Indian Journal of Agricultural Sciences **93** (4): 365–370, April 2023/Article
https://doi.org/10.56093/ijas.v93i4.133177

SSR marker-based genetic diversity and marker-trait association analysis in aromatic rice (*Oryza sativa*) landraces

MAUMITA BURMAN[1]*, SUNIL KUMAR NAIR[1], ABHINAV SAO[1], DEEPAK GAURAHA[1] and DEEPAK SHARMA[1]

Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh 492 012, India

Received: 12 February 2023; Accepted: 27 March 2023

ABSTRACT

The present study on SSR marker-based genetic diversity and marker-trait association analysis in aromatic rice (*Oryza sativa* L.) landraces was carried out at R. H. Richharia Research Laboratory, Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh during 2020. A total of 25 PCR-based simple sequence repeats (SSR) markers were evaluated in a set of 90 aromatic rice landraces along with 6 checks which includes 1 non-aromatic and 5 aromatic check varieties. Phenotypic data for marker-trait association analysis were taken for 24 yield attributing traits. Polymorphic Information Content (PIC) value ranged from 0.52 (RM316) to 0.79 (RM553) with a mean of 0.69 which reveals that all the markers used in this study were highly informative and useful for diversity analysis of a wide range of genotypes. Unweighted Pair Group Method with Arithmetic Mean (UPGMA) and Rho's similarity based cluster analysis revealed that all 5 aromatic check varieties falls in one cluster while the 1 non-aromatic check variety (Mahamaya) forms a separate cluster. Mixed linear model (MLM) was applied to perform genome-wide association mapping where 41 significant marker-trait associations were observed for 24 yield attributing traits. The potential markers identified in the study may provide new opportunities for rice breeder to improve yield and its attributing traits through marker assisted selection approach.

Keywords: Aromatic rice, Genetic diversity, Landraces, Molecular markers, Variability

Rice (*Oryza sativa* L.) is on the frontline in the fight against world hunger and poverty and is also a symbol of cultural identity and global unity. India abounds with a wealth of specialty rice, the prominent among which is the aromatic rice. Almost every state in India has its own collection of small and medium grained aromatic rice which is used in the preparation of delicacy during festival and on some special occasions. Palatability (good taste) is a trait of rice connected to quality, determined by the aroma, appearance, texture and taste (Susiyanti *et al.* 2020). Genetic diversity in aromatic rice landraces is quiet large as compared to other crop species and cultivars. SSR markers are particularly suitable for evaluating genetic diversity and relatedness among plant species. They are also suited for marker assisted selection (Rani and Adilakshmi 2011). SSR markers has been widely used in assessing genetic diversity in rice which can be used to identify potential parent in future aromatic rice breeding programs (Aljumaili *et al.* 2016).

Domestication and artificial selection pressure has changed a lot the genome composition and population structure of aromatic rice. In depth analysis of population structure is therefore required before initiating any breeding programme to get targeted improvement in the traits (Surapaneni *et al.* 2016). Population structure analysis reduces the type I and type II errors in association mapping between sub-groups that cause spurious association between molecular markers and trait of interest in autogamous species like rice. Assessment of genetic diversity and population structure in aromatic rice landraces will be very useful in finding out marker-trait associations by association mapping which utilizes ancestral recombination and natural genetic diversity available in the collected germplasm. So, the present study was carried out to understand the SSR marker-based genetic diversity and marker-trait association analysis in aromatic rice landraces.

MATERIALS AND METHODS

An experiment was conducted at Research cum Instructional Farm, Department of Genetics and Plant Breeding, College of Agriculture, Indira Gandhi Agricultural University, Raipur, Chhattisgarh during *kharif* 2019 and 2020. Mean phenotypic data of two years was used to perform marker-trait association analysis. The experimental material consists of 90 aromatic rice landraces belonging

[1]Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh.
*Corresponding author email: maumitaburman@gmail.com

to IGKV gene pool along with 6 checks which includes 1 non-aromatic and 5 aromatic check varieties (Supplementary Table 1).

