Original Research Article

Characterization of the Genetic Variability in Rice (Oryza sativa L.) and its assessment using Principal Component Analysis (PCA)

Abstract

Present investigation was conducted to estimate genetic parameters such as Genotypic Coefficient of Variation (GCV), Phenotypic Coefficient of Variation (PCV), heritability and Genetic Advance (GA) along with correlations, path coefficients, Genetic divergence and Principal Component Analysis (PCA) from data collected on 21 rice genotypes. The results revealed highly significant mean squares due to genotypes for all traits studied, indicating the existence of sufficient variation among the genotypes and therefore an ample scope for effective selection. High PCV, GCV, heritability and genetic advance as per cent of mean was observed for 1000-grain weight, seed yield per plant, and harvest index, indicating the effectiveness of direct selection for improvement of these traits. Thousand grain weight and harvest index had recorded positive and significant association with seed yield per plant. The results on path analysis also revealed high and positive direct effect for harvest index followed by flag leaf width, days to 50% flowering, days to maturity, plant height, panicle length, number of filled grains/panicle, number of spikelets/panicle and hence, these traits were identified as the most effective selection criteria for improvement of seed yield per plant in rice genotypes. The results on divergence analysis revealed that the genotypes were grouped into five clusters. Cluster I constituted maximum number of (17) genotypes. Maximum differences among the genotypes within the same cluster (intra-cluster) were shown by cluster I (18.33) followed by cluster II, III, IV, V showed zero intra -cluster distance. Cluster III and V (37.41) showed maximum inter cluster distance, suggesting that the genotypes constituted in these clusters may be used as parents for future hybridization programme. Principal Component Analysis (PCA) was utilized to estimate the relative contribution of various traits for total variability. Five components were found to possess eigen value more than 1.00. The PC1, PC2, PC3, PC4 and PC5 contributed 31.2, 14.5, 12.6, 10.5 and 10.1 per cent of variability. Together, they accounted for 78.9% of the variability of the genotypes used in the study has revealed the traits contributing for the variation.

Keywords: Rice, genetic parameters, correlation, path coefficient, seed yield per plant, divergence, principal component

1. Introduction

Rice is the most important cereal crop in the world. Rice *(Oryza sativa L.)* belongs to Gramineae family (Poaceae). The genus includes 24 species out of which 22 are wild and two of them, Oryza sativa and Oryza glaberimma are cultivated. Oryza sativa has a basic chromosome number (n) of 12 (2n=24). The cultivated varieties of *Oryza sativa* grouped in to three sub species, Indica, Japonica, and Javanica. Indica rice varieties are grown throughout the tropical and subtropical regions. Japonica varieties are grown throughout the temperate zone, and Javanica varieties are grown mainly in parts of Indonesia.

Rice (O*ryza sativa* L.) is one of the staple cereal crops of the world and it is one of the main sources of carbohydrate for nearly one half of the world population. It meets the calorie requirement of 50 percent of the population and provides livelihood to 160 million of rural poor [1]. Rice is grown under a variety of climatic and soil conditions. It is well adapted in areas ranging from below sea level (Kuttanad area of Kerala) to 3000 m high places in Nepal. Rice is grown in India under diverse climatic conditions (80N to 340N latitude), accounting for more than 40% of the worlds food grain production and providing direct employment to more than 70% of the rural population. Asia is considered to be 'rice bowl' of the world, and it produces and consumes more than 90% of world rice. In World, China, India, Indonesia, Bangladesh, Vietnam, and Thailand are the major rice producing countries [2]. In India, rice is grown in 43.78 million ha, the production level is 118.43 million tonnes, and the productivity is about 2705 kg/ha during 2019-2020 [3]. The demand for rice in India is estimated to increase to 185 million tonnes by 2030. Being the staple food for more than 65% of the people, national food security hinges on growth and stability of its production [4].

Information on heritability, genetic progress, and the extent of genetic diversity available in the experimental material for grain yield and yield characteristics would be extremely useful to breeders in selecting top genotypes to break the yield plateaus seen in rice. Correlation studies can be used to determine the relative importance of individual qualities. As a result, path coefficient analysis is used to get an understanding of a trait's direct and indirect contribution to yield, allowing the breeder to evaluate genetic traits according to their contribution [5]. The genetic variations seen between genetic stocks with regard to a single characteristic or a set of traits are referred to as genetic diversity. It's caused by the germplasm's intrinsic genetic differences. Knowing the nature of genetic divergence and the degree of genetic divergence can help in selecting parents for breeding programme.

Principal Component Analysis (PCA) is a useful tool for identifying plant traits that distinguish potential genotypes. Due to the regular variation that occurs in crop species, PCA helps in the removal of redundancy in data sets. Its objective is to extract useful information from the table, represent it as a set of new orthogonal variables known as major variables, and visualize similarity patterns between observations and parameters as map points. 'Eigen values' measure the relevance and importance of each factor to the total variance while each individual vector coefficient measures the degree to which each original variable is associated with each main component [6]. The current work aimed to estimate genetic variability, correlation, path analysis, and PCA in a set of 21 rice genotypes.

