Research Article

Genetic Diversity Estimation of Rice Genotypes based on Morphological and Quality Parameters through Principal Component Analysis

Salma Sharif[1], Rana Arsalan Javaid[2]*, Abid Majeed[2], Muhammad Shahzad Ahmed[2], Qurat ul Ain Sani[2], Faiza Siddique[2], Muhammad Arshad[3] and Niaz Ali[1]

[1]Department of Botany, Hazara University Mansehra, Khyber Pakhtunkhwa, Pakistan; [2]Rice Research Program, Crop Sciences Institute, National Agricultural Research Centre, Islamabad, Pakistan; [3]Crop Sciences Institute, National Agricultural Research Centre, Islamabad, Pakistan.

Abstract | For more than half of the world's population, rice (*Oryza Sativa* L., family=Poaceae, 2n=24) is the second most important food crop after wheat. It is produced in several states all around the world, while primarily being grown in river deltas of Asia and the Southeast Asia. Therefore, it is of major significance to food security in a growing amount of less privileged countries. A study was carried out to assess indigenous rice germplasm using qualitative and quantitative features, as well as to select attractive genotypes for use in breeding programs in the future. A set of 65 accessions were evaluated in the field of Crop Sciences Institute, NARC following factorial design with three replications. highly significant differences were observed for all the quantitative traits as days to heading, days to maturity, plant height, number of tillers, panicle length, flag leaf area, leaf length, culm length, culm diameter, number of grains per panicle, grain length, grain diameter, chlorophyll content, net differential vegetation index, thousand grain weight and grain yield. The overall mean for Grain yield ranged from 1.71 tons/ha to 6.18 tons/ha. Grand mean of all genotypes for Grain yield was 4.1 tons/ha. Maximum Grain yield was observed in GSR 11 followed by GSR 10 and GSR 44 that were followed by 5.88 and 5.74 tons/ha respectively. Significant variation was also found in case of qualitative characters as obvious groups were formed based on visual observations. In PCA, the first two components were found to contribute 33.902% of the total variability so, the biplot was created using the first two components. The results of Principal Component Analysis (PCA) matched those of the cluster analysis quite closely. Breeders may now use these findings to create high-yielding rice varieties as well as novel breeding techniques for rice development.

Received | August 05, 2023; **Accepted** | September 20, 2023; **Published** | September 27, 2023
***Correspondence** | Rana Arsalan Javaid, Rice Research Program Crop Sciences Institute, National Agricultural Research Centre, Islamabad, Pakistan. ranaarsalan48@gmail.com
Citation | Sharif, S., R.A. Javaid, A. Majeed, M.S. Ahmed, Q.A. Sani, F. Siddique, M. Arshad and N. Ali. 2023. Genetic diversity estimation of rice genotypes based on morphological and quality parameters through principal component analysis. *Pakistan Journal of Agricultural Research,* 36(3): 207-216.
DOI | https://dx.doi.org/10.17582/journal.pjar/2023/36.3.207.216
Keywords | Analysis of variance, Cluster analysis, Grain yield, Variability, Biplot

Introduction

Rice is of major importance to food security in a growing number of less privileged countries (Singh *et al*., 2018). Rice is a major crop in many nations, with cultivation ranging from the humid tropics to northeast China and southeast Australia, and from sea level to more than 2500 meters in Nepal and Bhutan's temperate areas. Rice is a nutrient-dense cereal crop, delivering 20% of the calories and 15% of the proteins consumed by the global population. Although it is Asia's primary source of carbohydrates and protein, it also contains minerals and fiber (Roychoudhury, 2020).

Plant genetic diversity is influenced by a variety of factors, including geographical, ecological, anthropogenic effects and breeding system (Saeed and Barozai, 2021). Germplasm collections that are diverse, are useful resources for crop development (Nice *et al*., 2016). In the case of rice, the genetic heterogeneity available in *O. sativa* germplasm has been primarily exploited in such breeding programs. Modern, high-yielding varieties have largely supplanted older kinds, particularly in the irrigated rice habitat, resulting in a smaller genetic base and hence increased genetic susceptibility (Singh *et al*., 2018).

It is critical to diversify the genetic base of this vital crop by introducing genes from various sources. As a result, untapped germplasm must be collected, exploited, and evaluated. In this respect, breeders attempt to characterize the germplasm accessions for various morphological and agronomic experiments, as well as to identify the germplasm diversity (Shrestha, 2016). Characterization of germplasm is critical for providing information on accessions attributes and ensuring that the germplasm collection is used to its full potential by end users. The collection and recording of data on the important traits that separate accessions within a species allows for easy and rapid phenotyping (Gangapur *et al*., 2014). Variations in characteristics within species, on the other hand, are vital for the development of stress-resistant cultivars and varieties. For long-term production, it is necessary to improve our understanding of adaptation in harsh conditions and the best use of the adaption features found in all breeds to pyramid them in a new single genotype through breeding interventions (Lutatenekwa *et al*., 2020).

The study has been carried out to analyze the diversity of rice germplasm and characterization of morphological features that will help to identify the potential genotypes to be employed in future breeding programs. Information regarding population genetic structure of rice cultivars may also be revealed in order to devise any crop improvement program with enhanced efficiency.

Materials and Methods

To determine the genetic diversity in advance rice lines for agronomic and quality traits, diverse rice germplasm was characterized in the field. A set of 65 accessions (GSR lines) were evaluated. The germplasm was collected and evaluated at Rice program Crop Sciences Institute (CSI), National Agricultural Research Centre Islamabad (33.6844°N, 73.0479°E). Seeds of all the rice genotypes were sown in the soil to develop nursery in research area of Crop Sciences Institute, NARC and then these seedlings were transplanted to the field. Nursery was sown on 15 June and transplanted to field on 16th July 2021. Factorial design was adopted with three replications for the execution of the experiment. There were three rows of each genotype Distance between the adjacent rows was 20 cm while between two plants in the same row it was also maintained 20cm. Data for the major quantitative and qualitative parameters (Table 1) was recorded and evaluated.

