



RESEARCH ARTICLE

Identification of markers for yield and yield-related traits using association mapping analysis in rice (*Oryza sativa* L.) under rainfed conditions

Surekha¹, Vishal Kumar Gupta^{2*}, S B Verulkar¹, Shrikant Yankanchi³ & Sunil Kumar Verma^{1*}

¹Department of Plant Molecular Biology and Biotechnology, Indira Gandhi Krishi Vishwavidyalaya, Raipur 492 012, Chhattisgarh, India

²Department of Genetics and Plant Breeding, Mankombu Sambasivan Swaminathan School of Agriculture, Centurion University of Technology and Management, Gajapati 761 211, Odisha, India

³Department of Plant Molecular Biology and Biotechnology, ICAR-Indian Institute of Maize Research, Ludhiana 141 004, Punjab, India

*Correspondence email - vishalgupta030794@gmail.com; sunil.verma244@gmail.com

Received: 26 December 2025; Accepted: 25 June 2026; Available online: Version 1.0: 28 September 2026

Cite this article: Surekha, Vishal KG, Verulkar SB, Shrikant Y, Sunil KV. Identification of markers for yield and yield-related traits using association mapping analysis in rice (*Oryza sativa* L.) under rainfed conditions. Plant Science Today (Early Access). <https://doi.org/10.14719/pst.13374>

Abstract

Rice (*Oryza sativa* L.) grain yield is determined by multiple factors, including ideal plant architecture, appropriate flowering time and adaptability to diverse agro-climatic conditions. However, inadequate water availability in major rice-growing regions has led to increased incidence of drought and other abiotic stresses. The present investigation was undertaken to identify marker-trait associations for yield and yield-related traits, as well as to assess genetic diversity and population structure under rainfed conditions. A set of 100 genetically diverse rice genotypes, classified into wild, cultivated and germplasm groups, were evaluated under rainfed conditions using 151 Simple Sequence Repeat (SSR) markers and marker-trait associations were analysed through general linear model (GLM) and mixed linear model (MLM) models using TASSEL and STRUCTURE software during 2020–2021 in the Agriculture Research cum Instructional farm and Department of Plant Molecular Biology and Biotechnology, College of Agriculture, Indira Gandhi Krishi Vishwavidyalaya, Raipur, India. A total of 414 alleles were amplified across 151 SSR markers (alleles/locus: 3 to 11 with mean 2.77). The polymorphism information content (PIC) values ranged from 0.019 to 0.498 (mean 0.263) the most informative markers were CS-62, RM28160, RGS-1 and RM-408 (PIC > 0.49). STRUCTURE Harvester (ΔK , $K = 3$) classified 100 genotypes into three sub-populations (POP1: 85; POP2: 13; POP3: 2). The GLM and MLM identified 90 and 42 significant marker-trait associations respectively (False discovery rate (FDR): 50.1 %), with 10 SSR markers associated with multiple yield-related traits under both models. Eight SSR markers (RM-242, RM-178, RM-154, RM-21, 18646, IND-1, CS-62 and PPC-1SET1) were tightly linked with grain yield across all linkage groups, confirming their suitability for marker-assisted selection. Sarmatia (250.05 g/plot) and IBD-1 (234.55 g/plot) are recommended as donor parents for breeding drought-tolerant rainfed rice varieties.

Keywords: abiotic stress; drought tolerance; genetic diversity; grain productivity; population structure; SSR markers

Introduction

Rice is one of the most important crops and crucial for global food security and the *Oryza* germplasm (which includes the genetic material of rice varieties) plays a vital role in rice breeding. The state of Chhattisgarh is known as the "rice bowl of India" due to presence of a rich diversity in rice cultivar. The area under rice cultivation in India is 51.27 million ha with production of 150.18 million tonnes and an average productivity of 2929 kg/ha whereas, the state of Chhattisgarh stands at 4.73 million ha in area under rice cultivation with production of 11.38 million tonnes and average productivity of 2409 kg/ha (1). Genetic diversity is essential for developing new varieties that are resilient to pests, diseases and changing environmental conditions. By harnessing the diversity within the *Oryza* germplasm, scientists and breeders can create rice varieties that are higher yielding, more nutritious and better adapted to

various growing conditions around the world (2). *Oryza sativa* L. (Asian rice) and *Oryza glaberrima* Steud. (African rice) are two distinct cultivated species within the genus *Oryza*. Both species belong to the AA genome with a diploid level ($2n = 24$), indicating they have similar genetic characteristics at the basic chromosome level. *O. sativa* (Asian rice) was domesticated independently in South Asia, particularly in regions around the foothills of the Himalayas. On the contrary *O. glaberrima* African rice domesticated independently in West Africa, especially in the region around the Niger River delta. Evaluating genetic diversity using DNA markers has become the standard due to its repeatability, stability and reliability. Several techniques have been developed and widely used for analysing the genetic improvement of rice varieties, based on DNA, such as the analysis of different marker utilising the polymerase chain reaction (PCR) (3). Assessment of morphological traits alone to measure genetic

diversity with traditional methods is much more difficult when compared to molecular marker-based methods, which required additional time, space and labour cost also highly influenced by environment (4). Conventional breeding methods for developing drought-resistant cultivars are often constrained by a limited understanding of the complex genetic mechanisms underlying drought tolerance in plants. Drought tolerance is typically polygenic in nature and is characterised by strong genotype $\times$ environment ($G \times E$) interactions, low heritability, weak correlations with yield-related traits and the occurrence of linkage drag. These methods primarily rely on crossing plants with desirable traits and selecting superior progeny over successive generations to enhance drought resistance (5).

Linkage mapping and association mapping based on linkage disequilibrium (LD) are the two principal approaches used for the identification of genes or quantitative trait loci (QTLs). Conventional linkage mapping using biparental populations is constrained by low mapping resolution, limited recombination events and the analysis of only two alleles per locus. In contrast, association mapping utilises natural genetic diversity and historical recombination events, thereby offering higher mapping resolution and improved detection of favourable alleles. Since its first successful application in maize, association mapping has become an effective tool for dissecting complex agronomic traits in crop plants (6). Recent advances in genome-wide association studies (GWAS) and marker-assisted breeding have facilitated the identification of markers linked with drought tolerance and yield traits in rice (7). However, earlier studies were restricted to biparental populations, smaller sample sizes or limited trait evaluation under controlled environments. Comprehensive marker-trait association studies conducted under rainfed field conditions in Chhattisgarh using genetically diverse rice germplasm are still limited. Furthermore, very few studies have simultaneously investigated multiple yield and yield-related traits using diverse rice accessions comprising wild, cultivated and germplasm groups to identify stable multi-trait associated Simple Sequence Repeat (SSR) markers suitable for marker-assisted selection under drought-prone rainfed ecosystems. Population structure and kinship among genotypes are major factors contributing to false-positive associations in marker-trait analysis. To overcome this limitation, mixed linear model (MLM)-based approaches incorporating both population structure (Q matrix) and kinship (K matrix) are considered more reliable than general linear model (GLM) based on Nei's genetic distance (8). Nevertheless, significant associations identified in one population require validation in independent populations before their effective utilisation in breeding programmes. The present study was undertaken based on the hypothesis that genetically diverse rice genotypes grown under rainfed conditions exhibit significant marker-trait associations for yield and yield-related traits. It was further hypothesised that SSR-based association mapping using general linear model (GLM) and mixed linear model (MLM) approaches can identify stable and reliable markers for marker-assisted breeding. In order to test this hypothesis, 100 diverse rice genotypes were evaluated for 18 yield and yield-related traits under rainfed field conditions. Genetic diversity and population structure were assessed using 151 SSR markers and STRUCTURE Harvester analysis. Furthermore, SSR markers significantly associated with yield and related traits were identified using GLM and MLM approaches for

their potential utilisation in marker-assisted selection and rice improvement programmes under rainfed ecosystems.

Materials and Methods

Plant material selection and its evaluation

In the present study 100 rice genotypes comprising with landraces and cultivated varieties sown in field condition at instructional cum research farm of Indira Gandhi Krishi Vishwavidyalaya (IGKV), Raipur, under rainfed condition following a randomised block design with two replications (Table 1). The genotypes are obtained from department of plant molecular biology and biotechnology, IGKV, Raipur, Chhattisgarh. Each genotype was sown with 2 rows of 2.4 m length was used during each season. These 100 lines are selected based on diverse genetic backgrounds to maximise allelic diversity, adaptation to rainfed and drought-prone conditions of Chhattisgarh, contrasting performance for grain yield and drought tolerance essential for association mapping. Genotypes such as Sarmatia and IBD1 were included for their superior grain yield under rainfed conditions, while WR41 represented the low-yielding wild background, providing the phenotypic contrast necessary for effective marker trait association detection. The 18 yield and attributing traits including plant height, days to 50 % flowering, panicle length, number of productive tillers, spikelet fertility (%), 100 grain weight and grain yield per plot measured after Protection of Plant Varieties and Farmers' Rights (PPV&FR) distinctiveness, uniformity and stability (DUS) guidelines for rice where five plant were selected randomly from each genotype and individual observations were recorded.

Climate and soil characteristics

The experiment was conducted during the wet cropping season of 2020–2021 at the Research Farm of Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh, under rainfed conditions. The region falls under a dry sub-humid climatic zone and weekly observations of maximum and minimum temperature were recorded throughout the crop growth period (Fig. 1). The experimental field had sandy loam soil. Soil samples collected from 0–15 cm depth were analysed for physicochemical properties. Soil pH ranged from 6.90 with normal electrical conductivity (0.23 dS m^{-1}) and organic carbon content of 0.49 %. Available nitrogen, phosphorus, potassium and sulphur ranged from 100–204, 5.4–39.5, 277.31–484.96 and 16.80–46.98 kg/ha respectively indicating suitable soil fertility for evaluation of rice genotypes under rainfed environments. No supplemental irrigation given to crop.

