to contribute 67.69% of total variability. Similar findings were reported earlier (13). The first PC (PC1) had the greatest contribution to divergence (22.10%). The traits, yield plant^{-1} (0.810), panicle length (0.613), grains panicle^{-1} (0.588), zinc content (0.529), test weight (0.407), iron content (0.382), head rice recovery% (0.335), ear bearing tillers plant^{-1} (0.304), amylose content (0.0243), days to maturity (0.227), and days to 50% flowering (0.209) had positively contributed to the variation in the principal component, whereas other studied characters had negatively contributed. Test weight (0.354), iron content (0.344), yield plant^{-1} (0198), grain panicle^{-1} (0160), amylose content (0.068), and protein content (0.060) were observed to be positively loaded, while other traits were observed to be negatively loaded for PC2, which described 16.60% of the total variance. The PC3 contributed 11.73 percent of total variability and the characters, amylose content (0.634), iron content (0.556), head rice recovery (%) (0.322), days to maturity (0.320), days to per cent flowering (0.269), test weight (0.078), yield plant^{-1} (0.077) and protein content (0.013) had contributed positively for the variation in the principal component, while other characters had contributed negatively. The characters, plant height (0.576), zinc content (0.370), iron content (0.346), ear bearing tillers plant^{-1} (0.340), amylose content (0.256), yield plant^{-1} (0.089), head rice recovery (%) (0.039) and panicle length (0.021) had positive loadings on PC4, whereas other traits had negative loadings. The PC5 contributed 7.80 per cent towards the total variability and the characters, namely, protein content (0.412), zinc content (0.362), ear bearing tillers plant^{-1} (0.289), amylose content (0.216), yield plant^{-1} (0.212), grains panicle^{-1} (0.117) and days to 50 per cent flowering (0.008) were perceived to contribute positively for the variability elucidated by the PC, while other parameters recorded negative loadings for the principal component. Results of PCA revealed yield plant^{-1}, iron content and amylose content as the highest contributing traits to the total variability. Earlier studies (14) on PCA in rice also reported yield per plant as significant contributor to variability. The scores of PCA for 31 high protein rice cultures were computed and measured as three axes as X, Y, and Z and the squared distance of each culture from these three axes was calculated and is depicted in Table 8. The two-dimensional scatter diagram was generated using the PCA scores for 31 cultures that were displayed in a graph (Fig. 4). These results were examined and 18 distinct groups were found. Among these, 11 clusters were found to be mono-genotypic. The pattern of distribution of the genotypes in these clusters was observed to be at random with no reference to geographic diversity as genotypes from different geographical regions were grouped in the same as well as different clusters. The results are in agreement with the findings of earlier workers (21 and 22). Maximum genetic diversity was observed between JAK 248-3 (11), JAK 638 (30) and JAK 611 (28). These genotypes were scattered relatively away from other cultures. The findings are also in agreement with the results obtained through Mahalanobis D^2 Statistic in the present study.

Table 7. Eigen values, proportion of total variance represented by first five principal components, cumulative per cent variance and component loading of different traits in high protein rice genotypes

Component	PC 1	PC 2	PC 3	PC 4	PC 5
Eigen Value (Root)	2.830	1.980	1.530	1.23	1.010
% Var. Exp.	22.100	16.600	11.730	9.460	7.800
Cum. Var. Exp.	22.10	38.700	50.430	59.890	67.690
Days to 50 per cent flowering	0.209	-0.739	0.269	-0.315	0.008
Days to maturity	0.227	-0.693	0.320	-0.298	-0.119
Plant height	-0.324	-0.438	-0.101	0.576	-0.422
Ear bearing tiller plant^{-1}	0.304	-0.619	-0.244	0.340	0.289
Panicle length	0.613	-0.037	-0.467	0.021	-0.336
Grains panicle^{-1}	0.588	0.160	-0.228	-0.510	0.117
Test weight	0.407	0.354	0.078	-0.004	-0.465
Head rice recovery %	0.335	-0.230	0.322	0.039	-0.057
Amylose content	0.243	0.068	0.634	0.256	0.216
Protein content	-0.629	0.060	0.013	-0.047	0.412
Iron content	0.382	0.344	0.556	0.346	-0.028
Zinc content	0.529	-0.095	-0.428	0.370	0.362
Amylose content	0.243	0.068	0.634	0.256	0.216
Grain yield plant^{-1}	0.810	0.198	0.077	0.089	0.212

