Indian J. Genet. Plant Breed., (2025); 85(3): 376-388
https://doi.org/10.31742/ISGPB.85.3.3
ISSN: 0975-6906
www.isgpb.org

RESEARCH ARTICLE

Evaluation of the genetic architecture utilizing simple sequence repeat (SSR) markers in deep water rice (*Oryza sativa* L.) landraces of Assam, India

D. Shephrou Helena, Megha Rohilla[1], Tapan Kumar Mondal[1], Priyanka Roy, Sushil Kumar, Dhiren Chowdhury[2] and Priyabrata Sen*

Abstract

Deep water rice (DWR) is an essential agricultural practice in flood-prone regions, supporting millions of farmers in Asia and Africa. However, its cultivation is under threat from changing climates, modern agricultural practices and socio-economic shifts. Although the state has quite a large collection of deep water rice (*Oryza sativa* L.), there is less exploration on the nature and extent of genetic diversity. Therefore, the present study was conducted to investigate the genetic diversity in a set of 92 deep water rice landraces by evaluating genetic polymorphism using 56 polymorphic SSR markers. A total of 139 alleles were detected, showing high polymorphism among all these diverse landraces. The major allele frequency of SSR loci comes in the range of 0.299 to 0.88. Expected heterozygosity varied from 0.21 and 0.74, whereas the observed heterozygosity ranged from 0.00 to 0.73. The PIC value ranged from 0.18 to 0.69 and the RM 206 marker was found to be most appropriate to discriminate among these landraces, owing to the highest polymorphic information content value of 0.69. AMOVA revealed that the principal molecular variance existed within populations (96%) rather than among populations (4%). The phylogenetic analysis clustered these accessions into 7 clusters, in which cluster II had a maximum of 27 genotypes, followed by cluster III and cluster I. Similarly, structure analysis based on Bayesian clustering grouped these diverse accessions into 7 sub-populations and also observed admixture in the accessions. The information accrued from the current study offers valuable insights for effective use in improving DWR varieties.

Keywords: SSR, PIC, Deep water rice, Flood, Genetic diversity, Polymorphism, SSR markers.

Introduction

Rice (*Oryza sativa* L.), being one of the most commonly cultivated crops in different regions and climates of the world, is known to have enormous genetic diversity. *Oryza sativa* (Asian cultivated rice) is known to be diversified into five subpopulations on basis of genetic markers: *indica*, *aus*, *aromatic*, *temperate japonica*, and *tropical japonica*. In addition, *indica* and *japonica* comprise *temperate japonica* and *tropical japonica*, *aus* is known to be close to *indica*, and *aromatic* is intermediate between *indica* and *japonica* (Garris et al. 2005). India is one of the centres for rice diversity, where rice accessions have a great reservoir of useful genes that rice breeders can use for rice crop improvement programs.

Deep water rice (DWR) is an essential agricultural practice in flood-prone regions, supporting millions of farmers in Asia and Africa. However, its cultivation is under threat from changing climates, modern agricultural practices and socio-economic shifts. It is grown in flooded conditions with water depths of more than 50 cm (20 inches) for at least a month (Matin et al. 2012). Every year, during the

Department of Agricultural Biotechnology, Assam Agricultural University (AAU), Jorhat 785 013, Assam, India.

[1]National Institute for Plant Biotechnology, Pusa Campus, New Delhi 110 012, India.

[2]Regional Agricultural Research Station, AAU, North Lakhimpur 787 001, Assam, India.

***Corresponding Author:** Priyabrata Sen, Department of Agricultural Biotechnology, Assam Agricultural University (AAU), Jorhat 785 013, Assam, India, E-Mail: pbsen14@gmail.com

How to cite this article: Helena D.S., Rohilla M., Mondal T.K., Roy P., Kumar S., Chowdhury D. and Sen P. 2025. Evaluation of the genetic architecture utilizing simple sequence repeat (SSR) markers in deep water rice (*Oryza sativa* L.) landraces of Assam, India. Indian J. Genet. Plant Breed., **85**(3): 376-388.

Source of support: Ministry of Tribal Affairs, Govt. of India (Ref. No. 201819-NFST-MAN-02172)

Conflict of interest: None.

Received: Nov. 2024 **Revised:** June 2025 **Accepted:** July 2025

Fig. 1. Flood-prone areas of Assam showing different districts from which deep water rice landraces were collected (indicated by circles)

rice cultivation season, Assam witnesses severe flooding in many districts, leading to a drastic decline in the seasonal rice yield. The reduction in the production of rice is a matter of concern as rice is the staple food of the region. Although the state has quite a large collection of deep water rice (*Oryza sativa* L.), there is less exploration on the nature and extent of genetic diversity. Continued efforts in research, adaptation, and preservation of traditional practices are vital for their sustainability. Therefore, there is an emerging need to counter the unpredictable scenario by selection potential rice cultivars for a breeding program to develop flood-tolerant rice varieties.

Molecular markers have been widely used to evaluate allelic variation to explain the genetic relationships within and among species. Evaluation of genetic diversity using DNA markers technology is non-destructive, requires a small amount of samples, is not affected by environmental factors, and does not require a large experimental setup and equipment for analysis (Kanawapee et al. 2011). New opportunities have been created with the development of DNA markers, of which simple sequence repeat markers (SSR) are the most commonly employed tool, having the ability to give a better genetic diversity spectrum due to their multi-allelic and highly polymorphic nature (Singh et al. 2016). As compared to other markers, PCR-based microsatellites are more cost-effective, efficient, abundant, co-dominant, highly reproducible and extend throughout the genome (Panaud et al. 1996). SSR markers are regions of DNA having repeated sequences of one to five nucleotides, often forming part of the non-coding regions widely dispersed throughout the rice genome (Temnykh et al. 2001). Studies on SSRs for the determination of genetic diversity in both wild and cultivated varieties of rice have been reported (Siwach et al. 2004). Despite having a high level of polymorphism, SSR can provide information on genetic diversity even among genotypes having a smaller number of markers (McCouch et al. 1997). The present investigation was undertaken with the aim of assessing the trend in genetic diversity in 92 landraces of deep water rice collected from various flood-prone districts of Assam using 74 SSR markers.

Materials and methods

Plant materials

Ninety two deep water rice (DWR) landraces, along with their respective local names, were collected from various flood-prone regions of Assam (Fig. 1) and preserved at the Regional Agricultural Research Station (RARS) in North Lakhimpur. These collected genetic resources (Table 1) were utilized as analytical material for SSR-assisted marker analysis. The experiments were conducted at the Regional Agricultural Research Station, AAU, North Lakhimpur, Assam, India, during *sali* season for 2 years. All the deep water rice landraces of Assam were evaluated at the molecular level using 74 SSR markers.

DNA extraction

Genomic DNA isolation was done from leaf tissue of 20 to 25-day-old seedlings following the method described by Dellaporta et al. (1983). The isolated DNA was quantified with the help of NanoDrop (Thermo Scientific) at 260 and 280 nm absorbance. The working sample of genomic DNA was diluted with sterile distilled water to a concentration of 20 ng μL^{-1}. The Polymerase Chain Reaction (PCR) amplification was performed in a reaction volume of 10 µL with a mixture composition of 2 µL of Template DNA, 0.1 µL of 2.5 mM dNTPs, 1-µL (for both forward and reverse primers) and 0.3 µL of 3U/µL Taq DNA polymerase.

Genotyping using SSR markers

In order to select effective polymorphic SSR markers for these selected 92 DWR landraces, a total of 74 SSR markers of dinucleotide repeats were utilized for the genotyping

