Genetic Characterization of coloured rice genotypes (*Oryza Sativa* L.) from different regions of India using molecular markers

[1]AMRUTA DESAI, [2]BHUSHAN AHER AND [1]SUBHASH N.
1Department of Agricultural Biotechnology, Anand Agricultural University, Anand.
2College of Agriculture, Anand Agricultural University, Jabugam

Abstract: Genetic diversity analysis of 24 coloured and 8 white rice genotypes were conducted usingSSR and InDel markers. Among them, SSR markers such as RM5, RM 210, RM 251, RM 261, RM 490, RM 3859 and InDel marker RID 12, RID 13, RID 19 and RID 25 were capable to distinguish coloured rice genotypes. The PIC values ranged from 0.16 (RM 232) to 0.78 (RM 258) and 0.52 (RID 25) and 0.80 (RID 12) for SSR and INDEL markers respectively. The DICE similarity coefficient based on SSR and InDel analysis was 0.89 and 1.00, respectively while the cophenetic correlation (r) value for cluster analysis of SSR and InDel was 0.89 and 0.71, respectively. Molecular variance (AMOVA), F statistics and genetic variation statistics revealed that coloured rice genotypes were more genetically diverse than white rice genotypes. In the case of both SSR and INDEL markers highest genetic diversity was observed between Pankhali 203 (White pericarp) and IRST 1 (Red pericarp) whereas least genetic diversity was observed between RRT GR12 (Red pericarp) and Krishna Kamod (White pericarp). However, it was also found that black and brownpericarp isgeneticallysimilar to red pericarp. It was suggested from the conducted experiment to adopt SSR markers as it gave better results than InDel even though all InDel markers used in present study were also able to distinguish coloured rice genotypes.

Key words: Red rice, SSR, INDEL, Polymorphism, Genetic diversity

Introduction

Rice (*Oryza sativa* L.) is one of the most important cereal crops in the world and India. Rice is also called as the " Grain of Life", because it is not only the staple food for more than 70 percent of the Indians but also a source of livelihood for about 120-150 million rural households (Dhurai *et al*., 2014). India stands second in world area (about 45 million hectare) and production of rice. It is a rich source of carbohydrates, energy, thiamine, pantothenic acid and folic acid but lacks important micronutrients especially iron (Fe) and zinc (Zn). Different colours, such as purple, brown, red and white occur in the pericarp of rice. Rice which contains a red aleurone layer is called red rice. The red pigment in rice grains is due to the presence of pigments known as proanthocyanidin. (Oki *et al*., 2002) Regardless of the problems associated with red rice as a weed, the red pigment is of interest for nutritional reasons and importantly "Biofortification" as the Iron and Zinc content of red rice is 2–3 times higher than that of white rice (Ramaiah and Rao, 1953).

For red coloration of pericarp two loci have been identified using classical genetic analysis *Rc* (brown pericarp and seed coat) and *Rd* (red pericarp and seed coat). These two loci accumulatively produce red seed colour whereas only *Rd* provides brown seeds. *Rc* alone lacks any phenotype (Kato and Lshikawa, 1921). A 14 bp deletion in the *Rc* gene changed the red seed of wild rice and gave rise to white seeds of modern rice. (Lee *et al*., 2009)This character is shared by a large majority of rice varieties, regardless of subspecies.

The rapid development of molecular markers provides an effective tool for studying genetic diversity and population differentiation of plant species (O'Hanlon *et al*., 2000). Molecular markers are based on the genetic variation at specific gene loci and hence it provides information about the amount and distribution of genetic diversity within and among populations (Buu and Lang., 1999). A wide variety of DNA markers such as Restriction Fragment Length Polymorphism (RFLP), Amplified Fragment Length Polymorphism (AFLP), Simple Sequence Repeat (SSR), Insertion/Deletion markers (INDEL), Single nucleotide polymorphism (SNP) etc. have been extensively used in rice for genetic diversity analysis, phylogenetic and evolutionary studies, mapping and tagging genes for quantitative traits of agronomic importance and Marker Assisted Selection (MAS).

Simple Sequence Repeat (SSRs) also referred as microsatellites, are allele-specific and co-dominant markers. Compared with other markers microsatellites are abundant, codominant and interspersed throughout the genome. These markers can detect a significantly higher degree of polymorphism in rice, which becomes ideal for studies on genetic diversity and intensive genetic mapping. SSR markers can estimate genetic diversity between cultivars e.g. between parents of a genepool or between plants extracted from a population or between populations (Vhora *et al*., 2013).The genetic diversity and genetic structure of natural populations of O. *rufipogon* in China using SSR markers and information was investigated by Zhou *et al*., (2003). Gealy *et al*., (2002) used DNA microsatellite marker techniques to distinguish among red rice types, rice cultivars, and red rice–cultivated rice hybrid derivatives (RC hybrids) and to assess the genetic variation among populations of these three *Oryza* groups in the southern United States. Vaughan *et al*., (2001) analyzed red rice samples with simple sequence repeat polymorphism markers distributed among 12 chromosomes.

InDels, short for insertion/deletion markers, are stings of mutated base pairs in DNA and are identified by their genomic position and their length (Zeng *et al*., 2013). They are useful because the chances of two InDel mutations same length occurring in the same genomic position are negligible and they can therefore be used to map common ancestry and origin studies.Insertion/deletion (InDel) markers were developed to help define break points across the QTL region associated with red grain (rg7.1) on chromosome 7 by Sweeney *et al*., (2006) Sweeney *et al*., (2007) genotyped a set of 440 geographically and genetically diverse rice cultivars using the rice InDel RID 12 primers to determine the frequency and distribution of the 14-bp deletion in the *Rc* gene that resulted in white pericarp in rice.

