



Variance analysis for grain number related traits and study of population structure in rice (*Oryza sativa*)

N MOHANTY¹, S K NAYAK², J KUMAR³, S MOHANTY¹, J MOLLA⁴ and L BEHERA^{1*}

ICAR-National Rice Research Institute, Cuttack, Odisha 753 006, India

Received: 13 April 2023; Accepted: 28 July 2023

ABSTRACT

Utilising average data from the rainy (*kharif*) seasons of 2021 and 2022, the genetic diversity of 188 recombinant inbred lines (RILs) of rice (*Oryza sativa* L.) was evaluated at the research farm of ICAR-National Rice Research Institute, Cuttack, Odisha. The grain number and associated 10 characteristics' coefficient of phenotypic variability was substantial, with the grain number per panicle having the greatest PCV and GCV. Principal component analysis was used to further identify the relationships and trends among the RILs. The first four primary components (74.58%) fully described the variability of all 10 features. Cultivars were divided into 8 groups based on the characters used to attribute grain number. To analyse genetic differences between RILs at the molecular level, 22 SSR markers were utilised and the PIC value was 0.709. A strong and significant relationship between them was shown by the structural analysis, with spikelet number and spikelet fertility per cent indicating a particularly strong correlation. Between clusters 1 and 3, there was considerable genetic variety, which offers great breeding options.

Keywords: Correlation, Cluster, Genetic diversity, Grain number, PCA

The global significance of scientific research to enhance rice (*Oryza sativa* L.) yield is a pressing concern in the face of escalating food demands. Estimates suggest that by 2050, the aggregate rice demand in Asia is projected to soar by a staggering 90%, a trajectory that falls short of meeting the nutritional needs of the anticipated 9 billion people. Crucial in rice cultivation, grain yield is intricately moulded by diverse factors. Among these are biomass, harvest index, and tillers per plant, as highlighted by Harsha *et al.* (2017). Equally influential are traits like 1000-grain weight and panicle length, studied by Deep *et al.* (2017) and Singh *et al.* (2020), respectively, revealing the intricate orchestration of grain yield. Efforts to design targeted breeding programmes underscore the necessity of harnessing existing variability. Genetic insights, encompassing genotypic coefficient of variation (GCV), phenotypic coefficient of variation (PCV), heritability, and genetic advancement, gleaned from diverse traits within recombinant inbred populations as studied by Mehmood *et al.* (2021) and Maurya *et al.* (2022), form the bedrock for focused advancements. Integral to yield enhancement

is the correlation coefficient between grain number and its component attributes, emphasized by Prasad *et al.* (2020). Comprehending the intricate trait interplay not only informs breeding strategies, but also holds potential for substantive yield amplification. Within crop improvement pursuits, the extent of genetic diversity assumes paramount significance. To provide robust selection guidance, an inclusive grasp of constituent trait interplay and their impact on grain yield is indispensable. Here, principal component Analysis (PCA) and cluster analysis, as illuminated by Tiwari *et al.* (2019), emerge as potent tools for elucidating interdependencies and prioritizing attributes for yield optimization. Against this backdrop, the current study embarks on a comprehensive exploration of the relationships underpinning yield-related traits-grain number, spikelet number, tiller number, and more-within Recombinant Inbred Lines (RILs) derived from high- and low-grain rice genotypes. The aim was to elucidate optimal plant attributes that, when selected for, indirectly pave the way for heightened rice grain numbers and, consequently, amplified yield potential.