Table 1. A list of deep water rice landraces used in the present study

S. No.	Genotype name	Code	IC numbers	Location	District	Latitude	Longitude
1	Rongadhar Kekuwabao	V1	NA	Ghilamara	Lakhimpur	27.31°N	94.41°E
2	Nasatibao	V2	NA	Katohaguri	Lakhimpur	27.94814°N	80.77777°E
3	Rangi-1	V3	IC352815	Potiamohaian	Lakhimpur	27.94814°N	80.77777°E
4	Saikiabao	V4	IC632952	North lakhmipur	Lakhimpur	27.22849°N	94.09444°E
5	Saikiabao 1	V5	NA	North lakhimpur	Lakhimpur	27.22849°N	94.09444°E
6	Amona 1	V6	IC591483	Khoga, North lakhimpur	Lakhimpur	27.143455°N	94.167678°E
7	Borjul	V7	IC450343	Pota mohaijan	Lakhimpur	27.94814°N	80.77777°E
8	Betu-1	V8	IC591481	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
9	Boga Amona-1	V9	NA	Kolabari,Gohpur	BiswanathChariali	26.88462°N	93.62262°E
10	Sunmoti 1	V10	IC591469	Gohpur	BiswanathChariali	26.88462°N	93.62262°E
11	Betu-2	V11	NA	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
12	Kekuwa-1	V12	NA	Ghilamora	Lakhimpur	27.31°N	94.41°E
13	Sonamukhi bao-2	V13	IC422972	Bogodebeli, Dhemaji	Dhemaji	27.47525°N	94.56184°E
14	Sworgfollabao	V14	IC6321947	Karson	Lakhimpur	27.94814°N	80.77777°E
15	Rongabao	V15	IC394600	Kotohaguri	Lakhimpur	27.94814°N	80.77777°E
16	Dhusuribao	V16	IC591511	Gohpur	BiswanathChariali	26.88462°N	93.62262°E
17	Ikorabao	V17	IC423043	Gorehoga	Lakhimpur	27.94814°N	80.77777°E
18	Joleebao 1	V18	IC632948	Gorehoga	Lakhimpur	27.94814°N	80.77777°E
19	Bokajohingia bao-2	V19	IC591487	Dhemaji	Dhemaji	27.47525°N	94.56184°E
20	Betnasalibao	V20	IC632949	Morigaon	Morigaon	26.25213°N	92.34157°E
21	Bogadhepabao	V21	NA	Kolabari,Gohpur	BiswanathChariali	26.88462°N	93.62262°E
22	Sanjulbao	V22	IC591462	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
23	Thiagothabao	V23	IC394861	Borpeta	Borpeta	26.3304209°N	91.0040547°E
24	Dolmora	V24	IC575044	Potiamohaijan	Lakhimpur	27.94814°N	80.77777°E
25	Bedelbao	V25	IC632950	Katohaguri	Lakhimpur	27.94814°N	80.77777°E
26	Dal bao	V26	IC423046	Ghilamora	Lakhimpur	27.31°N	94.41°E
27	Sonamukhi bao-1	V27	IC591490	Bogodebeli, Dhemaji	Dhemaji	27.47525°N	94.56184°E
28	Salkhosera	V28	IC464944	Chenga,Borpeta	Borpeta	26.2790111°N	91.1565196°E
29	Jengbao	V29	IC591494	Ghilamora	Lakhimpur	27.31°N	94.41°E
30	Khutijulbao	V30	IC591494	Potiamohaian	Lakhimpur	27.94814°N	80.77777°E
31	UjanniKhamtibao	V31	IC591506	Majuli	Majuli	26.96274°N	94.16202°E
32	Jeng bao-1	V32	IC575062	Ghilamora	Lakhimpur	27.31°N	94.41°E
33	Garo bao	V33	IC632951	Loharghatkamrup	Kamrup	25.58′52N	91.28′38E
34	Baolabao	V34	NA	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
35	HBJ Amona	V35	NA	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
36	Bogagothabao	V36	IC591465	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
37	Burulibao	V37	IC575065	Korsun	Lakhimpur	27.94814°N	80.77777°E
38	Bangla bao	V38	IC575264	Borpeta	Borpeta	26.3304209°N	91.0040547°E

39	Bezel bao	V39	IC632946	Kathohaguri	Lakhimpur	27.94814°N	80.77777°E
40	Khalburunibao	V40	IC591471	Borpeta	Borpeta	26.3304209°N	91.0040547°E
41	Chenimaguribao	V41	NA	Dolpatier	Lakhimpur	27.94814°N	80.77777°E
42	Koladhepabao	V42	IC632953	Gohpur	BiswanathChariali	26.88462°N	93.62262°E
43	Miyan bao-1	V43	NA	Gohpur	BiswanathChariali	26.88462°N	93.62262°E
44	Mahsuribao	V44	IC547090	Panigao	Lakhimpur	27.1497°N	94.1111°E
45	Garembao	V45	IC465206	Kutuha	Lakhimpur	27.94814°N	80.77777°E
46	Gumrajbao	V46	IC632954	Hajo	Kamrup Rural	26.25°N	91.53333°E
47	Konbao	V47	IC632955	Dhemaji	Dhemaji	27.47525°N	94.56184°E
48	Poranegheribao	V48	IC632956	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
49	Chengabao	V49	IC394572	Borpeta	Borpeta	26.3304209°N	91.0040547°E
50	Rayada B3	V50	IC380642	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
51	Betgutibao	V51	IC394280	Borbeta	Borpeta	26.3304209°N	91.0040547°E
52	Bogajulbao	V52	IC394956	Majuli	Majuli	26.96274°N	94.16202°E
53	Tar kekuwabao	V53	IC574901	Matmora,Dhakuwakhana	Lakhimpur	27.16539°N	94.11876°E
54	Herepibao	V54	IC575057	Dhakuwakhana	Lakhimpur	27.16539°N	94.11876°E
55	Hawaibao	V55	IC591472	Kesukhona	Dhemaji	27.47525°N	94.56184°E
56	Tulsibao	V56	IC591507	Gohpur	BiswanathChariali	26.88462°N	93.62262°E
57	Samrajbao	V57	IC380618	Borpeta	Borpeta	26.3304209°N	91.0040547°E
58	Panikekuwabao	V58	IC332895	Narowathan	Dhemaji	27.47525°N	94.56184°E
59	Bhuspuribao	V59	IC464420	Korson	Lakhimpur	27.94814°N	80.77777°E
60	Moimonsingiabao	V60	IC538319	Kolabari	BiswanathChariali	25.5132°N	89.8400°E
61	Beroi	V61	IC558260	Karson	Lakhimpur	27.94814°N	80.77777°E
62	Rangi-2	V62	IC591473	Khoga NL	Lakhimpur	27.143455°N	94.167678°E
63	Rangolibao	V63	IC591484	Kotohaguri	Lakhimpur	27.94814°N	80.77777°E
64	Jalprobhabao	V64	IC632957	Titabar	Jorhat	26.59223°N	94.18652°E
65	Panisalibao	V65	IC394939	Nalibari	Nalbari	26.4446°N	91.43824°E
66	India bao	V66	IC632958	Gorehoga	Lakhimpur	27.94814°N	80.77777°E
67	Gadhubamaguri	V67	IC632959	Dhemaji	Dhemaji	27.47525°N	94.56184°E
68	Balambao	V68	IC316458	Katohaguri,NL	Lakhimpur	27.143455°N	94.167678°E
69	Badal-1	V69	IC575075	Khoga,NL	Lakhimpur	27.143455°N	94.167678°E
70	Narayanpuriabao	V70	IC591488	Panigaon	Lakhimpur	27.1497°N	94.1111°E
71	Boga Amona	V71	IC591496	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
72	Khejengbao	V72	NA	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
73	Maguribao	V73	IC298347	Karson	Lakhimpur	27.94814°N	80.77777°E
74	Panindra	V74	IC575267	North lakhimpur	Lakhimpur	27.22849°N	94.09444°E
75	Padumoni	V75	IC575100	North lakhimpur	Lakhimpur	27.22849°N	94.09444°E
76	Panchanan	V76	IC575253	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
77	Basudev	V77	NA	Lakhmipur	Lakhimpur	27.94814°N	80.77777°E
78	Padmanath	V78	IC546548	North lakhmipur	Lakhimpur	27.22849°N	94.09444°E
79	Jolee-2	V79	IC632960	Machkhowa	Dhemaji	27.28616°N	94.44493°E

80	Joldubi-2	V80	IC395027	Mahajan	Lakhimpur	27.94814°N	80.77777°E
81	Pani kekuwa-2	V81	IC591464	Dhakuwakhana	Lakhimpur	27.16539°N	94.11876°E
82	Bogadhepa	V82	NA	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
83	Sonamukhi	V83	IC591490	Bogodebeli	Dhemaji	27.47525°N	94.56184°E
84	Nasati bao-1	V84	NA	Nahatigaokolabari	Lakhimpur	27.94814°N	80.77777°E
85	Negheri-1	V85	IC394535	Korson	Lakhimpur	27.94814°N	80.77777°E
86	Sworgfolla bao-1	V86	IC632947	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
87	Bedel bao-1	V87	IC394562	Khoga	Lakhimpur	27.143455°N	94.167678°E
88	Sonjulbao 1	V88	IC591462	Lakhimpur	Lakhimpur	27.94814°N	80.77777°E
89	Sonamukhi-3	V89	IC575132	Bogodebeli, Dhemaji	Dhemaji	27.47525°N	94.56184°E
90	Jul bao	V90	NA	Potiamohaijan	Lakhimpur	27.94814°N	80.77777°E
91	Amona-3	V91	IC394617	Karson, NL	Lakhimpur	27.22849°N	94.09444°E
92	Negheribao (TTB)	V92	IC394631	Korson, NL	Lakhimpur	27.22849°N	94.09444°E

IC = Indigeneous collection number, NA = Not available

(Supplementary Table S1). The SSR markers were selected from previously available genome sequence (physical) maps of rice (http://www.gramene.org) (McCouch et al. 2002). Seventy-four primers were utilized for the amplification of SSR loci in the current set of DWR landraces. The PCR amplification was carried out using 1-µL of diluted template DNA (20 ngµL^{-1}) of each line in 96-well PCR plates (AXYGEN). PCR reactions were carried out with the program: (1) 5-minute initial denaturation at 94°C, (2) 35 cycles of run, each with denaturation at 94 °C for 1-minute, annealing T_m of ± 2°C with respect to annealing temperature of different primer pairs (Table 2) for 1-minute and extension at 72°C for 1-minute , (3) the final step of extension at 72 °C for 5 minutes. The PCR products, along with a 100 bp ladder, were size fractioned and stained using ethidium bromide staining, on 3% MetaPhorTM (Lonza) agarose gels and visualized using gel documentation system (Alpha Innotech, USA).

