3 Biotech (2019) 9:217
https://doi.org/10.1007/s13205-019-1757-x

ORIGINAL ARTICLE

The potentiality of rice microsatellite markers in assessment of cross-species transferability and genetic diversity of rice and its wild relatives

Umakanta Ngangkham[1,3] · Sofini Dash[1] · Madhuchhanda Parida[1] · Sanghamitra Samantaray[1] · Devachandra Nongthombam[2] · Manoj Kumar Yadav[1] · Awadhesh Kumar[1] · Parameswaran Chidambaranathan[1] · Jawahar L. Katara[1] · Bhaskar C. Patra[1] · Lotan K. Bose[1]

Received: 1 December 2018 / Accepted: 11 May 2019 / Published online: 20 May 2019

Abstract

The main aim of this study is to assess the potentiality of SSR markers for the identification of the cross-species transferability frequency in a large set of the diverse genome types of wild relative rice along with cultivated rice. Here, we used 18 different rice genotypes representing nine different genome types with 70 SSR markers to investigate the potentiality of cross-species transferability rate. The overall cross-species transferability of SSR markers across the 18 rice genotypes ranged from 38.9% (RM280 and RM447) to 100% (RM490, RM318, RM279, RM18877 and RM20033, RM19303) with an average of 76.58%. Also, cross-species transferability across chromosome ranged from 54.4% (chromosome 4) to 86.5% (chromosome 2) with an average of 74.35%. The polymorphism information content of the markers varied from 0.198 (RM263) to 0.868 (RM510) with a mean of 0.549 ± 0.153, showing high discriminatory power. The highest rate of cross-transferability was observed in *O. rufipogon* (97%), The highest rate of cross-species transferability was in *O. rufipogon* (97.00%), followed by *O. glaberrima* (94.20%), *O. nivara* (92.80%), Swarna (92.80%), *O. longistaminata* (91.40%), *O. eichingeri* (90%), *O. barthii* (88.50%), *O. alta* (82.80%), *O. australiensis* (77.10%), *O. grandiglumis* (74.20%), *O. officinalis* (74.20%), *Zizania latifolia* (70.00%), *O. latifolia* (68.50%), *O. brachyantha* (62.80%), *Leersia perrieri* (57.10%) and *O. ridleyi* (41.40%) with least in *O. coarctata* (28.50%). A total of 341 alleles from 70 loci were detected with the number of alleles per locus ranged from 2 to 12. Based on dendrogram analysis, the AA genome groups was separated as distinct group from the rest of the genome types. Similarly, principal coordinate analysis and structure analysis clearly separated the AA genome type from the rest of the genome types. Through the analysis of molecular variance, more variance (51%) was observed among the individual, whereas less (14%) was observed among the population. Thus, our findings may offer a valuable resource for studying the genetic diversity and relationship to facilitate the understanding of the complex mechanism of the origin and evolutionary processes of different *Oryza* species and wild relative rice.

Keywords Rice · Wild relative rice · SSR markers · Cross-species transferability · *Oryza* · Microsatellites

Introduction

Rice (*Oryza sativa* L.) is the single most important staple food crop for more than half of the world's population (Ngangkham et al. 2010) and provides 20% of the calories consumed by the world population (Huang et al. 2012). However, as per the increasing world population, it is required to enhance the rice grain production up to 50% more by the year 2025 (Khush 2003) to meet the global food demands. Besides, the ever-increasing global population coupled with decreasing arable land and changing global climate demands sustainable rice improvement in terms of

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13205-019-1757-x) contains supplementary material, which is available to authorized users.

✉ Umakanta Ngangkham
ukbiotech@gmail.com

[1] ICAR-National Rice Research Institute, Cuttack, Odisha 753006, India

[2] College of Horticulture and Forestry, Central Agricultural University (Imphal), Pasighat 791102, Arunachal Pradesh, India

[3] Present Address: ICAR-Research Complex for North East Hill Region, Umiam 793103, India

grain yield. Furthermore, rice production and sustainability are continuously threatened by several biotic (pests and pathogens) and abiotic (drought, submergence, alkalinity and salinity) stresses. These stresses are becoming increasingly significant, particularly in the context of global climate change. Thus, to overcome these constraints and ensure continued food security, we need to develop genetically superior rice varieties with higher yield potential, possessing multiple resistances to biotic and abiotic stresses and with more nutritious grain quality for which intervention of modern tools and approaches is essentially required (Ngangkham et al. 2018a, b). Though the development and deployment of rice varieties introgressed with resistance genes are the most cost-effective and environmentally friendly approach (Yadav et al. 2017), the identification and selection of the pre-breeding rice lines with the authenticated desired traits are a prerequisite step. Rice domestication involved a complex process of selection by our ancestors, leading to the development of a new species, *O. sativa*, during the last 10,000 years (Izawa 2008). Nevertheless, it resulted in the selection of a limited number of plants or seeds from wild relative population, thereby causing yield stagnant, progressive susceptible to biotic and abiotic stresses in the generations that follow. Therefore, there is a need to mine out a number of genes from wild rice which are being adapted to diverse geographical environments and known as the excellent reservoir of variability for several traits, including resistance to biotic and abiotic stresses and quality and productivity traits (Patra et al. 2015; Jacquemin et al. 2013).

The genus *Oryza* comprises of two cultivated (*Oryza sativa* L. and *Oryza glaberrima* Steud.) and 22 wild species that categorize them into 10 genome types (Zhang et al. 2014) represented by six diploids (AA, BB, CC, EE, FF and GG) and four tetraploid (BBCC, CCDD, HHJJ and HHKK) genome types (Ge et al. 1999); these belong to the family Poaceae and tribe Oryzeae (Patra et al. 2015). These diverse species of *Oryza* genus facilitate a huge, rich source of genetic resources for the genetic improvement of domesticated rice cultivars and for studying the complex origin and evolutionary processes of wild rice. To tap these huge genetic resources, the International *Oryza* map Alignment Project (I-OMAP) was initiated to develop the reference quality sequences from representative of all 23 *Oryza* species, of which more than eight reference genome sequencing have been completed (Jacquemin et al. 2013). The genetic variability in wild rice populations is highly important for breeding programs and genetic crop improvement for biotic and abiotic stresses (Brondani et al. 2005). For instance, introgression of specific traits from wild to cultivated varieties may result in novel gene combinations for increasing the productivity of the crop and tolerance/resistance to several biotic and abiotic stresses (Ramiah 1953; Jena and Khush 1990). Recently, with the advancement of re-sequencing technology, huge amount of single-nucleotide variants have been generated and deposited in several rice databases, namely dbSNP at NCBI (Sherry et al. 2001), RiceVarMap (Zhao et al. 2014), Gramene (Tello-Ruiz et al. 2015), SNP-Seek (Mansueto et al. 2016), etc. However, SSR (simple sequence repeat) markers are considered as the primary choice for genotyping due to high density, codominant inheritance, high allelic diversity and highly reproducible as compared to the other PCR-based markers (Song et al. 2019). The availability of the complete whole genome sequence of rice has largely facilitated the identification of large numbers of SSR markers covering the entire genome to map a number of agronomically important QTLs in rice, which help in monitoring of alien gene introgression (Patra et al. 2015; Ngangkham et al. 2018b). To date, the use of microsatellite markers on cross-transferability in rice and its wild relatives is meager (Brondani et al. 2003; Gao et al. 2005; Ray et al. 2016). The present study was carried out to assess the potentiality of SSR markers developed from Nipponbare genome, for the identification of the cross-species amplification and transferability frequency in a large set of the diverse genome types of wild relative rice along with cultivated rice. Further, genetic diversity, genetic structure and relationships based on these SSR markers were investigated to understand the complex origin and evolutionary processes of different rice genome types which have been debated and discussed for quite a long time.

Materials and methods

Plant materials

A set of 18 rice accessions including three cultivated rice types and fifteen wild relative rice types representing nine different genome types (AA, CC, EE, FF, CCDD, HHJJ, KKLL and two NOWRs: non-*Oryza* wild rice) obtained from the National Gene Bank, ICAR-National Rice Research Institute, Cuttack, India, which is located in Odisha (85° 55′48″ E longitudes and 20° 26′35″ N latitude), were selected. The detailed list of the rice accessions used is given in Table 1. The leaf samples were collected, frozen in liquid nitrogen and stored at − 80 °C for further use of genomic DNA isolation.

Selection and scoring of SSR markers

A set of total 105 SSR markers that are evenly distributed across the whole 12 chromosomes of rice were collected from the Gramene marker database (http://archive.gramene.org/markers/) developed from Nipponbare (*japonica* rice) genome sequencing to test the transferability of the marker across the selected wild relative rice. The detail information

Table 1 List of 18 rice genotypes used in the present study

Sl no.	Samples	Genome types	Types
1	*Oryza sativa* sp. *Japonica* cultivar Nipponbare	AA	Asian cultivated
2	*Oryza sativa* sp. *Indica* cultivar Swarna	AA	Asian cultivated
3	*Oryza glaberrima*	AA	African cultivated
4	*Oryza nivara*	AA	Indian wild rice
5	*Oryza rufipogon*	AA	Wild relative rice
6	*Oryza longistaminata*	AA	Wild relative rice
7	*Oryza barthii*	AA	Wild relative rice
8	*Oryza eichingeri*	CC	Wild relative rice
9	*Oryza officinalis*	CC	Wild relative rice
10	*Oryza australiensis*	EE	Wild relative rice
11	*Oryza brachyantha*	FF	Wild relative rice
12	*Oryza latifolia*	CCDD	Wild relative rice
13	*Oryza alta*	CCDD	Wild relative rice
14	*Oryza grandiglumis*	CCDD	Wild relative rice
15	*Oryza ridleyi*	HHJJ	Wild relative rice
16	*Zizania latifolia*	NOWR[a]	Wild rice
17	*Leersia perrieri*	NOWR[a]	Wild rice
18	*Oryza coarctata*	KKLL	Wild relative rice

[a]NOWR—non-Oryza wild rice

about the 105 SSR markers used in the present study is given in Table S1. PCR amplification of specific microsatellite markers in different genotypes was considered to be successful when sharp bands in the expected size range were detected, or otherwise the absence of specific PCR band ranges was considered null allele of SSR marker. The analysis of cross-species amplification of SSR maskers was carried out by determining the expected range of PCR products in the other genome. We further evaluated the utility of two contrast transferability marker rates by comparing between highest transferability (HT) and lowest transferability (LT) of SSR markers using a set of 96 landraces of AA genome types.

