

Research Article

†The online version of this article has been updated since original publication. A notice detailing the changes has also been published at: <https://doi.org/10.1017/S1479262121000502>

Cite this article: Chettri D, Anandan A, Nagireddy RK, Sabarinathan S, Sathyanarayana N (2021). Genetic diversity in rice (*Oryza sativa* L.) landraces of Sikkim-Himalaya and early insight into their use in genome-wide association analysis. *Plant Genetic Resources: Characterization and Utilization* **19**, 347–356. <https://doi.org/10.1017/S1479262121000411>

Received: 16 December 2020

Revised: 9 August 2021

Accepted: 10 August 2021

First published online: 6 September 2021

Key words:

Association mapping; genetic diversity; landraces; population structure; rice; Sikkim-Himalaya

Author for correspondence:

N. Sathyanarayana,

E-mail: nsathyanarayana@cus.ac.in

Genetic diversity in rice (*Oryza sativa* L.) landraces of Sikkim-Himalaya and early insight into their use in genome-wide association analysis

Deepak Chettri¹, A. Anandan², Ranjit Kumar Nagireddy², S. Sabarinathan² and N. Sathyanarayana¹ 

¹Molecular Biology and Biotechnology Laboratory, Department of Botany, Sikkim University, Gangtok, East Sikkim 737102, India and ²Crop Improvement Division, ICAR-National Rice Research Institute (NRI), Bidhyadharpur, Cuttack, Odisha 753006, India

Abstract

The Sikkim-Himalaya represents one of the unique reservoirs of rice genetic resources in India owing to the presence of a large number of landraces adapted to extreme climatic and edaphic conditions. This valuable gene pool is now under threat due to the introduction of high-yielding varieties, wanting suitable genetic intervention for enhancing their productivity, and thus economic viability. Development of lodging resistant, high-yielding varieties tolerant to soil acidity through association mapping aided marker-assisted breeding programmes can help achieve this in a fast and efficient manner. But this requires information on genetic diversity, population structure, etc., of the germplasm collection, which is strikingly lacking. We, therefore, characterized a set of 53 rice landraces from Sikkim to address the above issues. The results revealed moderate diversity, poor divergence and high gene flow in our germplasm collection attesting its utility as an association mapping panel. Further, a total of 115 putative marker-trait associations ($P < 0.05$, $R^2 \geq 10\%$) were obtained using the general linear model and mixed linear models of which 25 were identical in both. Some of the associated markers were positioned in those regions where Qualitative Trait Loci have been previously identified in rice, providing credence to our results. The resources generated from this study will benefit the rice breeders from this region and elsewhere for targeting the yield and related traits, in addition to conservation efforts by the interested researchers.

Introduction

India is home to a large number of rice landraces (Pachauri *et al.*, 2013) with all the major rice-growing states having their preferred assemblage (Singh and Singh, 2003). Adaptation to wide-ranging edaphic and climatic factors has shaped >120,000 of them with extensive phenotypic variations (Khus, 1997). Rice accessions from northeastern (NE) states of India are regarded as the storehouse of agronomically important genes (Choudhury *et al.*, 2013; Das *et al.*, 2013) as they have endured extreme conditions of this mountainous terrain. It is estimated that about 10,000 such cultivars/landraces of ecological and cultural significance are preserved here (Hore, 2005; Das *et al.*, 2013) wanting in-depth evaluation on their offerings.

Sikkim has considerable antiquity in rice cultivation (Rahman and Karuppaiyan, 2011). Rice is grown here at an elevation ranging from foothills to 1700 m above sea level, and from sub-tropical to temperate climate. It is the second-largest crop after maize occupying 15% (10,327 ha) of the area under agriculture (Portel, 2015). As the state receives plenty of rainfall (3250 mm/annum), well distributed over 6 months from May to October, rice farming has emerged as the major source of livelihood for the majority of the indigenous communities inhabiting this pristine land. Landraces, which are grown in a traditional way in the hill slopes with and without terracing, dominate the rice production here with a 55–60% share in acreage (Rahman and Karuppaiyan, 2011). There are 70 plus such landraces reported so far (Sharma *et al.*, 2016) which await detailed characterization. However, barring a few scattered works (Rahman and Karuppaiyan, 2011; Choudhury *et al.*, 2013; Das *et al.*, 2013; Sharma *et al.*, 2016), not much has been done in this direction.

The majority of these landraces suffer from poor yield, lodging and late maturity although they boast other superior traits such as hardiness, pest tolerance/resistance, low input requirements and weed coexistence (Sharma *et al.*, 2015). In addition, the soils in Sikkim are highly acidic, leached and generally poor in fertility and water holding capacity (Das and Avasthe, 2016). Even though high-yielding varieties (HYV) are available (<https://icar-nrri.in/released-varieties>), their introduction in place of existing landraces/cultivars is perceived as detrimental



to the native gene pool owing to evident genetic erosion (Rathi *et al.*, 2018). Moreover, farmers still prefer landraces over HYVs due to inherent cultural preferences, and low input costs (Kapoor *et al.*, 2017). Thus, developing high-yielding, lodging resistant varieties tolerant to acidic soils by upgrading the prevailing landraces through marker-assisted backcross breeding (MAB) is the need of the hour to simultaneously achieve productivity and acceptability for the rice production in the Sikkim. More importantly, this will safeguard the local gene pool and their cropping system, the heritage which has survived for generations under the custodianship of the native people.

DNA markers have been the choice tool for the genetic characterization of rice genotypes (Karkousis *et al.*, 2003). Among them, simple sequence repeats (SSRs) are preferred due to polymorphic nature, abundance in the eukaryotic genome and ease of generating (Tautz, 1989; Morgante and Olivieri, 1993). In rice, SSRs have been used earlier for the assessment of genetic diversity (Rahman *et al.*, 2012; Choudhury *et al.*, 2013; Singh *et al.*, 2016; Suvi *et al.*, 2019), conservation planning (Sharma *et al.*, 2007), marker-assisted selection (Perez-Sackett *et al.*, 2011; Rani and Adilakshmi, 2011), cultivar identification and hybrid purity analysis, besides gene mapping (Weising *et al.*, 1997; Altaf-Khan *et al.*, 2006; Sarao *et al.*, 2010).

Further, any MAB initiative entails the identification of candidate genes/Qualitative Trait Loci (QTLs) for the traits of interest. Typically, linkage mapping and/or association analysis is used for this purpose. The major shortcoming in the linkage analysis is that only two alleles at any given locus can be considered in bi-parental crosses which provides a low mapping resolution (Flint-Garcia *et al.*, 2003). Alternatively, association mapping categorizes QTLs by exploiting the natural diversity and identifies important genes (Zhu *et al.*, 2008). It has been successfully used in several crop species (Abdurakhmonov and Abdugarimov, 2008) including rice (Famoso *et al.*, 2011; Huang *et al.*, 2012; Zhou *et al.*, 2012; Han and Huang, 2013). The present study was therefore conceived to (a) assess the genetic relationship and population structure among the 53 rice landraces and (b) to test their efficacy for the association analysis and advance early insights on the loci controlling the key agronomic (7) and nutritional quality traits (14).