Recognizing the importance of a better utilization of features associated with red rice in breeding programs, The communication deals with characterization of genes responsible for colour in red rice genotypes using molecular markers with the aim of assessing the genetic variability at molecular level among 24 genotypes of red rice and eight genotypes of white rice through DNA fingerprinting by SSR and InDel markers.

Materials and methods

Plant material and DNA extraction

The study was conducted at Plant Biotechnology Laboratory, Department of Agricultural Botany, B.A College of Agriculture, Anand Agricultural University, Anand. The seeds of 32 rice genotypes comprising of coloured rice (red, black, brown and white) used in the present study was obtained from the Main Rice Research Station (MRRS), Anand Agricultural University, Nawagam (Table 1). Few healthy seeds were sown and allowed to grow for three weeks in the pots. For the proper growth and emergence of healthy seedlings, these pots were watered regularly and proper light and temperature conditions were maintained. Total genomic DNA extraction from leaves of three weeks old seedlings were carried out by Cetyl trimethyl ammonium bromide (CTAB) method (Ahmadikhah *et al*., 2006). 300mg tissue samples were homogenized in liquid nitrogen prior mixing with 800 µl of extraction buffer and were incubated for 1 hour at 65º C in water bath. Later on Chloroform: isoamyl alcohol (24:1) was added in the tubes and centrifuged at 4ºC for 15 minutes at 12,000rpm. The supernatant was collected and washed 1-2 times with Chloroform: isoamyl alcohol (24:1) and kept for precipitation with absolute alcohol for overnight. The samples were centrifuged at 5,000 rpm for 15 min to get DNA in pellet form. The pellets were further washed with 70% alcohol and kept for drying and the quality was confirmed through Nanodrop N.D. 1000 software (ver 3.7.1). For SSR and INDEL analysis DNA amplification was performed in 25µl reaction volume containing 50ng genomic DNA, 10X polymerase buffer, 25mM dNTPs, 0.5 µl of each primer (10pmol), 1 unit of Taq DNA polymerase using Applied Biosystem Thermal Cyclers. The cycling conditions were: 1 cycle of 94°C for 7min followed by 40 cycles of 45 sec each of 94°C, 48°C, 72°C and finally 1 cycle of 72°C for 5min. Total of 10 µl aliquots of the amplification products loaded in 3% (w/v) agarose gel for electrophoresis in 1X TBE buffer and stained with Ethidium bromide and documented using SYNGENE GENESNAP G-BOX gel documentation system. These photographs were used to score the DNA bands for analysis. The gels were scored for computer analysis on the basis of the presence and absence of the amplified products. If a product was present in a genotype, it was designated as '1' and if absent; it was designated as '0'. A total of 21 SSR and 4 INDEL markers used for characterizing 32 rice genotypes. Both SSR and INDEL data were analyzed using NTSYS-PC (Numerical Taxonomical and Multivariant Analysis System computer package). The data were used to generate Jaccard's similarity coefficients for SSR and INDEL bands. The Jaccard's coefficients between each pair of accessions were used to construct a dendrogram using the unweighted pair group method with arithmetic averages (UPGMA). Analysis parameters Allele frequency, observed and expected heterozygosity were calculated by Genalex software version 6.3 (Peakall and Smouse, 2006). While genetic variation statistics were computed using the formulas described in the POPGENE (version 3.3) software user manual.

Microsatellite marker genomic analysis

Total 60 primers were screened for the study of molecular characterization of colored (red, brown, black) and white genotypes in rice, out of which 21 primers were polymorphic. These 21 primers were used for SSR analysis of 24 red and 8 white rice genotypes as discussed here and presented in Table 1.

