ISSN 1479-2621
Plant Genetic Resources: Characterization and Utilization (2014) **12**(3); 264–272
doi:10.1017/S1479262113000580

Chakhao (delicious) rice landraces (*Oryza sativa* L.) of North-east India: collection, conservation and characterization of genetic diversity

Somnath Roy[1]*, A. Banerjee[2], A. Pattanayak[2], S. S. Roy[3], R. S. Rathi[1], A. K. Misra[1], S. V. Ngachan[2] and K. C. Bansal[4]

[1]*National Bureau of Plant Genetic Resources (NBPGR), Regional Station, Umiam, Meghalaya 793 103, India,* [2]*ICAR Research Complex for NEH Region, Umiam, Meghalaya 793 103, India,* [3]*ICAR Research Complex for NEH Region, Manipur Centre, Imphal 795 004, India and* [4]*NBPGR, Pusa Campus, New Delhi 110 012, India*

Received 9 October 2013; Accepted 22 November 2013 – First published online 2 January 2014

Abstract

Rice landraces have been developed through artificial selection imposed by farmers during the long-term domestication process. Although the global rice diversity is well characterized, few studies have conducted an in-depth analysis of genetic diversity on a local scale. In India, there are many locally adapted non-Basmati aromatic rice landraces in which the pace of improvement is slow, despite their high economic values. The aromatic and quality rice landraces of Manipur, locally called *Chakhao* (delicious), are important, considering their high economic and cultural values. To conserve and encourage improvement of these landraces, we collected *Chakhao* accessions from eight districts of Manipur. The aim of the current study was to investigate the level of genetic diversity and structure of 37 *Chakhao* landraces based on genotyping with 47 microsatellite markers. The genetic diversity analysis revealed high gene diversity (0.673) within the *Chakhao* population, with values ranging from 0.303 (*Poireiton*) to 0.471 (mixed *Chakhao*). The *Chakhao* rice accessions could be divided into six subgroups based on genetic structure analyses. The population structure derived from the STRUCTURE analysis largely correlated with the farmers' classification of *Chakhao* landraces. The results of genetic diversity analyses and the indigenous knowledge of the names and use of *Chakhao* landraces would facilitate the conservation and utilization of this unique genetic resource.

Keywords: aromatic rice; *Chakhao*; conservation; genetic diversity; SSR

Introduction

Being the centre of rice diversity, India has a wealth of traditional landrace varieties. The landraces from the North-eastern (NE) states of the country are especially diverse morphologically and genetically and are worthy candidates for in-depth analyses (Das *et al.*, 2013; Roy *et al.*, 2013). The traditional farmers of the region grow many diverse rice varieties having special cultural values and qualities. Moreover, the region is also home to many locally adapted aromatic and quality rice landraces. In the current study, we focused on examining the genetic structure of the aromatic rice cultivars grown by the traditional farmers of Manipur state of India. These rice varieties are locally called *Chakhao*. The literal meaning of *Chakhao* in Manipuri language is delicious rice (*Chak* = rice and *ahoba* or *hao* = delicious). In the state, the farmers describe different

* Corresponding author. E-mail: sroypbr@gmail.com

types of *Chakhao* based on grain characteristics (*Amubi* = black; *Angoba* = white). The farmers grow a number of different types of *Chakhao* rice landraces within a short range of altitude differences on a small scale. The variability in the grain morphology of *Chakhao* rice landraces is remarkable. *Chakhao* rice is very special to the people of Manipur, as they use it in festivals and social ceremonies to prepare various unique dishes, namely *Ethe Tan* (a local *puri* made from black *Chakhao* rice flour in Chandel district), *Buhman Sang* (a local delicacy made from *Buhman* landraces in Churachandpur district) and *Utong Chak* (a special dish prepared within bamboo sticks in Chandel district). The cultivation of *Chakhao* landraces is declining, as the farmers prefer growing agronomically improved high-yielding varieties. Therefore, the assessment, documentation, analysis and conservation of the level of genetic diversity are essential for exploiting these rice landraces in variety development.

Although the genetic diversity of cultivated rice germplasm on a global scale has been well characterized using molecular markers (Yu *et al.*, 2003; Garris *et al.*, 2005; Caicedo *et al.*, 2007), important issues concerning crop genetic diversity and its relationship with local cultures can only be answered by rigorously studying rice germplasm collected from geographically isolated regions (see Thomson *et al.*, 2009). A number of studies have been conducted in the past to characterize subsets of rice germplasm, with molecular markers being used for analysing the genetic diversity within specific countries (Prashanth *et al.*, 2002; Jain *et al.*, 2004; Gao *et al.*, 2005; Thomson *et al.*, 2007; 2009). Recently, an analysis of genetic diversity in the rice germplasm of NE India has been carried out using microsatellite markers (Choudhury *et al.*, 2013; Das *et al.*, 2013). These studies have reported a high level of genetic diversity in the subsets of NE Indian rice landraces. However, an analysis of the genetic structure of *Chakhao* rice germplasm is yet to be carried out, which is essential for the systematic conservation of these landraces. The current study was carried out to determine the level of genetic diversity in a set of 37 *Chakhao* rice accessions collected from the state of Manipur, NE India, using 47 microsatellite markers.