2. Materials and methods

The experimental material for the present study comprised of 21 genotypes of rice collected from different Krishi Vigyan Kendra's, Andhra Pradesh, India. These genotypes were sown during *Kharif* 2021 at Department of Plant Breeding & Genetics at Agriculture Farm, School of Agriculture, Lovely Professional University, Jalandhar (Punjab), in a randomized block design with three replications. For transplanting, nursery was raised separately, and 27 days old seedlings were transplanted in the main field with a spacing of 20×15 cm. Each plot consists of four rows of 10 m length. Standard agronomic practices and plant protection measures were taken as per schedule. Five competitive plants were selected from each replication for 13 quantitative traits viz: days to 50% flowering, plant height (cm), panicle length (cm), number of tillers per plant, number of spikelets/panicle, flag leaf length (cm), flag leaf width (cm), number of filled grains/panicle, days to maturity, 1000-grain weight (gm), biological yield (gm), seed yield per plant (gm), harvest index (%). The data was subjected to Analysis of variance [7], Heritability [8], Genetic advance [9], Correlation coefficient analysis [10], Path coefficient analysis [5], Genetic diversity using D2 analysis [11], Principal component analysis [12]. The data of 21 Rice genotypes statistically analyzed using R software.

2. Results and discussion

Genetic variability

Analysis of variance (ANOVA) indicated that the mean sum of square (MSS) due to genotypes were highly significant for all the characters at 1% and 5% level of significance and it indicated the presence of consisted amount of genetic variability for all the traits (Table 1). As a result, there is a lot of scope for effective selection. Table 2 shows the results for yield and yield component characteristics in terms of mean, range, phenotypic coefficient of variation (PCV), genotypic coefficient of variation (GCV), heritability, and genetic advance as a percentage of mean. A perusal of these results revealed maximum range of variability for the trait number of spikelets/panicle (115.81-248.67) while minimum range (1.67-2.10) was recorded for flag leaf width. In the present study, the phenotypic coefficient of variation was greater than genotypic coefficient of variation for all the traits. It means the observed variation is not only due to the genotypes, but also due to suitable influence of environment. The phenotypic coefficient of variation and genotypic coefficient of variation were estimates classified into low (below 10%); medium (10-20%) and high (above 20%) in order to draw conclusions about these parameters. In the present study, there is only high PCV and GCV for 1000-grain weight, seed yield per plant and harvest index, indicating the existence of high amount of variability for effective selection towards improvement of these traits. Similar kind of results founded by [13] for 1000-grain weight, and [14, 15] for seed yield per plant. However, the magnitude of PCV and GCV were moderate for plant height, panicle length, number of tillers/plant, number of spikelets/panicle, Flag leaf length, Flag leaf width, Number of filled grains/panicle, indicating little scope for improvement of these traits through selection. Similar kind of results found by [16] for plant height, flag leaf width, panicle length, [17] for number of tillers/plant, number of spikelets/panicle, flag leaf length, [18] for number of filled grains/panicle. Low PCV and low GCV were observed for traits viz., days to 50% flowering, days to maturity and biological yield. Similar kind of results found by [14] for days to 50% flowering and [19] for days to maturity.

Heritability estimates for the various traits studied ranged from 72.38 (biological yield) to 99.04 (panicle length). High estimates of heritability (> 60%) were recorded for all the 13 traits studied. Similar kind of results found by [16, 20]. A perusal of the results on genetic advance as percent of mean revealed high values (>20%) for plant height, panicle length, number of tillers/plant, number of spikelets/panicle, flag leaf length, flag leaf width, number of filled grains/panicle, 1000-grain weight, seed yield per plant and harvest index. It indicates that maximum genetic gain can be determined by using these characters and where these characters are transferred from parent to their offsprings. Similar kind of results found by [16, 20] for plant height, panicle length, number of tillers/plant, flag leaf length, flag leaf width, 1000-grain weight, and seed yield per plant. Moderate genetic advance as percentage of (10-20%) mean was recorded for days to 50% flowering, days to maturity and biological yield. It indicates partially genetic gain by selection. Similar kind of results were found by [21] for days to 50% flowering and [19] for days to maturity.

High heritability coupled with high genetic advance as per cent of mean was recorded for plant height, panicle length, number of tillers/plant, number of spikelets/panicle, flag leaf length, flag leaf width, number of filled grains/panicle, 1000-grain weight, seed yield per plant and harvest index. Indicating that heritability observed was due to additive gene effects and therefore selection would be effective in improvement of these traits. High heritability coupled with moderate genetic advance observed in traits *viz.,* days to 50% flowering, days to maturity and biological yield. while moderate heritability with moderate genetic advance and low heritability with low genetic advance was not observed in any character. In the present study, high GCV and PCV coupled with high heritability and high genetic advance as per cent of mean was observed for 1000-grain weight, seed yield per plant and harvest index, indicating the preponderance of additive gene action and therefore scope for effective improvement of these traits through selection.

