Gene Reports 30 (2023) 101719

Contents lists available at ScienceDirect

Gene Reports

journal homepage: www.elsevier.com/locate/genrep

Assessment of allelic and genetic diversity, and population structure among farmers' rice varieties using microsatellite markers and morphological traits

Pritesh Sundar Roy [a,1], Shubhransu Nayak [b,1], Soma Samanta [a], Apurba Chhotaray [a], Soumya Mohanty [a], Sudhiranjan Dhua [a], Urmila Dhua [a], Bhaskar Chandra Patra [a], Kapil Kumar Tiwari [c], S.V. Amitha C.R. Mithra [d], Rameswar Prasad Sah [a], Lambodar Behera [a,*], Trilochan Mohapatra [e]

[a] ICAR-National Rice Research Institute, Cuttack 753006, Odisha, India
[b] Odisha Biodiversity Board, RPRC, Bhubaneswar, Odisha, India
[c] Bio Science Research Centre, SDAU, SK Nagar, Gujarat, India
[d] ICAR-National Research Centre for Plant Biotechnology, New Delhi 110012, India
[e] ICAR, Government of India, New Delhi, India

ARTICLE INFO

Edited by Dominic Voon

Keywords:
Allelic diversity
farmers' rice varieties
Genetic diversity
Microsatellite marker
Rice
Structure

ABSTRACT

Traditional rice varieties grown by the farmers serve as valuable genetic resources for future rice improvement. These varieties are highly adapted to varied agro-ecological conditions. However, they are rapidly lost because of the adoption of high-yielding varieties. The extent of allelic and genetic diversity present in the germplasm is a prerequisite for the improvement of any crop and conservation strategies under adverse impacts of climate. Farmers' rice varieties are usually poor yielders but are allelic treasurer for different traits, especially biotic and abiotic stresses, grain qualities, early seedling vigor, input use efficiency, etc. Therefore, the present study was aimed for a detailed understanding of allelic and genetic diversity, and population structure of 607 farmers' rice varieties using 36 fluorescently labeled microsatellite markers and 53 morphological traits. A total of 363 alleles was detected with an average of 10.33 alleles per locus and moderately high Nei's allelic/gene diversity (0.502) was detected. Polymorphic information content ranged from 0.685 to 0.987 with an average of 0.901. 34 unique, 236 rare, 84 low-frequency and 44 high-frequency alleles were detected. 53 morphological traits harbored a total of 195 variables with an average of 4.217 variables per trait. 50 out of 53 morphological traits showed polymorphism and highly significant differences among varieties. High genetic diversity was observed among 607 farmers' rice varieties both at molecular (0.653) and phenotypic (0.656) levels. The dendrogram based on both microsatellite markers and morphological traits grouped the 607 farmers' rice varieties into three major groups. A moderate population structure was observed with two independent subpopulations SP1 and SP2, which have membership percentages of 82.6 % and 17.4 %, fixation index values of 0.19 and 0.194, respectively. The AMOVA could explain 63 % of the total variation among varieties and 34 % within varieties. Our results showed that the farmers' rice varieties of Odisha harbored higher levels of both allelic and genetic diversity. Hence, these varieties would be useful for the identification of novel and elite alleles, and serve as a source of donors for the development of climate-smart varieties with improved grain yield and qualities, and input use efficiency, which would be sustainable in changing climate scenario conditions and improve farmers' income.

Abbreviations: AMOVA, Analysis of Molecular Variance; CTAB, Cetyltrimethylammonium Bromide; FAO, Food and Agriculture Organization; F_{IS}, Inbreeding Coefficient; F_{ST}, Fixation Index; FV, Farmers' Variety; He, Expected Allelic/Gene Diversity; Ho, Expected Homozygosity; MCMC, Markov Chain Monte Carlo; Na, Number of Alleles; Ne, Number of Effective Alleles; IPR, Intellectual Property Rights; PCR, Polymerase Chain Reaction; PCA, Principal Component Analysis; PCoA, Principal Co-ordinate Analysis; PIC, Polymorphism Information Content; PPV&FRA, Protection of Plant Varieties and Farmers'Rights Authority; RGNMS, Rice Genic Non-coding Microsatellite; Rs, Allelic Richness; SP, Subpopulation; SSR, Simple Sequence Repeat.

* Corresponding author.
E-mail addresses: lbehera.publi2018@gmail.com, lambodarjamujhadi@gmail.com (L. Behera).

[1] Contributed equally to the work.

https://doi.org/10.1016/j.genrep.2022.101719
Received 13 July 2022; Received in revised form 15 November 2022; Accepted 27 November 2022
Available online 2 December 2022

1. Introduction

Rice (*Oryza sativa* L) is the staple food for >50 % of the world population. It is one of the most important crops, which has potential towards achieving global food demand and food security in upcoming years (FAO (Food and Agriculture Organization), 2004). There has been a notable yield plateau and rice production after the green revolution era (Barker and Dawe, 2002). However, to address the issue of yield stagnation, attention must be taken for both large farmers cultivating high yielding modern varieties as well as small and marginal farmers who rely on local landraces. Landraces are being grown for decades with on-farm management practices and are being selected by farmers to suit their local environmental conditions and needs (Zeven, 1998). Moreover, the genetic diversity available in rice landraces is comparatively higher than other crop species. Such landraces cultivated, maintained by farmers associated with traditional knowledge are being named as 'farmers' varieties under the PPV&FR Act, 2001 (Nagarajan et al., 2010). These varieties are disappearing fast due to their low yield and introduction of high-yielding varieties over the years (Rauf et al., 2010). According to Nagarajan (2007), farmers' varieties are almost uniform, homogenous populations, distinct among themselves for various traits and are highly accepted by consumers. Since, the landraces are an intermediate group between cultivated rice and wild ancestors (Pusadee et al., 2009), they are in the state of evolution and are adapted to varied agro-ecological conditions (Singh and Singh, 2003). In fact, various useful genes for different valuable and complex traits are available in these landraces; they could be successfully utilized for crop improvement programs (Hanamaratti et al., 2008; Lisa et al., 2011; Amegan et al., 2020). The decline in cultivation leading to genetic erosion and a huge bottleneck of these rice accessions is a major concern prioritizing their collection, cultivation, characterization, and commercialization.

Rice is the predominant crop in Asian countries and Southern part of Asia linked to the major rice domestication center is known as 'food basket' (Callaway, 2014). Asia in general and India have an ample number of farmers' varieties, but many of them are still uncharacterized. Though, India is one of the largest exporters of rice, most of the population depends on local varieties for internal consumption. These varieties have enormous potential for wide adaptability, stress tolerance/resistance, cooking quality, nutritional values and medicinal properties (Taguchi-Shiobara et al., 2013; Lai, 1995; Frei and Becker, 2004; García Montecinos et al., 2011; Chang et al., 2010; Behera et al., 2012; Das and Das, 2014). Many of these varieties hold significance for their unique properties for the preparation of various value added products and hence are highly preferred in domestic markets (Rijal, 2010; Dela Cruz and Khush, 2000). Rice cultivation in the eastern Indian state, i.e., Odisha is primarily covered by tribal communities and small farmers who have their own sets of native rice varieties in almost all the districts (Das, 2012). These farmers have tremendous socio-religious belief for their traditional rice varieties and are reluctant to adapt to modern high-yielding varieties. Moreover, these farmers' varieties are named traditionally according to farmers perception and preference mostly linked to the local area. Due to lack of sufficient research attention, genetic background and potential of these valuable resources are still untapped and are on the verge of extinction. A large number of indigenous varieties available in Jeypore tract of Odisha is associated with the secondary center of origin of rice (Arunachalam et al., 2006) and are not characterized properly. Further, farmers' varieties of Odisha are cultivated in varied agro-climatic conditions, from extreme saline of the coastal region to drought affected western Odisha, different soil conditions and nine different agro-ecological zones (Das, 2012). Therefore, conservation, maintenance, legal protection and genetic enhancement of these farmers' varieties are prerequisite for economic upliftment of the poor farmers of Odisha.

Characterization of these farmers' varieties, both at molecular and morphological level could be of high significance to elucidate their relationship and domestication process, which would be further useful for their genetic enhancement and successful use in crop improvement programs. The farmers' varieties available in the natural population could serve as valuable material for broadening the narrow genetic base present in improved high yielding cultivars, thereby overcoming the yield gap.

The establishment of polymerase chain reaction strategy for several crops and wide adoption of biotechnology and bioinformatics techniques has supported characterization and assessment of genetic diversity within and among closely related crop species. Molecular markers have proven to be useful tools for varietal discrimination and investigation of genetic structure available within the germplasm set. In rice, out of many PCR markers, SSR markers are highly preferred as they are robust, co-dominant, highly informative, hyper variable, highly polymorphic, abundant, and cost effective (Chambers and Avoy, 2000). Moreover, with the generation of complete genetic map of rice and availability of abundance of saturated SSR markers throughout the genome (McCouch et al., 2002; International Rice Genome Sequencing Project, 2005), understanding the potential genetic variation and development of conservation strategies for elite, but endangered rice landraces have become quite feasible. Rice is the model plant for the study of grass genetics (Causse et al., 1994; Kurata et al., 1994). Several studies have been undertaken for identification and fingerprinting of rice varieties, characterization of genetic diversity in landraces and cultivated rices (Jain et al., 2004; Bhuyan et al., 2007; Kumar et al., 2010; Behera et al., 2012, 2013; Choudhury et al., 2013, 2014; Das et al., 2013; Singh et al., 2016b; Aljumaili et al., 2018; Pathaichidachote et al., 2019; Nilthong et al., 2020; Mohanty et al., 2021).

Genetic diversity of indigenous rice collections of North-Eastern regions of India, the region known for primary domestication of rice, have carried out by many researchers both by agro-morphological (Vairavan et al., 1973; Borkakati et al., 2000; Sarma and Pattanayak, 2009) and molecular markers (Glaszmann et al., 1989; Sarma and Bahar, 2005; Bhuyan et al., 2007; Singh et al., 2016a; Anupam et al., 2017; Vanlalsanga et al., 2019). But the potential of rice germplasm available in the eastern part of India has not been studied extensively. Therefore, we have characterized a set of under-exploited 607 farmers' varieties collected from different parts of Odisha at the molecular and phenotypic levels using 36 fluorescently labeled microsatellite markers and 53 morphological traits with 195 morphometric descriptors. Their characterization led to the understanding of their genetic structure providing a platform for their protection, addressing intellectual property rights (IPR), and their rational use as genetic resources for crop improvement.

2. Materials and methods

2.1. Plant materials

A total of 607 locally cultivated farmers' varieties of rice (*Oryza sativa* L.) were collected from different regions of Odisha (Fig. 1). Further, these collected rice varieties were categorized according to their 28 districts of collection and nine agro-climatic zones of Odisha. The variety name, locality and their agro-ecological area of cultivation are indicated in Supplementary Table S1. These 607 farmers' varieties are *indica* land races with varied agro-morphological, yield and quality traits, adaptions, etc., and were used for assessment of allelic and genetic diversity, and population structure using microsatellite markers and morphological traits.

2.2. Isolation of genomic DNAs and genotyping using fluorescently labeled microsatellite markers

Fresh leaf samples were harvested from one-month young seedlings and genomic DNAs were extracted from each farmer's rice variety following a modified cetyltrimethylammonium bromide (CTAB) extraction protocol (Dellaporta et al., 1983). Thirty-six SSR loci, three from each of 12 chromosomes of rice were selected for the genetic

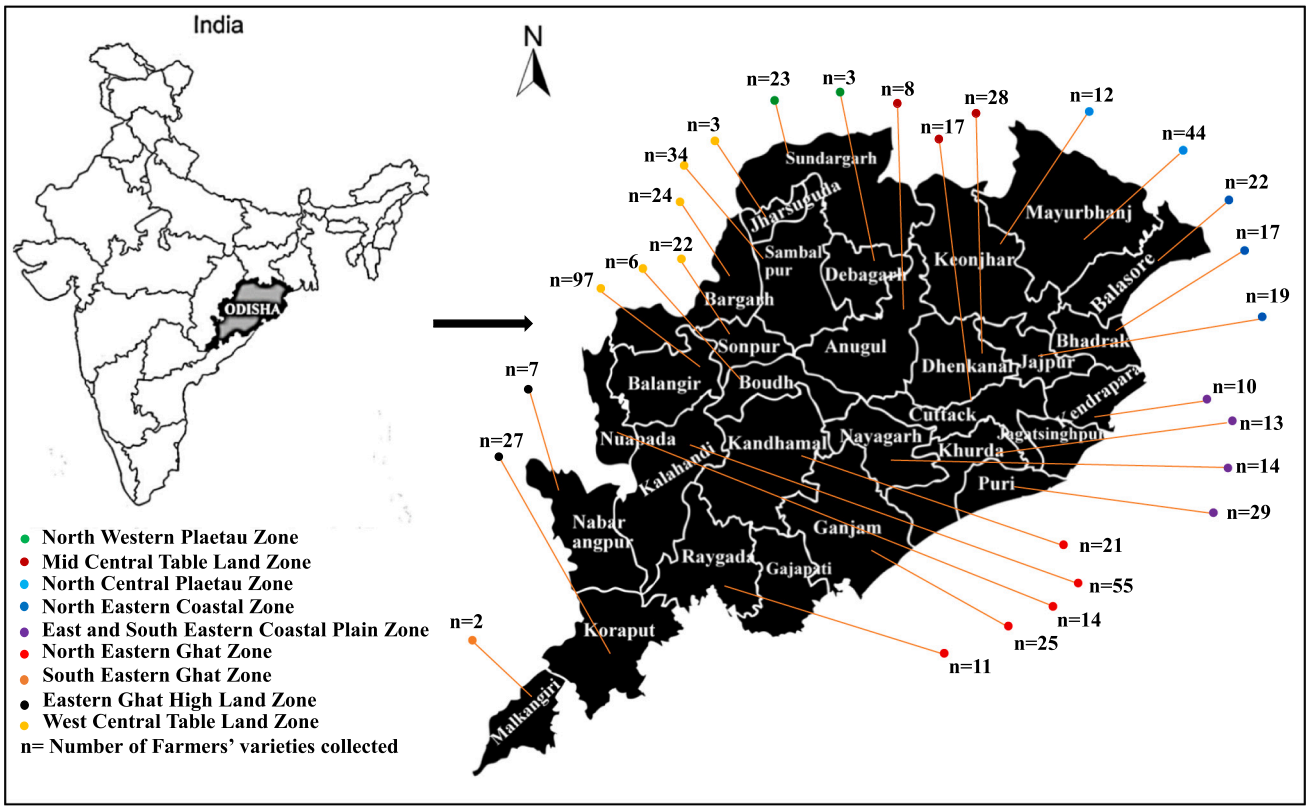


Fig. 1. Sampling sites of farmers rice varieties from twenty-eight districts of Odisha, India and information on their distribution by agro-climatic zones.

diversity analyses. Details of the SSR loci used in the present study are given in Supplementary Table S2. For multiplexing, primers were fluorescently labeled with four different dyes (FAM, VIC, NED, and PET) and the PCR amplification was performed according to the procedure described by Chen et al. (1997). The PCR amplifications were performed in a total reaction volume of 10 μl containing 20 ng genomic DNA, 2.0 picomoles of each forward and reverse primer, 1 μl of 10× buffer (0.1 M Tris, pH 9, 0.5 M KCl, 15 mM $MgCl_2$, 0.1 % gelatine), 200 μM each of dNTPs and 0.3 U of *Taq* DNA polymerase. The PCR condition was an initial denaturation at 94 °C for 5 min followed by 35 cycles of denaturation at 94 °C for 30 s, annealing (depending on the TM value of primer) at 50°-60 °C for 30 s, extension at 72 °C for 1 min and a final extension of 20 min at 72 °C. However, a reference blank was used in individual PCR reaction to avoid any unambiguous amplification due to contamination. The PCR products were mixed with the fluorescent dyes at 1:1:2:4 ration for FAM: VIC: NED: PET, respectively (Tiwari et al., 2015). Further, 8.9 μl of Hi di formamide and 0.2 μl of an internal size standard ROX500 (Applied Biosystems, USA) were mixed with 1.0 μl of mixed microsatellite sample, the samples were denatured at 95 °C for 5 min and run in ABI3730xl 96 capillary automated fragment analyzer system (ABI, Model 373xl) and the results were analyzed with GENE MAPPER 4.1 (Applied Biosystems 2009).

2.3. Allele scoring

The size of each intense amplified fragment for all SSR loci was determined by comparison with the size standard (100 bp DNA ladder) and scored by incremental numbering from the lowest molecular weight band to the progressively higher molecular weight bands to prepare the genotype matrix.