Genomic DNA isolation and PCR amplification

Genomic DNA was isolated from young leaf tissues using the cetyl trimethyl ammonium bromide (CTAB) method as described by Ngangkham et al. (2018b). The purified genomic DNA was diluted with nuclease-free water to the working concentration of 20 ng/µl for PCR amplification. The PCR mixture was prepared containing of 1× T*aq* buffer (10 mM Tris–HCl, 50 mM KCl, pH 8.3), 0.2 µM of each forward and reverse primers, 1.5 mM $MgCl_2$, 0.2 µM of each of dNTP, 20 ng template DNA and 1U of T*aq* DNA polymerase enzyme (DreamTaq, Thermo Scientific, USA). The PCR cycle program was set up by maintaining 94 °C of 5 min for initial denaturation followed by 35 cycles of 94 °C for 30 s, primers annealing for 30 s at 55 °C and elongation for 1 min at 72 °C, which is followed by a final elongation at 72 °C for 10 min. The PCR bands were separated by gel electrophoresis in 3.5 or 4% agarose gels (Lonza, USA) along with a 100-bp DNA ladder (DreamTaq, Thermo Scientific, USA) stained with ethidium bromide. The PCR bands were scored by documenting the image using a gel documentation system (Alpha Imager, USA).

Statistical analysis of scoring data of SSR markers

The scored genotypic data of 70 SSR markers were used for estimation of genetic diversity parameters such as the number of alleles per locus (*K*), observed heterozygosity (H_O), expected heterozygosity (H_E), polymorphism information content (PIC) and deviations from Hardy–Weinberg equilibrium (HWE) using the Cervus 3.0 program (Field Genetics Ltd, London, England). The Shannon's information index (*I*) and the effective number of alleles per locus (Ne) were calculated using POPGENE 32 (Yeh 1997). The patterns of allelic richness, private allele frequency and F_{ST} (genetic differentiation) were calculated using GenAlEx 6.502 software (Peakall and Smouse 2012). The GenAlEx 6.502 software was also used to compare the mean of genetic diversity parameters for eight genome types and to estimate the pair-wise F_{ST} to partition genetic diversity and compute the PCoA (Principle coordinate analysis) by plotting the eigenvector values in a scatter graph taking the first principal component and the second principal component as the axes. A hierarchical analysis of molecular variance (AMOVA) of rice accessions between the cultivated and wild relative rice was conducted using GenAlEx 6.502 software (Peakall and Smouse 2012). The genetic relationship analysis was conducted by estimating genetic distance and similarity coefficients. An unweighted neighbor joining un-rooted tree was constructed by the calculated NEI coefficient of dissimilarity index (Nei 1972) with a bootstrap value of 1000 using DARwin6 software (Perrier and Jacquemoud-Collet 2006).

The program STRUCTURE version 2.3.4 (Pritchard et al. 2000) was adopted to determine the presence of genetic structure with the number of clusters was run from $K=1$ to $K=10$, with 5 independent replications per *K* using the admixture model and correlated allele frequencies, a 200,000 burn-in period and 200,000 MCMC to determine approximations of posterior distributions. Then, the optimum *K* value was determined by estimating the peak value of ΔK using Evanno et al. (2005) method through the STRU CTURE HARVESTER program (Earl 2012). For each value of *K*, STRUCTURE produces a Q-matrix (QST) that lists the estimated membership coefficients for each genotype in each subgroup. An individual having more than 70% of its

genome fraction value under a particular *K* subgroup was assigned to that subgroup.

Results and discussion

The genetic diversity existing in the cultivated rice gene pool has been reportedly reduced drastically by domestication and long-term cultivation of a few varieties as compared to the wild species of rice (Krishnan et al. 2014; Song et al. 2003; Tanksley and McCouch 1997). However, a lot of efforts have been made to improve yield, grain quality and other agronomical traits by introgression of the alien genes from the secondary rice gene pool that are largely untapped yet to the cultivated rice through wide hybridization by rice breeders (Shakiba and Eizenga 2014; Vaughan et al. 2003; Ray et al. 2016). These efforts have been further facilitated by the development of large number of DNA-based markers to help in monitoring alien genome introgression from the wild to the cultivated species. In this study, an effort was made to identify the SSR markers that can be used to study introgression from a large set of diverse wild relative rice to the cultivated rice (*O. sativa* L.) by estimating the marker cross-species transferability.

Selection of SSR markers and Genotyping

PCR amplification of a specific microsatellite marker was considered to be useful marker only when sharp bands in the expected size range of cultivated *japonica* rice (Nipponbare) was detected, whereas failure to be amplified was ignored for the cross-species transferability analysis, which was used in the present studies. Initially, a set of 105 SSR markers was selected randomly from the 12 chromosomes of rice with the average of 8.83 markers per chromosome and range of 6 to 13 SSR markers per chromosomes. The highest number of markers was from chromosome 6 (13), followed by chromosomes 3 and 11 (11), chromosomes 5, 9 and 12 (9), chromosomes 2 and 4 (8), chromosomes 1, 7 and 8 (7) and chromosome 10 (6). Out of 105 SSR markers used for preliminary amplification test in Nipponbare DNA, only 70 SSR markers (66.67%) were found to be amplified with sharp bands of the expected size; hence, these 70 SSR markers were used for further studies (Fig. 1, Table S1). The remaining markers were not tried for PCR optimization as 70 SSR markers were considered enough for cross-species amplification analysis compared with other earlier reports. The range of such successful SSR amplification rate in the present studies is similar to the result reported by other researchers (Wang et al. 2014; Zheng et al. 2013). The mean percentage of SSR amplification ranged from 0% (chromosome 10) to 100% (chromosome 1). The highest percentage of amplified markers was found in chromosome 1 (100%) followed by chromosome 6 (92.30%), chromosome 2 (87.5%), chromosome 7 (85.71%), whereas > 50% amplification was observed in chromosomes 3, 5, 8 and 9. The chromosomes which showed < 50% amplification of SSR markers were found to be in chromosomes 4, 11 and 12. However, there was no amplification of any markers on chromosome 10; so, this was not included in the cross-transferability analysis. Since these SSR markers are located in both coding and non-coding regions of the genome, they would be useful for understanding the genetic evolutionary processes, genetic conservation studies and introgression of

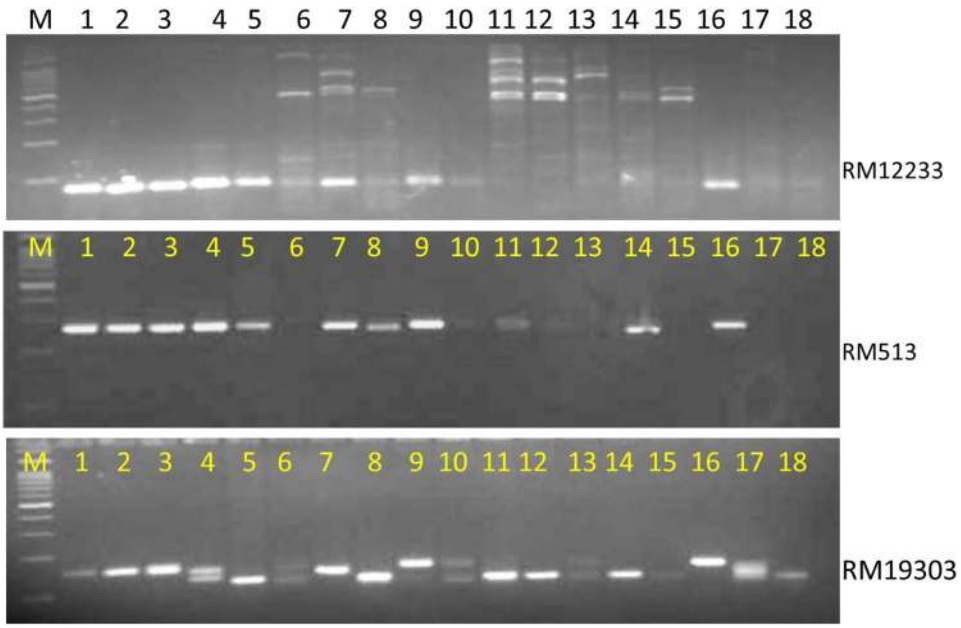

Fig. 1 Representatives of PCR amplified fragments of SSR markers in 18 rice genotypes. Lane: M1:100-bp DNA ladder, 1–18 represents rice germplasm listed in Table 1. Markers are shown in the right side of the figure

alien genes in the cultivated rice from the wild-related species for genetic improvement of rice breeding.