Materials and methods

Plant materials

The rice landraces for the present study were sourced from different agro-climatic regions in Sikkim through collection trips during the years 2017–19. The information on varietal preference, landrace name, traditional uses, etc., was acquired through structured interviews. Prior informed consent was obtained for the accessions collected from the farmers. The collection strategy involved capturing maximum variability through minimum representative accessions which were arrived from the data of the structured interview. Based on the preliminary screening, a total of 68 landraces were shortlisted. After eliminating the redundant collections, 53 were retained for the phenotypic and genetic characterization (online Supplementary Table S1). The plants were raised during the Kharif season of 2019 at the experimental farm situated at Nandok, East Sikkim, Sikkim state (27°17'36" N and 88°36'18" E) at an elevation of 972 m above sea level. Initially, seeds were sown in a raised bed nursery and, after 30 days, seedlings were transplanted in the field. The experiment

was laid out in a randomized block design with three replications. Each entry was transplanted in three rows with 20 cm spacing between row to row, and 14 cm between plant to plant. The standard agronomic practice was followed for crop growth.

Phenotyping

Twenty-one quantitative and qualitative traits of agronomic (7), quality (9) and nutritional (5) significance were evaluated (online Supplementary Table S2). For the seven agronomic traits such as days to 50% flowering, flag leaf length (cm), flag leaf width (cm), stem thickness (cm), plant height (cm), panicles per plant and number of effective tillers per plant, observations were recorded on five randomly selected plants as per the method prescribed by the National test guidelines for DUS (PPV & FR Act, 2007). For the measurement of quality traits such as grain length, grain width, grain thickness, kernel width and kernel length, 10 grains were chosen randomly and measured in millimetres (mm) using a digital Vernier caliper following the National test guidelines for DUS (PPV & FR Act, 2007). In the case of cooked kernels, 5 g of the dehusked samples were soaked in 15 ml water for 10 min and cooked for 20 min. After cooling, the grains were measured using a digital Vernier caliper. For 100-grain weight, three replicates of 10 whole grains, dried to 13% moisture content, were chosen randomly and weighed in grams (g), and then the average was multiplied by 10. Proximate analysis of crude protein, carbohydrate, crude fat and crude fibre was carried out following the protocols of the Association of Analytical Chemists (AOAC, 2004) and the American Association of Cereal Chemists (AACC, 2000), respectively. The amylose value was determined using the standard protocol (Juliano, 1979).

Genotyping

Genomic DNA was isolated from fresh young leaves (14–20 days) following the modified cetyl-trimethylammonium bromide method (Doyle and Doyle, 1987). For genotyping, 45 SSR primer pairs (online Supplementary Table S4) used previously for the association analysis (Islam *et al.*, 2018b) were downloaded from the Gramene's marker database (<http://www.gramene.org>) and primers were synthesized. These markers offered coverage on all the 12 chromosomes of the rice genome. DNA amplification was carried out in a BIORAD T100 thermal cycler (USA) with a reaction mixture containing 1 µl of genomic DNA (50 ng), 0.5 µl of each primer (10 pmol/µl), 6 µl of PCR master mix (Nucleo Spin Takara) and 2 µl of PCR grade water. PCR amplification was performed with an initial denaturation at 94 °C for 4 min followed by 35 cycles of denaturation, annealing and extension at 94 °C for 1 min, 55 °C for 1 min and 72 °C for 1 min with a final extension at 72 °C for 5 min. The PCR products were resolved on a 3% agarose gel using 1× Tris-boric acid-EDTA (TBE) buffer. After staining with the ethidium bromide, bands were visualized using the Gel Documentation system (BIORAD-Chemi Doc XRS+, USA). All PCR reactions were repeated at least twice before the data analysis.

Allele scoring

Post amplification, a cluster of two to five distinct bands was visible in the stained gel for the majority of the markers. The size of the most intensely amplified band of each microsatellite marker was determined by comparing it with the standard 50 bp DNA

ladder. For individual markers, the presence of an allele was recorded as '1' and the absence as '0'. The representative gel images for SSR marker RM 314 (a and a1) and RM 253 (c and c1) are provided in online Supplementary Fig. S1.

Data analysis

Power Marker V.3.25 (Liu and Muse, 2005) was used to calculate the major allele frequency (M_{AF}) and the polymorphism information content (PIC). The marker characteristics such as observed heterozygosity (H_o), expected heterozygosity (H_e), total number of alleles (N_a), effective number of alleles (N_e) and percentage of polymorphic loci (%P) were computed using GenAEx V.6.5 (Peakall and Smouse, 2012). The genetic diversity indices such as total genetic diversity (H_t), genetic diversity within populations (H_s), genetic differentiation (G_{st}), Shannon's information index (I) and Nei's gene diversity (h) were determined using POPGENE V.1.32 (Yeh *et al.*, 1999). The distance-based clustering was performed with an Un-Weighted Pair Group Method with Arithmetic mean (UPGMA) using FreeTree software (Pavlicek *et al.*, 1999). The Principal Coordinate Analysis (PCoA) plot was constructed using DARwin V.6.0 (Perrier and Jacquemoud-Collet, 2006). The robustness of the UPGMA dendrogram was tested with a bootstrap analysis using 1000 iterations.

Population structure

Using STRUCTURE V. 2.3.4 software (Pritchard *et al.*, 2000), the Bayesian model was implemented to ascertain the optimal genetic clusters within the 53 rice accessions. The admixtures model and correlated allele frequencies were applied for each run with 100,000 burn-in periods and 200,000 Markov Chain Monte Carlo (MCMC) simulations. The optimum K value was determined from 10 replicate runs for each value of K (Evanno *et al.*, 2005). The ΔK was based on the change in the log probability [$\ln P(D)$] of the data between successive K values. To obtain a clear peak at the ΔK , the output from the structure analysis was loaded onto STRUCTURE HARVESTER V. 6.0 (Earl and Von Holdt, 2012). Analyses of molecular variance (AMOVA) and pairwise F_{ST} were carried out by GenAEx V. 6.5.

Marker-trait association

Software TASSEL V. 5.0 (Bradbury *et al.*, 2007) was used to perform the marker-trait association (MTA) analysis. Both the models – the general linear model (GLM) and the mixed linear model (MLM) – were tested. The major limitation of the GLM is that it does not incorporate the kinship matrix (K matrix) but rather uses population structure (Q matrix) for finding a MTA which sometimes results in false positives. This limitation is overcome by the MLM which incorporates both Q and K matrix reducing spurious association (Yu *et al.*, 2006). The significance of MTA was tested at $P < 0.05$ and the correlation coefficient ($R^2 \geq 10\%$) was used to determine the phenotypic variation explained by each MTA.

Results

Marker characteristics

Totally 45 SSR primer pairs were used to genotype 53 rice landraces of which 42 (93.33%) were polymorphic, two (RM 25 and RM 455) were monomorphic, and one (RM 170) failed to amplify.

These 42 polymorphic SSR primers pairs were used in all the subsequent analyses. A total of 227 alleles were detected which ranged from 2 (in many) to 11 (RM 20) with an average of 5.262 alleles per locus. The major allele frequency (M_{AF}) was high with a mean value of 0.679 and ranged from 0.460 to 0.930. The number of different alleles (N_a) per locus varied from 1 to 9 with a mean value of 4.721, while the number of effective alleles (N_e) per locus varied from 1.210 to 6.384 with a mean of 3.344. The fact that $N_a > N_e$ indicates that only a handful of alleles contributed to the overall diversity. The expected heterozygosity (H_e) ranged from 0.175 (RM 322) to 0.842 (RM 284) with an average of 0.632. The observed heterozygosity (H_o) values had a mean value of 0.222, and ranged from 0.00 (RM 237, RM 513, RM 178, RM 277, RM 271, RM 44) to 0.98 (RM 514). The value of PIC, which measures the extent of polymorphism revealed by the markers based on the number of alleles per locus, ranged from 0.121 (RM 322) to 0.375 (RM 44) with an average of 0.323. While the major portion of the microsatellite markers (34) was moderately polymorphic with PIC values ranging between 0.25 and 0.50, eight markers generated low polymorphism with $PIC < 0.25$. The details of the marker attributes are provided in online Supplementary Table S3.