Table 1: *List of quantitative and qualitative characters observed for the study.*

Quantitative traits		Qualitative traits
Days to heading	Culm length	Leaf color
Days to maturity	Culm diameter	Flag leaf attitude
Plant Height	Grain length	Panicle attitude
No. of tillers	Grain diameter	Panicle Type
Panicle length	Chlorophyll content	Leaf spot incidence
Flag leaf area	Total grain weight	Awning and Awn color
Leaf length	Grain yield	Caryopsis shape
No of grains/panicle		Chalkiness

Statistical analysis

Plant characterization aids in the availability of information on desirable plant features, allowing for more selective breeding for environmental challenges and, as a result, achieving long-term efficient crop production. Analysis of Variance (ANOVA) and Least Significant Difference (LSD) following Steel

and Torrie 1980 was assessed for all genotypes for agronomical and morphological parameters using statistix 8.1 software to see if there are any statistically significant variations between the means of different genotypes under the same conditions. K-means clustering was adopted that divides a large number of n observations into k clusters, with genotypes in each cluster having statistically similar means, as indicated by the lloyd-forgy model (Sreevalsan, 2021). Principal component analysis was performed using XLSTAT software to figure out the major contributing factors towards total variability and distribution of diverse germplasm against the specific character. Principal component analysis, will reduce the complexity of a data collection by summarizing the data on distinct trends and patterns (Kherif and Latypova, 2020). It may lead towards objective based selection of the genotypes effectively.

Results and Discussion

Quantitative data analysis

ANOVA results revealed highly significant differences for all the quantitative traits as days to heading, days to maturity, plant height, number of tillers, panicle length, flag leaf area, leaf length, culm length, culm diameter, number of grains per panicle, grain length, grain diameter, chlorophyll content, net differential vegetation index, thousand grain weight and grain yield (Table 2).

Days to heading is a significant genetic locus in rice that controls photoperiod sensitivity and grain yield (Gao *et al*., 2014). The overall mean for DH ranged from 90-120 days in all genotypes with a grand mean of 101.9 days. Maximum days to heading were observed in GSR 16 followed by GSR 17 and GSR 56 that were 117 days, 114 days and 114 days on average in all the three replications. Lesser number of days to heading among the genotypes were observed in GSR 7, GSR 8, GSR 11 and GSR 57 taking only 92 and 91 days, respectively. Days to maturity is considered as one of the important traits for rice cultivation (Takai *et al*., 2019). Rice genotypes were found to be mature in a time interval of 124 to 162 days. Early maturing genotypes are preferred by the farmer to have two crops in a season. GSR 56 may be referred to as early maturing as it took only 124.33 days to get mature. While the other group falling in late maturing as GSR 31, GSR 41 and GSR 61 completed its cycle in 161.67 followed by 158.67 and 158.67 days, respectively (Table 3).

Table 2: *Analysis of variance for the quantitative characters of diverse rice germplasm.*

Source of variation	df	Days to heading	Days to maturity	Plant height	Panicle length	No. of tillers	Flag leaf area	Leaf length	Culm length	Culm dia	Grains/ panicle	Grain length	Grain Dia	Chlorophyl content	NDVI	TGW	Yield
Replication	2	1.745	513.33	8.867	0.246	2.436	36.26	0.0004	10.667	0.032	1296.1	1.056	0.001	4.167	0.761	78.13	39.02
Genotypes	64	119.38	101.78	159.89	11.185	14.051	141.2	0.195	45.332	0.486	1327.1	1.073	0.376	23.148	0.004	26.03	263.4
Error	128	2.388	28.453	18.356	4.184	6.847	28.74	0.035	5.651	0.16	679.17	0.253	0.051	8.687	2.593	1.985	48.52
Grand mean		101.91	151.18	103.08	27.03	11.195	32.78	0.71	70.57	3.32	153.05	7.87	2.299	35.907	0.64	25.84	41.96

Table 3: *Mean performance of rice genotypes for the morphological characters under observation.*

Gen	DH	DM	PH	NT	PL	FLA	CL	CD	G/P	GL	GD	CC	NDVI	TGW	GY
GSR-1	94	148	92.67	9.67	23.33	35.55	72.33	3.13	167.95	8.88	1.9	39.2	0.6	19.7	4.85
GSR-2	96	147	96.33	11.33	27.33	32.45	68.78	3.54	152.25	8.29	2.16	39.23	0.59	24	5.12
GSR-3	96	149	94.33	7.67	26.67	38.63	69.67	3.38	156.36	7.04	2.17	39.03	0.6	26.2	5.19
GSR-4	97	149	95.67	11	26.33	43.03	70	3.02	141.93	7.56	2.57	36.53	0.61	24.4	3.53
GSR-5	96	146	104.67	9.33	25.67	32.65	82.22	3.05	145.45	7.99	2.38	37.1	0.63	23	3.79
GSR-6	96	148	104	12	26.67	39.88	72.33	3.34	153.59	8.27	2.62	35.03	0.66	23.6	4.81
GSR-7	92	145	106.67	11.67	23	42.2	78.67	3.01	173.26	6.7	2.66	34.7	0.62	21.7	4.97
GSR-8	92	144	96.67	12.33	25.33	32.9	70	2.79	158.22	7.08	2.68	34.33	0.63	21.9	4.21
GSR-9	98	150	110.33	9.67	24.67	49.63	77.11	3.1	167.3	8	2.73	35.83	0.6	25.9	4.93
GSR-10	98	150	104.33	11.33	28.33	40.53	76.56	3.36	139.83	7.43	2.48	34.93	0.64	24.6	5.88
GSR-11	91	144	98.67	12	31	34.2	69	2.83	165.13	7.27	2.62	36.63	0.64	23.6	6.18
GSR-12	105	158	91	9	31.33	38.43	69.44	3	165.33	7.14	2.37	39.37	0.55	23.6	4.36
GSR-13	98	147	108	11.67	29	35.48	71.56	2.27	173.63	7.16	2.17	30.07	0.63	21	3.98
GSR-14	92	144	96.33	11	27.33	51.83	70.78	2.8	170.43	7.76	2.11	34.87	0.59	18.3	4.89

Table continued on next page.....