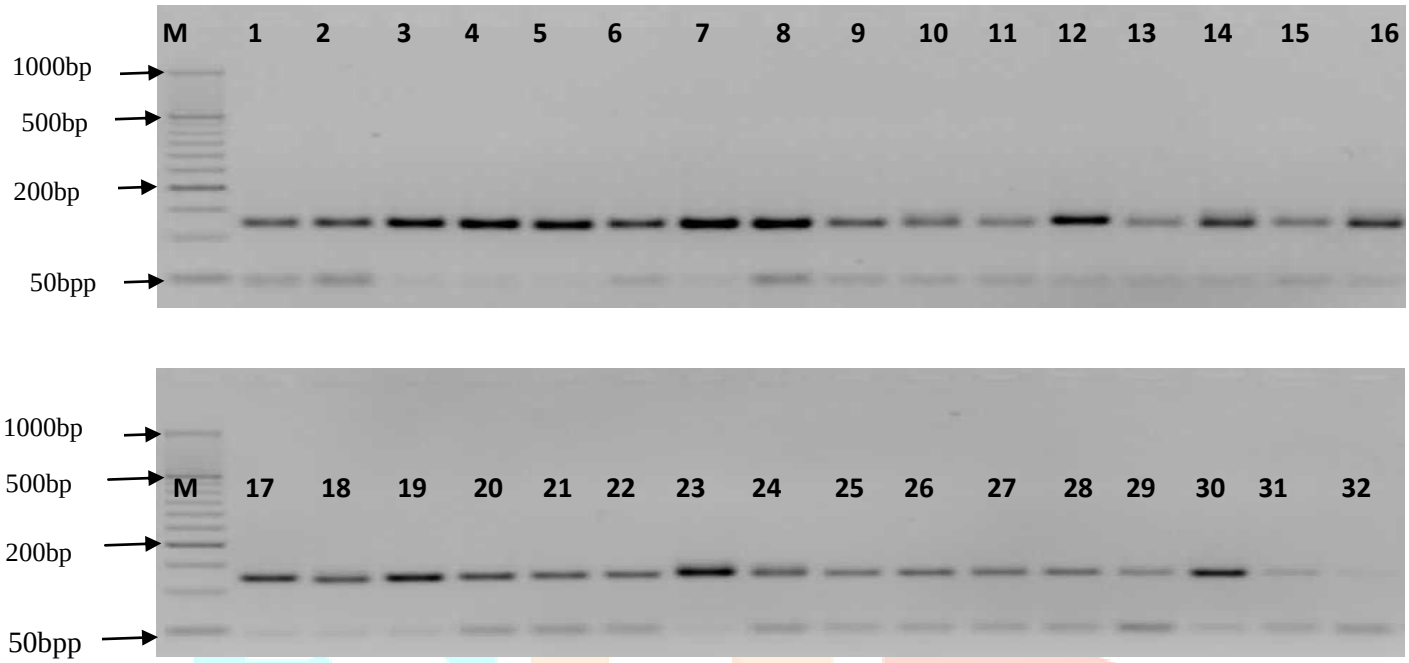


Figure 1: SSR Profile of RM 408.
M = 50 bp DNA ladder; 1=RRT IRST 1; 2=RRT IRST 2; 3=RRT IRST 4; 4=RRT IRST 7; 5=RRT IRST 10; 6=RRT IRST 13; 7=RRT IRST 14; 8=RRT IRST 16; 9=RRT IRST 19; 10=RRT IRST 39; 11=RRT IRST 41; 12=RRT IRST 44; 13=RRT IRST 45; 14=RRT IRST 47; 15=RRT IRST 48; 16=RRT IRST 50; 17=JOT 21224; 18=IR 81429-B-31; 19=HB-1; 20=GR-11 ; 21=GURJARI; 22=RRT GR-12; 23=KRISHNA KAMOD; 24=PANKHALI-203; 25=GAR-1; 26=LALKADA; 27=DHANHAR; 28=IR-64; 29=A(004); 30=077; 31=BC1; 32=TSP

The 21 SSR primers amplified a total of 575 bands with an average of 27.3 bands per marker. Based on the SSR data, cluster analysis was performed using genetic similarity values and a dendrogram was generated showing genetic relationships among these genotypes. The similarity coefficient range was found between 0.49 to 0.89, while the highest similarity index value of 0.89 was found between RRT IRST 48 and RRT IRST 50. The SSR marker RM215 and RM 225 produced maximum number of 70 bands, while RM53 produced minimum number of 10 bands. The SSR marker RM210 produced maximum number of seven alleles while RM10, RM22, RM214 and RM408 produced only a single allele. The average of allele amplified by 21 markers was 3.66. The highest PIC (Polymorphic Information Content) value obtained was 0.78 by marker RM258 while the lowest value was 0.16 by marker RM232. The details of amplification products are given in the Table 2.

Table 1: List of genotypes collected from Nawagam

Sr. No	White Rice Genotypes	Sr. No	Red Rice Genotypes	Sr. No	Red Rice Genotypes
1	RRT GR-11	**1**	RRT IRST 1	**13.**	RRT IRST 45
2	Gurjari	**2**	RRT IRST 2	**14.**	RRT IRST 47
3	Krishna Kamod	**3**	RRT IRST 4	**15.**	RRT IRST 48
4	Pankhali-203	**4**	RRT IRST 7	**16.**	RRT IRST 50
5.	GAR-1	**5**	RRT IRST 10	**17.**	JOT 21224
6.	Dhanhar	**6**	RRT IRST 13	**18.**	IR 81429-B-31
7.	IR-64	**7**	RRT IRST 14	**19.**	HB-1
8.	RRT GR-12	**8**	RRT IRST 16	**20.**	LALKADA
	-	**9**	RRT IRST 19	**21.**	A (004)
	-	**10**	RRT IRST 39	**22.**	077
	-	**11.**	RRT IRST 41	**23.**	BC1
	-	**12.**	RRT IRST 44	**24.**	TSP

Allele frequencies and polymorphism

Seventy two alleles were observed at the 21 microsatellite loci among the 32 genotypes sampled from two populations, averaging 3.66 per locus. The number of alleles per locus varied greatly among population, ranging from one at RM 10, RM 22, RM 214 and RM 408 to seven at RM 210. It was noted that all alleles obtained were polymorphic because they fulfilled the requirement of having allele frequencies below 0.99. It was observed that allele frequency varied between two populations hence the markers used in present study were successful in distinguishing both populations. Certain alleles showed higher frequency in red rice population. Allele having molecular weight 137 bp in marker RM 234 showed highest allelic variation while allele having molecular weight of 304, 297 and 281bp in marker RM 261 showed least allelic variation between two population.

