GMR
The Original

ASSOCIATION MAPPING OF KEY GRAIN QUALITY AND NUTRITIONAL TRAITS FOR RICE IMPROVEMENT

***Mangla Parikh, Ritu R Saxena, Bhawana Sharma, Parmeshwar Sahu, Omnarayan Verma and Ajay Tiwari**

Department of Genetics and Plant Breeding, College of Agriculture, IGKV, Raipur Chhattisgarh 492012 India
***Corresponding author**: Mangla Parikh, manglaparikh5@gmail.com

ABSTRACT
A Genome-Wide Association Studies (GWAS) was conducted using 192 diverse rice accessions to understand genetic relatedness, diversity and identify genomic regions associated with grain quality and nutritional traits by using 130 molecular markers. The mixed linear model (MLM) method of the TASSEL software was used to analyze the associations between markers and traits. A total of 715 alleles were amplified by 130 pairs of SSR primers. The average number of observed alleles, heterozygosity (H), and the polymorphism information content (PIC) per marker were 5.5, 0.55 and 0.52, respectively. One ninety-two rice accessions were divided into two groups by population structure followed by neighbour joining tree. A Genome-Wide Association Studies (GWAS) revealed 78 significant QTLs through marker trait association (MTAs) for 16 traits. Out of 78 QTLs, five have been reported earlier while rest are novel. Interestingly, 64 QTLs explained more than 10% phenotypic variance. In addition, association of RM 133 with hulling percentage, milling percentage, head rice recovery and Iron content are explained the possible occurrence of pleiotropism among these traits, similarly, marker Rc showed pleiotropic relationships among gelatinization temperature, Phenolics content and flavonoids content. The present study will open many avenues towards utilization of these QTLs in rice breeding for developing nutrition-rich high yielding varieties.

KEYWORDS: Rice, germplasm, SSR markers, genetic diversity, association analysis

INTRODUCTION

Rice is the predominant source of energy, protein and micronutrients but does not provide enough essential minerals nutrients such as Iron (Fe), Manganese (Mn), Zinc (Zn), (Bouis and Welch, 2010), Protein (Bryant *et al*, 2013) and Oil content. The range of nutrients composition of brown rice was reported as Fe (2 – 35 ppm), Zn (10 – 33 ppm), Mn (0.5 – 2.0 mg) (http://www.whfoods.com), Protein (6 – 8%) and Oil (2 – 4%). Increase in health consciousness and literacy percentage in recent years, people have greater awareness about the health and nutritional benefits of brown rice or pigmented rice. Several studies showed that brown and pigmented rice is rich in protein, vitamins, minerals and also phytochemicals such as phenolics, flavonoids, gamma oryzanols, tocopherols etc. which are essential for healthy life. (Chaudhary *et al*., 2018). The potential health benefits have support to cardiovascular, reproductive and nervous system and have antioxidative effects, anticancer, antidiabetic, anti – obesity, antimicrobial and neuroprotection (Khoo *et al*., 2017). Thus, prioritization is given to improve the concentration of these nutritional compounds to enrich the staple food. The rice biofortification can be use as tool to tackle hidden hunger by ensuring the balanced nutrients in adequate and affordable amounts in major staple food. The traditional landrace showed wide variation and found to have 3 to 4-fold higher contents for Vitamins, Minerals and protein (Pathak *et al.,*2017). Therefore, large opportunities exists to exploit this germplasm for nutritional rice breeding programme (Descalsota-Empleo *et al*., 2019). The grain nutritional quality and phytochemical traits in rice are highly complex and are governed by major and minor genes/quantitative trait loci. QTL detection is an important tool for identification of favourable alleles in population. The identification of quantitative trait loci is largely achieved through mapping. The identification and mapping of QTLs/genes are essential for exploiting the unadapted genotypes. Still there exists untapped genetic variation in germplasm, which were not utilized during evolution and domestication, can be exploit by mining desirable alleles. Traditional rice often possess beneficial QTLs or genes, could be exploited by developing nutritious rice varieties through genome wide mapping approach (Descalsota-Empleo *et al*., 2019).

Molecular breeding gives a new dimension over conventional breeding especially in dissecting the genetic mechanism behind important economic traits (Jonah *et al*., 2011). Molecular markers are very effective tools in detecting the allelic variation present in crop species as well as in furnishing the information about their evolutionary background. This technology facilitates to identify the chromosome location controlling quantitative traits known as QTL. Identification of such genomic region increase the efficiently and effectiveness of breeding programme (Kordrostami and Rahimi, 2015). SSR are most widely used markers because they are highly informative, codominant and multi – allelic in nature and are efficiently used in studies of population structure, genetic mapping, evolutionary biology, genetic diversity studies and identification of

new QTLs (Mason, 2015; Becerra, 2015). Association mapping is a powerful technique for unravelling the molecular basis for phenotypic diversity, using molecular markers in unstructured or loosely structured population. It led to the identification of new alleles/genes/QTLs of complex in natural population (Yang *et al*., 2025). Rice is amenable to GWAS largely because of its evolutionary and domestication history (McCouch *et al*., 2016). QTL identification is performed by examining the association of markers with the trait through linkage disequilibrium between marker and polymorphisms across a set of diverse germplasm (Zhu *et al.*, 2008). It also allows the identification of haplotypes representing different alleles of gene (Gupta *et al*., 2005). The another approach for QTL identification is bi-parental population method, but the limitation with this approach is that it restricts the transferability of the QTLs to other genetic backgrounds. This limitation may be resolved by adopting modified form of linkage disequilibrium mapping, known as genome-wide association analysis (GWAS). The existence of huge allelic diversity for grain quality and nutritional traits in rice, GWAS could be deployed as the most promising strategy of QTL mapping for various grain quality and nutritional traits with high precision (Huang and Han, 2014). Thus, this holistic approach is very effective in identification of causative alleles for grain nutritional traits with modest number of markers distributed over all the chromosomes.

The genome-wide association analysis was performed with diverse association panel evaluated for grain quality and nutritional traits in rice gene pool. Selected traits are considered most important in bio-fortification programme to develop nutri-rich rice varieties to combat mineral malnutrition. The main objective of the research study was to explore potential QTL/associated markers with traits of interest and exploiting them in bio-fortification breeding programme. The results of this study have significant scope in developing strategic marker aided breeding for introgression of grain quality and nutritional traits in rice.

MATERIALS AND METHODS

Germplasm and experimental details:

The research materials comprised of a panel of 192 diverse rice genotypes (Table 1) including 186 test genotypes and 6 improved check varieties viz., Aruna (red rice variety of Kerala), Krishnanjana (red rice variety of Kerala), Dubraj Selection -1 (Aromatic rice variety), Badshahbhog Selection-1 (Aromatic rice variety), Jaldubi (Submergence tolerant rice variety) and IGKV R-1 (Rajeshwari) (High yielding rice variety) (supplementary Table 1). All these accessions were procured from R.H. Richharia Rice Gene Bank, Department of Genetics and Plant Breeding, Indira Gandhi Krishi Vishwavidyalaya (IGKV), Raipur-492012 (Chhattisgarh) for the study.

The experiment was conducted at Research Farm, IGKV, Raipur, during wet (*Kharif)* season 2018 by following the Augmented Block Design (Federer 1956). Six improved checks were randomized within each block. All the genotypes were grown in six blocks (31 test genotypes and six check varieties). Each genotype was transplanted into four rows with spacing of 0.2m row-row and 0.15m plant-plant. Recommended package of practices were followed to raise the crop. After harvesting, all the genotypes were properly cleaned, dried under shade till the moisture content decreased to the recommended 12-14% and were stored at room temperature for four months prior to the grain quality and nutritional analysis.

Phenotyping for grain quality and nutritional traits

Ten grain quality parameters *viz.,* kernel length (KL), kernel breadth (KB), kernel length breadth ratio (KLB), elongation ratio (ER), hulling % (HULL %), milling % (MILL %), head rice recovery % (HRR %), alkali spreading value /gelatinization temperature (GT), gel consistency (GC) and amylose content (AC) were recorded as per standard protocols of Indian Institute of Rice Research (IIRR, Hyderabad, 2014). Six nutritional traits (micronutrients and phytochemicals) were also recorded in this study. Protein content (Lowry method, 1951) was estimated in brown rice powder. Iron (Fe) and zinc (Zn) contents were estimated in brown rice using energy dispersive X-ray fluorescence (ED-XRF) (OXFORD Instruments X-Supreme 8000, Abingdon, UK) (Rao *et al*., 2014). Rough rice samples were hulled by a single pass through standard rubber roll huller (Satake Engineering Co. Ltd. Tokyo, Japan) to eliminate the mineral contamination from huller machine. Similarly, milling of hulled rice was done in a standard miller to extract the bran fraction (Satake Engineering Co. Ltd. Tokyo, Japan). Three phytochemicals viz, Total phenolic content equivalent to µg of gallic acid (Bao *et al*., 2005), flavonoid content equivalent to µg of quercetin (Jia *et al*., 1999) and γ-oryzanol content (Bucci *et al*., 2003) were estimated from rice bran. Rice bran of each genotype was packed into zipper polythene, labelled properly and kept in 4^0C to avoid the harmful activities of lipase enzyme.

Bio statistical Analysis:

The analysis for augmented design was performed by the online software available at IASRI site, as per method given by Federer (1956). Descriptive statistics was analysed by software XLSTAT v 19.03. Phenotypic correlations among studied traits were estimated with the help of PAST v 3.14 software (Hammer *et al*., 2001). Mean data (adjusted mean) of all the traits were subjected to bio-statistical analysis and association mapping study.

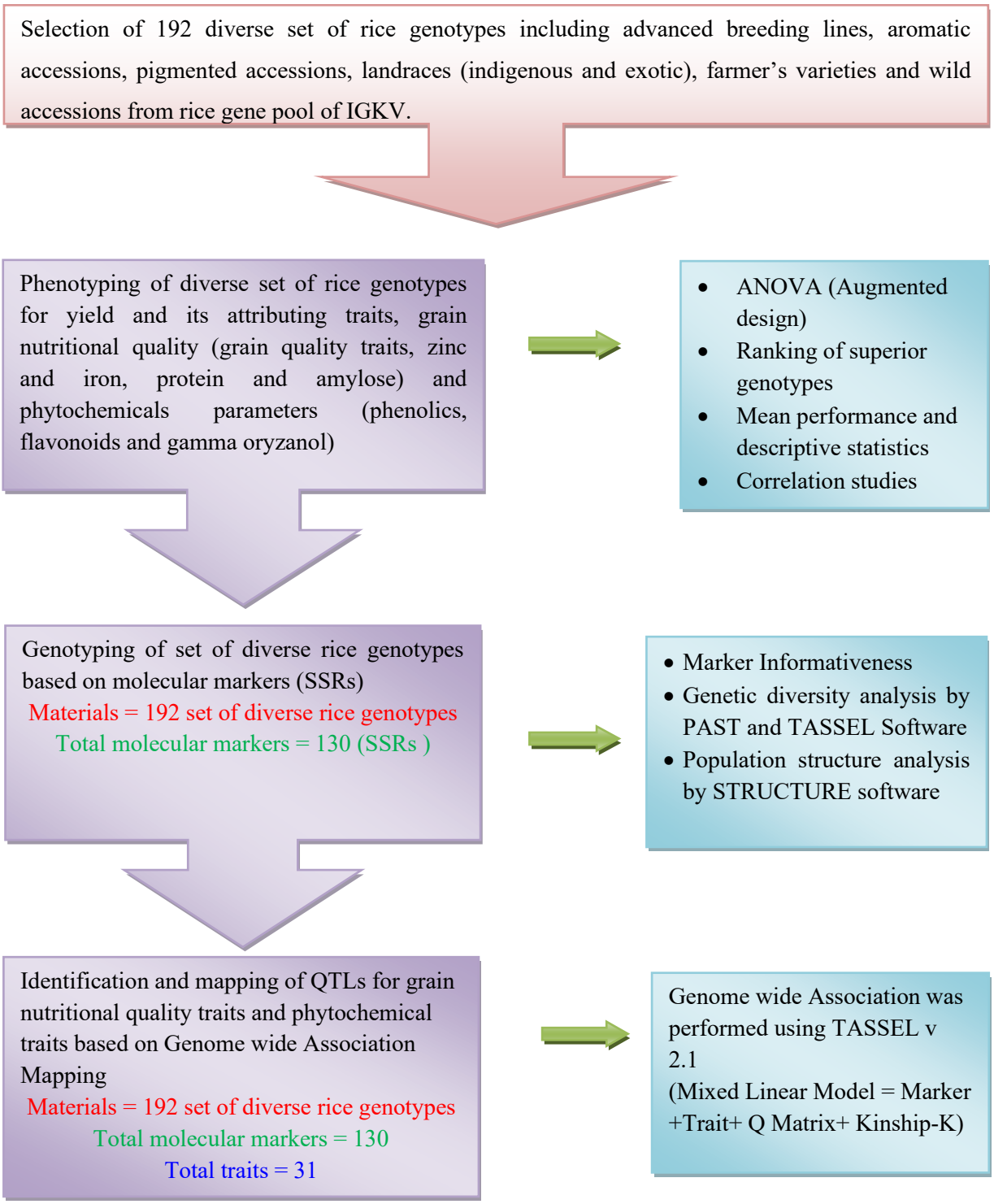


Fig. 1 Flow chart showing stepwise research activities

Genomic DNA isolation, PCR Analysis and Selection of SSR Markers

The genomic DNA was isolated from freshly collected leaves of four-week-old plants of each rice genotypes following CTAB protocol (Zheng *et al*., 1995). A total of 130 SSR markers were taken from Gramene marker database (http://www.gramene.org/markers/microstat/) available in the public domain and used for genotyping of the panel population. The DNA fragments were resolved in gel electrophoresis for quantification of the isolated DNA. The PCR amplification was carried out using the markers selected based on position covering the whole genome to illustrate the diversity and to identify the polymorphic loci among the 196 rice accessions. The PCR reactions were set to initial denaturation step (94ºC ,4 min), followed by 35cycles of denaturation (94ºC, 45 sec), annealing/extension (55ºC, 30 sec.), extension (72ºC, 45 sec.), final extension (72ºC ,7 min.) and store at 4ºC (infinity). PCR products of SSR markers were electrophoreses using 5% acrylamide gel solution (65 ml) containing ethidium bromide and 50 bp DNA ladder was used to determine the size of amplicon. The gel was run for 40-50 min. at 120-160Vcm-1 and photographed using Gel Documentation System (Syngene).