Materials and methods

Seed collection and conservation

Seeds from 37 *Chakhao* rice cultivars were collected from the districts of Manipur, India, during a collection trip

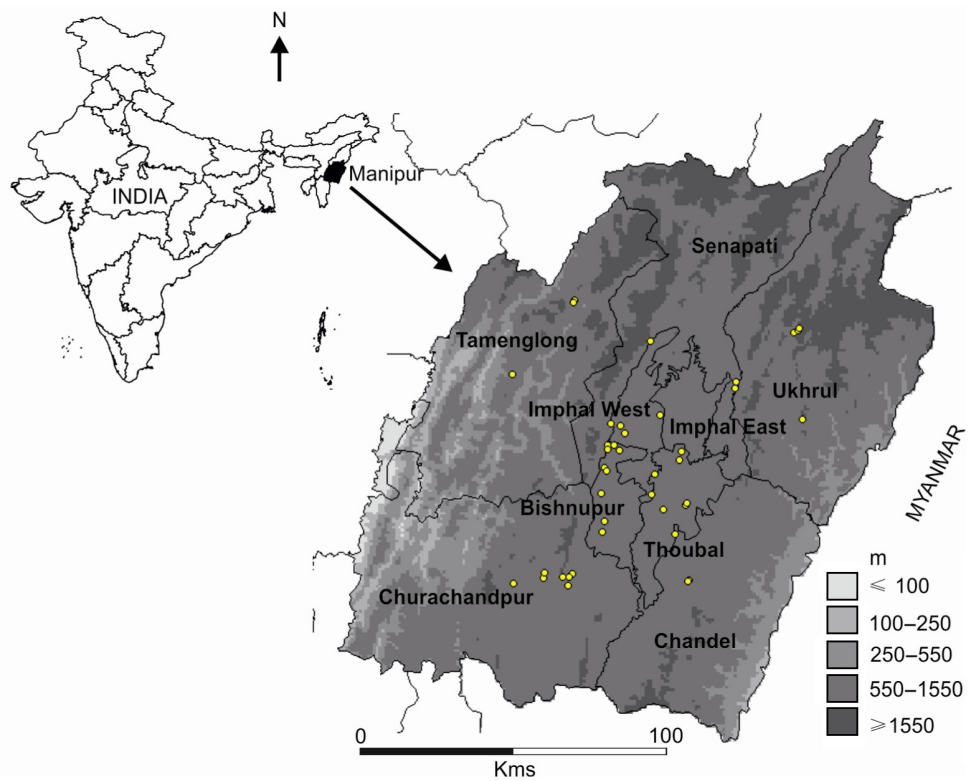

Fig. 1. Geographical distribution of the sampled sites in Manipur. A point indicates a site where collection has been made.

https://doi.org/10.1017/S1479262113000580 Published online by Cambridge University Press

from 4 to 13 November, 2011. Seeds were collected directly from the fields in each district (Fig. 1). In a few cases, freshly harvested seeds were collected from the farmers. The farmers were interviewed to gather information on the naming, use and other characteristics of the specific cultivars. The seeds of the rice accessions used in this study are conserved in the National Genebank of National Bureau of Plant Genetic Resources (NBPGR), New Delhi, India, and are publicly available for research purposes upon request and agreeing to an material transfer agreement (MTA).

Simple sequence repeat (SSR) genotyping

Five-day-old rice seedlings germinated from five well-developed seeds from a single plant of each accession were used for total genomic DNA extraction with a plant DNA extraction kit following the manufacturer's protocol (Qiagen, CA, USA). Forty-seven SSR primer pairs distributed across the rice genome were designed from the Gramene marker database (http:\\www.gramene.org\markers\microsat\). The name, chromosomal location, motif and annealing temperature of the SSR markers are given in Supplementary Table S1 (available online). Polymerase chain reactions (PCRs) were carried out using a mixture of total volume of 25 µl consisting of 50 ng of genomic DNA, 0.2 mM of dNTPs, 2.5 mM of $MgCl_2$, 2.5 µl of 10 × PCR Buffer, 2.5 pmol of forward and reverse primers, and 0.6 U of *Taq* DNA polymerase on a thermocycler (Mastercycler; Eppendorf, Hamburg, Germany). The following PCR protocol was used: a denaturation period of 5 min at 94°C, followed by 30 cycles of 30 s at 94°C, 30 s at a particular annealing temperature and 30 s at 72°C, and then 7 min at 72°C for final extension.

The PCR products were resolved in ethidium bromide (10 mg/ml)-stained 2% agarose gels. The molecular sizes (in nucleotide) of the amplified alleles were determined based on their migration relative to the 50 bp DNA Step ladder (Promega, Madison, WI, USA) using the PyElph-1.4 gel image analysis software (Pavel and Vasile, 2012). The band sizes of each marker were checked using the Gramene database (www.gramene.org\markers\microsat\). The band with the lowest molecular weight for each SSR marker was assigned allele number 1 and the progressively heavier bands were assigned numbers incrementally. For an individual marker, the genotypes were scored as homozygous/heterozygous for particular allele(s).