Scoring and cluster analysis

The amplified products were set for scoring analysis. If a product is present in a certain genotype, it is designated as '1' and for the absence as '0' for each genotype primer sequence. The binary data matrix generated using polymorphic SSR markers was further utilized for analysis using NTSYS-PC 2.10 software (Rohlf 2000). Parameters like total number of alleles, major allele frequency, gene diversity (H_e), and observed heterozygosity (H_o) were determined using Power Marker V3.0 (Herrera et al. 2008). Polymorphism information content (PIC) was evaluated by using the formula given by Anderson et al. (1993).

$$PIC = 1 - \sum_{j=1}^{n} P_{ij}^2$$

Where P^{ij} is the frequency of the j^{th} allele or band for the i^{th} marker and the summation extends over 'n' alleles or bands.

The coefficient of similarities and relatedness among the ninety-two deep water rice genotypes was validated using Jaccard's coefficient followed by clustering with UPGMA (unweighted pair group method with arithmetic mean) algorithm by implemented in DARwin 6.0 software (10,000 bootstrapping) (Perrier et al. 2003) to provide a general visualization of genetic relationship among these 92 DWR landraces. The principal coordinate analysis (PCoA) and analysis of molecular variance (AMOVA) were estimated using GenAlEx.

Population structure analysis

The population structure was analysed using a Bayesian-based model method in STRUCTURE 2.3.4 (Pritchard et al. 2000) to find out the genetic relationship among 92 DWR landraces. For reducing the computational period and easy handling, the STRUCTURE run was done using the command line Python program under a Linux environment. Five independent runs were performed by setting the hypothetical number of expected populations (K) from 1 to 10. The data were set with a burn-in period and Monte Carlo Markov Chain (MCMC) iterations values of 1,00,000 each. Web-based STRUCTURE HARVESTER 6.0 (Earl and Von Holdt 2012) was used to determine the most likely possible number of subpopulations (K) wherein ΔK statistics was subject to the rate of change in log probability (LnP(D)) between consecutive K values and generated the Evanno plot.

Results

Genetic diversity analysis of 92 DWR landraces using SSR markers

Seventy four SSR markers were used to assess the genetic diversity across 92 deep water rice (DWR) landraces. Out of 74 SSR markers, 56 generated reproducible and polymorphic

patterns. A total of 139 alleles were detected by using these 56 pairs of SSR primers with the allele number per locus ranging from 2-5 with an average of 2.48 alleles per marker. The maximum number of alleles was observed in the SSR marker, RM55 (Fig. 2). The level of polymorphism among the 92 rice genotypes was evaluated by calculating the allele frequency, number of alleles for each SSR locus, H_o, H_e and PIC values for each of the 56 SSR loci. Major allele frequency was found to range from 0.29 (RM206) to 0.88 (RM413) while H_e ranged from 0.21 (RM413) to 0.74 (RM206) with a mean of 0.47. The H_o ranged from 0 to 0.73, with a mean of 0.08. The overall PIC value ranged from a minimum of 0.18 (RM413) to a maximum of 0.69 (RM206) with an average value of 0.39. Low PIC values might be a result of closely related genotypes and high PIC values were an indication of highly diverse genotypes. The SSR marker RM 206 was found to be the most appropriate marker with the highest PIC value of 0.69, followed by RM 171 (0.67), RM 250 (0.65), RM 5 (0.56), RM 55 (0.55) and RM 412 (0.52) (Table 2).

Cluster analysis of SSRs

Similarity matrices of all the 92 accessions were generated using NTSYS-pc v2.10. Phylogeny analysis grouped all the 92 landraces into seven different clusters at a 0.64 coefficient. The first cluster comprised 14 accessions, where Amona 1 (v6) to Khalburuni bao (v40) were found to be closely related (Fig. 3). Cluster I has been divided into two subpopulations; one consists of Ujanni Khamti bao, Jeng bao-1, Garo bao and the other has Rangi 1, Amona 1 shared phylogenetically conserved single progenitor. Cluster II was considered to be the largest group, having 27 accessions, which was the maximum number. The second cluster was grouped into two subclusters, one starting from v16 to v22 and the other comprised genetic resources v46 to v59. In cluster II, Dhasuri bao and Ikora bao are closely related but distant relatedness to Joleebao 1, Bokajohingia bao, Betnasali bao and Bogadhepa bao. The dendrogram showed that cluster I and cluster II shared a common lineage. The cluster III represents the most cleared or unmixed clade, having 2 different subpopulations, one starting from v30 to v66 and the second from v68 to v76. India bao (v66) and Balam bao (v68) were found to be evolutionary conserved and shared common parentage. Clusters IV, V and VI are comprised of 9, 13 and 8 genotypes each. The last two accession v4 and v5 are phylogenetically separately evolved and grouped into cluster VII, which is found to be unrelated to other clades.

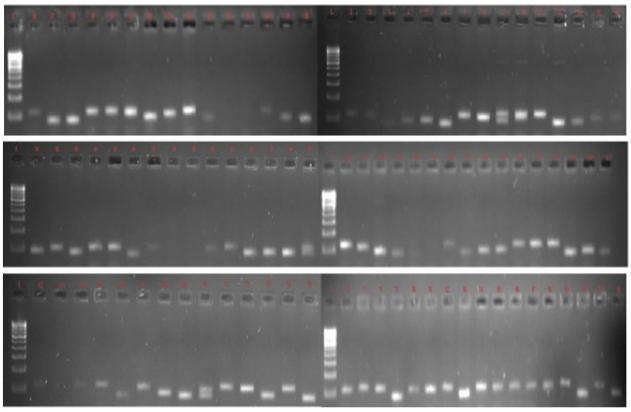

Fig. 2. Allelic variations in 92 deep warer rice germplasms using RM5 marker, Ladder=100 bp

Population structure analysis

The panel of 92 DWR landraces was used to imply the current number of clustered sub-populations. We have estimated sub-populations using LnP(D)-derived Δk for k from 3 to 10 with five independent runs (Fig. 4a). The log probability of the data (LnP(D) and ad hoc statistics ΔK from the evanno plot generated, both revealed the presence of seven possible sub-populations in the model-based population (Fig. 4b). Genotypes having the I_{st} value more than 0.9 were considered to be homozygous genotypes (pure) for that sub-population. The estimated F_{st} values were found to be extremely variable among the sub-populations, which were predicted based on STRUCTURE at ΔK = 7. According to F_{st} values (F_{st} > 0.5), 4 genotypes were included under

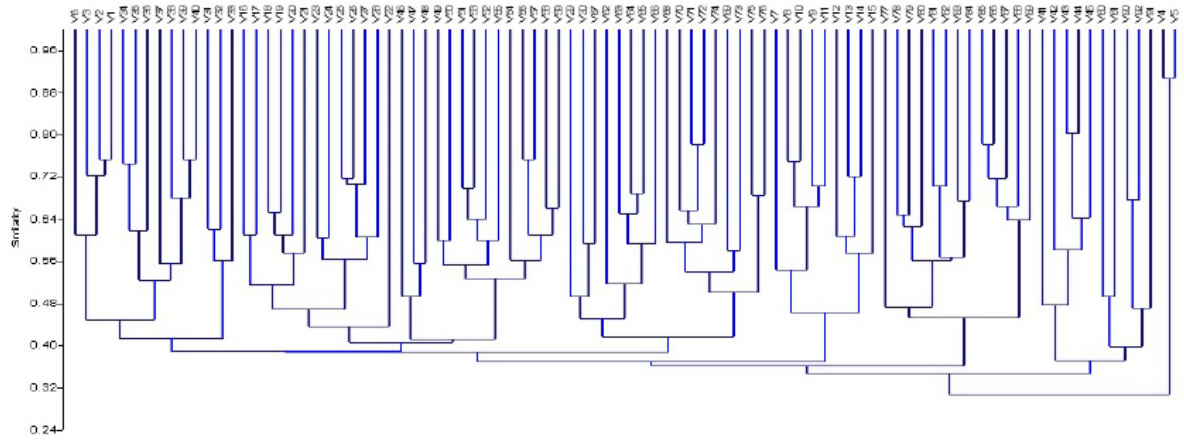

Fig. 3. Dendrogram showing genetic diversity among 92 DWR landraces based on the SSR markers

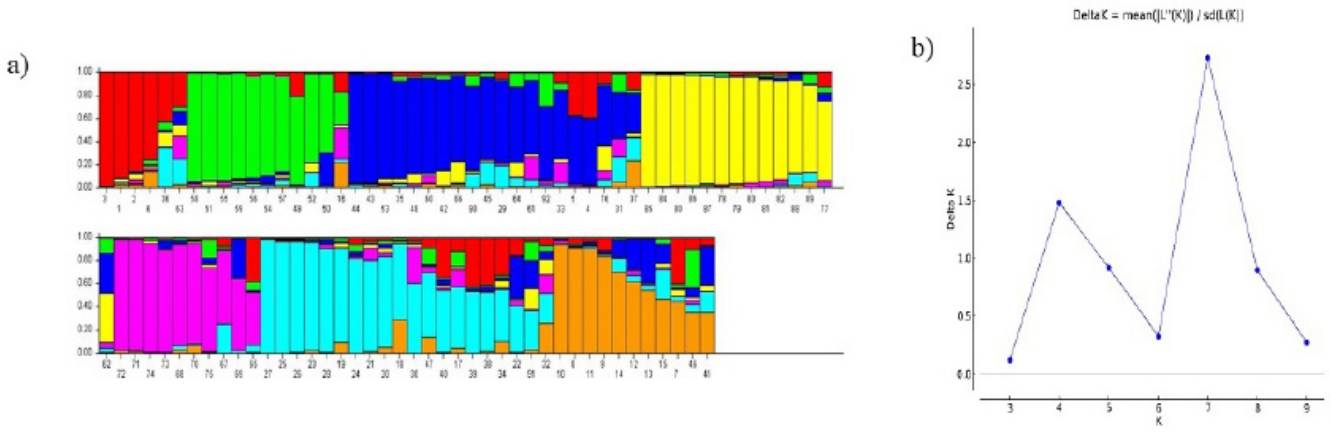

Fig. 4. Population structure analysis in 92 DWR landraces a) Structure plot at K=7 where each color represent seven different clustered population. b) A plot for ΔK for different numbers of clustered sub-populations using structure harvester and the highest peak ΔK ≥ 2.5 was obtained at K = 7