2.4. Molecular data analysis

The amplified bands/alleles were scored as present (1) or absent (0) for each genotype and primer combination. The data were entered into a binary matrix and subsequently analyzed using different computer software packages. The total number alleles, number of polymorphic alleles, number of unique alleles, number of rare alleles, number of low frequency alleles, number of high frequency alleles, number of multiple alleles and polymorphism information content (PIC) were calculated to assess the diversity of alleles of each marker locus. An allele that was observed in >30 % of the 607 rice varieties was a high frequency/abundant/common allele, while an allele having a frequency between 5 % and 30 % is called as a low frequency/intermediate allele. An allele that was observed in <5 % was considered to be rare allele. The polymorphism information content (PIC) for each SSR marker locus was calculated using the formula: $PICi = 1 - \sum_{j=1}^{n} (Pij)^2$, where n is the number of marker alleles for marker i and P_{ij} is the frequency of the j^{th} allele of marker I (Anderson et al., 1993). Gene/allelic diversity parameters viz., number of alleles (*Na*), effective number of alleles (*Ne*), expected homozygosity (Ho), expected heterozygosity (*He*) (Nei, 1973) and Shannon Index (*I*) were evaluated using POPGENE V1.32 (http://www.ualberta.ca/fyeh) with 1000 permutations. The allelic richness (*Rs*) for each SSR locus was measured using FSTAT 2.9.3 (Goudet, 2005). The Bayesian model-based clustering analysis of the varieties was used for determining the optimal number of genetic clusters found among rice varieties using the STRUCTURE software (Pritchard et al., 2000) with 1,00,000 burn in periods and 1,00,000 Markov Chain Monte Carlo (MCMC) replicates with ten independent runs (K) ranging from 1 to 10. The ΔK based on the change in the log probability of the data between successive K values was estimated using the parameters described by Evanno et al. (2005) using the software program Structure Harvester v6.0 (Earl and von Holdt, 2012) and population clusters were produced by the software Structure Plot developed by Ramasamy et al. (2014) (http://btismysore.in/strplot). Moreover, varieties were further grouped based on their collection by geographical location (28 districts) and the genetic diversity parameters of the varieties within each district were determined using POPGENE V1.32. The genetic variation within

and among the populations was calculated by the procedure of AMOVA (Analysis of Molecular Variance) using the software GeneALEx6 (Peakall and Smouse, 2006). Isolation by distance was evaluated by assessing the correlation matrix between pairwise genetic distance (Nei's) and geographic distance between districts matrices using a Mantel's test in the program IBD (Bohonak, 2002) with 1, 00,000 random permutations. In order to know the genetic relationship among 607 rice varieties, Jacquard's genetic dissimilarity coefficients among varieties were computed using NTSYS-pc(V2.02)(Rolf, 1998) while the neighbour-joining (NJ) tree was constructed based on the Nei's genetic distance using DARwin 6.0.9 (Perrier and Jacquemoud-Collet, 2006). The principal component analysis (PCA) and principal co-ordinate analysis (PCoA) were performed using PAST4.2 statistical package and XL-STAT, respectively.

2.5. Morphological characterization

The seeds of each of 607 rice farmers' varieties were sown in the nursery. 21 days old seedlings were transplanted in the experimental field of ICAR-NRRI, Cuttack, Odisha, India research farm at a spacing of 20 cm × 20 cm following Augmented design. Recommended agronomic practices such as manual weeding and fertilizer applications were adopted. Fertilizer doses of 80 kg N, 40 kg P_2O_5, and 40 kg K_2O per hectare were applied. 50 % N, 100 % P_2O_5, and 50 % of K_2O were applied as a basal dose. The remaining nitrogen was applied in equal two splits, one at active tillering and the second at the booting stage, while the remaining K_2O was applied at the booting stage. Pest control and water management were paid attention every day.

2.6. Data recording morphological traits and statistical analysis

All the 607 farmers' varieties were characterized for 195 morphometric descriptors of 53 DUS morphological traits from the early vegetative stage to post-harvest stage following descriptors for rice (IRRI-IBGR, 1980) (Supplementary Table S3). Data were recorded on 10 plants per variety. These phenotypic traits of all the varieties were subjected to statistical analysis, i.e., mean, standard deviation, standard error, variance, test for significance of each trait, and correlation among traits using SPSS (v. 23). In order to know the genetic relationship among 607 rice varieties, Jacquard's genetic dissimilarity coefficients among varieties were computed using NTSYS-pc(V2.02)(Rolf, 1998) while the neighbour-joining (NJ) tree was constructed based on the Nei's genetic distance using DARwin 6.0.9 (Perrier and Jacquemoud-Collet, 2006). The principal component analysis (PCA) and principal co-ordinate analysis (PCoA) were performed using PAST4.2 statistical package and XL-STAT, respectively.

3. Results

3.1. Allelic/gene diversity at microsatellite loci

a) Number of alleles

All the 36 rice microsatellite markers including 20 rice microsatellite (RM) and 16 rice genic non-coding microsatellite (RGNMS) markers revealed polymorphism between 607 farmers' rice varieties. A total of 363 alleles were amplified, all being polymorphic. The number of alleles

Table 1
Diversity parameters of microsatellite markers used in the study.

Microsatellite locus	Na	Ne	PIC	Ho	He	I	F_{IS}	F_{ST}	Rs
RM259	16	5.712	0.984	0.174	0.825	2.063	0.711	0.859	3.679
RM5	9	3.145	0.948	0.317	0.682	1.340	0.651	0.827	2.873
RGNMS135	18	3.365	0.975	0.297	0.703	1.696	0.584	0.793	3.025
RGNMS13	11	3.474	0.971	0.287	0.712	1.447	0.889	0.945	2.816
RGNMS249	10	3.769	0.953	0.265	0.735	1.528	0.540	0.772	3.093
RGNMS169	21	6.307	0.987	0.158	0.842	2.385	0.579	0.792	3.995
RM489	10	1.316	0.919	0.760	0.240	0.594	0.698	0.849	1.604
RM338	3	1.154	0.696	0.867	0.133	0.266	0.666	0.833	1.277
RGNMS1	3	1.106	0.706	0.904	0.096	0.215	1.000	1.000	1.179
RGNMS20	27	3.467	0.979	0.288	0.712	1.784	0.725	0.864	3.060
RM307	6	1.177	0.857	0.849	0.150	0.374	0.721	0.886	1.395
RM237	11	1.808	0.939	0.553	0.447	1.042	0.581	0.801	2.182
RM507	3	1.297	0.709	0.771	0.229	0.398	0.458	0.736	1.463
RGNMS190	3	1.083	0.685	0.923	0.077	0.179	0.265	0.705	1.160
RM433	11	5.026	0.974	0.198	0.801	1.889	0.648	0.827	3.503
RM133	5	1.751	0.885	0.571	0.429	0.696	0.965	0.983	1.893
RGNMS141	5	1.055	0.838	0.947	0.053	0.155	0.932	0.991	1.142
RGNMS167	18	4.016	0.981	0.248	0.751	1.779	0.754	0.877	3.107
RM125	5	1.536	0.837	0.651	0.349	0.718	0.481	0.740	1.782
RM11	11	2.131	0.948	0.469	0.531	1.147	0.701	0.856	2.350
RM118	4	2.151	0.874	0.464	0.535	0.972	0.880	0.940	2.281
RM25	11	5.484	0.973	0.182	0.818	1.841	0.640	0.821	3.499
RGNMS165	7	1.403	0.887	0.712	0.287	0.656	0.642	0.829	1.651
RM284	8	1.530	0.900	0.653	0.347	0.797	0.497	0.758	1.942
RM316	11	3.004	0.963	0.332	0.667	1.450	0.726	0.868	2.894
RGNMS198	8	1.789	0.916	0.559	0.441	0.914	0.628	0.818	2.157
RM552	23	4.384	0.983	0.227	0.772	1.987	0.426	0.721	3.414
RGNMS27	15	2.133	0.963	0.468	0.531	1.211	0.570	0.824	2.341
RGNMS23	11	4.249	0.971	0.235	0.765	1.727	0.639	0.829	3.308
RM171	6	1.883	0.877	0.531	0.469	0.927	0.437	0.736	2.175
RGNMS29	4	2.022	0.876	0.494	0.505	0.882	0.971	0.986	2.155
RM287	13	2.911	0.962	0.343	0.656	1.546	0.497	0.774	2.799
RM144	18	3.279	0.973	0.304	0.695	1.680	0.404	0.707	3.114
RGNMS38	3	1.370	0.760	0.730	0.270	0.471	1.000	1.000	1.602
RGNMS175	7	2.709	0.913	0.369	0.631	1.190	0.528	0.766	2.513
RM277	8	1.241	0.891	0.806	0.194	0.414	0.624	0.821	1.429
Mean	10.083	2.646	0.901	0.497	0.502	1.121	0.657	0.829	2.385

Na: Number of alleles, Ne: Number of effective alleles, PIC: Polymorphism information content, Ho: Expected homozygosity, He: Expected allelic/gene diversity, I: Shannon's information index, F_{IS}: Inbreeding coefficient, F_{ST}: Fixation index, Rs: Allelic richness.

per locus ranged from 3 (RM338, RGNMS1, RM507 and RGNMS190) to 27 (RGNMS20) with an average of 10.08 alleles (Table 1). Besides, RGNMS20, six of the markers were found to be amplifying >15 alleles, of which, RM522 amplified 23 alleles followed by RGNMS169, RM144, RGNMS167, RGNMS135 and RM259 with 21, 18, 18, 18 and 16 alleles, respectively. The number of effective alleles detected varied from 1.055 (RGNMS141) to 6.307 (RGNMS169) with an average of 2.65.

The SSR loci with simple tri-nucleotide repeat motifs amplified the highest number of alleles (average = 11.80, $n = 5$, Na = 59) followed by di-nucleotide repeat motifs (average = 10.58, $n = 26$, Na = 275), mixed-nucleotide repeat motifs (average = 8.50, n = 2, Na = 17) and tetra-nucleotide repeat motifs (average = 4.00, $n = 3$, Na = 12) (Supplementary Table S4). On an average, the RGNMS markers tended to be detected with higher number of alleles (average = 10.68, $n = 16$, Na = 171) over RM markers (average = 9.60, $n = 20$, Na = 192). The size of alleles varied from 80 bp (RM287, RGNMS23) to 348 bp (RM171). The size difference between the smallest allele and the largest allele varied from 2 bp (RGNMS38) to 89 bp (RM489) (Supplementary Table S5). A positive correlation ($r = 0.38$, $P = 0.027$) was found between number of alleles amplified per locus and number of repeats in simple motif of an SSR locus (Supplementary Table S6).

b) Allelic/gene diversity

The allele richness (Rs) varied from 1.142 (RGNMS141) to 3.995 (RGNMS169) with an average of 2.385. Moderately high Nei's allelic/gene diversity (Ne) was detected, which varied from 0.053 (RGNMS141) to 0.825 (RM259) with an average of 0.502 (Table 1). Rice varieties of similar names collected from same or different areas showed variability for one or more locus. For instance, three accessions of Akul collected from three different areas, i.e., Bargarh, Sambalpur and Bolangir showed allelic variation for RM259, RGNMS20, RM433 and RGNMS141. Similarly, four accessions of Jhilli collected from Bolangir showed allelic difference for 29 of the 36 tested loci. The expected heterozygosity (He) varied from 0.053 (RGNMS141) to 0.842 (RGNMS169) with an average of 0.502 per locus (Table 1). Similar results were detected for Shannon's information index (I). The inbreeding coefficient value (F_{IS}) was in line with a fixation index for the tested markers. The highest value of inbreeding coefficient (i.e., 1.0) was observed at RGNMS1 and RGNMS38 loci whereas, the lowest value (i.e., 0.265) was observed at RGNMS190 locus. Though, no significant difference was detected between the two sets of RGNMS markers for polymorphism information content, expected homozygosity and allele richness values, RGNMS markers were comparatively higher than RM markers for contributing Nei's gene diversity and Shannon's information index values in these varieties. The diversity parameters like polymorphism information content, Nei's allelic/gene diversity, Shannon's information index and allele richness were recorded high for di-nucleotide repeat motifs, whereas these were lowest for tetra-nucleotide repeat motifs.

c) Types of alleles (Unique, rare, low frequency, high frequency and multiple alleles)

A total of 236 (65.01 %) rare alleles (average = 6.56), ranging from 1 (RM118, RM338, RM507, RGNMS29 and RGNMS38) to 24 (RGNMS20) were observed. Moreover, 84 (23.14 %) low-frequency alleles and 44 (12.12 %) high-frequency alleles were recorded in our study. On an average, the RM markers amplified high number (53) of low frequency alleles as compared to RGNMS markers (31). All SSR locus amplified atleast one rare and one high-frequency allele, while 33 SSR loci amplified one low-frequency allele (Supplementary Table S5). A positive correlation was found between number of alleles amplified per locus and number of rare alleles ($r = 0.969$, $P < 0.0000001$) and also with number low frequency allele ($r = 0.666$, $P = 0.000009$) (Supplementary Table S6).

An allele which was detected only in one variety for any given locus was considered as a unique allele. A total of 34 (9.37 %) unique alleles were observed at 19 (52.78 %) SSR loci (Supplementary Table S5). A positive correlation was found between number of alleles amplified per locus and number of unique alleles ($r = 0.704$, $P = 0.000002$). The loci, RGNMS20 detected a maximum of 4 unique alleles and one in each variety, namely, Sunakhadi, Akul, Rangahajari, Haladi Sapuru (Table 2). Ten SSR loci RM5, RM11, RM144, RM171, RM259, RM316, RM433, RM507, RGNMS169 and RGNMS249 amplified one unique allele each in the farmers' varieties, Ramakrushna Bilasha, Damodar Bhoga, Patini, Machhakanta, Kalakrushna, Karpurmoti Mahipal, Abhirman, Kapa Anthi and Kalkati, respectively (Table 2). Three unique alleles each were detected by four microsatellite loci, RM552 (in Jaifulla, Jayaphula and Dhobgaini), RGNMS27 (in Basa Bhog, Luchai and Abhirman), RGNMS135 (in Bada Champa, Karpurjira and Kalkati) and RGNMS167 (in Magur Manji, Krushana Bhog and Haladi Sapuru). Further, two unique alleles were amplified by each of RM237 (in Bhasa Manik and Bunde), RM277 (in Pimpudi Basha and Parbat Jeera), RM284 (in Lalgundi and Ratan Mali) and RGNMS198 (in Kunda Dhan and Sathia) (Table 2). Both the marker sets, i.e., RGNMS and RM, each detected 17 unique alleles.

Thirty-three of 36 microsatellite loci amplified multiple alleles (Supplementary Table S5). Three markers RGNMS1, RGNMS141 and RGNMS190 did not amplify any multiple alleles. 597 out of 607 farmers' varieties showed the presence of multiple alleles at least at one of 33 loci. For any given marker locus, multiple alleles were detected in an average of 19 % varieties per locus, indicating the presence of high level of heterogeneity in the varieties. The number of varieties with multiple alleles varied from two (0.3 %) at RGNMS141 locus to 264 (43.5 %) at RM552 locus. A positive correlation ($r = 0.712$, $P < 0.000001$) was found between total number of alleles and frequency of varieties showing multiple alleles (Supplementary Table S6). Ten varieties, Basudha-R, Basudha, Bhogi, Bodo, Haldi Ropa, Kalakrushana, Kusuma, Patini, Ranga Luchai, Ranisaheband and Ratanchudi did not show multiple alleles at any of 36 SSR loci.

3.2. PIC values

The PIC value varied from 0.685 (RGNMS190) to 0.987 (RGNMS169) with an average of 0.901. Twenty-one loci showed the PIC value of >0.9. Microsatellite markers with simple di-nucleotide repeat motifs detected higher polymorphism (Mean: 0.918; $n = 26$) than those with tri-nucleotide repeat motifs (Mean: 0.882, $n = 5$) and tetra-

Table 2
Unique alleles amplified in farmers' rice varieties by microsatellite markers.

Sl. no.	Chrom#	Locus	UA	Name of farmers' rice varieties
1	1	RM259	1	Kalakrushna
2	1	RM5	1	Ramakrushna Bilasha
3	1	RGNMS135	3	Bada Champa, Karpurjira, Kalkati
4	2	RGNMS249	1	Kalkati
5	2	RGNMS169	1	Kapa Anthi
6	4	RGNMS20	4	Sunakhadi, Akul, Rangahajari, Haladi Sapuru
7	4	RM237	2	Bhasa Manik, Bunde
8	5	RM507	1	Abhirman
9	5	RM433	1	Mahipal
10	6	RGNMS167	3	Magur Manji, Krushana Bhog, Haladi Sapuru
11	7	RM11	1	Damodar Bhoga
12	8	RM284	2	Lalgundi, Ratan Mali
13	9	RM316	1	Karpurmoti
14	9	RGNMS198	2	Kunda Dhan, Sathia
15	9	RM552	3	Jaifulla, Jayaphula, Dhobgaini
16	10	RGNMS27	3	Basa Bhog, Luchai, Abhirman
17	10	RM171	1	Machhakanta
18	11	RM144	1	Patini
19	12	RM277	2	Pimpudi Basha, Parbat Jeera
		Total	34	

nucleotide (Mean: 0.782, $n = 3$) (Supplementary Table S4). A positive correlation was observed between the number of alleles amplified and PIC values ($r = 0.755$, $P < 0.0000001$) as well as the between number of repeats of simple motif and PIC value ($r = 0.397$, $P = 0.0165$) (Supplementary Table S6).

3.3. Homozygosity, heterozygosity and heterogeneity

The automated fragment analyzer system utilized in this study recognizes the fluorescence emission wavelength and provides different peaks for different amplicons of a given product. Since, rice is diploid; allele calling was opted for 2. For any given loci when a particular sample produced one peak was understood as homozygous for that locus. However, when two peaks were available and of (i) similar height, those were taken as heterozygous but (ii) peaks with highly different height were treated as heterogeneous.

The complete set of 607 farmers' varieties amplified a total of 25,396 alleles with an average of 41.84 alleles per variety. Two varieties Bhulo (Bargarh) and Ranisaheb (Bolangir) amplified lowest number (i.e., 31) alleles, whereas Parbat Jira (Kalahandi) amplified highest number alleles (i.e., 60 alleles). Five out of 607 varieties, i.e., FV-064, FV-345, FV-465, FV-513 and FV-517 were homozygous for all the 36 SSR loci. Further, 22 varieties showed heterozygosity at one or more loci. One variety, FV-147 showed heterozygosity for 9 loci (maximum) whereas another variety, FV-573 was heterozygous for only one locus. As expected in case of farmers' varieties, 580 varieties were detected to be heterogeneous from the difference of peaks generated by the tested set of markers. FV-509 was identified to be highly heterogeneous with an average difference in the peak of 1.31. Further, a total of 296 varieties were identified to be in a high heterogeneous group with an average difference in their peaks of >0.500. FV-482 showed heterogeneity for a maximum of 24 loci followed by FV-340 and FV-556, which were heterogeneous at 23 loci. Moreover, 82 varieties were identified to be heterogeneous with 50 % of the tested markers. RM552 could detect heterogeneity in maximum (269) number of varieties. However, markers like RGNMS1, RGNMS190 and RGNMS38 could not detect heterogeneity in any of the varieties.