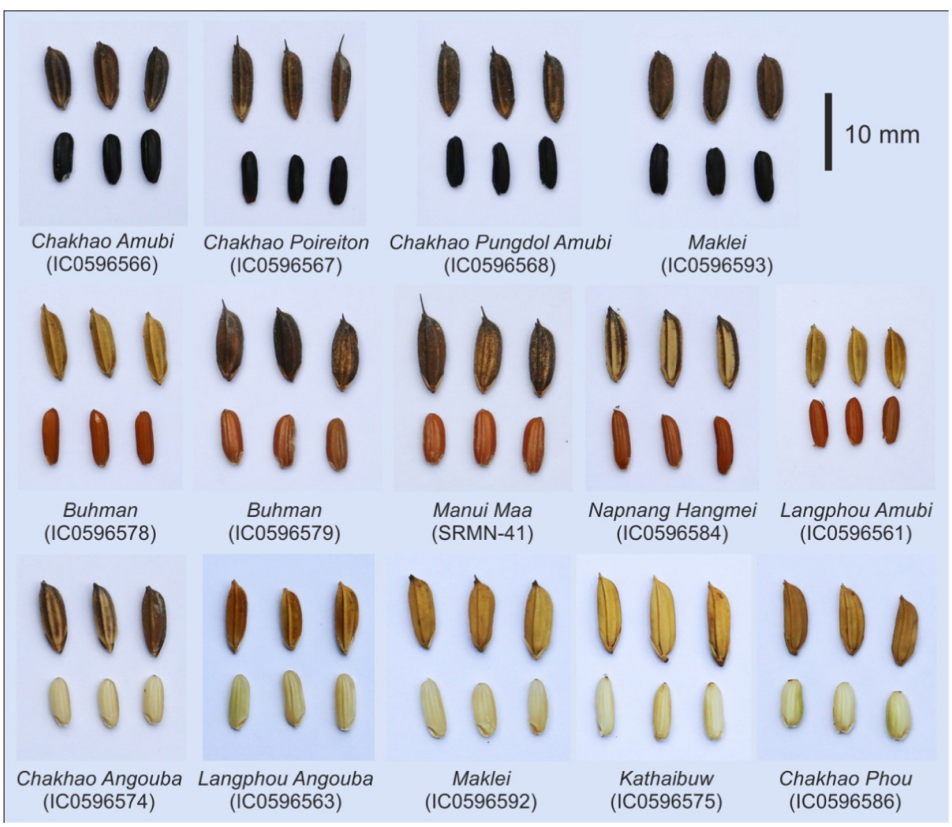

Fig. 2. Variation in the grain morphology of different *Chakhao* rice landraces collected from Manipur state of North-east India.

https://doi.org/10.1017/S1479262113000580 Published online by Cambridge University Press

Data analysis

Genetic parameters such as observed number of alleles (A), average number of alleles per locus (N_a), effective number of alleles (N_e), polymorphic information content (PIC), percentage of polymorphic loci, observed heterozygosity (H_o), expected heterozygosity (H_e), Shannon's diversity index (I) and number of private alleles were calculated using the PopGene (version 1.31, Yeh *et al.*, 1999) and GenAlEx (version 6.5, Peakall and Smouse, 2012) software packages. Pairwise genetic differentiation (F_{ST}) values among the subpopulations were estimated and an analysis of molecular variance (AMOVA) of the rice populations was carried out to partition the levels of genetic diversity among regions (valleys and hills) and among and within *Chakhao* types using the GenAlEx software. For drawing the dendrogram, genetic distance was calculated using the Chord distance matrix (Cavalli-Sforza and Edwards, 1967), followed by tree construction using neighbour-joining as implemented in the NTSYS-pc software (version 2.1, Rohlf, 2000). Principal coordinate analysis (PCoA) was carried out using the genetic distance matrix calculated in the GenAlEx software.

The population structure of *Chakhao* rice accessions was assessed using the model-based (Bayesian clustering) method implemented in STRUCTURE version 2.3.4 (Pritchard *et al.*, 2000). The number of subgroups (SGs) (K) was set from 1 to 9 based on models characterized by admixture and correlated allele frequencies. For each K, ten runs were performed separately, with 1,00,000 iterations

Table 1. List of the 37 *Chakhao* rice accessions of Manipur used in the current study