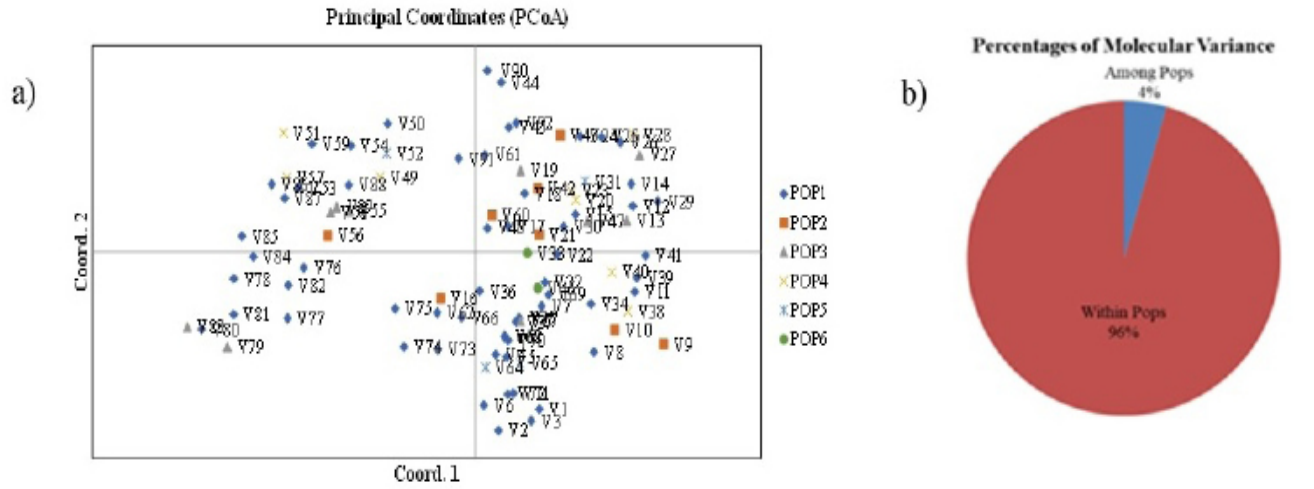

Fig. 5. Genetic diversity analysis in DWR landraces based on SSR markers (a) Principal coordinate analysis (PCoA) (b) Analysis of molecular variance (AMOVA)

sub-population I, Sub-population II contained 10 genotypes, Sub-population III contained 18 genotypes, Sub-population IV contained 13 genotypes, Sub-population V consisted of 9 genotypes, Sub-population VI comprised the highest number, i.e., 16 and Sub-population VII consisted of 7 genotypes showing both pure and admixture individuals and the remaining 15 genotypes were found to be highly ad mixture type. According to I_{st} values being more than 0.9, two pure genotypes in sub-population 1, four in sub-population II, three in sub-population III, eight in sub-population IV, three in sub-population V, four in sub-population VI and three in sub-population VII were found to be homozygous.

Principal coordinate analysis (PCoA) and analysis of molecular variance (AMOVA)

In order to understand the complementary idea of genetic relationship among the DWR landraces, PCoA has been done. PCoA exposed significant diversity among these DWR landraces. The analysis categorised all these accessions into six groups involving different genetic resources as pop1, pop2, pop3, pop4, pop5 and pop6. The results are roughly parallel to the population structure analysis and phylogenetic analysis. All six sub-population groups were dispersed in the PCoA plot (Fig. 5a). In the present investigation, the PCoA explained 8.59% variation at 1-axis, 7.61% variation at 2-axes, and 6.89% variation at 3-axes. The highest cumulative variations were observed to be 23.10% at 3 axes. These variations observed in our analysis confirmed that our DWR landraces are diverse and genetically distinct from each other.

To analyse the distribution of genetic diversity among and within the population, an AMOVA study has performed. Analysis of molecular variance (AMOVA) revealed high genetic variation within populations (96%) and low

Table 2. Genetic parameters of 56 polymorphic SSR markers in 92 DWR landraces

Marker-id	Major Allele Frequency (MAF)	No of Genotypes (Ng)	Number of alleles (Na)	Gene Diversity (H_e)	Heterozygosity (H_o)	Polymorphic information content (PIC)
RM1	0.5163	3.0000	2.0000	0.4995	0.0326	0.3747
RM5	0.4783	3.0000	3.0000	0.6352	0.0000	0.5636
RM250	0.4138	6.0000	4.0000	0.7039	0.0460	0.6526
RM55	0.5899	7.0000	5.0000	0.5947	0.0674	0.5529
RM85	0.5761	5.0000	3.0000	0.5320	0.1304	0.4406
RM130	0.5163	4.0000	3.0000	0.5147	0.0326	0.3980
RM514	0.5054	6.0000	3.0000	0.6144	0.7391	0.5414
RM241	0.5326	3.0000	2.0000	0.4979	0.0217	0.3739
RM252	0.7011	3.0000	2.0000	0.4191	0.0543	0.3313
RM13	0.6685	4.0000	3.0000	0.4982	0.0109	0.4475
RM412	0.4725	6.0000	3.0000	0.6057	0.0989	0.5235
RM541	0.6413	2.0000	2.0000	0.4601	0.0000	0.3542
RM11	0.5815	6.0000	3.0000	0.5741	0.4457	0.5110
RM125	0.8098	4.0000	3.0000	0.3120	0.3587	0.2696
RM214	0.7717	2.0000	2.0000	0.3523	0.0000	0.2903
RM25	0.7826	2.0000	2.0000	0.3403	0.0000	0.2824
RM153	0.7011	6.0000	3.0000	0.4560	0.2609	0.4037
RM377	0.7391	3.0000	2.0000	0.3856	0.0870	0.3113
RM433	0.8696	4.0000	4.0000	0.2318	0.2609	0.2136
RM205	0.8478	3.0000	2.0000	0.2580	0.0217	0.2247
RM215	0.5393	5.0000	3.0000	0.5424	0.5843	0.4443
RM206	0.2935	7.0000	4.0000	0.7455	0.0326	0.6980
RM19	0.5109	3.0000	2.0000	0.4998	0.0217	0.3749
RM235	0.6848	2.0000	2.0000	0.4317	0.6304	0.3385
OSR17	0.4565	4.0000	3.0000	0.6153	0.3043	0.5347
RM44	0.7500	5.0000	3.0000	0.4034	0.1848	0.3645
RM171	0.3804	4.0000	4.0000	0.7257	0.0000	0.6772
RM259	0.5924	3.0000	2.0000	0.4829	0.0109	0.3663
RM271	0.5761	2.0000	2.0000	0.4884	0.0000	0.3691
RM284	0.6522	2.0000	2.0000	0.4537	0.0000	0.3508
RM307	0.4176	8.0000	4.0000	0.6866	0.0549	0.6286
RM312	0.6087	2.0000	2.0000	0.4764	0.0000	0.3629
RM316	0.5163	3.0000	2.0000	0.4995	0.0326	0.3747
RM334	0.5326	2.0000	2.0000	0.4979	0.0000	0.3739
RM408	0.5435	2.0000	2.0000	0.4962	0.0000	0.3731
RM413	0.8804	2.0000	2.0000	0.2105	0.0000	0.1884
RM447	0.5000	2.0000	2.0000	0.5000	0.0000	0.3750
RM489	0.8533	4.0000	3.0000	0.2608	0.0109	0.2445
RM495	0.6087	2.0000	2.0000	0.4764	0.0000	0.3629
RM507	0.6222	2.0000	2.0000	0.4701	0.0000	0.3596
RM510	0.5652	2.0000	2.0000	0.4915	0.0000	0.3707
RM522	0.5978	3.0000	2.0000	0.4809	0.1739	0.3652
RM144	0.5934	2.0000	2.0000	0.4826	0.0000	0.3661

RM287	0.5435	2.0000	2.0000	0.4962	0.0000	0.3731
RM152	0.5435	2.0000	2.0000	0.4962	0.0000	0.3731
RM161	0.6630	2.0000	2.0000	0.4468	0.0000	0.3470
RM237	0.6923	2.0000	2.0000	0.4260	0.0000	0.3353
RM283	0.5652	2.0000	2.0000	0.4915	0.0000	0.3707
RM162	0.8587	2.0000	2.0000	0.2427	0.0000	0.2132
RM244	0.7011	3.0000	2.0000	0.4191	0.0109	0.3313
RM5686	0.5000	6.0000	3.0000	0.5964	0.1413	0.5150
RM474	0.6196	2.0000	2.0000	0.4714	0.0000	0.3603
RM249	0.6087	2.0000	2.0000	0.4764	0.0000	0.3629
RM547	0.5652	2.0000	2.0000	0.4915	0.0000	0.3707
RM280	0.6522	2.0000	2.0000	0.4537	0.0000	0.3508
RM276	0.6630	2.0000	2.0000	0.4468	0.0000	0.3470
Mean	0.6101	3.3214	2.4821	0.4796	0.0868	0.3924
Total	34.1672	186.0000	139.0000			

genetic differentiation among populations (4%) as shown in Fig.5b.The maximum F_{ST} value calculated between the subpopulations within and among populations was found to be 0. The values observed for PhiPT and Nm were 0.0439 and 5.444, respectively (Table 3).