3.4. Genetic diversity, cluster and population structure analysis

High genetic diversity was found among 607 farmers' varieties, which varied from 0.021 (FV180-FV185) to 0.988 (FV297-FV028, FV297-FV131) with an average of 0.653 based on the SSR data (Supplementary Table S7). The pair-wise genetic distance ranged from 1.000 (FV002-FV003, FV157-FV158, FV268-FV271, FV269-FV270, and FV269-FV271) to 6.000 (FV494-FV297, FV561-FV297, and FV593-FV297) with an average of 4.576 (Supplementary Table S8). The unrooted neighbour-joining dendrogram constructed based on Nei's genetic distance grouped these varieties into different three major clusters. Further, each major cluster is sub-grouped into a different number of sub-clusters (Fig. 2a). PCA analysis also grouped all the 607 varieties into 3 major clusters (Fig. 2b).

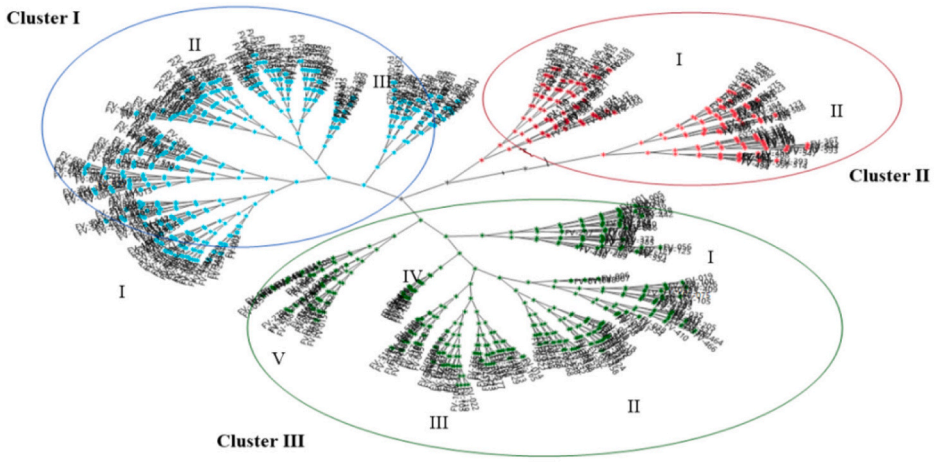
STRUCTURE analysis identified two subpopulations, SP1 and SP2 at K = 2 (Fig. 2c) among 607 farmers' rice varieties. The highest log likelihood value (K) calculated based on Evanno et al. (2005) was K = 2 (Supplementary Fig. 1). This indicated that the entire population could be divided into two sub-populations, i.e., SP 1 with a membership percentage of 82.6 %, while SP 2 with 17.4 %. The fixation index (F_{ST}) values of sub-populations were 0.190 and 0.194 for SP1 and SP2, respectively, with an average of 0.192 (allele frequency divergence = 0.211), indicating the existence of a moderate population structure in our population (Supplementary Table S9). However, an average genetic diversity of 0.422 and 0.552 were recorded between the varieties of subpopulation SP1 and SP2, respectively. Further, 486 out of 607 varieties were assigned to the subpopulation 1 with a membership probability of >0.8. 79 varieties were retained in subpopulation 2, while 42 varieties were detected to be admixtures with a membership probability of <0.8 (Fig. 2c). A relatively higher number of admixtures were detected in the sub-population 1 due to presence of higher number of varieties in this group. However, the AMOVA revealed significant molecular variation ($P < 0.001$) within the subpopulations identified from population structure analysis (for K = 2) of 607 landraces. Of the total variation, 25 %, 47 % and 28 % were recorded among the subpopulations, among the varieties of each subpopulation and within the entire set of varieties, respectively (Supplementary Table S10). When we increased the K value of STRUCTURE to 3, and 3 sub-populations were found (Fig. 2a), which was consistent with the NJ tree and PCA (Fig. 2a, b). 17.87 %, 2.27 %, 1.26 %, and 1.04 % of the molecular variances were explained by first, second, third, and fourth coordinates, respectively (Supplementary Table S11).

3.5. Genetic diversity based on geographical location

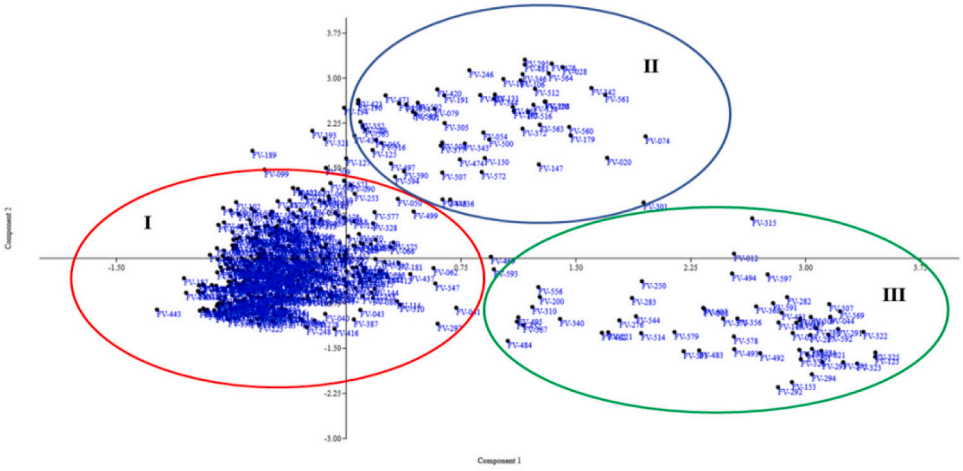
All the 607 farmers' varieties collected across different regions of Odisha were grouped into 28 districts (administrative geographical units) and 9 different agro-climatic zones (Gangopadhyay, 1991). Our collection represents 93 % of the total districts (i.e., 30) of Odisha. We separated the varieties according to their region of the collection to estimate the genetic diversity available within a particular district. We had the highest collection from Bolangir district (97 varieties) as compared to Jharsuguda and Deogarh (3 varieties)), and Malkangiri (2 varieties)). The total number of alleles, the number of effective alleles and the number of private alleles were positively related to the number of varieties collected from each district. On an average, the number of locally common alleles found in <25 % of the population ranged from 0.056 in Deogarh district to 2.028 in Malkangiri district (Supplementary Table S12). Since, only 2 varieties were available from Malkangiri district, a low level of genetic diversity (He = 0.257) was detected in the germplasm collection. However, despite having only 13 varieties, Khurdha district has maximum gene diversity (He = 0.563) followed by Rayagada (He = 0.555, 11 varieties) and Kalahandi (He = 0.553, 55 varieties). The Eastern Ghats zone comprising of Ganjam, Kalahandi, Kandhamal, Nuapada and Rayagada districts with an average possessed a highest gene diversity (He = 0.519) for the farmers' rice varieties grown in these regions. The gene flow within the set of varieties was recorded high (Nm = 0.167) in Nuapada and Jharsuguda district, while Jajpur district showed a low (Nm = 0.036) proportion of gene flow (Table 3). The fixation index was inversely related to gene flow within the varieties of a district (range: 0.87 in Jajpur to 0.60 in Jharsuguda). Further, varieties collected from three districts, i.e. Kalahandi, Bolangir and Sambalpur showed a 100 % polymorphism in the 36 SSR loci. However, in Nuapada district, a significantly lower percentage ($P = 0.944$) of polymorphism was detected. The analysis of molecular variance (AMOVA) could detect only 1 % and 2 % variation among 9 zones and 28 districts at $P > 0.001$, respectively (Table 4). Further, 63 % of the total variation was explained among individuals and 34 % was explained within individuals.

The pairwise Nei's unbiased genetic distance among districts ranged from 0.018 (Kalahandi and Bolangir) to 0.389 (Malkangiri and Rayagada) with an average of 0.100 (Supplementary Table S12). The neighbour-joining dendrogram broadly grouped the 28 districts in two clusters, i.e., I and II. Cluster I contained 15 districts, namely, Deogarh, Anugul, Bargarh, Koraput, Bolangir, Sundargarh, Nuapada, Kanhdhamal, Sambalpur, Boudh, Sonepur, Khurda, Rayagada, Kalahandi and Ganjam, whereas, 13 districts (Nayagarh, Cuttack, Dhenkanal, Keonjhar, Nabrangpur, Jharsuguda, Mayurbhanj, Balasore, Jajpur, Malkangiri, Kendrapada and Bhadrak) were grouped in Cluster II (Fig. 3). Cluster II was mainly comprised of North-Eastern districts of Odisha. Most of the districts situated in Southern and Western part of Odisha were clustered in Cluster I.

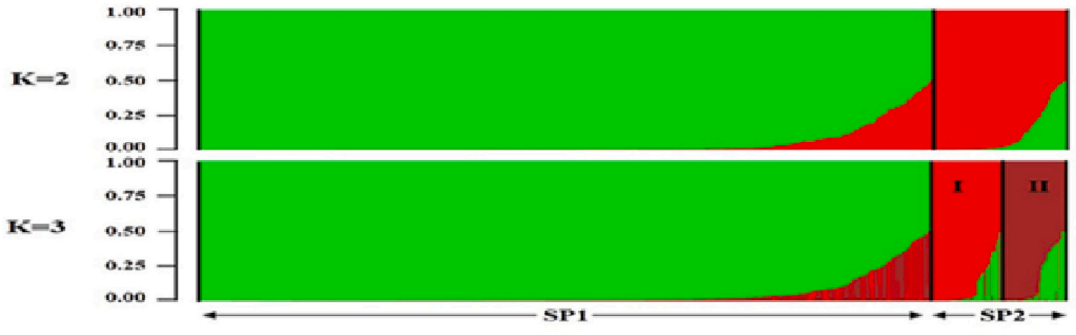
When we assigned STRUCTURE population to the varieties of individual district, 17 out of 28 districts were having varieties from both SP1



(a)



(b)



(c)

Fig. 2. Population structure and genetic relationship among 607 farmers' rice varieties based on the 36 microsatellite data.
a) NJ tree based on pairwise Nei's unbiased genetic distance; b) Two dimensional PCA analysis; and c) Bayesian model-based population STRUCTURE analysis.

Table 3
Genetic diversity parameters of 607 farmers' rice varieties based on 28 geographical districts of collection.

Districts	Co-ordinates	Na	NaF	NPA	NCA	He	F_{ST}	Nm	%P
Deogarh	21.53°N 84.73°E	65	1.81	0.00	0.06	0.30	0.85	0.05	61.11
Sundargarh	21.47°N 83.97°E	161	3.06	0.03	0.83	0.45	0.80	0.06	97.22
Keonjhar	21.63°N 85.58°E	112	2.53	0.03	0.42	0.38	0.80	0.06	75.00
Mayurbhanja	21.933°N 86.733°E	213	2.83	0.06	1.39	0.44	0.82	0.06	97.22
Balasore	21.49°N 86.93°E	139	2.61	0.03	0.58	0.39	0.80	0.06	86.11
Bhadrak	21.06°N 86.50°E	129	2.86	0.00	0.58	0.39	0.86	0.04	77.78
Jajpur	20.85°N 86.333°E	173	3.94	0.14	1.14	0.48	0.87	0.04	94.44
Kendrapada	20.525°N 86.475°E	117	3.25	0.00	0.50	0.40	0.73	0.09	83.33
Khurdha	20.18°N 85.62°E	156	3.36	0.06	0.97	0.56	0.82	0.06	97.22
Nayagarh	20.116°N 85.01°E	118	2.67	0.00	0.25	0.38	0.76	0.08	77.78
Puri	20.47°N 84.23°E	158	2.72	0.03	0.94	0.40	0.81	0.06	88.89
Ganjam	19.38°N 85.07°E	166	3.08	0.03	0.89	0.52	0.76	0.08	94.44
Kalahandi	20.083°N 83.2°E	232	3.61	0.06	1.86	0.55	0.84	0.05	100.00
Kandhamal	20.47°N 84.23°E	179	3.28	0.03	0.97	0.50	0.86	0.04	97.22
Nuapada	20.8167°N 82.5333°E	141	3.28	0.03	0.64	0.46	0.60	0.17	0.94
Raigada	19.17°N 83.42°E	150	3.44	0.00	1.08	0.56	0.83	0.05	97.22
Koraput	18.8083°N 82.7083°E	187	3.33	0.14	1.17	0.49	0.82	0.06	94.44
Nabarangpur	19.23°N 82.55°E	84	2.33	0.00	0.22	0.32	0.78	0.07	66.67
Malkangiri	18.35°N 81.90°E	58	1.61	0.00	0.25	0.26	0.81	0.06	52.78
Baragarh	21°20′N 83°37′E	173	2.92	0.08	0.89	0.46	0.80	0.06	97.22
Boalngir	20.72°N 83.48°E	260	3.53	0.19	2.03	0.52	0.82	0.06	100.00
Boudh	20.84°N 84.32°E	107	2.97	0.00	0.39	0.46	0.77	0.07	88.89
Jharsuguda	21.85°N 84.03°E	69	1.92	0.00	0.11	0.29	0.60	0.17	61.11
Sambalpur	19.48°N 85.49°E	201	3.22	0.08	1.36	0.50	0.78	0.07	100.00
Sonepur	20.83°N 83.92°E	160	3.17	0.08	0.86	0.50	0.80	0.06	97.22
Angul	20°47′50″N 85°1′26″E	123	3.42	0.03	0.56	0.48	0.78	0.07	88.89
Cuttack	20.27°N 85.52°E	129	2.89	0.00	0.64	0.39	0.78	0.07	83.33
Dhenkanal	20.67°N 85.6°E	181	2.89	0.03	1.03	0.42	0.84	0.05	88.89

Na: Number of alleles, NaF: No. of different alleles with a frequency ≥ 5 %, NPA: No. of private alleles unique to a single population, NCA: No. of locally common alleles (Freq. ≥ 5 %) found in 25 % or fewer populations, He: Unbiased Nei's genetic diversity, F_{ST}: Fixation index, Nm: Gene flow estimated from Fst = 0.25(1 − Fst) / Fst., % P: Percent of polymorphic alleles.

Table 4
Analysis of molecular variance (AMOVA) based on 28 districts and 9 agro-ecological zones.

Source	df	SS	MS	Est. Var.	%	F-statistics
Among zones	8	284.675	35.584	0.100	1 %	Frt = 0.011
Among districts	19	402.048	21.160	0.161	2 %	Fsr = 0.017
Among individuals	579	8662.159	14.961	5.904	63 %	Fst = 0.028
Within individuals	607	1913.000	3.152	3.152	34 %	Fis = 0.652
Total	1213	11,261.882	–	9.317	100 %	Fit = 0.662

df: Degree of freedom, SS: Sum of squares, MS: Mean of squares, Est. Var.: Estimated variation.

and SP2 (Supplementary Fig. 2). All the varieties from Balasore, Cuttack, Deogarh, Jharsuguda, Keonjhar, Malkangiri, Nabrangpur and Nayagarh belong to SP1. Varieties from Puri, Kendrapada and Bhadrak also represented SP1 with a 2, 1 and 1 admixture, respectively. None of the districts were detected to be having varieties of both the STRUCTURE population in an equal proportion. Population 1 was predominant in all the districts while Raigada district contained a maximum percentage (36.4 %) of varieties from SP2 group followed by Khurda district (33.3 %). The percentage of admixture varieties ranged between 2.3 % in Mayurbhanj and 16.7 % in Khurda. On an average, mostly the Southern and Western districts were having most varieties from SP2 in comparison to the Eastern and Northern districts. The isolation by distance based on Nei's genetic distance and geographical distance could not establish any significant co-relation ($r = 0.257, p < 0.021$) (Fig. 4). Further, when we removed a few of the out layers (districts not being separated from their nearby districts with a distinct geographical distance), it followed an isolation by distance pattern with $r = 0.512$ and $p < 0.001$.