Gen	DH	DM	PH	NT	PL	FLA	CL	CD	G/P	GL	GD	CC	NDVI	TGW	GY
GSR-15	98	151	107.67	11.33	29	42.53	72.33	2.9	155.81	8.21	2.25	33.7	0.62	33.2	4.06
GSR-16	120	147	115.33	12	27.33	39.65	73.44	3.72	163.97	7.48	2.49	31.43	0.64	29.1	3.9
GSR-17	120	147	97.67	11	22.33	32.55	68.67	3.02	175.44	8.16	2.84	37.1	0.62	20.4	4.21
GSR-18	97	149	128.67	9	32.33	22.43	93.11	2.95	138.41	7.46	2.92	38.7	0.57	28.4	3.75
GSR-19	96	147	119	5.33	27.33	94.38	79.11	4.81	267.97	7.68	2.83	35.37	0.61	29.9	5.38
GSR-20	96	145	102	10.67	27.67	74.2	64.22	3.62	186.03	7.82	2.99	32.17	0.63	27.1	4.81
GSR-21	97	150	107	12	26	36.85	73.11	2.64	157.61	8.92	1.89	32.77	0.57	24	4.63
GSR-22	103	149	106	11.67	30.33	35.25	71.11	3.04	176.83	9.2	2.08	36.67	0.59	25.9	4.17
GSR-23	101	153	107	12	27.33	34.2	72.22	2.93	142.01	8.89	1.8	33.77	0.61	21	3.53
GSR-24	105	158	106.67	13	28.33	43.03	71.56	3.4	144.35	9.7	1.87	33.4	0.59	29.3	3.64
GSR-25	104	160	107	11.67	26	43.05	71.44	3.8	166.34	9.7	2.05	35.07	0.64	27.5	4.74
GSR-26	101	159	96.33	12.67	30	32.55	73.22	3.61	146.46	8.69	2.59	31.27	0.57	22.6	3.26
GSR-27	100	158	104.67	15	30	28.25	70.89	3.31	126.34	9.86	2.41	30.5	0.58	28.1	2.86
GSR-28	100	159	98	8.33	28.33	27.65	67.89	3.59	149.62	10.37	2.32	33.6	0.6	25.6	3.07
GSR-29	102	159	104.67	13	29.67	29.83	71.22	3.11	108.21	8.83	1.82	32.77	0.6	28.1	3.72
GSR-30	96	147	105.33	13.67	28.67	26.75	69.33	3.35	140.66	7.47	1.34	32.37	0.62	22.9	4.21
GSR-31	102	167	100.33	14.67	28.33	25.58	71.22	3.48	165.1	5.96	2.1	37.6	0.62	23.1	4.55
GSR-32	109	158	107	14.67	29.67	30.7	71.44	3.27	156.39	7.57	2.04	33.8	0.6	25.3	3.9
GSR-33	104	153	108.67	10.67	28.33	27.33	67.89	3.78	139.16	7.9	2.06	34.23	0.64	22.6	4.21
GSR-34	104	153	110	11	29.33	20.83	73.44	3.23	151.05	8.81	2.06	36	0.63	22.6	3.75
GSR-35	103	161	107	11.67	30.33	28	71.67	3.57	143.91	9.37	2.08	33.3	0.68	22.4	4.21
GSR-36	104	160	106	11.67	29.67	42.68	71.67	3.25	135.9	8.85	2.1	36.73	0.67	25.4	3.49
GSR-37	106	160	100	11.33	28	21.55	70.11	3.41	141.38	8.44	2.09	36.33	0.61	24.4	3.26
GSR-38	116	160	95.67	11	27.33	26.75	65.67	3.5	149.95	9.51	1.95	35.53	0.58	23.1	2.77
GSR-39	104	159	109	10.33	25	24.5	69.22	2.96	127.76	7.67	2.25	35.93	0.62	24.9	3.25
GSR-40	103	160	105	15.67	25.67	31.55	65.33	3.13	124.33	8.14	2.21	36.07	0.64	25.6	2.12
GSR-41	102	161	104.67	9	27.33	39.28	70.11	3.61	134.36	9.14	2.42	33.6	0.63	29.4	1.71
GSR-42	107	160	97.67	13.67	28	32.7	76	3.73	132.08	7.42	2.21	32.63	0.67	24	4.02
GSR-43	109	155	86.67	13	28	34.65	65.33	3.4	145.65	7.26	2.06	32.67	0.6	25.5	5.27
GSR-44	105	158	91.33	11	25.33	26.03	65.78	3.25	143.15	7.69	2.28	35.67	0.59	28.2	5.74
GSR-45	105	161	100.67	8.33	29	24.38	64.67	3.56	140.74	8.85	2.26	36.6	0.52	26.8	2.92
GSR-46	103	159	104.33	9	27	23.98	66.67	2.91	139.26	8.35	1.88	36.23	0.52	20.6	2.99
GSR-47	103	161	106.33	12.33	30	29.03	69.78	3.35	145	8.79	2.19	36.67	0.6	24.7	4.85
GSR-48	105	150	79	12	30	26.68	67	3.15	156.57	7.47	2.03	39.57	0.58	22.8	2.75
GSR-49	108	153	98.33	6.33	25.67	29.18	72.11	3.67	149.71	7.64	2.17	34.97	0.6	25.8	4.81
GSR-50	92	148	105	20	29	30.6	57.33	2.59	110.58	8.55	1.94	39.17	0.57	19.6	2.61
GSR-51	105	160	93.67	11.67	30.67	24.88	68.33	3.35	126.53	9.93	2.09	39.57	0.66	25.9	4.32
GSR-52	107	153	106.33	10.67	25.67	24.3	72.78	3.79	181.9	10.23	1.91	36.67	0.65	28.3	5.31
GSR-53	107	153	107	9.33	28.33	25.75	66	3.51	163.41	9.64	2.96	39.03	0.54	22.8	4.97
GSR-54	107	155	111	15.67	27.33	26.03	69.89	3.18	158.26	9.7	2.39	37.37	0.59	24.8	5.31
GSR-55	109	158	87.67	13.33	32	33.33	72.78	3.42	188.3	9.54	2.27	32.97	0.57	22.8	4.02
GSR-56	117	125	108	10	27	31.13	70.22	4.28	172.22	7.61	2.17	38.47	0.62	26	4.55
GSR-57	91	148	108.33	9.33	28.33	25.18	60.44	2.72	139.07	7.05	1.98	36.03	0.52	22.9	3.66
GSR-58	109	155	109	13.67	27.67	25.15	71.11	3.95	141.53	10.18	2.4	33.8	0.62	28.2	3.87
GSR-59	107	153	107	13.33	25.67	30.9	73.89	4.13	185.27	10.43	2.24	38.67	0.68	25	4.25
GSR-60	106	155	96.33	9.33	26.67	41.7	74	4.02	155.09	10.64	2.22	41.63	0.64	26.2	3.98
GSR-61	100	161	109	7.33	26	35.38	68.78	3.73	178.4	8.76	3.05	39.33	0.57	23.9	3.74
GSR-62	99	160	116.67	9.33	26.67	27.68	66	3.15	154.64	9.85	2.68	36.23	0.64	19.2	3.62
GSR-63	122	149	106.33	9.67	20	34.15	75.44	3.43	194.67	9.6	2.66	40.17	0.66	26	3.4
GSR-64	94	146	114.33	9	21.67	32.98	91.67	3.26	113.56	7.23	3.61	43.4	0.63	28.2	3.63
GSR-65	94	146	114	11.33	25.67	33.35	83.67	3.03	127.56	7.37	3.35	39.83	0.66	29.3	3.78