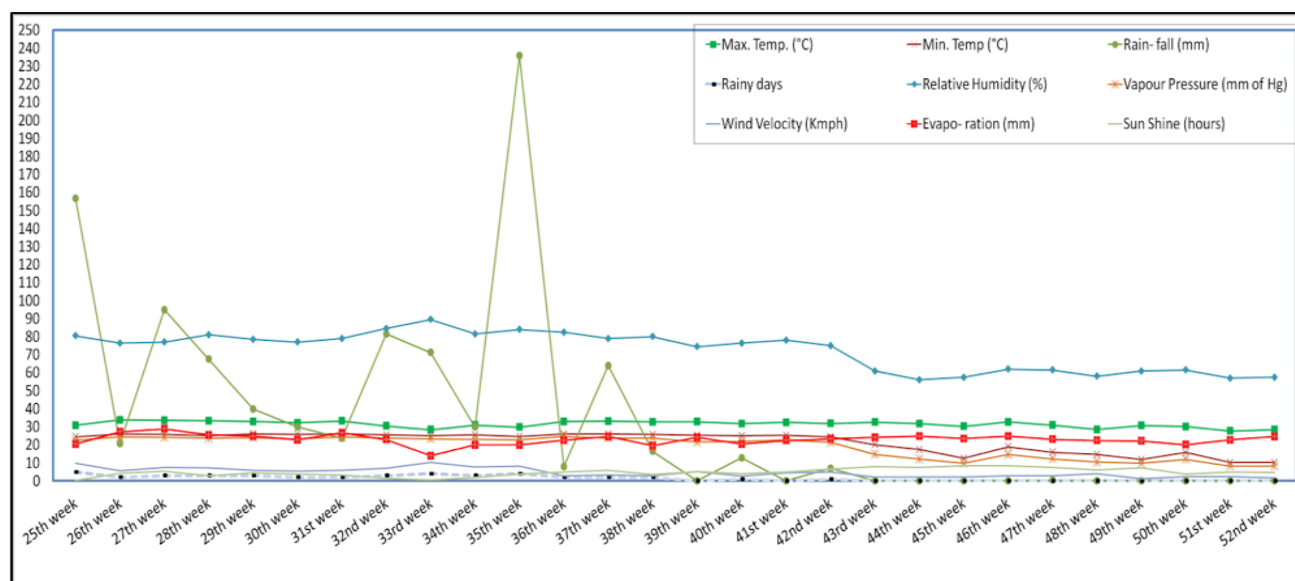
Molecular characterisation

Genomic DNA was isolated from fresh young leaf tissue (0.1 g) of each rice genotype following a modified cetyltrimethylammonium bromide (CTAB) method, purified by chloroform: isoamyl alcohol (24:1) extraction, precipitated with pre-chilled isopropanol, washed with 70 % ethanol and dissolved in 50–100 μl Tris–EDTA (TE) buffer. The DNA concentration and purity were assessed using a NanoDrop spectrophotometer (Thermo Scientific ND-1000) at 260 nm absorbance. Polymerase chain reaction (PCR) amplification was performed using 151 SSR primers in a 20 μl reaction volume containing genomic DNA (25 ng/ μl), deoxyribonucleotide triphosphates (dNTPs) (2.0 mM), primer pair (10 μM), 2 $\times$ polymerase chain reaction buffer and Thermus

Table 1. List of genotypes

Sl. No.	Source	Name of Variety	Sl. No.	Source	Name of variety
1	PMBB, IGKV	Annada	51	PMBB, IGKV	Sehra dabri
2	PMBB, IGKV	Abhaya	52	PMBB, IGKV	Tarunbhog
3	PMBB, IGKV	Azucena	53	PMBB, IGKV	Chapti Gurmata
4	PMBB, IGKV	ARB6	54	PMBB, IGKV	Kalam Gurmata (3053)
5	PMBB, IGKV	Bamleshwari	55	PMBB, IGKV	Srikamal
6	PMBB, IGKV	Buddha	56	PMBB, IGKV	Elayachi
7	PMBB, IGKV	Bakal	57	PMBB, IGKV	Bisni-I
8	PMBB, IGKV	Bhatapool	58	PMBB, IGKV	CGZR-1
9	PMBB, IGKV	Batroo	59	PMBB, IGKV	Basmati 370
10	PMBB, IGKV	CT 9993	60	PMBB, IGKV	Kalamata
11	PMBB, IGKV	Cross 116	61	PMBB, IGKV	Jhilli
12	PMBB, IGKV	Deshi Lal Dhan	62	PMBB, IGKV	Ramjiyawan
13	PMBB, IGKV	Dagad Deshi	63	PMBB, IGKV	RKVY-104
14	PMBB, IGKV	IR 62266	64	PMBB, IGKV	BAM 5926
15	PMBB, IGKV	IR 36	65	PMBB, IGKV	Moroberekan
16	PMBB, IGKV	IR42253	66	PMBB, IGKV	Naginna 22
17	PMBB, IGKV	IR64	67	PMBB, IGKV	GP-145-42
18	PMBB, IGKV	IR 55419-04	68	PMBB, IGKV	GP-145-43
19	PMBB, IGKV	Kranti	69	PMBB, IGKV	Kalokuchi
20	PMBB, IGKV	Lalmat	70	PMBB, IGKV	Shenong
21	PMBB, IGKV	Laloo-14	71	PMBB, IGKV	IR-55419
22	PMBB, IGKV	Mahamaya	72	PMBB, IGKV	R-RF-75
23	PMBB, IGKV	Samleshwari	73	RG 1, IGKV	Kadam Phool
24	PMBB, IGKV	Swarna	74	RG 2, IGKV	Ama Jhopa
25	PMBB, IGKV	Vandana	75	RG 3, IGKV	Koudi Dhull
26	PMBB, IGKV	IBD-1	76	RG 4, IGKV	Chhind Guchchi
27	PMBB, IGKV	Danteshwari	77	RG 5, IGKV	Nariyal Phool
28	PMBB, IGKV	Poornima	78	RG 6, IGKV	Ama Ruthi
29	PMBB, IGKV	Sahabhagi Dhan	79	RG 7, IGKV	Khajur Jhopa
30	PMBB, IGKV	MTU 10106	80	RG 8, IGKV	Naykain Jhaba
31	PMBB, IGKV	Aganni	81	RG 9, IGKV	Ram Laxman
32	PMBB, IGKV	Karma masuri	82	RG 10, IGKV	Do Dana
33	PMBB, IGKV	Safri 17	83	RG 11, IGKV	Pankhi
34	PMBB, IGKV	Dubraj	84	RG 12, IGKV	Sua Pankhi
35	PMBB, IGKV	BPT 5204 (Improved)	85	RG 13, IGKV	Maina Gali
36	PMBB, IGKV	Jitpiti	86	RG 14, IGKV	Parewa
37	PMBB, IGKV	PR 122	87	RG 56, IGKV	Chini Kapoor
38	PMBB, IGKV	WR 32	88	RG 16, IGKV	Pakshi Raj
39	PMBB, IGKV	WR 36	89	MED 4551, IGKV	Surmatia
40	PMBB, IGKV	WR 41	90	RG 18, IGKV	Parmal
41	PMBB, IGKV	WR 99	91	RG 19, IGKV	Dokra Mechha
42	PMBB, IGKV	WR 116	92	RG 65, IGKV	Jai Gundi
43	PMBB, IGKV	E-1702	93	RG 21, IGKV	Khatia Pati
44	PMBB, IGKV	Maheshwari	94	RG 66, IGKV	Tulsi Prasad
45	PMBB, IGKV	Durgeshwari	95	RG 26, IGKV	CR-1014
46	PMBB, IGKV	Shymala	96	RG 24, IGKV	Anterved
47	PMBB, IGKV	Rajeshwari	97	RG 95, IGKV	Layacha
48	PMBB, IGKV	Chandahasini	98	RG 26, IGKV	Tulsi Manjari
49	PMBB, IGKV	Jaldubi	99	RG 27, IGKV	Kera Ghul
50	PMBB, IGKV	Indira Sugandhit Dhan 1	100	RG28, IGKV	Kubri Mohar

Note: PMBB: Department of Plant Molecular Biology and Biotechnology; IGKV: Indira Gandhi Krishi Viswavidyalaya.

**Fig. 1.** Meteorological data recorded during 25th June to 31st December, 2020.

aquaticus DNA (Taq DNA) polymerase, with thermal cycling conditions of initial denaturation at 95 °C for 5 min, followed by 35 cycles of 95 °C for 1 min, 55 °C for 1 min and 72 °C for 1 min, with a final extension at 72 °C for 7 min. Amplified products were resolved on 5 % polyacrylamide gel (CBS Scientific) at 180 V using a 50 bp DNA ladder as molecular size standard, stained with ethidium bromide and visualised under UV light using a gel documentation system. Simple sequence repeat (SSR) bands were scored based on presence/absence and allele size and only distinct and reproducible bands were considered for assessment of genetic diversity, population structure and marker-trait association analysis (9).

The genotypic data were obtained after PAGE analysis. The gel images of markers RM19 and RM586 are presented in Fig. 2.