Marker data analysis and genetic diversity indices

The scoring of PCR-amplified DNA fragments was carried out based on band size (bp) as per Power Marker software format. Different indices of diversity viz., major allele frequency, number of alleles per locus, gene diversity, heterozygosity and polymorphic information content (PIC) value for each marker were estimated by unbiased estimator using POWERMERKER version 3.25 (Liu and Muse, 2005).

Analysis of Population structure and genetic diversity

Population structure was analysed in 192 diverse set of rice genotypes based on 130 microsatellite markers by following Bayesian statistical model in programme STRUCTURE version 2.3.4 (Pritchard *et al.*, 2000). STRUCTURE software was run with K-values varying 1 to 10, with 10 iteration for each K value to derive the ideal number of groups. A high throughput parameter set of burn-in period of the 100,000 steps with model set of "possibility of admixture and allele frequency correlated" followed by 100,000 Monte Carlo Markov Chain replicates (MCMC) during the running period. The result generated from structure analysis was used to identify optimum K value based on LnP(D) and Evanno's K method in STRUCTURE HARVESTER (Earl, 2012). The accessions with the probability ≥ 0.80 scores (inferred ancestry) were assigned as pure gene pool and less than 0.08 probability were adjuged as "admixture" group. Population differentiation was assessed by calculating pairwise Fst values and Nei's genetic distances for different population pairs. Genetic distance and phylogenetic analyses of rice accessions were conducted based on Rho's coefficient similarity matrix with the help of PAST v 3.15 (Hammer *et al.,* 2001), obtained from the proportion of shared SSR fragments. Dendogram and Unrooted neighbour joining (NJ) tree based on UPGMA algorithm were constructed by SAS v9.4 software (SAS, Inc., Cary, NC, USA).

Association Mapping for grain quality and nutritional traits

Genome-wide association mapping was performed using the TASSEL v3.0 software (Bradbury *et al*., 2007) by following the mixed linear model (MLM) using 50,000 time permutations for correlation of multiple testing (Pritchard *et al*., 2000). MLM analysis was preferred with molecular data, phenotypic data, population inferred ancestry (Q matrix) and kinship (K) matrix (Yu *et al*., 2006). Significant marker-trait associations were identified by considering the Bonferroni threshold and it was calculated by standard method in MSExcel (Moran, 2003). Bonferroni threshold level for this study was 1/130= 0.0076, where 130 is the total markers used for association tests for each trait in this study. Therefore, markers with the probability p value 0.0076 of less than this were considered as significant.

RESULTS

Genetic diversity indices and informativeness of the molecular markers

All the 192 diverse rice panel were genotyped using 130 microsatellite markers (SSR) for measuring allelic diversity. A total 715 allele (bands) were amplified for all the SSR markers. The number of alleles per loci varied from 1to 14 with an average 5.5 alleles per locus. The maximum number of alleles were detected for marker loci RM 334 (14) and the minimum was detected number of allele (1) was detected for a group of markers viz., RM 338, RM2578, RM14473, RM4796, RM24547 and HvSSR06-50. Polymorphic information content (PIC) value represents the relative informativeness of each marker and in the present study, the average PIC value was found to be 0.52. PIC values ranged between 0.020 for RM21889 and RM18 to 0.87 for RM334. Expected heterozygosity or Gene diversity index (H_e) estimated according to Nei (1973) varied from 0 to 0.88 (RM 334) with an average value of 0.55. The results indicated the presence of high allelic diversity in the genotypes studied .The major allele frequency varied from 0.199 (RM 334) to 1 with an average of 0.56 (Table 1).

Table 1: List of markers (SSRs, HvSSRs and InDels) along with their map positions and genetic diversity parmeters

S. No.	Marker name	Chromo-some No.	Map position (cM/bp)	Major allele frequency	Allele No.	Gene diversity (He)	PIC	Marker type
1	RM 495	1	2.8	0.333	5	0.729	0.680	SSR
2	RM 1	1	29.7	0.426	12	0.782	0.768	SSR
3	RM 283	1	31.4	0.524	7	0.662	0.625	SSR
4	RM 259	1	54.2	0.246	8	0.825	0.801	SSR
5	RM 312	1	71.6	0.705	4	0.440	0.378	SSR
6	RM 5	1	94.9	0.251	8	0.835	0.814	SSR
7	RM 237	1	115.2	0.241	7	0.830	0.807	SSR
8	RM 431	1	178.3	0.432	5	0.685	0.628	SSR
9	RM1360	1	30.5	0.455	4	0.628	0.552	SSR
10	RM580	1	68.2	0.534	7	0.543	0.445	SSR
11	RM10864	1	14235715	0.307	8	0.780	0.747	SSR
12	RM11162	1	21384304	0.197	9	0.863	0.848	SSR

13	RM11488	1	27887991	0.663	3	0.481	0.411	SSR
14	RM11694	1	32224855bp	0.745	5	0.427	0.407	SSR
15	RM11928	1	37910904bp	0.853	4	0.260	0.242	SSR
16	RM12119	1	40779925bp	0.599	4	0.577	0.528	SSR
17	RM 154	2	4.8	0.353	6	0.737	0.692	SSR
18	RM 452	2	58.4	0.531	4	0.580	0.499	SSR
19	RM 53	2	32.7	0.400	7	0.748	0.714	SSR
20	RM 29	2	68.9	0.948	3	0.100	0.098	SSR
21	RM 279	2	17.3	0.382	7	0.751	0.715	SSR
22	RM6842	2	124679bp	0.250	8	0.830	0.808	SSR
23	RM12720	2	6580532bp	0.677	4	0.503	0.467	SSR
24	RM12975	2	10289242bp	0.693	5	0.485	0.450	SSR
25	RM3858	2	19182948bp	0.339	6	0.761	0.724	SSR
26	RM2578	2	32250183bp	1.000	1	0.000	0.000	SSR
27	RM208	2	186.4	0.292	8	0.812	0.787	SSR
28	RM 489	3	29.2	0.461	9	0.605	0.527	SSR
29	OSR 13	3	53.1	0.277	8	0.821	0.799	SSR
30	RM 338	3	108.4	1.000	1	0.000	0.000	SSR
31	RM 514	3	216.4	0.307	9	0.803	0.776	SSR
32	RM 6832	3	88.95	0.272	10	0.817	0.793	SSR
33	RM 3562	3	11683187bp	0.729	4	0.438	0.406	SSR
34	RM14307	3	1019485bp	0.259	9	0.813	0.787	SSR
35	RM14437	3	3914049bp	1.000	1	0.000	0.000	SSR
36	RM14651	3	7221114bp	0.479	5	0.588	0.504	SSR
37	HvSSR03-09	3	3797148bp	0.349	8	0.754	0.715	HvSSR
38	RM15026	3	14459750bp	0.719	3	0.443	0.401	SSR
39	HvSSR03-40	3	15505710bp	0.534	5	0.613	0.553	HvSSR
40	RM15446	3	23080448bp	0.302	7	0.772	0.738	SSR
41	RM4796	3	27435768bp	1.000	1	0.000	0.000	SSR
42	RM15722	3	27658912bp	0.380	5	0.722	0.673	SSR
43	RM168	3	171.2	0.318	6	0.764	0.728	SSR
44	INDEL2	3	-	0.516	7	0.655	0.610	INDEL
45	INDEL8	3	-	0.854	2	0.249	0.218	INDEL
46	RM15767	3	28431374bp	0.990	2	0.021	0.020	SSR
47	RM411	3	127.9	0.458	3	0.622	0.543	SSR
48	RM 307	4	0	0.484	7	0.661	0.609	SSR
49	RM 124	4	150.1	0.599	4	0.527	0.444	SSR
50	Ra	4	-	0.832	4	0.295	0.279	SSR
51	RM16616	4	12023381bp	0.510	4	0.571	0.482	SSR
52	RM16738	4	16410449	0.579	7	0.628	0.603	SSR
53	HvSSR04-32	4	22674757bp	0.442	7	0.724	0.690	HvSSR
54	RM5506	4	33273406bp	0.286	10	0.840	0.823	SSR
55	RM 507	5	0	0.864	4	0.242	0.225	SSR
56	RM 413	5	26.7	0.416	9	0.752	0.723	SSR
57	RM 161	5	96.9	0.864	3	0.239	0.217	SSR
58	RM 178	5	118.8	0.542	4	0.575	0.495	SSR
59	RM 334	5	141.8	0.199	14	0.882	0.871	SSR
60	RM 87	5	129.2	0.529	5	0.647	0.602	SSR
61	RM17990	5	4664936bp	0.677	3	0.474	0.411	SSR
62	HvSSR05-24	5	8283540bp	0.542	3	0.559	0.471	HvSSR
63	RM18600	5	18997773bp	0.356	7	0.771	0.738	SSR
64	RM18908	5	24015384bp	0.682	4	0.497	0.461	SSR
65	RM18	5	90.4	0.250	8	0.814	0.788	SSR
66	RM 133	6	0	0.573	3	0.502	0.391	SSR
67	RM 510	6	20.8	0.859	3	0.249	0.231	SSR
68	RM 454	6	99.3	0.995	2	0.010	0.010	SSR

69	RM 162	6	108.3	0.896	5	0.195	0.190	SSR
70	RM19255	6	534728bp	0.714	3	0.415	0.337	SSR
71	RM19454	6	3945764bp	0.547	6	0.650	0.617	SSR
72	RM20066	6	16475032bp	0.440	5	0.678	0.623	SSR
73	HvSSR06-50	6	23162026bp	1.000	1	0.000	0.000	HvSSR
74	RM5509	6	27828305bp	0.479	9	0.732	0.713	SSR
75	RM20730	6	30786736bp	0.234	9	0.835	0.814	SSR
76	RM30	6	125.4	0.755	4	0.377	0.317	SSR
77	RM 125	7	24.8	0.356	8	0.777	0.747	SSR
78	RM 11	7	47	0.411	6	0.740	0.704	SSR
79	RM 455	7	65.7	0.974	3	0.051	0.051	SSR
80	RM 118	7	96.9	0.526	5	0.616	0.551	SSR
81	RM 418	7	42.1	0.225	7	0.824	0.801	SSR
82	Rc	7	-	0.511	4	0.614	0.543	SSR
83	RM20895	7	1559317bp	0.911	4	0.164	0.156	SSR
84	RM21182	7	5768137bp	0.365	5	0.726	0.680	SSR
85	HvSSR07-40	7	21870543bp	0.365	5	0.759	0.723	HvSSR
86	RM21889	7	23826962bp	0.990	2	0.021	0.020	SSR
87	RM8044	7	24194491bp	0.318	8	0.812	0.789	SSR
88	RM22165	7	29406921bp	0.651	5	0.484	0.406	SSR
89	RM214	7	34.7	0.724	6	0.451	0.425	SSR
90	RM 408	8	0	0.464	6	0.671	0.615	SSR
91	RM 152	8	9.4	0.557	6	0.643	0.613	SSR
92	RM 44	8	60.9	0.714	6	0.462	0.434	SSR
93	RM 284	8	83.7	0.529	7	0.635	0.584	SSR
94	RM 433	8	116	0.386	7	0.715	0.668	SSR
95	RM 447	8	124.6	0.272	11	0.803	0.776	SSR
96	RM 339	8	72.2	0.854	5	0.264	0.254	SSR
97	RM22579	8	5967354bp	0.442	8	0.723	0.687	SSR
98	RM8271	8	7611709bp	0.406	6	0.740	0.704	SSR
99	RM22994	8	17501335bp	0.714	4	0.455	0.418	SSR
100	HvSSR08-37	8	18233120bp	0.708	5	0.473	0.446	HvSSR
101	RM3155	8	27772501bp	0.740	4	0.428	0.400	SSR
102	RM 316	9	1.8	0.474	8	0.673	0.623	SSR
103	RM 105	9	32.1	0.450	5	0.697	0.652	SSR
104	RM 215	9	99.4	0.547	6	0.636	0.595	SSR
105	RM 326	9	1.9	0.807	3	0.327	0.300	SSR
106	RM23676	9	780072bp	0.469	5	0.689	0.644	SSR
107	RM24353	9	15505057bp	0.628	5	0.569	0.540	SSR
108	RM24547	9	18312692bp	1.000	1	0.000	0.000	SSR
109	RM2705	9	22482806bp	0.573	4	0.569	0.499	SSR
110	RM 271	10	59.4	0.392	8	0.758	0.726	SSR
111	RM 171	10	73	0.323	4	0.719	0.665	SSR
112	RM 484	10	97.3	0.865	4	0.238	0.216	SSR
113	RM 228	10	87.6	0.663	6	0.533	0.508	SSR
114	RM24873	10	272650bp	0.885	4	0.211	0.203	SSR
115	RM25317	10	12429023bp	0.708	3	0.438	0.378	SSR
116	RM25486	10	15905373bp	0.630	7	0.567	0.537	SSR
117	RM 552	11	40.6	0.589	7	0.614	0.585	SSR
118	RM 536	11	55.1	0.474	6	0.645	0.582	SSR
119	RM 287	11	68.6	0.375	8	0.789	0.765	SSR
120	RM25993	11	780582bp	0.495	3	0.510	0.390	SSR
121	HvSSR11-13	11	5919351bp	0.339	7	0.809	0.787	HvSSR
122	RM26416	11	9233362bp	0.385	5	0.727	0.680	SSR
123	RM26781	11	17544779bp	0.700	5	0.463	0.417	SSR
124	RM27089	11	23733178bp	0.745	5	0.423	0.397	SSR