Serial no.	Collector code	Accession number (IC)	Vernacular name	Kernel length class[a]	Kernel shape[a]	Aroma[b]
1	SRMN-1	0596559	*Chakhao poireiton*	Long	Medium	High
2	SRMN-2	0596560	*Chakhao angouba*	Short	Bold	High
3	SRMN-3	NA	*Chakhao amubi*	Medium	Medium	High
4	SRMN-4	0596561	*Langphou amubi*	Medium	Medium	Low
5	SRMN-7	0596563	*Langphou angouba*	Long	Medium	Low
6	SRMN-8	0596564	*Chakhao*	Short	Bold	Medium
7	SRMN-9	0596565	*Chakhao poireiton*	Medium	Medium	High
8	SRMN-10	0596566	*Chakhao amubi*	Medium	Medium	High
9	SRMN-11	0596567	*Chakhao poireiton*	Long	Slender	High
10	SRMN-12	0596568	*Chakhao pungdol amubi*	Long	Slender	High
11	SRMN-13	0596569	*Chakhao poireiton*	Long	Slender	High
12	SRMN-14	0596570	*Chakhao amubi*	Medium	Medium	High
13	SRMN-15	0596571	*Chakhao angouba*	Long	Medium	Medium
14	SRMN-17	0596573	*Chakhao amubi*	Medium	Medium	High
15	SRMN-18	0596574	*Chakhao angouba*	Long	Slender	High
16	SRMN-19	0596575	*Kathaibuw*	Long	Medium	Low
17	SRMN-20	0596576	*Buhman*	Long	Medium	Medium
18	SRMN-21	0596577	*Buhman*	Long	Medium	High
19	SRMN-22	0596578	*Buhman*	Long	Slender	Medium
20	SRMN-23	0596579	*Buhman*	Long	Medium	High
21	SRMN-24	0596580	*Buhman*	Long	Medium	Low
22	SRMN-25	0596581	*Buhman/Idaw*	Extra long	Slender	Medium
23	SRMN-26	0596582	*Buhman*	Long	Medium	High
24	SRMN-27	0596583	*Napnang hangmei*	Long	Medium	Medium
25	SRMN-28	0596584	*Napnang hangmei*	Long	Slender	Medium
26	SRMN-29	0596585	*Chakhao phou*	Long	Medium	High
27	SRMN-30	0596586	*Chakhao phou*	Medium	Bold	Medium
28	SRMN-31	0596587	*Chakhao poireiton*	Long	Medium	High
29	SRMN-32	0596588	*Chakhao*	Medium	Slender	High
30	SRMN-33	0596589	*Chakhao amubi*	Medium	Slender	High
31	SRMN-34	0596590	*Chakhao poireiton*	Medium	Medium	High
32	SRMN-36	NA	*Maklei (Chakhao)*	Long	Medium	High
33	SRMN-37	0596591	*Maklei*	Medium	Medium	Medium
34	SRMN-38	0596592	*Maklei*	Long	Medium	Medium
35	SRMN-39	0596593	*Maklei*	Medium	Medium	Medium
36	SRMN-40	0596594	*Maklei*	Medium	Medium	Medium
37	SRMN-41	NA	*Manui maa*	Long	Medium	Medium

IC, indigenous collection; NA, not available.
[a] Determined following Cruz and Khush (2000). [b] Detected following Nagaraju *et al.* (1991).

https://doi.org/10.1017/S1479262113000580 Published online by Cambridge University Press

Table 2. Genetic diversity parameters of different types of *Chakhao* rice varieties[a]

Chakhao type	Number of accessions	N_a	*P* (%)	H_e	H_o	*I*	Number of private alleles	F_{ST}
Types								0.443
Poireiton	6	1.723 (0.498)	70	0.303 (0.031)	0	0.439 (0.044)	6	–
Amubi	7	2.468 (0.830)	89	0.431 (0.028)	0.030 (0.009)	0.707 (0.050)	14	–
Angouba	4	2.085 (0.803)	75	0.372 (0.036)	0	0.590 (0.060)	3	–
Buhman	7	1.511 (0.505)	51	0.341 (0.030)	0	0.535 (0.047)	6	–
Chakhao	8	3.191 (0.825)	100	0.471 (0.026)	0.016 (0.006)	0.828 (0.050)	29	–
Maklei	5	1.830 (0.433)	81	0.391 (0.028)	0.021 (0.009)	0.552 (0.040)	11	–
Regions								0.138
Valley	17	3.468 (1.316)	96	0.550 (0.025)	0.013 (0.009)	0.979 (0.057)	52	0.300
Hill	20	3.936 (1.389)	100	0.609 (0.021)	0.012 (0.008)	1.112 (0.052)	74	0.372
Total	37	5.043 (1.876)	78	0.673 (0.019)	0.012 (0.006)	1.327 (0.056)	69	–

N_a, average number of alleles per locus; *P*, percentage of polymorphic loci; H_e, expected heterozygosity (Nei's gene diversity); H_o, observed heterozygosity; *I*, Shannon's information index; F_{ST}, genetic differentiation.
[a] Numbers in parentheses indicate standard deviation.

carried out for each run after a burn-in period of 1,00,000 iterations. The number of clusters (*K*) was set from 1 to 9 with ten independent runs. The 'Structure Harvester' (version 6.0, Earl and von Holdt, 2011) was used to summarize the STRUCTURE results and to find the optimum *K* value implementing the parameters of Evanno *et al.* (2005). The ΔK value was based on the change in the log probability of the data between successive *K* values.

Results

A wide range of variations were observed in the grain morphology of the collected aromatic rice landraces of Manipur (Fig. 2). All the known aromatic rice landraces such as *Chakhao Poireiton*, *Chakhao Amubi* and *Chakhao Angouba* (as recognized by the farmers) were collected along with other *Chakhao* types such as *Buhman*, *Maklei*, *Chakhao Phou* and *Napnang Hangmei*. Variations in grain shape, size and aroma are given in Table 1.