Table 2: Results of SSR analysis

Sr. No	Marker	Chromosome No.	No of alleles amplified	Molecular Weight Range	Total no. of bands	PIC value
1	RM 5	1	3	96-113	12	0.48
2	RM 10	7	1	159	13	1.00
3	RM 22	3	1	194	26	1.00
4	RM 53	2	3	160-182	10	0.46
5	RM 190	6	4	100-124	24	0.64
6	RM 210	8	7	106-146	43	0.78
7	RM 214	7	1	112	26	1.00
8	RM 215	10	3	83-148	70	0.57
9	RM 221	2	2	101-192	44	0.39
10	RM 225	6	3	101-148	70	0.57
11	RM232	3	2	158-174	22	0.16
12	RM 234	7	4	137-156	15	0.72
13	RM 242	9	6	210-270	18	0.75

14	RM 251	3	3	127-162	18	0.20
15	RM 258	10	5	454-148	18	0.78
16	RM 261	4	5	260-338	20	0.62
17	RM 408	8	1	128	28	1.00
18	RM 481	7	6	130-212	29	0.65
19	RM 490	1	6	87-140	13	0.72
20	RM 3859	7	4	172-233	30	0.64
21	RM 7121	7	2	53-61	26	0.41
Total			72	-	575	-
Average	-		3.42	-	27.38	0.72

RM 234, RM 215 and RM 251 were successful in distinguishing red and white rice genotypes these results compliment to results obtained by Estorninos *et al*., (2000). RM 5 was also able to distinguish red genotypes. This result was supported by Vaughan *et al*., (2001). From the present data it was studied that RM 258, RM 261, RM 481 and RM 490 were also successful in distinguishing red and white rice genotypes. RM 221 showed similar allele frequency at two loci of both populations hence it shows a good correlation to both populations.

The level of polymorphism among the genotypes was evaluated by calculating the PIC values for each 21 microsatellite markers evaluated. The highest PIC (Polymorphic Information Content) value obtained was 0.78 of marker RM258 and RM 210 while the lowest value was 0.16 by marker RM232. High PIC values can be attributed to the use of more informative markers. It was also found that the higher the PIC value of a locus, the higher the number of alleles detected.

The PIC values indicated that RM 258 and RM 210 might be the best marker for diversity analysis of rice genotypes, followed by RM 234, RM 481, RM190 and RM3859. RM 221 and RM232 were likely the least powerful markers. Observed and expected heterozygocity among population was also estimated. Expected heterozygocity is an estimate of the extent of genetic variability in the population while observed heterozygocity is an estimate of observed genetic variability in the population. Observed and expected heterozygocity varied among population to a higher extent. Observed heterozygocity was higher than expected heterozygocity in 15 cases out of 21 markers. Observed heterozygocity was higher in red rice genotypes in majority cases. Hence red rice genotypes were genetically more diverse. Deviation in observed heterozygocity is population specific in all cases.

From the present study, it can be said that the best markers to characterize red rice genotypes are RM 210 and RM 258 followed by RM 5, RM 251, RM 261, RM 490 and RM 3859.

Population genetic structure

AMOVA is a method used for studying molecular variation within a species. Total genetic variation is partitioned by analysis of molecular variance (AMOVA) into two levels: between populations and among individuals, and then summarized as F-statistics (Excoffier et al., 1992; Peakall et al., 1995). For microsatellites markers, genetic variation occouring among populations between populations ($p < 0.05$) and within populations ($p < 0.05$) was also detected.

AMOVA analysis for microsatellite data showed that the two populations under observation (*viz*. Red rice and white rice) contributed only 1% variation while 99% variation was contributed by distinct genotypes used in the present study. This confirmed the result obtained by observed heterozygocity and showed the diverse genetic nature of red rice genotypes. This Permutation test (based on 999 permutations) suggested that the overall ΦPT was not significantly different from the null distribution (ΦPT = 0.010, P = 0.130), which indicates the differences among genotypes are non significant.

Cao *et al*., 2006 conducted similar study on the Panshan weedy populations. It was found that the weedy populations (coloured rice) possessed the highest level of genetic diversity. This result supports the present study.

The F statistics also allows the analysis of the structures of subdivided populations. Therefore it may be used to measure the genetic distance among subpopulations, that are not intermating will have different allele frequencies to those of the total population. The statistical indexes involved measure:

Fis = the deficiency or excess of average heterozygotes in each population under observation.

Fst = the degree of gene differentiation among populations on the basis of allele frequencies.

Fit = the deficiency or excess of average heterozygotes in a group of populations under observation.

In the two populations of *O. sativa* L. detected by microsatellite loci, *Fst* was 0.05, indicating that the total genetic variation existed because of differences among populations was less. It also indicates that allelic frequencies are not deviating much from the Hardy–Weinberg equilibrium. The allele frequency differentiation among the two populations is very low (*Fst* = 0.005) and is probably a result of many random mating. Hence overall *Fst* and *Fit* reflected the effect of inbreeding for self pollinated species such as rice.

The results based on both *Fst* and AMOVA demonstrated that major genetic variation existed within red rice populations. Broad genetic variation in the red rice gene pool may provide more opportunities for selecting beneficial genetic resources for rice breeding.

Table 3: Analysis of molecular variance (AMOVA) for 32 rice populations was analyzed. 24 individuals from population 1 and eight individuals from population 2 were evaluated for 72 alleles amplified by 21 SSR primer pairs.