3.6. Phenotypic diversity

The basic statistics for all the 53 phenotypic traits like minimum, maximum, mean, median, standard deviation, standard error, variance, and coefficient of variation (CV) and significance were calculated and were given in Table 5. The frequency distributions of all the traits were classified based on Descriptors for Rice (IRRI-IBGR 1980). 53 traits harbored a total of 195 variables with an average of 4.217 per trait. 50 (94.34 %) out of 53 traits showed variations among 607 farmers' rice varieties. Three traits, leaf ligule, male sterility, and presence of amylose in endosperm did not reveal any difference among farmers' rice varieties. In the polymorphic traits, the number of variables within a trait studied varied from 2 (anthocyanin colouration, anthocyanin colouration of leaf sheath, leaf auricle, leaf collar, anthocyanin colouration of leaf, leaf ligule, male sterility, anthocyanin colouration of nodes, anthocyanin colouration of internodes, awn of panicle, presence of secondary branching in panicle, phenol reaction of lemma, presence of amylose in endosperm and aroma in decorticated grain) to 9 (colour of lemma and palea and colour of decorticated grain) (Table 5). Coefficients of variance (CV) varied from 0.0 (leaf ligule, male sterility and presence of amylose in endosperm) to 131.19 (decorticated grain: Aroma) with an average of 49.51. The median indicates the predominant phenotypic traits among 607 genotypes (Table 5, Supplementary Table 13). For example, green (median:1) was the dominant sheath colour while dark (median: 7) was the dominant leaf colour. Similarly, the absence of anthocyanin in the leaf and leaf sheath, the presence of pubescence strongly in the leaf blade, the presence of auricles in the leaf, etc. are dominant phenotypes (Table 5, Supplementary Table 13). Supplementary Table 13 shows the number and percentage of genotypes showing different morphological/phenotypic traits. High variations (0–100 %) were observed in phenotypic traits. 100.00 % of genotypes showed the presence of ligule in the leaf and amylose in endosperm while all the 607 genotypes showed the absence of male sterility (Supplementary Table 13). Fifteen phenotypic traits were exhibited by >80.00 % and <99.99 % of 607 genotypes. These traits are the absence

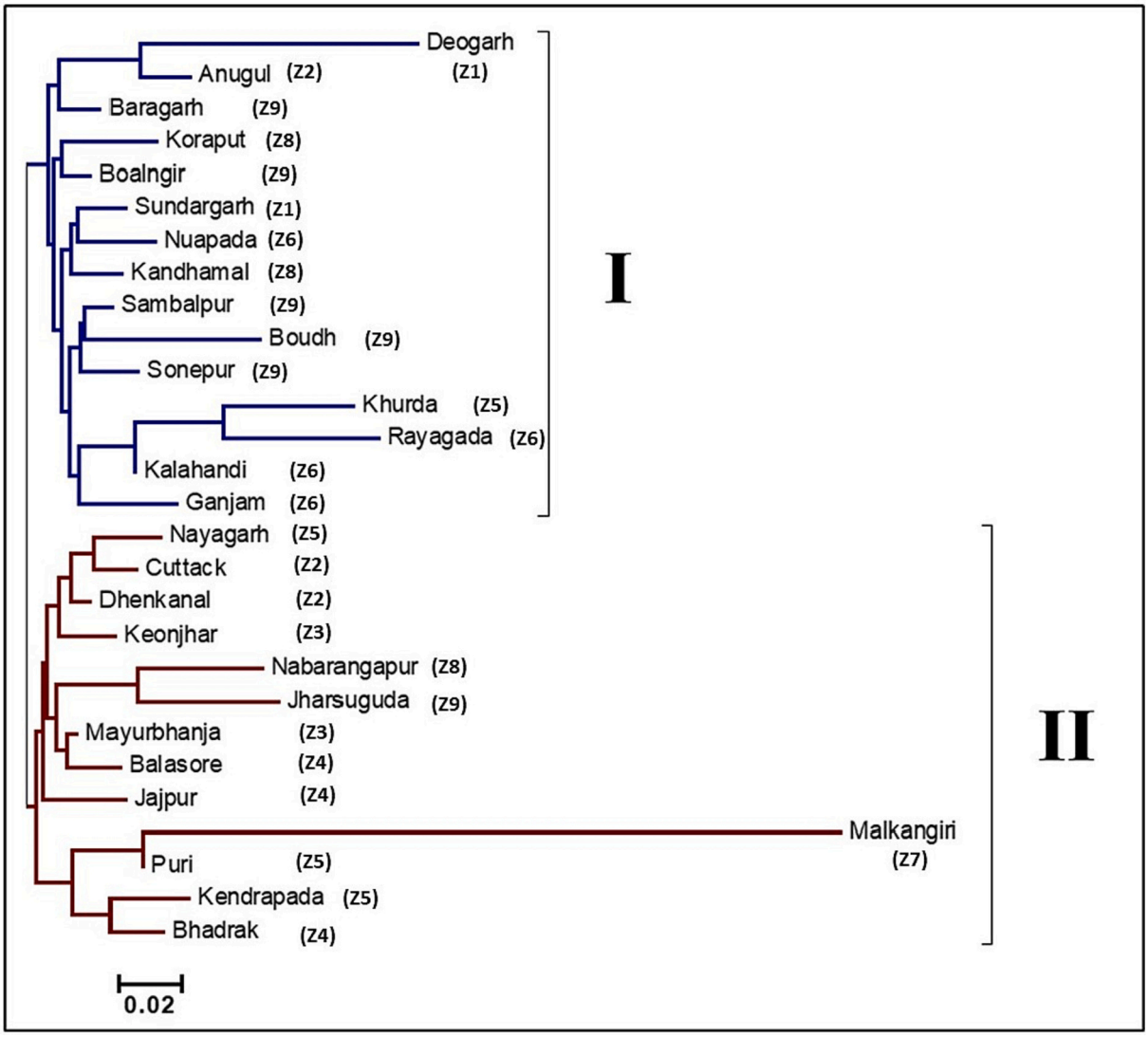


Fig. 3. Unrooted neighbour-joining dendrogram showing genetic relationship among 28 geographical districts based on Nei's pairwise unbiased genetic distance based on the 36 microsatellite data.
Numbers in parenthesis indicates nine agro-climatic zones (Z1 to Z9): Z1-North Western Platetue Zone, Z2-Mild Central Table Land Zone, Z3-North Central Plaetau Zone, Z4-North Eastern Coastal Zone, Z5-East and South Eastern Coastal Plain Zone, Z6-North Eastern Ghat Zone, Z7-South Eastern Ghat Zone, Z8-Eastern Ghat High Land Zone, Z9-West Central Table Land Zone.

of anthocyanin colour in leaf (80.23 %), presence of auricles in the leaf (99.84 %), presence of collar in the leaf (99.44 %), split ligule in the leaf (84.02 %), white ligule colour in the leaf (81.55 %), long (>45 cm) leaf blade (88.30 %), absence of anthocyanin colour in keel lemma (90.77 %), absence of anthocyanin colour in apex lemma (90.77 %), presence of thick (>0.55 cm) stem (86.99 %), panicle (11−20) number per plant (85.67 %), absence of awn in the panicle (85.50 %), presence of secondary branching (99.84 %), straw colour of sterile lemma (82.21 %), amylose content (20–25 %) in the endosperm, and absence of aroma in grains (85.97 %)(Supplementary Table 13). Analysis of variance showed highly significant differences ($p < 0.001$) among 607 farmers' rice varieties for all the 50 polymorphic agro-morphological traits.

All the 50 polymorphic morphological traits showed either significantly positive or negative Pearson's correlation coefficient with the few other traits except three non-polymorphic traits (leaf ligule, male sterility, and presence of amylose in endosperm) (Supplementary Table S14). For example, the sheath colour of basal leaf showed a significant positive correlation with nine traits while a significant negative correlation with five traits. No correlation was found in leaf ligule, male sterility, and presence of amylose in endosperm with other traits. In the present study, 53 morphological traits were measured among the farmers' varieties and analyzed for correlation to know the relationship among these traits. The correlation analysis revealed a significant relationship between grain weight, grain length, grain width, panicle length, and panicle branching with most of the leaf, stem, panicle, and grain traits (Supplementary Table S14).

High genetic diversity was detected among 607 farmers' varieties, which varied from 0.035 (FV167-FV475) to 0.991 (FV400-FV404) with an average of 0.656 (Supplementary Table S15). The Nei's genetic distance between varieties varied from 3.0 (FV123-FV273) to 41.01 (FV199-FV202, and FV404-FV408) with an average of 18.43 (Supplementary Table S16). The neighbour-joining tree constructed based on Nei's genetic distance grouped the 607 varieties into 3 major clusters (I, II and III). Cluster III was highly diverse with 5 sub-clusters (Ia, Ib, Ic, Id, Ie) followed by cluster I having 3 sub-clusters (IIa, IIb, IIc), and cluster III having 3 sub-clusters (IIa, IIb, IIc) (Fig. 5a). The principal component analysis also grouped varieties into 3 major clusters (Fig. 5b). Each of the first, second, third, and fourth principal components explained about 68.71 %, 4.22 %, 1.54 %, and 0.63 % of the total variance, respectively (Supplementary Table S17). The first four principal components

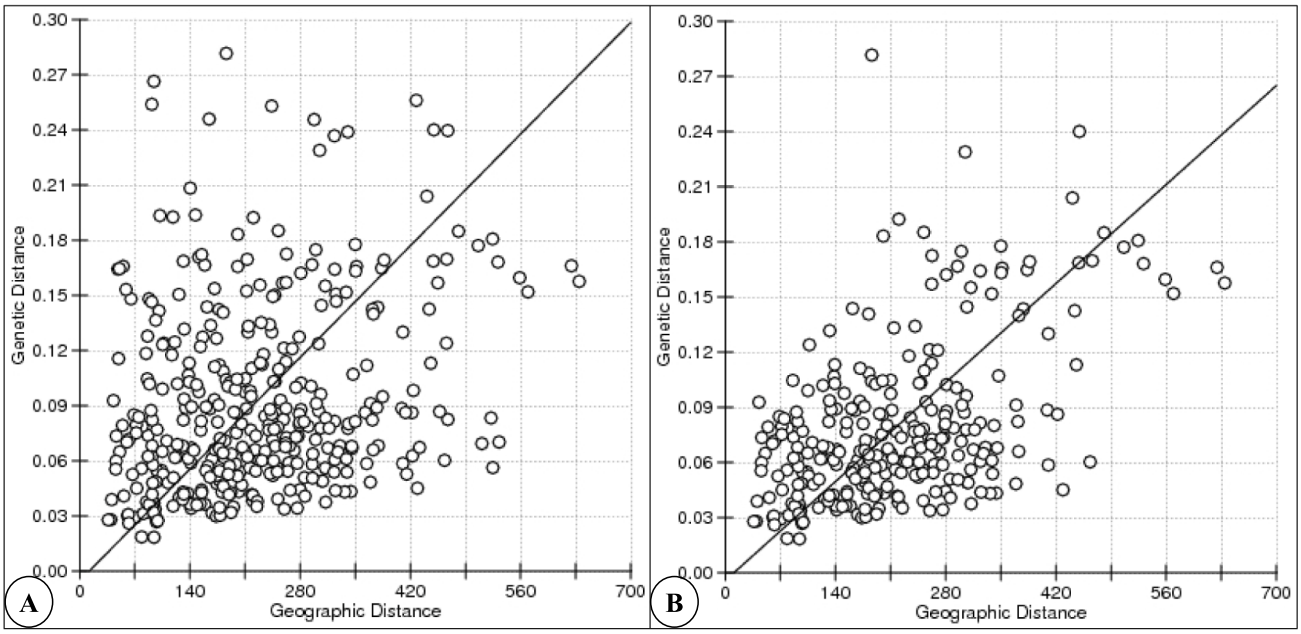


Fig. 4. Genetic isolation by distance for correlation between Nei's genetic distance and geographic distance (km) from 10,0000 randomizations based on the 36 microsatellite data.
A: non-significant correlation with $r = 0.2567$, $P < 0.0210$; B: significant correlation with $r = 0.5122$, $P < 0.0010$.

explained about 75.1 % of the total phenotypic variance. The two-dimensional diagram of 53 morphological traits resulting from principal co-ordinate analysis (PCoA) based on the 607 rice farmers varieties indicated that 4, 19, 24, and 6 traits are dominant in coordinate I, II, II, and IV, respectively (Fig. 6).

4. Discussion

Farmers' rice varieties constitute the intermediate gene pool and are under recent evolution process (Kumar et al., 2010). These have proved to be the most suitable material for mining of novel allele(s) for different traits over years. However, one needs to have a diverse set of landraces to fulfill the aforesaid purpose. Genetic drift, gene flow, natural and artificial selection have resulted in a huge bottleneck in rice over thousands of years of domestication. This has a profound effect on the loss of valuable traits which is a major threat for future crop improvement (Zhu et al., 2007; Caicedo et al., 2007). However, genetic variation has been limited into a certain geographical entity due to genetic isolation. Since landraces are best adapted and perform well in their native area of cultivation, their characterization is of potential use leading to genetic enhancement and strengthening their conservation strategies (Parzies et al., 2004; Roy et al., 2015). In fact, the molecular and morphological level differentiations of farmers' varieties from a particular state like Odisha, India, of which Jeypore tract is associated with the secondary center of origin for rice, would provide insights into their genetic makeup and population structure. Odisha, itself represents 9 different agro-climatic zones and farmers from each of the areas have their own choice set of germplasm according to their preferences and ethnic values (Pachauri et al., 2010; Das, 2012). Therefore, landraces of Odisha maintained by the farmers over decades hold tremendous importance for their characterization.

4.1. Allelic/gene diversity

A total of 363 alleles was detected with 36 SSR loci with an average of 10.08 per locus, which is significantly higher than reported by others (Jain et al., 2006; Ram et al., 2007; Behera et al., 2012, 2013; Choudhary et al., 2013; Zhang et al., 2013; Das et al., 2013; Roy et al., 2015; Salgotra et al., 2015; Shankar et al., 2015; Ahmed et al., 2016; Chandra et al., 2017; Aljumaili et al., 2018; Islam et al., 2018; Jasim et al., 2018; Kumar et al., 2018; Dwivedi et al., 2019; Pathaichidachote et al., 2019; Vanlalsanga et al., 2019; Amegan et al., 2020; Nilthong et al., 2020; Yogi et al., 2020; Embate et al., 2021). Ali et al. (2011) detected a similar number of alleles per locus in the diverse collection of 409 Asian rice accessions using fluorescently labeled 36 SSR markers and ABI Prism DNA Analyzer. Courtois et al. (2012) reported 189 alleles in 425 European rice germplasm collection using 25 SSR markers with an average of 7.56 alleles per locus. Jain et al. (2004) detected 3 to 22 alleles at microsatellite locus in Basmati and non-Basmati varieties using fluorescent techniques in their study. Tiwari et al. (2015) detected 3 to 29 alleles with an average of 12.4 in Indian mini-core rice germplasm using the same panel of fluorescently labeled microsatellite markers. Park et al. (2019) detected 11 to 47 alleles with an average of 25.6 allele per SSR locus in 548 coloured rice germplasm collected from Korea and other countries using fluorescent technique and HAD-GT12™ Genetic Analyzer System. Prasad et al. (2020) observed 2 to 5 alleles with an average of 3.04 per locus in 208 non-Basmati aromatic rice accessions of India using 55 SSR markers. Naaz et al. (2022) evaluated 25 rice accessions including basmati and non-basmati aromatic and non-aromatic popular rice genotypes using 30 SSR markers. The number of alleles varied from 2 to 9 with an average of 4.14. The number of alleles detected by a single microsatellite locus varied from 1 to 31 depending upon the techniques and rice accession used in the studies (Blair et al., 2002; Jain et al., 2004; Garris et al., 2005; Jayamani et al., 2007; Thomson et al., 2009; Kaushik et al., 2011; Behera et al., 2012, 2013; Roy et al., 2015; Ahmed et al., 2016; Aljumaili et al., 2018; Islam et al., 2018; Pathaichidachote et al., 2019; Vanlalsanga et al., 2019; Amegan et al., 2020; Nilthong et al., 2020; Embate et al., 2021; Andarini et al., 2022). This reflects that high allelic diversity present in our collection and demonstrated the potential of the selected set of SSR markers for molecular dissection of rice. The sequence level alteration during the process of domestication resulting in insertion and deletion could be the primary factor for the huge difference in the number of alleles per marker (3 to 27) detected in our study (Gross et al., 2010; Sang and Ge, 2013). We found a positive correlation between number of alleles amplified per locus and number of repeats in simple motif of an SSR

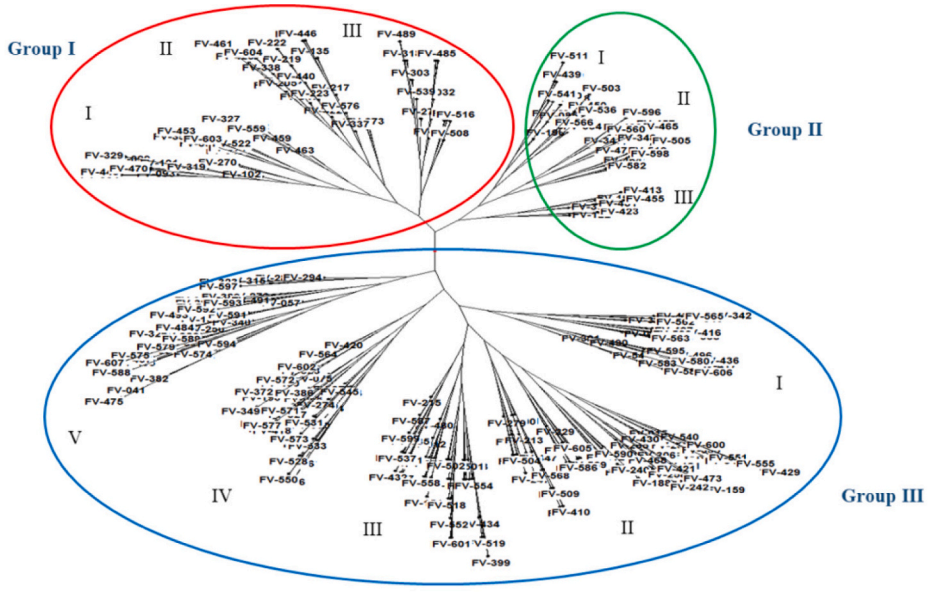
Table 5
Descriptive statistics for 53 morphological traits in 607 farmers' rice varieties.