Statistical analysis of genotyping data

Estimating the genetic diversity of rice germplasm molecular study was executed, which includes DNA isolation, quantification and dilution of DNA, PCR amplification using microsatellite primers, electrophoresis using polyacrylamide gel, scoring and analysis of data. Statistical analyses of genotypic and association mapping data were performed using STRUCTURE software to assess genetic diversity within the rice population. Marker-trait association analysis for yield-related traits was carried out using TASSEL software. Phenotypic traits, markers on rice chromosomes. Polymorphism information content (PIC) was calculated for each SSR marker using the formula (10)

$$PIC_i = 1 - \sum_{j=1}^n p_j^2$$

Where, P_j is the frequency of the j^{th} allele for marker i . The PIC values range from 0 to 1, markers with $PIC > 0.5$ are considered highly informative, 0.25–0.5 moderately informative and < 0.25 less

informative (11).

Gene diversity (Expected heterozygosity, H_e) was estimated as:

$$H_e = 1 - \sum_{i=1}^k p_i^2$$

Where, P_i is the frequency of the i^{th} allele and k is the total number of alleles per locus. Major allele frequency (MAF) was recorded as the frequency of the most common allele at each locus. All diversity parameters were computed using Power Marker v3.25 (12).

Population structure analysis

Population structure among the rice genotypes was inferred using STRUCTURE software v2.3.4 (13) based on a Bayesian model-based clustering approach. The number of sub-populations (K) was tested from $K = 1$ to $K = 10$, with 10 runs per K value, using a burn-in period of 100000 iterations followed by 100000 Markov Chain Monte Carlo (MCMC) replications under the admixture model with correlated allele frequencies. The optimal K value was determined using the ΔK method of (14) through STRUCTURE Harvester (15). Membership probability (Q matrix) of each genotype to each sub-population was assigned based on the highest Q value (threshold $Q \geq 0.60$).

Kinship analysis

A kinship matrix (K matrix) representing pair-wise genetic relatedness among the 100 genotypes was estimated using TASSEL v3.0 (16) based on the identity-by-state (IBS) method. The K matrix was used as a covariate in the mixed linear model (MLM) to control for cryptic relatedness and minimise false-positive associations.

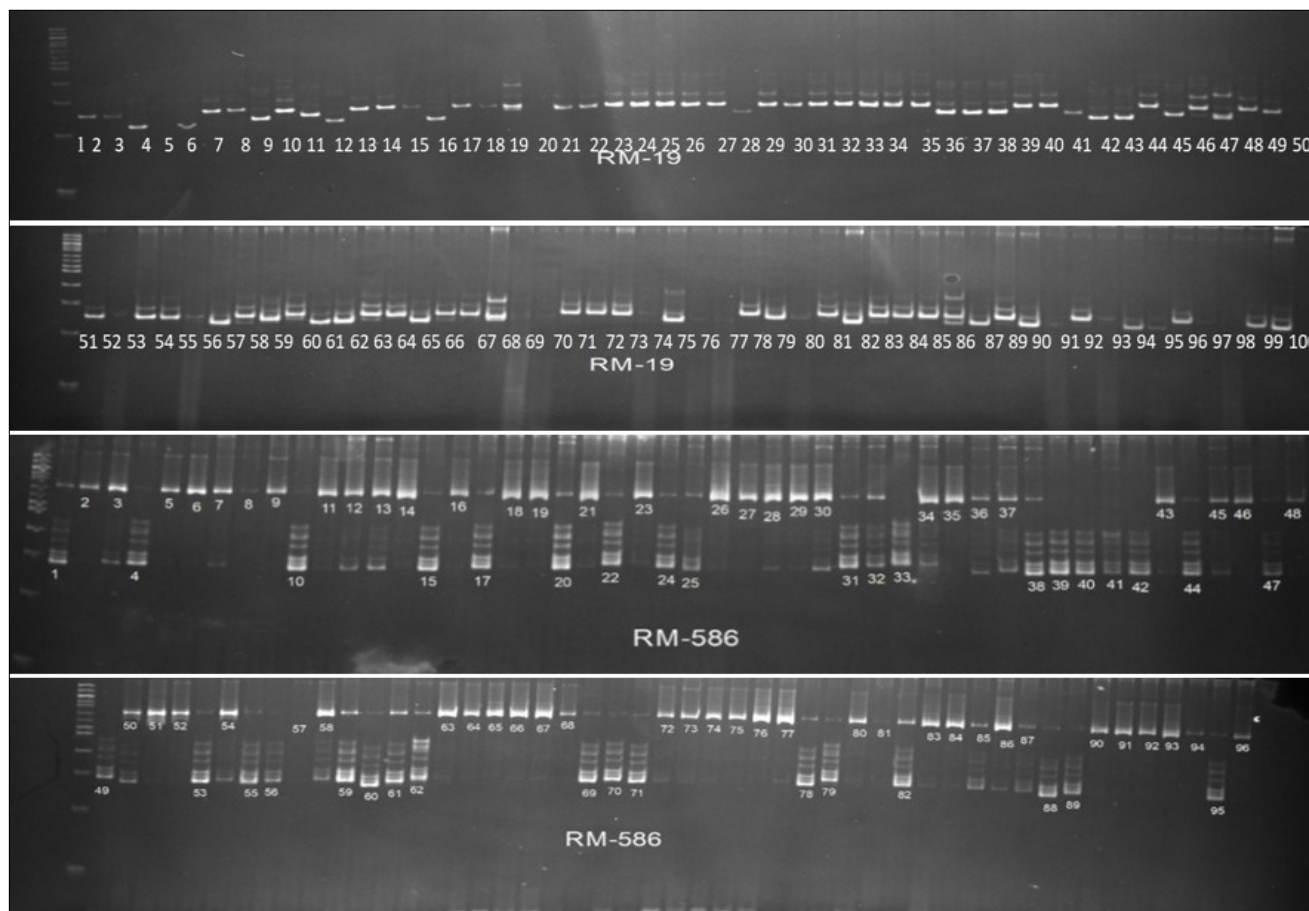


Fig. 2. Gel electrophoresis picture showing different banding pattern of genotype.

Marker-trait association analysis

Marker-trait association analysis for 18 yield and yield-related traits was carried out using TASSEL v3.0 employing two models:

General linear model (GLM) incorporates population structure (Q matrix) as a covariate:

$$Y = X\beta + Qv + \varepsilon$$

Where, Y is the phenotypic value, X is the marker genotype, β is the marker effect, Q is the population structure matrix, v is the fixed effect of sub-population and ε is the residual error.

Mixed linear model (MLM) incorporates both population structure (Q matrix) and kinship (K matrix) to further reduce false-positive associations

$$Y = X\beta + Qv + Zu + \varepsilon$$

Where, Z is the incidence matrix, u is the random genetic effect with variance $\text{Var}(u) = 2K\sigma_g^2$, K is the kinship matrix, σ_g^2 is the genetic variance and $\varepsilon \sim N(0, \sigma_e^2)$. Marker trait associations were considered significant at $p \leq 0.001$ with R^2 values indicating the proportion of phenotypic variance explained by the marker. False discovery rate (FDR) correction was applied to control for multiple testing (17). Only markers consistently significant in both GLM and MLM were considered reliable for further recommendation in marker-assisted selection (MAS).

Results and Discussion

Phenotypic analysis

The present investigation assessed 18 yield and yield-related traits across 100 diverse rice genotypes under rainfed field conditions at IGKV, Raipur during kharif 2020–2021. Wide phenotypic variation was recorded among the genotypes for all traits studied, reflecting the rich genetic diversity present in the germplasm panel comprising wild accessions, cultivated varieties and germplasm lines. Among the genotypes evaluated, the highest grain yield was recorded in Sarmatia (250.05 g/plot), followed by IBD-1 (234.55 g/plot), while the lowest grain yield was recorded in WR 41 (20.3 g/plot). This wide range in grain yield highlights the substantial phenotypic diversity existing within the panel and confirms the suitability of the germplasm for association mapping studies. The genotypes with superior grain yield such as Sarmatia and IBD-1 can serve as potential donor parents for future breeding programmes targeting drought-prone rainfed ecosystems. Such phenotypic variation is a prerequisite for successful association mapping and marker-assisted selection.

Genotypic analysis

Genomic DNA was isolated of 100 rice genotypes using leaf sample followed after CTAB method which is a common technique in molecular biology and genetics research (18). The concentration and purity of the extracted genomic DNA were assessed using a spectrophotometer, such as nanodrop. The DNA concentration is determined after measuring absorbance at 260 nm (A260) and purity is assessed by the A260/A280 and A260/A230 ratios. (NanoDrop, 2000c). Reaction mixture (20 ml) for PCR with gene specific 151 primers, DNA 25 ng/ul- 2.0ml, dNTPs (2.0 mM) 4.0ml, primer (10 mM) forward and reverse primers (F + R) 1.2ml, PCR reaction buffer (2 ×) 10.0ml, Taq DNA polymerase 0.5ml, sterile distilled H2O- 0.5ml, sterile distilled H2O- 2.3ml, the reaction

mixture was briefly centrifuged to ensure thorough mixing of all components. Then 0.20 mL PCR tubes were loaded in a thermal cycler. The diversity in banding patterns confirmed the polymorphic nature of the SSR markers and their ability to discriminate among the diverse genotypes. Five percent vertical polyacrylamide gels were employed to achieve improved separation and visualisation of small-sized PCR-amplified products, owing to their superior resolving power.