Experimental design and crop management: Nursery sowing was done in well prepared raised seed bed in first week of July 2019 and 2020. Twenty 8-days-old seedlings were transplanted in well puddle field in Augmented Block Design. The plant to plant and row to row distance was maintained 15 cm and 20 cm, respectively. The distance between each block was maintained at 50 cm. Five random but robust plants from inner rows were tagged from each plot for data collection. A total of 24 yield attributing traits, viz. days to 50% flowering, plant height (cm), number of effective tillers per plant, length of leaf blade (cm), width of leaf blade (cm), panicle length (cm), panicle weight (g), panicle harvest index, filled grains per panicle, unfilled grains per panicle, total grains per panicle, spikelet fertility percentage, grain yield per plant (g), thousand grain weight (g), grain length (mm), grain breadth (mm), grain length/ breadth ratio, decorticated grain length (mm), decorticated grain breadth (mm), decorticated grain length/breadth ratio, biological yield, harvest index, hulling percentage and milling percentage were measured at particular stages of rice plant following the minimal descriptor of rice. The observations recorded were statistically analyzed using PAST v3.14 software.

SSR markers based genotyping: SSR markers based genotyping work was carried out at R. H. Richharia Research Laboratory, Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh, during 2020. The genomic DNA was isolated from freshly collected leaves of each rice genotypes using protocol of CTAB Method of DNA extraction for rice (Doyle and Doyle 1987). The first 6 SSR markers were selected from different published literatures and gramene database with maintaining the genome wide distribution of markers. The remaining 19 SSR markers were taken from a panel of 50 standard well distributed SSR markers as given in Gramene marker database (http://www.gramene.org/markers/microstat/) (Table 1).

PCR amplification using SSR primers: PCR amplification was carried out in 10 µl mixture containing 1 µl DNA (50 ng/µl), 1 µl of 10X PCR buffer with $MgCl_2$, 1 µl of dNTPs, 1 µl of primer (forward and reverse) and 0.2 µl of 3U of Taq polymerase. The reaction volume was made up to 10 µl

Table 1 List of 25 SSR markers used for genotyping

Primer	Forward sequence	Reverse sequence	Chromosome number
Aro 7	ATTTGCCTCCTGAGTCTG	GAGGATGGGGAAGATAAA	8
FM badh2- E2A	CCTCTGCTTCTGCCTCTGAT	GATTGCGCGGAGGTACTTG	8
FM badh2- E7	GGTTGCATTTACTGGGAGTT	CAGTGAAACAGGCTGTCAAG	8
ARSSR -3	GACACGCACCTCTGTCTAGC	GTTTAATTGGTGAGGAAGTGG	8
BADEX7 -5	TGTTTTCTGTTAGGTTGCATT	ATCCACAGAAATTTGGAAAC	8
AR-5+AR-3	TTGTTTGGAGCTTGCTGATG	ACCAGAGCAGCTGAAATAT	8
RM 342	CCATCCTCCTACTTCAATGAAG	ACTATGCAGTGGTGTCACCC	8
RM 223	GAGTGAGCTTGGGCTGAAAC	GAAGGCAAGTCTTGGCACTG	8
RM 553	AACTCCACATGATTCCACCC	GAGAAGGTGGTTGCAGAAGC	9
RM 120	CACACAAGCCCTGTCTCACGACC	CGCTGCGTCATGAGTATGTA	11
RM 2997	TAAGAAGTAATTTAGTCAAC	TAGAATAGCTATTCTATACC	10
RM 133	TTGGATTGTTTTGCTGGTCGC	GGAACACGGGGTCGGAAGCGAC	6
RM 124	ATCGTCTGCGTTGCGGCTGCTG	CATGGATCACCGAGCTCCCCCC	4
RM 125	ATCAGCAGCCATGGCAGCGACC	AGGGGATCATGTGCCGAAGGCC	7
RM 118	CCAATCGGAGCCACCGGAGAGC	CACATCCTCCAGCGACGCCGAG	7
RM 42	ATCCTACCGCTGACCATGAG	TTTGGTCTACGTGGCGTACA	8
RM 277	CGGTCAAATCATCACCTGAC	CAAGGCTTGCAAGGGAAG	12
RM 428	AACAGATGGCATCGTCTTCC	CGCTGCATCCACTACTGTTG	1
RM 316	CTAGTTGGGCATACGATGGC	ACGCTTATATGTTACGTCAAC	9
RM 338	CACAGGAGCAGGAGAAGAGC	GGCAAACCGATCACTCAGTC	3
RM 17	TGCCCTGTTATTTTCTTCTCTC	GGTGATCCTTTCCCATTTCA	12
RM 84	TAAGGGTCCATCCACAAGATG	TTGCAATGCAGCTAGAGTAC	1
RM 243	GATCTGCAGACTGCAGTTGC	AGCTGCAACGATGTTGTCC	1
RM 55	CCGTCGCCGTAGTAGAGAAG	TCCCGGTTATTTTAAGGCG	3
RM 23120	AACTGTTGGATCGACAAGACCTTCC	ACGGCGTTAAGCTAGACAGACAGAGC	8