The PIC of the markers refers to the relative value of each marker regarding the amount of polymorphism exhibited, ranged from 0.198 (RM263) to 0.868 (RM510) with a mean of 0.549 ± 0.153 which demonstrates the high discriminatory power of these markers (Table 2). Further, 70 SSR markers showing high level of polymorphism would be effectively utilized in fingerprinting, genetic diversity, genetic conservation, etc. All the selected 70 SSR markers were polymorphic with 100% polymorphic loci revealed a high level of genetic diversity among the 18 rice accessions. However, 42 SSR markers (60%) were found to be more informative by exhibiting *PIC* value > 0.5 which could also be useful to determine genetic diversity, gene introgression, genetic conservation, etc. This high level of polymorphism was very similar to polymorphism level obtained in *Paspalum* (Wang et al. 2006) and bread wheat (Gupta et al. 2003). However, the level of polymorphism detection depends on the genetic divergence of the sample studied and such high level of polymorphism in the present study could be explained partly by the presence of polyploidy (tetraploid) and some polymorphic loci might not be homolocus due to mis-primings across species (Wang et al. 2006).

Cross-transferability of SSR markers across the genomes

Though SSR markers show their potentiality in studying the cross-transferability among wild relative rice and cultivated rice, a few information is available in the identification of such SSR markers where only one or few diverse genome were used (Ray et al. 2016; Li et al. 2008; Gao et al. 2005; Brondani et al. 2003). However, in this study, a total of fifteen diverse wild rice relatives with distinct genome types along with all the cultivated rice species (*indica, japonica* and *glaberrima*) were taken using a large set of SSR markers covering the entire genome for the first time. Besides, SSR markers was developed for cross-species transferability analysis using cultivated rice genome sequence data (Nipponbare) which could help in gene introgression, genome evolution and genome relationships among the *Oryza* species. Since development of DNA-based markers for a particular crop is expensive and time-consuming, the study of cross-species transferability of markers in related genera/species is a viable alternative for immediate intervention for genetic improvement of crops. The success of cross-species transferability of markers depends on the kind of markers selected along with evolutionary relatedness of the species selected (Dayanandan et al. 1997). Here, transferability of SSR marker was carried out to determine the transferability of SSR markers. Since the SSR markers were designed using Nipponbare genome database and randomly selected from the 12 chromosomes covering the whole genome of rice, occurrence of ascertainment bias in the present study was avoided (Gao et al. 2005). The cross-species amplification of SSR markers across the 18 rice genotypes which was estimated as the total number of SSR markers amplified in a particular genome type was ranged from 38.90% (RM280 and RM447) to 100% (RM490, RM318, RM279, RM18877 and RM20033, RM19303) with an average of 76.58% which is corroborated to the other reports (Wang et al. 2014; Alexander et al. 2018; Hernández et al. 2001; Gupta et al. 2003); this range was found to be higher as compared to the *Melilotus* with 61.40% (Yan et al. 2017), *Paspalum* with 61% (Wang et al. 2006), Siberian wildrye with 49.11% (Zhou et al. 2016) and *Cucumis* with 12.70% (Fernandez-Silva 2008). Such differences in the rate of cross-species transferability could be explained by genetic relatedness and PCR primers used (Wang et al. 2006). The overall cross-species transferability of SSR marker which is the rate of amplification of markers across the genome types varied from 54.40% (chromosome 4) to 86.50% (chromosome 2) chromosome-wise with an average of 74.35% (Table 3). The lowest cross-transferability markers such as RM280 and RM447 are localized on chromosome 4 and chromosome 8, respectively. On the other hand, the highest cross-transferability markers were found to be six in number, viz. RM490, RM318, RM279, RM18877, RM20033 and RM19303, and found to be amplified in all the 18 genotypes with 100% transferability. The RM490 are localized at 6.67 Mb on chromosome 1, RM318 at 29.6 Mb and RM279 at 2.88 Mb on chromosome 2, RM18877 at 23.57 Mb on chromosome 5, RM20033 at 29.52 Mb and RM19303 at 27.41 Mb on chromosome 6. These markers would be quite useful as a source of genetic markers of other distantly related rice genus and simultaneously help in studying the gene introgression, genetic mapping, evolutionary, phylogenetic, etc. The polymorphism information content (PIC), expected heterozygosity (H_E) and the number of allele (K) of the HT markers showed 0.563, 0.633 and 4, respectively, whereas in the case of LT, these three were observed as 0.460, 0.559 and 3, respectively. This comparison result of the two marker groups indicates that these SSR markers would yield similar information for polymorphism in cultivated rice, which would be useful for genetic improvement of cultivated rice as they have distinct transferability among the genome groups of wild relative rice. These markers with wide transferability would be highly valuable for developing the molecular discrimination of *Oryza* species and wild relative rice which is still difficult to differentiate them based on morphological and physiological traits.

The highest rate of cross-species transferability of 70 SSR markers among the different genotypes or genomes was in *O. rufipogon* (97.00%), followed by *O. glaberrima* (94.20%), *O. nivara* (92.80%), Swarna (92.80%), *O.*

Table 2 Genetic diversity analysis for 70 SSR markers in 18 rice genotypes

Locus	*K*	H_O	H_E	PIC	Ne	*I*	F_{ST}
RM490	3	0.778	0.552	0.441	2.160	0.854	0.158
RM529	3	0.067	0.384	0.323	1.590	0.639	0.725
RM1	4	0.357	0.458	0.392	1.790	0.808	0.707
RM595	3	0.188	0.373	0.327	1.566	0.656	0.609
RM513	2	0	0.331	0.269	1.471	0.500	0.817
RM12233	2	0.471	0.371	0.295	1.562	0.546	0.506
RM1349	3	0.2	0.542	0.424	2.062	0.824	0.853
RM263	3	0.077	0.218	0.198	1.266	0.431	0.930
RM318	2	0.722	0.513	0.374	1.994	0.692	0.128
RM110	7	0.6	0.747	0.694	3.600	1.578	0.559
RM12634	7	0.25	0.766	0.707	3.879	1.576	0.656
RM279	6	0.833	0.794	0.735	4.378	1.577	0.438
RM550	5	0.571	0.526	0.478	2.031	1.037	0.623
RM3586	7	0.714	0.778	0.716	4.000	1.607	0.284
RM135	3	0.538	0.532	0.458	2.049	0.883	0.694
RM16238	6	0.3	0.7	0.609	2.985	1.326	0.739
RM15203	4	0.8	0.737	0.645	3.333	1.280	0.688
RM13709	6	0.357	0.534	0.493	2.063	1.118	0.752
RM15809	3	0.313	0.615	0.522	2.473	0.990	0.721
RM14368	5	0.111	0.804	0.719	4.154	1.494	0.792
RM15539	3	0.455	0.394	0.344	1.603	0.689	0.677
RM15490	6	0.643	0.762	0.692	3.769	1.488	0.665
RM168	8	0.412	0.786	0.729	4.219	1.674	0.542
RM280	6	1	0.835	0.744	4.455	1.631	0.688
RM16925	6	0.375	0.476	0.437	1.855	0.996	0.559
RM17349	4	0.375	0.642	0.547	2.510	1.103	0.807
RM480	6	0.5	0.786	0.717	4.056	1.560	0.496
RM163	5	0.188	0.655	0.585	2.738	1.224	0.844
RM574	2	0	0.508	0.371	1.969	0.685*	1.000
RM1248	3	0.25	0.575	0.474	2.256	0.913	0.599
RM18983	4	0.375	0.692	0.582	2.844	1.163	0.734
RM18948	4	0.294	0.704	0.635	3.159	1.268	0.540
RM18877	5	0	0.692	0.617	3.057	1.292**	0.535
RM527	3	0	0.556	0.447	2.170	0.865*	0.686
RM136	3	0.059	0.508	0.397	1.973	0.769	0.734
RM20724	6	0.5	0.817	0.733	4.267	1.602	0.785
RM528	7	0.375	0.746	0.68	3.606	1.520	0.657
RM20316	8	0.313	0.724	0.675	3.346	1.579	0.606
RM20033	7	0.278	0.673	0.63	2.893	1.452	0.556
RM20077	5	0.059	0.658	0.59	2.766	1.229**	0.727
RM20388	9	0.647	0.715	0.664	3.266	1.587	0.463
RM20733	5	0	0.727	0.649	3.270	1.367	0.867
RM19303	4	0	0.584	0.522	2.314	1.059	0.721
RM20704	6	0.824	0.624	0.571	2.535	1.254	0.221
RM510	12	0.615	0.914	0.868	8.244	2.291	0.616
RM4986	4	0.556	0.739	0.642	3.306	1.276	0.642
RM336	8	0.333	0.752	0.689	3.447	1.638	0.759
RM22085	6	0.5	0.675	0.599	2.723	1.331	0.804
RM21131	5	0.6	0.492	0.441	1.907	0.956	0.646
RM21778	3	0.059	0.651	0.557	2.714	1.045**	0.689

Table 2 (continued)

Locus	*K*	H_O	H_E	PIC	Ne	*I*	F_{ST}
RM22060	4	0.4	0.553	0.48	2.105	0.982	0.754
RM447	4	0.429	0.582	0.502	2.178	1.029	0.787
RM404	6	0.364	0.732	0.656	3.315	1.430	0.669
RM22914	3	0.235	0.519	0.437	2.014	0.848	0.536
RM419	2	0	0.423	0.325	1.690	0.598	0.776
RM195	7	0.583	0.692	0.636	2.969	1.456	0.618
RM23901	8	0.462	0.868	0.813	6.036	1.911	0.624
RM566	3	0	0.542	0.436	2.110	0.846	0.746
RM296	3	0.071	0.553	0.424	2.142	0.822*	0.725
RM205	3	0.182	0.623	0.526	2.469	0.995	0.829
RM5899	7	0.615	0.782	0.716	4.024	1.596	0.598
RM24033	10	0.875	0.873	0.829	6.481	2.054	0.476
RM26333	3	0.067	0.393	0.342	1.613	0.683	0.813
RM26423	5	0.417	0.587	0.525	2.286	1.117	0.739
RM27124	4	0.143	0.267	0.246	1.347	0.559	0.762
RM224	5	0.8	0.784	0.718	4.128	1.495	0.408
RM27099	6	0.214	0.648	0.592	2.667	1.302	0.805
RM28207	2	0.077	0.409	0.316	1.649	0.583	0.825
RM27691	5	0.4	0.701	0.627	3.103	1.301	0.675
RM28011	4	0.1	0.721	0.632	3.175	1.258	0.781
Mean ± SD	4.871 ± 2.063	0.361 ± 0.259	0.623 ± 0.152	0.549 ± 0.153	2.873 ± 1.220	1.168 ± 0.395	0.660 ± 0.020