DH (days to heading), DM (days to maturity), PH (plant height) in cm, NT (number of tillers), PL (panicle length) in cm, FLA (flag leaf area) in cm^2, LL (leaf length) in cm, CL (culm length) in cm, CD (culm diameter) in mm, G/P (grains per panicle), GL (grain length) in mm, GD (grain diameter) in mm, CC (chlorophyll content), NDVI (net differential vegetation index), TGW (thousand grain weight) in gms, and GY (grain yield) in T/ha.

The overall mean for plant height ranged from 79 cm to 119 cm with a grand mean of 103.08cm. Minimum plant height was observed in GSR 48 that was 79 cm. In LSD analysis, the critical value for comparison among the genotypes was 6.9218. It has been reported that rice plants range in size from dwarf mutants 0.3 to 0.4 m tall to more than 7 m tall floating variety (Chang and Bardenas, 1965). The morphology and varietal characteristics of the rice plant (Vol. 4). Int. Rice Res. Inst. The majority of commercial types are between 1 and 2m tall (Dhungana *et al*., 2022). Panicle length is one element of panicle design that is commonly quantified in terms of yield (Ahmed *et al*., 2016). The overall mean for panicle length ranged from 20cm to 30cm. Grand mean for all genotypes for panicle length was 27.03cm. The number of tillers per plant influence the number of panicles, which is an important factor in grain yield (Promsomboon *et al*., 2019). The overall mean for number of tillers ranged from 5 to 18 with a grand mean of 11.19 for all genotypes for number of tillers (Table 3).

Grand mean for the flag leaf area was 32.7 cm^2 with a variable range in between 20.8 to 47.5 cm^2. A lot of variation was observed for flag leaf area both within the replications and among the genotypes. The most essential component of rice (*Oryza sativa* L.) plant architecture is leaf length, which has a direct impact on yield (Wang and Li, 2005). Maximum length of leaf was observed in GSR 42 followed by GSR 30 and GSR 12 that was1.30cm followed by 1.17cm and 1.17cm, respectively. Minimum leaf length was observed in GSR 31 that was 0.27cm. In LSD analysis, the critical value for comparison of leaf length among the genotypes was 0.3004. The culm is one of the most significant agronomic features, as it determines lodging resistance and the final architecture of crop plants (Ogi *et al*., 1993; Wang *et al*., 2017). The length of the culm (CL) is made up of various extended internode lengths (EILs) (44). Minimum culm length was observed in GSR 50 that was 57.3 cm while the maximum was in GSR 65 that was 77.6cm with a grand mean of 70.5cm. The Culm diameter was variegating in between 2.2mm and 4.4 mm with a grand mean of 3.32mm (Table 3).

The number of grains per plant is being determined by the number of tillers and panicles, which are important factors in grain yield (Li *et al*., 2022). Significant variation was observed among the genotypes which is governed by various factors. Variation was ranging from 108 to 228 grains per plant with a grand mean of 153.05. A lot of variation was observed for grain length and grain diameter as it was fluctuating in between 5.9 to 8.9mm and 1.30 to 3.15mm respectively. GSR 45, GSR 34 and GSR 63 were found to have longer grains with a length measurement of 8.85mm, 8.81mm and 8.69mm, respectively. The amount of chlorophyll in leaves can reveal a lot about a plant's physiological state (Wang *et al*., 2014). The overall mean for chlorophyll content ranged from 30 to 44. Grand mean of all genotypes for chlorophyll content was 35.91. Maximum chlorophyll content was observed in GSR 64 followed by GSR 60 and GSR 63 that were 43.40 followed by 41.63 and 40.17 respectively. Minimum chlorophyll content was observed in GSR 13 that was 30.05. One of the most widely used vegetation indices is the NDVI. It has been used for staple crop management (González- and Mayorga, 2018). The overall mean for NDVI ranged from 0.55 to 0.71 with a grand mean of a 0.64. Maximum NDVI was observed in GSR 59 followed by GSR 35 and GSR 36 that were 0.71 followed by 0.71 and 0.70, respectively. Minimum NDVI was observed in GSR 45 and GSR 46 that was 0.55 (Table 3).