RM, Rice marker.

using autoclaved distilled water. The PCR reactions were performed under the following conditions: (i) pre-denature at 94ºC for 5 min; (ii) 35 cycles of run, each followed by denature at 94º C for 45 sec, anneal at 55ºC for 30 sec and extension at 72ºC for 45 sec; (iii) Final extension at 72ºC for 7 min. For better resolution of PCR amplified products five percent polyacrylamide gels (vertical) were used for electrophoresis in 1X TBE. For sizing the fragment DNA ladder of 50 bp was loaded and run alongside the samples. The gels were scanned with the help of BIO-RAD Gel Doc XR System.

Marker informativeness and genetic diversity parameters: The genetic diversity parameters, viz. major allele frequency, number of alleles per locus, gene diversity, heterozygosity and polymorphic information content (PIC) value were measured for each marker using software POWERMARKER version 3.25 (Liu and Muse 2005).

SSR marker based genetic diversity analysis: The marker data matrix was subjected to cluster analysis by following UPGMA (Unweighted pair group method with arithmetic mean) method based on Rho's similarity coefficient as genetic distance measurement with the help of PAST v3.15 (Hammer *et al.* 2001).

Population structure analysis through Bayesian approach: Population structure was analyzed using programme STRUCTURE version 2.3.4 (Pritchard *et al.* 2000). Each run length was specified with a burning period of 100,000 steps followed by 100,000 Monte Carlo Markov Chain replicates (MCMC) with model set of "possibility of admixture and allele frequency correlated". Five independent runs were checked by fixing the number of subpopulations (K) from K= 1 to 10. The accession assigned to single group with the probability >80%, whereas the genotypes those with < 80% probability were adjudged as "admixture" group. The most likely K value was measured based on Evanno's delta K method (Evanno *et al.* 2005) by Online software tool "STRUCTURE HARVESTER" (Earl and HoldtVon 2012).

Marker-trait association analysis by association mapping: Association mapping for yield attributing traits was performed by Mixed Linear Model (MLM) using the software TASSEL v2.3 (Bradbury *et al.* 2007). The significant marker-trait associations were indicated by P-value with corresponding R^2 estimated for each marker as the percentage of the total variation explained. To test the P-value, Bonferroni threshold value was estimated. Bonferroni threshold obtained as 1/n, where n is the number of markers used in the study (Moran 2003).

RESULTS AND DISCUSSION

SSR marker profile in aromatic rice landraces: In this study, among the 25 PCR-based simple sequence repeats molecular markers, 18 markers showed polymorphism. So, 18 polymorphic markers were used for genotyping 96 rice genotypes which produce a total of 108 alleles. The Polymorphic Information Content (PIC) value ranged from 0.52 (RM316) to 0.79 (RM553) with a mean of 0.69. The summary statistics of 18 SSR markers are given in Table 2.

Table 2 SSR marker informativeness used in present study

Marker	Chromosome number	Major allele frequency	Allele	Gene diversity	Heterozygosity	PIC
RM84	1	0.39	9	0.73	0.00	0.69
RM277	12	0.30	7	0.79	0.00	0.76
RM553	9	0.30	8	0.81	0.00	0.78
RM243	1	0.27	8	0.80	0.00	0.78
RM55	3	0.26	8	0.79	0.00	0.76
RM125	7	0.28	7	0.79	0.70	0.76
RM124	4	0.38	5	0.73	0.00	0.68
RM316	9	0.45	3	0.61	0.00	0.52
RM118	7	0.37	5	0.75	0.00	0.72
RM338	3	0.27	7	0.80	0.00	0.78
RM17	12	0.38	4	0.73	0.00	0.68
Aro7	8	0.29	7	0.78	0.00	0.78
FMbadh2-E2A	8	0.43	4	0.66	0.00	0.60
Badex7-5	8	0.41	4	0.69	0.00	0.63
ARSSR3	8	0.49	4	0.64	0.00	0.58
FMbadh2-E7	8	0.35	5	0.77	0.00	0.73
RM342	8	0.42	8	0.75	0.71	0.72
RM223	8	0.45	5	0.62	0.00	0.54
Mean		0.36	6	0.74	0.07	0.69

RM, Rice marker.