Sl no	Traits	Min	Max	Mean	Median	SD[a]	SE[a]	Variance	CV[a]	*P* value
1	Basal leaf: Sheath colour:	q	4.00	1.590	1.00	1.089	0.044	1.186	68.509	0.000
2	Leaf: Intensity of green colour	3.00	7.00	6.245	7.00	1.004	0.041	1.007	16.070	0.000
3	Anthocyanin colouration	1.00	9.00	2.582	1.00	3.189	0.129	10.168	123.519	0.000
4	Leaf Sheath: Anthocyanin colouration	1.00	9.00	3.109	1.00	3.528	0.143	12.436	113.319	0.000
5	Leaf: Pubescence of blade surface	3.00	9.00	6.759	7.00	1.002	0.041	1.005	14.829	0.000
6	Leaf: Auricles	1.00	9.00	8.987	9.00	0.325	0.013	0.105	3.613	0.000
7	Leaf: Anthocyanin colouration of auricles	1.00	3.00	1.376	1.00	0.743	0.030	0.552	53.997	0.000
8	Leaf: Collar	1.00	9.00	8.987	9.00	0.325	0.013	0.105	3.613	0.000
9	Leaf: Anthocyanin colouration of collar	1.00	9.00	2.806	1.00	3.347	0.136	11.203	119.301	0.000
10	Leaf: Ligule	9.00	9.00	9.000	9.00	0.000	0.000	0.000	0.000	–
11	Leaf: Shape of Ligule	2.00	3.00	2.840	3.00	0.367	0.015	0.134	12.912	0.000
12	Leaf: Colour of Ligule	1.00	3.00	1.259	1.00	0.584	0.024	0.341	46.367	0.000
13	Leaf: Length of blade	3.00	7.00	6.759	7.00	0.671	0.027	0.450	9.927	0.000
14	Leaf: Width of blade	1.00	7.00	4.012	5.00	1.007	0.041	1.028	25.295	0.000
15	Culm: Attitude	1.00	7.00	2.717	3.00	1.065	0.043	1.134	39.201	0.000
16	Time of heading (50 % of plants with panicles)	1.00	9.00	7.738	9.00	1.611	0.065	2.596	20.823	0.000
17	Flag leaf: Attitude of blade (early observation)	1.00	5.00	1.623	1.00	0.975	0.040	0.951	60.111	0.000
18	Spikelet: Density of pubescence of lemma	1.00	9.00	5.824	5.00	1.313	0.053	1.723	22.539	0.000
19	Male sterility	1.00	1.00	1.000	1.00	0.000	0.000	0.000	0.000	–
20	Lemma: Anthocyanin colouration of keel	1.00	9.00	1.494	1.00	1.643	0.067	2.699	109.951	0.000
21	Lemma: Anthocyanin colouration of area below apex	1.00	9.00	1.652	1.00	1.949	0.079	3.798	117.943	0.000
22	Lemma: Anthocyanin colouration of apex	1.00	9.00	3.211	1.00	3.297	0.134	10.873	102.695	0.000
23	Spikelet: Colour of stigma	1.00	5.00	2.071	1.00	1.765	0.072	3.115	85.234	0.000
24	Stem: Thickness	3.00	7.00	6.651	7.00	0.966	0.039	0.943	14.592	0.000
25	Stem: Length (excluding panicle; excluding floating rice)	1.00	9.00	5.626	7.00	2.150	0.088	4.673	38.502	0.000
26	Stem: Anthocyanin colouration of nodes	1.00	9.00	3.056	1.00	3.499	0.142	12.208	114.207	0.000
27	Stem: Anthocyanin colouration of internodes	1.00	9.00	3.096	1.00	3.520	0.143	12.394	113.726	0.000
28	Panicle: Length of main axis	3.00	9.00	6.766	7.00	1.074	0.044	1.153	15.871	0.000
29	Flag leaf: Attitude of blade (late observation)	1.00	7.00	3.154	3.00	0.985	0.043	1.115	34.038	0.000
30	Panicle: Curvature of main axis	1.00	7.00	2.974	3.00	1.114	0.045	1.240	37.451	0.000
31	Panicle: Number per plant	3.00	7.00	4.799	5.00	0.731	0.031	0.571	15.763	0.000
32	Spikelet: Colour of tip of lemma	1.00	6.00	2.265	1.00	1.805	0.074	3.310	80.137	0.000
33	Lemma and Palea: Colour	1.00	9.00	2.439	1.00	2.058	0.084	4.263	86.079	0.000
34	Panicle: Awns	1.00	9.00	2.160	1.00	2.819	0.114	7.946	130.518	0.000
35	Panicle: Presence of secondary branching	1.00	9.00	8.987	9.00	0.325	0.013	0.105	3.613	0.000
36	Panicle: Secondary branching	1.00	3.00	1.817	2.00	0.461	0.019	0.212	25.361	0.000
37	Panicle: Attitude of branches	1.00	9.00	3.626	3.00	1.597	0.065	2.551	44.051	0.000
38	Panicle: Exertion	3.00	7.00	5.965	7.00	1.193	0.048	1.423	19.996	0.000
39	Time maturity (days)	1.00	9.00	7.738	9.00	1.611	0.065	2.596	20.823	0.000
40	Leaf: Senescence	3.00	7.00	4.477	5.00	1.298	0.054	1.793	30.078	0.000
41	Sterile lemma: Colour	1.00	5.00	1.325	1.00	0.839	0.034	0.705	63.379	0.000
42	Grain: Weight of fully developed grains 1000	1.00	9.00	4.891	5.00	2.000	0.082	4.034	41.133	0.000
43	Grain: Length	1.00	9.00	3.807	3.00	1.096	0.045	1.202	28.797	0.000
44	Grain: Width	1.00	9.00	4.855	5.00	1.539	0.062	2.368	31.698	0.000
45	Grain: Phenol reaction of lemma	1.00	9.00	3.544	1.00	3.729	0.151	13.902	105.217	0.000
46	Decorticated grain: Length	1.00	9.00	4.239	5.00	2.265	0.092	5.129	53.429	0.000
47	Decorticated grain: Width	3.00	7.00	5.369	5.00	1.181	0.048	1.395	21.998	0.000
48	Decorticated grain: Shape (in lateral view)	1.00	6.00	3.614	4.00	1.201	0.049	1.442	33.222	0.000
49	Decorticated grain: Colour	1.00	6.00	3.094	2.00	2.223	0.090	4.940	71.838	0.000
50	Endosperm: Presence of amylose	9.00	9.00	9.000	9.00	0.000	0.000	0.000	0.000	–
51	Endosperm: Content of amylose	3.00	7.00	4.838	5.00	0.704	0.030	0.534	15.127	0.000
52	Gelatinization temperature through alkali spreading value	1.00	7.00	3.722	3.00	1.200	0.049	1.439	32.232	0.000
53	Decorticated grain: Aroma	1.00	9.00	2.122	1.00	2.780	0.113	7.725	131.190	0.000
	Total	90.00	393.00	223.657	206.00	78.762	3.206	169.92	2623.834	0.000
	Mean	1.70	7.42	4.22	3.89	1.486	0.06	3.206	49.51	0.000

[a] SD-Standard deviation, SE-Standard error, CV-Coefficient of variation.

locus. Similar to our observation, Ni et al. (2002) and Yu et al. (2003) found a positive correlation between number of alleles amplified and number of repeats within a microsatellite marker. However, Nagaraju et al. (2002), Juneja et al. (2006) and Behera et al. (2012, 2013) found no direct correlation between the number of repeats and the number of alleles detected in aromatic and wild rice materials, respectively.

Thirty-four (9.36 %) unique alleles were identified from 607 farmers' varieties. Unique alleles are important because they may be diagnostic of a particular variety and useful for breeding purposes. Behera et al. (2012) identified 11 unique alleles (6.51 %) among medicinal rice cultivars at 25 of 30 SSR loci. Davierwala et al. (2000) identified many alleles specific to elite cultivars of India using microsatellite markers. Islam et al. (2018) identified 11 (7.86 %) unique alleles at 11 out of 45 SSR loci in 113 aromatic rice germplasm of Bangladesh. Dwivedi et al. (2019) identified two unique/genotype-specific alleles amplified by two SSR markers RM7173 and RM101 in HPR 2761 and Vasumati, respectively while assessing genetic diversity in basmati and non-basmati aromatic rice genotypes of the northern hill region. Similarly, others also detected unique alleles both in cultivated and wild rices (Davierwala et al., 2000; Saini et al., 2004; Giarrocco et al., 2003; Wong et al., 2009; Behera et al., 2013; Gour et al., 2017). Rare and low frequency alleles are an important source of genetic diversity. Occurrence of rare alleles in rice cultivars may have resulted from the unequal crossing over, translocations or other type mutations. We observed high proportion of rare alleles (65.01 %) and low frequency alleles (23.14 %) in farmers' rice varieties. Similar results were obtained by others (Behera et al., 2012, 2013; Aljumaili et al., 2018; Islam et al., 2018). Tiwari et al. (2015) detected 41.38 % rare alleles in Indian mini-core rice germplasm using



(a)

(b)

Fig. 5. Genetic relationship among 607 farmers' rice varieties based on the 53 morphological traits.
a: NJ tree based on pairwise Nei's unbiased genetic distance; b: Two dimensional PCA analysis. The neighbour-joining tree constructed based on Nei's genetic distance grouped the 607 farmers' rice varieties into 3 major clusters (I, II and III). Cluster III was highly diverse with 5 sub clusters (Ia, Ib, Ic, Id, Ie) followed by cluster I having 3 sub clusters (IIa, IIb, IIc), and cluster III having 3 sub clusters (IIa, IIb, IIc). The numbers in the plot correspond to the rice varieties given in Table 1. b: Principal component analysis also shows 3 clusters.

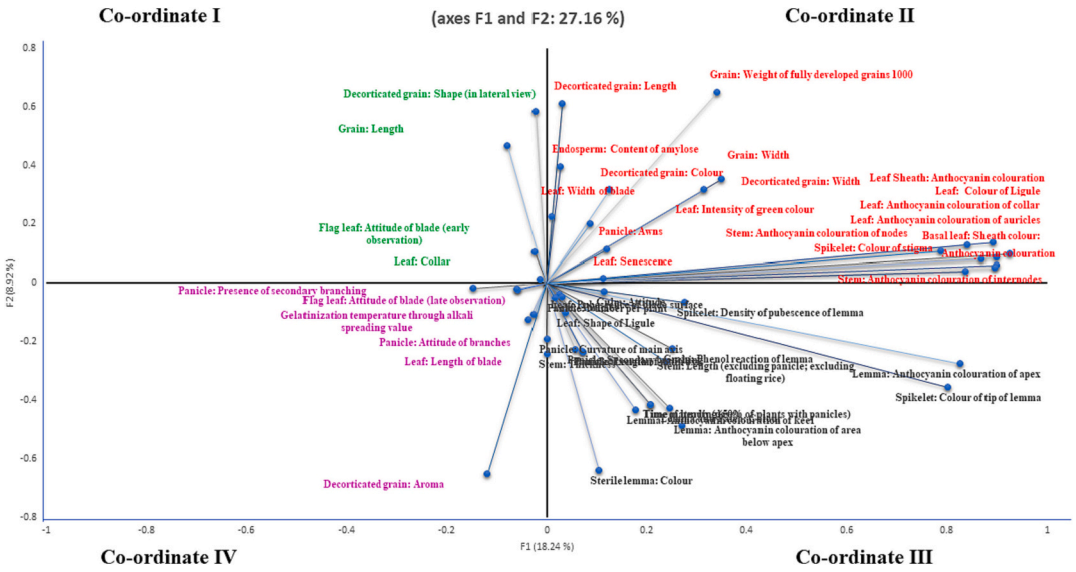


Fig. 6. Two-dimensional diagram of 53 morphological traits resulting from principal co-ordinate analysis (PCoA) in the 607 rice farmers' rice varieties.
Co-ordinate I: decorated grain shape, grain length, flag leaf: attitude of blade(early observation) and leaf collar (4 traits).
Co-ordinate II: decorated grain length, grain: weight of fully developed grains 1000, grain width, etc. (19 traits).
Co-odinate III: sterile lemma colour, spikelet: colour of tip of lemma, spikelet: density of pubescence of lemma, etc. (24 traits).
Co-ordinate IV: decorated grain aroma, leaf: length of blade, panicle: attitude of branches, etc. (6 traits).

same panel of fluorescently labeled microsatellite markers. Islam et al. (2018) detected 52 (37.14) rare alleles at 35 out of 45 SSR loci in 113 aromatic rice germplasm of Bangladesh. Kumar et al. (2018) identified six (9.84 %) rare alleles in 24 aromatic rice genotypes (both basmati and non-basmati) using 25 SSR markers. 597 out of 607 farmers' rice varieties showed the presence of multiple alleles at least at one SSR locus. Rice is a self-pollinated, diploid crop species. The microsatellites usually reveal the single copy and allelic heterogeneity is rare in respect of pure line varieties. Hence, the presence of multiple alleles in cultivars is generally an indication of seed mixtures, mixture of pure lines or residual heterozygosity (Jain et al., 2004). This phenomenon is quite common in landraces that contribute to their broad genetic plasticity to adapt themselves to different agro-climatic conditions in traditional farming systems (Olufowote et al., 1997). Similar to our observation others obtained multiple alleles both in cultivated and wild rices (Garland et al., 1999; Giarrocco et al., 2003; Yu et al., 2003; Lu et al., 2005; Jain et al., 2006; Juneja et al., 2006; Jayamani et al., 2007: Behera et al., 2012, 2013). Further, varieties having similar names collected from same or different areas were detected with allelic variation. This could be possibly due to seed exchange and further selection of base material conducted by farmers to suit their local practices (Roy et al., 2015).

The polymorphism information content (PIC) value reflects the discriminating or resolving power of the marker. The average PIC value of our study for the 36 tested loci, i.e., 0.901 is relatively higher than the reports of Das et al. (2013), who observed an average PIC of 0.75 with 23 microsatellite loci in Eastern and North Eastern rice landraces of India. Naaz et al. (2022) evaluated 25 rice accessions including Basmati and non-Basmati aromatic and non-aromatic popular rice genotypes using 30 SSR markers. They observed lower PIC values which ranged from 0.07 (RM507) to 0.83 (RM514) with an average of 0.46. Similarly, Prasad et al. (2020) observed lower PIC values, which ranged from 0.171(RM477) to 0.721(RM276) with an average of 0.54 in 208 non-basmati aromatic rice accessions of India using 55 SSR markers. The PIC value was also significantly higher than the previous reports (i.e., 0.46: Lu et al., 2005; 0.64: Juneja et al., 2006; 0.64: Joshi et al., 2010; 0.49: Courtois et al., 2012; 0.811: Behera et al., 2012; 0.66: Behera et al., 2013; 0.62: Roy et al., 2015; 0.405: Salgotra et al., 2015; 0.54: Shankar et al., 2015; 0.74: Tiwari et al., 2015; 0.671; Ahmed et al., 2016; 0.62: Ashraf et al., 2016; 0.63: Aljumaili et al., 2018; 0.291: Islam et al., 2018; 0.63: Jasim et al., 2018; 0.465: Kumar et al., 2018; 0.348: Dwivedi et al., 2019; 0.36: Mishra et al., 2019; 0.56: Pathaichidachote et al., 2019; 0.59: Vanlalsanga et al., 2019; 0.53: Amegan et al., 2020; 0.605: Nilthong et al., 2020; 0.479: Yogi et al., 2020; 0. 31: Embate et al., 2021; 0.62: Andarini et al., 2022). However, Park et al. (2019) observed higher PIC values (i.e., 0.913) in 548 coloured rice germplasm collected from Korea and other countries using fluorescent technique. Similarly, Chandra et al. (2017) detected higher PIC values (i.e., 0.946) in 30 rice genotypes including basmati, non-basmati aromatic, and non-aromatic elite cultivars using 20 SSR markers. The high PIC value detected in our study could be due to the inclusion of fluorescently labeled RGNMS markers along with RM markers and the presence of high genetic diversity. We detected a positive correlation between the total number of alleles/loci and their PIC values, which agrees with Jain et al. (2006) and Behera et al. (2012, 2013). The total number of alleles and PIC values of microsatellite markers with different repeat motifs varies in different germplasm set used. SSR loci having mixed repeats showed higher rates of polymorphism as compared to that of trinucleotide, di-nucleotide and tetra-nucleotide repeat motifs (Kaushik et al., 2011; Behera et al., 2013). In contrast, we detected higher polymorphism for di-nucleotide repeats than tri, tetra and mixed motif repeats, which is similar to the results of Cho et al. (2000) and Jain et al. (2004).

Ecological and evolutionary impacts have led to the accumulation of rich genetic diversity in the farmers' rice varieties over years of selection (Hamrick and Godt, 1996; Choudhury et al., 2013). The allelic/gene diversity ($He = 0.502$) detected with all the 36 loci is higher than the Thai rice collection ($N = 101$, $He = 0.476$, Chakhonkaen et al., 2012), Taiwan rice varieties ($N = 34$, $He = 0.41$, Lin et al., 2012), North-East Indian rice core collection ($N = 701$, $He = 0.382$, Choudhury et al., 2014), Assam upland rice collection ($N = 100$, $He = 0.393$, Rathi et al., 2014), aromatic rice accessions of Malaysia ($N = 53$, $He = 0.36$, Ahmed et al., 2016), Korean rice germplasm ($N = 130$, $He = 0.08$, Amegan et al., 2020), but lower than aromatic and quality rice accessions of North-eastern regions of India ($N = 107$, $He = 0.67$, Roy et al., 2015), Indian rice mini-core rice germplasm ($N = 98$, $He = 0.758$, Tiwari et al., 2015). Thai and exotic rice accessions ($N = 167$, $He = 0.59$, Pathaichidachote et al., 2019), and the upland rice varieties of Thailand (($N = 98$, $He = 0.664$, Nilthong et al., 2020). Roy et al. (2015) assessed allelic and

genetic diversity among 107 aromatic/quality rice accessions collected from north-eastern states of India and a set of 67 structured global rice accessions representing *indica, aus, aromatic, temperate japonica, tropical japonica,* and admixed cultivars originating from India and its neighbouring countries using 30 microsatellite markers. Because of inclusion of different subspecies and ecotypes, moderately high allelic/gene diversity (0.67) was observed. Tiwari et al. (2015) developed mini-core rice germplasm containing 98 varieties from a core set of 6912 rice accessions collected from the National Gene Bank of India based on the genotypic data of the same set of 36 fluorescently labeled microsatellite markers. Mini-core retained 94 % of alleles present in the core set. Hence, higher allelic/gene diversity was observed. Pathaichidachote et al. (2019) evaluated gene and genetic diversity among 167 Thai and exotic rice accessions which included both *indica* and *japonica* subspecies with 49 SSR markers. The inclusion rice accession of *indica* and *japonica* subspecies might be responsible for moderately high gene diversity (0.59). The *indica* group of rice is reported to be more diverse than the *japonica* group (Yonemaru et al., 2012). However, the gene diversity in our collection was comparable with the worldwide *indica* collection (*He* = 0.55) (Garris et al., 2005). In our study, also these markers were validated as highly informative, which could serve as a reference set of markers for the assessment of genetic diversity in Indian rice germplasm. Though the diversity values detected with RGNMS markers were higher than that of RM marker, there was no significant difference between the two sets of markers. However, higher genetic diversity detected in the genic non-coding regions provides insights into the functional genetic alteration in farmers' rice varieties and understands their evolutionary adaptation potential to various agro-ecological conditions (Parida et al., 2009). Markers with tri-nucleotide repeats amplified the higher number of alleles (average: 12) but a high level of genetic diversity was detected with the di-nucleotide repeat markers.