The overall mean for total grain weight ranged from 19 to 35. Grand mean of all genotypes for total grain weight was 25.84gm. Maximum total grain weight was observed in GSR 15 followed by GSR 19 and GSR 41 that were 34.26gm followed by 30.96gm and 30.46gm respectively. Minimum total grain weight was observed in GSR 14 that was 19.36. Grain yield in rice is a complex agronomic trait. The number of panicles, the number of grains per panicle, the grain weight, and the grain filling all affect rice grain yield (or the percentage of ripened grains per total grains). The overall mean for grain yield ranged from 1.71 tons/ha to 6.18 tons/ha. Grand mean of all genotypes for grain yield was 4.1 tons/ha. Maximum grain yield was observed in GSR 11 followed by GSR 10 and GSR 44 that were followed by 5.88 and 5.74 tons/ha, respectively. Minimum Grain yield was observed in GSR 41 that was 17.65. Standard deviation for this trait was recorded as 9.370. In LSD analysis, the critical value for comparison of grain yield among the genotypes was 11.254. Also, the three genotypes with maximum grain yield were found to be lying in the same group along with other genotypes such as GSR 54, GSR 52, GSR 43, GSR 3 and GSR 2 (Table 3).

Qualitative data analysis

Significant variation was also found in case of qualitative characters as obvious groups were formed based on visual observations. In terms of leaf color most of the genotypes showed yellowish green color. 22 genotypes showed yellowish color, 30 genotypes showed yellowish green color while light green color was observed in remaining 13 genotypes. Rice genotypes were evaluated for flag leaf attitude as erect, semi erect or horizontal. 47 of the total genotypes showed erect attitude of flag leaf, 15 of them showed semi erect type of flag leaf attitude while in remaining three genotypes horizontal type of flag leaf attitude was observed (Figure 1).

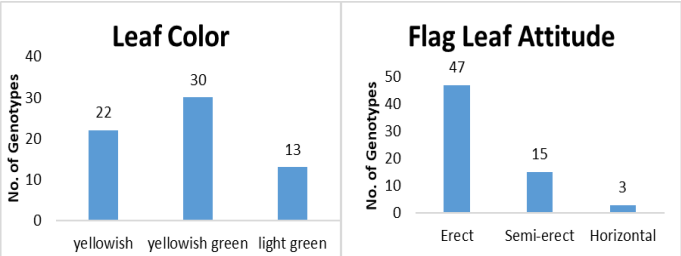

Figure 1: *Distribution of 65 GSR lines for leaf color and attitude.*

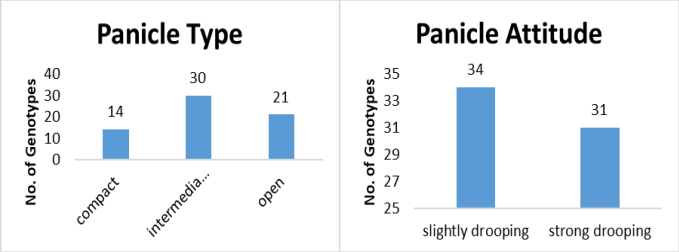

Figure 2: *Distribution of 65 GSR lines for panicle type and attitude.*

Panicle attitude is either slightly drooping or strong drooping depending on the genotype. As in these genotypes 34 rice genotypes have shown slightly drooping type of panicle attitude while 31 of them showed strong drooping type of panicle attitude. Most of the genotypes showed intermediate type of panicle. 30 genotypes among 65 showed intermediate panicle type, 21 genotypes showed open panicle type while compact type of panicle was observed in 14 genotypes (Figure 2). Awning in rice ranges from absent, short and fully awned to long and fully awned. Most of the genotypes under study showed no awning (47 genotypes). 15 of the rest were short and fully awned while only three genotypes were long and fully awned. Awn color was classified as straw, brown and gold. 47 of the genotypes had no awning so out of remaining 18 genotypes 11 showed straw colored awns, 5 were having brown awns while only 2 had gold awns (Figure 3).

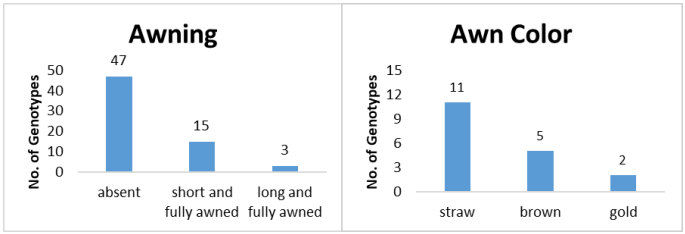

Figure 3: *Observation of diverse rice for Awn presence and coloring.*

Caryopsis shape was classified as round, semi round, half spindle shape and spindle shape. In this data, 25 genotypes were having semi round caryopsis shape, 22 had half spindle shape, and 15 of them showed spindle shape while only three showed round caryopsis shape. Chalkiness is a serious concern in rice for quality estimation. In the study 23 genotypes showed no chalkiness, 21 showed small chalkiness, 9 of them showed medium chalkiness while 12 of them showed considerable amount of chalkiness (Figure 4).