Genetic diversity

The mean values of Nei's gene diversity (h) and Shannon's information index (I) in our landrace collection were 0.217 and 0.354 respectively signifying moderate genetic diversity. The mean value of total gene diversity ($H_t = 0.322$) further supported this. Two populations were assumed based on the presence/absence of aroma trait of which the first population (population-1) comprised of 16 aromatic accessions, while the second population (population-2) represented 37 non-aromatic accessions. In population-1, about 97.67% of SSR markers were polymorphic, while 96.51% were polymorphic in population-2. Among the two groups, the aromatic group with $I = 0.365$ and $h = 0.223$ was slightly more diverse *vis-a-vis* a non-aromatic group with $I = 0.343$ and $h = 0.209$. The summary of genetic diversity is presented in Table 1.

The UPGMA dendrogram resolved 53 landraces, with significant bootstrap support, into two major clusters (Fig. 1) – CL-I and CL-II – with a single outlier accession. Of these, CL-I included 48 accessions belonging to both aromatic and non-aromatic groups and was further divided into two smaller groups – sub-cluster-IA (SCL-IA) and sub-cluster-IB (SCL-IB). Interestingly, all the aromatic accessions such as Raja bara, Ram jeera, Ram bhog, Rudhua, Shyam jeera, Krishna bhog, Kalo dhan, Kalo nunia-I, Kalo nunia-II, Kaanchi, Dharnali and Birinful grouped within SCL-IB with a few non-aromatic rice accessions, while SCL-IA was entirely composed of non-aromatic rice accessions (Table 2). The CL-II encompassed only four non-aromatic accessions namely Taprey, Thamba, Thulo tulashi and Tulashi. Doodh kalami was an outlier in UPGMA cluster. Overall, the grouping did not fully correspond to either the geographical origin or aroma trait except for the noticeable influence of aroma in the grouping of SCL-IB accessions. A similar dispersion was observed in the PCoA (Fig. 2) plot where the first two axes accounted for 14.82 and 8.86% of the variation, respectively.

Population structure

The population structure analysis using STRUCTURE revealed $K = 3$ as the highest log-likelihood value signifying three genetic

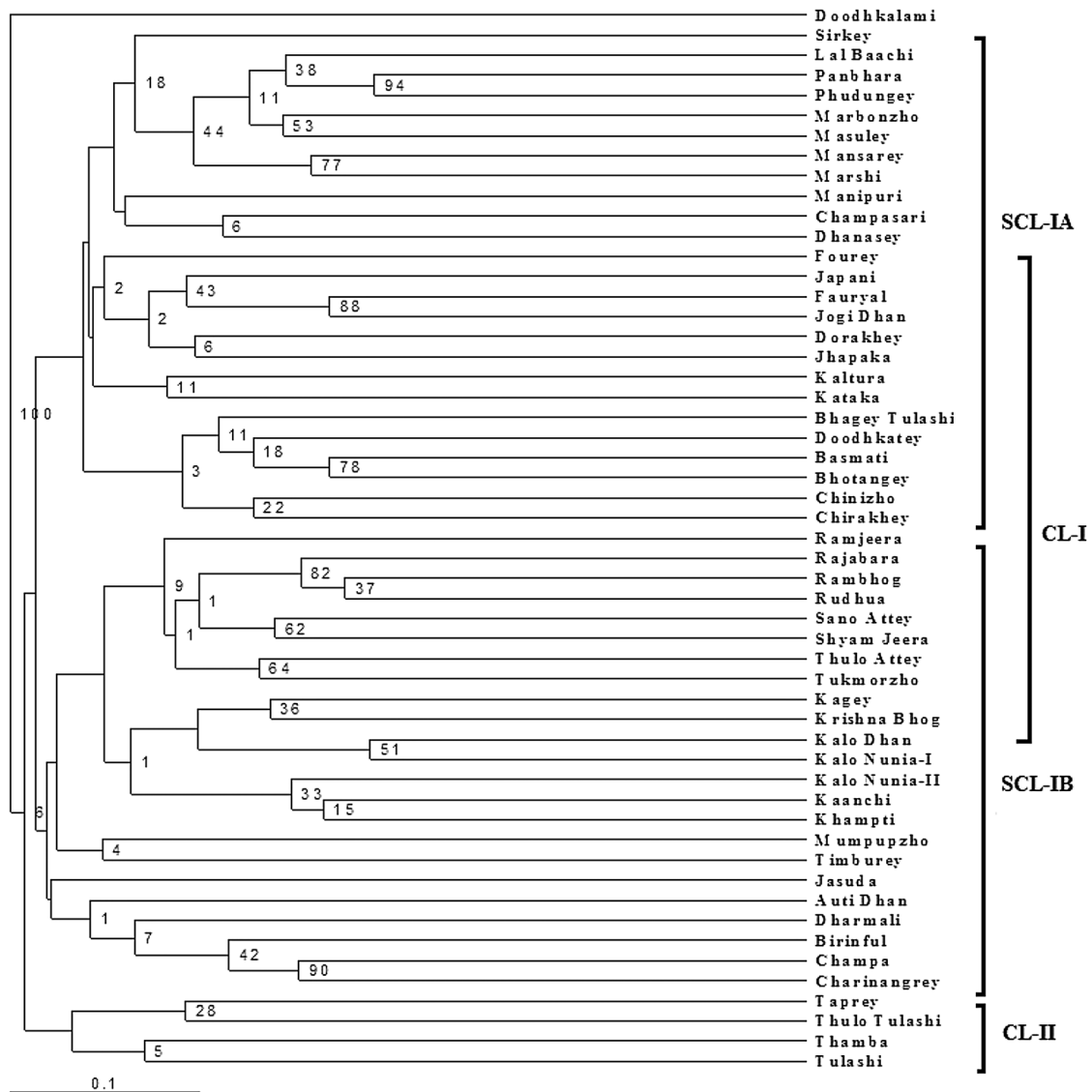


Fig. 1. UPGMA dendrogram depicting the relationship among 53 rice landraces of Sikkim. The number at the node of each branch indicates the bootstrap value.

Table 1. Summary of genetic diversity statistics of 53 rice landraces of Sikkim

	<i>N_a</i>	<i>N_e</i>	<i>h</i>	<i>I</i>	<i>H_o</i>	<i>H_e</i>	% <i>P</i>
Population 1	5.047	3.473	0.223	0.365	0.209	0.648	97.67%
Population 2	4.395	3.215	0.209	0.343	0.225	0.616	96.51%
Mean	4.721	3.344	0.217	0.354	0.217	0.632	97.09%
S.E.	0.326	0.129	0.137	0.186	0.008	0.016	0.580

N_a, total number of alleles per locus; *N_e*, number of effective alleles per locus; *I*, Shannon's information index; *H_o*, observed heterozygosity; *H_e*, expected heterozygosity; SE, standard error.

groups within our landrace collection (Fig. 3). These were named RC-I, RC-II and RC-III, respectively. Based on the *Q*-value (membership proportions), accessions were further designated as pure (*Q*-value $\geq 80\%$) and admixtures (*Q*-value $< 80\%$) which revealed 39 pure and 14 admixtures in our collection. Out of the 39 pure accessions, 19 were grouped with RC-I with five aromatic accessions, 13 aligned in RC-II with nine aromatic accessions, while

RC-III included seven non-aromatic accessions only. The analysis of molecular variance (AMOVA) showed 99% of the molecular variance was contributed by within-population differences, while only 1% was due to variation between the populations. The results of AMOVA displayed highly significant genetic differences ($P < 0.001$) among the individuals and within the individuals.