MATERIALS AND METHODS

This study aimed to evaluate phenotypic traits in 188 recombinant inbred lines (RILs) (F_8) of rice developed from a cross between 2 different genotypes. The study was carried out over 2 rainy (*kharif*) seasons (2021 and 2022) at ICAR-National Rice Research Institute, Cuttack, Odisha. The 2 parent genotypes have distinct characteristics, with PDK

¹ICAR-National Rice Research Institute, Cuttack, Odisha;
²Maharaja Sriram Chandra Bhanjadeo University, Mayurbhanj, Baripada, Odisha; ³Central Rainfed Upland Rice Research Station, ICAR- National Rice Research Institute, Hazaribagh, Jharkhand; ⁴Ghatal Rabindra Satabarsiki Mahavidyalaya, Medinipur, West Bengal. *Corresponding author email: l_behera@yahoo.com

Shriram having a high grain number and bunchy panicles but short and slender grains, while Heera has a low grain number but bold grains. The recommended doses of P_2O_5 and K_2O (40 kg/ha) along with a half-dose of nitrogen were applied at the time of final field preparation before transplanting. The remaining 40 kg of nitrogen was split 50/50 at the tillering and grain-filling stages. The study recorded 10 morphological traits, viz. days to 50% flowering (DFF), plant height (PH), tiller number (TN), panicle number (PN), panicle length (PL), grain number (GN), chaff number (CN), spikelet number (SN), spikelet fertility percentage (SFP), and thousand-grain weight (TGW). DNA was extracted from 25-day-old seedlings using the CTAB-modified method and PCR amplification was done using 22 polymorphic single sequence repeat (SSR) markers. The PCR-amplified products were separated on a 3% agarose gel containing ETBR and visualized using a gel documentation system. The bands were scored as 1 (present) or 0 (absence) to create a raw data matrix for future analysis. The data were taken from two seasons to calculate mean values for phenotypic traits and ANOVA, and diversity studies. Phenotypic correlation and PCA analysis were performed using the PAST 3.0 version software. Genetic diversity, allele frequency, gene diversity, observed heterozygosity and polymorphic information content (PIC) were calculated using POWER MARKER Ver3.25. A UPGMA dendrogram was created using DARwin 6.12 based on the Neighbour-Joining (NJ) approach to illustrate the genetic links between RILs. Population structure was analysed using Structure v2.3.4 software, and the K value was selected from 1 to 10 with 3 replications. The structure result was uploaded to "structure harvester," an online platform to obtain the final population structure.

RESULTS AND DISCUSSION

Descriptive statistics: The mean and range values for various traits related to grain number in parents and 188 RILs during *kharif* seasons of 2021 and 2022 showed significant variation. These traits include DFF, PH, TN, PN, PL, GN, CN, SN, SFP% and TGW. For parents Heera and PDK Shriram, the mean values for these traits were observed to vary widely, indicating significant genetic variation between the two parents (Table 1). The RILs exhibited a wide range of variation for these traits, with mean values ranging from 9.39 for tiller numbers per plant to 198.86 for spikelet numbers per panicle. The range of values for days to 50% flowering was observed to be from 71 to 132 days, while the range for 1000-grain weight was from 9.40 g to 33.50 g. Overall, these findings suggest that the genetic makeup of the RILs is diverse, and there is a high degree of variation among them. These results can be valuable in breeding programmes aimed at improving grain yield in rice and other crops.

Genetic Variability: This study examined the relationship between PCV and GCV values for different traits in the RIL population. It found that PCV was higher than GCV for all traits, except for tiller number (Table 1),

Table 1 Descriptive statistics, Genetic parameter estimates, heritability in the broad sense, and genetic advancement of RILs along with parents for grain number and related traits