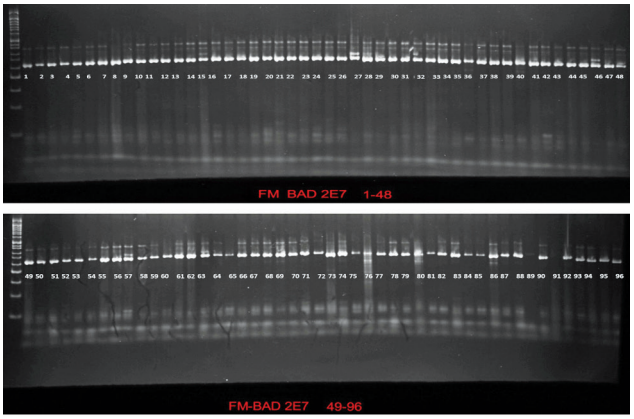

Fig 1 Polymorphism shown by the SSR markers.

The results of the present analysis having average PIC value of 0.69 reveals that all the markers used in this study were highly informative and useful for diversity analysis of a wide range of genotypes because of their highly polymorphic nature (Fig 1). Shamim *et al.* (2016) also reported average PIC value of 0.66 while performing SSR marker based characterization and divergence analysis.

Molecular markers based genetic diversity analysis: The cluster analysis showed a great diversity among aromatic rice landraces. Unweighted Pair Group Method with Arithmetic Mean (UPGMA) and Rho's similarity based Cluster analysis grouped the 96 rice genotypes into 5 Clusters which comprise of 3 major Clusters and 2 minor Clusters comprising of only a single genotype at 0.80 Rho's similarity level. The non-aromatic check variety Mahamaya is genetically dissimilar from all other aromatic accessions and thus forms a single Cluster whereas, the 5 aromatic check varieties were grouped in one cluster. The cluster dendrogram showed that the Rho's similarity coefficient matrix varies from 0.68 to 0.99 (Fig 2). Shamim *et al.* (2016) and Talukdar *et al.* (2017) also used UPGMA clustering method and grouped the studied rice accessions into three major clusters. Fig 2 clearly depicts that all the 5 aromatic checks marked in red forms one cluster while non-aromatic check (Mahamaya) marked in blue forms a single cluster.

Population structure analysis: The highest score for log likelihood reached at K=2 (Fig 3), 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 subpopulations (SG1 and SG2).

Marker-trait association analysis for yield attributing traits: To test the significant marker-trait association Bonferroni threshold was estimated as 1/18 = 0.05, where 18 is the number of markers used in the study (Moran 2003). Therefore P value 0.05 or less than 0.05 for the markers were considered as significant. Mixed linear model (MLM) was used to perform genome-wide association mapping for assessment of novel marker-trait combinations. A total of 41 significant marker-trait associations were detected for 24 different traits in MLM model (Table 3).

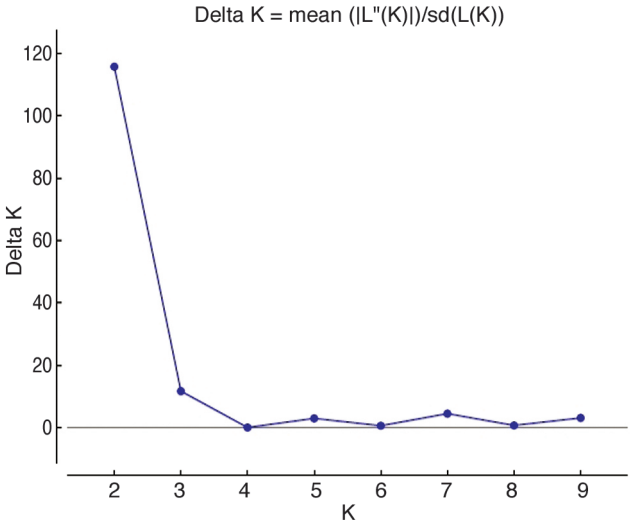

Fig 3 The graph showing the estimated membership fraction for K=2.