Mean ± SD = Means followed by standard deviations; *K* = number of alleles per locus; H_o = observed heterozygosity; H_E = expected heterozygosity; PIC = polymorphism information content; Ne = effective number of alleles per locus; *I* = Shannon's information index; *F* = genetic differentiation

* and ** Significant deviations from the Hardy–Weinberg equilibrium at *P* value < 0.05 and < 0.01, respectively

longistaminata (91.40%), *O. eichingeri* (90%), *O. barthii* (88.50%), *O. alta* (82.80%), *O. australiensis* (77.10%), *O. grandiglumis* (74.20%), *O. officinalis* (74.20%), *Zizania latifolia* (70.00%), *O. latifolia* (68.50%), *O. brachyantha* (62.80%), *Leersia perrieri* (57.10%) and *O. ridleyi* (41.40%). The least rate of cross-transferability of 70 SSR markers was found to be in the most distantly related *O. coarctata*, a salt-tolerant wild rice growing in coastal area with only 20 SSR markers amplification (28.50%) (Table 3). The cross-species amplification of 70 SSR markers among the seven AA genome types was found to be highest with a range of 88.50% (*O. barthii*) to 100% (Nipponbare) with an average of 93.82%. The extensive cross-species amplification of SSR in *O. rufipogon* (97.00%) exhibited maximum allele sharing with the cultivated rice which also supports the evolution of cultivated rice, *O. sativa* from the commonly recognized wild progenitor, *O. rufipogon.* Though a similar percentage of primer conservations have already been reported in earlier studies in AA rice genome types (Gao et al. 2005; Wu and Tanksley 1993, Panaud et al. 1996), the number of primers used in the present study is comparatively more. Such high rates of cross-transferability among the AA genome types indicated the close relationship within the species of the *Oryza* genus. In case of more distantly related CC genome rice group, represented here as *O. eichingeri* and *O. officinalis,* a slight decrease in cross-species amplification with an average of 82.10% was observed. This is similar to the earlier report (Panaud et al. 1996) where the average cross-species amplification was 84% using 25 microsatellite loci in four rice genotypes: *O. rufipogon, O. nivara, O. officinalis and O. glaberrima*. Subsequently, Li et al. (2008) also observed such high cross-transferability (84.60%) between the *Oryza sativa* and *O. officinalis* revealing highly conservative and good synteny between the homologous chromosomes of these two species (Tan et al. 2005). Their observation along with the our present cross-species amplification rate (82%) in CC genome reassures the evolutionary closeness and conservation between the AA and CC genomes as compared to other genome type (Tan et al. 2005; Li et al. 2008). However, the difference in the success of cross-species transferability may be further influenced by the types of DNA markers used which was observed by Ray et al. (2016) after carrying out the cross-transferability analysis between the *O. sativa* and *O. brachyantha* and found drastic difference in cross-transferability between STS and CAPS markers (84.78%) and SSR (0.75%). Moreover, it was observed that the success

Table 3 Analysis of cross-transferability of SSR markers across different rice genomes

Markers	AA								CC			CCDD			
	Nippon-bare	Swarna	*O. nivara*	*O. rufipogan*	*O. barthii*	*O longis-taminata*	*O. glaber-rima*	Transfer-ability (%)	*O. eich-ingeri*	*O. offici-nalis*	Transfer-ability (%)	*O. lati-folia*	*O. alta*	*O. grandi-glumis*	Transfer-ability (%)
RM490	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM529	+	−	+	+	+	+	+	85.7	−	−	0	+	+	+	100
RM1	+	+	+	+	+	+	+	100	+	+	100	+	+	−	66.7
RM595	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM513	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM12233	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM1349	+	+	+	+	−	+	+	85.7	+	−	50	+	−	−	33.3
RM263	+	+	+	+	+	+	+	100	+	−	50	+	+	+	100
RM318	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM110	+	+	+	+	+	+	+	100	+	−	50	−	+	+	66.7
RM12634	+	+	+	+	+	+	+	100	+	+	100	−	+	+	66.7
RM279	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM13709	+	+	+	+	+	+	+	100	+	−	50	+	+	+	100
RM550	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM3586	+	+	+	+	+	+	+	100	−	+	50	+	+	+	100
RM135	+	+	+	+	+	−	+	85.7	+	+	100	−	+	−	33.3
RM16238	+	+	+	+	+	+	+	100	+	−	50	−	+	−	33.3
RM15203	+	−	−	+	−	−	+	42.9	+	+	100	+	+	+	100
RM15809	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM14368	+	+	+	+	−	+	+	85.7	+	+	100	−	+	−	33.3
RM15539	+	+	+	+	+	+	+	100	+	−	50	+	−	−	33.3
RM15490	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM168	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM280	+	+	−	+	+	+	+	85.7	−	−	0	−	−	−	0
RM16925	+	+	−	+	+	+	+	85.7	+	+	100	+	+	+	100
RM17349	+	+	+	+	+	+	+	100	+	−	50	−	−	−	0
RM480	+	−	+	+	+	−	−	57.1	−	+	50	+	−	−	33.3
RM163	+	+	+	+	+	+	+	100	+	−	50	+	+	+	100
RM574	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM1248	+	+	+	+	+	+	+	100	+	+	100	−	+	+	66.7
RM18983	+	+	+	+	−	+	−	71.4	+	−	50	−	−	+	33.3
RM18948	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM18877	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100

Table 3 (continued)

Markers	AA								CC			CCDD			
	Nippon-bare	Swarna	*O. nivara*	*O. rufipogan*	*O. barthii*	*O longis-taminata*	*O. glaber-rima*	Transfer-ability (%)	*O. eich-ingeri*	*O. offici-nalis*	Transfer-ability (%)	*O. lati-folia*	*O. alta*	*O. grandi-glumis*	Transfer-ability (%)
RM527	+	+	+	+	+	−	+	85.7	+	+	100	+	+	+	100
RM136	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM20724	+	−	−	−	+	−	+	85.7	+	−	50	−	+	+	66.7
RM528	+	+	+	+	+	+	+	100	+	+	100	−	+	+	66.7
RM20316	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM20033	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM20077	+	+	+	+	+	+	+	100	+	+	100	+	−	+	66.7
RM20388	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM20733	+	+	+	+	+	+	+	100	+	+	100	+	−	−	33.3
RM19303	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM20704	+	+	+	+	−	+	+	85.0	+	+	100	+	+	+	100
RM510	+	+	+	−	+	+	+	85.7	+	+	100	+	+	+	100
RM4986	+	+	+	+	+	+	+	100	−	−	0	−	+	−	33.3
RM336	+	+	+	+	+	+	−	85.7	−	−	0	−	−	−	0
RM22085	+	+	+	+	+	+	+	100	+	−	50	−	−	−	0
RM21131	+	+	+	+	+	+	+	100	+	+	100	−	+	+	33.3
RM21778	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM22060	+	+	+	+	−	+	+	85.7	+	−	50	+	−	−	33.3
RM447	+	−	+	+	−	−	−	42.9	+	+	100	−	+	+	66.7
RM404	+	+	+	+	+	+	+	100	+	+	100	−	+	−	33.3
RM22914	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM419	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM195	+	+	+	+	+	+	+	100	+	+	100	−	+	−	33.3
RM23901	+	+	+	+	+	+	+	100	+	+	100	+	−	−	33.3
RM566	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM296	+	+	+	+	+	+	+	100	+	+	100	−	+	+	66.7
RM205	+	+	+	+	+	+	+	100	+	+	100	−	+	+	66.7
RM5899	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM24033	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM26333	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM26423	+	+	+	+	+	+	+	100	+	+	100	−	+	+	66.7
RM27124	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM224	+	+	−	+	+	+	+	85.7	+	+	100	+	+	+	100

Table 3 (continued)

Markers	AA								CC			CCDD			
	Nippon-bare	Swarna	*O. nivara*	*O. rufipogan*	*O. barthii*	*O longis-taminata*	*O. glaber-rima*	Transfer-ability (%)	*O. eich-ingeri*	*O. offici-nalis*	Transfer-ability (%)	*O. lati-folia*	*O. alta*	*O. grandi-glumis*	Transfer-ability (%)
RM27099	+	+	+	+	−	+	+	85.7	+	−	50	+	+	+	100
RM28207	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM27691	+	+	+	+	+	+	+	100	+	+	100	+	+	+	100
RM28011	+	+	+	+	+	+	+	100	−	−	0	−	+	−	33.3
Cross-species amplification (%)	70 (100%)	65 (92.8%)	65 (92.8%)	68 (97.1%)	62 (88.5%)	64 (91.4%)	66 (94.2%)		63 (90%)	52 (74.2%)		48 (68.5%)	58 (82.8%)	52 (74.2%)	