Discussion

Understanding the genetic diversity of crop plants is vital for conservation, breeding, and sustainability. India is considered a secondary centre of origin for cultivated rice, and its north-eastern region, particularly Assam and the Eastern Himalayas, is recognized as a hotspot of rice diversity (Singh et al. 2016). Genetic diversity analysis utilizing various genomic resources helps to understand the evolutionary relationship among them. The identification of genetic resources and their evolutionary relationship assists the crucial step of breeding, parental choice and tolerance towards biotic stress as well as abiotic stress (Assefa et al. 2019). Deep water rice, cultivated in regions submerged for several months, represents an underexploited resource in modern breeding. Its adaptation to prolonged inundation has led to unique physiological traits, such as rapid internode elongation and anaerobic germination. In India, DWR is mainly grown in eastern Uttar Pradesh, Bihar, Bengal, and Assam. However, systematic studies on its genetic architecture have been limited. Rohilla et al. (2020) reported significant intra-population variability among DWR landraces from Assam using SNP markers, and our SSR-based data further confirm the heterogeneous genetic background of this rice type. The genotypes used in the present study as a genomic resource are primarily grown in different flood-prone areas of Assam, India. The DWR landraces play a fundamental role in the socio-cultural life of the people of Assam, India, which provides food

Table 3. Summary of AMOVA in 92 DWR landraces using SSR markers

Source of variation	Degrees of freedom (df)	Sum of squares (SS)	Mean of squares (MS)	Estimated variation CV	Percentage of variation (%)
Among Pops	5	126.008	25.202	0.603	4%
Within Pops	178	2332.633	13.105	13.105	96%
Total	183	2458.641		13.707	100%
[a]PhiPT	0.0439				
[b]Nm	5.444				
Statistic	Value		*p value*		
Fst	0.043		0.019		
Fst max	0.528				
Fst	0.083				

p-values estimates are based on 9999 permutations.
[a]PhiPT = AP/ (WP + AP) = AP/TOT
[b]Nm = ((1/PhiPT) − 1)/4

and nutrition to 30 million people in a condition where no other crops can be cultivated due to seasonal flood stress. DWR, commonly known as Bao Dhan, withstands water stagnation of beyond 50 cm that remains for periods of 2-4 months in the growing season. The average yield of these landraces is poor because of the variable ecosystem that stays right from the seedling to the harvesting stage of the crop (Rohilla et al. 2019). Globally, several studies corroborate the high genetic variability of traditional rice types. For example, Thomson et al. (2007) analyzed rice germplasm using SSR markers and reported widespread diversity across different ecotypes. More recent genome-wide studies using SNP and DArTseq markers have also revealed substantial differentiation among deep water, lowland, and upland rice (McNally et al. 2009; Mansueto et al. 2017). These efforts mirror present findings and underscore the value of traditional DWR germplasm in global rice improvement. A comprehensive study by Travis et al. (2015) assessing the genetic diversity of rice accessions from Bangladesh, Assam, and West Bengal using high-density genotyping revealed that accessions from Assam exhibit notable genetic uniqueness and heterogeneity. Their work highlighted the importance of local adaptation and farmer selection in shaping the genetic structure of landraces from this region, corroborating our findings of extensive intra-population variability and the presence of distinct subpopulations in deep water rice of Assam. In the current study, 56 different highly polymorphic SSR markers were utilized to study the genetic diversity among the total of 92 DWR landraces. The key purpose of the investigation was to establish genetic diversity, relationship and population structure to categorise these DWR landraces as a suitable parent for hybridisation. For the development of new high-yielding deep water tolerant varieties, frequent inbreeding and domestication are major hurdles and as such, knowledge about genetic diversity would go a long way in overcoming these issues. Previously, genome-wide genetic diversity analysis has been done in 94 DWR landraces of Assam by using a 50K SNP genic chip, which revealed huge genetic differences among these genotypes, some of which are common to our panel (Rohilla et al. 2020). The use of SSR markers in this study revealed substantial genetic variation among 92 DWR landraces collected from flood-prone regions of Assam. The findings align with earlier reports that SSR markers, owing to their high reproducibility, co-dominant inheritance, and genome-wide distribution, are effective in detecting polymorphism among rice genotypes (Panaud et al. 1996; Temnykh et al. 2001). PCR-based SSR marker techniques were proven to be an effective tool to understand genomic divergence, which is employed for molecular characterisation between various DWR landraces. The polymorphic pattern was spawned by 56 sets of pruned polymorphic markers, which revealed an extreme level of diversification, polymorphism and evolutionary conserved clades among these DWR accessions. Our study discovered a total of 139 alleles with 2 to 5 alleles (average of 2.48) using SSR markers. The genetic diversity ranged from 0.21 to 0.74, with a mean of 0.47 and the value of MAF varied from 0.29 (RM206) to 0.88 (RM413). The experimental heterozygosity level detected ranged from 0.00 to 0.73, with a mean value of 0.08.

The genotypes assessed in this study showed substantial allelic richness and heterozygosity. A total of 139 alleles across 56 SSR loci (mean = 2.48) and high gene diversity values (He = 0.47) reinforce the region's genetic wealth. Such variability is essential for identifying trait-specific donors for breeding programs aimed at enhancing submergence tolerance, yield, and grain quality. The average PIC value ranged from the lowest 0.18 (RM413) to the highest 0.69 (RM206), with an average value of 0.39. The genotypes with low PIC value were phylogenetically closely related and the higher the value of PIC, the greater is the diversity among the genotypes. The greatest diversity was achieved by the polymorphic marker RM206 with the PIC value of 0.6980 was found to be the most informative marker followed by RM171 (0.67), RM250 (0.65), RM5 (0.56), RM55 (0.55) and RM 412 (0.52) because markers with PIC values of 0.5 or more are regarded as having higher discriminatory power for genetic diversity analysis (Hildebrand et al. 1992). Similar results for mean PIC were reported in the 83 landraces of northeast India (Das et al. 2013). Similarly, the mean PIC values in the range of 0.37 and 0.38 were obtained in sets of 14 improved rice varieties and 27 landraces of rice collected from different areas of seven Indian states (Pachauri et al. 2013), which had lesser PIC content than ours. The PIC value in the range of 0.23 and 0.25 (Choudhary et al. 2013) was also found in Assam rice and whole northeast rice collections, respectively. In a separate study high average PIC value (0.61) in a population structure of 54 rice accessions using 14 polymorphic SSR markers indicated substantial polymorphism. Despite a narrow genetic base with two main sub-populations, significant variation existed among individuals (Suvi et al. 2019). All the reports on PIC showed that the accessions used in the present study material were found to have higher genetic diversity. The clustering of accessions into seven genetic groups through UPGMA and STRUCTURE analyses revealed admixture and ancestral lineage sharing, which is typical for landraces maintained under traditional farming. The high within-population variance (96%) observed in AMOVA indicates that localized selection and farmer-mediated seed exchange have preserved extensive diversity within regions. Kimwemwe et al. (2023) reported an assessment of the genetic diversity and population structure of 94 rice genotypes from the Democratic Republic of Congo using 8,389 DArTseq-based SNP markers. Verma et al. (2024) also carried out population structure analysis, distance-based neighbor joining cluster and principal coordinate analysis

in rice using genotypic data, grouping the genotypes into two sub-populations. Further, analysis of molecular variance and pairwise FST values revealed significant differentiation among all the pairs of sub-populations. Results revealed moderate marker informativeness, five genetic sub-populations, and significant genetic differentiation (Fst = 0.52) with limited gene flow (Nm = 0.23). High genetic diversity was evident, suggesting valuable potential for future rice breeding and improvement programs.

On average, more than 60% percent of heterozygosity values were precise and more reliable for population genetics and the higher heterozygosity contributes to the chances of having a huge genetic variation (Lin et al. 2020). The SSR assisted markers such as RM235, RM215, and RM514 correlate 65% of heterozygosity (H_o) (0.630, 0.5843, 0.7391, respectively) and thus direct the polymorphic ability in deep-water rice variety, while gene diversity (H_e) ranged from 0.21 to 0.74, showing huge diversity among DWR landraces. The genotypes in each subpopulation play a key role in the observation and help the study panel to identify the unevenness at loci and improve allele richness. The disproportionate land races in the subpopulation prevent the panel study from finding a larger number the private or novel alleles. The private allele is crucial for characterising the variability among the genotypes and channelling the key adaptive genes among the DWR germplasm.