Allele diversity and population structure

Genetic association study among 100 accessions were analysed using 151 SSR markers spanning all rice chromosomes, a total of 414 alleles were amplified. The number of alleles per locus varied from 3 to 11, with an average of 2.77 alleles per locus. This allelic diversity reflects the broad genetic base of the germplasm panel, which included wild types, cultivated varieties and diverse germplasm accessions. A higher number of alleles per locus indicates greater discriminatory power of the markers and contributes to more reliable marker-trait association analysis. The allelic richness observed in the present study is consistent with previous reports where SSR markers amplified multiple alleles in diverse rice germplasm panels, confirming the utility of SSR markers for genetic diversity assessment and association mapping (19).

Polymorphism information content (PIC) of SSR markers

The relative information of each SSR marker was assessed using PIC values, which ranged from 0.019 (PPC-3 SET-3) to 0.498 (CS-62) with an average of 0.263 across all 151 markers (Table 2). The markers CS-62 (PIC = 0.498), RM28160 (PIC = 0.496), RGS-1 (PIC = 0.496) and RM-408 (PIC = 0.496) exhibited the highest PIC values, indicating their high polymorphism and strong discriminatory ability among the 100 rice genotypes (Fig. 3). Markers with PIC values close to 0.5 are classified as highly informative making them particularly suitable for genetic diversity studies and association analyses. In contrast, markers with PIC < 0.25 such as DREB-1 (0.020), GLN-2 (0.020) and NR-1 (0.020) were less informative, indicating limited polymorphism within the studied panel. The average PIC value of 0.263 observed in the present study indicates a medium level of genetic diversity, which is comparable to earlier observations where medium genetic diversity (PIC = 0.39) was reported among studied rice genotypes ranging from 0.516 to 0.92 with an average of 0.704, possibly due to a larger number of alleles detected in wider germplasm panels with greater genetic divergence (20).

Diversity in rice genotype

Phylogenetic analysis based on Nei's genetic distance revealed three major clusters among the 100 rice accessions (Fig. 4). Cluster I was the smallest, comprising 8 genotypes, while Cluster III contained 17 genotypes. Cluster II was the largest, with two sub-groups accommodating 13 and 63 genotypes respectively, totalling 76 accessions. The grouping pattern largely separated the wild accessions from the cultivated varieties and germplasm lines, reflecting their evolutionary distances. The large number of genotypes clustering in Cluster II sub-groups suggests that most of the cultivated landraces and improved varieties share a closer genetic background, which is expected given the common geographic origin of the Chhattisgarh rice germplasm. These findings are broadly consistent with earlier cluster analysis (21), in which Cluster II comprised 75 accessions in a study of Indian rice landraces and improved varieties. The phylogenetic clustering also validates the selection of a genetically diverse panel for association mapping, since marker-trait associations identified across genetically distinct sub-groups are more likely to represent true

Table 2. List of simple sequence repeat markers with their polymorphism information content value and number of alleles

Sl. No.	Markers	No. of alleles	PIC Value	Sl. No.	Markers	No. of alleles	PIC Value
1	RM-1	8	0.20	76	PPC-2A-SET-5	3	0.39
2	RM-5	1	0.40	77	PPC-2B SET-2	2	0.26
3	RM-11	11	0.31	78	PPC-2B SET-4	1	0.11
4	RM-19	6	0.25	79	PPC-2B SET-5	2	0.30
5	RM-25	3	0.37	80	PPC-3 SET-3	1	0.02
6	RM-125	1	0.13	81	RGS-1	2	0.50
7	RM-154	4	0.26	82	SEQ-3-1	2	0.08
8	RM-162	2	0.16	83	GW-8 band 1	1	0.28
9	RM-171	2	0.49	84	GW-8 band 2	1	0.16
10	RM-178	2	0.25	85	18640	3	0.32
11	RM-201	4	0.26	86	18646	3	0.34
12	RM-142	4	0.31	87	18654	2	0.30
13	RM259	5	0.27	88	18655	4	0.13
14	RM-271	2	0.45	89	24579	1	0.23
15	RM-283	3	0.16	90	RM 11940	1	0.10
16	RM-284	2	0.30	91	RM 12003	4	0.33
17	RM-307	3	0.27	92	RM 12010	1	0.06
18	RM-312	2	0.44	93	RM 12020	1	0.35
19	RM-316	4	0.25	94	RM 12091	5	0.33
20	RM-408	2	0.50	95	RM 124	1	0.11
21	RM-413	4	0.34	96	RM 1261	5	0.24
22	RM-433	3	0.31	97	RM 13	4	0.34
23	RM-435	1	0.06	98	RM 136	4	0.29
24	RM-454	2	0.24	99	RM 1361	1	0.46
25	RM-484	1	0.39	100	RM 161	2	0.23
26	RM-540	3	0.47	101	RM 18626	1	0.20
27	RM-574	4	0.28	102	RM 18639	3	0.35
28	RM-1233	4	0.30	103	RM 18641	2	0.28
29	RM-22710	3	0.36	104	RM 21	6	0.23
30	OSR-13	4	0.31	105	RM 212	3	0.40
31	Xa-13 Pro	2	0.05	106	RM 219	3	0.42
32	RM 3739 band 1	4	0.33	107	RM 236	2	0.44
33	RM 505 band 1	4	0.14	108	RM 237	5	0.25
34	RM 24569 band 2	6	0.17	109	RM 240	4	0.33
35	RM 27489 band 1	2	0.38	110	RM 24568	5	0.20
36	RM 27489 band 2	2	0.31	111	RM 24571	3	0.34
37	RM 18652 band 1	3	0.39	112	RM 24575	1	0.06
38	RM 18652 band 2	2	0.08	113	RM 27492	5	0.32
39	AMT 3-2	1	0.11	114	RM 27493	9	0.18
40	AMT-4	1	0.10	115	RM 28048	2	0.45
41	APO-1	2	0.10	116	RM 28076	3	0.28
42	CS-62	2	0.50	117	RM 28086	3	0.39
43	DREB-1	1	0.02	118	RM 28089	3	0.31
44	DREB-2	1	0.27	119	RM 28099	2	0.39
45	DREB-2B	2	0.33	120	RM 28130	3	0.35
46	DREB-3	1	0.06	121	RM 28157	4	0.30
47	GL	1	0.08	122	RM 28160	2	0.50
48	GLN-1-2	1	0.04	123	RM 28199	2	0.37
49	GLN-2	1	0.02	124	RM 3412	6	0.23
50	GLT-2	2	0.09	125	RM 3483	4	0.31
51	GS-5	1	0.28	126	RM 3520	5	0.25
52	GW-5	1	0.10	127	RM 3825	4	0.25
53	IND	2	0.47	128	RM 416	2	0.44
54	IND-1	2	0.40	129	RM 431	3	0.30
55	IND-10	2	0.48	130	RM 44	2	0.41
56	N-1	1	0.15	131	RM 447	4	0.26
57	N-12	2	0.46	132	RM 455	2	0.12
58	N-8	2	0.49	133	RM 472	4	0.28
59	N-16	1	0.04	134	RM 474	6	0.19
60	N-18	1	0.06	135	RM 486	3	0.36
61	N 2.1-E	1	0.06	136	RM 511	4	0.25
62	N2-1-P	1	0.02	137	RM 518	3	0.37
63	N2-3B-2	1	0.18	138	RM 520	4	0.29
64	NOD-1	4	0.30	139	RM 547	5	0.27
65	NOD-3	4	0.15	140	RM 561	2	0.49
66	NIR-2	3	0.33	141	RM 574	3	0.32
67	NR-1	1	0.02	142	RM 5759	1	0.15
68	NR-1-102178	2	0.49	143	RM 6504	3	0.33
69	NR2-121810	1	0.02	144	RM 7195	7	0.18
70	NRT 1-2	1	0.02	145	RM 7311	5	0.26
71	NRT2.3B2	1	0.20	146	RM 8215	4	0.34
72	OSGLU-1	1	0.13	147	RM 85	5	0.29
73	PPC 3 SET 1	2	0.41	148	RM-42	2	0.45
74	PPC-1 SET 1	3	0.12	149	RM-515	5	0.25
75	PPC-1 SET-3	1	0.10				

Note: PIC: polymorphism information content.