Markers	*O. australiensis* (EE)	*O. brachyantha* (FF)	*O. ridleyi* (HHJJ)	*O. coarctata* (KKLL)	*Z. latifolia* (NOWR)	*L. perrieri* (NOWR)	Overall cross-transferability (%)	
							Marker-wise	Chromosome-wise
RM490	+	+	+	+	+	+	100.0	Chr. 01 (83.33)
RM529	+	+	−	−	+	−	83.3	
RM1	+	+	−	−	−	+	77.8	
RM595	+	+	+	−	+	−	88.9	
RM513	−	+	+	−	+	−	83.3	
RM12233	+	+	+	−	+	+	94.4	
RM1349	−	+	−	−	−	+	55.6	
RM263	+	+	−	−	+	−	77.8	Chr. 02 (86.5)
RM318	+	+	+	+	+	+	100.0	
RM110	+	+	−	+	+	+	83.3	
RM12634	+	+	−	+	+	+	88.9	
RM279	+	+	+	+	+	+	100.0	
RM13709	+	−	−	−	+	+	77.8	
RM550	−	−	−	−	+	+	77.8	

Table 3 (continued)

Markers	*O. australiensis* (EE)	*O. brachyantha* (FF)	*O. ridleyi* (HHJJ)	*O. coarctata* (KKLL)	*Z. latifolia* (NOWR)	*L. perrieri* (NOWR)	Overall cross-transferability (%)	
							Marker-wise	Chromosome-wise
RM3586	−	−	+	+	−	+	77.8	Chr. 03 (68.52)
RM135	+	+	+	−	+	+	77.8	
RM16238	−	+	−−	−	−	−	55.6	
RM15203	+	−	−	−	+	−	55.6	
RM15809	+	+	+	−	+	−	88.9	
RM14368	−	−	−	−	−	−	50.0	
RM15539	−	+	−	−	+	−	61.1	
RM15490	−	+	−	−	+	−	55.6	
RM168	+	+	+	−	+	+	94.4	
RM280	+	+	−	−	−	−	38.9	Chr. 04 (57.4)
RM16925	+	+	−	+	+	+	88.9	
RM17349	−	−	−	−	−	−	44.4	
RM480	+	−	+	+	+	+	61.1	Chr. 05 (80.95)
RM163	+	−	+	+	+	+	88.9	
RM574	+	+	+	−	+	−	88.9	
RM1248	+	−	+	+	+	+	88.9	
RM18983	−	−	+	−	−	−	44.4	
RM18948	+	+	+	−	+	+	94.4	
RM18877	+	+	+	+	+	+	100.0	
RM527	+	+	+	−	+	+	88.9	Chr. 06 (85.64)
RM136	+	+	+	+	+	−	94.4	
RM20724	+	−	−	−	−	+	44.4	
RM528	+	+	+	−	+	+	88.9	
RM20316	+	+	+	−	+	−	88.9	
RM20033	+	+	+	+	+	+	100.0	
RM20077	+	+	+	+	+	+	94.4	
RM20388	+	+	−	+	+	+	94.4	
RM20733	−	−	−	−	−	+	61.1	
RM19303	+	+	+	+	+	+	100.0	
RM20704	+	+	+	+	+	+	94.4	
RM510	+	−	−	−	+	+	77.8	

Table 3 (continued)

Markers	*O. australiensis* (EE)	*O. brachyantha* (FF)	*O. ridleyi* (HHJJ)	*O. coarctata* (KKLL)	*Z. latifolia* (NOWR)	*L. perrieri* (NOWR)	Overall cross-transferability (%)	
							Marker-wise	Chromosome-wise
RM4986	+	−	−	−	−	−	50.0	Chr. 07 (62.96)
RM336	+	+	−	−	−	+	50.0	
RM22085	−	−	−	−	−	−	44.4	
RM21131	+	+	−	−	+	+	83.3	
RM21778	+	+	−	+	+	+	94.4	
RM22060	−	−	+	−	+	−	55.6	
RM447	+	−	−	−	−	−	38.9	Chr. 08 (67.78)
RM404	+	−	−	−	−	−	61.1	
RM22914	+	+	+	−	+	+	94.4	
RM419	+	+	−	−	−	−	77.8	
RM195	+	−	−	−	+	−	66.7	
RM23901	+	+	−	−	−	+	72.2	Chr. 09 (80.55)
RM566	+	+	+	−	+	+	94.4	
RM296	+	+	−	−	−	+	77.8	
RM205	+	−	−	−	−	−	61.1	
RM5899	+	−	−	−	+	+	83.3	
RM24033	+	+	−	+	+	+	94.4	
RM26333	+	+	−	−	+	−	83.3	Chr. 11 (77.77)
RM26423	+	−	−	−	−	−	66.7	
RM27124	−	+	−	−	+	−	77.8	
RM224	+	−	−	+	+	+	83.3	
RM27099	+	+	−	−	+	+	77.8	Chr. 12 (69.45)
RM28207	−	−	−	−	+	−	72.2	
RM27691	−	−	+	−	+	−	77.8	
RM28011	+	−	−	−	−	−	50.0	
Cross-species amplification (%)	54 (77.1%)	44 (62.8%)	29 (41.4%)	20 (28.5%)	49 (70%)	40 (57.1%)		

rate of cross-species transferability gradually declines with the increasing genetic distances among the species. Therefore, in most distantly related tetraploid genome CCDD group of the three species, *O. latifolia, O. alta* and *O. grandiglumis* showed a range of 68.50% to 82.80% with an average of 75.16% cross-species amplification which indicates the degree of relative distinctness of this tetraploid genome from the AA genome of rice.

Among the 70 SSR markers, five SSR markers, namely RM529, RM280, RM4986, RM336 and RM28011, did not amplify in the CC genome groups. In case of the CCDD genome group, three SSR markers, namely RM280, RM336 and RM22085, failed to amplify. Similarly, there were many unique SSR markers in different genome types like EE, FF, HHJJ, KKLL, etc., which would be immediately useful to differentiate the rice genome types as well as rice species, genetic diversity, genetic relationships, biodiversity conservation and effectively further utilized for the identification of rice genome types. These SSR markers developed from Nipponbare genotype have served as excellent sources for markers across the *Oryza* species and NOWR.

Genetic diversity parameters analysis

The genetic variability assessment in wild rice populations is highly important for breeding programs and genetic improvement of crop for biotic and abiotic stresses (Brondani et al. 2005). For instance, introgression of specific traits from wild to cultivated varieties may result in novel gene combinations which may increase the productivity of the crop and tolerance/resistance to several biotic and abiotic stresses. The present study was not aimed at investigating the genetic diversity; a total of 341 alleles from 70 loci were detected (Table 2). The number of alleles per locus (K) ranged from 2 to 12 with an average of 4.871 ± 2.063 which was found to be similar to previous reports on *Oresitrophe* and *Mukdenia* (Liu et al. 2018) but lower than *O. glaberrima* (Yelome et al. 2018), *Melilotus* (Yan et al. 2017) and higher from the results reported in rice (Wang et al. 2014). The RM510 marker mapped at 2.83 Mb on chromosome 6 showed the highest number of alleles (12) with mean observed and expected heterozygosity were found to be 0.361 ± 0.259 (H_O) and 0.623 ± 0.152 (H_E), respectively. The genetic diversity pattern based on H_E in the present study was obtained as moderate which are similar to the other reports in rice germplasm (Choudhury et al. 2013; Nachimuthu et al. 2015), outcrossing plant species ($H_E = 0.65$) (Liu et al. 2018; Nybom 2004; Yuan et al. 2014). Such moderate level of genetic diversity pattern detected from 18 rice accessions may relate to diverse or distant genome types, ploidy level, inbreeding which is maintained by its mode of reproduction, self-pollination along with cross-incompatibility. The maximum number of SSR markers (81.42%) showed less value of observed heterozygosity (H_O) than their corresponding expected heterozygosity (H_E) which might be due to self-pollination.

The effective number of alleles per locus (Ne) showed a range of 1.266 (RM263) to 8.244 (RM510) with an average of 2.873 ± 1.220. The Shannon's information index (*I*) ranged from 0.431 (RM263) to 2.291 (RM510) with a mean of 1.168 ± 0.395. Significant deviation from the Hardy–Weinberg equilibrium (P value < 0.05) was observed for the six markers such as RM574, RM18877, RM527, RM20077, RM21778 and RM296 (Table 2) yielding extreme low heterozygosity ($H_O < 0.05$) which might be due to deficit of heterozygotes. The genetic differentiation for 70 SSR markers ranged from 0.128 (RM318) to 1.0 (RM574) with an average of 0.660 ± 0.020.

The results of the comparison in mean genetic diversity parameters among the eight genomes groups using these 70 SSR markers showed wide variations (Table 4, Figure S1). The highest *N* (6.243), Na (3.086), Ne (2.284) and *I* (0.871) among the genome groups were observed in AA genome groups, whereas the lowest of *N* (0.286), Na (0.371), Ne (0.371) and *I* (0.059) were observed in KKLL genome group; this large variation might be due to sample size and ecological factors which greatly influences on genetic differentiation among the wild rice populations (Orn et al. 2015 and Ishii et al. 2011). The percent of polymorphic loci (%P) ranged from 8.57% in KKLL genome to 97.14% in the AA genome group. The H_O of AA genome type groups showed reduction as compared to other genome type groups which might be due to the presence of three cultivated rice types (Nipponbare, Swarna and *O. glaberrima*) that have been domesticated resulting in the reduction in genetic variation. Obtaining low genetic diversity parameters in KKLL, HHJJ, FF and EE might be due to a single individual in their respective genome groups and perennial growth habits. The difference between the H_O and H_E was small in all the genome groups except in AA and HHJJ genome groups. The overall low average gene diversity in the present study may be due to self-pollination and majority of them are exhibiting perennial growth habit. The frequency of private alleles ranged from 0.814 (AA genome) to 0.057 (KKLL genome) with an average of 0.261 ± 0.059. The allelic pattern based on 70 markers showed high diversity indexes in AA genome type as compared to rest of the genome type.