The remarkable size difference between smallest and largest allele detected with the tested loci supports the idea that microsatellite markers are highly heterogeneous in nature and are therefore widely used in plant characterization (Agarwal et al., 2008; Grover and Sharma, 2011). Gene flow and cross-pollination even in the self-pollinated crop are the major aspects of evolution, which often leads to a heterozygous nature of plant (Chen et al., 2004). The varieties like Parbat Jira (Kalahandi), Krushna Bhog (Sundargarh) and Sapur (Koraput) were highly heterozygous at 22 loci and could possibly be the products of such events. Most the farmers' rice varieties screened in this study were in heterogeneous (580 varieties) state followed by heterozygous (22 varsities) and only a few were homozygous (5 varieties) for the tested loci. These varieties being selected and maintained by traditional farmers over the ages accumulate heterogeneity due to their natural phenomenon of evolution and constitute major gene pool of rice (Deb, 2005).

The allelic comparison of farmers' varieties with Indian mini-core collection highlighted that genetic variation available in the present farmers' variety collection is quite comparable with the mini-core collection (Tiwari et al., 2015). This is important in the context that these heritages should be conserved and exploited to harness different valuable genes possibly available within them.

4.2. Genetic diversity and cluster analysis among farmers' rice varieties

A higher degree of genetic diversity was observed among 607 farmers' rice varieties both at molecular (0.653) and phenotypic (0.656) levels. The estimated genetic diversity among 607 farmers at molecular is higher than the report of Kim et al. (2014), revealing the rich genetic diversity that exists in farmers' rice varieties. Embate et al. (2021) observed higher genetic diversity (i.e., 0.88) in 43 pigmented traditional rice varieties of Philippines using 47 agro-morphological (22 qualitative and 25 quantitative) traits, grain chemical properties, reaction to diseases and insects, and 64 SSR markers. Amegan et al. (2020) revealed lower (0.58) genetic diversity in 130 Korean rice germplasm using SSR markers. 53 morphological traits harbored a total of 195 variables with an average of 4.217 variables per trait. However, Ahmed et al. (2016) identified higher (0.77) genetic diversity in 31 farmers' varieties of Bangladesh using 45 SSR markers. Similarly, Tiwari et al. (2015) identified a higher level of genetic diversity (0.76) in the Indian mini-core rice germplasm. Few studies are available, where farmers' grown landraces/traditional rice varieties have been studied extensively both at molecular and phenotypic levels. In our study, 50 out of 53 morphological traits showed polymorphism and significant differences among 607 farmers' rice varieties. Ahmad et al. (2015) studied genetic diversity among 42 coloured rice genotypes using 25 microsatellite markers and 15 agro-morphological traits. All the 15 agro-morphological traits showed significant differences at $p = 0.05$ and $p = 0.01$, indicating the presence of diversity among the 42 coloured rice genotypes. Hien et al. (2007) assessed genetic diversity among 36 aromatic rice cultivars collected from Asia using 22 morphological characters with 101 morphometric descriptors. Except for ligule colour, the rests of the traits were polymorphic. Ray et al. (2013) indicated the presence of immense phenotypic diversity among 414 farmers' landraces based on the 29 phenotypic traits. Barhate et al. (2021) assessed genetic diversity among 45 aromatic genotypes based on 13 quantitative and three cooking quality traits. All the phenotypic traits showed highly significant differences among genotypes.

The neighbour-joining tree constructed based on Nei's genetic distance grouped the 607 farmers' varieties into 3 major clusters (I, II and III) based on the both molecular and phenotypic data. All the rice varieties could be clearly differentiated. Rice varieties of similar names collected from same or different areas showed variability for one or more locus. For instance, three accessions of Akul collected from three different areas, i.e., Bargarh, Sambalpur and Bolangir showed allelic variation at four microsatellite loci, RM259, RGNMS20, RM433 and RGNMS141. Principal component analysis also grouped varieties into 3 major clusters based on both molecular and phenotypic data. First four principal components explained about 30.55 % and 75.1 % of the total variance at molecular and phenotypic levels, respectively. Ahmed et al. (2016) studied genetic diversity among 31 rice genotypes with similar names of Bangladesh using 45 microsatellite markers, and 14 agro-morphological and physico-chemical traits. The analysis revealed that no duplicate genotype was existed. The PCA analysis indicated that the first five components accounted for 82.9 % of the total variations. Ahmad et al. (2015) studied genetic diversity among 42 coloured rice genotypes using SSRs and agro-morphological traits. Clustered analysis grouped the 42 coloured rice genotypes into 7 and 4 clusters, respectively based on the molecular and agro-morphological data. Further, first three principle components of PCA explained 75.28 % of total variations at molecular level while 83.76 % of total variation was explained by first four components of PCA at phenotypic level. Islam et al. (2018) assessed the genetic diversity of 113 aromatic rice germplasm of Bangladesh based on the 12 phenotypic traits and molecular data (45 SSR). Similar to our result, cluster analysis classified into three major groups both at molecular and phenotypic levels which was further confirmed by STRUCTURE and PCA analysis. Prasad et al. (2020) evaluated the genetic diversity of 208 non-Basmati aromatic rice accessions of India using 46 morphological, agronomical, grain quality traits, reaction to pests, and 55 SSR markers. Cluster analysis based on \ Jaccard's similarity coefficients grouped these genotypes into three main clusters. Six, 81, and 121 genotypes were included in clusters I, II, and III, respectively. Cluster II was further grouped into nine sub-clusters while cluster III was grouped into eight sub-clusters. Further, based on the molecular, morphological, agronomical, and quality traits data, all 208 aromatic genotypes were grouped into 20 clusters using Gowers' similarity index. 45 aromatic genotypes could be grouped into 10 clusters based on the 13 quantitative and three cooking quality traits (Barhate et al., 2021). Manjunath et al. (2021) assessed the genetic diversity among 16 aromatic landraces of the Wayanad district of Kerala, one Basmati and two non-aromatic (Aathira and Uma) genotypes using 86

SSR markers. Cluster analysis could differentiate Basmati, Gandhakasal, Jeerakasala, Aathira, and Uma from each other and grouped into five clusters. All the 12 Gandhakasala morphotypes belonged to Cluster III while all three Jeerakasala morphotypes belonged to Cluster IV. Aathira, Uma, and Basmati belonged to Cluster I, II, and V, respectively.

4.3. Population structure

Based on population STRUCTURE analysis, we could differentiate the entire collection into 2 sub-populations which is similar to the reports of Zhang et al. (2011) in the rice core collection from 20 different provinces and Choudhury et al. (2013) in NE Indian rice collection. However, Tiwari et al. (2015) have identified 3 sub-populations in a mini core of Indian rice collections. Varieties from SP1 group were predominant with a membership percentage of 82.6, which might be due to common origin and single domestication process for the varieties of this group. Moreover, 42 admixtures were identified that represented a higher rate of cross-pollination among the varieties of separate sub-populations. A significant molecular variation of 47 % was recorded among individuals, which is in agreement with the earlier studies of Choudhury et al. (2014) who could detect 48 % variation among individuals of NE India. Roy et al. (2015) identified three populations in 107 aromatic and quality rice accessions of North-eastern regions of India. Salgotra et al. (2015) identified 5 subpopulations (A, B, C, D, and E) in 142 basmati rice genotypes consisting of landraces, farmer's cultivars, elite cultivars, and advanced breeding lines from Basmati growing regions of India using 40 SSR markers. Similarly, Islam et al. (2018) revealed the existence of population structure in 113 aromatic rice germplasm of Bangladesh. They identified three populations (P1, P2 and P3) with a majority of germplasm in P1. This grouping agreed with genetic distance-based clustering and PCA. Pathaichidachote et al. (2019) and Vanlalsanga et al. (2019) identified two sub-populations (P1 and P2) in 167 Thai and exotic rice accessions, and 65 local rice cultivars of North-eastern regions of India, respectively. Nilthong et al. (2020) observed two sub-populations (P1 and P2) in 98 upland rice varieties of Thailand. However, Park et al. (2019) detected 7 subpopulations among 548 coloured rice germplasm collected from Korea and other countries. The sub-population identified in our study was in agreement with distance-based clustering and was validated by principal co-ordinate analysis. However, varieties from SP2 were separated into two distinct groups when K value was increased from 2 to 3 in population STRUCTURE. Yogi et al. (2020) evaluated 13 rice genotypes including 4 aerobic, 6 low-land basmati, and three high-yielding *indica* rice genotypes using 129 SSR markers. They identified two groups/subpopulations through both PCA and STRUCTURE analysis. Mishra et al. (2019) assessed the genetic diversity of 35 rice accessions including 34 aromatic rice from different parts of India, 10 nonaromatic and one aus (N22) rice genotypes using 55 SSR markers and 17 agro-morphological and yield-related traits. STRUCTURE analysis identified three distinct groups; *indica* (nonaromatic), aus (Nagina 22), and aromatic.

4.4. Genetic diversity of farmers' rice varieties associated with a particular district

Every district of a particular state has its own set of rice landraces according to the farmers' preference to suit the agro-ecological conditions and cultural practices (Das, 2012). Estimation of genetic diversity of a particular district could provide insights into the genetic potential of that area. Moreover, this would lead to understanding the seed exchange and spread of rice germplasm. Out of 28 districts, Bolangir rice collection had the highest number of private alleles followed by Koraput and Jajpur. Bargarh, the nearby district of Bolangir is known as the rice bowl of Odisha. Rapid globalization and the introduction of high-yielding rice varieties have resulted in huge genetic erosion in many parts of Odisha (Deb, 2005). Since, Bolangir district is mainly comprised of rural communities who are much dependent on their local varieties, the highest number of farmers' varieties (97) was available from Bolangir. The varieties collected from Khurdha, Rayagada, Kalahandi, Ganjam, Bolangir, Sambalpur, Sonepur and Koraput were highly diverse with higher genetic diversity value suggesting the need for further exploitation of rice landraces in these regions. Since, Eastern Ghats regions of Odisha are the major biodiversity hot spot (Nayak et al., 2013; Panda et al., 2013), a higher level of diversity for farmers' rice varieties was detected in Eastern-Ghats belt. The western part of Odisha is mainly comprised of tribal communities. Therefore, the impact and adoption of high yielding varieties are comparatively less in these areas (Das, 2012; Padhi and Panigrahi, 2011). Rice landraces cultivated in these parts have undergone a higher rate of artificial selection leading to many alternative forms of a particular variety and hence farmers' varieties collected from Kalahandi, Bolangir and Sambalpur showed 100 % polymorphism for all the 36 markers. We could detect a very less percentage of variation among zones and districts, which could be due to frequent seed exchange between the farmers of Odisha (Pusadee et al., 2009). However, the farmers' varieties of Odisha are highly differentiated which was supported by AMOVA, where 63 % of the total variation was due to differences among varieties. Many of these varieties and the knowledge related to their cultural practices are being spread from one area to the nearby area (Sirabanchongkran et al., 2004; Poudel et al., 2015) which was justified in our study also as a lowest genetic distance was detected between Kalahandi and Bolangir. Both districts share common geographical boundaries. Further, the NJ dendrogram clustered the coastal and eastern districts together. Most of the districts present in western Odisha and close to nearby state Chhattisgarh were grouped in a separate cluster. Considering Jeypore tract as the Secondary Centre for Origin of rice and the regions of western Odisha, Jharkhand and Chhattisgarh as the Centre of origin of *aus* ecotypes of rice (Sharma et al., 2000), the evolution of rice in the western belt of Odisha and then their spread to other parts could be possible. Similar results were recorded by Roy et al. (2015) for the evolution and spread of aromatic short grain rice landraces in Odisha. Moreover, a significant isolation by distance pattern was also detected which justified the probable spread pattern of farmers' varieties in Odisha which were genetically isolated during their spread over the time of domestication. Similar to our study, Singh et al. (2016b) observed 63 % of the total variation among 729 Indian rice varieties using 36 HvSSR markers. Pusadee et al. (2009) also reported isolation by distance in a Thai local rice variety, Bue Chomee. The assignment of STRUCTURE population demonstrated that varieties of 17 districts represented both the identified sub-populations in different proportion and the majority of them were western and southern of Odisha. Further, a higher number of admixed individuals were detected in the western region of the state, suggesting higher rate of introgression in these regions. Therefore, western Odisha having a comparatively higher genetic diversity of farmers' varieties of rice holds enormous potential for their characterization, conservation and utilization in crop improvement.

The grain yield in rice is significantly attributed to panicle and grain traits such as grain weight (Anilkumar et al., 2022), grain length (Nayak et al., 2022), grain width, panicle length (Sah et al., 2022), and panicle branching (Sah et al., 2022), etc. Yield-attributing traits had correlations with other morphological traits due to the multigenic nature of the traits with small and cumulative variance. The significant correlations between grain weight, grain length, grain width, panicle length, and panicle branching with most of the leaf, stem, panicle, and grain traits indicate relationships among the traits for their expression. This is a favorable relation for a plant breeder for selective mating and the selection of segregants for targeted traits. Besides, the selection of traits such as relationships also helps the indirect selection of the favorable traits in breeding pipelines. Such multi-traits correlation helps in finding the best donor and favorable alleles for yield and yield-attributing traits in the population for breeding use. The farmers' varieties that could be used as potential donors for yield improvement are Akula Jalgudi, Jangali Jata, Kakudi Manji, Suantuti. These varieties had high grain

weight, medium grain length, narrow to broad grain width, bear of secondary branches in panicles, strong to cluster secondary branches, medium to long panicle axis length, and medium to high number of panicles per plant.

Presently, we are evaluating these farmers' varieties (FV) against different biotic and abiotic stresses, grain qualities, and nutrient use efficiency. Recently, a research group from our institute evaluated these 600 farmers' rice varieties against brown planthopper using the standard seed box screening method at the seedling stage. Based on the initial screening result, a panel of 106 FV was selected for further evaluation following different parameters of BPH resistance. These varieties were genotyped with 87 gene-linked markers associated with 34 BPH resistance genes. 18 varieties were identified as highly resistant (SES score 1), while 22 were moderately resistant (SES score 3). 10 markers were found to be associated with BPH resistance genes. *Bph6* and *Bph30* exhibited strong resistance to BPH by having a significant association with different phenotypic parameters of BPH resistance (Anant et al., 2021). Further, 600 farmers' rice varieties were tested for the presence of endophytes. A total of 141 endophytes were identified. Some of them enhanced the rice plant growth significantly. Two endophytes were found to be highly effective against the causal organisms of sheath blight (*Rhizoctonia solani*) and seedling blight (*Fusarium* sp.). About 90–100 % inhibition of growth of these pathogens was observed (data not sown). These varieties are being presently used in different breeding and allele mining programs to improve yield, grain quality, and climate resiliency, which is required for sustainable production.

5. Future prospects

Genetic diversity is a prerequisite for breeding plants for desirable traits. Traditional and extinct varieties have ancestry genes that spontaneously mutate throughout the course of long-term cultivation. These mutations are random and cause a huge genetic variation both at the morphological and molecular levels. In this study, the variation observed in the farmers' varieties both at molecular and morphological levels is obvious. This spontaneous variation or mutation has the advantage of allowing the plant to endure the current climatic situation for the long term for survival. Thus, only those mutated plants survived which were coping with the climate vulnerability. As a result, these plants serve as a repository for many unique genes needed for the development of commercial cultivars with increased productivity and climatic resilience. The presence of variation among the farmers' varieties would serve as a source of donors for different traits contributing to yield and survivability under stress. Because, states like Odisha have diverse rice ecology including irrigated to rainfed food-prone ecology, thus bear several kinds of genetic variation in germplasm collected over the states. Besides, it reveals the nature of variation and farmers' varietal dynamics cultivated over a region.