Genetic diversity

The whole set of 47 SSR markers detected a total of 237 alleles across 37 *Chakhao* rice accessions (Supplementary Table S2, available online) with an average of 5.04 alleles per locus (size ranging from 73 to 343 bp). The highest number of alleles was scored at the locus RM552 (11 alleles) and the lowest was scored at the loci RM55 and RM178 (two alleles each). The PIC values varied from 0.317 (RM178) to 0.868 (RM552) with an average of 0.63. High gene diversity (H_e) was observed for several SSR loci such as RM552 (0.892), RM240 (0.874) and RM80 (0.837). Among the 237 identified alleles, 69 were private; that is, a given allele was identified in only one accession (Table 2). The highest number of private alleles was observed in *Chakhao* (mixed) population (29) followed by *Amubi* (14) and *Maklei* (11).

Judging from the overall genetic parameters (Table 2), considerable genetic diversity was found among the *Chakhao* landraces ($H_e = 0.673$, $I = 1.327$). Genetic diversity was not uniformly distributed across the populations. The highest level of genetic diversity was found in *Chakhao* (mixed) population ($H_e = 0.471$, $I = 0.828$), whereas the lowest was recorded in *Poireiton* ($H_e = 0.303$, $I = 0.439$). Considering the regions (valley and hill) from where the rice accessions were collected, the largest genetic diversity was recorded in hilly regions (data not shown). F_{ST} among the *Chakhao* types ($F_{ST} = 0.443$) was higher than that among the regions ($F_{ST} = 0.138$). F_{ST} values among the populations of valleys and hills were similar to F_{ST} values

Table 3. Results of the analysis of molecular variance based on the 47 SSR loci of 37 *Chakhao* rice accessions

Source	df	MS	CV	% Total	*P* value
Among regions	1	316.716	7.295	10	≥0.001
Among *Chakhao* types	4	175.105	22.041	31	≥0.001
Within *Chakhao* types	31	41.983	41.983	59	≥0.001
Total	36		71.320	100	

df, Degree of freedom; MS, mean sum of squares; CV, variance component estimates; % Total, percentage of total variation.

https://doi.org/10.1017/S1479262113000580 Published online by Cambridge University Press

0.300 and 0.372, respectively. The AMOVA results revealed statistically significant differentiation with 10% variation among the regions, 31% among the *Chakhao* types and 59% within the types (Table 3).

Genetic cluster

The genetic distance-based neighbour-joining-based phylogeny divided the *Chakhao* germplasm into two broad groups (Fig. 3(a)). The *Poireiton* and *Amubi* accessions having black kernel were genetically similar and grouped in Cluster 1. The *Angouba* and *Buhman* accessions possessing white/red kernels grouped in Cluster 2. During the collection trip, some ambiguities were observed in the classification of *Angouba* and *Buhman* cultivars by the farmers of Churachandpur district. Cluster 1 also included two *Angouba* accessions. The *Maklei* and *Manui Maa* and *Napnang Hangmei*, *Chakhao Phou*, and *Kathaibuw* accessions formed subclusters within Cluster 1. Similarly, the PCoA using pairwise genetic distances revealed a clear separation of the *Poireiton*, *Amubi*, *Angouba* and *Buhman* accessions from the other *Chakhao* cultivars (Fig. 3(b)), similar to that in the cluster analysis. The results of cluster analysis and PCoA revealed a significant population structure in the *Chakhao* germplasm.

Population structure

An analysis of the population structure of the 37 *Chakhao* landraces identified the most appropriate grouping with six SGs. The number of SGs (K) was identified based on the ΔK *versus* K plot, which showed an evident knee at $K = 6$ (Supplementary Fig. S1, available online). Using a membership probability threshold of 0.60, eight accessions (*Chakhao Phou*, *Napnang Hangmei*, *Chakhao* and *Maklei*) were assigned to SG1, three accessions (*Angouba* and *Pungdol Amubi*) to SG2, nine accessions (*Buhman*, *Langphou Angouba* and *Chakhao*) to SG3, eight accessions of *Poireiton* and *Amubi* to SG4, four accessions of *Amubi* to SG5, and four accessions (three *Maklei* and one *Manui Maa*) to SG6 (Fig. 3(c)). The STRUCTURE results indicated that the inferred population structure mostly correlated with the farmers' classification of the *Chakhao* landraces. The assignments of the 37 *Chakhao* accessions to six SGs derived from STRUCTURE were consistent with their grouping in the neighbour-joining tree.

Discussion

The current study estimated the genetic relatedness among 37 *Chakhao* rice landraces from Manipur state of NE India using 47 SSR markers and analysed the population structure among the landraces. The PCR amplification results with all the markers are consistent with those reported previously (Garris *et al.*, 2005). The N_a value obtained in the current study was 5.04. This value is lower when compared with a value of 7.8 across a set of 52 aromatic rice cultivars from India (Jain *et al.*, 2004). The N_a value obtained in the current study was also lower than that reported across diverse sets of global rice accessions (11.9, Xu *et al.*, 2004; 13.0, Thomson *et al.*, 2007; and 12.4, Borba *et al.*, 2009). Recently, Choudhury *et al.* (2013) and Das *et al.* (2013) have reported higher N_a values (13.57 and 7.9, respectively) across sets of indigenous rice cultivars of eastern and NE Indian states. However, the N_a value obtained in the current study is comparable to that (5.5) recorded in a set of 183 rice accessions collected from 18 villages on the island of Borneo, Indonesia (Thomson *et al.*, 2009). The average PIC value (0.63 per marker) was similar to the value (0.66) reported within a set of traditional and improved rice varieties of Indonesia (Thomson *et al.*, 2007). A similar PIC value (0.6) was also recorded by Jain *et al.* (2004) across a set of aromatic rice cultivars distributed throughout India using 30 SSR markers. Das *et al.* (2013) recorded an average PIC value of 0.57 within the rice accessions collected from different NE Indian states using 23 trait-linked SSR markers.