Source	df	SS	MS	Est. Var.	%
Between Pops	1	1.115	1.115	0.010	1
Within Pops	30	29.792	0.993	0.993	99
Total	31	30.906		1.003	100
Stats	Value		P(rand >= data)		
ΦPT	0.010		0.130		

Genetic variability

Effective number of alleles is the number of alleles that can be present in a population. From the present data, 1.3948 alleles would be expected in a locus in each population used in present study. Gene diversity ranged from 0.500 to 0.000 while Shannon's information index data ranged from 0.69 to 0.000. From the above data it can be concluded that the genotypes used in this study were distinct from each other and hence can be used in breeding program for the purpose of improvement in nutritional quality. Similar results were found by Cao *et al*., (2006), they used Simple sequence repeat (SSR) markers to estimate the genetic diversity of 30 weedy rice populations from Liaoning, each containing about 30 selected rice varieties. Prathepha (2011) had evaluated ninety-nine weedy rice accessions from four populations by means of four SSR markers and had obtained similar results. However the result shows discrepancy from a previous study by Yu et al. (2008), where low genetic diversity (He = 0.053) was found in a total of 42 weedy rice individuals from Liaoning.

SSR cluster analysis

The Cluster analysis divided the genotypes into two major groups A and B. The first major group (A) included IRST1, IRST2, IRST4, IRST7, IRST10, IRST13, IRST14, IRST16, IRST19, IRST39, IRST41, IRST44, IRST45, IRST47, IRST48 and IRST50. In this group IRST48 and IRST50 showed highest genetic similarity. First cluster (A) was further sub-divided into two clusters first minor cluster (A1) comprised of IRST1, IRST2, IRST4, IRST7, IRST10, IRST13, IRST14, IRST16, IRST19, IRST39, IRST44, IRST45, IRST47, IRST48 and IRST50 whereas second minor cluster (A2) included IRST41. Second major cluster (B) contained JOT-21224, IR-81429-B-31, HB1, GR-11, Gurjari, GAR-11, Lalkada, Dhanhar, IR-64, A (004), 077, BC1 and TSP. Second major cluster (B) included two minor clusters first minor cluster (B1) included JOT-21224, IR-81429-B-31, HB1, GR-11, Gurjari, GAR-11, Lalkada, IR-64, A(004) and BC1 while second minor cluster (B2) included Lalkada, BC1 and TSP. Second major cluster depicted that black and brown rice is genetically similar and related to red rice. Highest genetic diversity was shown between IRST1 and Krishna Kamod whereas least genetic diversity was seen between IRST1 and IRST2 (Fig. 2).

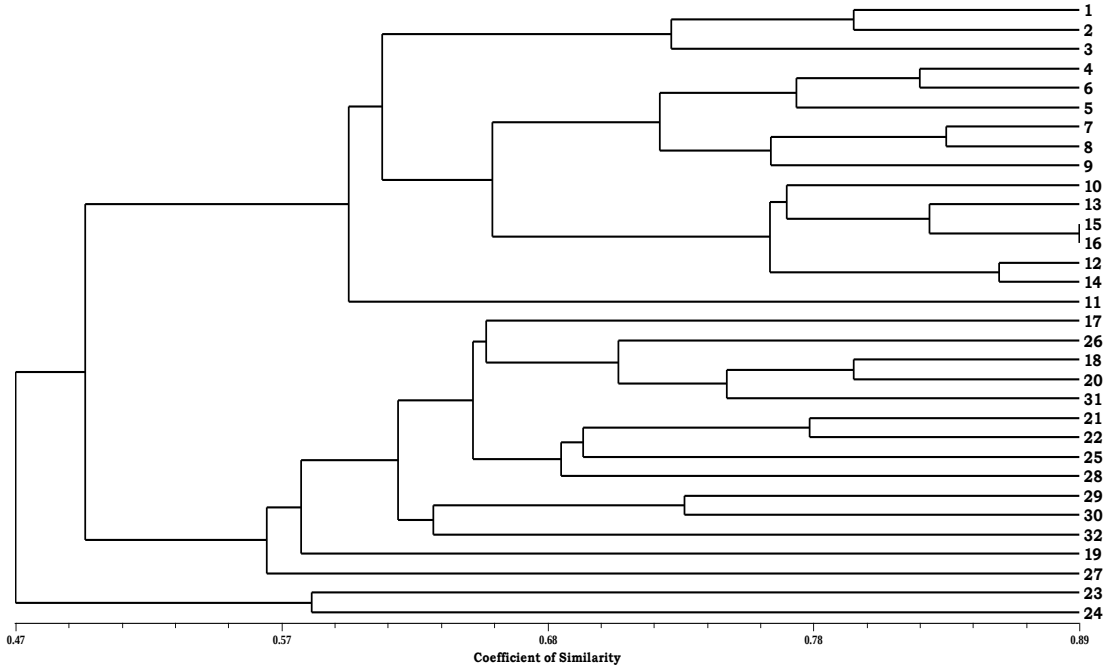
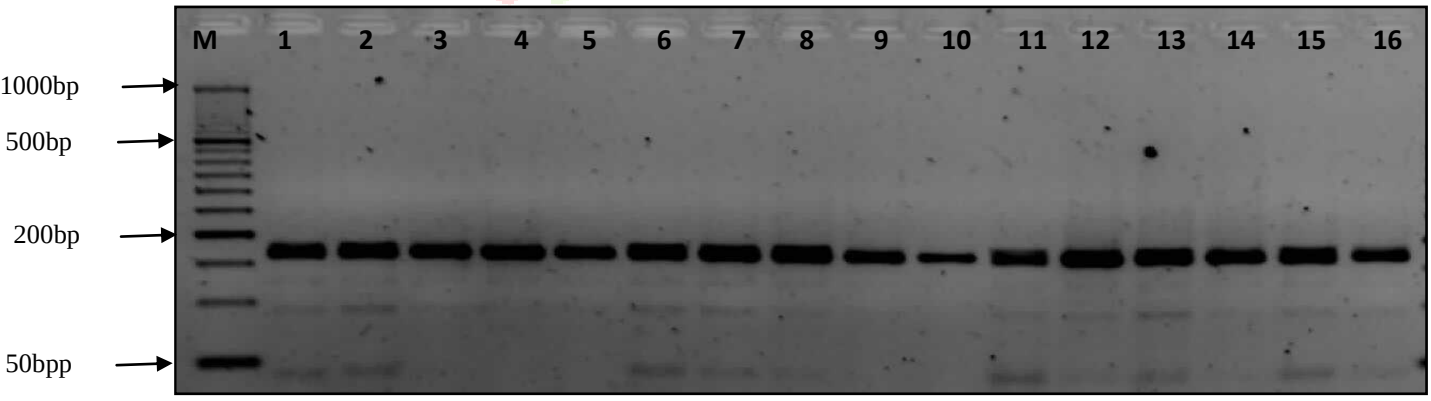