Correlation

The correlation analysis among the thirteen characters studied is presented in Table 3. In the present findings, the magnitudes of genotypic correlation coefficient were greater than corresponding phenotypic correlation coefficient in general. In genotypic correlation coefficient seed yield per plant exhibited positive and high significant correlation with the 1000-grain weight. While harvest index had showed positive and significance correlation for seed yield per plant. Other all the characters had shown non significance and positive association with the seed yield per plant. Flag leaf width had shown negative correlation with the seed yield. where significance had not been shown for the negative correlation with seed yield per plant. Similar findings were found with seed yield per plant yield exhibit for 1000-grain weight by [22, 23] and similar findings for harvest index [24, 25, 26]. Seed yield per plant exhibited positive and high significant correlation with the characters namely 1000-grain weight and harvest index. While no observations are shown positive and significance only at 5 % level and days to 50% flowering, plant height, panicle length, number of tillers/plant, number of spikelets/panicle, flag leaf length, number of filled grains/panicle, days to maturity, and biological yield shown positive and no significance to the seed yield per plant. Harvesting index had shown positive significant correlation with seed yield. Similar findings were found with seed yield per plant yield exhibit for 1000-grain weight and harvest index [26]. The results revealed that there is scope for simultaneous improvement of these traits through selection.

Path analysis

The correlation coefficient is partitioned into direct and indirect effects with the help of path coefficient analysis (Table 4). The highest positive direct effect on seed yield per plant was exhibited by harvest index followed by flag leaf width, days to 50% flowering, days to maturity, plant height, panicle length, number of filled grains/panicle and number of spikelets/panicle. Negative direct effect on seed yield per plant was exhibited by flag leaf length followed by 1000-grain weight, number of tillers/plant, and biological yield. The highest positive indirect effect on seed yield per plant was exhibited by harvest index, 1000-grain weight, panicle length, number of tillers/plant, days to maturity, number of filled grains/panicle, days to 50% flowering, flag leaf length, number of spikelets/panicle, plant height and biological yield followed by flag leaf width had shown indirect negative effect on seed yield. Similar findings were found by [27] for days to 50% flowering, plant height, number of filled grains/panicle, [28] for flag leaf length, [22] for1000-grain weight. A positive and high direct effect of a trait on grain yield reveals the effectiveness of a trait for direct selection. High direct effect along with positive and high indirect through other traits provides a better chance for a character to be selected through breeding programmes. Therefore, harvest index, flag leaf width, days to 50% flowering, days to maturity, plant height, panicle length, number of filled grains/panicle and number of spikelets/panicle are the traits which can be directed in selection programmes to obtain higher grain yield.

Genetic divergence

Genetic divergence analysis based on Mahalanobis D^2 statistics, twenty-one genotypes under study were grouped into five clusters (Table 5). Among five clusters, cluster I was the biggest with 17 genotypes followed by cluster II (1 genotype), cluster III (1 genotype), IV (1 genotype) and V (1 genotype). The discrimination of genotypes into discrete clusters suggested presence of high degree of genetic diversity in the material evaluated. Presence of substantial genetic diversity among the parental material screened in the present study indicated that this material may serve as good source for selecting the diverse parents for hybridization programme. In conformity to the present investigation, similar findings were found by [29, 30, 31, 32]. The inter and intra average distances among nine clusters were computed and have been presented in Table 6. Divergence analysis exhibited Maximum differences among the genotypes within the same cluster (intra-cluster) were shown by cluster I (18.33) followed by cluster II, III, IV, V showed zero intra -cluster distance. Cluster III and V (37.41) showed maximum inter cluster distance followed by that between cluster I and V, cluster II and V, cluster IV and V, cluster III and IV, cluster II and III, cluster I and II, cluster I and IV, cluster I and III and cluster II and IV. To realize much variability and high heterotic effect, [33, 34] recommended that parents should be selected from two clusters had wider inter cluster distance.

The cluster mean values showed a wide range of variations for all of the characters undertaken in the study (Table 7). The maximum cluster mean for days to 50% flowering in cluster V, plant height in cluster V, panicle length in cluster V, number of tillers/plant in cluster V, number of spikelets/panicle in cluster III, flag leaf length in cluster V, flag leaf width in cluster IV, number of filled grains/panicle in cluster IV, days to maturity in cluster II, 1000-grain weight in cluster IV , biological yield in cluster II, seed yield per plant in cluster V, harvest index in cluster V. The above observation confirms wide variation from one cluster to another in respect of cluster mean, which indicated that genotypes having distinct different mean performance for various characters were separated into different clusters. The five clusters in the aforesaid divergence analysis contained frequently the genotypes of heterogeneous origin. Although the genotype originated in same place or geographic region were also found to be grouped together in same cluster. The instances of grouping of genotypes of different origin or geographical region also in same cluster were observed in case of all the five clusters. This suggests lack of parallelism between genetic and geographic diversity. Therefore, the selection of parental material for hybridization programme simply based on geographic diversity may not be a successful exercise for the choice of suitable divergent parents. Selection based on genetic divergence analysis would be more rewarding than the choice made based on geographic diversity.

The contribution of each trait to total divergence is presented in table 8. Among the traits studied, flag leaf width contributed maximum divergence (24.76%) followed by harvest index (22.86%), 1000-grain weight (20.48%), flag leaf length (12.38%), and panicle length (10.00%). The minimum percentage of contribution was observed in plant height (0.48%) number of filled grains/panicle (0.48%), days to 50% flowering (0.48%) followed by seed yield per plant (0.95%), days to maturity (1.43%), number of tillers/plant (2.86%), number of spikelets/panicle (2.86%) and biological yield contributed zero percent divergence. The traits *viz.,* flag leaf width, harvest index, 1000-grain weight, flag leaf length, and panicle length contributed 90.48 per cent towards total divergence. Hence, these characters should be given importance during hybridization and selection in the segregating population.