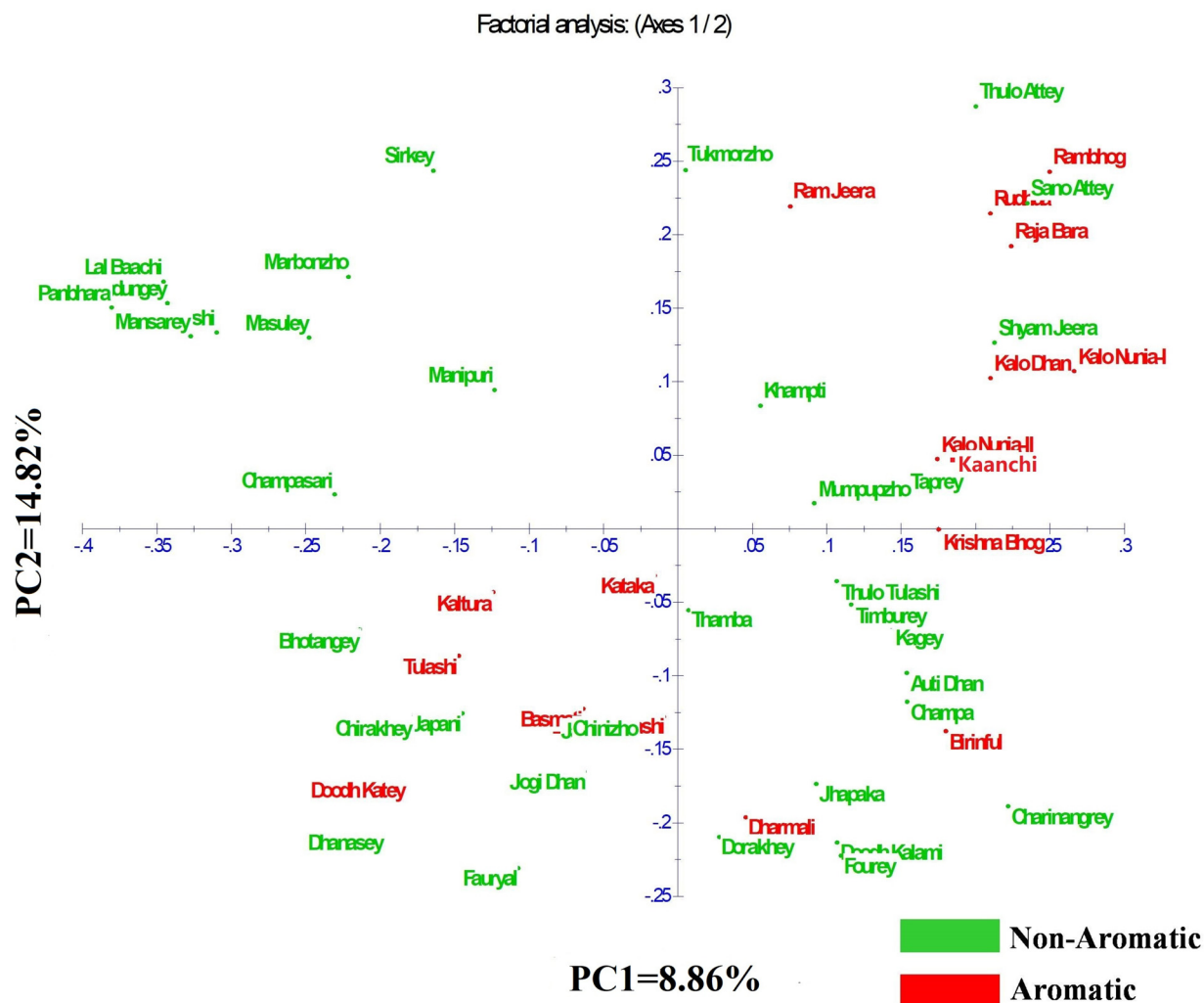


Fig. 2. Principal Coordinate Analysis (PCoA) of tested germplasm spread over coordinates.

Table 2. Grouping of accessions in UPGMA clusters

Cluster	Sub-cluster	No. of accessions	Name of the landraces
I	IA	25	Sirkey, Lalbaachi, Panbhara Phudungey, Marbonzho, Masuley, Mansarey, Marshi, Manipuri, Chamapasari, Dhanasey, Fourey, Fauryal, Japani, Jogi dhan, Dorakhey, Jhapaka, Kultura, Kataka, Bhagey tulashi, Doodh katey, Basmati, Bhotangey, Chinizho, Chirakhey
	IB	23	<i>Aromatic:</i> Raja bara, Ram jeera, Ram bhog, Rudhua, Shyam jeera, Krishna bhog, Kalo dhan, Kalo nunia-I and II, Kaanchi, Dharmali, and Birinful <i>Non-aromatic:</i> Sano attey, Thulo attey, Tukmorzho, Kagey, Timburey, Jasuda, Anti dhan, Dharmali, Champa, Khampati, Mumpupzho, Charinanagrey
II		4	Taprey, Thulo tulashi, Thamba, Tulashi

Marker-trait association

The MTA analysis for the two trait groups – biochemical and yield-related – was carried out using 42 polymorphic SSR markers across 53 rice accessions. Counting both GLM (62) and MLM (53), a total of 115 significant MTAs were observed ($P < 0.05$) for 21 traits which were contributed by 71% (30) of the SSR markers. In total, 25 MTAs for 17 traits were identical in both the models (Table 3), and the remaining five MTAs were detected in either of them.

In the case of GLM, a total of 62 significant MTAs ($P < 0.05$, $R^2 \geq 10\%$) were found which ranged from a minimum of 1 to a maximum of 7. The minimum number of MTAs was detected for 100-grain weight, kernel width after cooking, crude fat and amylose content while the maximum number of MTAs was reported for grain thickness. The correlation value (R^2) for GLM ranged from 10 to 26%. The highest R^2 value (26%) was between carbohydrate content and marker RM 528. In the case of MLM, 53 significant ($P < 0.05$, $R^2 \geq 10\%$) MTAs were detected.

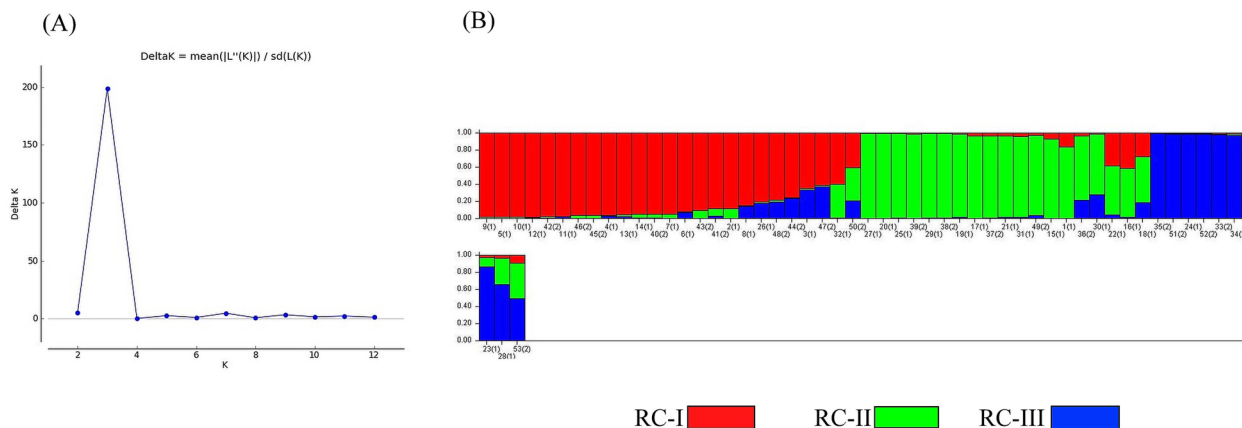


Fig. 3. Model-based population structure analysis of 53 rice landraces. (A) ΔK showing the presence of three hypothetical subpopulations. (B) Population structure for 53 rice landraces of Sikkim sorted by kinship (Q) matrix at $K=3$.