6. Conclusion

The present study provides a better understanding of allelic and genetic diversity and population structure of 607 farmers' rice varieties. Higher genetic diversity among 607 farmers' rice varieties was detected both at molecular (0.653) and phenotypic (0.656) levels, revealing the rich genetic diversity that exists within farmers' rice varieties. Moderately high allelic/gene diversity (0.507) was detected. The number of alleles varied from 3 to 27 alleles with an average of 10.2 per locus. 34 unique, 236 rare, 84 low-frequency, and 44 high-frequency alleles were detected. Based on the Nei genetic distance, all the 607 varieties were grouped into three major clusters, which were also validated by PCA analysis both at molecular and phenotypic levels. Further, based on population STRUCTURE analysis, we could differentiate the entire collection into three subpopulations at $K = 3$. These varieties are highly differentiated, where 63 % of the total variation was because of differences among varieties. Hence, these varieties should be conserved and exploited to harness different valuable genes/alleles available within them. Since landraces are best adapted to different climatic conditions, their characterization is of potential use for strengthening their conservation strategies and genetic enhancement, and use them as donors in breeding programs for the development of climate-resilient rice varieties with a higher yield to cope with climatic changing scenario conditions and improve farmers' income.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We thank Director, ICAR-National Rice Research Institute, Cuttack for providing necessary facilities.

CRediT authorship contribution statement

The study was designed by TM, SRD, PSR and LB. SRD, UD and BCP collected the farmers' rice varieties. SN, SS and SRD carried out morphological trait characterization. SN, SS, AC, KKT, SVACRM carried out genotyping work. PSR, SM, RPS and LB analyzed the data and performed the statistical analysis. PSR, LB and TM have drafted the manuscript. All authors read and approved the final manuscript.

Funding source

The financial supports provided by PPV and FRA, Ministry of Agriculture and Farmers Welfare, New Delhi, Government of India are gratefully acknowledged. The funding agency has no role in the design of the study and collection, analysis, interpretation of data and writing manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.genrep.2022.101719.

References

Agarwal, M., Shrivastava, N., Padh, H., 2008. Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep. 27 (4), 617–631.

Ahmad, F., Hanafi, M.M., Hakim, M.A., Rafii, M.Y., Arolu, I.W., Akmar Abdullah, S.N., 2015. Genetic divergence and heritability of 42 coloured upland Rice genotypes (Oryza sativa) as revealed by microsatellites marker and agro-morphological traits. PLoS ONE 10 (9), e0138246.

Ahmed, M.S.U., Zamani, K., Bashar, Md.K., Shamsuddin, A.K.M., 2016. Agro-morphological, physico-chemical and molecular characterization of rice germplasm with similar names of Bangladesh. Rice Sci. 23 (4), 211–218.

Ali, M.L., McClung, A.M., Jia, M.H., Kimball, J.A., McCouch, S.R., Georgia, C.E., 2011. A Rice diversity panel evaluated for genetic and agro-morphological diversity between subpopulations and its geographic distribution. Crop Sci. 51, 2021–2035.

Aljumaili, S.J., Rafii, M.Y., Latif, M.A., Sakimin, S.Z., Arolu, I.W., Miah, G., 2018. Genetic diversity of aromatic rice germplasm revealed by SSR markers. Hindawi BioMed Res. Int. 218, 1–11.

Amegan, E., Efisue, A., Akoroda, M., Shittu, A., Tonegnikes, F., 2020. Genetic diversity of korean rice (Oryza sativa L.) germplasm for yield and yield related traits for adoption in rice farming system in Nigeria. Int. J. Genet. Genom. 8 (1), 19–28.

Anant, A.K., Guru-Pirasanna-Pandi, G., Jena, M., Chandrakar, G., Parameshwaran, C., Raghu, S., Gowda, B.S., Annamalai, M., Patil, N., Adak, T., Ramasamy, N., Rath, P.C., 2021. Genetic dissection and identification of candidate genes for brown planthopper, Nilaparvata lugens (Delphacidae: Hemiptera) resistance in farmers' varieties of rice in Odisha. Crop Prot. 144, 105600.

Andarini, Y.N., Suwarno, W.B., Aswidinnoor, H., Kurniawan, H., 2022. Genetic relationship of pigmented rice (*Oryza sativa* L.) collected from Eastern Indonesia

based on morpho-agronomical traits and SSR markers. AIP Conf. Proc. 2462 (020023) https://doi.org/10.1063/5.0075706.
Anderson, J.A., Churchill, G.A., Autrique, J.E., Tanksley, S.D., Sorrels, M.E., 1993. Optimizing parental selection for genetic maps. Genome 36, 181–186.
Anilkumar, C., Sah, R.P., Muhammed Azharudheen, T.P., Behera, S., Singh, N., Prakash, N.R., Sunitha, N.C., Devanna, B.N., Marndi, B.C., Patra, B.C., Nair, S.K., 2022. Understanding complex genetic architecture of rice grain weight through QTL-meta analysis and candidate gene identification. Scient. Rep. 12 (1), 1–13.
Anupam, A., Imam, J., Quatadah, S., Siddaiah, A., Das, S.P., Variar, M., Mandal, N.P., 2017. Genetic diversity analysis of Rice germplasm in Tripura state of Northeast India using drought and blast linked markers. Rice Sci. 24 (1), 10–20.
Arunachalam, V.A., Chaudhury, S.S., Sarangi, S.K., Ray, T., Mohanty, B.P., Nambi, V.A., Mishra, S., 2006. In: Rising on Rice: The Story of Jeypore. A manual published by MS Swaminathan Research Foundation, Chennai, Tamil Nadu, India, pp. 1–39.
Ashraf, H., Husaini, A.M., Bhat, M.A., Parray, G.A., Khan, S., Ganai, N.A., 2016. SSR based genetic diversity of pigmented and aromatic rice (Oryza sativa L.) genotypes of the western Himalayan region of India. Physiol. Mol. Biol. Plants 22 (4), 547–555.
Barhate, K.K., Jadhav, M.S., Bhavsar, V.V., 2021. Genetic diversity analysis in aromatic lines of rice (Oryza sativa L.). J. Pharmacogn Phytochem. 10 (3), 367–370.
Barker, R., Dawe, D., 2002. The transformation of the Asian rice economy and directions for future research: the need for increased productivity. In: Sombilla, M., Hossain, M., Hardy, B. (Eds.), Developments in the Asian Rice Economy. International Rice Research Institute, Los BanosPhilippines.
Behera, L., Patra, B.C., Sahu, R.K., Nanda, A., Sahu, S.C., Patnaik, A., Rao, G.J.N., Singh, O.N., 2012. Assessment of genetic diversity in medicinal Rices using microsatellite markers. Austr. J. Crop Sci. 6 (9), 1369–1376.
Behera, L., Mohanty, S., Pradhan, S.K., Singh, S., Singh, O.N., Sahu, R.K., Sahu, S.C., Dash, S.K., Mohapatra, T., 2013. Assessment of genetic diversity of rainfed lowland rice genotypes using microsatellite markers. Ind. J. Genet. 73 (2), 142–152.
Bhuyan, N., Borah, B.K., Sarma, R.N., 2007. Genetic diversity analysis in traditional lowland rice (Oryza sativa L.) of Assam using RAPD and ISSR markers. Curr. Sci. 93, 967–972.
Blair, M.W., Hedetale, V., McCouch, S.R., 2002. Fluorescent-labeled microsatellite panels useful for detecting allelic diversity in cultivated rice (Oryza sativa L.). Theor. Appl. Genet. 105, 449–457.
Bohonak, J.A., 2002. IBD (isolation by distance): a program for analyses of isolation by distance. J. Hered. 93, 153–154.
Borkakati, R.P., Borah, P., Deka, P.C., 2000. Genetic divergence in photoperiod-insensitive autumn rice germplasm of Northeast India. In: Khush, G.S., Brar, D.S., Hardy, B. (Eds.), Advances in Rice Genetics. International Rice Research Institute, Los Banos, Philippines, pp. 74–76.
Caicedo, A.L., Williamson, S.H., Hernandez, R.D., Boyko, A., Feled-Alon, A., York, T.L., Polato, N.R., Olsen, K.M., Nielsen, R., McCouch, S.R., Bustamante, C.D., Purugganan, M.D., 2007. Genome-wide patterns of nucleotide polymorphism in domesticated rice. PLoS Genet. 3 (9), e163.
Callaway, E., 2014. Domestication: the birth of rice. Nature 514, S58–S59.
Causse, M.A., Fulton, T.M., Cho, Y.G., Ahn, S.N., Chunwongse, J., Wu, K., Xiao, J., Yu, Z., Ronald, P.C., Hanington, S.E., Second, G., McCouch, S.R., Steven, D., Tanksley, S.D., 1994. Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138, 1251–1274.
Chakhonkaen, S., Pitnjam, K., Saisuk, W., Ukoskit, K., Muangprom, A., 2012. Genetic structure of thai rice and rice accessions obtained from the international rice research institute. Rice 5, 19.
Chambers, M., Avoy, M., 2000. Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. Mol. Ecol. 9, 1037–1048.
Chandra, H., Yadava, Y.K., Bishnoi, P., Kumar, P., 2017. Genetic diversity analysis using microsatellite based simple sequence repeats (SSR) markers in Basmati Indica Rice (Oryza sativa L.). Vegetos 30 (Special), 2017.
Chang, H., Yu, B., Yu, X.P., Yi, L., Chen, C.Y., Mi, M.T., Hui, C., Bin, Y., Xiaoping, Y., Long, Y., Chunye, C., Mantian, M., Wenhua, L., 2010. Anticancer activities of an anthocyanin-rich extract from black rice against breast cancer cells in vitro and in vivo. Nutr. Cancer 62 (8), 1128–1136.
Chen, X.S., Temnykh, Y.X., Cho, Y.G., McCouch, S.R., 1997. Development of a microsatellite framework map providing genome wide coverage in rice (Oryza sativa L.). Theor. Appl. Genet. 95, 55–567.
Chen, L.J., Lee, D.S., Song, Z.P., Suh, H.S., Lu, R.B., 2004. Gene flow from cultivated rice (Oryza sativa) to its weedy and wild relatives. Ann. Bot. 93, 67–73.
Cho, Y.G., Ishii, T., Temnykh, S., Chen, X., Lipovich, L., McCouch, S.R., 2000. Diversity of microsatellites derived from genomic libraries and Gene Bank sequences in rice (Oryza sativa L.). Theor. Appl. Genet. 100, 713–722.
Choudhary, G., Ranjitkumar, N., Surapaneni, M., Deborah, D.A., Kumar, V.A., Ghanta, A., Siddiq, E.A., Vemireddy, L., 2013. Molecular genetic diversity of major indian rice cultivars over decadal periods. PLoS ONE 8 (6), e66197.
Choudhury, B., Khan, M.L., Dayanandan, S., 2013. Genetic structure and diversity of indigenous rice (Oryza sativa) varieties in the eastern himalayan region of Northeast India. Springerplus 2, 228.
Choudhury, D.R., Singh, N., Singh, A.K., Kumar, S., Srinivasan, K., Tyagi, R.K., Ahmad, A., Singh, N.K., Singh, R.A., 2014. Analysis of genetic diversity and population structure of rice germplasm from north-eastern region of India and development of a core germplasm set. PLoS ONE 9 (11), e113094.
Courtois, B., Frouin, J., Greco, R.A., Bruschi, G., Droc, G., 2012. Genetic diversity and population structure in a european collection of Rice. Crop Sci. 52, 1663–1675.
Das, S.R., 2012. Rice in Odisha. In: IRRI Technical Bulletin No. 16. International Rice Research Institute, Los Baños, Philippines, pp. 1–31.
Das, T., Das, A.K., 2014. Inventory of the traditional rice varieties in farming system of southern Assam: a case study. Indian J. Tradit. Knowl. 13 (1), 157–163.
Das, B., Sengupta, S., Parida, S.K., Roy, B., Ghosh, M., Prasad, M., Ghose, T.K., 2013. Genetic diversity and population structure of rice landraces from eastern and north eastern states of India. BMC Genet. 14, 71.
Davierwala, A.P., Chowdari, K.V., Kumar, S., Reddy, A.P.K., Ranjekar, P.K., Gupta, V.S., 2000. Use of three different marker systems to estimate genetic diversity of indian elite rice varieties. Genet. 108, 269–284.
Deb, D., 2005. In: Seedsof Tradition, Seeds of Future, Folk Rice Varieties of Eastern India. Research Foundation for Science Technology & Ecology, A-60 Hauz Khas, New Delhi 11001, pp. 1–17.
Dela Cruz, N., Khush, G.S., 2000. In: Singh, R.K., Singh, U.S., Khush, G.S. (Eds.), Aromatic Rices. Oxford & IBH Publishing Co., New Delhi, pp. 15–28.
Dellaporta, S.L., Wood, J., Hicks, J.B., 1983. A plant DNA mini-preparation version II. Plant Mol. Biol. Rep. 1, 19–21.
Dwivedi, A., Rathour, R., Basandrai, D., Sarial, A.K., 2019. Molecular genetic diversity analysis using SSR markers of basmati rice (Oryza sativa L.) genotypes of northern hill region. Ind. J. Cereal Res. 11 (3), 224–230.
Earl, D.A., von Holdt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4 (2), 359–361.
Embate, M.V.G., Calayugan, M.I.C., Gentallan Jr., R.P., Cruz, P.C.T., Hernandez, J.E., Borromeo, T.H., 2021. Genetic diversity of selected pigmented traditional rice (Oryza sativa L.) varieties from Mindanao, Philippines using agromorphological traits and simple sequence repeats markers. J. Crop. Sci. Biotechnol. 24, 259–277.
Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620.
FAO (Food and Agriculture Organization), 2004. Rice Is Life. FAO, Italy.
Frei, M., Becker, K., 2004. Agro-biodiversity in subsistence-oriented farming systems in a Philippine upland region: nutritional considerations. Biodivers. Conserv. 13 (8), 1591–1610.
Gangopadhyay, S., 1991. Orissa. In: Ghosh, S.P. (Ed.), Agro-climaticZone-specific Research, Indian Perspective Under NARP. ICAR, New Delhi, pp. 280–313.
García Montecinos, K.L., Godoy Godoy, J.A., Carrillo Centeno, P.M., Pachón, H., 2011. Sensory evaluation of the Azucena rice (Oryza sativa) variety in Nicaragua's Región Autónoma del Atlántico Norte. Perspect. Nutr. Hum. 13 (2), 135–146.
Garland, S.H., Lewin, L., Abedinia, M., Henry, R., Blackeney, A., 1999. The use of microsatellite polymorphisms for the identification of Australian breeding lines of rice (Oryza saitva L.). Euphytica 108, 53–63.
Garris, A.J., Tai, T.H., Coburn, J., Kresovich, S., McCouch, S.R., 2005. Genetic structure and diversity in Oryza sativa L. Genet. 169, 1631–1638.
Giarrocco, L.E., Marassi, M.A., Salerno, G.L., 2003. Assessment of the genetic diversity in Argentine rice cultivars with SSR markers. Crop Sci. 47, 853–858.
Glaszmann, J.C., Benyayer, P., Arnaud, M., 1989. Genetic divergence among rices from Northeast India. In: Rice Genetics Newsletter. Genetic Resources Section, National Institute of Genetics, Mishima, Japan, p. 63.
Goudet, J., 2005. FSTAT(V2.9.4)[Online]. Website last modified on August 23, 2005 (accessed on August 18, 2020) Available at. http://www2.unil.ch/popgen/softwares/fstat.htm.
Gour, L., Koutu, G.K., Singh, S.K., Mishra, D.K., 2017. Genetic potency identification of indigenous rice (Oryza sativa L.) lines using quality assessment and SSR marker analysis. Int. J. Chem. Stud. 5 (4), 97–101.
Gross, B.L., Reagon, M., Hsu, S.C., Caicedo, A.L., Jia, Y., Olsen, K.M., 2010. Seeing red: the origin of grain pigmentation in US weedy rice. Mol. Ecol. 19, 3380–3393.
Grover, A., Sharma, P.C., 2011. Is spatial occurrence of microsatellites in the genome a determinant of their function and dynamics contributing to genome evolution? Curr. Sci. 100 (6), 859–869.
Hamrick, J.L., Godt, M.J., 1996. Conservation genetics of endangered plant species. In: Avise, J.C., Hamrick, J.L. (Eds.), Conservation Genetics: Case Histories from Nature. Chapman & Hall, New York, pp. 281–304.
Hanamaratti, N.G., Prashanthi, S.K., Salimath, P.M., Hanchinal, R.R., Mohankumar, H. D., Parameshwarappa, K.G., Raikar, S.D., 2008. Traditional landraces of rice in Karnataka: reservoirs of valuable traits. Curr. Sci. 94, 242–247.
Hien, N.L., Sarhadi, W.A., Oikawa, Y., Hirata, Y., 2007. Genetic diversity of morphological responses and the relationships among Asia aromatic rice (Oryza sativa L.) cultivars. Tropics 16 (4), 343–355.
International Rice Genome Sequencing Project, 2005. The map-based of the rice genome. Nature 436, 793–800.
Islam, M.Z., Khalequzzaman, M., Bashar, M.K., Ivy, N.A., Mian, M.A.K., Pittendrigh, B.R., Haque, M.M., Ali, M.P., 2018. Variability assessment of aromatic rice germplasm by pheno-genomic traits and population structure analysis. Sci. Rep. 8, 9911.
Jain, S., Jain, R.K., McCouch, S.R., 2004. Genetic analysis of Indian aromatic and quality rice (Oryza sativa) germplasm using panels of fluorescently-labeled microsatellite markers. Theor. Appl. Genet. 109, 965–977.
Jain, N., Jain, S., Saini, N., Jain, R.K., 2006. SSR analysis of chromosome 8 regions associated with aroma and cooked kernel elongation in Basmati rice. Euphytica 152, 259–273.
Jasim, A.S., Rafi, M.Y., Latif, M.A., Sakimin, S.Z., Arolu, I.W., Miah, G., 2018. Genetic diversity of aromatic rice germplasm revealed by SSR markers. Hindawi Biomed. Res. Int. 2018, 1–11.
Jayamani, P., Negrao, S., Martins, M., Macas, B., Oliveira, M.M., 2007. Genetic relatedness of Portuguese rice accessions from diverse origin as assessed by microsatellite markers. Crop Sci. 47, 879–886.
Joshi, R.K., Subudhi, E., Kar, B., Nayak, S., 2010. Comparative genetic analysis of lowland rice cultivars of India using microsatellite markers. Biores. Bull. 4, 213–223.
Juneja, S., Das, A., Joshi, S.V., Sharma, S., Vikal, Y., Patra, B.C., Bharaj, T.S., Sidhu, J.S., Singh, K., 2006. Oryza nivara Sharma et Shastry, the progenitor of O. sativa L.