Principle component analysis

PCA identifies traits that contribute to most of the variation within a group of genotypes. In this study, the first five components recorded an eigenvalue greater than 1.00 (Table 9). They together contributed 78.9% of the total variation. Principal component 1 (PC1), with an eigenvalue of 4.05, contributed 31.2% of the total variation, while PC2 with an eigenvalue of 1.88 accounted for 14.5% of the total variation, PC3 had an eigenvalue of 1.63 accounted for 12.6% of the total variation, PC4 had an eigenvalue of 1.36 accounted for 10.5% of the total variation and PC5 had an eigenvalue 1.31 accounted for 10.1% contributed to the observed variability. In the first principal component, traits days to days to 50% flowering, panicle length, 1000-grain weight, seed yield per plant and harvest index contributed greater for total variation due to their high loading. The Plant height, flag leaf length, 1000-grain weight, seed yield per plant and harvest index are the important traits in the second principal component. The days to 50% flowering, plant height, days to maturity and biological yield are the important traits in third principal component. In the fourth principal component, traits panicle length, number of spikelets/panicle, flag leaf width and biological yield contributed for total variation. In the fifth principal component, number of spikelets/panicle, flag leaf width, number of filled grains/panicle and biological yield is the greatest contributor to the observed variation. Similar kind of results were found by [35, 36, 37]. Based on PCA most of the important yield and yield attributing traits were present in PC5 (Table 10) and from first three PCs it was cleared that the days to Days to 50% flowering, Plant height and panicle length is high. Rotated component matrix revealed that first five PCs are representing maximum variability (78.9%) hence, the traits falling in these five PCs must be given due importance in rice breeding programme. The positive and negative loading shows the presence of positive and negative correlation trends between the components and the variables. Therefore, the below mentioned characters which load high positively or negatively contributed more to the diversity and they were the ones that most differentiated the clusters.

4. Conclusion

The findings indicate that the genotypes studied have adequate genetic variability. Characters viz., 1000-grain weight and harvest index had recorded positive and significant association with grain yield per plant. Path analysis revealed high and positive direct effect for harvest index followed by flag leaf width, days to 50% flowering, days to maturity, plant height, panicle length, number of filled grains/panicle, number of spikelets/panicle and hence, these traits were identified as the most effective selection criteria for improvement of seed yield per plant in rice genotypes. Genetic diversity analysis was carried out through D2 analysis in order to assess the genetic divergence among genotypes under study. Based on D2 analysis twenty-one rice genotypes were grouped into five clusters. The results on divergence analysis revealed that Cluster I constituted maximum number of genotypes. Maximum differences among the genotypes within the same cluster (intra-cluster) were shown by cluster I followed by cluster II, III, IV, V that showed zero intra -cluster distance. Cluster III and V showed maximum inter cluster distance, suggesting that the genotypes constituted in these clusters may be used as parents for future hybridization programme. Further, PCA revealed characters viz., days to 50% flowering, panicle length, 1000-grain weight, seed yield per plant, and harvest index contributed for most of the variation in the studied genotypes. Hence the results will be of greater benefit to identify parents for improving traits for future hybridization programme.

Table 1: Analysis of variance for seed yield and yield components in rice

Source of variation	D.F	Mean sum of square												
		DF	**PH**	**PL**	**NTPP**	**NSPP**	**FLL**	**FW**	**NFGP**	**DM**	**TGW**	**BY**	**SYPP**	**HI**
Replication	2	23.53	26.39	0.005	0.66	550.51	0.36	0.00	470.82	10.50	0.04	57.84	24.86	77.33
Treatment	20	175.88**	1046.84**	63.21**	14.97**	2498.18**	70.144**	0.198	1548.87**	214.19**	83.52**	180.70**	126.71**	205.81**
Error	40	6.58	88.34	0.60	0.53	139.01	0.88	0.00	110.67	14.16	1.19	49.91	6.40	19.40

Table 2.: Genetic parameters of variability for yield and yield attributing characters in Rice

Characters	**Range**			**Coefficient of variance**		**H^2b**	**GAM 5%**
	Min	**Max**	**Mean**	**PCV**	**GCV**		
Days to 50% flowering	88.667	115.00	100.33	7.63	7.48	96.25	15.18
Plant height (cm)	77.1	147.4	110.31	16.11	15.42	91.56	35.23
Panicle length (cm)	20.53	43.66	24.95	18.34	18.25	99.04	9.36
Number of tillers/plant	8.87	18.3	13.53	16.22	15.92	96.40	4.43
Number of spikelets/panicle	115.81	248.67	195.21	14.71	14.30	94.44	56.13
Flag leaf length (cm)	22.93	43.00	30.63	15.75	15.65	98.74	9.83
Flag leaf width(cm)	1.67	2.10	1.67	15.37	15.29	99.01	0.52
Number of filled grains/panicle	75.19	231.93	168.51	13.47	12.98	92.85	43.46
Days to maturity	110.82	145.2	130.03	6.49	6.27	93.39	16.25
1000-grain weight	12.28	29.51	20.37	25.61	25.43	98.57	10.71
Biological yield	58.87	102.17	86.66	9.10	7.74	72.38	11.57
Seed yield per plant	17.9	36.3	27.07	23.86	23.25	94.95	12.71
Harvest index	20.78	54.38	31.70	25.60	24.37	90.57	15.45