Table 3. Marker-trait associations obtained from the general linear model (GLM) and the mixed linear model (MLM)

Trait	Marker	Chr. No.	GLM		MLM	
			P-value	R ²	P-value	R ²
Grain length	RM 5	1	0.001	0.19	0.015	0.12
Grain width	RM 408	8	0.014	0.11	0.012	0.14
Grain thickness	OSR 13	3	0.023	0.11	0.018	0.11
	RM 551	4	0.020	0.11	0.024	0.10
	RM 10	7	0.008	0.14	0.013	0.13
100-grain weight	RM 489	3	0.006	0.15	0.029	0.11
Kernel breadth	RM 5	1	0.001	0.18	0.009	0.14
	RM 215	9	0.010	0.12	0.025	0.11
Kernel length	RM 489	3	0.009	0.14	0.004	0.22
	RM 44	8	0.005	0.20	0.016	0.16
Kernel thickness	RM 283	1	0.012	0.13	0.014	0.13
	RM 10	7	0.005	0.15	0.009	0.15
Kernel length after cooking	RM 452	2	0.001	0.17	0.017	0.12
Number of effective tillers	OSR 13	3	0.024	0.11	0.028	0.10
Number of panicles per plant	RM 551	4	0.015	0.11	0.021	0.11
	RM 447	8	0.002	0.17	0.024	0.10
	RM 515	8	0.001	0.18	0.022	0.11
Flag leaf length	RM 271	10	0.017	0.10	0.017	0.12
Flag leaf width	OSR 13	3	0.015	0.11	0.008	0.15
Plant height	RM 514	3	0.001	0.18	0.010	0.14
Protein	RM 6	2	0.010	0.12	0.006	0.16
Carbohydrate	RM 447	8	0.015	0.10	0.021	0.11
	RM 223	8	0.006	0.14	0.021	0.13
Crude fat	RM 118	7	0.020	0.10	0.021	0.11
Crude fibre	RM 408	8	0.009	0.13	0.008	0.16

The minimum number of MTA (1) was detected for grain width, kernel width after cooking, kernel length after cooking, number of effective tillers, stem thickness and amylose value. The maximum number of MTAs (4) were detected for 100-grain weight, number

of panicles per plant, flag leaf width, carbohydrate content and crude fibre content. The correlation value (R^2) for MLM ranged from a minimum of 10% in several MTAs to a maximum of 22%. The highest R^2 value (22%) was found between the kernel

length and marker RM 489. In addition, the following seven markers were associated with two or more than two traits: RM 5 (associated with grain width and kernel width); RM 408 (associated with grain length and crude fibre); OSR 13 (associated with grain thickness and effective tillers per plant); RM 551 (associated with grain thickness and number of panicle per plant); RM 10 (associated with grain thickness and kernel thickness); RM 489 (associated with 100-grain weight and kernel length); RM 447 (associated with the number of panicles per plant and carbohydrate content). Further, the markers RM 447, RM 6, RM 408 and RM 118 were found to be associated with carbohydrate, crude protein, crude fibre and crude fat contents. Similarly, markers OSR13, RM 551 and RM 10 were associated with grain thickness, and RM 489, RM 215 and RM 44 were associated with kernel length only. Likewise, RM 551, RM 447 and RM 515 were associated with the number of panicles per plant.

Discussion

Rice has great antiquity in the Sikkim-Himalayan region as evident from the etymological significance of its old name 'Denzong' meaning 'the valley of rice'. Rice farming has been an inspiring enterprise here as it has evolved in response to the dynamic needs of a multi-ethnic society of diverse food habits, limited landholding; inhabiting difficult topographies of a harsh climate. Their enduring creative labour, for centuries, has led to the selection of an array of landraces, which occupy a significant niche in the local rice production system (Rahman and Karuppaiyan, 2011). However, of late, like other Himalayan states, rice production in Sikkim is witnessing a steady shift from the traditional to modern agro-ecosystem. Consequently, landraces are being slowly faced out by the high-yielding varieties (FS&ADD, 2018–19). It is thus cautioned that – in the absence of appropriate scientific intervention – indigenous rice landraces and their farming technologies in the region might become history with time (Kambewa *et al.*, 1997; Zuberi, 1997).

The majority of these landraces suffer from low yield and lodging (Sharma *et al.*, 2016). Besides, high soil acidity is adversely impacting their production potential (Das and Avasthe, 2016). Identifying favourable markers/alleles for yield, tolerance to high soil acidity and lodging resistance via association mapping can provide fast and efficient means to strategize marker-assisted breeding programmes to improve these traits. Before that, there is a need to adequately characterize the local germplasm and generate information related to diversity and population structure. Hence, the main aim of the present work was to determine the genetic diversity and relationship among the rice landraces of Sikkim. In addition, we have explored the prospects of association mapping through an extended experiment.

In our genetic diversity estimation, the number of alleles ranged from 2 to 11 with a mean value of 5.27 per locus. This suggests superior allelic diversity in our germplasm when compared with aromatic rice varieties of Assam (3.7; Talukdar *et al.*, 2017), rice varieties released in India during 1940–2013 (3.11; Singh *et al.*, 2016) as well as elite rice varieties of Bangladesh (4.18; Rahman *et al.*, 2012) and Malaysia (4.09; Aljumaili *et al.*, 2018); but lower as compared with eastern Himalayan landraces (Choudhury *et al.*, 2013; 7.9 – Das *et al.*, 2013; 8.49 – Roy *et al.*, 2016). The variability in the number of alleles/locus observed in pan-Himalayan studies, including ours, reflects the diverse nature of the rice germplasm of this region. Besides, the number of effective alleles per locus ranged from 1.21 to 6.84,

with a mean of 3.34 – the value close to (3.77) that reported by Chen *et al.* (2017) indicating a good number of alleles contributed to the genetic diversity. The major allele frequency across all the 42 SSR primers ranged from 0.46 (RM 277) to 0.93 (RM 322) with an average of 0.679 (online Supplementary Table S3) which is higher than the previous studies on Indian (0.53; Upadhyay *et al.*, 2012) and Korean landraces (0.5; Li *et al.*, 2014). The greater number of alleles indicates the usefulness of the selected SSR markers for genetic studies.

The PIC of a marker is a good measure of their usefulness for the linkage analysis (Elston, 2014). Often, PIC values are the reflection of allelic diversity among the varieties (Meti *et al.*, 2013). In our study, a moderate PIC value with a mean of 0.32 (range 0.121–0.375) was observed for 53 accessions. A similar PIC value (0.37) was reported earlier by Choudhury *et al.* (2013) in rice landraces of NE states of India. Saha *et al.* (2013) and Pachauri *et al.* (2013) also found the mean PIC values of 0.37 and 0.38 in their study on different Indian rice varieties. However, higher PIC values up to 0.74 and 0.65 have also been reported in some other studies from the eastern Himalayan region (Das *et al.*, 2013; Roy *et al.*, 2016). This may be due to specific variability held by these landraces. Further, as PIC and inbreeding coefficient (FIS) are also the functions of how heterozygosity is partitioned within and among the accessions based on differences in allele frequencies (Muluaem *et al.*, 2018), it can be inferred that a large majority of the loci targeted by our markers are homozygous.