In the current study, the highest gene diversity (H_e) value was recorded in the subpopulation of *Chakhao* (mixed) landraces (0.47) and the lowest value was observed in that of *Poireiton* landraces (0.303). It is obvious that gene diversity in the *Chakhao* (mixed) subpopulation was due to the inclusion of different *Chakhao* landraces within this group. Higher H_e values were also recorded in both *Amubi* and *Maklei* subpopulations. The *Chakhao* landraces had a lower gene diversity value when compared with the value of 0.78 across a set of indigenous rice landraces of NE India. However, the value was higher than that recorded in agronomically improved varieties (0.46) of the region (Choudhury *et al.*, 2013). Although the *Chakhao* landraces represent a subset of total genetic diversity in NE India, the average gene diversity recorded in the current study was at par with that recorded across diverse sets of rice germplasm of Indonesia (0.68, Thomson *et al.*, 2007) and China (0.71, Tu *et al.*, 2007) and across sets of global rice collections (0.55, Garris *et al.*, 2005; 0.64, Borba *et al.*, 2009). The gene diversity value obtained in the current study was higher than that recorded in a set 183 rice landraces (0.49) from the island of Borneo, Indonesia (Thomson *et al.*, 2009). Generally, the cultivars sampled from a small geographical region would be inherently less diverse as they are grown in a similar set of

https://doi.org/10.1017/S1479262113000580 Published online by Cambridge University Press