Table.3 Genotypic and Phenotypic correlation among 13 different characters in 21 rice genotypes

S.N.	Characters	DFF	PH	PL	NTPP	NSPP	FLL	FLW	NFGPP	DM	TGW	BY	HI	SYPP
1	Days to 50% flowering	1.0000 1.0000	0.1245 0.1349	0.4882* 0.4768*	0.1457 0.1400	0.0517 0.0462	0.5950** 0.5788*	-0.2321 -0.2286	0.0785 0.0832	0.7497** 0.7150**	0.4334* 0.4122	-0.0399 -0.0153	0.1914 0.1948	0.1846 0.1988
2	Plant height (cm)		1.0000 1.0000	0.3619 0.3473	-0.0508 -0.0539	0.0006 -0.0114	0.6588** 0.6212*	-0.1075 -0.1046	0.3814 0.3435	-0.1545 -0.1321	0.0303 0.0149	-0.2787 -0.2181	0.1780 0.1748	0.0752 0.0808
3	Panicle length (cm)			1.0000 1.0000	0.4546* 0.4484*	-0.0148 -0.0127	0.5942** 0.5874**	-0.0250 -0.0249	0.2658 0.2570	0.3818 0.3670	0.4036 0.3989	-0.1056 -0.0962	0.4450* 0.4259	0.3966 0.3868

4	Number of tillers/plant				1.0000 1.0000	-0.0399 -0.0479	0.0568 0.0538	0.0817 0.0805	0.0755 0.0782	0.1731 0.1596	0.2991 0.2928	-0.2873 -0.2461	0.4455* 0.4188	0.3515 0.3355
5	Number of spikelets/panicle					1.0000 1.0000	0.1142 0.1131	-0.1950 -0.1869	0.3177 0.2906	0.2550 0.2433	-0.0749 -0.0735	0.1342 0.0976	0.0485 0.0541	0.1141 0.1160
6	Flag leaf length(cm)						1.0000 1.0000	-0.0039 -0.0038	0.3550 0.3441	0.3158 0.3087	0.1376 0.1384	-0.2432 -0.2073	0.2221 0.2150	0.1343 0.1353
7	Flag leaf width (cm)							1.0000 1.0000	-0.1228 -0.1238	-0.3473 -0.3350	-0.0661 -0.0641	0.4417 0.3770	-0.1847 -0.1747	-0.0365 -0.0357
8	Number of filled grains/panicle								1.0000 1.0000	0.1621 0.1441	0.1959 0.1822	0.0191 -0.0101	0.3326 0.3186	0.3420 0.3240
9	Days to maturity									1.0000 1.0000	0.3391 0.3258	0.2022 0.1764	0.2628 0.2399	0.3514 0.3321
10	1000 grain weight										1.0000 1.0000	-0.1212 -0.0940	0.8084** 0.7468	0.7844** 0.7468**
11	Biological Yield											1.0000 1.0000	-0.2509 -0.3122	0.0309 0.0379
12	Harvesting index												1.0000 1.0000	0.9582* 0.9319**

__*, ** indicate significant at 5% and 1% probability level, respectively.__

DFF= Days to 50 % flowering, PH = Plant height (cm), PL = Panicle length (cm), NTPP = Numbers of tillers per plant, NSPP = Numbers of spikelets per panicle, FLL = Flag leaf length (cm), FLW = Flag leaf width (cm), NFGPP = Numbers of filled grains per panicle, DM =Days to maturity, TGW = 1000 grains weight (g), BY = Biological yield per plant (g), HI = Harvest Index (%), SYPP =Seed yield per plant (g)

Table 4 Genotypic Direct and indirect effect of different characters on seed yield per plant in Rice