Genetic relationship analysis

All 70 SSR markers were used to infer the phylogenetic relationships among the 18 rice genotypes with diverse genome types to explore evolutionary relationships and genome evolution using the genotyping data in DARwin5 software. The genotyping data were used to estimate the distance matrix using the Jaccard's similarity coefficient and a

Table 4 Comparison of mean genetic diversity parameters for eight genome groups using 70 SSR markers

Population	*n*	*N*	Na	Ne	*I*	*Ho*	H_E	*F*	*%P*	PA
AA	7	6.243	3.086	2.284	0.871	0.310	0.500	0.428	97.14	0.814
CC	2	1.686	2.014	1.910	0.605	0.479	0.400	-0.198	77.14	0.300
EE	1	0.700	0.957	0.957	0.178	0.257	0.129	-1.000	25.71	0.071
FF	1	0.814	1.257	1.257	0.307	0.443	0.221	-1.000	44.29	0.157
CCDD	3	2.214	1.800	1.639	0.506	0.388	0.333	-0.154	65.71	0.214
HHJJ	1	0.414	0.571	0.571	0.109	0.157	0.079	-1.000	15.71	0.071
NOWR	2	1.443	1.729	1.688	0.485	0.371	0.321	-0.140	60.00	0.400
KKLL	1	0.286	0.371	0.371	0.059	0.086	0.043	-1.000	8.57	0.057
Mean ± SE	2.25	1.725 ± 0.08	1.473 ± 0.05	1.335 ± 0.04	0.390 ± 0.01	0.311 ± 0.01	0.253 ± 0.01	-0.219 ± 0.03	97.14 ± 49.2	0.261 ± 0.05

n = Number of individual per population; *N* = number of alleles per locus; Ne = number of effective alleles per locus; *I* = Shannon's information index; H_o = observed heterozygosity; H_E = expected heterozygosity; *F* = fixation index; *%P* = percentage of polymorphic loci; PA = frequency of private alleles (number of allele unique to a single population); Mean ± SE = Means followed by standard errors

corresponding dendrogram was developed. The robust dendrogram divided the 18 genotypes into three major clusters I, II and III, which are consistent with the established evolutionary relationships among the *Oryza* genus (Wang et al. 2014; Zhang et al. 2014; Zhu et al. 2014); this shows the usefulness of the presently used markers in genetic analyses (Fig. 2). In cluster I, there were eight rice accessions which were further separated into two sub-clusters, IA and IB. In IA, all the seven AA genome types were clustered agreeing the results of the highly conserved genome structure of AA genome sequencing data analysis (Zhang et al. 2014), whereas in IB sub-cluster, there was only *O. eichengiri* with CC genome type. Across the AA genomes, there were fairly conservation of genome size, number of genes, the contents of TEs and genomic architecture and gene colinearity (Zhang et al. 2014). The IA cluster was further distinctly divided into two groups A1 and A2 as Asian and African branches similar to other reports (Zhang et al. 2014; Stein, et al. 2018) supporting the hypothesis of two independent origins of *japonica* and *indica* rice. Stein et al. (2018) have also shown a single-species phylogeny and two independent origins of *japonica* and *indica* rice using 6015 single-copy orthologs from ten *Oryza* genome including *L. perrieri*. Here, the phylogenetic tree also clearly distinguished *japonica* and *indica* subspecies from each other. Interestingly, A1 consists of only Asian rice such as cultivated rice: Nipponbare, Swarna, and their progenitor wild relative rice: *O. nivara* (annual) and *O. rufipogon* (perennial), whereas A2 contains only African rice like cultivated *O. glaberrima* and its progenitor *O. barttii* (annual) and *O. longistaminata* (perennial) suggesting their recent speciation and adaptation in distinct environments in Africa (Zhang et al. 2014; Kim et al. 2015). In the cluster II, there were nine rice accessions which were further classified into two major groups IIA and IIB. In IIA cluster, the subgroup IIA1 contains four genotypes with a majority of CC or DD genome types such as *O. grandiglumis* (CCDD), *O. alta* (CCDD), *O. australiensis* (EE) and *O. officinalis* (CC). This closed connection between the *officinalis* complex (CCDD) and EE genome types has already also been reported by Yamaki et al. (2013). On the other hand, IIA2 cluster has two accessions with *Z. latifolia* (unknown) and *O. latifolia* (CCDD). In IIB cluster, three accessions such as *O. coarctata* (KKLL), *L. perrieri* (unknown) and *O. ridleyi* (HHJJ) were clustered together. The most distant accession, *O. brachyantha* was separated as an outgroup (Zhang et al. 2014; Chen et al. 2013) from the rest of the 17 accessions as Cluster III. Among the rice genome, O. *brachyantha* has the smallest genome size with a limited number of retrotransposons (Chen et al. 2013; Zhang and Gao 2017).

To establish the genetic relationships of these 18 rice accessions based on the 70 SSR markers, the PCoA (Principle coordinate analysis) using GenAlEx 6.502 software was

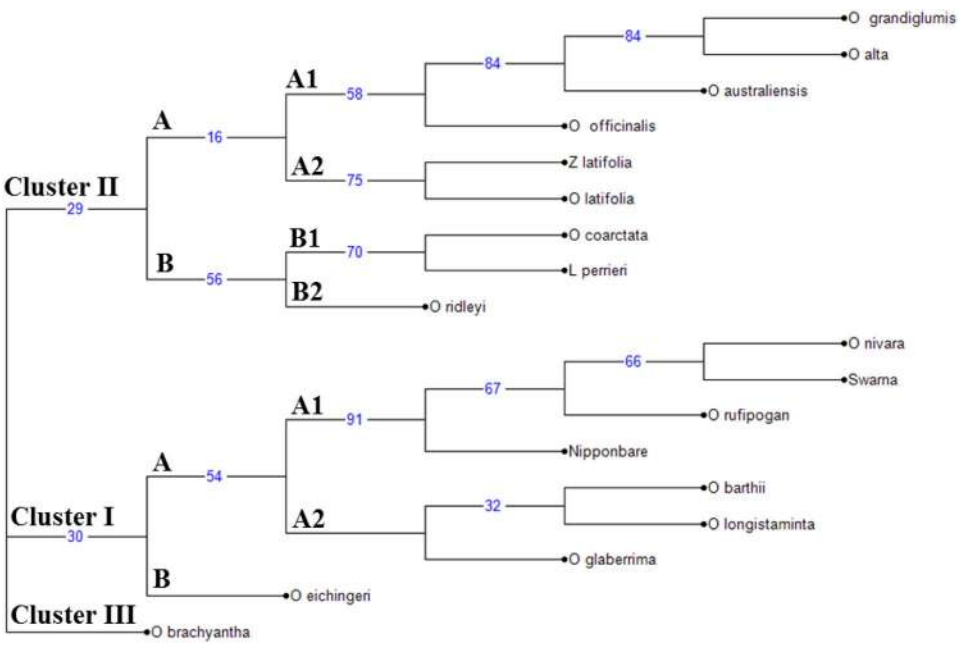

Fig. 2 Clustered analysis of 18 rice genotypes based on 70 SSR markers

further constructed. A scatter plot generated from the PCoA analysis showed that the first two components accounted for 20.17% and 11.06% of the genetic variation which resulted in a total genetic variation of 31.23% (Fig. 3). These scatter plots showed a clear separation of AA genome type groups from the rest of the genome types which is found to be concurrent with the result of the phylogenetic relationship analysis.

Genetic structure analysis

The genotypic data of 70 SSR markers were used to assess the presence of population genetic structure in 18 rice accessions using Structure software. Using STRUCTURE HARVESTER program, the peak plateau of *ad hoc* measure ΔK was detected at $K=7$ (Fig. 4). Population structure analysis in different rice diversity panels has already indicated the existence of two to eight subpopulations in rice (Nachimuthu et al. 2015). At $K=7$, the entire 18 rice accessions were distributed into seven subgroups, viz. SG1, SG2, SG3, SG4, SG5, SG6, and rest of them as admixtures (AD) (Table S2). The SG1 contains only four AA genomes such as Nipponbare, Swarna, *O. nivara* and *O. rufipogon*, suggesting that AA genome groups of seven accessions used in the present study are quite closely related genetically. It is, however, possible to further detect substructure using sufficient markers. The other subgroups are SG2 (*O. latifolia*; CCDD), SG3 (*O. alta*; CCDD and *O. grandiglumis*; CCDD),