125	RM21	11	85.7	0.432	6	0.716	0.675	SSR
126	RM 19	12	20.9	0.245	8	0.832	0.811	SSR
127	RM 277	12	57.2	0.854	3	0.256	0.236	SSR
128	RM27428	12	598151bp	0.771	5	0.383	0.356	SSR
129	RM7601	12	3553768bp	0.693	5	0.470	0.422	SSR
130	RM2734	12	26153853bp	0.422	5	0.733	0.696	SSR

Microsatellite marker based genetic diversity analysis

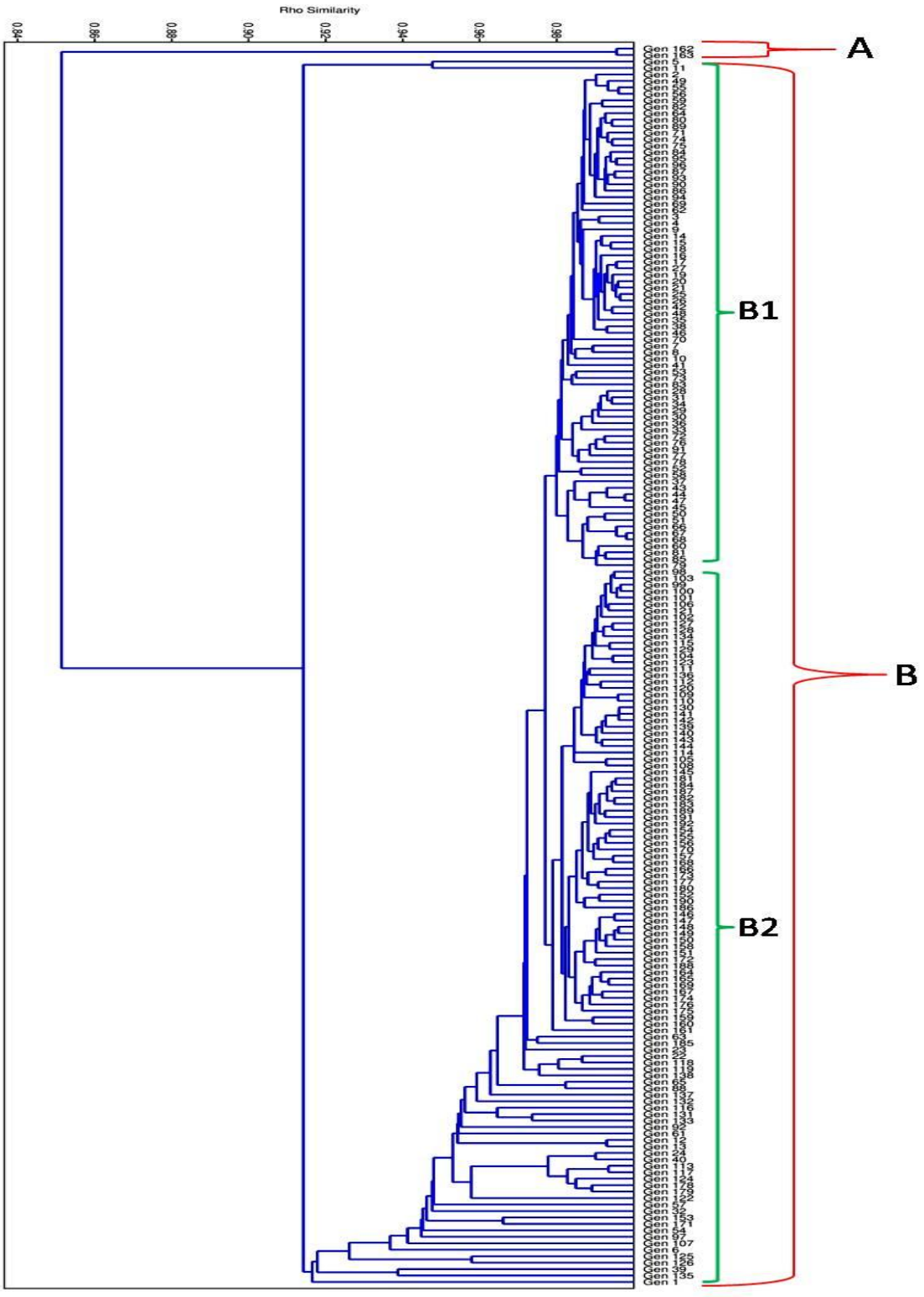
The present investigation was undertaken to assess the level of relatedness in the study material comprised of exotic and indigenous germplasm, wild rice accessions and improved genotypes and to better understand the association between genetic diversity and geographic factors. The cluster dendogram showed that the Rho similarity coefficient matrix varied from 0.85 to 0.99. Although the genetic diversity was observed to be low, SSRs proved to be an efficient marker in estimating the genetic diversity of rice genotypes.

The dendogram grouped the 192 rice genotypes into two major groups A and B at 85% similarity level. Group A comprised of only two genotypes *Oryza latifolia* and *Oryza officinalis,* which are wild species of rice, while rest 190 genotypes were grouped in the cluster B, consisting of *Oryza sativa, Oryza nivara* and weedy species *Oryza sativa* var. *fatua* (Table 2). The two clusters were differentiated based on the distinct morphological traits possessed by genotypes of each cluster. The cluster B has further divided into two sub clusters *viz*., B1 and B2, of which cluster B1 have predominantly of *indica* type accession from different geographic origin *i.e.* Indonesia, Thailand, China, Vietnam, India, Philippines and improved genotypes. The inclusion of such genotypes in same cluster indicated the involvement of diverse genotypes in modern breeding programme in developing suitable varieties by crossing genotypes of different origins for important yield and its attributing traits.

Population structure analysis

In the present investigation an attempt was made to define the population structure of 192 diverse set of rice genotypes that comprised of exotic and indigenous landraces, aromatic accessions, red rice accessions, advanced breeding lines, farmers varieties and wild rice accessions (*Oryza nivara, Oryza latifolia* and *Oryza officinalis*). A set of 49 SSR markers was taken from a panel of 50 standard SSR markers utilized by generation challenged programme, for the analysis of diversity in rice. These markers distributed evenly with excellent genome coverage, were used to generate genotypic data on 192 diverse rice genotypes. These markers help to establish the genetic relatedness among the genotypes because they exhibit significant polymorphism across rice genome despite of being less in number (McCouch *et al*., 1997). Model based (Bayesian) cluster software programme STRUCTURE version 2.3.4 was applied to work out the genetic relatedness among the diverse rice genotypes. Simulations were determined through admixture model with the value of K ranging from 1 to 10 with 10 iterations for all 192 genotypes which identified the significant population structure at K = 2 (Fig.3). The comparison for LnP(D) and Evanno's value was made by increasing the value of K from 2 to 10. The highest score for log likelihood was reached at K=2, represented by a sharp peak and value for ΔK was also found highest at K=2 using structure harvester which suggested that the population can be grouped into two sub populations (SG1 and SG2). The bar plot for inferred population structure at K =2 is represented in Fig. 4.

On the basis of membership fraction the genotypes under each sub population were categorize as pure genotypes or admixtures. The genotypes with score ≥ 80% were assigned as pure genotypes and < 0.80 as admixtures (Fig.4). In the present investigation the genetic diversity in the population structure of diverse set of rice genotypes was checked at K = 2 to K = 5. At K = 2, among the 192 genotypes, 96 genotypes were pure and 15 were admix in sub population 1 (having < 80% inferred ancestry value from SG 1 or SG 2), while 69 were pure accessions and 12 were admix in sub population 2. The parameters obtained from the population structure analysis suggested that the population was more clear at K = 2. The inferred ancestry value (Q value) of the genotypes are given in Table 2. Fst statistics or fixation index values worked out to study the genetic variability in sub population.


Rho Similarity
0.84
0.86
0.88
0.90
0.92
0.94
0.96
0.98
A
B1
B
B2

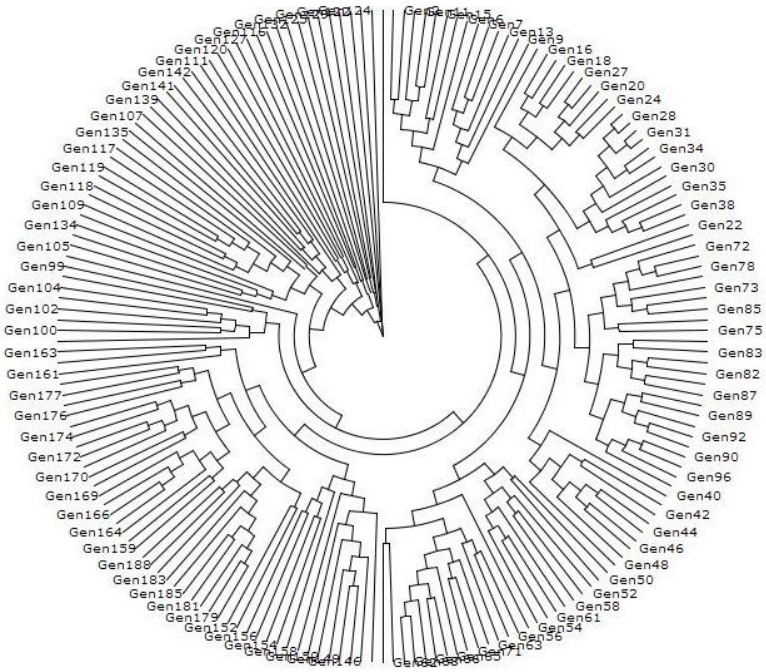


Fig. 2 Dendogram based on UPGMA cluster analysis of 192 diverse rice genotypes using 130 microsatellite markers

The specific mean Fst value for SG-1 and SG-2 was ranged between 0.1270 and 0.1548, respectively with an average of 0.1409 indicating moderate population structure (Table 3). Thus, sub group-2 was said to be more structured population followed by SG-1.