Figure 2: Dendrogram showing clustering of 32 Rice genotypes constructed using UPGMA based on DICE similarity coefficient obtained from SSR analysis

1=RRT IRST 1; 2=RRT IRST 2; 3=RRT IRST 4; 4=RRT IRST 7; 5=RRT IRST 10; 6=RRT IRST 13; 7=RRT IRST 14; 8=RRT IRST 16; 9=RRT IRST 19; 10=RRT IRST 39; 11=RRT IRST 41; 12=RRT IRST 44; 13=RRT IRST 45; 14=RRT IRST 47; 15=RRT IRST 48; 16=RRT IRST 50; 17=JOT 21224; 18=IR 81429-B-31; 19=HB-1; 20=GR-11 ; 21=GURJARI; 22=RRT GR-12; 23=KRISHNA KAMOD; 24=PANKHALI-203; 25=GAR-1; 26=LALKADA; 27=DHANHAR; 28=IR-64; 29=A(004); 30=077; 31=BC1; 32=TSP.

Insertion Deletion Analysis (InDel) analysis

Total of 5 InDel markers for gene Rc was present in chromosome 7. Screening of these markers with 24 red rice genotypes and 8 white rice genotypes was conducted, out of which 4 InDel markers showed expected results. Highest numbers of polymorphic bands were observed in RID13 while RID19 produced minimum number of bands. PIC (Polymorphic Information Content) value ranged from 0.52 in RID25 to 0.80 in RID12. The 4 InDel markers amplified a total of 156 bands with an average of 39 bands per marker.



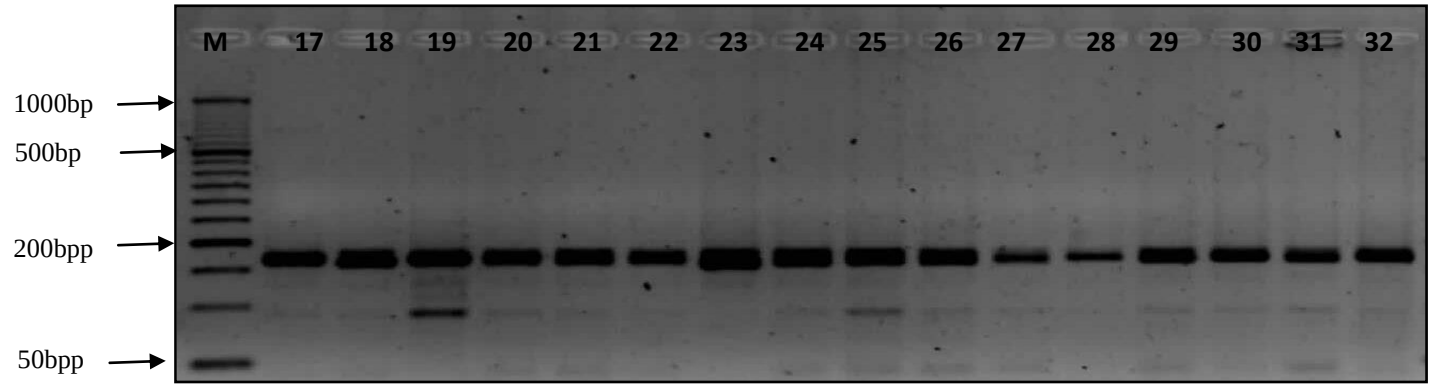


Figure 3: InDEL Profile of RID 19.
M = 50 bp DNA ladder; 1=RRT IRST 1; 2=RRT IRST 2; 3=RRT IRST 4; 4=RRT IRST 7; 5=RRT IRST 10; 6=RRT IRST 13; 7=RRT IRST 14; 8=RRT IRST 16; 9=RRT IRST 19; 10=RRT IRST 39; 11=RRT IRST 41; 12=RRT IRST 44; 13=RRT IRST 45; 14=RRT IRST 47; 15=RRT IRST 48; 16=RRT IRST 50; 17=JOT 21224; 18=IR 81429-B-31; 19=HB-1; 20=GR-11 ; 21=GURJARI; 22=RRT GR-12; 23=KRISHNA KAMOD; 24=PANKHALI-203; 25=GAR-1; 26=LALKADA; 27=DHANHAR; 28=IR-64; 29=A(004); 30=077; 31=BC1; 32=TSP.