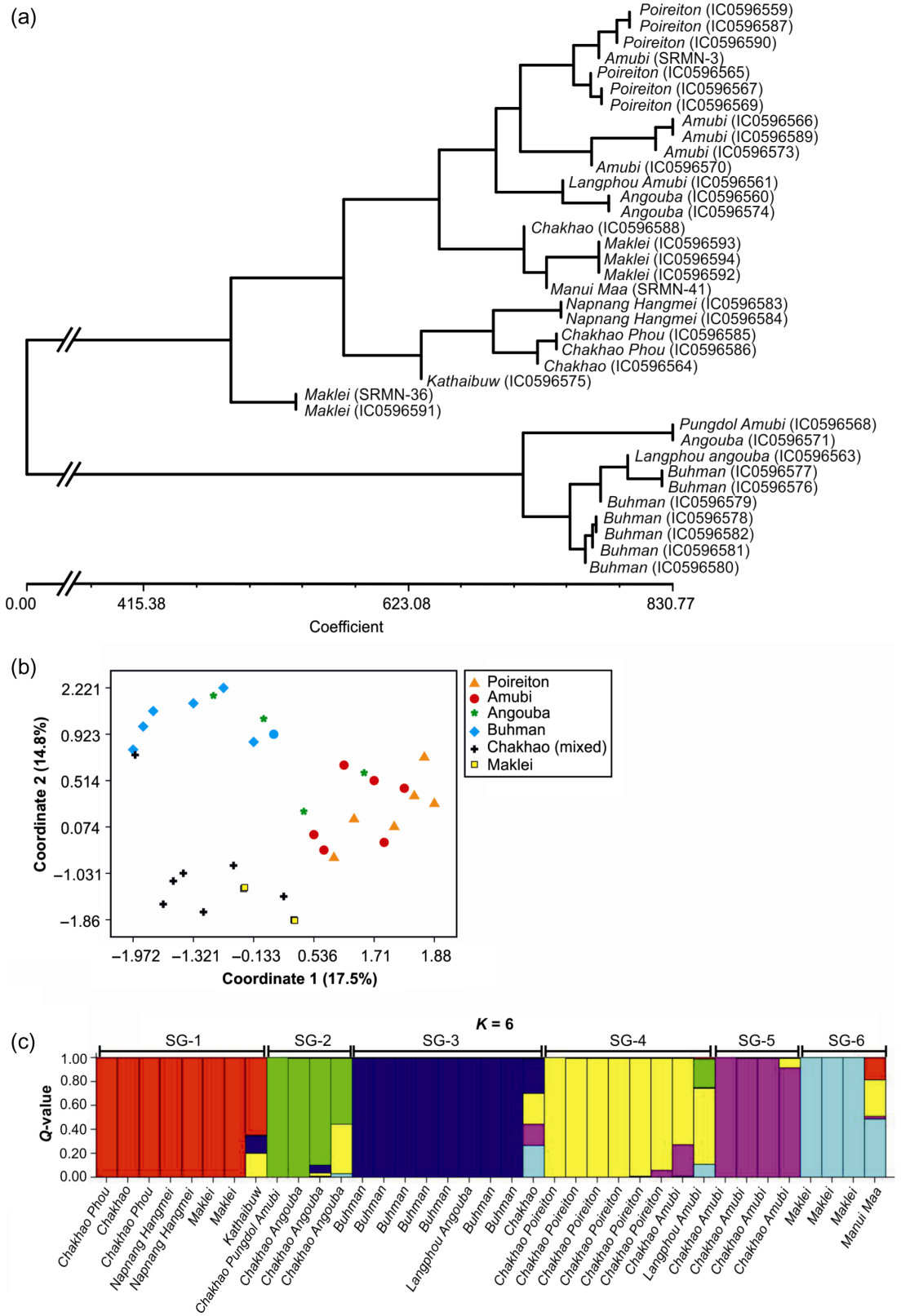

Fig. 3. Genetic structure of the *Chakhao* rice accessions. (a) Dendrogram of 37 *Chakhao* accessions by neighbour-joining clustering; (b) principal coordinate analysis based on variation in 47 simple sequence repeats (SSR); and (c) six subgroups (SGs) inferred from the STRUCTURE analysis. The vertical of each SG indicates the membership coefficients (*Q*). The SGs (1 to 6) are indicated on the top of the bar plot.

https://doi.org/10.1017/S1479262113000580 Published online by Cambridge University Press

environmental conditions, whereas global collections of germplasm being sampled from a wider range of geographically and ecologically distinct regions will exhibit greater genetic diversity due to divergent selection pressures. The high genetic diversity in the current set of 37 *Chakhao* accessions may be a reflection of the prevalent diverse agro-climatic conditions and diverse cultural practices followed by the farmers of the region. However, further characterization of more landraces from this region would help in establishing this association. In general, the genetic diversity maintained in a plant species is considered a function of its ecological and evolutionary history (Hamrick and Godt, 1996). The conservation of high gene diversity within the *Chakhao* landraces could be the outcome of traditional farming systems and varied human preferences in Manipur.

The results based on both F_{ST} and AMOVA confirmed that a major genetic variation exists within different types of *Chakhao* rice landraces. A moderately high estimate of F_{ST} (0.443) considering the *Chakhao* types indicated less admixed ancestry among the cultivars. However, when analysing F_{ST} for the regions, a value of 0.138 indicated the presence of admixed ancestry within the regions. Similarly, Thomson *et al.* (2009) reported a lack of geographical differences within rice landraces collected from three different regions within the island of Borneo, Indonesia. It was also proposed that the lack of seed exchange among the traditional farmers has prevented any significant genetic differences occurring over time. The F_{ST} results are also supported by the AMOVA results, which indicated that 31% of the total variation was due to differentiation among the *Chakhao* types. It is to be noted that 59% of the total variation is explained by differentiation within the types, i.e. among the different *Chakhao* landraces. Similarly, Choudhury *et al.* (2013) also reported that differentiation among rice varieties explained 66% of the total genetic variation in a set of diverse rice genotypes comprising indigenous cultivars, agronomically improved varieties and wild relatives from NE India.

The cluster analysis revealed obvious grouping of the *Chakhao* accessions corresponding to different cultivar types. The black-kernelled landraces such as *Chakhao Poireiton* and *Chakhao Amubi* had a close genetic relationship, whereas white-kernelled *Angouba* landraces were genetically similar to the *Buhman* accessions having white or red kernels. Both cluster analysis and PCoA showed a clear grouping of *Poireiton*, *Amubi*, *Angouba* and *Buhman*. The other *Chakhao* cultivars such as *Maklei*, *Napnang Hangmei*, *Chakhao Phou* and *Chakhao* were found to be genetically diverse from the major *Chakhao* cultivars such as *Poireiton*, *Amubi*, *Buhman* and *Angouba*. Population structure is an important component of association analyses between molecular markers and traits. All the three population structure approaches, neighbour-joining phylogeny, PCoA and STRUCTURE analysis, revealed a significant genetic structure within the *Chakhao* accessions.

During the collection trip, attempts were made to collect all the available *Chakhao* cultivars from the state. Multiple accessions of each cultivar with the same name were collected with the idea that each accession is inherently valuable for the conservation of rice germplasm. It was found that the landraces with the same name collected from adjacent areas were often genetically similar and could be considered members of the same cultivar, while this was not true for cultivars with the same name collected from distant areas (Thomson *et al.*, 2009). The current study showed that the farmers, who are the custodians of the traditional landraces, are very efficient at conserving different cultivars within a small geographical region.

Conclusions

In conclusion, the high gene diversity detected among various *Chakhao* cultivars is comparable to that detected in rice populations in various parts of the world. There is also a pronounced genetic structure among different types of *Chakhao*, which could be associated with limited gene flow and introgression with different rice varieties, mainly due to geographical isolation of the state of Manipur. The current study demonstrated that the traditional farmers of Manipur are efficiently preserving several types of aromatic rice varieties within a small geographical region over time. The genetic characterization of these landraces, which are a set of dynamic gene pool of traditional varieties adapted to the local environment and in close association with human cultural preferences, offers an important foundation for conserving these valuable genetic resources and exploiting genetic diversity for improving rice varieties.