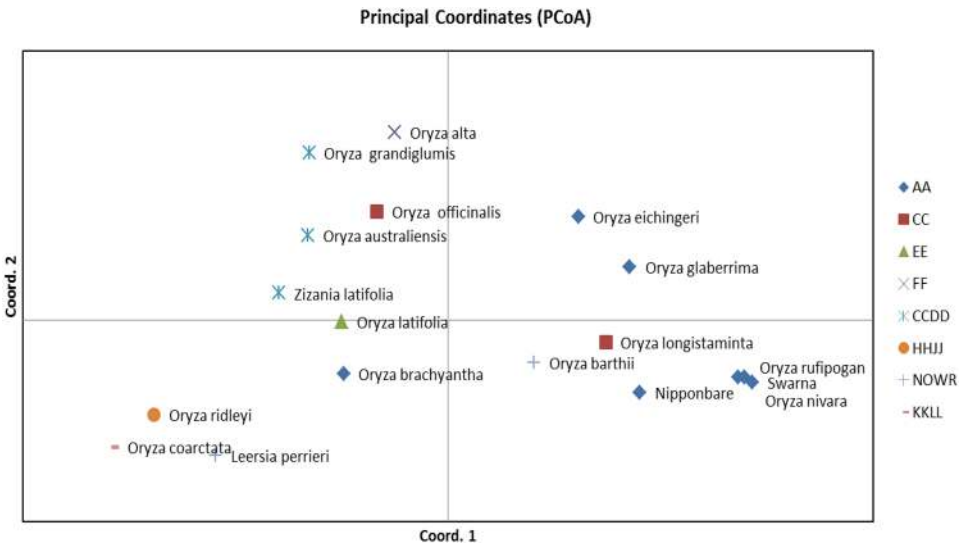

Fig. 3 Two-dimensional PCoA display of 18 rice genotypes based on 70 markers. Coord 1 and Coord 2 represent first and second coordinates, respectively. The two PCoA axes accounted for 20.17% and 11.06% of the genetic variation among populations

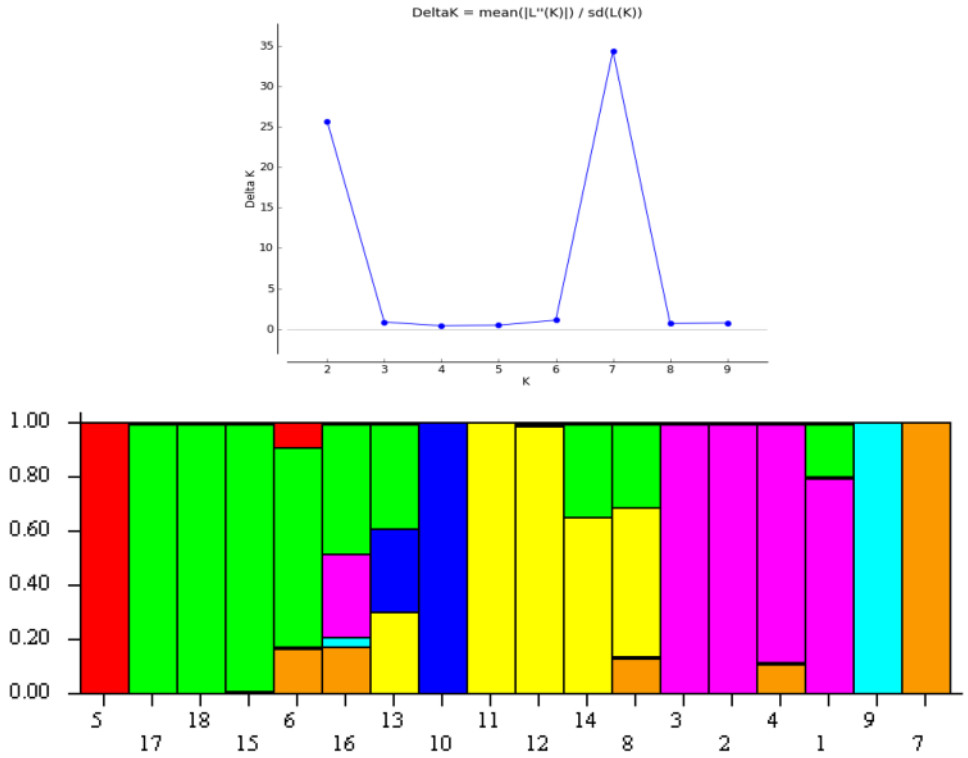

Fig. 4 Population structure of 18 germplasm based on 70 markers and the maximum of ad hoc measure ΔK determined by structure harvester was found to be $K=7$, which indicated that the entire population can be grouped into 7 subgroups. Different colors within group indicate the proportion of shared ancestry with other group which has the same color with the admixture. 1–18 represent rice germplasm listed in Table 1

SG4 (*O. ridleyi*; HHJJ, *L. perrieri*; NOWR and *O. coarctata*; KKLL), SG5 (*O. longistaminata*; AA) and SG6 (*O. officinalis*; CC). On the other hand, remaining accessions were categorized as admixtures (AD) and found to be six genotypes such as *O. eichingeri*; CC, *O. brachyantha*; FF, *O. glaberrima*; AA, *Z. latifolia*; NOWR, *O. australiensis*; EE, *O. barthi*; AA. Therefore, the $K=7$ structure was characterized by highly admixed individuals of the various genome types revealing exchange of genome content. Consequently, as a whole, the structure analysis in the 18 rice accessions could differentiate AA genome from the rest of the genome types which was further separated as a distinct population in SG1 revealing the complicated origin and evolution of different rice genomes.

To conduct the analysis of molecular variance (AMOVA), the 18 rice accessions were grouped into two, viz. cultivated rice (CR: *Oryza sativa*—Nipponbare (spp. *japonica*), Swarna (spp. *indica*) and *O. glaberrima*), and the remaining fifteen rice accessions as wild relative rice (WRR). It was found that more variance (51%) was observed among the individual, whereas among the population, it was less (14%) indicating different origins (Fig. 4 and Table 5). This result shows an extensive genetic diversity among 18 rice accessions which is expected due to diverse and distinct genome types. Therefore, the main contribution to the genetic variation in the 18 genotypes was due to variation within genotypes, while differences among populations had only 14% contribution to the total genetic variation. The pair-wise fixation indices (F_{ST}) among the eight populations varied from moderate (0.193) to high (0.900) genetic differentiation (Table 6) with the highest value (0.900) observed in between the HHJJ and KKLL, whereas the lowest (0.193) was found to be between the AA and CC.

Table 5 Results of analysis of molecular variance (AMOVA) between the cultivated and wild relative rice

Source	*df*	SS	MS	Est. Var.	Percentage of variation
Among Pops	1	73.139	73.139	3.623	14
Among Indiv	16	590.500	36.906	13.717	51
Within Indiv	18	170.500	9.472	9.472	35
Total	35	834.139		26.813	100

df degree of freedom; SS: sum of squares; MS: mean squares; Est. Var. = estimated variance (P value > 0.001)

Conclusion

With the availability of complete and high-quality genome sequencing data of Nipponbare genotype in the public domain, a large set of genomic SSR was developed and validated earlier by many rice researchers. However, there is no report of cross-species transferability of these markers among the diverse rice genome types. In the present study, it was conducted for the first time the cross-species

Table 6 Pair-wise F_{ST} estimates among the eight populations of rice

Population	AA	CC	EE	FF	CCDD	HHJJ	NOWR	KKLL
AA	0.000							
CC	0.193	0.000						
EE	0.416	0.457	0.000					
FF	0.325	0.393	0.660	0.000				
CCDD	0.248	0.302	0.533	0.353	0.000			
HHJJ	0.551	0.627	0.811	0.755	0.629	0.000		
NOWR	0.270	0.353	0.590	0.514	0.392	0.688	0.000	
KKLL	0.616	0.678	0.867	0.798	0.706	0.900	0.723	0.000

*F*st Values below diagonal

transferability among more than eight genome types using a set of 70 SSR markers developed from cultivated rice which will provide an important resources for conservation, gene introgression, genome evolution and genome relationship among the *Oryza* species. The informative SSR markers with higher cross-species transferability rate among the different genome types identified from the present study will be useful in future studies of rice population structure, genetic diversity, MAS and identification of genome type. The present study would further facilitate the understanding of complex mechanism of the origin and evolutionary processes of different *Oryza* species and enhance the valuable resources for rice genetic improvement through marker assisted breeding in cultivated rice.

Acknowledgments The authors are extremely grateful to the Director, ICAR-National Rice Research Institute, Cuttack, India, for his support and facilitation for carrying out the research work successfully.

Compliance with ethical standards

Competing interests The authors declare that they have no potential competing interests.

References

Alexander LW, Thammina CS, Kramer M (2018) Cross-transferability of SSR markers in *Osmanthus*. Genet Resour Crop Evol 65(1):125–136

Brondani C, Rangel PHN, Borba TCO, Brondani RPV (2003) Transferability of microsatellite and sequence tagged site markers in *Oryza* species. Hereditas 138:187–192

Brondani RPV, Zucchi MI, Brondani C, Rangel PHN, Borba TCDO, Rangel PN et al (2005) Genetic structure of wild rice *Oryza glumaepatula* populations in three Brazilian biomes using microsatellite markers. Genetica 125:115–123

Chen J, Huang Q, Gao D, Wang J, Lang Y, Liu T et al (2013) Whole-genome sequencing of *Oryza brachyantha* reveals mechanisms underlying *Oryza* genome evolution. Nat Commun 4:1595

Choudhury B, Khan ML, Dayanandan S (2013) Genetic structure and diversity of indigenous rice varieties (*Oryza sativa*) in Eastern Himalayan region of Northeast India. Springerplus 2:228–237

Dayanandan S, Bawa K, Kesseli R (1997) Conservation of microsatellites among tropical trees (Leguminosae). Am J Bot 84:1658–1663

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4(2):359–361

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14(8):2611–2620

Fernandez-Silva I, Eduardo I, Blanca J, Esteras C, Pico B, Nuez F et al (2008) Bin mapping of genomic and EST-derived SSRs in melon (*Cucumis melo* L.). Theor Appl Genet 118(1):139–150

Gao LZ, Zhang CH, Jia JZ, Dong YS (2005) Cross-species transferability of rice microsatellites in its wild relatives and the potential for conservation genetic studies. Genet Resour Crop Evol 52:931–940

Ge S, Sang T, Lu BR, Hong DY (1999) Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proc Natl Acad Sci USA 96:14400–14405