Table 8. PCA scores of high protein rice genotypes

S. No.	Genotype	PCA I	PCA II	PCA III
		X Vector	Y Vector	Z Vector
1	JAK 14	-16.783	2.052	7.589
2	JAK 17	-13.274	-2.317	9.667
3	JAK 25	-16.388	-3.707	14.359
4	JAK 90	-18.268	-3.220	8.860
5	JAK 108	-14.134	-1.561	11.169
6	JAK 120	-16.417	-2.441	11.447
7	JAK 124	-18.517	-0.859	11.217
8	JAK 153	-15.128	1.624	11.481
9	JAK 163	-20.099	-1.435	8.501
10	JAK 247	-13.951	-1.437	11.346
11	JAK 248-3	-13.159	-6.932	16.538
12	JAK 287	-18.714	-4.872	12.888
13	JAK 341-2	-17.412	-4.195	14.002
14	JAK 355	-19.417	-1.538	13.006
15	JAK 374	-13.266	-0.919	14.111
16	JAK 377-3	-16.318	-2.566	10.969
17	JAK 390	-21.506	-0.314	12.384
18	JAK 400	-18.239	-0.737	15.568
19	JAK 423	-16.858	-1.287	16.147
20	JAK 424	-20.337	-1.478	17.508
21	JAK 440	-20.570	-0.723	12.920
22	JAK 453	-16.071	-2.443	11.766
23	JAK 486	-15.148	-0.644	9.136

24	JAK 513-1	-18.494	-3.248	13.057
25	JAK 519	-16.893	-0.892	13.101
26	JAK 552	-15.330	1.154	11.439
27	JAK 595-1	-17.756	-1.844	10.873
28	JAK 611	-19.686	4.142	8.752
29	JAK 625	-14.564	-3.273	15.639
30	JAK 638	-14.264	3.718	8.262
31	CR DHAN 310	-17.786	-2.383	10.001

Conclusion

Mahalanobis D^2 analysis resulted in the grouping of 31 rice genotypes into six clusters. Cluster I was observed to be the largest with 19 genotypes, followed by Cluster III with five genotypes. Maximum inter-cluster distance was observed between genotypes of Cluster V (JAK 14, JAK 611, JAK 638 and JAK 552) and with JAK 248-3 of Cluster VI, while intra-cluster distance was noticed to be the maximum for the genotypes in Cluster V. Cluster VI had recorded maximum grain yield plant^{-1}, panicle length, grains panicle^{-1}, test weight and zinc content. The Cluster V had recorded highest protein content. Further, maximum contribution towards divergence was noticed by grain yield plant^{-1}(21.60%) followed by iron content (10.54). Studies on PCA also revealed grain yield plant^{-1} and iron content to contribute maximum towards existing variability. The results on genetic divergence and PCA indicated the need for hybridization of the rice landraces, JAK 611, JAK 638 with JAK 248-3 for the production of heterotic and desirable recombinants with high grain yield and protein content.

Acknowledgements

All authors thank Acharya N.G. Ranga Agricultural University and ICAR-Indian Institute of Rice Research towards providing of facilities for the research.

Authors contributions

BB carried out the studies and drafted the manuscript. YS conceptualized the study and participated in its design and coordination. JAK supervised and provided the resources for research study. TS performed the statistical analysis and provided the required guidance for preparation of the manuscript. All authors reviewed and approved the final manuscript.

Compliance with ethical standards

Conflict of interest Authors do not have any conflict of interests to declare.

Ethical issues: None.