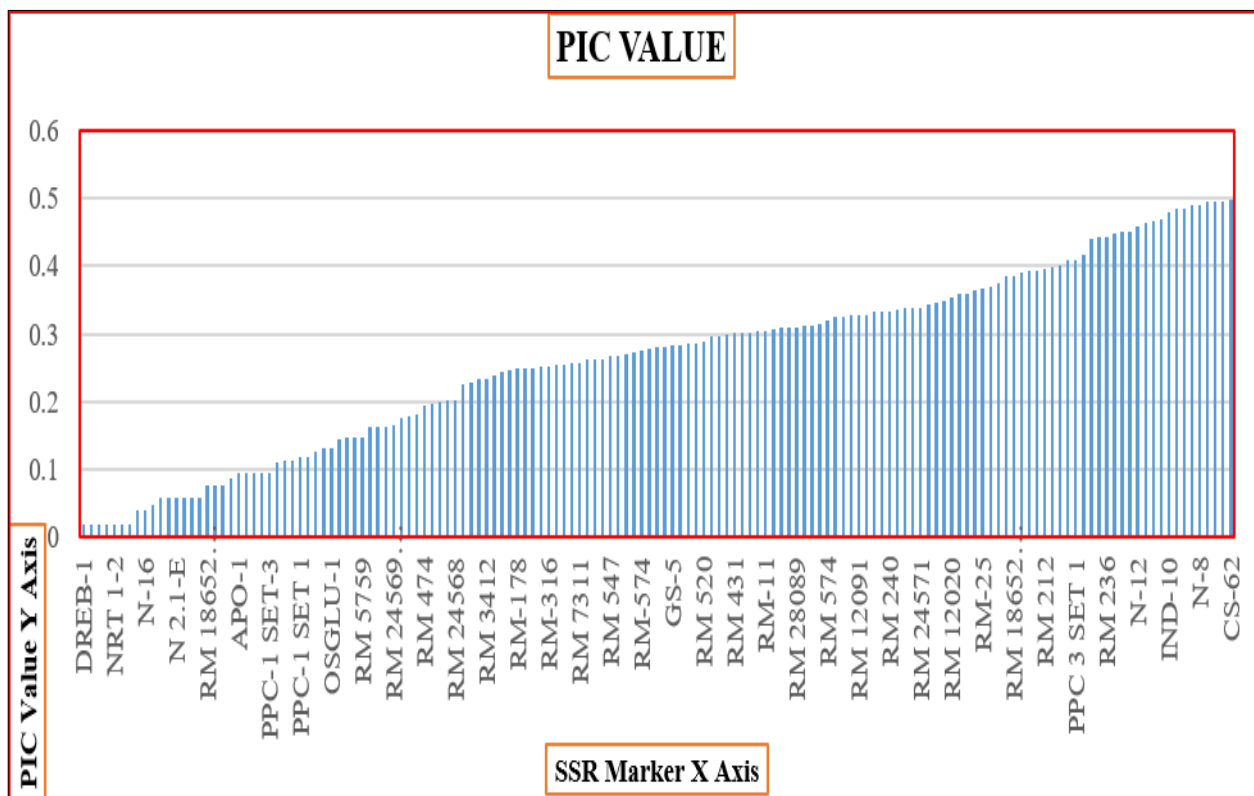


Fig. 3. Graphical representation of polymorphism information content value of polymorphic simple sequence repeat marker.

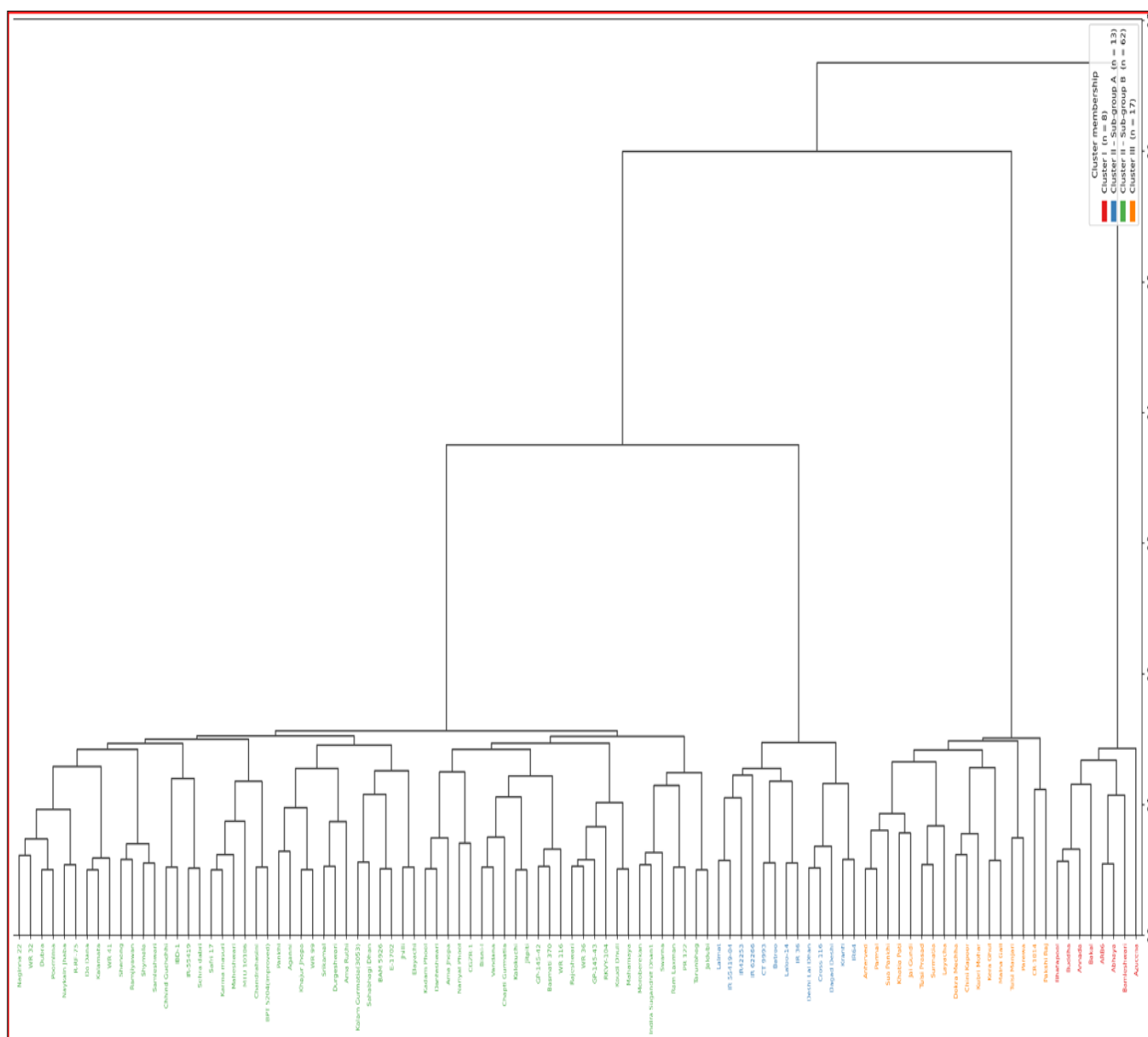


Fig. 4. Phylogenetic tree based on Nei's genetic distance has three major clusters perform.

associations rather than population-structure-induced artefacts.

Structure population

Analysis based on structure program carried out genetic diversity analysis revealed an optimal K value of 3, classifying the rice genotypes into three clusters corresponding to wild, cultivated and germplasm types (Fig. 4 and 5). Out of 100 genotypes, 85 genotypes come under POP1 (population-1), 13 genotypes (Sl. No. 67, 68, 69, 71, 73, 75, 83, 85, 88, 91, 92, 97, 98) remaining comes under POP2 comprised of 2 genotypes (Sl. No. 99, 100) comes under POP3 Fig. 5. A bar plot illustrating the diversity of rice accessions with different colours likely representing different genetic clusters or groups. A ternary plot showing the diversity at $k = 3$ obtained using structure software. Each point represents a rice accession and its position within the triangle indicates a line plot depicting the diversity in rice genotypes, obtained from structure harvester. This graph is used to determine the optimal number of genetic clusters in the dataset (Fig. 6). Genetic variability plays a central role in breeding programs. Insights from germplasm evaluation of rice landraces and the genetic analysis of 18 traits provided the rationale for conducting the current study. The Q matrix derived from this analysis was used as a covariate in the GLM model to correct for population stratification, while both Q and kinship (K) matrices were incorporated in the MLM model to further minimise false-positive associations. The identification of three sub-populations in the present study is in agreement with previous studies reporting population structures ranging from 2 to 8 sub-populations in rice (22). Earlier studies have detected two and four subgroups in study populations of 425 and 91 rice landrace accessions respectively (23), indicating that the number of sub-populations varies with the breadth of germplasm sampled.

Marker-trait association

In this study, 151 SSR markers were analysed, revealing their associations with various traits. A total of 90 and 42 marker-trait associations were identified using the GLM and MLM models respectively (Tables 2 and 3). Among these, 10 markers were found to be associated with multiple traits. Researchers identified several markers associated with rice yield-related traits, including genomic regions on chromosomes 1, 4, 8 and 9 linked to grain yield under stress conditions, as well as a novel QTL, DRO3, for root growth angle (RGA) located on chromosome 7 (24).

General linear model

Large genotypic and phenotypic variation was observed among the diverse rice accessions. Genome wide association analysis was carried out using the software TASSEL v 3.0.2. Association analysis was performed using both GLM and MLM employing population parameters previously determined (P3D) method (Table 4). Based on the p and R^2 values, in this study the 151 SSR markers were analysed for 100 rice accessions were found to be significant in GLM approach and the number of associations substantially reduced to while applying MLM approach with a false discovery rate of 50.1 % over GLM. Based on Nei's genetic distance, neighbour joining tree was constructed using TASSEL 3.0 (Table 5). Marker-trait association genome wide association analysis revealed out of 90 GLM and 52 markers associated traits at > 0.05 level of significance show (black colour) and 27 markers associated < 0.05 highly significant show (green colour) and < 0.01 show 11 marker were very highly significant the table indicate red colour. This is further evidenced by the false discovery rate (FDR) of 50.1 % in GLM over MLM, emphasising the need to apply the more

stringent MLM approach for reliable marker identification. Similar observations were reported in previous research (25), who detected numerous marker-trait associations using GLM in diverse rice cultivars, many of which were not validated under more stringent models. The GLM model, which accounts for population structure through the Q matrix, effectively reduced spurious associations attributable to sub-population stratification. The higher number of associations detected under GLM compared to MLM is consistent with the general observation that GLM tends to inflate associations when cryptic relatedness among individuals is not fully accounted for (26).