Gupta PK, Rustgi S, Sharma S, Singh R, Kumar N, Balyan HS (2003) Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. Mol Genet Genomics 270(4):315–323

Hernández P, Dorado G, Laurie DA, Martin A, Snape JW (2001) Microsatellites and RFLP probes from maize are efficient sources of molecular markers for the biomass energy crop *Miscanthus*. Theor Appl Genet 102(4):616–622

Huang X, Zhao Y, Li C, Wang A, Zhao Q, Li W et al (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nat Genet 44(1):32–39

Ishii T, Hiraoka T, Kanzaki T, Akimoto M, Shishido R, Ishikawa R (2011) Evaluation of genetic variation among wild populations and local varieties of rice. Rice 4:170–177

Izawa T (2008) The process of rice domestication: a new model based on recent data. Rice 1(2):127–134

Jacquemin J, Bhatia D, Singh K, Wing RA (2013) The International Oryza Map Alignment Project: development of a genus-wide comparative genomics platform to help solve the 9 billion-people question. Curr Opin Plant Biol 16(2):147–156

Jena KK, Khush GS (1990) Introgression of genes from *O. officinalis* Wall. ex Watt to cultivated rice. O. sativa L. Theor Appl Genet 80:737–745

Khush G (2003) Productivity improvements in rice. Nutr Rev 61:S114–S116

Kim K, Lee SC, Lee J, Yu Y, Yang K, Choi BS et al (2015) Complete chloroplast and ribosomal sequences for 30 accessions elucidate evolution of *Oryza* AA genome species. Sci Rep 5:15655

Krishnan SG, Waters DLE, Henry RJ (2014) Australian wild rice reveals pre-domestication origin of polymorphism deserts in rice genome. PLoS ONE 9(6):e98843

Li G, Hu W, Qin R, Jin H, Tan G, Zhu L, He G (2008) Simple sequence repeat analyses of interspecific hybrids and MAALs of *Oryza officinalis* and *Oryza sativa*. Genetica 134(2):169–180

Liu L, Wang Y, He P, Li P, Lee J, Soltis DE, Fu C (2018) Chloroplast genome analyses and genomic resource development for epilithic sister genera *Oresitrophe* and *Mukdenia* (Saxifragaceae), using genome skimming data. BMC Genomics 19(1):235

Mansueto L, Fuentes RR, Borja FN, Detras J, Abriol-Santos JM, Chebotarov D et al (2016) Rice SNP-seek database update: new SNPs, indels, and queries. Nucleic Acids Res 45(D1):D1075–D1081

Nachimuthu SVV, Muthurajan R, Duraialaguraja S, Sivakami R, Pandian BA, Ponniah G et al (2015) Analysis of population structure and genetic diversity in rice germplasm using SSR markers: an initiative towards association mapping of agronomic traits in *Oryza*. Rice 8:30

Nei M (1972) Genetic distance between populations. Am Nat 106(949):283–292

Ngangkham U, Parida SK, De S, Kumar AR, Singh AK, Singh NK, Mohapatra T (2010) Genic markers for Wild Abortive (WA) based male sterility and its fertility restoration in rice. Mol Breed 26:275–292

Ngangkham U, Samantaray S, Yadav MK, Kumar A, Chidambaranathan P, Katara JL (2018a) Effect of multiple allelic combinations of genes on regulating grain size in rice. Plos ONE 13(1):e0190684

Ngangkham U, Nath M, Dokku P, Mithra SA, Ramamurthy S, Singh NK et al (2018b) An EMS-induced new sequence variant, TEMS5032, in the coding region of SRS3 gene leads to shorter grain length in rice (*Oryza sativa* L.). J Appl Genet 59(4):377–389

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol Ecol 13(5):1143–1155

Orn C, Shishido R, Akimoto M, Ishikawa R, Htun TM, Nonomura KI et al (2015) Evaluation of genetic variation among wild rice populations in Cambodia. Breed Sci 65(5):430–437

Panaud O, Chen X, McCouch SR (1996) Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). Mol Gen Genet 252(5):597–607

Patra BC, Ray S, Ngangkham U, Mohapatra T (2015) Rice. In: Upadhyaya HD, Singh M (eds) Genetic and genomic resources for grain cereals improvement. Elsevier Academic Press, Cambridge

Peakall RO, Smouse PE (2012) GenAlex 6.5: genetic analysis in Excel. Population genetic software for teaching and research. Bioinformatics 28:2537–2539

Perrier X, Jacquemoud-Collet JP (2006) DARwin software (http://darwin.cirad.fr/darwin)

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2):945–959

Ramiah K (1953) Rice breeding and genetics. ICAR, New Delhi

Ray S, Bose LK, Ray J, Ngangkham U, Katara JL, Samantaray S et al (2016) Development and validation of cross-transferable and polymorphic DNA markers for detecting alien genome introgression in *Oryza sativa* from *Oryza brachyantha*. Mol Genet Genomics 291(4):1783–1794

Shakiba E, Eizenga GC (2014) Unraveling the secrets of rice wild species. In: Rice-germplasm, genetics and improvement. InTech

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29(1):308–311

Song ZP, Xu X, Wang B, Chen JK, Lu BR (2003) Genetic diversity in the northernmost *Oryza rufipogon* populations estimated by SSR markers. Theor Appl Genet 107(8):1492–1499

Song S, Tian D, Zhang Z, Hu S, Yu J (2019) Rice genomics: over the past two decades and into the future. Genom Proteomics Bioinf 1:1. https://doi.org/10.1016/j.gpb.2019.01.001

Stein JC, Yu Y, Copetti D, Zwickl DJ, Zhang L, Zhang C et al (2018) Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus *Oryza*. Nat Genet 50(2):285

Tan G, Jin H, Li G, He R, Zhu L, He G (2005) Production and characterization of a complete set of individual chromosome additions from *Oryza officinalis* to *Oryza sativa* using RFLP and GISH analyses. Theor Appl Genet 111(8):1585–1595

Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277(5329):1063–1066

Tello-Ruiz MK, Stein J, Wei S, Preece J, Olson A, Naithani S et al (2015) Gramene 2016: comparative plant genomics and pathway resources. Nucleic Acids Res 44(D1):D1133–D1140

Vaughan DA, Morishima H, Kadowaki K (2003) Diversity in the *Oryza* genus. Curr Opin Plant Biol 6(2):139–146

Wang ML, Chen ZB, Barkley NA, Newman ML, Kim W, Raymer P et al (2006) Characterization of seashore *Paspalum* (*Paspalum vaginatum* Swartz) germplasm by transferred SSRs from wheat, maize and sorghum. Genet Resour Crop Evol 53(4):779–791

Wang C, Liu X, Peng S, Xu Q, Yuan X, Feng Y et al (2014) Development of novel microsatellite markers for the BBCC *Oryza* genome (Poaceae) using high-throughput sequencing technology. PloSone 9(3):e91826

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241(1–2):225–235

Yadav MK, Aravindan S, Ngangkham U, Shubudhi HN, Bag MK, Totan Adak T et al (2017) Use of molecular markers in identification and characterization of resistance to rice blast in India. Plos ONE 12(4):e0176236

Yamaki S, Ohyanagi H, Yamasaki M, Eiguchi M, Miyabayashi T, Kubo T et al (2013) Development of INDEL markers to discriminate all genome types rapidly in the genus *Oryza*. Breed Sci 63(3):246–254

Yan Z, Wu F, Luo K, Zhao Y, Yan Q, Zhang Y et al (2017) Cross-species transferability of EST-SSR markers developed from the transcriptome of *Melilotus* and their application to population genetics research. Sci Rep 7(1):17959

Yeh FC (1997) Population genetic analysis of co-dominant and dominant marker and quantitative traits. Belgian J Bot 130:129–157

Yelome OI, Audenaert K, Landschoot S, Dansi A, Vanhove, Silue D et al (2018) Analysis of population structure and genetic diversity reveals gene flow and geographic patterns in cultivated rice (*O. sativa* and *O. glaberrima*) in West Africa. Euphytica 214:215

Yuan N, Sun Y, Comes HP, Fu CX, Qiu YX (2014) Understanding population structure and historical demography in a conservation context: population genetics of the endangered *Kirengeshoma palmata* (Hydrangeaceae). Am J Bot 101(3):521–529

Zhang QJ, Gao LZ (2017) Rapid and recent evolution of LTR retrotransposons drives rice genome evolution during the speciation of AA-genome *Oryza* species. G3: Genes Genom Genet 7(6):1875–1885

Zhang QJ, Zhu T, Xia EH, Shi C, Liu YL, Zhang Y et al (2014) Rapid diversification of five *Oryza* AA genomes associated with rice adaptation. Proc Natl Acad Sci USA 111(46):E4954–E4962

Zhao H, Yao W, Ouyang Y, Yang W, Wang G, Lian X et al (2014) RiceVarMap: a comprehensive database of rice genomic variations. Nucleic Acids Res 43(D1):D1018–D1022

Zheng X, Pan C, Diao Y, You Y, Yang C, Hu Z (2013) Development of microsatellite markers by transcriptome sequencing in two species of *Amorphophallus* (Araceae). BMC Genomics 14(1):490

Zhou Q, Luo D, Ma L, Xie W, Wang Y, Wang Y, Liu Z (2016) Development and cross-species transferability of EST-SSR markers in Siberian Wildrye (*Elymus sibiricus* L.) using Illumina sequencing. Sci Rep 6:20549

Zhu T, Xu PZ, Liu JP, Peng S, Mo XC, Gao LZ (2014) Phylogenetic relationships and genome divergence among the AA-genome species of the genus *Oryza* as revealed by 53 nuclear genes and 16 intergenic regions. Mol Phylogenet Evol 70:348–361