References

1. Prasad, Rama Krishna, Suneetha, Y. and Srinivas. T. Genetic diversity studies in rice (*Oryza sativa* L.). Electronic Journal of Plant Breeding. 2018; 9(4): 1335-1341. https://doi.org/10.5958/0975-928X.2018.00166.7
2. FAOSTAT. Food and Agricultural Organization of United Nation. 2005. Rome, Italy.
3. Gearing M.E. Good as gold: Can golden rice and other biofortified crops prevent malnutrition. Science in the News, Harvard University. https://sitn.hms.harvard.edu/flash/2015/good-as-gold-can-golden-rice-and-other-biofortified-crops-prevent-malnutrition/
4. Sahu S.K., Kumar S.G., Bhat B.V., Premarajan K.C., Sarkar S., Roy G., Joseph N. Malnutrition among under-five children in India and strategies for control. Journal of Natural Science, Biology and Medicine. 2015; Jan; 6(1):18-23. https://doi.org/10.4103/0976-9668.149072
5. Faizan U., Rouster A.S. Nutrition and Hydration Requirements in Children and Adults. 2022 Aug 29. *In*: StatPearls. Treasure Island (FL): StatPearls Publishing; 2022 Jan–. PMID: 32965878.
6. Nitrayová, S. ., Brestenský, M. ., & Patráš, P. Comparison of two methods of protein quality evaluation in rice, rye and barley as food protein sources in human nutrition. Potravinarstvo Slovak Journal of Food Sciences, 2018; 12(1), 762–766. https://doi.org/10.5219/991
7. Hossain S., Maksudul Haque M.D., Rahman J. Genetic Variability, Correlation and Path Coefficient Analysis of Morphological Traits in some Extinct Local Aman Rice (*Oryza sativa* L). J Rice Res 2015 (3):158. https://doi.org/10.4172/2375-4338.1000158
8. .V. Arunachalam, "Genetic Distance in Plant Breeding," Indian Journal of Genetics and Plant Breeding, Vol. 41, No. 2, 1981, pp. 226-236.
9. V. G. Panse and P. V. Sukhatme, Statistical Methods for Agricultural Workers, 2nd Edition, Indian Council of Agricultural Research, New Delhi, 1967.
10. Mahalanobis, P.C. A Statistical Study at Chinese Head Measurement. Journal of Asiatic Society, Bangladesh, 1928; 25: 301-337.
11. Banfield C.F. Multivariate analysis in genstat. Journal of Statistical Computation and Simulation.1978; 6(3-4): 211-22. https://doi.org/10.1080/00949657808810190
12. Wilks S.S. Criterian of generalization in the analysis of variance. Biometrical Journal. 1932; 24 (4): 471-494 . https://doi.org/10.2307/2331979
13. Rao CR. Advanced Statistical Methods in Biometrical Research. John Wiley and Sons Inc., New York. 1952; 236-272. https://www.scirp.org/(S(czeh 2tfqyw2orz553k1w0r45))/reference/ReferencesPapers.aspx?ReferenceID=1001255 . Rao, C.R. Advanced Statistical Methods in Biometric Research, John Wiley and Sons, New York, 1952.
14. Singh RK and Chaudhary BD. Biometrical Methods in Quantitative Genetics Analysis. Kalyani Publishers. New Delhi. 1997; 215-218. https://www.scirp.org/(S(lz5mqp453edsnp55rrgjct55))/reference/ReferencesPapers.aspx?ReferenceID=1435674 Singh, R.K. and Chaudhary, B.D. Biometrical Methods in Quantitative Genetic Analysis. Kalyani Publishers, New Delhi, 1995; 215-218.
15. Jackson JE. A user's guide to principal components. John Wiley and Sons. New York, 1991.https://doi.org/10.1002/0471725331
16. Ehrenberg A.S.C. Data Reduction Analysis and Interpreting Sta-

tistical Data. John Wiley and Sons Inc., New York. 1975; this is a book

17. Sudeepthi K., Srinivas T., Ravikumar BNVSR J.D, Nafeez Umar S.K. Genetic divergence studies for yield and yield component traits in rice (*Oryza sativa* L.). Multilogic in Science. 2020b. 9: 415 -8. https://www.ycjournal.net/Multilogicinscience/ResearchPapers.aspx

18. Singh K.S., Suneetha Y., Kumar G.V., D. Rao V.S., Raja S., Srinivas T. Genetic divergence studies for yield and quality traits in coloured rice. Journal of Pharmacognosy and Phytochemistry. 2020;9(4):1234-40. https://doi.org/10.22271/phyto.2020.v9.i4q.11906

19. Sri Lakshmi M, Suneetha Y and Srinivas T. 2021. Genetic diversity analysis for grain yield and yield components in rice. International Journal of Chemical Studies. 2021; 9(1):1386-1389. https://doi.org/10.22271/chemi.2021.v9.i1t.11416

20. Karuppaiyan R, Kapoor C, Gopi R. Variability, heritability and genetic divergence in lowland rice genotypes under the mid-hills of Sikkim. *Oryza.* 2013:50(1):81-4.