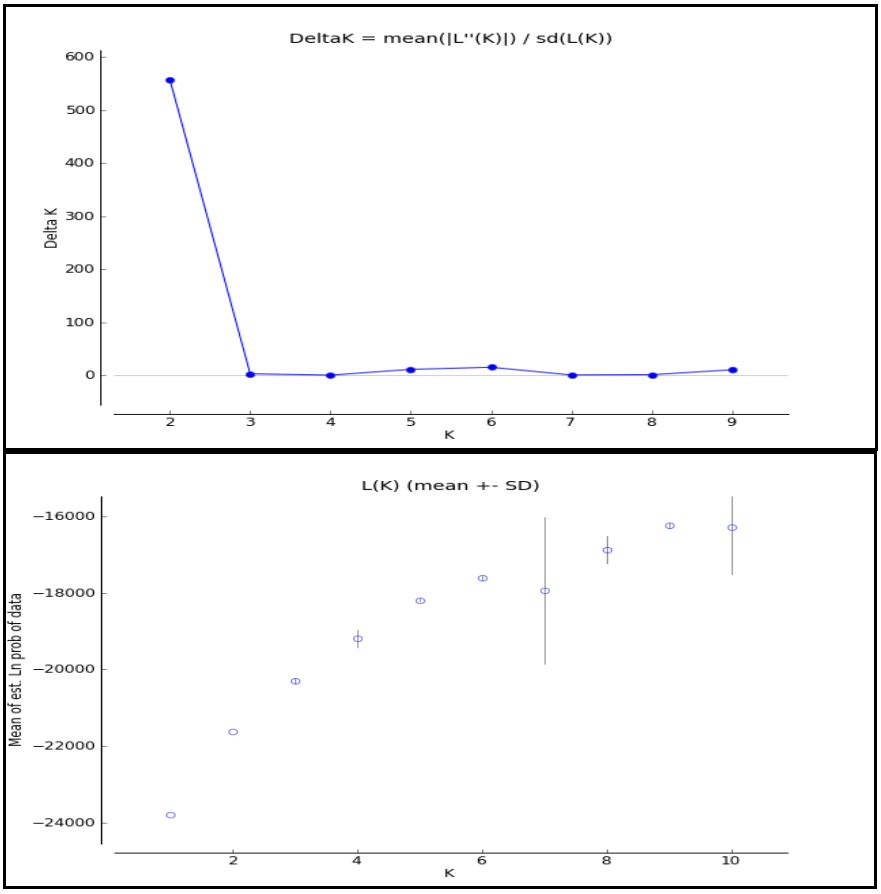
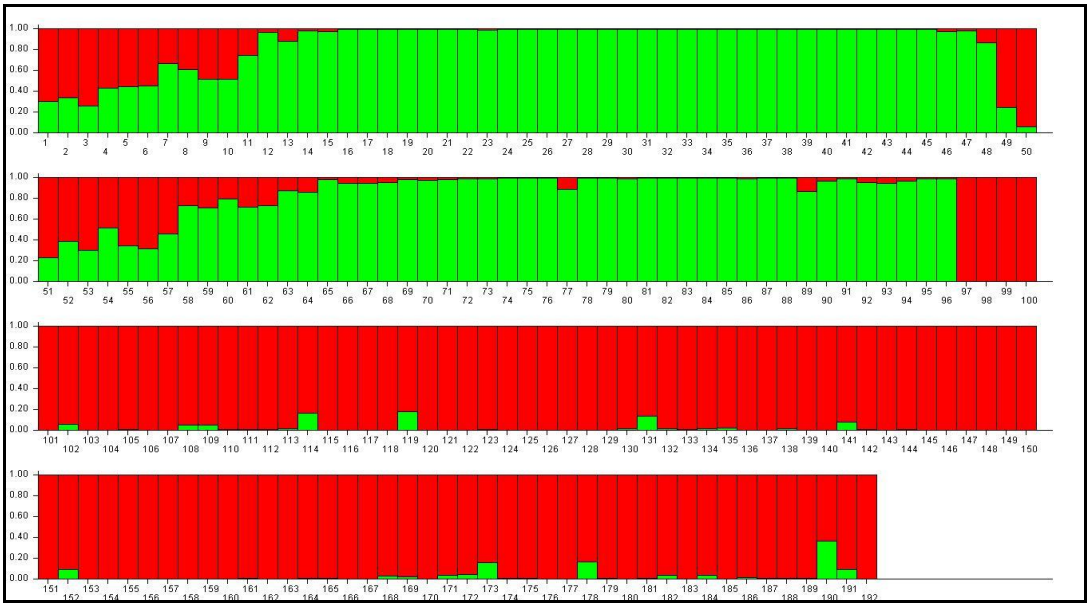
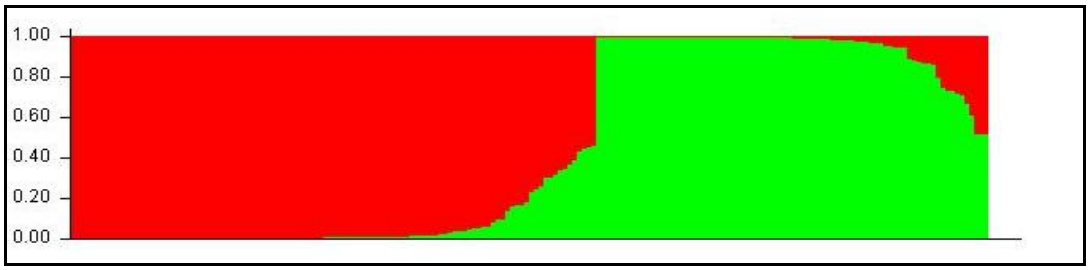


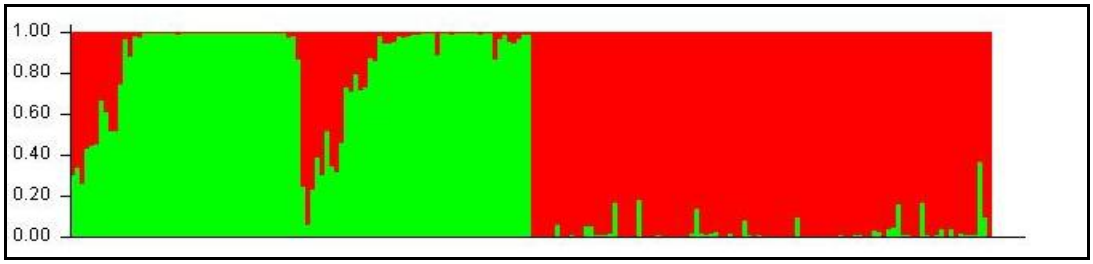
Fig. 3 Estimation of population using LnP(D) derived delta K for determining optimum number of subpopulations.



A. Group wise individual view of population structure of 192 diverse rice genotypes



B. Short by color view of population structure of 192 diverse rice genotypes



C. Single line view of population structure of 192 diverse rice genotypes

Fig. 4 Distribution pattern of 192 rice genotypes based on 130 molecular markers within sub-populations.

Table 2 Population structure subgroup of diverse rice genotypes based on Inferred ancestry values (Q-value)

Genotype number	Name of genotypes	Inferred ancestry		Structure subgroup
		Q1	Q2	
Gen1	Longku labat (Indonesia)	0.693	0.307	Admixture SG 1
Gen2	Hasan Serai (Iran)	0.661	0.339	Admixture SG 1
Gen3	NiiawHawm (Pant nagar)	0.739	0.261	Admixture SG 1
Gen4	Lua Nhe Den (Thailand)	0.569	0.431	Admixture SG 1
Gen5	Hawm Jan (Indonesia)	0.556	0.444	Admixture SG 1
Gen6	Hung-mi-hsiang-ma-Tsan (Vietnam)	0.545	0.455	Admixture SG 1
Gen7	Guinata (Thailand)	0.332	0.668	Admixture SG 2
Gen8	Dawleuang (China)	0.389	0.611	Admixture SG 2
Gen9	Bongcay (Vietnam)	0.48	0.52	Admixture SG 2
Gen10	Binirhen (Phillipines)	0.484	0.516	Admixture SG 2

Gen11	Lalbasmati (Jammu & Kashmir)	0.257	0.743	Admixture SG 2
Gen12	Kalikhasa (Assam)	0.031	0.969	SG 2
Gen13	Kamini Joha (Assam)	0.116	0.884	SG 2
Gen14	BAS 837 (US patent line)	0.017	0.983	SG 2
Gen15	Ayepyaung (Myanmar)	0.027	0.973	SG 2
Gen16	Dubraj (D:1251)	0.006	0.994	SG 2
Gen17	Luchai (L:246)	0.004	0.996	SG 2
Gen18	Mancha (M:1028)	0.004	0.996	SG 2
Gen19	Luchai (L:1099)	0.004	0.996	SG 2
Gen20	Mahipal (M:27A)	0.003	0.997	SG 2
Gen21	Umari (CGR:17451)	0.004	0.996	SG 2
Gen22	Kekai (K:927 II)	0.005	0.995	SG 2
Gen23	BanthaLuchai (B:2733)	0.014	0.986	SG 2
Gen24	Mokdo (M:550)	0.003	0.997	SG 2
Gen25	Shri kamal (S:660 I)	0.004	0.996	SG 2
Gen26	Kankadiya (K:18 II)	0.003	0.997	SG 2
Gen27	Khuddi (K: 1128 IV)	0.003	0.997	SG 2
Gen28	Jira Dhan (J:53)	0.003	0.997	SG 2
Gen29	Javaphool (J:333)	0.003	0.997	SG 2
Gen30	Maharaji (M:504)	0.003	0.997	SG 2
Gen31	Fundri (F:28)	0.002	0.998	SG 2
Gen32	Banspatri (B:728)	0.004	0.996	SG 2
Gen33	Anterved (A:217)	0.003	0.997	SG 2
Gen34	Jaigundi (J:248)	0.004	0.996	SG 2
Gen35	Chinnor (C:151)	0.003	0.997	SG 2
Gen36	Keraghul (K:2034)	0.003	0.997	SG 2
Gen37	R 2053-202-2-145-1	0.003	0.997	SG 2
Gen38	R 1607-321-1-34-1	0.007	0.993	SG 2
Gen39	R 1661-1372-1-601-1	0.004	0.996	SG 2
Gen40	R 1700-2247-1-2313-1	0.007	0.993	SG 2
Gen41	Improved Chepti gurmatiya	0.004	0.996	SG 2
Gen42	Improved Dokra-dokri	0.003	0.997	SG 2
Gen43	R 1779-321-1-112-1	0.005	0.995	SG 2
Gen44	R 2054-685-1-205-1	0.007	0.993	SG 2
Gen45	R 1882-306-4-243-1	0.006	0.994	SG 2
Gen46	R 1973-206-2-86-1	0.026	0.974	SG 2
Gen47	Sonagathi	0.017	0.983	SG 2
Gen48	KodhaPhool (Farmer’s variety)	0.132	0.868	SG 2
Gen49	Balamsar (Farmer’s variety)	0.753	0.247	Admixture SG 1
Gen50	Harikhuta dhan (Farmer’s variety)	0.939	0.061	SG 1
Gen51	Barabali (Farmer’s variety)	0.768	0.232	Admixture SG 1
Gen52	Lahunchi (Farmer’s variety)	0.612	0.388	Admixture SG 1
Gen53	Bagdisona (Farmer’s variety)	0.695	0.305	Admixture SG 1
Gen54	Netakalani (Farmer’s variety)	0.479	0.521	Admixture SG 1
Gen55	Damrubaba-3(Farmer’s variety)	0.651	0.349	Admixture SG 1
Gen56	Nardha (Farmer’s variety)	0.681	0.319	Admixture SG 1
Gen57	Mehar dhan (Farmer’s variety)	0.537	0.463	Admixture SG 1
Gen58	Kata mehar (Farmer’s variety)	0.266	0.734	Admixture SG 2
Gen59	Goindi (Farmer’s variety)	0.287	0.713	Admixture SG 2
Gen60	Luchai (Farmer’s variety)	0.201	0.799	Admixture SG 2
Gen61	Tetirpakhi (Farmer’s variety)	0.285	0.715	Admixture SG 2
Gen62	Hadrasal (Farmer’s variety)	0.268	0.732	Admixture SG 2
Gen63	Dubraj (Red rice)	0.122	0.878	Admixture SG 2
Gen64	Baihaguda (Red rice)	0.139	0.861	Admixture SG 2
Gen65	Bhathaguda (Red rice)	0.015	0.985	SG 2
Gen66	IC: 41843 (Black rice)	0.057	0.943	SG 2
Gen67	IC: 466901 (Black rice)	0.054	0.946	SG 2
Gen68	R-RKM-1 (Black rice)	0.044	0.956	SG 2
Gen69	Ateya (CGR : 2)	0.015	0.985	SG 2
Gen70	Dubraj 11 (CGR : 12230)	0.022	0.978	SG 2