The population structure analysis divided these DWR genotypes into seven different sub-populations on the basis of polymorphic markers. The presence of admixture type in each sub-population suggested that all these genotypes undergo multiple recombination procedures over the years and adapting themselves resulted in a high level of diversity among them. The genotypes are adapted under deep-water conditions and therefore might be evolutionarily and genetically similar in nature under similar selection pressure. Similarly, a study of population structure analysis divided 94 DWR genotypes into the 4 sub-populations consisting of pure and admixture type genotypes by using a 50K SNP rice genic chip as genotyping criterion (Rohilla et al. 2020). AMOVA's result in the study concludes a high level of genetic diversity (96%) within subpopulations and a low (4%) variation among subpopulations. The variations were significant conferring to the partitioning value (p <0.001). The feasible elucidation for high variation within groups is the assortment of respective agronomically important attributes in breeding programs. A PhiPT value greater than 0.15 is considered significant in differentiating populations (Frankham et al. 2002). Low PhiPT value of 0.049 between the sub-populations correlates with the less genetic divergence (4%) of genotypes among the populations. The low level of diversity among the population is directly proportional to the higher gene flow and low genetic exchange among subpopulations (Babu et al. 2014). The higher value of Nm suggests, thus far, less-exploited genotypes such as novel alleles or genetically introgressed traits, which are the most preferred for diversity studies. The various convention factors, such as random migration, population size, selection, and mutation rate, should be kept in mind to estimate Nm indirectly from the PhiPT. It has been reported that an Nm value less than 1 indicates less gene exchange among populations, whereas in our study, the Nm values of 5.4 suggested that a high genetic exchange or high gene flow occurs and leads to less genetic differentiation among populations. PCoA and the population structure analysis performed with 56 polymorphic SSR markers were found to be correlated with the resultant tree obtained and show significance genetic differentiation of the study group.

Breeders should take meticulous effort to manifest a high Nm value, depicting recurrent gene flow among the subpopulations. Study panel should be highly precautious while inferring the Nm value indirectly, even if it is still significant to determine the magnitude of gene flow. The understanding of the genetic divergence in the deep–water rice germplasm would foster future study as well as the identification of the stocks that are stress tolerant and maintain plant water molecular dynamics. In conclusion, the present study analyses the pattern of divergence that exists in a population of 92 DWR genotypes. Based on various statistical methods, we identified seven subgroups within 92 DWR, which can be utilized as a great source for establishing an association mapping panel. The mean number of alleles per locus and genetic diversity indicated the existence of broad genetic collections. The result of the structure analysis comes in parallel with the clustering method of neighbor-joining tree and PCoA. So, the outcome of this study which identifies the genetic diversity of these unexplored deep-water rice (DWR) landraces collected from different districts of Assam, can be utilized for association analysis, parental line selection, mapping population development in breeding programs and hybrid variety production by exploiting the natural genetic variation existing in this population. By employing various statistical methods, we were able to identify seven distinct subgroups within the 92 DWR genotypes, which can serve as an excellent resource for establishing an association mapping panel. The findings of Sar et al. (2024) explored the genetic structure and yield-related traits of 181 *aus* rice accessions using over 399,000 SNP markers and GWAS analysis. Six genetically distinct subpopulations with geographic patterns were identified. Key yield and plant architecture traits were associated with specific principal components. GWAS revealed OsSAC1, OsGLT1, and OsPUP4/Big Grain 3 as major genes influencing yield traits. These findings enhance understanding of aus rice diversity and provide valuable genetic targets for rice improvement. Therefore, this study provides valuable insights into the genetic diversity of these DWR genotypes,

which can be effectively utilized in association analysis, parental line selection, mapping population development in breeding programs, and the production of hybrid varieties by harnessing the natural genetic variation present in this population.

Overall, the unique genetic architecture of DWR landraces of Assam, as revealed by SSR analysis, highlights their potential as valuable pre-breeding materials. The presence of both rare and common alleles, high PIC markers such as RM206, and the indication of distinct subpopulations are assets for marker-assisted breeding, especially for traits like submergence tolerance, early vigor, and elongation ability under flood stress (Hattori et al. 2009). Future work should integrate high-throughput genotyping platforms and genome-wide association studies (GWAS) to map important adaptive loci in DWR. Preservation and characterization of such traditional varieties are crucial for developing climate-resilient rice that caters to food security in flood-affected regions.

Supplementary material

Supplementary Table S1 is provided, which can be accessed at www.isgpb.org

Authors' contribution

Conceptualization of research (PS, TKM); Designing of the experiments (PS); Contribution of experimental materials (DC); Execution of field/lab experiments and data collection (DC, DSH, MR, PR); Analysis of data and interpretation (SK, DSH, MR); Preparation of the manuscript (MR, DSH, PS, TKM).

Acknowledgments

Fellowship to DSH from the Ministry of Tribal Affairs, Govt. of India (Ref. No. 201819-NFST-MAN-02172) is thankfully acknowledged**.**

References

Anderson J.A., Churchill G.A., Autrique J.E., Tanksley S.D. and Sorrells M.E. 1993. Optimizing parental selection for genetic linkage maps. Genome, **36**: 181–186. https://doi.org/10.1139/g93-024

Assefa T., Mahama A.A., Brown A.V. and Cannon E.K.S. 2019. A review of breeding objectives, genomic resources, and marker-assisted methods in common bean (*Phaseolus vulgaris* L.). Mol. Breed., **39**: 20. https://doi.org/10.1007/s11032-018-0920-0

Babu B.K., Meena V., Agarwal V. and Agrawal P. 2014. Population structure and genetic diversity analysis of Indian and exotic rice (*Oryza sativa* L.) accessions using SSR markers. Mol. Biol. Rep., **41**(7): 4329–39. https://doi: 10.1007/s11033-014-3304-5

Choudhary B., Khan M.L. and Dayanandan S. 2013. Genetic structure and diversity of indigenous rice (*Oryza sativa*) varieties in the eastern Himalayan region of northeast India. Springer Plus, **2**: 228. http://www.springerplus.com/content/2/1/228

Das B., Sengupta S., Parida S.K., Roy B., Ghosh M., Prasad M. and Ghose T.K. 2013. Genetic diversity and population structure of rice landraces from Eastern and North Eastern States of India. BMC Genet., **14**: 7110. https://doi.org/10.1186/1471-2156-14-71

Dellaporta S.L., Wood J. and Hicks J.B. 1983. A plant DNA minipreparation: Version II. Plant Mol. Biol. Rep., **1**: 19–21. https://doi.org/10.1007/BF02712670

Earl D.A. and VonHoldt B.M. 2012. Structure Harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Cons. Genet. Res., **4**: 359-361.https://doi.org/10.1007/s12686-011-9548-7

Frankham R., Ballou J.D. and Briscoe D.A. 2002. Introduction to conservation genetics. Cambridge University Press. doi:10.1017/S0016672304216913

Garris A.J., Tai T.H., Coburn J. Kresovich S. and McCouch S. 2005. Genetic structure and diversity in *Oryza sativa* L. Genetics, **169**: 1631–1638. doi: 10.1534/genetics.104.035642

Hattori Y., Nagai K., Furukawa S., Song X.J., Kawano R., Sakakibara H., Wu J., Matsumoto T., Yoshimura A., Kitano H. and Matsuoka M. 2009. The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. Nature, **460** (7258): 1026-1030. doi: https://doi.org/10.1038/nature08258

Herrera T.G., Duque D.P., Almeida I.P., Nunez G.T., Pieters A.J., Martinez C.P. and Tohm J.M. 2008. Assessment of genetic diversity in Venezuelan rice cultivars using simple sequence repeats markers. Elect. J. Biotechnol., **11**(5):3-4.http://dx.doi.org/10.4067/S0717-34582008000500003

Hildebrand C.E., Torney D.C. and Wagner R.P. 1992. Informativeness of polymorphic DNA markers. Los Alamos Sci., **20**: 100-102. doi:10.1007/978-1-62703-389-3_18

Kanawapee N., Sanitchon J., Srihaban P. and Theerakulpisut P. 2011. Genetic diversity analysis of rice cultivars (Oryza sativa L.) differing in salinity tolerance based on RAPD and SSR markers. Elect. J. Biotechnol.,**14**: 1-14. doi:10.2225/vol14-issue6-fulltext-4

Kimwemwe, P. K., Bukomarhe, C. B., Mamati, E. G., Githiri, S. M., Civava, R. M., Mignouna, J., Kimani, W., and Fofana, M. 2023. Population Structure and Genetic Diversity of Rice (*Oryza sativa* L.) Germplasm from the Democratic Republic of Congo (DRC) Using DArTseq-Derived Single Nucleotide Polymorphism (SNP). *Agronomy*, **13:** 1906. https://doi.org/10.3390/agronomy13071906

Lin E., Zhuang H., Yu J., Liu X., Huang H., Zhu M. and Tong Z. 2020. Genome survey of Chinese fir (*Cunninghamia lanceolata*): Identification of genomic SSRs and demonstration of their utility in genetic diversity analysis. Sci. Rep., **10**: 4698. https://doi.org/10.1038/s41598-020-61611-0