Our landraces revealed a gene diversity estimate of 0.217 which is much lower than 0.51 reported by Umakanth *et al.* (2017) for rice landraces of northeast India. The mean observed heterozygosity was 0.217 which is analogous to other studies on Indian rice varieties (Choudhury *et al.*, 2014; Nachimuthu *et al.*, 2015). Shannon's information index (*I*) ranged from 0.343 to 0.365 with an average of 0.354. This is lower than the findings of Aljumaili *et al.* (2018) and Suvi *et al.* (2019) who reported an index of 0.88 and 0.82, respectively. The moderate value of Shannon's information index is another indication of the moderate genetic diversity among the on-hand germplasm. Further, the inbreeding coefficient (FIS) represents the average deviation of the population's genotypic proportions from Hardy–Weinberg equilibrium for a locus. The FIS values in our study revealed that only three of the 42 markers (RM 215, OSR13 and RM 514) showed higher heterozygotes (–0.081, –0.085 and –0.27).

In our UPGMA analysis, landraces from the same geographic locations, or sharing similar traits such as aroma, did not correspond to specific groups indicating the nominal influence of these attributes on the genetic relationship. Similar results have been observed for 67 hill rice accessions of NE India (Roy *et al.*, 2016) as well as 48 traditional aromatic rice landraces from eastern India (Meti *et al.*, 2013). According to Muluaem *et al.* (2018), this may be due to their common descent as they are obtained from the nearby areas. Conversely, high gene flow aided by the cross-pollination might have also caused this (Musyoki *et al.*, 2015). However, the accretion of a significant number of aromatic accessions in a particular sub-group of UPGMA calls for further investigation on the role of aroma in shaping the phylogenetic relationships in rice landraces.

Population structure analysis using STRUCTURE revealed the highest ΔK value at $K = 3$ indicating 53 rice accessions represented three sub-populations (Fig. 3). They revealed 39 pure and 14 admixtures. The AMOVA showed significant genetic differences ($P \leq 0.001$) among and within individuals and lower variation

among the population groups. Of the total genetic variation, 70% was due to variation among the individuals. A similar trend was reported even in earlier studies (Anandan *et al.*, 2016; Roy *et al.*, 2016; Islam *et al.*, 2018a). The presence of such variability provides for generating wide crosses and creating the desirable heterotic group in the base breeding population (Alam *et al.*, 2015). On the other hand, the contribution of variation among the population groups was 1% indicating only a small collection within a given source captured the genetic diversity present in the test accessions. Further, the low mean value of F_{ST} (0.011) and G_{ST} (0.028) signify poor genetic differentiation (Wright, 1978) and the allele frequencies between population groups are likely to be similar. This was further supported by a very high estimate of gene flow (Nm : 3.96). Put together, these results point to a close evolutionary lineage of the rice accessions of Sikkim-Himalayas or extensive out-crossing between the accessions (Nuijten *et al.*, 2009). It is plausible that the exchange of rice accessions between the farmers may have also contributed to gene flow across different landraces.

In our association analysis, the majority of the associations detected in the MLM were also supported by the GLM approach with 25 MTAs for 17 traits being common in both (Table 3). In earlier works on the association analysis, Zhou *et al.* (2012) reported a total of 16 MTAs ($P < 0.01$) in the japonica rice varieties collected from China. Similarly, Anandan *et al.* (2016) reported 16 MTAs in 629 rice accessions using 39 SSR markers for the identification of genes associated with the early seedling vigour. Furthermore, Patil *et al.* (2014) found 13 significant SSR markers associated with the 18 agronomic traits in 58 rice accessions of Chhattisgarh, India.

Among the MTAs revealed, markers RM 489 and RM 408 were significantly associated with QTL controlling the grain length (GL) and grain width (GW). This is consistent with the findings of Huggins *et al.* (2019) who reported a strong association of RM 489 (chromosome 3) with grain length in brown rice, and grain width in rough rice of USDA National small grains collections. Similarly, RM 6, RM 215 and RM 44 located on chromosomes 2, 9 and 8, respectively, were significantly associated with the protein content, kernel breadth (KB) and kernel length (KL). A related earlier study by Septiningsih *et al.* (2003) had revealed RM 6 and RM 215 to have associated with the grains per plant (GPP) and panicle length (PL). These parallels between our results and similar studies in rice from across the world suggest the reliability of our MTAs for further testing in the relevant breeding programmes. For instance, markers RM 271 and OSR 13 mapped on chromosomes 10 and 3 are significantly associated with the flag leaf length (FLL) and flag leaf width (FLW) in our study. An earlier work reporting QTLs for leaf width (Anandan *et al.*, 2016) also established their involvement in leaf-related traits. Thus, it is worthwhile examining whether they could be used to improve leaf traits in Thulo Attey – a popular rice landrace of Sikkim. Similarly, the other MTAs could be explored as potential leads for testing in the relevant breeding programmes.

Conclusion

In conclusion, moderate diversity, low divergence and high gene flow among the rice landraces of Sikkim-Himalaya signify their prospects for association mapping. The reliability of the deployed SSR markers implies their utility in genetic studies. The identified MTA signals scope for marker-assisted breeding aiming at the improvement of locally relevant traits in these landraces.

However, given the unique requirements of the region, a participatory plant breeding approach is likely to provide the right direction for these efforts. The future work must also focus on comprehensive screening of the on-hand germplasm for genes/markers related to up-land agriculture and other traits relevant to this region.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S1479262121000411>.

Acknowledgement. The authors acknowledge the financial support from the Department of Biotechnology (DBT), Government of India (Grant No. BCIL/NER-BPMC/2016). They also thank The Director, National Rice Research Institute (NRRI) Cuttack, Odisha, India for extending the necessary facilities for this work. Further, all the authors duly acknowledge the contribution of local farmers of the Sikkim state for providing the rice germplasm/accession for this study.