Mixed linear model

The total 42 markers identified and 7 markers associated traits at > 0.05 level of significance shows (black colour) and 33 markers associated < 0.05 highly significant shows (green colour) and < 0.01 show 2 marker highly significant which highlighted red colour in the table, associated (p -value < 0.05) with respective trait. Two markers identified which are significantly (Table 6) associated (p -value > 0.05) highlighted in red colour. Comparison of the present results with earlier rice association studies showed that 16 and 10 SSR markers were significantly associated with economic selection value (ESV) traits using the general linear model (GLM) and mixed linear model (MLM) respectively. Marker alleles located on chromosome 2 were linked to shoot dry weight at 28 days after sowing (DAS) and vigour index at both 14 and 28 DAS. This is consistent with earlier association studies in rice which showed that MLM substantially reduced spurious associations while retaining biologically meaningful marker-trait relationships (27). The present results are also comparable to a previous study (28), which reported 16 and 10 SSR markers significantly associated with early seedling vigour (ESV) traits using GLM and MLM respectively, with marker alleles on chromosome 2 linked to shoot dry weight and vigour index, confirming that MLM produces a more conservative and reliable set of associations for practical use in breeding. Enhancing seedling growth rate is therefore valuable for identifying rice genotypes adapted to direct-seeded conditions through marker-assisted selection (29).

Multi-trait associated marker

Among all identified associations, 10 SSR markers RM-284, RM-28086, RM-3412, RM-307, PPC-2, RM-515, RM-201, RM-7195, RM-171 and CS-62 demonstrated significant association with multiple traits simultaneously under both GLM and MLM models (Table 4). Multi-trait associated markers are particularly valuable in plant breeding because they can simultaneously improve multiple agronomic traits through a single selection event, thereby enhancing the efficiency of marker-assisted selection (MAS) programmes. Such pleiotropic marker associations may indicate that these loci are located within genomic regions controlling metabolic pathways common to multiple yield-determining processes, such as carbon assimilation, panicle architecture and grain filling. These grain-yield-linked markers are of direct practical significance for rice improvement under rainfed conditions. The marker RM-242 has previously been associated with root length in rice, while RM-511, RM-1261 and RM-28130 have demonstrated significant relationships with grain yield under terminal drought conditions (30), corroborating the relevance of the markers identified in the present study. Similarly, a previous study (31) identified novel QTLs for drought tolerance at chromosomal positions overlapping with regions highlighted in the present study, further validating the biological relevance of these associations. The CS-62, which showed both the highest PIC value (0.498) and association with multiple yield traits including grain yield, emerges as a particularly

Table 3. List of markers

Sl. No.	Markers	Sl. No.	Markers	Sl. No.	Markers	Sl. No.	Markers
1	RM-1	39	AMY-4	77	PPC-1 SET-3	115	RM 27493
2	RM-5	40	APO-1	78	PPC-2A-SET-5	116	RM 28048
3	RM-11	41	CS-62	79	PPC-2B SET-2	117	RM 28076
4	RM-19	42	DREB-1	80	PPC-2B SET-4	118	RM 28086
5	RM-25	43	DREB-2	81	PPC-2B SET-5	119	RM 28089
6	RM-125	44	DREB-2B	82	PPC-3 SET-3	120	RM 28099
7	RM-154	45	DREB-3	83	RGS-1	121	RM 28130
8	RM-162	46	GL	84	SEQ-3-1	122	RM 28153
9	RM-171	47	GLN-1-2	85	GW-8	123	RM 28157
10	RM-178	48	GLN-2	86	18640	124	RM 28160
11	RM-201	49	GLT-1	87	18646	125	RM 28199
12	RM-142	50	GLT-2	88	18654	126	RM 3412
13	RM-259	51	GS-5	89	18655	127	RM 3483
14	RM-271	52	GW-5	90	24579	128	RM 3520
15	RM-283	53	IND	91	RM 11940	139	RM 3825
16	RM-284	54	IND-1	92	RM 12003	130	RM 416
17	RM-307	55	IND-10	93	RM12010	131	RM 431
18/	RM-312	56	N-1	94	RM12020	132	RM 44
19	RM-316	57	N-12	95	RM12091	133	RM 447
20	RM-408	58	N-8	96	RM 124	134	RM 455
21	RM-413	59	N-16	97	RM 1261	135	RM 472
22	RM-433	60	N-18	98	RM 13	136	RM 474
23	RM-434	61	N 2.1 E	99	RM 136	137	RM 486
24	RM-454	62	N2-1-P	100	RM 1361	138	RM 511
25	RM-484	63	N2-3A-1	101	RM 161	139	RM 518
26	RM-540	64	N2-3B-2	102	RM 18626	140	RM 520
27	RM-574	65	NOD-1	103	RM 18639	141	RM 547
28	RM-1233	66	NOD-3	104	RM 18641	142	RM 561
29	RM-22710	67	NIR-2	105	RM 21	143	RM 574
30	OSR-13	68	NR-1	106	RM 212	144	RM 5759
31	Xa-13Pro	69	NR-1-102178	107	RM 219	145	RM 6504
32	RM3739	70	NR2-121810	108	RM 236	146	RM 7195
33	RM-505	71	NRT 1-2	109	RM 237	147	RM 7311
34	RM-24569	72	NRT2.3B2	110	RM 240	148	RM 8215
35	RM-27489	73	OSGLU-1	111	RM 24568	149	RM 85
36	RM-18652	74	OSGLU-2	112	RM 24571	150	RM-42
37	AMT 3-2	75	PPC 3 SET 1	113	RM 24575	151	RM 515
38	AMT 2-2	76	PPC-1 SET 1	114	RM 27492		

Table 4. Significant marker-trait associations based on general linear model and mixed linear model ($p < 0.05$) with multiple traits

Trait	locus	C#	p-marker	Req. marker
PH, PL, RL, RW, GWI, GW, GL	RM284	8	0.02	0.15
FLL, PH, PL, RL, RW, FLW, SP, FP, GWI, GW, GL	RM28086	12	0.02	0.19
FLL, PL, RW, GWI, GL	RM3412	1	0.37	0.44
PH, PL, RL, RW, SP, FP, GW, GL, GWI	PPC-2SET2	9	0.01	0.24
GL, GW, GWI, RW, RL, PL, FLL	RM515	8	0.0	0.65
DFF, FP, SP	RM307	4	0.01	0.28
DTF, FP, SP	RM201	9	0.01	0.27
SP, FP	RM7195	12	0.02	0.36
FLL, PH, PL, SP, FP	RM171	10	0.04	0.18
TY, TPP, BY	CS-62	5	0.05	0.10

Note: C# = Chromosome number; p = Marker p-value; Req. marker = The representative marker for that association; DTF: days to flowering; PH: plant height (cm); FLL: flag leaf length (cm); FLW: flag leaf width (cm); PL: panicle length (cm); TPP: total panicles per plot; NFG: number of filled grains; NUFG: number of unfilled grains; TNG: total number of grains; FP: fertility percentage; SP: sterility percentage; BYP: biological yield per plot (g); GYP: grain yield per plot (g); HGW: hundred grain weight (g); GL: grain length (cm); GW: grain width (cm); RL: rice length (cm); RW: rice width (cm).

Table 5. Significant marker trait association based on general linear model

Trait	Locus	C#	p-marker	R ² marker	Trait	Locus	C#	p- marker	R ² marker
DTF	RM-201	9	0.008	0.1965		RM-242	9	0.09	0.2442
	RM-307	4	0.005	0.1749		IND	4	0.198	0.2194
	RM-408	4	0.018	0.1462		IND-1	5	0.081	0.3238
	RM-219	9	0.037	0.1593		PPC-1SET1	2	0.106	0.2089
	RM-28076	12	0.008	0.2460	BY	18646	5	0.042	0.2374
	RM-28089	12	0.002	0.2384		RM-316	9	0.108	0.3150
	RM-171	10	0.293	0.2056		RM-154	2	0.062	0.2537
PH	PPC-2B-SET-2	9	0.271	0.2119		RM-178	5	0.144	0.1485
	RM-271	10	0.014	0.1646		RM259	1	0.257	0.2028
	RM-284	8	0.043	0.1396	NFS	RM-237	1	0.131	0.2392
	RM-408	8	0.001	0.2117	NUS	RM-505band1	7	0.062	0.1386
	RM-413	5	0.063	0.1763		18640	5	0.058	0.1808
	RM-212	1	0.001	0.2650		RM-171	10	0.042	0.1757
	RM-28086	12	0.024	0.1897		RM-201	9	0.024	0.2487
	RM-472	1	0.043	0.2231	SFP	RM-307	4	0.037	0.2199
	RM-486	1	0.034	0.1938		PPC-2BSET-2	9	0.008	0.2134
	RM-6504	1	0.053	0.2007		RM-28086	12	0.142	0.1890
FLL	RM-8215	12	0.001	0.3275		RM-7195	12	0.023	0.3636
	RM-171	10	0.152	0.2262		RM-171	10	0.036	0.1795
	RM-28086	12	0.074	0.3649		RM-201	9	0.01	0.2724
	RM-515	8	0.142	0.384		RM-307	4	0.007	0.2751
	RM-28086	12	0.119	0.3455	SSP	PPC-2BSET-2	9	0.005	0.2439
FLW	RM-515	8	0.17	0.3746		RM-28086	12	0.071	0.2142
	RM-21	11	0.017	0.2825		RM-28089	12	0.131	0.1906
PL	RM-171	10	0.314	0.2641		RM-7195	12	0.131	0.4158
	PPC-2B-SET-2	9	0.355	0.2465		RM-284	8	0.302	0.2628
	RM-28086	12	0.072	0.6027		PPC-3SET-1	1	0.363	0.2553
	RM-3412	1	0.517	0.3844		PPC-2B-SET-2	9	0.286	0.2665
	RM-515	8	0.119	0.5974	GW	RM-28086	12	0.004	0.7092
TPP	RM-284	8	0.021	0.1511		RM-3412	1	0.462	0.3932
	RM-312	1	0.017	0.1536		RM-515	8	0.035	0.6534
	APO-1	6	0.025	0.1353		RGS-1	3	0.08	0.1165
	RM-178	5	0.064	0.1284		RM-413	5	0.063	0.1761
GW	RM-284	8	0.283	0.2240		RM-284	8	0.315	0.2683
	PPC-2B-SET-2	9	0.269	0.2246		PPC-2B-SET-2	9	0.302	0.2679
	RM-28086	12	0.011	0.5394		RM-28086	12	0.047	0.6595
	RM-515	8	0.01	0.5904		RM-3412	1	0.369	0.4412
	RM-284	8	0.335	0.2737	RL	RM-515	8	0.044	0.7084
GL	PPC-2B-SET-2	9	0.324	0.2738		RM-284	8	0.314	0.2329
	RM-28086	12	0.039	0.6928		PPC-2B-SET-2	9	0.29	0.2407
	RM-3412	1	0.356	0.4536		RM-28086	12	0.049	0.5369
	RM-472	1	0.479	0.3968		RM-3412	1	0.476	0.3795
GY	RM-515	8	0.044	0.7184		RM-515	8	0.05	0.5622
	18646	5	0.17	0.1565	TSP	RM-44	8	0.997	0.0551
	RM-21	11	0.049	0.2923		RM-237	1	0.995	0.1148