Gen71	Newari (CGR : 4056)	0.017	0.983	SG 2
Gen72	Badshahbhog (CGR : 10919)	0.012	0.988	SG 2
Gen73	Satha Dhan (CGR : 394)	0.014	0.986	SG 2
Gen74	Piso (CGR : 16109)	0.004	0.996	SG 2
Gen75	Newara (CGR : 4053)	0.004	0.996	SG 2
Gen76	Lonkti monchhi(CGR : 16804)	0.004	0.996	SG 2
Gen77	Kanji Local (CGR : 174)	0.109	0.891	SG 2
Gen78	Selection of Badshahbhog (CGR : 17760)	0.005	0.995	SG 2
Gen79	Fara (CGR : 113)	0.006	0.994	SG 2
Gen80	Dondagi (CGR : 12135)	0.009	0.991	SG 2
Gen81	Kadam phool (CGR : 943)	0.005	0.995	SG 2
Gen82	Khatiya pati (CGR : 14230)	0.005	0.995	SG 2
Gen83	Mekara ghol (CGR : 270)	0.006	0.994	SG 2
Gen84	Dokra dokri(CGR : 12126)	0.003	0.997	SG 2
Gen85	Chhatri (CGR : 669)	0.007	0.993	SG 2
Gen86	Sugarkand (CGR : 17213)	0.013	0.987	SG 2
Gen87	Angur guchcha (CGR : 1803)	0.003	0.997	SG 2
Gen88	Parmal (CGR : 15971)	0.006	0.994	SG 2
Gen89	Lakhouwal (CGR : 1128)	0.133	0.867	SG 2
Gen90	Anjan (CGR : 1814)	0.032	0.968	SG 2
Gen91	Anjana (CGR : 1816)	0.008	0.992	SG 2
Gen92	Mani gurmatia(CGR : 2730)	0.044	0.956	SG 2
Gen93	Makado (CGR : 3831)	0.055	0.945	SG 2
Gen94	Moti basmati (CGR : 5817)	0.029	0.971	SG 2
Gen95	Bauwara (CGR : 5854)	0.01	0.99	SG 2
Gen96	Cross 116 (CGR : 6366)	0.009	0.991	SG 2
Gen97	Kotari (CGR : 7805)	0.997	0.003	SG 1
Gen98	Krishna Koliyari (CGR : 7812)	0.993	0.007	SG 1
Gen99	Malpa (CGR : 8409)	0.996	0.004	SG 1
Gen100	Niwari- I (CGR : 8748)	0.996	0.004	SG 1
Gen101	IC: 116076 (CGR : 10007)	0.997	0.003	SG 1
Gen102	Nagbel (CGR : 4003)	0.939	0.061	SG 1
Gen103	Surmatia (CGR : 4909)	0.996	0.004	SG 1
Gen104	IC: 459172	0.995	0.005	SG 1
Gen105	IC: 460160	0.987	0.013	SG 1
Gen106	IC: 74637 A1	0.995	0.005	SG 1
Gen107	IC: 277830	0.994	0.006	SG 1
Gen108	IC: 323957	0.946	0.054	SG 1
Gen109	IC: 453927	0.948	0.052	SG 1
Gen110	IC: 457989	0.992	0.008	SG 1
Gen111	IC: 459147	0.991	0.009	SG 1
Gen112	IC: 449793	0.987	0.013	SG 1
Gen113	IC: 124346	0.981	0.019	SG 1
Gen114	IC: 124366	0.831	0.169	SG 1
Gen115	IC: 125267	0.996	0.004	SG 1
Gen116	IC: 133146	0.993	0.007	SG 1
Gen117	IC: 135827	0.994	0.006	SG 1
Gen118	IC: 135877	0.996	0.004	SG 1
Gen119	IC: 123505	0.815	0.185	SG 1
Gen120	IC: 124525	0.996	0.004	SG 1
Gen121	IC: 206322	0.994	0.006	SG 1
Gen122	IC: 206693	0.994	0.006	SG 1
Gen123	IC: 206754	0.991	0.009	SG 1
Gen124	IC: 206866	0.994	0.006	SG 1
Gen125	IC: 206615	0.994	0.006	SG 1
Gen126	IC: 331668	0.993	0.007	SG 1
Gen127	IC: 379109	0.995	0.005	SG 1
Gen128	IC: 379122	0.994	0.006	SG 1
Gen129	IC: 296890	0.996	0.004	SG 1

Gen130	IC: 388204	0.981	0.019	SG 1
Gen131	IC: 388737	0.858	0.142	SG 1
Gen132	IC: 389351	0.983	0.017	SG 1
Gen133	IC: 389509	0.991	0.009	SG 1
Gen134	IC: 389838	0.984	0.016	SG 1
Gen135	IC: 390299	0.978	0.022	SG 1
Gen136	IC: 435091	0.995	0.005	SG 1
Gen137	IC: 435541	0.994	0.006	SG 1
Gen138	IC: 435559	0.984	0.016	SG 1
Gen139	CR 3969-24-1-2-1-1 (IR 73907-753-2-3/ Pratiksya)	0.994	0.006	SG 1
Gen140	OR 2487-13 (OR 2076-2/Ashoka228)	0.993	0.007	SG 1
Gen141	NPT 14-10 (NPT 29/R 296)	0.921	0.079	SG 1
Gen142	JAYA	0.988	0.012	SG 1
Gen143	CR 3561-3-2-1-1-1-1 (Surendra/Annapurna)	0.995	0.005	SG 1
Gen144	Swarna	0.99	0.01	SG 1
Gen145	CR 3856-44-22-2-1-10-1-5 (IR 73963-86-1-5-2-2/CR2324-1)	0.996	0.004	SG 1
Gen146	OR 2573-11 (Birupa/ IR 76561-AC-8-8)	0.997	0.003	SG 1
Gen147	YNP 7060 (NPG 6/NPG 15)	0.998	0.002	SG 1
Gen148	CR 3856-29-14-2-1-1-7-1(IR 73963-86-1-5-2-2/CR 2324-1)	0.997	0.003	SG 1
Gen149	PA 6444	0.998	0.002	SG 1
Gen150	CR 3504-12-2-1-1-1-1 (IR 36/Birupa)	0.998	0.002	SG 1
Gen151	NDR 359	0.997	0.003	SG 1
Gen152	Loktimachhi	0.904	0.096	SG 1
Gen153	Swarna sub-1	0.997	0.003	SG 1
Gen154	Madhuraj-55	0.997	0.003	SG 1
Gen155	Barhasaal	0.996	0.004	SG 1
Gen156	Bisni	0.997	0.003	SG 1
Gen157	Tarunbhog	0.996	0.004	SG 1
Gen158	Chhattisgarh Zinc Rice-1	0.997	0.003	SG 1
Gen159	Lalloo 14	0.996	0.004	SG 1
Gen160	Lalmati	0.997	0.003	SG 1
Gen161	Kanak jira	0.992	0.008	SG 1
Gen162	O. officinalis	0.998	0.002	SG 1
Gen163	O. latifolia	0.998	0.002	SG 1
Gen164	*O. sativa* var. *fatua* (IC: 301549)	0.992	0.008	SG 1
Gen165	*O. sativa* var. *fatua* (IC: 301558)	0.989	0.011	SG 1
Gen166	*O. nivara* (IC: 301559)	0.995	0.005	SG 1
Gen167	*O. nivara* (IC: 301547)	0.996	0.004	SG 1
Gen168	*O. nivara* (IC: 301555)	0.968	0.032	SG 1
Gen169	*O. nivara* (IC: 301581)	0.972	0.028	SG 1
Gen170	*O. nivara* (IC: 301584)	0.994	0.006	SG 1
Gen171	*O. nivara* (IC: 301589)	0.962	0.038	SG 1
Gen172	*O. nivara* (IC: 301622)	0.954	0.046	SG 1
Gen173	*O. nivara* (IC: 301625)	0.839	0.161	SG 1
Gen174	*O. nivara* (IC: 301626)	0.989	0.011	SG 1
Gen175	*O. nivara* (IC: 301668)	0.992	0.008	SG 1
Gen176	Chanda (C: 287)	0.996	0.004	SG 1
Gen177	Bhantha luchai (B: 2733)	0.993	0.007	SG 1
Gen178	Karhani (K : 1276)	0.83	0.17	SG 1
Gen179	Adhan chilpa (A : 701)	0.986	0.014	SG 1
Gen180	Katarni bhog (K : 415 II)	0.994	0.006	SG 1
Gen181	Laxmibhog (L : 708)	0.991	0.009	SG 1
Gen182	Barounda Offtype (BOT : 61 II)	0.962	0.038	SG 1
Gen183	Chhatri (C : 54 I)	0.995	0.005	SG 1
Gen184	Umariya chudi (U : 229)	0.964	0.036	SG 1

Gen185	Dubraj (D : 1289)	0.993	0.007	SG 1
Gen186	Muchchhan moti (IC: 387442)	0.983	0.017	SG 1
Gen187	Aruna	0.992	0.008	SG 1
Gen188	Krishnanjana	0.991	0.009	SG 1
Gen189	Dubraj Selection-1	0.991	0.009	SG 1
Gen190	Badshahbhog Selection-1	0.629	0.371	Admixture
Gen191	Jaldubi	0.905	0.095	SG 1
Gen192	IGKV R1 (Rajeshwari)	0.996	0.004	SG 1

The mean alpha value was observed 0.086 for the population (Table 3), which was relatively small for the present population and revealed the presence of only few admixtures. If the value obtained for alpha is zero it indicates that most of the genotypes in the studied material belongs to separate group and if the value of alpha is greater than 1 it shows that population consist of maximum number of admixtures. The value of allele frequency divergence between two populations was 0.097 (Table 4).

Table 3 Mean Fst values and alpha value of population

Particulars	**Value**
Mean value of Fst_1	0.1270
Mean value of Fst_2	0.1548
Mean value of alpha	0.0860

Table 4 Allele frequency (divergence among population based on net nucleotide distance), computed using point estimates of P and overall proportion of membership of the sample in each of two clusters

Sub groups	**K 1**	**K 2**
K 1	0	0.0973
K 2	0.0973	0
Inferred clusters (Membership proportion)	0.564	0.436

The allele frequency value showed that the net nucleotide distance between SG 1 and SG 2 was less (0.0973) and the genotypes of both sub group majorly contributed for population diversity as their membership proportion was found to be high 0.564 and 0.436 respectively (Table 4). These two sub groups might be distinct due to several factors that included different agro-morphological traits, different nutritional and phytochemical parameters, different geographic origin and adaptation to different ecosystem and due to seasonal differentiation of rice genotypes.

Results of population structure were further confirmed by neighbour- joining tree. Neighbour-joining tree of 192 diverse rice genotypes was constructed using allele size matrix of 130 molecular markers through software TASSELv4.2. which categorized the rice genotypes into two main groups (SG-1 and SG-2) and supports the results of population structure analysis (Fig.5). The genotypes that mixed ancestry were grouped into separate cluster. The subgroup-1 and subgroup-2 both comprised of mixed genotypes i.e. exotic germplasm, Indian germplasm as well as improved genotypes suggesting that this might be due to the inclusion of common ancestors in their genesis.

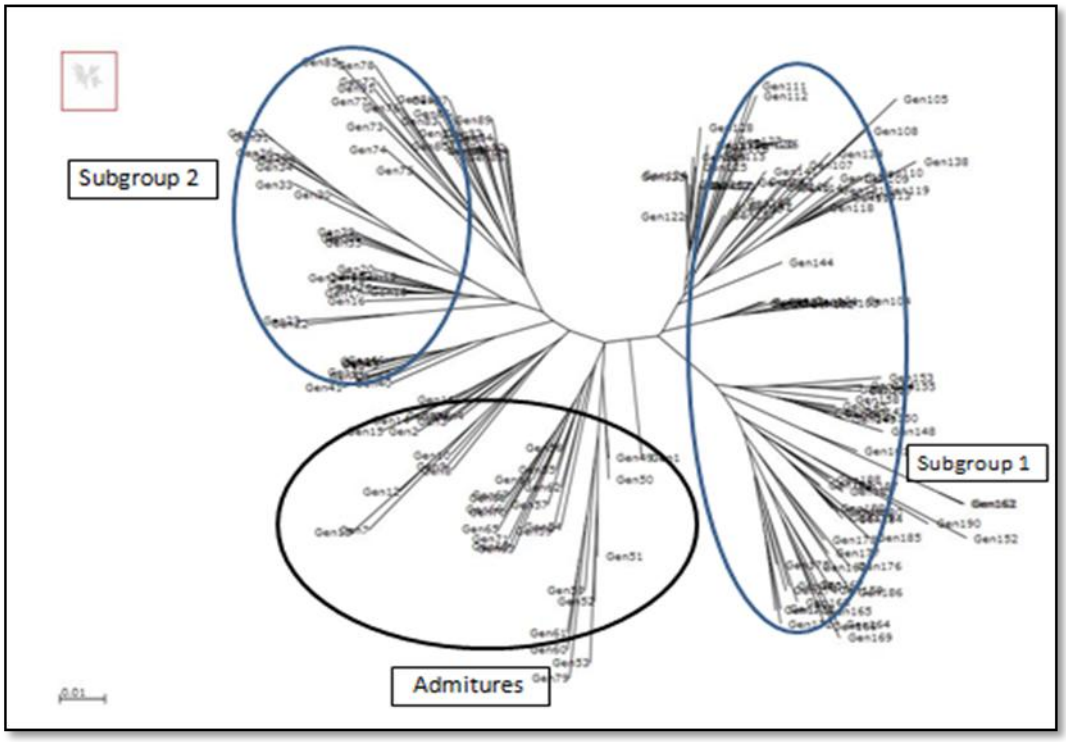


Fig 5 Neighbouring joining tree showing genetic relatedness among 192 diverse rice genotypes based on 130 SSR loci

Genome wide association mapping grain for quality traits and nutritional parameters

A total of 192 diverse set of rice genotypes were used for genome wide scanning through 130 molecular markers. The phenotypic data were generated on 10 grain quality traits and 6 nutritional parameters. The association mapping or LD mapping was performed by using phenotypic and genotypic data as source to analyse marker-trait association using two statistical models viz., general linear model (GLM) and mixed linear model (MLM/ Q+K model). Mixed linear model was estimated by taking into account of both population inferred ancestry (Q- matrix) and Kinship matrix (K) using software TASSELv2.1.

To test the significant marker-trait association Bonferroni threshold was estimated as 1/130 = 0.0076, where 130 is the number of markers used in the study (Moran, 2003). Therefore, the P value 0.0076 or less than 0.0076 for the markers were considered as significant (Table 5) for both GLM and MLM. Through MLM approach a total of 78 significant marker-trait association were detected for 16 different traits with R^2 (total variation explained in percentage) ranging from 44.45% to 7.0%. The associations obtained through MLM analysis were more effective and can be exploited further in smart breeding programme. Hence, in the present investigation more emphasis was given to MLM results and their detailed descriptions. Details of trait and their associated markers are described as follow:

Kernel length

Result revealed that only two markers RM15446 (chromosome-3) and RM228 (chromosome-10) were associated significantly with kernel length and explained 17.07% and 11.27% phenotypic variations, respectively.

Kernel breadth

The marker-trait association revealed significant association of two markers RM16738 (chromosome-4) and RM11694 (chromosome-1) with kernel breadth that described 9.69% and 8.62% phenotypic variation, respectively.