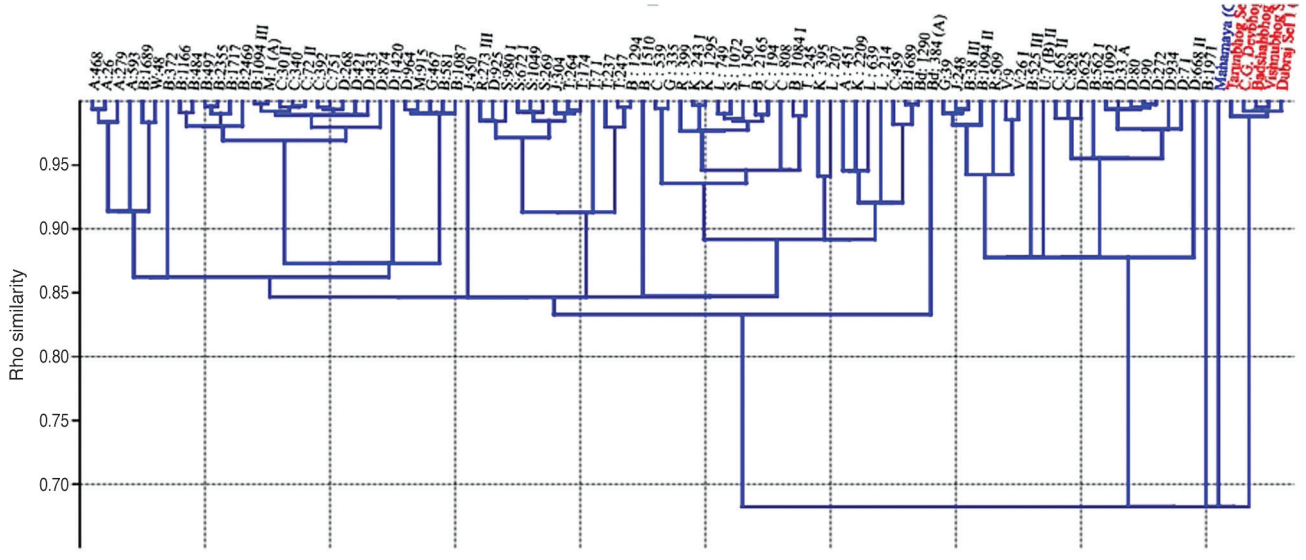

Fig 2 SSR markers based UPGMA dendrogram.

Table 3 Details of significant markers identified in association with 24 yield attributing traits

Trait	Marker	Chromosome no.	Map position (cM)	Marker P value	Marker R^2 value
Days to 50% flowering	RM125	7	24.8	0.0043	0.3779
Days to 50% flowering	RM124	4	150.1	0.0384	0.2418
Days to 50% flowering	Aro7	8	80.3	0.008	0.3125
Days to 50% flowering	Badex7-5	8	80.3	0.0219	0.2602
Plant height	RM316	9	1.8	0.0467	0.0729
Effective tillers per plant	RM277	12	57.2	0.0144	0.1882
Effective tillers per plant	RM338	3	108.4	0.0214	0.1808
Effective tillers per plant	RM17	12	108	0.0184	0.135
Length of leaf blade	RM243	1	57.3	0.0395	0.15
Width of leaf blade	RM84	1	18.8	0.0198	0.2092
Width of leaf blade	RM243	1	57.3	0.022	0.1944
Width of leaf blade	RM342	8	78.4	0.0037	0.2927
Panicle length	RM17	12	108	0.0478	0.0787
Panicle length	RM342	8	78.4	0.0287	0.2196
Panicle weight	RM277	12	57.2	0.0026	0.2071
Filled grains per panicle	RM124	4	150.1	9.72E-04	0.2029
Filled grains per panicle	RM338	3	108.4	0.0107	0.1706
Unfilled grains per panicle	RM118	7	130.5	0.0072	0.1363
Total grains per panicle	RM124	4	150.1	3.73E-04	0.2143
Total grains per panicle	RM338	3	108.4	0.0488	0.1308
Total grains per panicle	RM17	12	108	0.0194	0.1037
Spikelet fertility percentage	RM118	7	130.5	0.0047	0.1477
Grain yield	RM277	12	57.2	5.06E-04	0.2415
Thousand grain weight	RM277	12	57.2	9.91E-04	0.2871
Grain length	RM553	9	76.7	0.0138	0.1843
Grain length	RM124	4	150.1	0.0075	0.1548
Grain breadth	RM125	7	24.8	0.0383	0.2417
Grain L/B ratio	RM553	9	76.7	0.003	0.2142
Grain L/B ratio	RM125	7	24.8	0.0483	0.19
Decorticated grain length	RM277	12	57.2	0.0384	0.1558
Decorticated grain length	RM124	4	150.1	0.0101	0.1678
Decorticated grain breadth	RM243	1	57.3	0.0305	0.1588
Decorticated grain breadth	Badex7-5	8	80.3	0.0175	0.1093
Decorticated grain breadth	RM342	8	78.4	0.0108	0.2613
Decorticated grain L/B ratio	RM124	4	150.1	0.0313	0.1235
Biological yield	RM277	12	57.2	7.92E-04	0.224
Biological yield	RM316	9	1.8	0.0347	0.071
Harvest index	RM338	3	108.4	0.0239	0.1699
Hulling percentage	RM553	9	76.7	0.0248	0.19
Hulling percentage	Aro7	8	80.3	0.0015	0.3273
Milling percentage	RM84	1	18.8	0.024	0.2057