References

- AACC (2000) Approved Methods of the American Association of Cereal Chemists, United States: American Association of Cereal Chemists.
- Abdurakhmonov IY and Abdurakimov A (2008) Application of association mapping to understand the genetic diversity of plant germplasm resources. *International Journal of Plant Genomics* **2008**, 574927.
- Alam MA, Juraimi AS, Rafii MY, Hamid AA, Arolo IW and Latif MA (2015) Application of EST-SSR marker in the detection of genetic variation among purslane (*Portulaca oleracea* L.) accessions. *Revista Brasileira de Botanica* **38**, 119–129.
- Aljumaili SJ, Rafii MY, Latif MA, Sakimin SZ, Arolo IW and Miah G (2018) Genetic diversity of aromatic rice germplasm revealed by SSR markers. *BioMed Research International* **2018**, 1–11.
- Altaf-Khan M, Qureshi SN, Saha S, Jenkins JN, Brubaker CL and Reddy OUK (2006) Usefulness of SSR derived from tetraploid *Gossypium* spp. For analyses of diploid *Gossypium* spp. *Journal of Crop Improvement* **16**, 1–20.
- Anandan A, Anumalla M, Pradhan SK and Ali J (2016) Population structure, diversity and trait association analysis in rice (*Oryza sativa* L.) germplasm for early seedling vigor (ESV) using trait linked SSR markers. *PLoS ONE* **11**, e0152406.
- AOAC (2004) *Official Methods of Analysis*, 15th Edn. Washington, DC, USA: Association of Official Analytical Chemists.
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y and Buckler ES (2007) TASSEL, software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635.
- Chen C, He W, Nassirou TY, Nsabiymva A, Dong X, Adedze YMN and Jin D (2017) Molecular characterization and genetic diversity of different accessions of *Oryza sativa* and *Oryza glaberrima*. *Electronic Journal of Biotechnology* **30**, 48–57.
- Choudhury B, Khan ML 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.
- Choudhury DR, Singh N, Singh AK, Kumar S, Srinivasan K and Tyagi RK (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**, e113094.
- Das B, Sengupta S, Parida SK, Roy B, Ghosh M, Prasad M and Ghose TK (2013) Genetic diversity and population structure of rice landraces from Eastern and North Eastern States of India. *BMC Genetics* **14**, 71. <https://doi.org/10.1186/1471-2156-14-71>
- Das SK and Avasthe RK (2016) Soil acidity management for higher yields in organic farming. In Avasthe RK, Pradhan Y and Bhutia K (eds), *Transforming Rural Sikkim: Year Round Organic Crop Production*. Tadong, Gangtok, Sikkim: Sikkim Organic Mission, Govt. of Sikkim and ICAR Research Complex, Sikkim Centre, pp. 219–222.
- Doyle JJ and Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**, 11–15.

- Earl DA and Von Holdt BM** (2012) Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conservation Genetic Resources* **4**, 359–361.
- Elston RC** (2014) Polymorphism Information Content. Wiley StatsRef: Statistics Reference Online. <https://doi.org/10.1002/9781118445112.stat05425>
- Evanno G, Regnaut S and Goudet J** (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14**, 2611–2620.
- Famoso AN, Zhao K, Clark RT, Tung C-W, Wright MH, Bustamante C, Kochian LV and McCouch SR** (2011) Genetic Architecture of Aluminum Tolerance in Rice (*Oryza sativa*) Determined through Genome-Wide Association Analysis and QTL Mapping. *PLoS Genetics* **7**(8), e1002221. DOI 10.1371/journal.pgen.1002221
- Flint-Garcia SA, Thornsberry JM and Buckler ES** (2003) Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology* **54**, 357–374.
- Food Supply and Agriculture Development Department (FSADD) Govt. of Sikkim**, Annual Progress Report. (2018–19).
- Gramene**. Available at <http://www.gramene.org/> (Accessed 22 January 2019).
- Han B and Huang X** (2013) Sequencing-based genome-wide association study in rice. *Current Opinion in Plant Biology* **16**, 133–138.
- Hore DK** (2005) Rice diversity collection, conservation, and management in Northeastern India. *Genetic Resources and Crop Evolution* **52**, 1129–1140.
- Huang XH, Zhao Y, Wei XH, Li CY, Wang AH, Zhao Q, Li WJ, Guo YL, Deng LW, Zhu CR, Fan DL, Lu YQ, Weng QJ, Liu KY, Zhou TY, Jing YF, Si LZ, Dong GJ, Huang T, Lu TT, Feng Q, Qian Q, Li JY and Han B** (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature Genetics* **44**, 32–39.
- Huggins TD, Chen MH, Fjellstrom RG, Jackson AK, McClung AM and Edwards JD** (2019) Association analysis of three diverse rice (*Oryza sativa* L.) germplasm collections for loci regulating grain quality traits. *The Plant Genome* **12**, 170085. doi: 10.3835/plantgenome.2017.09.0085
- ICAR-National Rice Research Institute**: High yielding rice varieties and hybrids developed at NRRI. Available at <https://icar-nrri.in/released-varieties> (Accessed 15 May 2021).
- Islam MZ, Khalequzzaman M, Prince MFRK, Siddique MA, Rashid MSMH, Ahmed MSU, Pittendrigh BR and Ali MP** (2018a) Diversity and population structure of red rice germplasm in Bangladesh. *PLoS ONE* **13**, e0196096.
- Islam MZ, Khalequzzaman M and Bashar MK** (2018b) Variability assessment of aromatic rice germplasm by pheno-genomic traits and population structure analysis. *Scientific Reports* **8**, 9911.
- Juliano B** (1979) *Amylose Analysis – A Review*. In: *Chemical Aspects of Rice Grain Quality*. Los Banos: International Rice Research Institute, pp 251–260.
- Kambewa BMD, Mfitilodze MW, Huttner K, Wollny CBA and Phoya RKD** (1997) Ethno veterinary medicine: alternatives for livestock development. In Mathias E, Ranganekar DV and McCorkle CM (eds), *Proceedings of an International Conference held in Pune, India*. On November 4–6 (1999): Selected papers BAIF Development Research Foundation, Pune, India **1**, 60–66.
- Kapoor C, Avasthe RK, Chettri PK, Gopi R and Kalita H** (2017) On-farm status and collection of rice (*Oryza sativa* L.) genetic resources of Sikkim Himalayas. *Indian Journal of Plant Genetics Resources* **30**: 48–54. doi: 10.5958/0976-1926.2017.00005.5
- Karkousis A, Barr AR, Chalmers KJ, Ablett GA, Holton TA, Henry RJ, Lim P and Langridge P** (2003) Potential of SSR markers for plant breeding and variety identification in Australian barley germplasm. *Australian Journal of Agricultural Research* **54**, 1197–1210.
- Khush GS** (1997) Origin, dispersal, cultivation and variation of rice. In Sasaki T and Moore G (eds), *Oryza: From Molecule to Plant*. Dordrecht: Springer **35**, 24–34. <https://doi.org/10.1007/978-94-011-5794-0-3>
- Li FP, Lee YS, Kwon SW, Li G and Park YJ** (2014) Analysis of genetic diversity and trait correlation among Korean landrace rice (*Oryza sativa* L.). *Genetics and Molecular Research* **13**, 6316–6331.
- Liu K and Muse SV** (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* **21**, 2128–2129. doi: 10.1093/bioinformatics/bti282
- Meti N, Samal KC, Bastia DN and Rout GR** (2013) Genetic diversity analysis in aromatic rice accessions using microsatellite-based simple sequence repeats (SSR) marker. *African Journal of Biotechnology* **12**, 4238–4250.
- Morgante M and Olivieri A** (1993) PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal* **3**, 175–182.
- Mulalem T, Mekbib F, Shimelis H, Gebre E and Amelwork B** (2018) Genetic diversity of yam (*Dioscorea* spp.) landrace collections from Ethiopia using simple sequence repeat markers. *Australian Journal of Crop Science* **12**, 1223–1230.
- Musyoki MA, Kioko WF, Mathew NP, Muriira KG, Wavinya ND, Felix M, Chemtai LR, Mwenda NS, Kiambi MJ and Ngithi NL** (2015) Genetic diversity studies on selected rice (*Oryza sativa* L.) genotypes based on amylose content and gelatinization temperature. *Advances in Crop Science and Technology* **3**, 193. DOI 10.4172/2329-8863.1000193
- Nachimuthu VV, Muthurajan R, Duraijalaguraja S, Sivakami R, Pandian BA and Ponniah G** (2015) Analysis of population structure and genetic diversity in rice germplasm using SSR markers: an initiative towards association mapping of agronomic traits in *Oryza sativa*. *Rice* **8**, 30–54.
- Nuijten E, Van Treuren R, Struik PC, Mokuwa A, Okry F, Teeken B and Richards P** (2009) Evidence for the emergence of new rice types of inter-specific hybrid origin in West African farmers' fields. *PLoS ONE* **4**, e7335.
- Pachauri V, Taneja N, Vikram P, Singh NK and Singh S** (2013) Molecular and morphological characterization of Indian farmers rice varieties (*Oryza sativa* L.). *Australian Journal of Crop Science* **7**, 923–932.
- Patil AH, Premi V, Sahu V, Dubey M, Sahu GR and Chandel G** (2014) Identification of elite rice germplasm lines for grain protein content, SSR based genotyping and DNA fingerprinting. *International Journal of Plant Animal and Environment* **1**, 127–136.
- Pavlicek A, Hrdá S and Flegr J** (1999) Free tree: freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jackknife analysis of the tree robustness. Application in the RAPD analysis of genus *Frenkelia*. *Folia Biologica (Praha)* **45**, 97–99.
- Peakall R and Smouse PE** (2012) GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**, 2537–2539.
- Perez-Sackett PT, Cianzio SR, Kara PC, Aviles M and Palmer RG** (2011) QTL mapping of whitefly resistance in soybean. *Journal of Crop Improvement* **25**, 134–150.
- Perrier X and Jacquemond-Collet J** (2006) DARwin software. Version 5.0. Available at <http://Darwin-cirad.fr/darwin>.
- Portel BK** (2015) *Economics of organic crops and food crops in Sikkim* (Mphil Thesis). Sikkim University.
- Pritchard J, Stephens M, Rosenberg N and Donnelly P** (2000) Association mapping in structured populations. *The American Journal of Human Genetics* **67**, 170–180.
- Protection of Plant Varieties and Farmers' Right Authority (PPV&FRA), Government of India** (2007) Guidelines for the conduct of test for distinctiveness, uniformity and stability on rice (*Oryza sativa* L.). *Plant Variety Journal of India* **1**, 1–25.
- Rahman H and Karuppaiyan R** (2011) Agrobiodiversity of Sikkim. In Arrawatia ML and Tambe S (eds), *Biodiversity of Sikkim: Exploring and Conserving A Global Hotspot*. Gangtok, Sikkim: Information and Publications Department, Govt. of Sikkim, pp. 403–426.
- Rahman MM, Rasaul MG, Hossain MA, Iftekharuddaula KM and Hasegawa H** (2012) Molecular characterization and genetic diversity analysis of rice (*Oryza sativa* L.) using SSR markers. *Journal of Crop Improvement* **26**, 244–257.
- Rani MG and Adilakshmi D** (2011) Genetic analysis of blast resistance in rice with simple sequence repeats (SSR). *Journal of Crop Improvement* **25**, 232–238.
- Rathi RS, Singh SK and Misra AK** (2018) Collection and characterization of trait specific multi-crop germplasm from Sikkim. *Indian Journal of Hill Farming* **31**, 274–280.
- Roy S, Marndi BC, Mawkhlieng B, Banerjee A, Yadav RM, Misra AK and Bansal KC** (2016) Genetic diversity and structure in hill rice (*Oryza sativa* L.) landraces from the North-Eastern Himalayas of India. *BMC Genetics* **17**, 107.