Mansueto, L., Fuentes, R.R., Borja, F.N., Detras, J., Abriol-Santos, J.M., Chebotarov, D., Sanciangco, M., Palis, K., Copetti, D., Poliakov, A. and Dubchak, I. 2017. Rice SNP-seek database update: new SNPs, indels, and queries. Nucleic Acids Res., **45** (D1): D1075-D1081.https://doi.org/10.1093/nar/gkw1135

Matin S., Ashrafuzzaman M., Islam M.M., Sikdar S.U. and Zobayer N. 2012. Molecular marker based (SSR) genetic diversity analysis in deep water rice germplasms of Bangladesh. Int. J. Biosci., **2**: 64-72. https://doi.org/10.1007/s12298-016-0377-8

McCouch S., Teytelman L., Xu Y., Lobos K.B., Clare K., Walton M., Fu B., Maghirang R., Li Z. and Xing Y. 2002. Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.) DNA Res., **9**: 199-207. doi: 10.1007/BF02172406

McCouch S.R., Chen X., Panaud O., Temnykh S., Xu Y., Cho Y.G., Huang N., Ishii T. and Blair M. 1997. Microsatellite marker

development, mapping and applications in rice genetics and breeding. Plant Mol. Biol., **35**: 89–99. https://doi.org/10.1007/978-94-011-5794-0_9

McNally, K.L., Childs, K.L., Bohnert, R., Davidson, R.M., Zhao, K., Ulat, V.J., Zeller, G., Clark, R.M., Hoen, D.R., Bureau, T.E. and Stokowski, R. 2009. Genome-wide SNP variation reveals relationships among landraces and modern varieties of rice. Proc. the National Academy of Sciences, **106** (30): 12273-12278.https://doi.org/10.1073/pnas.0900992106

Pachauri V., Taneja N., Vikram P., Singh N.K. and Singh S. 2013. Molecular and morphological characterization of Indian farmers rice varieties (*Oryza sativa* L.). Aust. J. Crop Sci.,**7**(7): 923.

Panaud O., Chenayd X. and McCouch S.R. 1996. Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.) Mol. Genet. Genom., **252**: 597-607. doi: 10.1007/BF02172406

Perrier X., Flori A. and Bonnot F. 2003. Methods for Data Analysis. In: Hamon Seguin P, M., Perrier X., Glazmann J.C., Eds. Genetic Diversity of Cultivated Tropical Plants, Science Publisher Inc and Cirad, Montpellier, 31-63.

Pritchard J.K., Stephens M. and Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics, **155**: 945–959. doi: 10.1093/genetics/155.2.945

Rohilla M., Roy P., Chowdhury D., Sharma K.K., Saikia P., Sen P., Singh N.K. and Mondal T.K. 2019. Bao Dhan of Assam: organically grown indigenous rice slated to increase farmer's income. Curr. Sci., **116**:5.https://www.jstor.org/stable/27137924

Rohilla M., Singh N., Mazumder A., Sen P., Roy P., Chowdhury D., Singh N.K. and Mondal T.K. 2020. Genome wide association studies using 50K rice genic SNP chip unveil genetic architecture for anaerobic germination of deep water rice population of Assam, India. Mol. Genet Genom., **295**: 1211–1226. https://doi.org/10.1007/s00438-020-01690-w

Rohlf F. 2000. NTSYS-PC, Numerical Taxonomy System for the PC Exeter Software, Version 2.1. Applied Biostatistics Inc., Setauket.

Sar P., Gupta S., Behera M., Chakraborty K., Ngangkham U., Verma B.C., Banerjee A., Hanjagi P.S., Bhaduri D., Shil S., Kumar J., Mandal N.P., Kole P.C., Purugganan M.D. and Roy S. 2024. Exploring Genetic Diversity within *aus* Rice Germplasm: Insights into the Variations in Agro-morphological Traits. Rice, **17**: 20. https://doi.org/10.1186/s12284-024-00700-4

Singh B., Singh N., Mishra S., Tripathi K., Singh B.P., Rai V., Singh A.K. and Singh N.K. 2018. Morphological and molecular data reveal three distinct populations of Indian wild rice *Oryza rufpogon* Grif. Species Complex. Front. Plant Sci., **9**: 123. https://doi.org/10.3389/fpls.2018.00123

Singh N., Choudhury D., Tiwari G., Singh A.K., Kumar S., Srinivasan K., Tyagi R.K., Sharma A., Singh N. and Singh R. 2016. Genetic diversity trend in Indian rice varieties: an analysis using SSR markers. BMC Genet., **17**: 127. https://doi.org/10.1186/s12863-016-0437-7

Siwach P., Jain S., Saini N., Chowdhury V.K. and Jain R.K. 2004. Allelic diversity among Basmati and non-Basmati long-grain indica rice varieties using microsatellite markers. J. Plant Biochem. Biotechnol., **13**: 25-32. https://doi.org/10.1007/BF03263186

Temnykh S., DeClerck G., Lukashova A., Lipovich L., Cartinhour S. and McCouch S.R. 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res., **11**: 1441–1452. doi: 10.1101/gr.184001

Thomson, M.J., Septiningsih, E.M., Suwardjo, F., Santoso, T.J., Silitonga, T.S. and McCouch, S.R. 2007. Genetic diversity analysis of traditional and improved Indonesian rice (*Oryza sativa* L.) germplasm using microsatellite markers. Theor. Appl. Genet.,**114**: 559–568.https://doi.org/10.1007/s00122-006-0457-1.

Travis, A.J., Norton, G.J., Datta, S., Sarma, R., Dasgupta, T., Savio, F.L., Macaulay, M., Hedley, P.E., McNally, K.L., Sumon, M.H. and Islam, M.R. 2015. Assessing the genetic diversity of rice originating from Bangladesh, Assam and West Bengal. Rice, 8: 1-9. https://doi.org/10.1186/s12284-015-0068-z

Verma H., Theunuo S., Das S.P., Kumar A., Aochen C., Devi L., Sarika K., Baishya L.K., Rajkhowa D.J., Kalita H., Mishra V.K., Behera P.P. and Sarma R.N. 2024. Genetic diversity and marker-trait association analysis for grain quality, yield and yield attributes in hill rice (*Oryza sativa* L.). Indian J. Genet. Plant Breed., **84**(4): 587-601.https://doi.org/10.31742/ISGPB.84.4.9

Supplementary Table S1. Details of the SSR primers used for genotyping

S. No	Primer-id	Sequence	T_m	PCR product (bp)	Chromosome no.
1	RM1	F- 5′GCGAAAACACAATGCAAAAA-3′	55	113	1
		R- 3′GCGTTGGTTGGACCTGAC-5′			
2	RM5	F- 5′TGCAACTTCTAGCTGCTCGA-3′	57	113	1
		R- 3′GCATCCGATCTTGATGGG-5′			
3	RM250	F- 5′GGTTCAAACCAAGCTGATCA-3′	55	153	2
		R- 3′GATGAAGGCCTTCCACGCAG -5′			
4	RM55	F- 5′CGCAGTTGTGGATTTCAGTG-3′	55	226	3
		R- 3′TGCTCAACGTTTGACTGTCC-5′			
5	RM85	F- 5′CCAAAGATGAAACCTGGATTG-3′	55	107	3
		R- 3′GCACAAGGTGAGCAGTCC -5′			
6	RM130	F- 5′TGTTGCTTGCCCTCACGCGAAG-3′	55	85	3
		R- 3′GGTCGCGTGCTTGGTTTGGTTC-5′			
7	RM514	F- 5′AGATTGATCTCCCATTCCCC-3′	55	259	3
		R- 3′CACGAGCATATTACTAGTGG -5′			
8	RM241	F- 5′GAGCCAAATAAGATCGCTGA-3′	55	138	4
		R- 3′TGCAAGCAGCAGATTTAGTG -5′			
9	RM252	F- 5′TTCGCTGACGTGATAGGTTG-3′	55	216	4
		R- 3′ATGACTTGATCCCGAGAACG -5′			
10	RM13	F- 5′TCCAACATGGCAAGAGAGAG-3′	55	141	5
		R- 3′GGTGGCATTCGATTCCAG -5′			
11	RM412	F- 5′CACTTGAGAAAGTTAGTGCAGC-3′	55	198	6
		R-3″CCCAAACACACCCAAATAC -5′			
12	RM541	F- 5′TATAACCGACCTCAGTGCCC-3′	55	158	6
		R- 3′CCTTACTCCCATGCCATGAG 5″			
13	RM11	F- 5′TCTCCTCTTCCCCCGATC -3′	55	140	7
		R- 3′ATAGCGGGCGAGGCTTAG -5′			
14	RM125	F- 5′ATCAGCAGCCATGGCAGCGACC -3′	55	127	7
		R- 3′AGGGGATCATGTGCCGAAGGCC -5′			
15	RM214	F- 5′CTGATGATAGAAACCTCTTCTC -3′	55	112	7
		R- 3′AAGAACAGCTGACTTCACAA -5′			
16	RM25	F- 5′GGAAAGAATGATCTTTTCATGG -3′	55	146	8
		R- 3′CTACCATCAAAACCAATGTTC -5′			
17	RM153	F- 5′GCCTCGAGCATCATCATCAG -3′	55	201	5
		R- 3′ATCAACCTGCACTTGCCTGG -5′			
18	RM337	F- 5′GTAGGAAAGGAAGGGCAGAG -3′	55	192	8
		R- 3′CGATAGATAGCTAGATGTGGCC -5′			
19	RM433	F- 5′TGCGCTGAACTAAACACAGC -3′	55	224	8
		R- 3′AGACAAACCTGGCCATTCAC -5′			
20	RM205	F- 5′CTGGTTCTGTATGGGAGCAG -3′	55	122	9
		R- 3′CTGGCCCTTCACGTTTCAGTG -5′			