Identification of markers associated with multiple traits: In our study, out of 41 significant marker trait association for grain yield and its attributing traits, 38 markers were having higher R^2 value i.e. >10% indicating highly significant level of association between marker and traits. Results also revealed that rice marker RM277 located on chromosome 12 at position 57.2 cM is significantly associated with grain yield and its attributing traits like thousand grain weight, biological yield, panicle weight, effective tillers per plant and decorticated grain length. Swamy *et al.* (2017) also used this marker ($P<0.05$) to identify significant marker-trait association for grain yield, plant height and days to 50% flowering. They also observed significant marker-trait association for grain yield on chromosome 12 which is in close harmony with the present study. Talukdar *et al.* (2017) also used $P<0.05$ to identify significant marker-trait associations and found a total of 29 significant marker-trait associations for 10 agro-morphological characters.

The present study showed ample amount of genetic variability in studied rice landraces. The potential markers identified in the present investigation may provide new opportunities for rice breeder to improve yield and its attributing traits through marker assisted selection approach, thus making the conventional breeding faster and efficient.

REFERENCES

Aljumaili S J, Rafii M Y, Latif M A, Sakimin S Z, Arolu I W and Miah G. 2016. Genetic diversity of aromatic rice germplasm revealed by SSR markers. *BioMed Research International* **3**: 1–11.

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**: 2633–35.

Doyle J J and Doyle J L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.

Earl D A and HoldtVon B M. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the evanno method. *Conservation Genetics Resources* **4**(2): 359–61.

Evanno G, Regnaut S and Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–20.

Hammer Q, Harper D A T and Ryan P D. 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4**: 9.

Liu K and Muse S V. 2005. PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics* **21**: 2128–29.

Moran M D. 2003. Arguments for rejecting the sequential Bonferroni in ecological studies. *Oikos* **100**: 403–05.

Pritchard J K, Stephens M and Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–59.

Rani M G and Adilakshmi D. 2011. Genetic analysis of blast resistance in rice with simple sequence repeats (SSR). *Journal of Crop Improvement* **25**: 232–38.

Shamim M Z, Manzar H, Sharma V K and Kumar P. 2016. Microsatellite marker based characterization and divergence analysis among rice varieties. *Indian Journal of Biotechnology* **15**: 182–89.

Surapaneni M, Balakrishnan D, Mesapogu S, Raju A K, Rao Y V and Neelamraju S. 2016. Genetic characterization and population structure of Indian rice cultivars and wild genotypes using core set markers. *Biotechnology* **6**: 95–106.

Susiyanti S, Rusmana R, Maryani Y, Sjaifuddin S, Krisdianto N and Syabana M A. 2020. The physicochemical properties of several Indonesian rice varieties. *Biotropia - The Southeast Asian Journal of Tropical Biology* **27**: 41–50.

Swamy B P M, Shamsudin N A A, Rahman S N A, Mauleon R, Ratnam W, Cruz M T S and Kumar A. 2017. Association mapping of yield and yield relatedtraits under reproductive stage drought stree in rice (*Oryza sativa* L.). *Rice* **10**: 21–28.

Talukdar P R, Rathi S, Pathak K, Chetia S K and Sarma R N. 2017. Population structure and marker-trait association in indigenous aromatic rice. *Rice Science* **24**: 145–54.