subspecies indica harbours rich genetic diversity as measured by SSR markers. Curr. Sci. 91, 1079–1085.
Kaushik, A., Jain, S., McCouch, S.R., Jain, R.K., 2011. Phylogenetic relationships among various groups of rice (Oryza sativa L.) as revealed by microsatellite and transposable element-based marker analysis. Ind. J. Genet. Plant Breed. 71, 139–150.
Kim, H.J., Jeong, E.G., Ahn, S.N., Doyle, J., Singh, N., Greenberg, A.J., Won, Y.J., McCouch, S.R., 2014. Nuclear and chloroplast diversity and phenotypic distribution of rice (Oryza sativa L.) germplasm from the democratic people's republic of Korea (DPRK; North Korea). Rice 7, 7.
Kumar, S., Bisht, I.S., Bhat, K.V., 2010. Assessment of genetic diversity among rice (Oryza sativa L.) landrace populations under traditional production using microsatellite (SSR) markers Electron. J. Plant Breed. 1 (4), 474–483.
Kumar, V., Rastogi, N.K., Saxsena, R.R., Mandal, I.S.N., 2018. Molecular characterization and genetic diversity analysis of aromatic rice (Oryza sativa L.) genotypes using SSR markers. J. Crop Weed 14 (3), 94–99.
Kurata, N., Nagamura, Y., Yamamoto, K., Harushima, Y., Sue, N., Wu, J., et al., 1994. A 300 kilobase interval genetic map of rice including 883 expressed sequences. Nat. Genet. 8, 365–372.
Lai, J.H., 1995. Rare black glutinous rice germplasm in eastern Guizhou province. Crop Genet. Resour. 2, 53–54.
Lin, H.Y., Wu, Y.P., Hour, A.L., Ho, S.W., Wei, F.J., Hsing, Y.C., Lin, Y.R., 2012. Genetic diversity of rice germplasm used in Taiwan breeding programs. Bot. Stud. 53, 363–376.
Lisa, L.A., Elias, S.M., Rahman, M.S., Shahid, S., Iwasaki, T., Hasan, A.K.M.M., Kosuge, K., Fukami, Y., Seraj, Z.I., 2011. Physiology and gene expression of the rice landrace under salt stress. Funct. Plant Biol. 38 (4), 282–292.
Lu, H., Redus, M.A., Coburn, J.R., Rutger, J.N., McCouch, S.R., Tai, T.H., 2005. Population structure and breeding patterns of 145 US rice cultivars based on SSR marker analysis. Crop Sci. 45, 66–76.
Manjunath, G.A., Elsy, C.R., Joseph, J., Rose Mary Francies, R.M., 2021. Molecular characterization and genetic diversity analysis of aromatic rice (Oryza sativa L.) landraces using SSR markers. Electronic J. Plant Breed. 12 (1), 576–582.
McCouch, S.R., Teytelman, L., Xu, Y., Lobos, K.B., Clare, K., Walton, M., Fu, B., Maghirang, R., Li, Z., Xing, Y., Zhang, Q., Kono, I., Yano, M., Fjellstrom, R., DeClerck, G., Schneider, D., Cartinhour, S., Ware, D., 2002. Development and Mapping of 2240 New SSR Markers for Rice (Oryza sativa L.). DNA Res. 9, 199–207.
Mishra, A., Kumar, P., Shamim, M.D., Tiwari, K.K., Fatima, P., Srivastava, D., Singh, R., Yadav, P., 2019. Genetic diversity and population structure analysis of Asian and African aromatic rice (Oryza sativa L.) genotypes. J. Genet. 98, 92.
Mohanty, S., Donde, R., Das, S., Panda, D., Mishra, B., Pradhan, S.K., Dash, S.K., Swain, P., Behera, L., 2021. Utilization of genetic diversity and population structure to reveal prospective drought-tolerant donors in rice. Gene Rep. 23, 101151.
Naaz, S., Pandey, V., Yadav, H.K., 2022. Evaluation of genetic diversity in rice (Oryza sativa L. ssp. Indica) accessions using SSR marker. Vegetos. https://doi.org/10.1007/s42535-022-00378-3, 12 Sep 2022 12:38:53 Z.
Nagarajan, S., 2007. Geographical indications and agriculture related intellectual property rights issues. Curr. Sci. 92, 167–171.
Nagarajan, S., Trivedi, R.K., Raj Ganesh, D.S., Singh, A.K., 2010. India registers plant varieties under PPV&FR Act, 2001. Curr. Sci. 99 (6), 723–725.
Nagaraju, J., Kathirvel, M., Ramesh Kumar, R., Siddiq, E.A., Hasnain, S.E., 2002. Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR-PCR and SSR markers. Proc. Nat. Acad. SciUSA 99, 5836–5841.
Nayak, S., Reddy, A.M., Ponmurugan, P., Devi, A., Ayyappadasan, G., Upreti, D.K., 2013. Eastern Ghats' biodiversity reserves with unexplored lichen wealth. Curr. Sci. 104 (7), 821–825.
Nayak, A.K., Anilkumar, C., Behera, S., Sah, R.P., Lavanya, G.R., Kumar, A., Behera, L., Tp, M.A., 2022. Genetic dissection of grain size traits through genome-wide association study based on genic markers in rice. Rice Sci. 29 (5), 462–472.
Nei, M., 1973. Analysis of gene diversity in subdivided populations. Proc. Nat. Acad. SciUSA 70, 3321–3323.
Ni, J., Colowit, P.M., Mackill, D.J., 2002. Evaluation of genetic diversity in rice subspecies using microsatellite markers. Crop Sci. 42, 601–607.
Nilthong, S., Chukeatirote, E., Nilthong, R., 2020. Assessment of genetic diversity in Thai upland rice varieties using SSR markers. Austr. J. Crop Sci. 14 (04), 597–604.
Olufowote, J.O., Xu, Y., Chen, X., Park, W.D., Beachell, H.M., Dilday, R.H., Goto, M., McCouch, S.R., 1997. Comparative evaluation of within-cultivar variation of rice (Oryza sativa L.) using microsatellite and RFLP markers. Genome 40, 370–378.
Pachauri, V., Singh, M.K., Singh, A.K., Singh, S., Shakeel, N.A., Singh, V.P., Singh, N.K., 2010. Origin and Genetic Diversity of Aromatic Rice Varieties, Molecular Breeding and Chemical and Genetic Basis of Rice Aroma. J. Plant Biochem. Biotechnol. 19 (2), 127–143.
Padhi, S., Panigrahi, N., 2011. Brief 10: Tribal Movements and Livelihoods: Recent Developments in Orissa-A Working Paper by Indian Institute of Public Administration (IIPA) and Chronic Poverty Research Centre (CPRC), New Delhi.
Panda, T., Mishra, N., Mohanty, R.B., 2013. Diversity of Some Threatened Indigenous Rice Varieties Cultivated in OdishaIndia. Environ. Nat. Resour. J. 11 (2), 41–57.
Parida, S.K., Dalal, V., Singh, A.K., Singh, N.K., Mohapatra, T., 2009. Genic non-coding microsatellites in the rice genome: characterization, marker design and use in assessing genetic and evolutionary relationships among domesticated groups. BMC Genom. 10, 140.
Park, J.R., Yang, W.T., Kwon, Y.S., Kim, H.N., Kim, K.M., Doh-Hoon Kim, D.H., 2019. Assessment of the genetic diversity of rice germplasms characterized by black-purple and red pericarp color using simple sequence repeat markers. Plants 8, 471.
Parzies, H.K., Spoor, W., Ennos, R.A., 2004. Inferring seed exchange between farmers from population genetic structure of barley landrace Arabi Aswad from Northern Syria. Genet. Resour. Crop. Evol. 51, 471–478.
Pathaichidachote, W., Panyawut, N., Sikaewtung, K., Patarapuwadoli, S., Muangprom, A., 2019. Genetic diversity and allelic frequency of selected Thai and exotic rice germplasm using SSR markers. Rice Sci. 26 (6), 393–403.
Peakall, R., Smouse, P.E., 2006. Genalex 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295.
Perrier, X., Jacquemoud-Collet, J.P., 2006. DARwin Software. http://darwin.cirad.fr/darwin.
Poudel, D., Sthapit, B., Shrestha, P., 2015. An analysis of social seed network and its contribution to on-farm conservation of crop genetic diversity in Nepal. Int. J. Biodivers. 2015, 1–13.
Prasad, G.S.V., Padmavathi, G., Suneetha, K., Madhav, M.S., Muralidharan, K., 2020. Assessment of diversity of Indian aromatic rice germplasm collections for morphological, agronomical, quality traits and molecular characters to identify a core set for crop improvement. CABI Agric. Biosci. 1, 13.
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genet. 155, 945–959.
Pusadee, T., Jamjod, S., Chiang, Y.C., Rerkasem, B., Schaal, B.A., 2009. Genetic structure and isolation by distance in a landrace of Thai rice. Proc. Natl. Acad. Sci. U. S. A. 106 (33), 13880–13885.
Ram, S.G., Thiruvengadam, V., Vinod, K.K., 2007. Genetic diversity among cultivars, landraces and wild relatives of rice as revealed by microsatellite markers. J. Appl. Genet. 48, 337–345.
Ramasamy, R.K., Ramasamy, S., Bindroo, B.B., Naik, V.G., 2014. STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. Springerplus 3 (1), 1–3.
Rathi, S., Yadav, R.N.S., Pathak, K., Sarma, R.N., 2014. Genetic diversity in upland rice of Assam assessed through SSR markers. Indian J. Genet. 74 (2), 238–242.
Rauf, S., da Silva, J.T., Khan, A.A., Naveed, A., 2010. Consequences of plant breeding on genetic diversity. Int. J. Plant Breed. 41, 1–21.
Ray, A., Deb, D., Ray, R., Chattopadhayay, B., 2013. Phenotypic characters of rice landraces reveal independent lineages of short-grain aromatic indica rice. AoB PLANTS 5, plt032. https://doi.org/10.1093/aobpla/plt032.
Rijal, D.K., 2010. Role of food tradition in conserving crop landraces on-farm. J. Agric. Environ. 11, 107–119.
Rolf, F.J., 1998. NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, Version 2.02. Exeter Software Publishing, Setauket, New York.
Roy, S., Banerjee, A., Mawkhlieng, B., Misra, A.K., Pattanayak, A., Harish, G.D., Singh, S. K., Ngachan, S.V., Bansal, K.C., 2015. Genetic diversity and population structure in aromatic and quality rice (Oryza sativa L.) land races from North-Eastern India. PLoS ONE 10 (6), e0129607.
Sah, R.P., Nayak, A.K., Chandrappa, A., Behera, S., Azharudheen, T.P.M., Lavanya, G.R., 2022. cgSSR marker-based genome-wide association study identified genomic regions for panicle characters and yield in rice (Oryza sativa L.). Website last modified on August 23, 2005 (accessed on August 18, 2020) Available at J. Sci. Food Agril.. https://doi.org/10.1002/jsfa.12183.
Saini, N., Jain, N., Jain, S., Jain, R.K., 2004. Assessment of genetic diversity within and among Basmati and non-Basmati rice varieties using AFLP, ISSR and SSR markers. Euphytica 140, 133–146.
Salgotra, R.K., Gupta, B.B., Bhat, J.A., Sharma, S., 2015. Genetic diversity and population structure of basmati rice (Oryza sativa L.) germplasm collected from north western Himalayas using trait linked SSR markers. PLoS ONE 10 (7), e0131858.
Sang, T., Ge, S., 2013. Understanding rice domestication and implications for cultivar improvement. Curr. Opin. Plant Biol. 16, 139–146.
Sarma, R.N., Bahar, B., 2005. Genetic variation of bora rice (glutinous rice) of Assam as revealed by RAPDs. Plant Genet. Resour. Newsl. 144, 34–38.
Sarma, B.K., Pattanayak, A., 2009. Rice Diversity of Northeast India. Millenium Publishers, Guwahati, India.
Shankar, A.V., Ravi, S.D., Rukh, A., 2015. Recognition and separation of home-grown nonBasmati aromatic rice genotypes of India using microsatellite markers. Int. J. Agric. Sci. 5 (5), 758–763.
Sharma, S.D., Tripathy, S., Biswal, J., 2000. Origin of O. sativa and its ecotypes. In: Nanda, J.S. (Ed.), Rice Breeding and Genetics: Research Priorities and Challenges. Science Publishers, Enfield and Oxford and IBH, New Delhi, pp. 349–369.
Singh, R.K., Singh, U.S., 2003. Scented rices of India: an introduction. In: A Treatiseon the Scented Rices of India, 1st edition, pp. 1–4.
Singh, N., Majumder, S., Singh, S., Vikram, S., 2016a. Genetic diversity in Indian rice germplasm set using phenotypic and genotypic variables simultaneously. Indian J. Genet. 76 (3), 246–254.
Singh, N., Roy Choudhury, D., Tiwari, G., Singh, A.K., Kumar, S., Srinivasan, K., Tyagi, R. K., Sharma, A.D., Singh, N.K., Singh, R., 2016b. Genetic diversity trend in Indian rice varieties: an analysis using SSR markers. BMC Genet. 17, 127.
Sirabanchongkran, A., Yimyam, N., Boonma, W., Rerkasem, K., Coffey, K., Pinedo-Vasquez, M., Padoch, C., 2004. Varietal turnover and seed exchange: Implications for conservation of rice genetic diversity on-farm. Int. Rice Res. Note. 29 (2), 18–19.
Taguchi-Shiobara, F., Ozaki, H., Sato, H., Maeda, H., Kojima, Y., Ebitani, T., Yano, M., 2013. Mapping and validation of QTLs for rice sheath blight resistance. Breed. Sci. 63 (3), 301–308.
Thomson, M.J., Polato, N.R., Prasetiyono, J., Trijatmiko, K.R., Silitonga, T.S., McCouch, S.R., 2009. Genetic diversity of isolated populations of Indonesian landraces of rice (Oryza sativa L.), collected in East Kalimantan on the Island of Borneo. Rice 2, 80–92.
Tiwari, K.K., Singh, A., Pattnaik, S., Sandhu, M., Kaur, S., Jain, S., Tiwari, S., Mehrotra, S., Anumalla, M., Samal, R., et al., 2015. Identification of a diverse mini-

core panel of Indian rice germplasm based on genotyping using microsatellite markers. Plant Breed 134 (2), 164–171.
Vairavan, S., Siddiq, E.A., Arunachalam, V., Swaminathan, M.S., 1973. A study on the nature of genetic divergence in rice from Assam and Northeast Himalayas. Theor. Appl. Genet. 43, 213–221.
Vanlalsanga, Singh, S.P., Sing, Y.T., 2019. Rice of Northeast India harbor rich genetic diversity as measured by SSR markers and Zn/Fe content. BMC Genet. 20, 79.
Wong, S.C., Yiu, P.H., Bong, S.T.W., Lee, H.H., Neoh, P.N.P., Rajan, A., 2009. Analysis of Sarawak Bario Rice Diversity Using Microsatellite Markers. Am. J. Agric. Biol. Sci. 4, 298–304.
Yogi, R., Kumar, N., Kumar, R., Jain, R.K., 2020. Genetic diversity analysis among important rice (Oryza sativa L.) genotypes using SSR markers. Adv. Biores. 11 (2), 68–74.
Yonemaru, J., Yamamoto, T., Ebana, K., Yamamoto, E., Nagasaki, H., Shibaya, T., Yano, M., 2012. Genome-wide haplotype changes produced by artificial selection during modern rice breeding in Japan. PLoS ONE 7, e32982.
Yu, S.B., Xu, W.J., Vijayakumar, C.H.M., Ali, J., Fu, B.Y., Xu, J.L., Jian, Y.Z., Marghirang, R., Domingo, J., Aquino, C., Virmani, S.S., Li, Z.K., 2003. Molecular diversity and multi locus organization of the parental lines used in the International Rice Molecular Breeding Program. Theor. Appl. Genet. 108, 131–140.
Zeven, A.C., 1998. Landraces: a review of definitions and classifications. Euphytica 104, 127–139.
Zhang, P., Li, J., Li, X., Liu, X., Zhao, X., 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.
Zhang, Li-na, Cao, Gui-lan, Han, Long-zhi, 2013. Genetic diversity of rice landraces from lowland and upland accessions of China. Rice Sci. 20 (4), 259–266.
Zhu, Q.H., Zheng, X.M., Luo, J.C., Gaut, B.S., Ge, S., 2007. Multilocus analysis of nucleotide variation of Oryza sativa and its wild relatives: severe bottleneck during domestication of rice. Mol. Biol. Evol. 24, 875–888.

Further reading

IRRI-IBPGR, 1980. Descriptors for Rice (Oryza sativa L.). IBPGR-IRRI Rice Advisory Committee, IRRI, P.O. Box 933, Manila, Philippines.
Subedi, S., Rosyara, U.R., Adhikari, B.B., Ojha, B.R., Ghimire, D.P., Dhakal, D.D., Gurung, H.B., Pandey, S., 2011. Participatory crop improvement: effect of farmers' selection criteria on diversity of rice grown in uplands of Nepal. J. Agric. Res. 49 (1), 83–96.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Bio. Evol. 30 (12), 2725–2729.