Trait	Mean						Range				GCV	PCV	H2 %	GA %	GAM	CD	CV %
	PDK Shriram			Heera			RILs		Range								
	2021	2022	2021	2022	2021	2022	2021	2022	2021	2022							
DFF	114	119	61	58	95.73	94.3	64.00–132.00	57.5–115	11.816	11.961	0.977	22.805	22.547	2.817	1.851		
PH	110	115.3	70.7	78	112.3	107	70.67–163.33	71.17–162.2	14.499	14.753	0.966	32.239	29.354	4.749	2.681		
TN	10.3	10.67	6	6.67	9.39	7.76	6.00–23.00	4.67–13	13.824	20.494	0.348	1.299	30.033	1.638	11.965		
PN	9	10	5.5	6	8.04	8.04	5.00–13.33	4.5–11.5	16.427	20.051	0.671	2.333	27.723	1.552	11.497		
PPL	26.9	27.17	23.8	24.57	23.96	24.1	19.87–35.10	20.33–29.83	17.852	12.713	0.459	2.617	5.613	1.176	3.028		
GN	328	322.3	68.9	72	157.4	150	34.44–330.89	44.67–318.2	35.358	37.519	0.893	98.196	68.822	41.236	15.935		
CN	62.5	81.67	27.8	34.67	41.47	44.1	16.11–134.89	16.5–128.2	51.183	54.453	0.887	42.552	99.187	17.291	24.019		
SN	391	404	96.7	106.7	198.9	194	78.67–396.56	79.17–399.7	28.567	30.527	0.879	100.109	52.734	42.395	13.551		
SFP	84	79.79	71.4	67.5	78.28	76.7	29.55–93.96	36.37–89.66	15.709	16.149	0.947	24.213	31.481	4.6	3.721		
TGW	18	19.5	26.9	28.7	22.54	22.5	9.40–33.50	8.4–37.5	20.279	21.652	0.878	8.823	39.155	2.685	7.418		

Trait details are given under Materials and Methods.

suggesting that environmental factors had a relatively minor impact on the expression of these traits (Prasad *et al.* 2020, Prajapati *et al.* 2022). The study also found that some traits, such as days to 50% flowering, plant height, spikelet fertility, chaff number, 1000-grain weight, and grain number, exhibited high heritability and high genetic progress, suggesting that they are good candidates for selection in plant breeding programmes. However, traits such as tiller number and panicle length showed very low heritability and low genetic progress, indicating that they may not be good targets for selection. The study emphasizes the importance of understanding the genetic and environmental factors that contribute to the expression of various traits in a population, as this knowledge can help in identifying the traits that can be improved through genetic selection and developing strategies to enhance crop productivity and yield (Mehmood *et al.* 2021, Zhang *et al.* 2022 and Maurya *et al.* 2022). According to Pachauri *et al.* (2017), the nature of gene action can affect the relationship between heritability and genetic advancement for different traits in plants.

Analysis of variance (ANOVA), phenotypic correlation and principal component analysis: This study found significant variations in all traits, with panicle number, spikelet number, and 1000-grain weight being adversely significant. Similar findings have been reported by Amegan *et al.* (2020) and Mia *et al.* (2022). Pearson correlation coefficients were used to analyse the data and found that most traits were positively and significantly correlated with grain number, except for tiller number and spikelets per panicle (Fig 1). The degree of linkage among traits was deemed important in understanding their relationship with grain number. The study also found that the environment had a suppressing effect on the expression of traits at the phenotypic level. Positive associations between different traits and grain number were discussed, such as tiller number, panicle length, and grain number (Lakshmi *et al.* 2017, Bagudam *et al.* 2018, Ghimire and Mahat 2019, Bhargava *et al.* 2021, Prajapati *et al.* 2022)

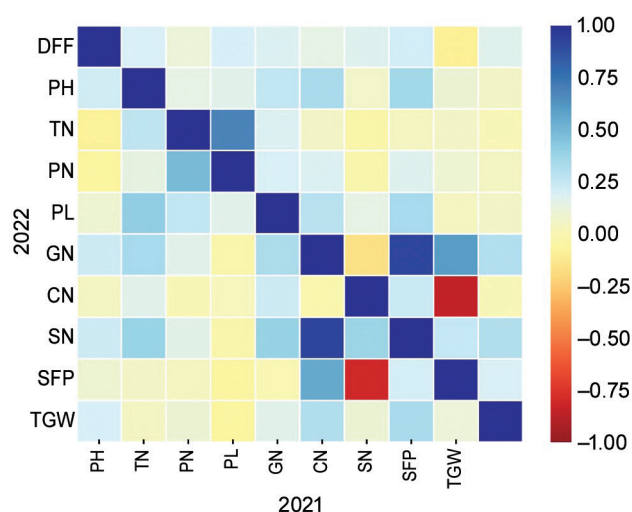


Fig 1 Significance of Pearson correlation coefficients among 10 phenotypic traits (Based on data over 2 years).