The InDel marker RID12 produced maximum alleles while RID13, RID19 and RID25 produced equal of alleles. The average no of allele amplified by 4 markers were 4. Based on the data, cluster analysis was performed using genetic similarity values and a dendrogram was generated showing genetic relationships among these genotypes. The similarity coefficient range was found 0.29 to 1.00. The details of amplification products are given in the Table 3.

Table 4: Results of InDel marker analysis

Sr. No	Marker	Chromosome No.	No of alleles amplified	Molecular Weight Range	Total no of bands	PIC value
1	RID12	7	7	160-254	36	0.80
2	RID13	7	3	273-378	52	0.59
3	RID19	7	3	159-172	32	0.56
4	RID25	7	3	251-292	36	0.52
Total			16	-	156	2.47
Average	-		4	-	39	0.61

Allele frequencies and polymorphism

Sixteen alleles were observed by four InDel markers among the 32 genotypes sampled from two populations, averaging 4 per locus. The number of alleles ranged from three at RID 13, RID 19 and RID 25 to seven at RID 12. It was noted that all alleles obtained were polymorphic because they fulfilled the requirement of having allele frequencies below 0.99. It was observed that allele frequency varied between two populations. Certain alleles showed higher frequency in red rice population. Allele having molecular weight 286 bp in marker RID 13 showed highest allelic variation while allele having molecular weight of 245 bp in marker RID 12 showed least allelic variation between two populations. All four InDel markers were successful in distinguishing red and white rice genotypes these results complimented to results obtained by Sweeny *et al*., (2006).

The level of polymorphism among the genotypes was evaluated by calculating the PIC values for each 21 microsatellite markers evaluated. The highest PIC (Polymorphic Information Content) value obtained was 0.80 of marker RID 12 while the lowest value was 0.52 by marker RID 25.

The PIC values indicated that RID 12 might be the best marker for diversity analysis of rice genotypes, followed by RID 13, RID 19 and RID 25. Observed and expected heterozygocity varied among population to a higher extent. Observed heterozygocity was higher than expected heterozygocity in all four markers. Observed heterozygocity was higher in red rice genotypes in all cases. Deviation in observed heterozygocity is population specific in all cases. From the

present study, it can be said that the best InDel marker to characterize red rice genotypes is RID 12 followed by RID 13, RID 19, and RID 25.

Population genetic structure

AMOVA analysis showed that the two populations under observation contributed 24% variation while 76% variation was contributed by distinct genotypes used in the present study. This result varies from AMOVA obtained from microsatellite data as InDel markers used in the present study were restricted to the coloration gene Rc, which is mutated and hence non functional in white rice genotypes. Hence 24% variation between two population was contributed due *Rc* gene. This Permutation test (based on 999 permutations) suggest that the overall ΦPT was not significantly different from the null distribution (ΦPT = 0.245, P=0.010), which indicates the differences among genotypes are non significant.

In the two populations of *O. sativa* L. detected by InDel markers, *Fst* was 0.124, indicating that the total genetic variation existed because of differences among populations was more. Low *Fis* value (*Fis* = -0.124) indicates that there was less deviation from the Hardy Weinberg's Equation.

The results based on both *Fst* and AMOVA demonstrated that major genetic variation existed within red rice populations and also between populations for genes responsible for coloration. Variation between populations can be beneficial to characterize both populations for introgression study. Broad genetic variation in the red rice gene pool may provide more opportunities for selecting beneficial genetic resources for rice breeding.

Table 5: Analysis of molecular variance (AMOVA) for 32 rice populations was analyzed. 24 individuals from population 1 and eight individuals from population 2 were evaluated for 16 alleles amplified by four InDel primer pairs.

Source	df	SS	MS	Est. Var.	%
Between Pops	1	3.813	3.813	0.253	24%
Within Pops	30	23.375	0.779	0.779	76%
Total	31	27.188		1.032	100%
Stats	Value		P(rand >= data)		
ΦPT	0.245		0.010		

Genetic variability

From POPGENE analysis, out 16 alleles identified for 4 InDel markers, the lowest gene frequency, 0.0312 and highest gene frequency 0.9688 was observed. . This data also shows a high variation in the gene frequency obtained by four InDel markers.

Effective number of alleles is the number of alleles that can be present in a population. From the present data, 1.3948 alleles would be expected in a locus in each population used in present study. Gene diversity ranged from 0.0312 to 0.9688 while Shannon's information index data ranged from 0.6616 to 0.1391. From the above data it can be concluded that the genotypes used in this study were distinct from each other. This result is supported by microsatellite result and data obtained by AMOVA for InDel marker.

Cluster Analysis (InDel)

The similarity matrix index based on DICE coefficient of 32 rice genotypes ranged from 0.29 to 1.00. The level of polymorphism among the genotypes was evaluated by calculating the number and PIC values for each 4 InDel markers evaluated. Among 4 InDel markers used the PIC value ranged from 0.52 (RID25) which was lowest and highest PIC value of 0.80 was obtained with RID12.