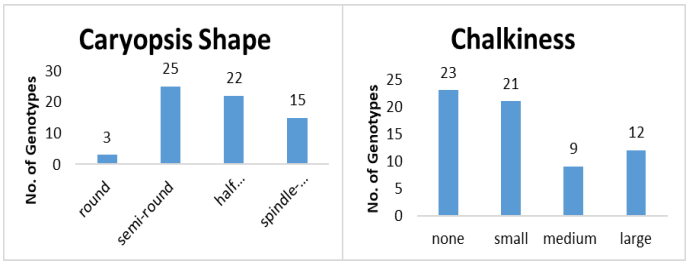

Figure 4: *Characterization of diverse rice germplasm for grain structure.*

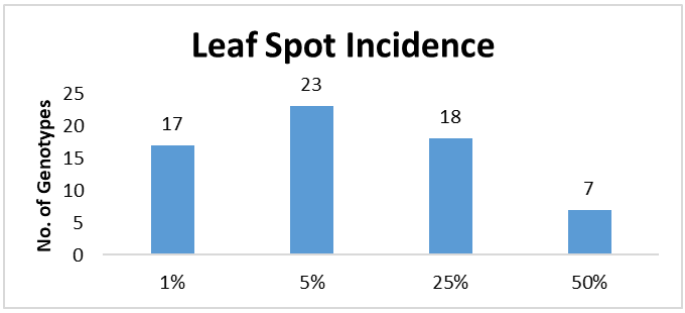

Figure 5: *Scoring of leaf spot found on the leaves of diverse rice material.*

Leaf spot is a common fungal disease that affects rice plants. It is caused by various species of fungi, including *Bipolaris oryzae, Cercospora oryzae*, and others. Leaf spot can reduce rice yield and quality if not managed properly (Uppala and Zhou, 2018). Leaf spot incidence was categorized in 4 classes depending on the percentage of incidence varying from 1%, 5%, 25% and 50%. 23 of the genotypes showed 5% leaf spot incidence, 25% leaf spot incidence was found in 18 genotypes, 1% leaf spot incidence was found in 17 genotypes and 50% of leaf spot incidence was observed in 7 of the genotypes (Figure 5).

Principal component analysis (PCA)

Principal component analysis (PCA) is the process of computing the principal components and using them to change the basis of the data, often simply using the first few and disregarding the rest. PCA is utilized in exploratory data analysis and predictive model development (Jolliffe and Cadima, 2016). PCA is a statistical process that allows you to summarize the information content of big data tables using a smaller collection of "summary indices" that can be viewed and examined more readily (Eriksson, 2020).

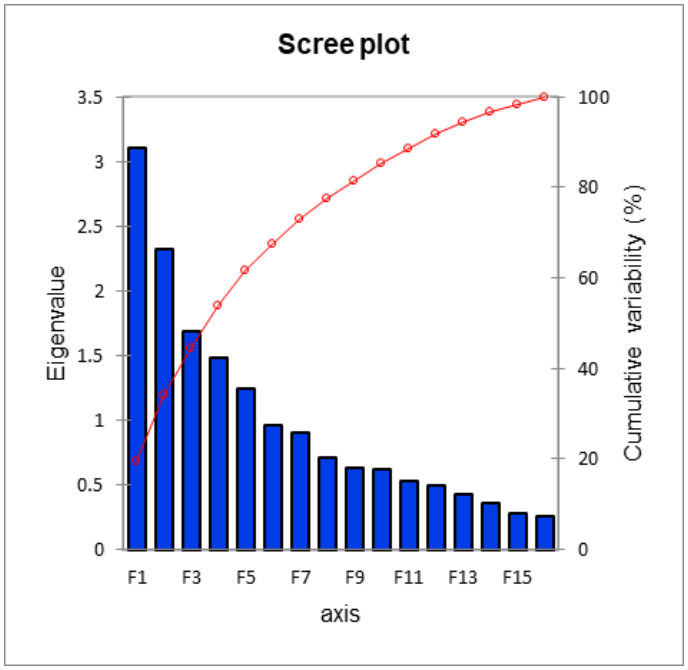

Figure 6: *Scree plot for the diverse rice genotypes for variability contribution.*

The first two components were found to contribute 33.902% of the total variability (Figure 6) so, the biplot was created using the first two components. PC 1 exhibited 19.394% of the total variability and explained major component weight by flag leaf area, culm length, number of grains per panicle, grain diameter and grain yield (Figure 7). In the first component highest weights were found for culm length (0.627) and grain diameter (0.615) indicating that PC1 was mostly associated with culm length and grain diameter (Figure 6). The nature and strength of relationship between two parameters can easily be depicted from Biplot. The length of line and angle between two lines represent the magnitude and type of association between them, respectively (Sathish and Senapati, 2017). Culm length was found to have maximum factor load in first principal component followed by grain diameter. Genotypes GSR 19, GSR 63, GSR 64 and GSR 7 were found to be scattered around culm length, grain diameter and grain yield so these can be selected for higher grain diameter as well as grain yield on the basis of PCA.

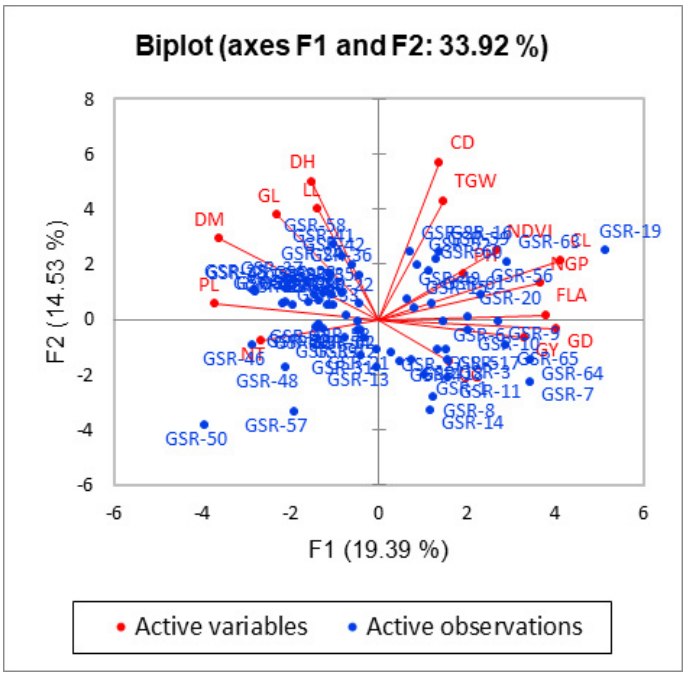

Figure 7: *Biplot of 1st and 2nd principal component for morphological parameters.*