Principal component analysis was used to analyze the genetic variance among genotypes with respect to various phenotypic traits. Three main axes, PC1, PC2, and PC3, had eigen values greater than one together accounted for 74.58% of the total variance (Table 2). A scatterplot type utilised in PCA is called a biplot. In this unique figure, the original data is represented by principal components that, when combined with loading vectors and PC scores, account for the bulk of the data variation. Specific traits that were strongly associated with each principal component were grain number, spikelet number and spikelet fertility percentage (Supplementary Fig 1). All traits are considered important factors responsible for grain number, and a thorough assortment procedure could be intended to bring about the rapid advance of the dependent variables by electing lines as of PC1 (Mohanty *et al.* 2017, Thakur and Sarma 2023)

Cluster analysis based on the phenotypic traits and molecular markers: Clustering is the task of grouping a set of genotypes in such a way that objects are in the same group. RILs developed from PDK Shriram and Heera were statistically analysed for diversity and similarity in their quantitative traits using cluster analysis. Major cluster I was divided into two sub-clusters, IA and IB (Supplementary Fig 2). This cluster contained 160 RILs included 21 RILs showing similar quantitative characters to Heera. Genetic variability among the parents has a great influence on the variability spectrum of segregating lines, making diversified RILs preferred for breeding approaches to get desired varieties (Mia *et al.* 2022, Thakur and Sarma 2023).

The graphic shows the neighbour-joining (NJ) tree that was created as a result of the genetic distance between 188 RILs and parents (Fig 2). It was divided into 1A, 1B and 2A, 2B, with cluster 1 having 118 genotypes along with Heera and cluster 2 having 72 genotypes with PDK Shriram. 41

Table 2 Principal component analysis (PCA) of 10 phenotypic traits in RILs (Based on data over 2 years)

Trait	PC 1	PC 2	PC 3	PC 4
DFF	0.19	0.3	-0.01	0.33
PH	0.3	0.16	-0.03	-0.42
TN	0.2	0.17	0.64	0.08
PN	0.27	0.18	0.58	0.13
PL	0.28	0.23	0	-0.39
GN	0.53	-0.13	-0.22	-0.05
CN	-0.08	0.63	-0.25	0.03
SN	0.49	0.12	-0.31	-0.04
SFP	0.32	-0.57	0.08	-0.02
TGW	0.24	-0.01	-0.21	0.73
Eigen value	2.9	1.98	1.59	0.98
Per cent variance	29.04	19.83	15.9	9.81
Cumulative variance per cent	29.04	48.87	64.77	74.58

Trait details are given under Materials and Methods.

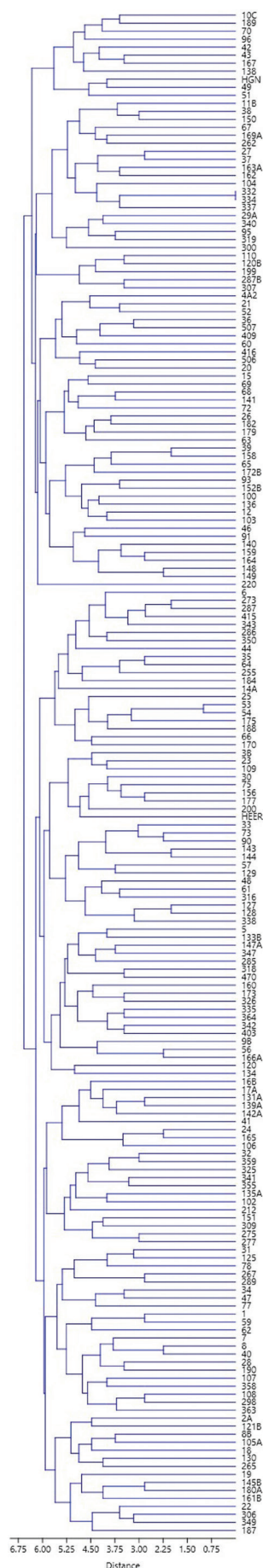


Fig 2 UPGMA cluster showing genetic relationship among RILs and parents based on the molecular data.