S.N.	Characters	DFF	PH	PL	NTPP	NSPP	FLL	FLW	NFGPP	DM	TGW	BY	HI	SYPP
1	Days to 50% flowering	**0.8142**	0.1014	0.3975	0.1186	0.0421	0.4845	-0.1890	0.0639	0.6104	0.3529	-0.0325	0.1558	0.1846
2	Plant height (cm)	0.0487	**0.3913**	0.1416	-0.0199	0.0002	0.2578	-0.0421	0.1492	-0.0605	0.0119	-0.1091	0.0696	0.0752
3	Panicle length (cm)	0.1286	0.0953	**0.2635**	0.1198	-0.0039	0.1566	-0.0066	0.0700	0.1006	0.1063	-0.0278	0.1173	0.3966
4	Number of tillers/plant	-0.0904	0.0315	-0.2820	**-0.6203**	0.0248	-0.0352	-0.0507	-0.0468	-0.1074	-0.1856	0.1782	-0.2763	0.3515
5	Number of spikelets/panicle	0.0061	0.0001	-0.0017	-0.0047	**0.1173**	0.0134	-0.0229	0.0373	0.0299	-0.0088	0.0157	0.0057	0.1141
6	Flag leaf length(cm)	-0.7893	-0.8740	-0.7882	-0.0753	-0.1515	**-1.3265**	0.0051	-0.4709	-0.4189	-0.1825	0.3227	-0.2946	0.1343
7	Flag leaf width (cm)	-0.2151	-0.0996	-0.0231	0.0757	-0.1807	-0.0036	**0.9266**	-0.1138	-0.3218	-0.0612	0.4093	-0.1711	-0.0365
8	Number of filled grains/panicle	0.0129	0.0627	0.0437	0.0124	0.0522	0.0583	-0.0202	**0.1643**	0.0266	0.0322	0.0031	0.0547	0.3420
9	Days to maturity	0.3247	-0.0669	0.1654	0.0750	0.1104	0.1368	-0.1504	0.0702	**0.4331**	0.1469	0.0876	0.1138	0.3514
10	1000 grain weight	-0.4331	-0.0303	-0.4033	-0.2989	0.0749	-0.1375	0.0660	-0.1957	-0.3388	**-0.9992**	0.1211	-0.8078	0.7844
11	Biological Yield	0.0187	0.1302	0.0493	0.1342	-0.0627	0.1136	-0.2063	-0.0089	-0.0944	0.0566	**-0.4671**	0.1172	0.0309
12	Harvesting index	0.3586	0.3335	0.8339	0.8348	0.0910	0.4162	-0.3461	0.6232	0.4924	1.5149	-0.4703	**1.8740**	0.9582

R SQUARE = 0.4202 RESIDUAL EFFECT = 0.7614, Bold figures indicate direct effect

Table 5: Clustering pattern of 21 rice genotypes & based on Mahalanobis's D^2 statistic

Cluster No.	No. of Genotypes	Genotypes
Cluster 1	17	NLR-3049, UMA, EXP-092, WGL-44, PR-21, RNR-15048, MTU-1153, MTU-1224, JGL-24423, KNM-118, HMT-999, RNR-1446, RGL-2537, KNM-1638, MTU-1064, MTU-1010, ADT-39
Cluster 2	1	PLA-1100
Cluster 3	1	BPT-5204
Cluster 4	1	SAMPAT SWARNA
Cluster 5	1	SWARNA KRANTI

Table 6: Estimates of average intra and inter-cluster distances for the five clusters in rice.

Clusters	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Cluster 1	**18.33**	23.28	20.85	21.04	34.11
Cluster 2		**0.00**	25.92	20.09	33.04
Cluster 3			**0.00**	29.47	37.41
Cluster 4				**0.00**	30.60
Cluster 5					**0.00**

Table 7: Cluster means for different characters in Rice

Characters	Days to 50% flowering	Plant height (cm)	Panicle length (cm)	Number of tillers/plant	Number of spikelets/panicle	Flag leaf length (cm)	Flag leaf width (cm)	Number of filled grains/panicle	Days to maturity	1000 Grain weight (gm)	Biological Yield (gm)	Seed yield per plant	Harvesting index
Cluster 1	98.29	116.36	24.15	13.32	194.36	29.90	1.71	165.77	127.27	19.60	86.16	26.42	31.11
Cluster 2	110.33	113.39	26.00	15.07	202.34	26.77	1.32	188.51	145.20	28.57	89.33	32.92	37.07
Cluster 3	104.33	88.50	20.53	16.20	216.27	29.64	1.30	156.67	136.00	14.39	72.59	18.40	25.35
Cluster 4	106.67	119.60	24.60	13.13	202.35	37.00	1.96	192.00	140.99	29.52	84.67	35.22	41.91
Cluster 5	115.00	134.40	43.67	18.30	192.18	43.00	1.54	185.03	145.00	26.97	79.54	36.30	45.93

Table 8: Contribution of 13 traits of Rice towards divergence

Sr. No.	**Source**	**Times Ranked 1st**	**Contribution %**
1	Days to 50% Flowering	1	0.48 %
2	Plant height (cm)	1	0.48 %
3	Panicle length (cm)	21	10.00 %
4	Number of tillers/plant	6	2.86 %
5	Number of spikelets/panicle	6	2.86 %
6	Flag leaf length (cm)	26	12.38 %
7	Flag leaf width (cm)	52	24.76 %
8	Number of filled grains/panicle	1	0.48 %
9	Days to maturity	3	1.43 %
10	1000 Grain weight (gm)	43	20.48 %
11	Biological Yield (gm)	0	0.00 %
12	Seed yield per plant	2	0.95 %
13	Harvesting index	48	22.86 %

Table 9: Eigen values, % variance and cumulative variance of Rice genotypes

Traits	**Principal component**	**Eigen value**	**Variation (%)**	**Cumulative variance (%)**
Days to 50% flowering	PC1	**4.0589**	0.312	0.312
Plant height	PC2	**1.8826**	0.145	0.457
Panicle length	PC3	**1.6360**	0.126	0.583
Number of tillers/plant	PC4	**1.3601**	0.105	0.688
Number of spikelets/panicle	PC5	**1.3120**	0.101	0.788
Flag leaf length	PC6	0.9041	0.070	0.858
Flag leaf width	PC7	0.5715	0.044	0.902
Number of filled grain/ panicle	PC8	0.4587	0.035	0.937
Days to maturity	PC9	0.3127	0.024	0.961
1000-grain weight	PC10	0.2811	0.022	0.961
Biological yield	PC11	0.1281	0.010	0.993
Seed yield per plant	PC12	0.0903	0.007	1.000
Harvest index	PC13	0.0037	0.000	1.000