Second principal component (PC2) explained 14.531% of total variability and found to be positively and significantly correlated with parameters days to heading, leaf length, culm diameter and thousand grain weight. In the second component highest weights for were found for culm diameter (0.750) and days to heading (0.660). As a result, genotypes situated at the biplot's extremities must be picked in order to achieve maximum genetic diversity retained in a new population after constructing a selection cycle. Principal component analysis has shown the genetic diversity of the germplasm. Plant breeders routinely evaluate and characterize germplasm, and the PCA tool, cluster, and multivariate statistical analysis give a valuable means of evaluating morphological variability within and between germplasm collections (Baranwal *et al*., 2013). These techniques are being utilized to detect significant changes in germplasm and extent of departure among crop species and they are valuable for evaluating prospective breeding value.

Cluster analysis

Cluster analysis is a set of statistical classification techniques in which cases are sorted into groups (clusters) based on a certain measure of similarity, so that the objects in one cluster are very similar to one another and quite distinct from the cases in other

Table 4: *Categorization of diverse rice germplasm into smaller number of homogeneous groups.*

Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V
G1, G7, G9, G11, G13, G14, G17, G19, G20, G22, G52, G55, G56, G59, G61, G63	G2, G3, G4, G5, G6, G8, G10, G12, G43, G44, G48, G49, G60	G15, G16, G21, G25, G31, G32, G53, G54, G62	G18, G27, G29, G36, G39, G40, G41, G50, G57, G64, G65	G23, G24, G26, G28, G30, G33, G34, G35, G37, G38, G42, G45, G46, G47, G51, G58
16	13	9	11	16

clusters (Steinbach *et al*., 2005). One of the simplest type of cluster analysis is k-means cluster analysis. On a table of raw data, k-means cluster analysis is performed, with each row representing an item and the columns representing the objects' quantitative attributes. Clustering variables refer to certain quantitative properties (Steinley, 2006).

Based on the data collected for morphological criteria, K-mean cluster analysis classified 65 different genotypes into five unique clusters. A collection of genotypes with similar performance was discovered in each cluster. The genotypes in one cluster were all closely related and didn't differ significantly. Because the quantity of genotypes in each cluster differs from that of other clusters, genotype distribution is unquestionably influenced by parameter values rather than an equal distribution of genotypes. Cluster 1 had 16 genotypes, cluster 2 had 13 genotypes, cluster 3 had 9 genotypes, cluster 4 had 11 genotypes, and cluster 5 had 16 genotypes. Clusters 1 and 5 had the most genotypes, each with 16 genotypes, while cluster 3 had the fewest genotypes, with only 9 genotypes grouped together (Table 4).

Conclusions and Recommendations

The study helped in finding out the Genetic diversity in 65 Rice genotypes. It aided in morphological characterization of diverse Rice germplasm. Suitable genotype for application in Agriculture enhancement programs were investigated. Out of 65 genotypes, the genotype performing better in terms of morphological traits were analysed which will be exploited in respective crossing schemes. There is enough diversity within the germplasm studied here, however for stability of traits these lines must be analyzed for multi-location studies. Lines with maximum diversity may be studied for additional traits e.g. high temperature, salinity and drought stress etc. Lines with highest yield as GSR11, GSR10 and GSR 41 may be crossed back to other adapted lines as parental lines for enhancing yield potential.

Acknowledgements

We are thankful to Rice Program Crop Sciences Institute NARC Islamabad for provision of germplasm and experimental site with field operational services. We are grateful to the PSDP PE Rice Project for financial assistance and Department of Botany Hazara University Mansehra for theoretical assistance and guidance.

Novelty Statement

The diverse rice germplasm has been characterized based on morphological and quality parameters. The potential genotypes will be gone through extensive selection cycles for picking the best ones and others can be exploited in the breeding programs

Author's Contribution

Salma Sharif: Conducted experiment, recorded data and drafted manuscript.
Rana Arsalan Javaid: Performed statistical analysis and scientific writing.
Abid Majeed and Muhammad Shahzad Ahmed: Assisted in field experimentation and reviewed paper.
Qurat ul Ain Sani: Took part in data recording and write-up.
Faiza Siddique: Participated field activities and data collection.
Muhammad Arshad and Niaz Ali: Supervised the study.

Conflict of interest

The authors have declared no conflict of interest.

References

Ahmed, M.Z., B. Gul, M.A. Khan and K.N. Watanabe. 2016. Characterization and function of sodium exchanger genes in *Aeluropus lagopoides* under NaCl stress. In: Halophytes for food security in dry lands. Academic Press.

pp. 1-16. https://doi.org/10.1016/B978-0-12-801854-5.00001-7

Baranwal, D., V. Mishra an T. Singh. 2013. Genetic diversity based on cluster and principal component analyses for yield and its contributing characters in wheat (*Triticum aestivum* L.). Madras Agric. J., 100: 320–323

Chang, T.T. and E.A. Bardenas. 1965. The morphology and varietal characteristics of the rice plant. Int. Rice Res. Inst., Vol. 4.

Dhungana, B., J. Gautam, A. Adhikari, P. Ale, A. Adhikari, S. Subedi and K.H. Dhakal. 2022. Varietal evaluation and preference analysis of sixteen released rice varieties in Bhojad, Chitwan, Nepal.

Eriksson, L., 2020. What is principal component analysis (PCA) and how it is used. Umetrics Suite Blog. Available online: https://blog.umetrics.com/what-is-principal-component-analysis-pca-andhow-it-is-used.