- Saha SM, Naveed SA and Arif M (2013) Genetic diversity in Basmati and non-Basmati rice varieties based on microsatellite markers. *Pakistan Journal of Botany* **45**, 423–431.
- Sarao NK, Vikal Y, Singh K, Joshi MA and Sharma RC (2010) SSR marker-based DNA fingerprinting and cultivar identification of rice (*Oryza sativa* L.) in Punjab state of India. *Plant Genetic Resources* **8**, 42–44. doi: 10.1017/S1479262109990128
- Septiningsih EM, Prasetyono J, Lubis E, Tai TH, Tjubaryat T, Moeljopawiro S and McCouch SR (2003) Identification of quantitative trait loci for yield and yield components in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. *Theoretical and Applied Genetics* **107**, 1419–1432. DOI 10.1007/s00122-003-1373-2
- Sharma RC, Chaudhary NK, Ojha B, Yadav L, Pandey MP and Shrestha SM (2007) Variation in rice landraces adapted to the lowlands and hills in Nepal. *Plant Genetic Resources* **5**, 120–127. doi: 10.1017/S1479262107837828
- Sharma L, Pradhan A, Bhutia KD, Tamang DL and Bhutia SG (2015) Local crop germplasm and landraces diversity in West Sikkim. *Journal of Agricultural Technology* **2**, 57–59.
- Sharma G, Partap U, Sharma E, Rasul G and Awasthe RK (2016) Agrobiodiversity in the Sikkim Himalaya: sociocultural significance, status, practices, and challenges. ICIMOD Working Paper 2016/5 Kathmandu: ICIMOD.
- Singh RK and Singh US (2003) *A Treatise on Scented Rices of India*. New Delhi: Kalyani Publishers.
- Singh N, Choudhury DR, Tiwari G, Singh AK, Kumar S, Srinivasan K, Tyagi RK, Sharma AD, Singh NK and Singh R (2016) Genetic diversity trend in Indian rice varieties: an analysis using SSR markers. *BMC Genetics* **17**, 217. DOI 10.1186/s12863-016-0437-7
- Suvi WT, Shimelis H, Liang M, Mathew I and Shayanowako AIT (2019) Assessment of the genetic diversity and population structure of rice accessions using SSR markers. *Acta Agriculturae Scandinavica, Section B – Soil & Plant Science* **70**, 76–86. DOI 10.1080/09064710.2019.1670859
- Talukdar PR, Rathi S, Pathak K, Chetia SK and Sharma RN (2017) Population structure and marker-trait association in indigenous aromatic rice. *Rice Science* **24**, 145–154.
- Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**, 6463–6471.
- Umakanth B, Vishalakshi B, Sathish Kumar P, Rama Devi SJS, Bhadana VP, Senguttuvel P, Kumar S, Sharma SK, Sharma PK, Prasad MS and Madhav MS (2017) Diverse rice landraces of North-East India enables the identification of novel genetic resources for magnaporthe resistance. *Frontiers in Plant Science* **8**, 1500. doi: 10.3389/fpls.2017.01500
- Upadhyay P, Neeraja CN, Kole C and Singh VK (2012) Population structure and genetic diversity in popular rice varieties of India as evidenced from SSR analysis. *Biochemical Genetics* **50**, 770–783.
- Weising K, Winter P, Hüttel B and Kahl G (1997) Microsatellite markers for molecular breeding. *Journal of Crop Production* **1**, 113–143.
- Wright S (1978) *Evolution and the Genetics of Population Vol 4: Variability within and among Natural Populations*. Chicago: University of Chicago Press.
- Yeh FC, Yang RC, Boyle T, Ye ZH and Mao JX (1999) POPGENE, Version 1.32: the user-friendly software for population genetic analysis. Edmonton, AB: Molecular Biology and Biotechnology Centre, University of Alberta.
- Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JE, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S and Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* **38**, 203–208.
- Zhou J, You A, Ma Z, Zhu L and He G (2012) Association analysis of important agronomic traits in japonica rice germplasm. *African Journal of Biotechnology* **11**, 2957–2970.
- Zhu C, Gore M, Buckler ES and Yu J (2008) Status and prospects of association mapping in plants. *The Plant Genome* **1**, 5–20.
- Zuberi MI (1997) Present state of the Ethno veterinary system in northwestern Bangladesh. Ethno veterinary medicine: alternatives for livestock development. In Mathias E, Ranganekar DV and McCorkle CM (eds), *Proceedings of an International Conference held in Pune, India*. On November 4–6 (1999): Selected papers BAIF Development Research Foundation, Pune, India. **1**, 53–56.