Supplementary material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S1479262113000580

Acknowledgements

This study was funded by the Indian Council of Agricultural Research (ICAR). The authors are grateful to the

farmers of Manipur for providing necessary information and seed samples used in this study. They also thank the scientific staff of ICAR Research Complex for NEH Region, Manipur, particularly N Prakash, G Singh and LT Monsang.

References

Borba TCDO, Brondani RPV, Rangei PHN and Brondani C (2009) Microsatellite marker-mediated analysis of the EMBRAPA rice core collection genetic diversity. *Genetics* 137: 293–304.

Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR, Bustamante CD and Purugganan MD (2007) Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS Genetics* 3: e163.

Cavalli-Sforza LL and Edwards AWF (1967) Phylogenetic analysis. Models and estimation procedures. *American Journal of Human Genetics* 19: 233–257.

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. *SpringerPlus* 2: 228.

Cruz DN and Khush GS (2000) Rice grain quality evaluation procedures. In: Singh RK, Singh US and Khush GS (eds) *Aromatic Rices*. New Delhi: Oxford & IBH Co. Pvt. Ltd, pp. 15–28.

Das B, Sengupta S, Parida SK, Roy B, Ghosh M, Prasad M and Ghosh TK (2013) Genetic diversity and population structure of rice landraces from Eastern and North Eastern states of India. *BMC Genetics* 14: 71.

Earl DA and von Holdt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Available at http://taylor0.biology.ucla.edu/structureHarvester/

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.

Gao Li-Zhi, Zhang Chi-Hong, Chang Li-Ping, Jia Ji-Zeng, Qiu Zong-En and Dong Yu-Shen (2005) Microsatellite diversity within *Oryza sativa* with emphasis on *indica–japonica* divergence. *Genetics Research* 85: 1–14.

Garris AJ, Tai TH, Coburn J, Kresovich S and McCouch S (2005) Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169: 1631–1638.

Hamrick JL and Godt MJ (1996) Conservation genetics of endangered plant species. In: Avise JC and Hamrick JL (eds) *Conservation Genetics: Case Histories from Nature*. New York: Chapman & Hall, pp. 281–304.

Jain S, Jain KR and McCouch SR (2004) Genetic analysis of Indian aromatic and quality rice (*Oryza sativa* L.) germplasm using panels of fluorescently-labeled microsatellite markers. *Theoretical and Applied Genetics* 109: 965–977.

Nagaraju M, Mohanty KK, Chaudhary D and Gangadharan CA (1991) Simple technique to detect scent in rice. *Oryza* 28: 109–110.

Pavel AB and Vasile CI (2012) PyElph – a software tool for gel images analysis and phylogenetics. *BMC Bioinformatics* 13: 9.

Peakall R and Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics* 28: 2537–2539.

Prashanth SR, Parani M, Mohanty BP, Talame V, Tuberosa R and Parida A (2002) Genetic diversity in cultivars and landraces of *Oryza sativa* subsp. *indica* as revealed by AFLP markers. *Genome* 45: 451–459.

Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Rohlf FJ (2000) *NTSYSpc 2.1, Numerical Taxonomy and Multivariate Analysis System*. Setauket, NY: Exeter Software.

Roy S, Rathi RS, Misra AK, Bhatt BP and Bhandari DC (2013) Phenotypic characterization of indigenous rice (*Oryza sativa L.*) germplasm collected from the state of Nagaland, India. *Plant Genetic Resources: Characterization and utilization*. DOI 10.1017/S1479262113000282

Thomson MJ, Septiningsih EM, Suwardjo F, Santoso TJ, Silitonga TS and McCouch SR (2007) Genetic diversity analysis of traditional and improved Indonesian rice (*Oryza sativa* L.) germplasm using microsatellite markers. *Theoretical and Applied Genetics* 14: 559–568.

Thomson MJ, Polato NR, Prasetiyono J, Trijatmiko KR, Silitonga TS and McCouch SR (2009) Genetic diversity of isolated populations of Indonesian landraces of rice (*Oryza sativa* L.) collected in East Kalimantan on the Island of Borneo. *Rice* 2: 80–92.

Tu M, Lu BR, Zhu Y and Wang Y (2007) Abundant within-varietal genetic diversity in rice germplasm from Yunnan province of China revealed by SSR fingerprints. *Biochemical Genetics* 45: 789–801.

Xu YB, Beachell H and McCouch SR (2004) A marker-based approach to broadening the genetic base of rice in the USA. *Crop Science* 44: 1947–1959.

Yeh FC, Yang R and Boyle T (1999) *POPGENE. Version 1.31. Microsoft Window-based Freeware for Population Genetic Analysis*. Edmonton, AB: University of Alberta.

Yu SB, Xu WJ, Vijayakumar CHM, Ali J, Fu BY, Xu JL, Jiang YZ, Marghirang R, Domingo J, Aquino C, Virmani SS and Li ZK (2003) Molecular diversity and multilocus organization of the parental lines used in the International Rice Molecular Breeding Program. *Theoretical and Applied Genetics* 108: 131–140.