Gangapur, D.R., K. Thiyagarajan, S. Magadum and D. Malarvizhi. 2014. Genetic diversity analysis of rice genotypes for yield and yield attributing traits. Environ. Ecol., 32(3): 803-807.

Gao, H., M. Jin, X.M. Zheng, J. Chen, D. Yuan, Y. Xin and J. Wan. 2014. Days to heading, a major quantitative locus determining photoperiod sensitivity and regional adaptation in rice. Proc. Natl. Acad. Sci., 111(46): 16337-16342. https://doi.org/10.1073/pnas.1418204111

González-Betancourt, M., and Z.L. Mayorga-Ruíz. 2018. Normalized difference vegetation index for rice management in El Espinal, Colombia. Dyna, 85(205): 47-56.

Jolliffe, I.T. and J. Cadima. 2016. Principal component analysis: A review and recent developments. Philos. Trans. R. Soc. A: Math., Phys. Eng. Sci., 374(2065): 20150202. https://doi.org/10.1098/rsta.2015.0202

Kherif, F. and A. Latypova. 2020. Principal component analysis. In: Machine Learning. Academic Press. pp. 209-225. https://doi.org/10.1016/B978-0-12-815739-8.00012-2

Li, G., B. Xu, Y. Zhang, Y. Xu, N.U. Khan, J. Xie and Z. Li. 2022. RGN1 controls grain number and shapes panicle architecture in rice. Plant Biotechnol. J., 20(1): 158. https://doi.org/10.1111/pbi.13702

Lutatenekwa, D.L., E.J. Mtengeti and G.M. Msalya. 2020. A review of plant characterization: First step towards sustainable forage production in challenging environments. Afr. J. Plant Sci., 14(9): 350-357. https://doi.org/10.5897/AJPS2020.2041

Nice, L.M., B.J. Steffenson, G.L. Brown-Guedira, E.D. Akhunov, C. Liu, T.J. Kono and G.J. Muehlbauer. 2016. Development and genetic characterization of an advanced backcross-nested association mapping (AB-NAM) population of wild× cultivated barley. Genetics, 203(3): 1453-1467. https://doi.org/10.1534/genetics.116.190736

Ogi, Y., H. Kato, K. Maruyama and F. Kikuchi. 1993. The effects on the culm length and other agronomic characters caused by semi dwarfing genes at the sd-1 locus in rice. Jpn. J. Breed., 43(2): 267-275. https://doi.org/10.1270/jsbbs1951.43.267

Promsomboon, P., R. Sennoi, T. Puthmee, R. Marubodee, W. Ruanpan and S. Promsomboon. 2019. Effect of seedlings numbers per hill on the growth and yield of Kum Bangpra rice variety (*Oryza sativa* L.). Int. J. Agric. Technol., 15(1): 103-112.

Roychoudhury, A., 2020. Rice research for quality improvement: Genomics and genetic engineering. Springer. https://doi.org/10.1007/978-981-15-4120-9

Saeed, S. and M.Y.K. Barozai. 2021. A review on genetic diversity of wild plants by using different genetic markers. Pure Appl. Biol., 1(3): 68-71. https://doi.org/10.19045/bspab.2012.13004

Sathish, V. and B. Senapati. 2017. Principal component analysis of morphological and yield attributing traits in advanced breeding lines of rice grown at rainfed shallow lowland condition of west Bengal. Int. J. Biol. Res. Dev., 7(3): 1-6. https://doi.org/10.24247/ijbrdoct20171

Shrestha, J., 2016. Cluster analysis of maize inbred lines. J. Nepal Agric. Res. Counc., 2: 33-36. https://doi.org/10.3126/jnarc.v2i0.16119

Singh, P.K., K. Venkatesan and T.P. Swarnam. 2018. Rice genetic resources in tropical islands. In Biodiversity and climate change adaptation in tropical Islands. Academic Press. pp. 355-384. https://doi.org/10.1016/B978-0-12-813064-3.00012-0

Sreevalsan-Nair, J., 2021. K-means clustering. In: Encyclopedia of mathematical geosciences. Cham: Springer International Publishing. pp. 1-3. https://doi.org/10.1007/978-3-030-26050-7_171-1

Steel, R.G.D. and J.H. Torrie. 1980. Principles and procedures of statistics. New York: McGraw-Hill.

Steinbach, M., V. Kumar and P. Tan. 2005. Cluster analysis: Basic concepts and algorithms. Introduction to data mining, 1st edn. Pearson Addison Wesley.

Steinley, D., 2006. K-means clustering: A half-century synthesis. Br. J. Math. Stat. Psychol., 59(1): 1-34. https://doi.org/10.1348/000711005X48266

Takai, T., P. Lumanglas, E.V. Simon, Y. Arai-Sanoh, H. Asai and N. Kobayashi. 2019. Identifying key traits in high-yielding rice cultivars for adaptability to both temperate and tropical environments. Crop J., 7(5): 685-693. https://doi.org/10.1016/j.cj.2019.06.004

Uppala, S. and X.G. Zhou. 2018. Field efficacy of fungicides for management of sheath blight and narrow brown leaf spot of rice. Crop Prot., 104: 72-77. https://doi.org/10.1016/j.cropro.2017.10.017

Wang, J., S. Gang, L. Yang, H. Zheng, J. Sun, H. Liu and D. Zou. 2017. Markers associated with culm length and elongated internode length in Japonica rice. Crop Sci., 57(5): 2329-2344. https://doi.org/10.2135/cropsci2016.06.0533

Wang, Y. and J. Li. 2005. The plant architecture of rice (*Oryza sativa*). Plant Mol. Biol., 59(1): 75-84. https://doi.org/10.1007/s11103-004-4038-x

Wang, Y., D. Wang, P. Shi and K. Omasa. 2014. Estimating rice chlorophyll content and leaf nitrogen concentration with a digital still color camera under natural light. Plant Methods, 10(1): 1-11. https://doi.org/10.1186/1746-4811-10-36