Note: GLM: general linear model; DTF: days to flowering; PH: plant height (cm); FLL: flag leaf length (cm); FLW: leaf width (cm); PL: panicle length (cm); TPP: total panicles per plant; GW: grain width (mm); GL: grain length (mm); GY: grain yield (g/m²); BY: biological yield (g/m²); NFS: number of filled spikelets; NUS: number of unfilled spikelets; SFP: spikelet fertility percentage; SSP: spikelet sterility percentage; GWL: grain weight index; RL: rice length (mm); TSP: total spikelets per panicle.

Table 6. Significant marker-trait associations based on mixed linear model (Q + K) model ($p < 0.05$)

Trait	Marker	C#	p- marker	R ² marker	Trait	Marker	C#	p- marker	R ² marker
DTF	RM-28160	12	0.01	0.05	BY	18646	5	0.002	0.26
	RM-8215	12	0.05	0.11		RM-312	1	0.02	0.11
PH	RM-408	8	0.05	0.08		CS-62	5	0.05	0.10
	RM-212	1	0.05	0.09		RM-547	8	0.06	0.17
FLL	RM-28086	12	0.05	0.42	GY	RM-242	9	0.01	0.15
	RM-28086	12	0.05	0.36		RM-178	5	0.02	0.11
FLW	RM-28086	12	0.05	0.35		PPC-1SET1	22	0.02	0.15
	RM-8215	12	0.03	0.16		CS-62	5	0.03	0.09
PL	RM-28199	12	0.05	0.12	GW	RM-154	2	0.03	0.16
	RM-515	8	0.05	0.60		IND-1	5	0.03	0.12
	RM-416	3	0.04	0.09		RM-515	8	0.02	0.59
TPP	RM-178	5	0.04	0.09		RM-28086	12	0.02	0.54
	CS-62	5	0.05	0.10	RL	RM-518	4	0.04	0.14
	RM-12091	1	0.06	0.18		RM-11	7	0.06	0.12
SSP	RM-505	7	0.03	0.14		RM-515	8	0.02	0.71
	RM-547	8	0.03	0.19		RM-28086	12	0.02	0.66
GL	RM-3739	12	0.05	0.23	RW	RM-413	5	0.04	0.21
	RM-11	7	0.08	0.19		RGS-1	3	0.03	0.09
	RM-28086	12	0.03	0.69		RM-3739	12	0.27	0.17
GWI	RM-515	8	0.03	0.72		RM-515	8	0.04	0.56
	RM-28086	12	0.004	0.71					
	RM-515	8	0.04	0.65					

Note: MLM: mixed linear model; DTF: days to flowering; PH: plant height (cm); FLL: flag leaf length (cm); FLW: leaf width (cm); PL: panicle length (cm); TPP: total panicles per plant; SSP: spikelet sterility percentage; GL: grain length (mm); GWI: grain weight index; BY: biological yield (g/m²); GY: grain yield (g/m²); GW: grain width (mm); RL: rice length (mm); RW: rice width (mm).

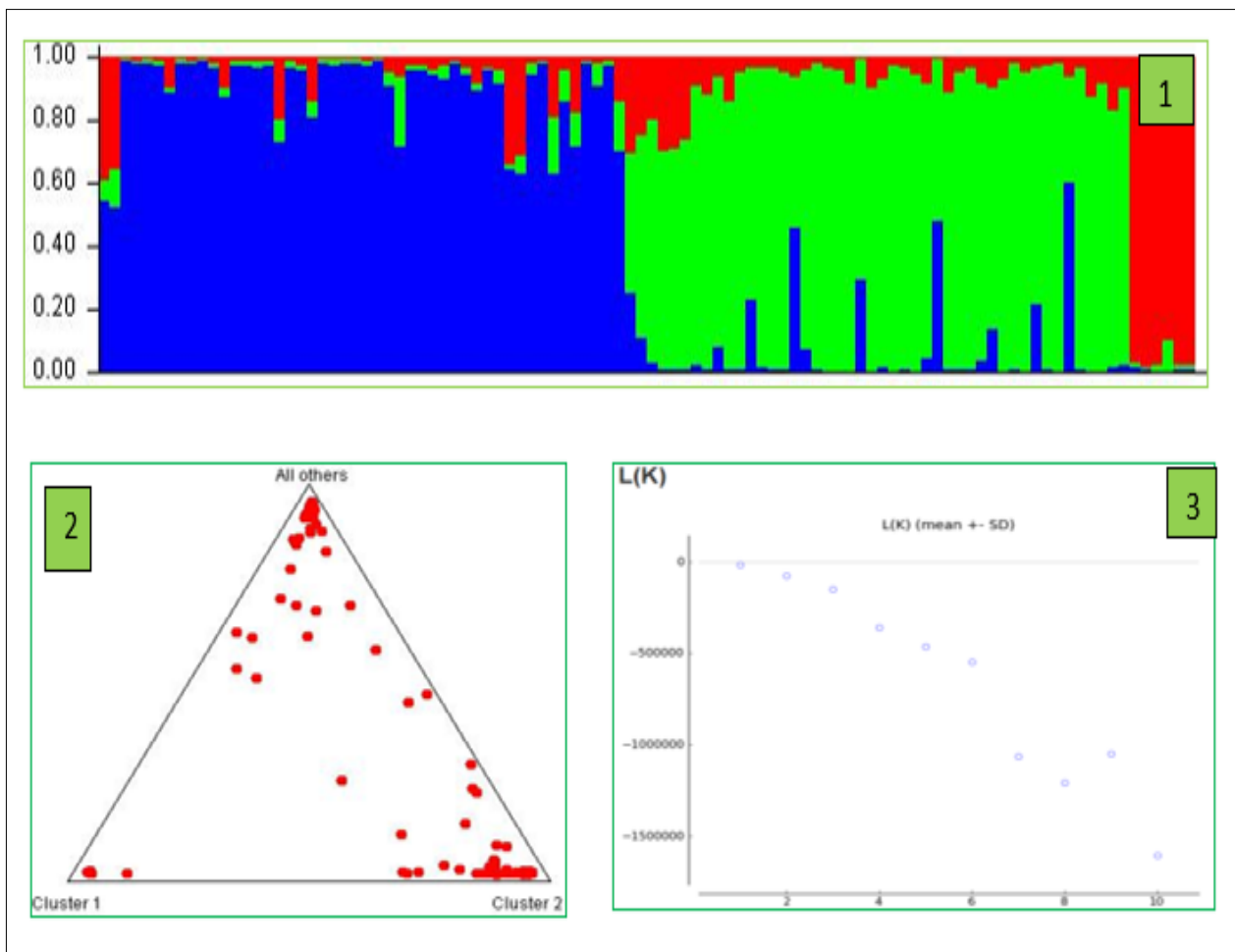
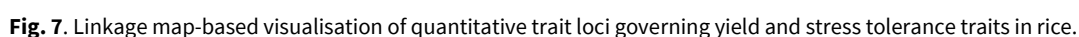
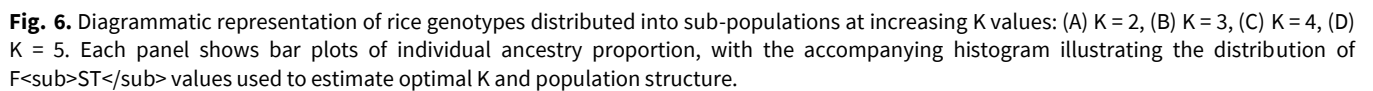


Fig. 5. Illustration of genetic diversity among rice accessions using STRUCTURE-based visualization: (1) bar plot showing individual ancestry proportions/genetic clusters; (2) a ternary plot displaying genome proportions among three clusters (K = 3); (3) a line plot of L(K) values used to identify the optimal number of clusters.



promising marker for MAS in drought-prone rice environments.