RILS genotypes were included in Heera type and 11 RILS counted as PDK Shriram type. The selection of parental traces for similar hybridization and genetic enhancement can benefit rice breeders with this information (Mohanty *et al.* 2017, 2018).

Similar clustering was observed based on phenotypic and molecular data. This study identified that all 22 SSR primers were polymorphic between RILs. The average major allele frequency was 0.571, with a range of 0.497 to 0.791 (Table 3). The marker with the highest major allele frequency was RM27879, whereas the markers with the lowest major allele frequencies were RM11738 and RM14294. The average heterozygosity number (0.506) was accurate, and the allelic diversity and frequency were reflected in the PIC value (0.709). RM26343 and RM17784 were shown to be the most effective markers, with a PIC value of 0.5 or higher. The markers with a PIC value of 0.5 or higher are extremely beneficial and very instructive. The findings could be useful for rice breeders and geneticists in developing new rice varieties with desirable traits. Similar results were also reported by Anik *et al.* (2021) and Ogunbayo *et al.* (2022).

Structure analysis: Based on a log mean probability and change in log probability (ΔK), the structure harvester revealed $k = 8$, indicating there are 8 sub-populations among the 188 RILs (Fig 3). This result shows similarities with the dendrogram result. The population structure of 188 RILs was examined using software for STRUCTURE analysis. The mapping population under examination contains 8 sub-populations, according to the greatest K value, which was discovered at $K = 8$, and was therefore chosen to illustrate the genetic makeup of the 188 genotypes. In two distinct rice diversity panels, two sub-populations were detected and the structure was examined by Zhang *et al.* (2022).

A lot of genetic variation in the parents and RILs was identified. This moves the traits in question in a more positive direction. For all of the traits under study, the results showed that PCV was higher than GCV, which may be related to the context in which these traits are expressed. Prajapati *et al.* (2022) studied genetic variability and trait correlation in F_2 segregating rice crosses, finding significant genetic variability and potential for improvement through selection. Maurya *et al.* (2022) found high variability in plant height, grain width and spikelet, with high biological yield heritability. Phenotypic correlation was stronger than genotypic. When designing a breeding programme to boost yield, breeders should keep this in mind. All the traits investigated had a significant positive correlation with grain number, with the exception of chaff and tiller numbers. The first four PCA components, which together accounted for 74.58% of the overall variation, provided an assessment of the proportional contribution of different variables to the genotypes under consideration. Saha *et al.* (2022) identified five clusters, explaining 72.9% of variance in the principal component by enabling early, high-yielding rice varieties with premium grain quality.



Fig 3 Maximum number of sub-populations estimated at K = 8 (Population structure plot on single line with Q matrix).

Table 3 Polymorphic information and heterozygosity of RILs using 22 SSR polymorphic markers

Marker	Chr#	MAF	Ho	PIC	Marker	Chr#	MAF	Ho	PIC
RM11522	1	0.518	0.540	0.722	RM8107	6	0.681	0.450	0.697
RM11738	1	0.497	0.510	0.742	HVSSR75	6	0.55	0.547	0.705
RM13854	2	0.602	0.499	0.721	RM20783	7	0.581	0.507	0.725
RM13861	2	0.534	0.545	0.714	RM22143	7	0.513	0.533	0.728
RM14294	3	0.497	0.515	0.740	RM22263	8	0.524	0.535	0.724
RM251	3	0.521	0.620	0.633	RM264	8	0.503	0.555	0.714
RM231	3	0.524	0.535	0.724	RM484	10	0.66	0.459	0.711
RM16947	4	0.607	0.481	0.733	RM26343	11	0.508	0.505	0.745
RM16952	4	0.628	0.482	0.718	RM26474	11	0.576	0.493	0.739
RM17784	5	0.518	0.504	0.745	RM28616	12	0.644	0.485	0.656
RM18795	5	0.581	0.491	0.738	RM27879	12	0.791	0.349	0.530
Mean		0.571	0.506	0.709	Mean		0.571	0.506	0.709

MAF, major allele frequency; Ho, heterosis; PIC, polymorphic information content.