21	RM215	F- 5′CAAAATGGAGCAGCAAGAGC -3′ R- 3′TGAGCACCTCCTTCTCTGTAG -5′	55	148	9
22	RM206	F- 5′CCCATGCGTTTAACTATTCT -3′ R- 3′CGTTCCATCGATCCGTATGG -5′	55	147	11
23	RM19	F- 5′CAAAAACAGAGCAGATGAC -3′ R- 3′CTCAAGATGGACGCCAAGA -5′	55	226	12
24	RM235	F- 5′AGAAGCTAGGGCTAACGAAC -3′ R- 3′TCACCTGGTCAGCCTCTTTC -5′	55	124	12
25	OSR13	F- 5′CATTTGTGCGTCACGGAGTA -3′ R- 3′AGCCACAGCGCCCATCTCTC -5′	53	112	3
26	RM44	F- 5′ACGGGCAATCCGAACAACC -3′ R- 3′TCGGGAAAACCTACCCTACC -5′	53	99	8
27	RM171	F- 5′AACGCGAGGACACGTACTTAC -3′ R- 3′ACGAGATACGTACGCCTTTG -5′	55	328	10
28	RM259	F- 5′TGGAGTTTGAGAGGAGGG -3′ R- 3′CTTGTTGCATGGTGCCATGT -5′	55	162	1
29	RM271	F- 5′TCAGATCTACAATTCCATCC -3′ R- 3′TCGGTGAGACCTAGAGAGCC -5′	55	101	10
30	RM284	F- 5′ATCTCTGATACTCCATCCATCC -3′ R- 3′CCTGTACGTTGATCCGAAGC -5′	55	141	8
31	RM307	F- 5′GTACTACCGACCTACCGTTCAC -3′ R- 3′CTGCTATGCATGAACTGCTC -5′	55	174	4
32	RM312	F- 5′GTATGCATATTTGATAAGAG -3′ R- 3′AAGTCACCGAGTTTACCTTC -5′	55	97	1
33	RM316	F- 5′CTAGTTGGGCATACGATGGC -3′ R- 3′ACGCTTATATGTTACGTCAAC -5′	55	192	9
34	RM334	F-5 «GTTCAGTGTTCAGTGCCACC -3′ R- 3′GACTTTGATCTTTGGTGGACG -5′	55	182	5
35	RM408	F- 5′CAACGAGCTAACTTCCGTCC -3′ R- 3′ACTGCTACTTGGGTAGCTGACC -5′	55	128	8
36	RM413	F- 5′GGCGATTCTTGGATGAAGAG -3′ R- 3′TCCCCACCAATCTTGTCTTC -5′	53	79	5
37	RM447	F- 5′CCCTTGTGCTGTCTCCTCTC -3′ R- 3′ACGGGCTTCTTCTCCTTCTC -5′	55	111	8
38	RM489	F- 5′ACTTGAGACGATCGGACACC -3′ R- 3′TCACCCATGGATGTTGTCAG -5′	55	271	3
39	RM495	F- 5′AATCCAAGGTGCAGAGATGG -3′ R- 3′AACGATGACGAACACAACC -5′	55	159	1
40	RM507	F- 5′CTTAAGCTCCAGCCGAAATG -3′ R- 3′CTCACCCTCATCATCGCC -5′	55	258	5

41	RM510	F- 5′AACCGGATTAGTTTCTCGCC -3′ R- 3′TGAGGACGACGAGCAGATTC -5′	57	122	6
42	RM552	F- 5′CGCAGTTGTGGATTTCAGTG -3′ R- 3′TGCTCAACGTTTGACTGTCC -5′	55	195	11
43	RM144	F- 5′TGCCCTGGCGCAAATTTGATCC -3′ R- 3′GCTAGAGGAGATCAGATGGTAGTGCATG-5′	57	237	11
44	RM287	F- 5′TTCCCTGTTAAGAGAGAAATC -3′ R- 3′GTGTATTTGGTGAAAGCAAC -5′	55	118	11
45	RM152	F- 5′GAAACCACCACACCTCACCG -3′ R- 3′CCGTAGACCTTCTTGAAGTAG -5′	53	151	8
46	RM161	F- 5′TGCAGATGAGAAGCGGCGCCTC -3′ R- 3′TGTGTCATCAGACGGCGCTCCG -5′	61	187	5
47	RM237	F- 5′CAAATCCCGACTGCTGTCC -3′ R- 3′TGGGAAGAGAGCACTACAGC -5′	55	130	1
48	RM283	F- 5′GTCTACATGTACCCTTGTTGGG -3′ R- 3′CGGCATGAGAGTCTGTGATG -5′	61	151	1
49	RM162	F- 5′GCCAGCAAAACCAGGGATCCGG -3′ R- 3′CAAGGTCTTGTGCGGCTTGCGG -5′	61	229	6
50	RM244	F- 5′CCGACTGTTCGTCCTTATCA -3′ R- 3′CTGCTCTCGGGTGAACGT -5′	55	163	10
51	RM5686	F- 5′CTCTTCTATGCATATTGCCA -3′ R- 3′ATAAACTGAGGGGCGATATA -5′	50	149	3
52	RM474	F- 5′AAGATGTACGGGTGGCATTC -3′ R- 3′TATGAGCTGGTGAGCAATGG -5′	55	252	10
53	RM249	F- 5′GGCGTAAAGGTTTTGCATGT -3′ R- 3′ATGATGCCATGAAGGTCAGC -5′	55	121	5
54	RM547	F- 5′TAGGTTGGCAGACCTTTTCG -3′ R- 3′GTCAAGATCATCCTCGTAGCG -5′	55	235	8
55	RM280	F- 5′ACACGATCCACTTTGCGC -3′ R- 3′TGTGTCTTGAGCAGCCAGG -5′	55	155	4
56	RM276	F- 5′CTCAACGTTGACACCTCGTG -3′ R- 3′TCCTCCATCGAGCAGTATCA -5′	55	149	6
57	RM463	F- 5′TTCCCCTCCTTTTATGGTGC-3′ R- 3′TGTTCTCCTCAGTCACTGCG-5′	55	192	12
58	RM118	F- 5′CCAATCGGAGCCACCGGAGAGC-3′ R- 3′ CACATCCTCCAGCGACGCCGAG-5′	67	156	7
59	RM124	F- 5′ CTCGATCCCCTAGCTCTC-3′ R- 3′TCACCTCGTTCTCGATCC-5′	55	162	4
60	RM127	F- 5′ GTGGGATAGCTGCGTCGCGTCG-3′ R- 3′AGGCCAGGGTGTTGGCATGCTG-5′	55	223	4

61	RM133	F- 5′TTGGATTGTTTTGCTGGCTCGC-3′ R- 3′GGAACACGGGGTCGGAAGCGAC-5′	61	230	6
62	RM134	F- 5′ACAAGGCCGCGAGAGGATTCCG-3′ R- 3′GCTCTCCGGTGGCTCCGATTGG-5′	55	93	7
63	RM154	F- 5′ACCCTCTCCGCCTCGCCTCCTC-3′ R- 3′CTCCTCCTCCTGCGACCGCTCC-5′	61	183	2
64	RM178	F- 5′TCGCGTGAAAGATAAGCGGCGC-3′ R- 3′GATCACCGTTCCCTCCGCCTGC-5′	67	117	5
65	RM277	F- 5′CGGTCAAATCATCACCTGAC-3′ R- 3′CAAGGCTTGCAAGGGAAG-5′	55	124	12
66	RM321	F- 5′CCAACACTGCCACTCTGTTC-3′ R- 3′GAGGATGGACACCTTGATCG-5′	55	200	9
67	RM332	F- 5′GCGAAGGCGAAGGTGAAG-3′ R- 3′CATGAGTGATCTCACTCACCC-5′	55	183	11
68	RM338	F- 5′CACAGGAGCAGGAGAAGAGC-3′ R- 3′GGCAAACCGATCACTCAGTC-5′	55	183	3
69	RM431	F- 5′TCCTGCGAACTGAAGAGTTG-3′ R- 3′AGAGCAAAACCCTGGTTCAC-5′	55	251	1
70	RM452	F- 5′CTGATCGAGAGCGTTAAGGG-3′ R- 3′GGGATCAAACCACGTTTCTG-5′	55	209	2
71	RM454	F- 5′CTCAAGCTTAGCTGCTGCTG-3′ R- 3′GTGATCAGTGCACCATAGCG-5′	54	268	6
72	RM455	F- 5′AACAACCCACCACCTGTCTC-3′ R- 3′AGAAGGAAAAGGGCTCGATC-5′	55	131	7
73	RM484	F- 5′TCTCCCTCCTCACCATTGTC-3′ R- 3′TGCTGCCCTCTCTCTCTCTC-5′	55	299	10
74	RM536	F- 5′TCTCTCCTCTTGTTTGGCTC-3′ R- 3′ACACACCAACACGACCACAC-5′	50	243	11