Kernel length breadth ratio

Only single marker RM16738 (chromosome-4) was found associated with kernel length breadth ratio and explained 15.65% of the phenotypic variations.

Hulling percentage

Five major loci were identified significantly associated with the hulling percentage trait and were distributed on chromosome-5 (RM18908), chromosome-6 (RM133 and RM510), chromosome-7 (RM21182) and chromosome-8 (RM339). These markers accounting the phenotypic variation ranging from 7.09% to 12.86%. Among these marker RM339 showed the highest phenotypic variation 12.86%.

Milling percentage
Five loci with significant association were detected with milling percentage and were located on chromosome-5 (RM17990 and RM18908), chromosome-6 (RM133), chromosome-7 (RM21182) and chromosome-8 (RM339). The phenotypic variance explained by these markers were varied from 9.91% (RM133) to 15.19% (RM339).
Head rice recovery percentage
For head rice recovery percentage three markers were identified on chromosome-2 (RM452), chromosome-6 (RM133) and chromosome-8 (RM8271) having significant association and with R^2 being 14.96%, 10.22% and 12.75%, respectively.
Kernel elongation ratio
For kernel elongation ratio three markers were identified on chromosome-4 (RM307), chromosome-5 (RM18600) and chromosome-12 (RM7601) having significant association and with R^2 being 12.14%, 9.55% and 8.56%, respectively.
Gel consistency
For gel consistency three markers were identified on chromosome-1 (RM10864), chromosome-2 (RM1) and chromosome-9 (RM105) having significant association and explaining 10.9%, 16.30% and 7.23%, phenotypic variation, respectively.
Gelatinization temperature
Only single marker Rc (chromosome-7) was found associated with gelatinization temperature and described about 15.47% of variations (R^2).
Amylose content
The marker-trait association revealed the putative association of three markers with amylose content and were distributed on chromosome-1 (RM10864), chromosome-2 (RM3858) and chromosome-4 (RM124). The marker RM10864 described highest phenotypic variation of 13.09% followed by marker RM3858 (10.54%) and RM124 (9.23%).
Protein content
Protein content was also found significantly associated with three markers distributed on chromosome-5 (RM18908) and chromosome-6 (RM5509 and RM510), explaining the phenotypic variation of 13.13%, 23.48% and 12.56, respectively.
Zinc content
Only single marker HvSSR05-24 (chromosome-5) was found associated with zinc content and described about 17.96% of variations (R^2).
Iron content
MLM analysis of marker-trait association revealed 36 markers associated with iron content locating on all the 12 chromosomes. The range of phenotypic variance explained (PVE) was from 17.69 (RM29)% to 44.45% (RM44).
Phenolics content
The results revealed that the markers namely RM17990 on chromosome-5 and Rc and RM21182 on chromosome-7 were significantly associated with free phenolics content. These markers accounting for 17.14%, 19.59% and 19.92% of phenotypic variations (R^2), respectively.
Flavonoid content
Five loci with significant association were detected with flavonoid content and were located on chromosome-1 (RM259), chromosome-3 (RM15446), chromosome-5 (RM17990), chromosome-6 (RM19454) and chromosome-7 (Rc). The phenotypic variance explained by these markers were varied from 13.55% (RM15446) to 29.04% (RM19454).
Oryzanol content
Two markers were identified to significantly associated on chromosome-4 (RM16738) and chromosome-6 (RM20730), explaining 8.76% and 27.27% phenotypic variation (R^2), respectively

Table.5 List of significant marker trait association (MTAs) (Bonferroni threshold < 0.0076) identified in association with 16 grain nutritional quality traits based on Q+K model in diverse rice panel.

S.No	Trait	Marker Name	Chromosome No	Map position	Marker P value	Inlikelihood	Marker R^2 value
1.	KL	RM15446	3	23080448bp	0.0059	-1.38E+02	0.121
2.	KL	RM228	10	87.6	0.0033	-1.61E+02	0.1127
3.	KB	RM16738	4	16410449	6.07E-05	-7.27E+01	0.1565
4.	KB	RM11694	1	32224855bp	0.0063	-6.02E+00	0.0862
5.	KLB	RM16738	4	16410449	0.0059	-5.05E+00	0.0969
6.	HULL %	RM 339	8	72.2	2.35E-04	-6.03E+02	0.1286
7.	HULL %	RM18908	5	24015384bp	0.001	-6.09E+02	0.1258
8.	HULL %	RM 133	6	0	0.0049	-6.14E+02	0.0666
9.	HULL %	RM 510	6	20.8	0.007	-6.15E+02	0.07

10.	HULL %	RM21182	7	5768137bp	0.0071	-6.06E+02	0.0991
11.	MILL %	RM 339	8	72.2	4.37E-04	-6.41E+02	0.1519
12.	MILL %	RM21182	7	5768137bp	0.0011	-6.41E+02	0.145
13.	MILL %	RM 133	6	0	0.0017	-6.51E+02	0.0991
14.	MILL %	RM17990	5	4664936bp	0.0023	-6.52E+02	0.1145
15.	MILL %	RM18908	5	24015384bp	0.0026	-6.47E+02	0.1247
16.	HRR %	RM8271	8	7611709bp	0.0011	-6.86E+02	0.1275
17.	HRR %	RM 133	6	0	0.0024	-6.99E+02	0.1022
18.	HRR %	RM 452	2	58.4	0.0039	-6.96E+02	0.1496
19	ER	RM 307	4	0	4.71E-04	42.8063	0.1214
20	ER	RM7601	12	3553768bp	0.003	41.3722	0.0856
21	ER	RM18600	5	18997773bp	0.0067	38.9064	0.0955
22	GC	RM 1	1	29.7	8.22E-04	-8.02E+02	0.163
23	GC	RM10864	1	14235715bp	0.0031	-8.32E+02	0.109
24	GC	RM 105	9	32.1	0.0073	-8.44E+02	0.0723
25	GT	Rc	7	-	7.85E-04	-2.85E+02	0.1547
26	AC	RM10864	1	14235715bp	5.41E-04	-5.96E+02	0.1309
27	AC	RM 124	4	150.1	0.0072	-6.15E+02	0.0923
28	AC	RM3858	2	6580532bp	0.0075	-6.07E+02	0.1054
29	PC	RM5509	6	27828305bp	8.71E-05	-3.66E+02	0.2348
30	PC	RM 510	6	20.8	0.0018	-3.87E+02	0.1256
31	PC	RM18908	5	24015384bp	0.005	-3.85E+02	0.1313
32	ZnC	HvSSR05-24	5	8283540bp	0.0047	-5.01E+02	0.1796
33	FeC	RM 44	8	60.9	9.32E-16	-3.40E+02	0.4445
34	FeC	RM21889	7	23826962bp	8.48E-14	-3.61E+02	0.3655
35	FeC	RM15767	3	28431374bp	8.48E-14	-3.61E+02	0.3655
36	FeC	RM 484	10	97.3	3.46E-13	-3.56E+02	0.3867
37	FeC	RM 507	5	0	6.54E-13	-3.54E+02	0.3827
38	FeC	RM25993	11	780582bp	9.02E-13	-3.59E+02	0.3678
39	FeC	RM 307	4	0	3.08E-12	-3.48E+02	0.4119
40	FeC	RM30	6	125.4	3.78E-12	-3.57E+02	0.3722
41	FeC	RM 124	4	150.1	4.25E-12	-3.56E+02	0.3775
42	FeC	RM 408	8	0	6.41E-12	-3.52E+02	0.3891
43	FeC	RM 284	8	83.7	6.72E-12	-3.45E+02	0.4134
44	FeC	RM 178	5	118.8	7.96E-12	-3.58E+02	0.3676
45	FeC	RM 489	3	29.2	8.26E-12	-3.45E+02	0.4185
46	FeC	RM1360	1	30.5	1.12E-11	-3.55E+02	0.3744
47	FeC	RM286	1	31.4	1.82E-11	-3.47E+02	0.3968
48	FeC	RM 215	9	99.4	1.83E-11	-3.53E+02	0.3827
49	FeC	RM22165	7	29406921bp	1.99E-11	-3.55E+02	0.3683
50	FeC	RM27428	12	598151bp	2.14E-11	-3.54E+02	0.3722
51	FeC	RM 6832	3	88.95	3.10E-11	-3.40E+02	0.4176
52	FeC	RM214	7	34.7	3.90E-11	-3.53E+02	0.378
53	FeC	RM 228	10	87.6	6.25E-11	-3.50E+02	0.3732
54	FeC	HvSSR04-32	4	22674757bp	2.25E-10	-3.48E+02	0.3882
55	FeC	RM6842	2	124679bp	2.44E-10	-3.50E+02	0.3893
56	FeC	RM5506	4	33273406bp	2.88E-10	-3.43E+02	0.413
57	FeC	RM 125	7	24.8	4.62E-10	-3.47E+02	0.3902
58	FeC	RM 447	8	124.6	1.06E-09	-3.40E+02	0.4016
59	FeC	RM 161	5	96.9	1.53E-07	-3.68E+02	0.3092
60	FeC	RM 133	6	0	6.68E-07	-3.69E+02	0.2753
61	FeC	HvSSR03-40	3	15505710bp	4.58E-04	-3.68E+02	0.2225
62	FeC	RM 536	11	55.1	0.0021	-3.68E+02	0.2406
63	FeC	RM19454	6	3945764bp	0.0029	-3.68E+02	0.1944
64	FeC	RM 118	7	96.9	0.003	-3.67E+02	0.2403
65	FeC	HvSSR03-09	3	3797148bp	0.003	-3.63E+02	0.2291
66	FeC	RM 162	6	108.3	0.0032	-3.70E+02	0.1904
67	FeC	RM 29	2	68.9	0.0041	-3.76E+02	0.1769
68	FeC	RM 334	5	141.8	0.0055	-3.48E+02	0.309
69	PheC	Rc	7	43.8-44.4cM	3.68E-08	-1.41E+03	0.1959

70	PheC	RM17990	5	4664936bp	4.51E-04	-1.44E+03	0.1714
71	PheC	RM21182	7	5768137bp	9.34E-04	-1.43E+03	0.1992
72	FC	Rc	7	43.8-44.4cM	1.25E-08	-1.69E+03	0.2338
73	FC	RM15446	3	23080448bp	4.30E-05	-1.69E+03	0.1355
74	FC	RM19454	6	3945764bp	0.0014	-1.70E+03	0.2904
75	FC	RM259	1	54.2	0.0046	-1.68E+03	0.1647
76	FC	RM17990	5	4664936bp	0.0071	-1.73E+03	0.1809
77	OC	RM20730	6	30786736bp	8.02E-06	-6.95E+02	0.2727
78	OC	RM16738	4	16410449	0.0062	-7.04E+02	0.0876

KL = Kernel length	HRR% = Head rice recovery percentage	ZnC = Zinc content
KB = Kernel breadth	GC = Gel consistency	FeC = Iron content
L/B = Kernel length breadth ratio	GT = Gelatinization temperature	PheC = Phenolics content
ER = Elongation ratio	AC = Amylose content	FC = Flavonoid content
HULL% = Hulling percentage	PC = Protein content	OC = Oryzanol content
MILL% = Milling percentage		

Pleiotropic QTL/ multiple trait associations

Among these QTL with significant marker trait association, 14 were found pleiotropic in nature. The markers RM133 showed highest number of association with four traits. Five markers viz., Rc, RM 16738, RM 17990, RM 18908 and RM 21182 had shown association with three traits and eight markers viz., RM 307, RM 124, RM 510, RM 339, RM 228, RM 10864, RM 15446 and RM 19454 shown association with three traits (Table

4.13). These MTAs will be helpful in marker assisted selection.

Markers	**Pleotropic effects**
RM307	Kernel elongation ratio, Iron content
RM124	Amylose content, Iron content
RM133	Hulling percentage, milling percentage, head rice recovery, Iron content
RM510	Hulling percentage, Protein content
RM339	Hulling percentage, Milling percentage
RM228	Kernel length, Iron content
Rc	Gelatinization temperature, Phenolics content, flavonoids content
RM10864	Gel consistency, Amylose content
RM15446	Kernel length, flavonoids content
RM16738	Kernel breadth, Kernel length breadth ratio, Oryzanol content
RM17990	milling percentage, phenolics content, flavonoid content
RM18908	Hulling percentage, milling percentage, protein content
RM19454	Iron content, flavonoids content
RM21182	Hulling percentage, milling percentage, phenolics content

DISCUSSION

Identification of genomic regions associated with grain quality and nutritional traits is a pre-requisite for deploying them in breeding program to enhance the quality breeding. Improving grain quality and nutritional characters in rice has drawn the attention of researchers since it has a significant impact on consumer needs, market demand, and the sustainability of rice food value chain. Several researchers attempted to map the genomic regions controlling these traits and identify the underlying genes (Alkahtani, 2022; Padmashree, 2023). The present study is conducted to perform genetic diversity studies, population structure analysis and marker-trait association on a panel of 192 diverse set of rice genotypes.