There were three main groups for the genotypes and three for the molecular group in the morphological dendrogram. Significant F_1 heterosis may occur from crosses between these distant clusters. The PIC values ranged from a low of 0.530 (RM27879) to a high of 0.745 (RM17784, RM26343), and the average was 0.506, demonstrating that they were quite useful for figuring out the rate of polymorphism of a marker at a certain locus. Ogunbayo *et al.* (2022) found moderate to substantial genetic diversity in rice, with eight PC axes explaining 75.13% variation with 10 polymorphic SSR markers. Morpho-agronomic traits should be included in diversity studies. Given the significant genetic variation discovered between clusters 1 and 3, genotypes from these two clusters alone would be excellent candidates for genetic advancement in breeding programmes. Mia *et al.* (2022) found local landraces yield higher than modern cultivars, with plant height and grain weight being significant contributors. Clustering analysis suggests potential for future improvement programmes.

These studies collectively contribute to the understanding of rice genetic diversity, trait associations and clustering patterns. They provide valuable insights for rice breeders and researchers working towards developing high-yield rice varieties with superior grain quality, potentially enhancing global food security.

REFERENCES

- Amegan E, Andrew E, Malachy A, Afeez S and Fiot T. 2020. Genetic diversity of Korean rice (*Oryza sativa* L.) germplasm for yield and yield related traits for adoption in rice farming system in Nigeria. *International Journal of Genetics and Genomics* 8(1): 19–28.
- Anik T R, Islam M A, Uddin M I, Islam M M, Rashid M, Hossain M M, Razia S and Haque M S. 2021. Screening and molecular analysis of some rice (*Oryza sativa* L.) genotypes for drought tolerance at seedling stage. *JAPS: Journal of Animal and Plant Sciences* 31(6).
- Bagudam R, Eswari K B, Badri J and Rao P R. 2018. Correlation and path analysis for yield and its component traits in npt core set of rice (*Oryza sativa* L.). *International Journal of Current Microbiology and Applied Science* 7(9): 97–108.
- Bhargava K, Shivani D, Pushpavalli S N C V L, Sundaram R M, Beulah P and Senguttuvel P. 2021. Genetic variability, correlation and path coefficient analysis in segregating population of rice. *Electronic Journal of Plant Breeding* 12(2): 549–55.
- Deep R, Saha S, Loitongbam B, Singh S and Singh P K. 2017. Genetic variability study for yield and yield components in rice (*Oryza sativa* L.). *International Journal of Agriculture, Environment and Biotechnology* 10(2): 171–76.
- Ghimire N H and Mahat P M. 2019. Genetic variability and correlation coefficients of major traits in cold tolerance rice (*oriza sativa* L.) under mountain environment of Nepal.