Table 10: Principal Component values of rotation component matrix

Traits	**Principal components**				
	PC1	**PC2**	**PC3**	**PC4**	**PC5**
Days to 50% flowering	**0.313**	0.244	**0.358**	-0.209	-0.270
Plant height	0.170	**0.437**	**-0.413**	-0.024	0.169
Panicle length	**0.366**	0.131	-0.067	**-0.313**	-0.028
Number of tillers/plant	0.227	-0.259	-0.151	-0.178	-0.183
Number of spikelets/panicle	0.071	0.146	0.279	**0.454**	**0.359**
Flag leaf length	0.294	**0.487**	-0.097	-0.214	0.054
Flag leaf width	-0.120	-0.162	-0.112	**-0.612**	**0.404**
Number of filled grain/ panicle	0.226	0.169	-0.103	0.262	**0.481**
Days to maturity	0.286	0.081	**0.563**	0.005	-0.143
1000-grain weight	**0.364**	**-0.329**	0.018	-0.039	-0.026
Biological yield	-0.108	-0.078	**0.458**	**-0.304**	**0.505**
Seed yield per plant	**0.378**	**-0.365**	-0.015	0.086	0.242
Harvest index	**0.402**	**-0.312**	-0.191	0.176	0.049

References

1. Santha, V. and Karthikeyan. (2016). Genetic variability Studies for Yield and Yield Components in Rice *(Oryza Sativa L.). Research Article Vegetos, 29*(4): 63-68*.*
2. Coats, B. (2003). Global rice production. *Rice origin, history, technology and production*, 247-470.
3. Agriculture statistics.2020. Pocket book on Agricultural statistics, ministry of Agricultural Government of India.48.
4. Shamim, Mohammad & SHEKH, A. & Pandey, Vyas & Patel, Haridas & Lunagaria, Manoj. (2012). Simulating the phenology, growth and yield of aromatic rice cultivars using CERES Rice model under different environments. *Journal of agrometeorology*. 14. 31-34. 10.54386/jam.v14i1.1376.
5. Dewey, D. R., and Lu. K. H. (1959). Correlation and path coefficient analysis of crested wheat grass seed production, *Agrion. J*., 51: 515-518.
6. Nachimuthu VV, Robin S, Sudhakar D, Raveendran M, Rajeswari S, Manonmani S. Evaluation of rice genetic diversity and variability in a population pannel by principal component analysis. *Indian J Sci. Technol 2014*;*7*(10):1555-1562.
7. Panse VG, Sukhatme PV. *Statistical Methods for Agricultural Workers*. ICAR, New Delhi 1978, 103-108.
8. Hanson, C.H., H.F. Robinson and R.E. Comstock. 1956. Biometrical studies on yield in segregating population of Korean lespedesa. *Agron. J.* 48:268-272.
9. Johnson HW, HF Robinson and RE Comstock. 1955. Estimates of genetic and environmental variability in soybeans. *Agronomy Journal 47*: 314-318.
10. Miller PJ, Willianis JC, Robinson HF and Comostock RE 1958. Estimates of genotypic and environmental variance and covariance in upland cotton and their implication in selection. *Agron. J. 50*: 126-131.
11. Mahalanobis, P. C. (1936). On the generalized distance in statistics", proceeding of the national institute of sciences of India. 2 (1), 49-55.
12. Ingebritsen, Steven & LYON, R.. (1985). Principal component analysis of multitemporal image pairs. International Journal of Remote Sensing - *INT J REMOTE SENS. 6.* 687-696. 10.1080/01431168508948 49.
13. Parihar, R., Sharma, D. J., Agrawal, A. P., & Minz, M. G. (2017). Genetic variability, heritability and genetic advance studies in aromatic short grain rice (Oryza sativa L.) genotypes. *Phytojournal. com.*
14. Singh, B., Gauraha, D., Sao, A., & Nair, S. K. (2021). Assessment of genetic variability, heritability and genetic advance for yield and quality traits in advanced breeding lines of rice (Oryza sativa L.). *The Pharma Innovation*, *10*(8), 1627-1630.
15. Lipi, L. F., Hasan, M. J., Akter, A., Quddus, M. R., Biswas, P. L., Ansari, A., & Akter, S. (2020). Genetic variation, heritability and genetic advance in some promising rice hybrids. *SAARC Journal of Agriculture*, *18*(2), 39-49.
16. Keerthiraj, B., & Biju, S. (2020). Genetic variability, heritability and genetic advance of yield and lodging-related traits in rice (Oryza sativa L.). *Electronic Journal of Plant Breeding*, *11*(04), 1093-1098.
17. Sumanth, V., Suresh, B. G., Ram, B. J., & Srujana, G. (2017). Estimation of genetic variability, heritability and genetic advance for grain yield components in rice (Oryza sativa L.). *Journal of Pharmacognosy and Phytochemistry*, *6*(4), 1437-1439.
18. Islam, M. A., Raffi, S. A., Hossain, M. A., & Hasan, A. K. (2015). Analysis of genetic variability, heritability and genetic advance for yield and yield associated traits in some promising advanced lines of rice. *Progressive Agriculture*, *26*(1), 26-31.
19. Dhurai, S. Y., Bhati, P. K. and Saroj, S. K., 2014. Studies on genetic variability for yield and quality characters in rice (Oryza sativa L.) under integrated fertilizer management. *The Bioscan. 9*(2): 745-748.
20. Sandeep, S., Sujatha, M., Subbarao, L. V., & Neeraja, C. N. (2018). Genetic variability, heritability and genetic advance studies in rice (Oryza sativa L.). *International Journal of Current Microbiology and Applied Science*, *7*(12), 3719-3727.