Linkage map

The chromosomal locations of markers were plotted on a linkage map constructed using QTL cartographer presented in Fig. 7. Eight SSR markers RM-242, RM-178, RM-154, RM-21, 18646, IND-1, CS-62 and PPC-1 SET1 were found to be tightly linked with grain yield and covered the entire linkage groups. The researchers highlighted that the high-density linkage map is useful for the identification of quantitative trait loci responsible for various traits like yield, stress tolerance, disease and insect resistance and other traits in rice and many other crops (32). These markers, particularly those tightly linked to grain yield and stress-tolerance traits, will be useful in assembling favourable alleles into elite local lines through marker-assisted breeding to develop superior drought-tolerant varieties for rainfed Chhattisgarh conditions (33).

Conclusion

The present study revealed considerable genetic variation among the rice genotypes evaluated under rainfed condition indicating significant scope for improving drought tolerance and grain yield. From the genotypes tested, Sarmatia and IBD-1 were the best in terms of grain yield and could be used as valuable donor parents. Association mapping using Simple Sequence Repeat (SSR) markers identified several significant marker-trait associations under both general linear model (GLM) and mixed linear model (MLM) models and markers RM-242, RM-178, RM-154, RM-21, 18646, IND-1, CS-62 and PPC-1SET1 showed strong association with grain yield and related traits. Multi-trait associated markers identified in this study can be utilised in marker-assisted selection for developing high-yielding and stress-tolerant rice varieties. Overall, the findings provide valuable genetic resources and molecular tools for accelerating rice improvement under rainfed conditions.

Acknowledgements

The authors gratefully acknowledge the technical assistance provided by the Department of Plant Molecular Biology and Biotechnology, Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh, India.

Authors' contributions

Financial and technical assistance for this study were provided by the Department of Plant Molecular Biology and Biotechnology, Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh, India. S, SKV and SBV carried out the genotyping and phenotyping experiments, performed statistical analysis and contributed to drafting the manuscript. SKV conceptualised the study, supervised the genotyping and phenotyping work and contributed to manuscript preparation. S, SKV, SY and VKG assisted in the review and editing of the manuscript. All authors read and approved the final manuscript.

Compliance with ethical standards

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

Ethical issues: None

References

1. Ministry of Agriculture & Farmers Welfare, Government of India. Second advance estimates of production of major agricultural crops 2024–25. New Delhi: Government of India; 2025.
2. Awad-Allah MMA, Shafie WWM, Alsubeie MS, Alatawi A, Safhi FA, ALshamrani SM, et al. Utilisation of genetic resources, genetic diversity and genetic variability for selecting new restorer lines of rice (*Oryza sativa* L.). *Genes* (Basel). 2022;13(12):2227. <https://doi.org/10.3390/genes13122227>
3. Sweeney M, McCouch S. The complex history of the domestication of rice. *Ann Bot*. 2007;100(5):951–7. <https://doi.org/10.1093/aob/mcm048>
4. Saini PS. Candidate gene-based association mapping of rice (*Oryza sativa* L.) germplasm accessions for qDTY1.1 and qDTY12.1. Raipur: Indira Gandhi Krishi Vishwavidyalaya; 2020.
5. Zeng L, Kwon TR, Liu X, Wilson C, Grieve CM, Gregorio GB. Genetic diversity analysed by microsatellite markers among rice (*Oryza sativa* L.) genotypes with different adaptations to saline soils. *Plant Sci*. 2004;166(5):1275–85. <https://doi.org/10.1016/j.plantsci.2004.01.005>
6. Yi Y, Hassan MA, Cheng X, Li Y, Liu H, Fang W, et al. QTL mapping and analysis for drought tolerance in rice by genome-wide association study. *Front Plant Sci*. 2023;14:1223782. <https://doi.org/10.3389/fpls.2023.1223782>
7. Hassan MA, Dahu N, Hongning T, Qian Z, Yueming Y, Yiru L, et al. Drought stress in rice: morpho-physiological and molecular responses and marker-assisted breeding. *Front Plant Sci*. 2023;14:1215371. <https://doi.org/10.3389/fpls.2023.1215371>
8. Nei M. Genetic distance between populations. *Am Nat*. 1972;106(949):283–92. <https://doi.org/10.1086/282771>
9. Doyle JJ, Doyle JL. Isolation of plant DNA from fresh tissue. *Focus*. 1990;12:13–5. <https://doi.org/10.2307/2419362>
10. Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME. Optimizing parental selection for genetic linkage maps. *Genome*. 1993;36(1):181–6. <https://doi.org/10.1139/g93-024>
11. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet*. 1980;32(3):314–31.
12. Liu K, Muse SV. PowerMarker: integrated analysis environment for genetic marker data. *Bioinformatics*. 2005;21(9):2128–9. <https://doi.org/10.1093/bioinformatics/bti282>
13. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155(2):945–59. <https://doi.org/10.1093/genetics/155.2.945>
14. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 2005;14(8):2611–20. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
15. Earl DA, VonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour*. 2012;4(2):359–61. <https://doi.org/10.1007/s12686-011-9548-7>
16. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*. 2007;23(19):2633–5. <https://doi.org/10.1093/bioinformatics/btm308>
17. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Stat Methodol*. 1995;57(1):289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
18. Green MR, Sambrook J. Polyacrylamide gel electrophoresis. *Cold*

- Spring Harb Protoc. 2020;(12). <https://doi.org/10.1101/pdb.prot100412>
19. Jin L, Lu Y, Xiao P, Sun M, Corke H, Bao J. Genetic diversity and population structure of a diverse set of rice germplasm for association mapping. *Theor Appl Genet.* 2010;121(3):475–87. <https://doi.org/10.1007/s00122-010-1324-7>
 20. Donde R, Mohapatra S, Baksh SY, Padhy B, Mukherjee M, Roy S, et al. Identification of QTLs for high grain yield and component traits in new plant types of rice. *PLoS One.* 2020;15(7):e0227785. <https://doi.org/10.1371/journal.pone.0227785>
 21. Kumbhar SD, Kulwal PL, Patil JV, Sarawate CD, Gaikwad AP, Jadhav AS. Genetic diversity and population structure in landraces and improved rice varieties from India. *Rice Sci.* 2015;22(3):99–107. <https://doi.org/10.1016/j.rsci.2015.05.013>
 22. Courtois B, Frouin J, Greco R, Bruschi G, Droc G, Hamelin C, et al. Genetic diversity and population structure in a European collection of rice. *Crop Sci.* 2012;52(4):1663–75. <https://doi.org/10.2135/cropsci2011.11.0588>
 23. Das B, Sengupta S, Prasad M, Ghose TK. Genetic diversity and population structure of rice landraces from eastern and north-eastern India using SSR markers. *Mol Biol Rep.* 2013;40(2):1119–30. <https://doi.org/10.1007/s11033-012-2197-3>
 24. Uga Y, Okuno K, Yano M. Fine mapping of Sta1, a quantitative trait locus determining stele transversal area, on rice chromosome 9. *Mol Breed.* 2010;26(3):533–8. <https://doi.org/10.1007/s11032-010-9450-0>
 25. Agrama HA, Eizenga GC, Yan W. Association mapping of yield and its components in rice cultivars. *Mol Breed.* 2007;19(4):341–56. <https://doi.org/10.1007/s11032-006-9066-6>
 26. Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet.* 2006;38(2):203–8. <https://doi.org/10.1038/ng1702>
 27. Nachimuthu VV, Robin S, Sudhakar D, Raveendran M, Rajeswari S, Manonmani S. Evaluation of rice genetic diversity and population structure using SSR markers and association analysis for grain yield traits. *J Genet.* 2015;94(2):215–26. <https://doi.org/10.1007/s12041-015-0521-3>
 28. Anandan A, Anumalla M, Pradhan SK, Ali J. Population structure, diversity and trait association analysis in rice (*Oryza sativa* L.) germplasm for early seedling vigour using trait linked SSR markers. *PLoS One.* 2016;11(3):e0152406. <https://doi.org/10.1371/journal.pone.0152406>
 29. Zhou L, Wang JK, Yi Q, Wang YZ, Zhu YG, Zhang ZH. Quantitative trait loci for seedling vigour in rice under field conditions. *Field Crops Res.* 2006;100:294–301. <https://doi.org/10.1016/j.fcr.2006.08.004>
 30. Freeg HA, Anis GB, Abo-Shousha AA, El-Banna AN, Sabagh AE. Genetic diversity among some rice (*Oryza sativa* L.) genotypes with different drought tolerance based on SSR markers. *Cercet Agron Mold.* 2016;49(3):39–50. <https://doi.org/10.1515/cerce-2016-0024>
 31. Ghazy MI, EL-Naem SA, Hefeina AG, et al. Genome-wide association study of rice diversity panel reveals new QTLs for tolerance to water deficit under the Egyptian conditions. *Rice.* 2024;17:29. <https://doi.org/10.1186/s12284-024-00703-1>
 32. Tanksley SD. Mapping polygenes. *Annu Rev Genet.* 1993;27(1):205–33. <https://doi.org/10.1146/annurev.ge.27.120193.001225>
 33. Emon RM. Molecular markers, marker assisted selection for rice in relation to biotic and abiotic stress. *Discov Agric.* 2025;3:102. <https://doi.org/10.1007/s44279-025-00265-w>

Additional information

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

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

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

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

Copyright: © The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited (<https://creativecommons.org/licenses/by/4.0/>)

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.