- International Journal of Applied Sciences and Biotechnology* 7(4): 445–52.
- Harsha D I, S Kumar and M Talha. 2017. Assessment of genetic variability and inter-character association studies in rice genotypes (*Oryza sativa* L.). *International Journal of Current Microbiology and Applied Science* 6: 2041–46.
- Lakshmi L, Rao M V B, Raju C S and Reddy S N. 2017. Variability, correlation and path analysis in advanced generation of aromatic rice. *International Journal of Current Microbiology and Applied Science* 6(7): 1798–1806.
- Maurya R K, Dwivedi D K, Khan N A, Giri S P and Dixit S. 2022. Genetic variability studies for qualitative and quantitative traits in rice (*Oryza sativa* L.). *The Pharma Innovation Journal* 11(5): 1140–43.
- Mehmood S, Uddin I, Ullah I, Mohamed H I, Basit A, Khan M N, Hussain Shah S S and Rehman U A. 2021. Agro-morphological and genetic diversity studies in rice (*Oryza sativa* L.) germplasm using microsatellite markers. *Molecular Biology Reports* 48: 7179–92.
- Mia S, Ahmed N U, Islam M Z, Rashad M M I, Islam M M and Zaman A M. 2022. Genetic diversity and yield performance among T. Aman rice (*Oryza sativa* L.) landraces in Barishal region of Bangladesh. *Journal of Crop Science and Biotechnology* 1–10.
- Mohanty S, Mohanty N, Sulakshana S, Pradhan S K, Dash S K and Behera L. 2017. Assessment of polymorphism at molecular level, association studies, multivariate analysis and genetic diversity among recombinant inbred lines of rice (*Oryza sativa* L.). *ORYZA-An International Journal on Rice* 54(2): 174–85.
- Mohanty S, Kumar J, Mohanty N, Gouda G, Pradhan S K, Dash S K and L Behera. 2018. Assessment of genetic variability in recombinant inbred lines of rice (*Oryza sativa* L.) using phenotypic traits and molecular markers. *New Agriculturist* 29(1): 1–12.
- Ogunbayo A, Ojo D K, Sanni K A, Tayo S, Afeez S, Akinwale M G, Adetoro N A, Adetumbi J A, Amusa O D and Adekoya M A. 2022. Assessment of genetic diversity among intra and interspecific lowland rice using morpho-agronomic traits and SSR markers. *Journal of Genetic Engineering and Biotechnology Research* 4(1): 140–64.
- Pachauri A K, Sarawgi A K, Bhandarkar S and Ojha G C. 2017. Agro-morphological characterization and morphological based genetic diversity analysis of rice (*Oryza sativa* L.) germplasm. *Journal of Pharmacognosy and Phytochemistry* 6(6): 75–80.
- Prajapati M R, Bala M, Patel V P, Patel R K, Sushmitha U S, Kyada A D and Patel D P. 2022. Analysis of genetic variability and correlation for yield and its attributing traits in F₂ population of rice (*Oryza sativa* L.). *Electronic Journal of Plant Breeding* 13(3): 983–90.
- Prasad S K, Chakraborty M, Singh D N, Kumari P and Ekka J P. 2020. Genetic variability, correlation and path coefficient studies in F₂ generation of rice (*Oryza sativa* L.). *International Journal of Chemical Studies* 8(4): 3116–20.
- Saha S R, Ferdousi A, Hassan L, Haque A, Begum S N, Yasmin F and Akram W. 2022. Rice landraces from haor areas of Bangladesh possess greater genetic diversity as revealed by morpho-molecular approaches along with grain quality traits. *Cogent Food and Agriculture* 8(1): 2075130.
- Singh S K, Habde S, Singh D K, Khaire A, Mounika K and Majhi P K. 2020. Studies on character association and path analysis studies for yield, grain quality and nutritional traits in F₂ population of rice (*Oryza sativa* L.). *Electronic Journal of Plant Breeding* 11(3): 969–75.
- Thakur K and Sarma M K. 2023. Genetic diversity and principal component analysis in cultivated rice (*Oryza sativa*) varieties of Assam. *The Indian Journal of Agricultural Sciences* 93(2): 145–50.
- Tiwari D N, Tripathi S R, Tripathi M P, Khatri N and Bastola B R. 2019. Genetic variability and correlation coefficients of major traits in early maturing rice under rainfed lowland environments of Nepal. *Advances in Agriculture* 1: 1–9.
- Zhang Yu, Qiao He, Zhou Xi, Zheng S, Wang Ye, Peiji Li and Wang Yue. 2022. Genetic diversity and population structure of 93 rice cultivars (lines) (*Oryza sativa* Xian group) in Qinba in China by 3 types of genetic markers. *BMC Genomics* 23(1): 550.