21. Ajmera, S., Kumar, S. S., & Ravindrababu, V. (2017). Evaluation of genetic variability, heritability and genetic advance for yield and yield components in rice genotypes. *International Journal of Current Microbiology and Applied Sciences*, *6*(10), 1657-1664.
22. Shrestha, N., Poudel, A., Sharma, S., Parajuli, A., Budhathoki, S., & Shrestha, K. (2018). Correlation Coefficient and Path Analysis of Advance Rice Genotypes in Central Mid-hills of Nepal. *International Journal of Research in Agricultural Sciences*, *5*(3), 2348-3997.
23. Rashid, K. A. S. H. I. F., Kahliq, I., Farooq, M. O., & Ahsan, M. Z. (2014). Correlation and cluster analysis of some yield and yield related traits in Rice (Oryza sativa). *Journal of Recent Advances in Agriculture*, *2*(6), 271-276.
24. Gour, L., Maurya, S. B., Koutu, G. K., Singh, S. K., Shukla, S. S., & Mishra, D. K. (2017). Characterization of rice (Oryza sativa L.) genotypes using principal component analysis including scree plot & rotated component matrix. *International Journal of Chemical Studies*, *5*(4), 975-83.
25. Moosavi, M., Ranjbar, G., Zarrini, H. N., & Gilani, A. (2015, January). Correlation between morphological and physiological traits and path analysis of grain yield in rice genotypes under Khuzestan conditions. *In Biological Forum* (Vol. 7, No. 1, p. 43). Research Trend.
26. Idris, A. E., Justin, F. J., Dagash, Y. M. I., & Abuali, A. I. (2012). Genetic variability and inter relationship between yield and yield components in some rice genotypes. *American Journal of Experimental Agriculture*, *2*(2), 233.
27. Ekka, R. E., Sarawgi, A. K., & Kanwar, R. R. (2011). Correlation and path analysis in traditional rice accessions of Chhattisgarh. *Journal of rice research*, *4*(1), 11-18.
28. Reddy, G. E., Suresh, B. G., Sravan, T., & Reddy, P. A. (2013). Interrelationship and cause-effect analysis of rice genotypes in north east plain zone. *The Bioscan*, *8*(4), 1141-1144.
29. Koli, G. K., Upadhyay, S., Rathi, S. R., Gupta, S., Kumawat, G., Singh, P. K., & Sinha, B. (2022). Study of genetic divergence for yield and yield attributing traits of wild germplasm in rice. *The Pharma Innovation Journal*;*11*(2): 472-474.
30. Ayenew, A., Dejene, T., & Worede, F. (2020). Genetic divergence analyses of lowland rice genotypes in Northwestern Ethiopia. *African Journal of Plant Science*, *14*(4), 165-171.
31. Ranjith, P., Sahu, S., Dash, S. K., Bastia, D. N., & Pradhan, B. D. (2018). Genetic diversity studies in Rice (Oryza sativa L.). *Journal of Pharmacognosy and Phytochemistry*, *7*(2), 2529-2531.).
32. Umesh, J., H.K., Sravan, T. and Waza, S.A. (2016). Genetic divergence analysis for yield and quality traits in some indigenous genotypes of basmati rice (Oryza sativa L). *Green Farming, 7* (2): 345-348.
33. Mishra LK, Sarawgi AK, Mishra RK. Genetic diversity for morphological and quality traits in rice (Oryza sativa L.). *Adv. in Plant Sci*. 2003; 16(1):287-293.
34. Chaturvedi HP, Maurya DM. Genetic divergence analysis in rice (Oryza sativa L.). *Advances Pl. Sci*. 2005; 8(1):349-353.
35. Kashyap, A. and Vijay Kumar Yadav. 2020. Principal Component Analysis and Character Association for Yield Components in Rice (Oryza sativa L.) Genotypes of Salt Tolerance under Alkaline Condition.*Int.J.Curr.Microbiol.App.Sci. 9*(10): 481-495.
36. Tejaswini, K. L. Y., Manukonda, S., Kumar, B. R., Rao, P. R., & Raju, S. K. (2018). Application of principal component analysis for rice f5 families characterization and evaluation. *Emergent Life Sciences Research*, *4*, 72-84.
37. Wijayawardhana, H. C. D., Herath, H. M. V. G., Weerasinghe, P. A., & Herath, H. M. D. A. K. (2015). Morphological variation in selected Sri Lankan rice (Oryza sativa L.) accessions in relation to the vegetative parameters. *Tropical Agricultural Research* Vol. 26 (2): 380 – 389.