A total of 130 SSR, HvSSR and InDel markers used for genotyping and association mapping studies in a diverse set of rice genotypes. All these molecular markers were dispersed well covering the whole genome of rice. The high allelic diversity, based on microsatellite markers, was observed. A total of 48 highly informative markers were identified on the basis of diversity parameters of markers viz., PIC value (> 0.60), allelic diversity (>0.6) and number of alleles (≥ 6). These SSR markers could be used to expand the genetic architecture of the current genotypes. Magnitude of molecular diversity existing within the present diverse rice panel is comparable with the earlier reports in rice germplasm (Padmashree *et al*., 2023; Nayak *et al*., 2022; Sahu *et al*. 2017, Salem and Sallam,2016). These markers would be highly fruitful and could be exploited for the genetic studies and QTL mapping studies further.

It is essential to understand about the genetic structure and diversity of rice germplasm. The molecular markers proved to be an efficient tool in dissecting the genetic architecture of gene pool. The UPGMA algorithm and Rho's similarity-based cluster analysis grouped the 192 genotypes into two main groups (A and B) at 0.85

similarity coefficient. Group A comprised of only two genotypes *Oryza latifolia* and *Oryza officinalis,* which are wild species of rice, while rest 190 genotypes were grouped in the cluster B, consisting of *Oryza sativa, Oryza nivara* and weedy species *Oryza sativa* var. *fatua*. The cluster B further divided into two sub clusters *viz*., B1 and B2. The cluster B1 have predominantly of *indica* type accession with different geographic origin. On the other hand, cluster B2, consisted mainly of *Oryza nivara* accessions, weedy rice (*Oryza sativa* var *fatua*) and few indigenous landraces. Based on SSR profiles the rice genotypes were not distinctly classify on the basis of geographic origin and pedigree of genotypes. This discrepancy was due to fact that the markers selected for the study was random and not trait specific as well as they reveal only small part of genotypic and phenotypic association with the complex traits. The result was supported by Upadhyay *et al*. (2012) and Das *et al*. (2013). Roy *et al* (2016); Das *et al*. (2013) and Verma *et al*. (2015) found two distinct sub groups within rice germplasm, whereas Nachimuthu *et al*. (2015), Talukdar *et al*. (2017), Beccara *et al*. (2017) and Sahu *et al*. (2017) grouped the rice accessions into three distinct clusters using UPGMA clustering algorithm. The diverse panel of rice genotypes taken in the present study has enormous importance because of the properties they possessed. These accessions possessed nutritional grain quality, phytochemical properties, better agronomic performance and resistant to biotic and abiotic stresses that belong to different geographic origin. The estimation of the distant genetic relation in diverse germplasm is crucial for detection of novel allele and its further utilization in breeding and also to broaden the genetic base of modern varieties.

The model-based approach by structure analysis revealed two subpopulations most of which were pure with few admixtures. The genetic structure of populations has been previously reported in rice to vary from 2 to 8 subpopulations. The genetic structure of rice has previously been documented (Agrama *et al*., 2007; Zhang *et al*., 2007). Courtois *et al*., 2012; Salem *et al.,* 2024 and Fatamatuzzohora *et al*., 2024 have detected two subgroups in their study population of 425, 27 and 50 accessions of rice respectively, whereas Das *et al.,* 2013 had detected 91 accessions of rice landraces. Agrama *et al*. (2007) described the population structure among 103 rice accessions and detected 8 subpopulations with 123 SSR markers. Zhang *et al.* (2007) analyzed the genetic structure of rice in Guizhou province, China, and revealed that the rice germplasms in Guizhou could be divided into 7 subpopulations.

Results of population structure analysis (Figs 2, 3 and 4) revealed only two major clusters and indicated a clear genetic divergence based on origin and breeding history of the rice genotypes. In the present study, 27 rice accessions were categorized as admixtures with the stringent threshold of 80%. Based on criterion of maximum membership probabilities, 192 accessions were grouped into two populations of 111 and 81 genotypes each. This finding is in accordance with the previous studies where they obtained two distinct subgroups in their core collection of rice (Courtois *et al*., 2012; Zhang *et al*., 2011). In the current rice collection panel, the use of trait linked SSR markers might also be the reason for the formation of less number of groups. In our study, an Fst value for SG-1 and SG-2 was ranged between 0.1270 and 0.1548, respectively with an average of 0.1409 indicating moderate population structure. Thus, sub group-2 was said to be more structured population followed by SG-1.

Results of population structure were further confirmed by neighbour- joining tree. Neighbour-joining tree of 192 diverse rice genotypes was constructed using allele size matrix of 130 molecular markers which categorized the rice genotypes into two main groups (SG-1 and SG-2). The genotypes that mixed ancestry were grouped into separate cluster. The subgroup-1 and subgroup-2 both comprised of mixed genotypes i.e. exotic germplasm, Indian germplasm as well as improved genotypes suggesting that this might be due to the inclusion of common ancestors in their genesis. On the basis of the geographical origin, no significant grouping was observed. The results suggested that the markers selected for the present investigation were not able to differentiate the genotypes based on geographic origin and genetic diversity of the genotypes. Hence, it is advisable to include more number of markers in the study.

GWAS is a powerful tool to identify loci associated with traits. There are two models used in association study to find the marker-trait association MTAs), one is GLM and the other one is MLM. MLM is more accurate than GLM to find the MTAs. So in the present study MLM approach was utilized to find MTA. Using a diverse panel of 192 rice accessions and 130 molecular markers, association mapping of grain quality and nutritional-related traits revealed 78 significant marker-trait associations (MTA) with R^2 ranging from 7.0% to 44.45% using MLM analysis in the present study. The R^2 value can indeed be important as strong evidence of high marker-trait association. The *p-value* determines the association of QTLs with markers and the R^2 value predicts the magnitude of the QTL effects (Ashfaq *et al*. 2023). The first GWAS associated with rice grain quality traits were conducted by Huang *et al.* (2010) by using 517 rice diverse landraces that genotyped by ~3.6 million SNPs. GWAS detected major and minor quantitative trait loci (QTLs) correlated with the grain quality traits. GWAS also fine mapped of QTLs that lead to identify genes regulated QTLs of grain quality traits and guide to develop rice varieties with high-quality grains to satisfy the rice consumers preferences. In our study, GWAS detected multiple major loci for all the traits. In total, 78 QTLs were identified for grain quality and nutritional traits (Table 5). They were distributed on all the chromosomes and the PVE of these QTLs varied from 7% to 44.45%. The highest numbers of QTLs were identified for iron content and flavonoids content. It is also interesting to note that out of 78 QTLs, association of five markers (RM 536, RM 489 and RM 215 with FeC; Rc with PheC and Rc with FC) were previously reported in rice. While rest of 73 QTLs are novel and reported first time in this study. Marker RM 536 was significantly associated with iron content (FeC) located in

chromosome -11 (F *et al*., 2014). Sabouri et al. (2021) and Islam et al. (2020) also reported the association of RM 489 (FeC) and RM 215 with QTLFe9 (FeC) on chromosome 3 and 9 respectively. Similarly, Shao *et al*. (2011) reported the association of Rc with phenolic content and flavonoid content at position 43.8-44.4cM in chromosome-7 which is comparable to present reports. These five QTLs for three above traits have been validated co-incidentally through the present investigation.

The GWAS analysis identified RM510, located on Chromosome 6 at 2.83 Mb as significantly associated with grain protein content. This association indicates that RM510 is in strong linkage disequilibrium with a nearby QTL or functional gene influencing grain protein biosynthesis during grain filling. A comparable finding was reported by Sabouri *et al*. (2021), who detected a significant association of RM217 with protein content on Chromosome 6 at 17.9 Mb. The recurrence of protein-related associations on Chromosome 6 across independent studies & independent populations suggests that chromosome 6 contains major, stable QTLs governing grain protein content in rice. Together, these results support the presence of multiple protein-content QTLs on Chromosome 6 and highlight RM510 and RM217 as potential marker-assisted selection targets for improving grain nutritional quality in rice.

In the present study, SSR marker RM489 positioned on Chromosome 3 at 29.2 cM, showed a significant association with grain iron content. This result aligns with the findings of Sabouri *et al.* (2021), who also reported RM489 as a marker linked to iron content at the same chromosome and same position. This concordance suggests that the genomic region tagged by RM489 is a true and stable QTL for iron content, consistently detectable across different germplasm sets and experimental conditions. In addition, marker RM 215 present on chromosome 9 was also found to be associated with grain iron content and was found identical to previous reported QTL (QTL Fe9) in research of Islam *et al*., 2020. In nearby position Suman *et al.* (2020) detected QTL linked with marker RM 257.Therefore, RM489 and RM 215 can be considered a reliable molecular marker for selecting high–iron rice genotypes and can play an important role in biofortification breeding programs aimed at improving micronutrient content in rice.

The marker, RM118 mapped to Chromosome 7 at 96.9 cM, showed a significant association with grain iron content, indicating that this marker is in linkage disequilibrium with a QTL influencing iron accumulation in rice grain. A comparable association was also reported by Kumar *et al*. (2017) for RM234 for same trait and which is also located on Chromosome 7 at 88.2 cM. The detection of iron-related QTLs within a similar chromosomal region (88–97 cM) across independent studies suggests that Chromosome 7 harbors a consistent QTL cluster controlling iron deposition in rice grains. The proximity of RM118 and RM234 implies that both markers may be tagging the same QTL region or closely linked loci involved in Fe accumulation in rice grain.

GC was found to be associated with three markers on chromosome 1 and 9. This study revealed an association of RM 1 (Chr-1) with gel consistency on chromosome 1 in the rice experimental material. Suman *et al.* (2020) also reported the association of RM 246 with gel consistency located on chromosome 1. Swamy *et al.* (2012) identified a QTL denoted as GC2.1 located on chromosome 2, which was associated with the trait of GC. Wang *et al.* (2007) have identified the presence of QGC2 on chromosome 2, which has an impact on the trait of GC. Mapping efforts identified different QTLs for GC on chromosome 2. The present study could not detect the association of markers on chromosome 2 linked to GC. After adequate validation the above markers can be used for improving cooking quality of rice.

Three markers showed significant association with HRR per cent on chromosomes 2, 6 and 8. Among which the marker RM 452 showed the maximum value for trait variability 14.96% on chromosome 2 at 58.4cM. A comparable association was also reported by Suman *et al*. 2020 for RM 262 for the same trait and which is also located on chromosome 2 at 30.5cM .RM 133 was found common loci for hulling percentage, milling percentage and head rice recovery thus, this genomic region provides interesting targets for follow-up studies to enhance our understanding of the genetics of grain quality in rice. Apart from these fourteen pleotropic loci were also identified. The co-occurrence or pleiotropy of markers associated with diverse attributes related to grain quality and nutritional quality provides the opportunity for improvement of these traits through marker assisted selection and QTL pyramiding.

CONCLUSION

Rice genetic resources are the asset and are fuel for breeding program which serve as a source of natural genes in the development of rice varieties. The genetic diversity, population structure and association analysis of 192 diverse set of were analyzed using 130 SSR markers widely distributed on 12 chromosomes. The present study demonstrates the potential of highly informative and selective SSR markers for genetic diversity analysis and genome wide association studies in the diverse set of rice genotypes. Genetic dissection of diverse rice genotypes revealed enormous genetic variation and diversity for grain quality and nutritional grain traits. Population stratification-based grouping divide the rice panel into two sub-groups. This study reported 80 novel QTLs for grain quality and nutritional grain traits, while five earlier reported QTLs were also validated. This study also provides a direction for breeding efforts in the selection of parents from the current collection with potential for novel genes or QTLs in rice breeding for developing nutri-rich high yielding varieties.

REFERENCES

1. Agrama, H. A., Eizenga, G. C. and Yan, W. 2007. Association mapping of yield and its components in rice cultivars. Mol Breed., 19(14): 341–356.
2. Alkahtani J. 2022. Genome-Wide Association Study of Grain Quality Traits in Rice Detected Genomic Regions of High-Quality Rice for Increasing Rice Consumption Biosci., Biotech. Res. Asia, 19(2): 333-346.
3. Ashfaq, M.; Rasheed, A.; Zhu, R.; Ali, M.; Javed, M.A.; Anwar, A.; Tabassum, J.; Shaheen, S.; Wu, X. 2023. Genome-Wide Association Mapping for Yield and Yield-Related Traits in Rice (*Oryza sativa* L.) Using SNPs Markers. Genes, 14, 1089. https://doi.org/10.3390/ genes14051089.
4. Bao JS, Cai Y, Sun M, Wang G, Corke H. 2005. Anthocyanins, flavonols, and free radical scavenging activity of Chinese bayberry (Myrica rubra) extracts and their color properties and stability. J. Agric. Food Chem., 53: 2327-2332.
5. Becerra, V., Paredes, M., Ferreira, M.E., Gutierrez, E. and Diaz, L.M. 2017. Assessment of the genetic diversity and population structure in *japonica* rice germplasm used in breeding in Chile, with SSR markers. Chil. J. Agr. Res., 77(1): 15-26.
6. Becerra, V., Paredes, M., Guieerrez, E. and Rojo, C. 2015. Genetic diversity, identification and certification of Chilean rice varieties using molecular markers. Chil. J. Agr. Res., 75(3): 267-274.
7. Bouis, H.E. and WQelch, R.M. 2010. Biofortifiction: a sustainable agricultural strategy for reducing micronutrient malnutrition in the global south. Crop Sci., 50: S20-S32.
8. Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y. and Buckler, E.S. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633-2635.
9. Bryant, R.J., Jackson, A.K., Yeaster, K.M., Yan, W.G., McClung, A.M., and Fjellstrom, R.G. 2013. Genetic variation and association mapping of protein concentration in brown rice using diverse rice germplasm collection. Cereal Chem., 90: 445-452.
10. Bucci, R., Magri, A.D. Magri, A.L. and Marini F. 2003. Comparison of three spectrophotometric methods for the determination of gamma-oryzanol in rice bran oil. Analytical and Bioanalytical Chemistry, 375 (8): 1254-1259.
11. Chaudhary, P., Sharma, A., Singh, B. and Nagpal, A. 2018. Bioactivities of phytochemicals present in tomato. Journal of Food Science and Technology, 55(6): 3221-3230.
12. Courtois, B., Frouin, J., Greco, R., Bruschi, G., Droc, G. and Hamerlin, C. 2012. Genetic diversity and population structure in a European collection of rice. Crop Sci., 52: 1663-1675.
13. Das, B., Sengupta,S., Parida, S.K., Roy, B., Ghosh, M., Prasad, M. and Ghose, T.K. 2013. Genetic diversity and population structure of rice landraces from Eastern and North Eastern States of India. BMC Genet, 14 (1): 71-78.
14. Descalsota-Empleo, G.I., Noraziyah, A.A.S., Navea, I.P., Chung, C., Dwiyanti, M.S., Labios, R.J.D., Ikmal, A.M., Juanillas, V.M., Inabangan-Asilo, M.A., Amparado, A., Reinke, R., Cruz, C.M.V., Chin, J.H. and Swamy, B.P.M. 2019. Genetic Dissection of Grain Nutritional Traits and Leaf Blight Resistance in Rice. Genes, 10(1): 284-305.
15. Earl, D.A. and Holdt von, B.M. 2012. STRUCTURE HARVESTER: A Website and Program for Visualizing STRUCTURE Output and Implementing the Evanno Method. Conservation Genetics Resources, 4(2): 359-361.
16. Fatamatuzzohora, M., Islam, M.S., Rabbee, M.F. Hossain, M.S., Kang, S.G. and Matin, M.N. 2024. Genetic variation, population structure, and cluster analysis of rice (*Oryza sativa* L.) genotypes using morphological characteristics and molecular markers. *CEREAL RESEARCH COMMUNICATIONS,* 52: 1751–1768.
17. Federer, W. T. 1956. Augmented (or Hoonuiaku) designs. Hawaiian Planters' Record 55:191-208.
18. Gupta, P.K., Rastogi, S. and Kulwal, P.L. 2005. Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Mol. Biol., 57: 461-485.
19. Hammer, Q., Harper, D.A.T. and Ryan, P.D. 2001. Paleontological statistics software package for education and data analysis. Palaeontol. Electronica, 4(1): 9.
20. Huang X H, Han B. 2014. Natural variations and genome-wide association studies in crop plants. Annu Rev Plant Biol, 65: 531–551.
21. Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., Lu, T., Zhang, Z., Li, M., Fan, D., Guo, Y., Wang, L., Deng, L., Li, W., Lu, Y., Weng, Q., Liu, K., Huang, T., Zhou, T., Jing, Y., Li, W., Lin, Z., Buckler, E.S., Qian, Q., Zhang, Q.F., Li, J., and Han, B. 2010. Genome-wide association studies of 14 agronomic traits in rice landraces. Nat. Genet., 42(11): 961-967.
22. Islam M Z, Arifuzzaman M, Banik S, Hossain M A, Ferdous J, Khalequzzaman M, Pittendrigh B R, Tomita M, Ali M P. 2020. Mapping QTLs underpin nutrition components in aromatic rice germplasm. PLoS One, 15(6): e0234395.
23. Jia, Z., Mengcheng, T., Jianming, W., 1999. The determination of flavonoid contents in mulberry and their scavenging effects on superoxide radicals. Food Chemistry, 64: 555–559.
24. Jonah, P., Bello, L., Lucky, O., Midau, A. and Moruppa, S. 2011. Review: the importance of molecular markers in plant breeding programmes. Glob. J. Sci. Front. Res. 11: eV-ves1.

25. Kiranmayi, S. L., K. Manorama,K., Tripura Venkata, V. G. N. , Radhika, K., Cheralu, C., Roja, V. and Sarla, N. 2014. Identification of markers associated with iron and zinc concentration in recombinant inbred lines of brown rice. Indian J. Genet., 74(4): 423-429.
26. Kordrostami, M. and Rahimi, M. 2015. Molecular markers in plants: Concepts and applications. Genetics in the Third Millennium.13: 4024-4031.
27. Kumar, A., Maurya, S., Vishwakarma, A., Patil, A.H. and Chandel, G. 2017. Identification of DNA Markers Associated with Grain Iron, Zinc and Protein Contents in Rice (*Oryza sativa* L.). Trends in Biosciences, 10(38): 7881-7889.
28. Liu, K. and Muse, S.V. 2005. PowerMarker: Integrated Analysis Environment for Genetic Marker Data. Bioinformatics, 21: 2128-2129.
29. Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J.1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem., 193: 265-275.
30. Mason, A. 2015. SSR genotyping. Methods in molecular biology, 1245: 77-89.
31. McCouch, S.R., Chen, X.L., Panaud, O., Temnykh, S., Xu, Y.B., Cho, Y.G., Huang, N., Ishii, T. and Blair, M. 1997. Microsatellite marker development, mapping and applications in rice genetics and breeding. Plant Mol. Biol., 35: 89-99.
32. McCouch, S.R., Wright, M.H., Tung, C.W., Maron, L.G., McNally, K.L., Fitzgerald, M., Singh, N., DeClerck, G., Agosto-Perez, F. and Korniliev, P. 2016. Open access resources for genome-wide association mapping in rice. Nat. Commun.,7: 10- 53.
33. Moran, M. D. 2003. Arguments for rejecting the sequential Bonferroni in ecological studies. Oikos, 100: 403-405.
34. Nachimuthu, V. V., Muthurajan, R., Duraialaguraja, S., Sivakami, R., Pandian, B. A., Ponniah, G., Gunasekaran, K., Swaminathan, M., Suji, K. K. and Sabariappan, R. 2015. Analysis of population structure and genetic diversity in rice germplasm using SSR markers: An initiative towards association mapping of agronomic traits in *Oryza sativa*. Rice, 8: 30.
35. Nayak, L., Panda, D., Dash, G. K., Lal, M. K., Swain, P., Baig, M. J. 2022. A chloroplast Glycolate catabolic pathway bypassing the endogenous photorespiratory cycle enhances photosynthesis, biomass and yield in rice (*Oryza sativ*a L.). Plant Sci. 314, 111103. doi: 10.1016/j.plantsci.2021.111103.
36. Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Nat. Aca. Sci., USA, 70: 3321–3323.
37. Padmashree, R., V. Reddy, K. M. Barbadikar, M. Honnappa, N. D. Magar, D. Balakrishnan, C. Gireesh, A. M. Siddaiah, J. Badri, R. Lokesha, Y. M. Ramesha, P. Senguttuvel, J. R. Diwan, M. S. Madhav and R. M. Sundaram 2023. G×E analysis to Identify the Stable High-yielding Rice Lines among a Set of Selected Germplasm Panel. Int. J. Environ. Clim. Change, 13(7): 10-23. doi: 10.9734/IJECC/2023/v13i71849.
38. Pathak, K., Wasifur R. S., Bhagawati, S. and Gogoi, B. 2017. Assessment of nutritive and antioxidant properties of some indigenous pigmented hill rice (*Oryza sativa* L.) cultivars of Assam. Indian J. Agric. Res., 51 (3): 48-59.
39. Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics, 155: 945-959.
40. Rao, D. S., Babu, P. M., Swarnalatha, P., Kota, S., Bhadana, V. P., Varaprasad, S., Surekha, K., Neeraja, C. N. and Babu, V. R. 2014. Assessmentof grain zinc and iron variability in rice germplasm using energy dispersive X-ray fluorescence spectrophotometer. Journal of Rice Research, 7(1): 45-52.
41. Roy, S., Marndi, B.C., Mawkhlieng, B., Banerjee, A., Yadav, R.M., Misra, A.K. and Bansal, K.C. 2016. Genetic diversity and structure in hill rice (*Oryza sativa* L.) landraces from North Eastern Himalayas of India. BMC Genet., 17(10): 1-15.
42. Sabouri, A., Nasiri, E., Esfahani, M., & Forghani, A. 2021. SSR marker-based study of the effects of genomic regions on Fe, Mn, Zn, and protein content in a rice diversity panel. Journal of Plant Biochemistry and Biotechnology, 30(3): 504–514. https://doi.org/10.1007/s13562-020-00637-x
43. Sahu, P. K., Mondal, S., Sharma, D., Vishwakarma, G., Kumar, V. and Das, B. K. 2017. InDel marker based genetic differentiation and genetic diversity in traditional rice (Oryza sativa L.) landracFes of Chhattisgarh, India. PloS one, 12(11), e0188864.
44. Salem, K. F.M. and Sallam, A. 2016. Analysis of population structure and genetic diversity of Egyptian and exotic rice (*Oryza sativa* L.) genotypes. C. R. Biologies 339(1) : 1-9.
45. Salem, K.F.M., Safhi, F.A., Alwutayd, K.M., Abozahra, M.S., Almohisen, I.A.A., Alsharari, S.F. 2024. Analysis of genetic diversity, population structure and phylogenetic relationships of rice (*Oryza sativa* L.) cultivars using SSR markers. Genetic Resources and Crop Evolution. 71 (5):2213–27. https://doi.org/10.1007/s10722-023-01789-0.
46. Shao, Y., Jin, L., Zhang, G., Lu, Y., Shen, Y. and Bao, J. 2011. Association mapping of grain color, phenolic content, flavonoid content and antioxidant capacity in dehulled rice. Theor Appl Genet., 122(5): 1005-16.
47. Suman, K., Madhubabu, P., Rathod, R., Rao, D.S., Rojarani, A., Prashant, S., Subbarao, L.V., Ravindrababu, V., Neeraja, C.N., 2020. Variation of grain quality characters and marker-trait association in rice (Oryza sativa L.). J. Genet. 99 (1): 1–12.

48. Swamy, B.P., Kaladhar, K., Shobha, R.N., Prasad, G.S., Viraktamath, B.C., Reddy, G.A. 2012. QTL analysis for grain quality traits in 2 BC2F2 populations derived from crosses between Oryza sativa cv Swarna and 2 accessions of *O. nivara*. J Hered., 103(3): 442–452. https://doi.org/10.1093/jhered/esr145 PMID: 22312119.
49. Talukdar, P. R., Rathi, S., Pathak, K., Chetia, S., Ranjan Baruah, A. and Sarma, R.N. 2017. QTL Analysis in Aromatic Rice of Assam, India. J. Rice Res., 5(4): 186-192.
50. Upadhyay, P., Neeraja, C.N., Kole, C. and Singh, V.K. 2012. Population structure and genetic diversity in popular rice varieties of India as evidenced from SSR analysis. Biochemical Genetics, 50(9): 770-783.
51. Verma, H., Pathak, K., Rathi, S. and Sharma, R.N. 2015. Association analysis for grain quality traits in rice. Indian J. Genet. Plant Breed., 75(4): 506-509.
52. Wang, L. Q., Liu, W.J., Xu, Y., He, Y.Q., Luo, L.J., Xing, Y.J., Xu, C.G and Zhang, Q. 2007. Genetic basis of 17 traits and viscosity parameters characterizing the eating and cooking quality of rice grain. Theoretical and Applied Genetics. 115: 463-476.
53. Yang, L., He, W., Zhu, Y., Lv, Y., Li, Y. and Zhang, Q. 2025. GWAS meta-analysis using a graph-based pan-genome enhanced gene mining efficiency for agronomic traits in rice. Nat. Commun. 16, 3171. doi: 10.1038/s41467-025-58081-1.
54. Yu, J., Pressoir, G., Briggs, W. H., Vroh, B. I., Yamasaki, M. and Doebley, J. F. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet., 38: 203-208.
55. Zhang, D., Zhang, H., Wei, X., Qi, Y., Wang, M., Sun. J., Ding, L., Tang, S., Cao, Y. and Wang, X. 2007. Genetic structure and diversity of Oryza sativa L. in Guizhou, China. Chinese Sci. Bull., 52(3): 343-351.
56. Zhang, P., Li, J., Li, X., Zhao, X. and Lu, Y. 2011. Population structure and genetic diversity in a rice core collection (*Oryza sativa* L.) investigated with SSR markers. PLoS One, 6(12): e27565.
57. Zheng, K.L., Huang, N., Bennett, J. and Khush, G.S. 1995. PCR-Based Marker Assisted Selection in Rice Breeding. International Rice Research Institute, Manila, the Philippines,12: 1-24.
58. Zhu, C., Gore, M., Buckler, E.S. and Yu, J.2008. Status and prospects of association mapping in plants. Plant Genome, 1: 5-20.
Web Sources: http://www.whfoods.com and http://www.gramene.org/markers/microstat/