The dendrogram divides the genotypes into two major groups. The first major (A) group included IRST1, IRST2, IRST4, IRST7, IRST10, IRST13, IRST14, IRST16, IRST19, IRST39, IRST41, IRST44, IRST45, IRST47, IRST48, IRST50, A(004), JOT 21224 and IR-64. Line between the genotypes IRST1 and IRST39, IRST14 and IRST48, IRST 9 and A (004), IRST 47 and IRST50, JOT 21224 and IR-64, IRST4, IRST7 and IRST11, IRST10 and IRST13 indicates that they have highest genetic similarity. First cluster (A) was further sub-divided into two minor clusters, first minor cluster (A1) comprised IRST1, IRST2, IRST14, IRST19, IRST39, IRST44, IRST45, IRST47, IRST48, IRST50, A(004), JOT 21224 and IR-64., whereas second (A2) comprised IRST4, IRST7, IRST 41, IRST10, IRST13 and IRST16. Second major cluster (B) contained GR-11, GURJARI, GR-12, Krishna Kamod, P-203, GAR-1, Lalkada, Dhanhar, A(004), 077. Second major cluster (B) included three minor clusters, first minor cluster (B1) comprised GR-11, Gurjari, GR-12, Krishna

Kamod, P-203, GAR-1.Second minor cluster (B2) included Lalkada, Dhanhar, A (004), 077, BC1 and TSP. Line between GR-11 and Gurjari, GR-12, Krishna Kamod and P-203, Dhanhar and IR-64 indicates that they have highest genetic similarity. First major cluster depicted that black and brown rice is genetically related to red rice (fig 4).

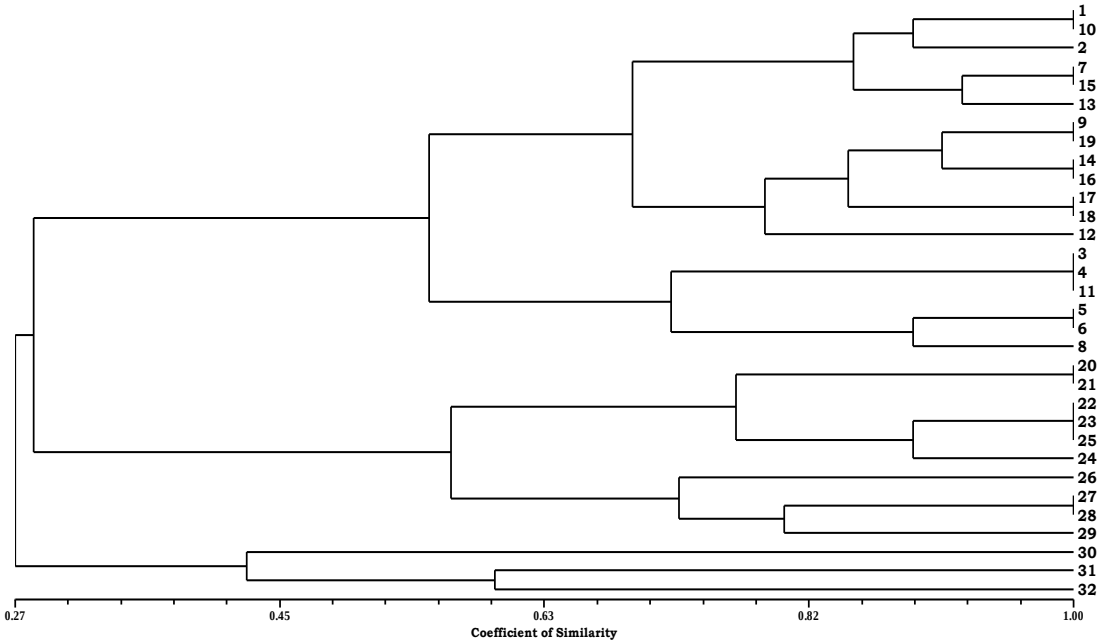


Figure 4: Dendrogram showing clustering of 32 Rice genotypes constructed using UPGMA based on DICE similarity coefficient obtained from InDel analysis.

1=RRT IRST 1; 2=RRT IRST 2; 3=RRT IRST 4; 4=RRT IRST 7; 5=RRT IRST 10; 6=RRT IRST 13; 7=RRT IRST 14; 8=RRT IRST 16; 9=RRT IRST 19; 10=RRT IRST 39; 11=RRT IRST 41; 12=RRT IRST 44; 13=RRT IRST 45; 14=RRT IRST 47; 15=RRT IRST 48; 16=RRT IRST 50; 17=JOT 21224; 18=IR 81429-B-31; 19=HB-1; 20=GR-11 ; 21=GURJARI; 22=RRT GR-12; 23=KRISHNA KAMOD; 24=PANKHALI-203; 25=GAR-1; 26=LALKADA; 27=DHANHAR; 28=IR-64; 29=A(004); 30=077; 31=BC1; 32=TSP.

The correlation between the similarity values measured using two marker systems

The similarity index values for SSR ranged from 0.49 to 0.89. Highest similarity (0.89) was observed between IRST 48 and IRST 50 Similarity index for InDel ranged from 0.29 to 1.00. Highest similarity (1.00) was found between IRST 1 and IRST 4 Thus it can be concluded that there was less diversity detected by InDel as it showed high similarity value.

Mantels and Cophenetic Correlation test

The values of Mantels test correlation among two marker systems were 0.48 between SSR and INDEL. The test indicated that the clusters formed based on INDEL and SSR marker was not conserved since matrix correlation value was 0.49 against the minimum required value of 0.80. Cophenetic correlation value for SSR was found to be 0.89 and 0.81 for InDel. These cophenetic correlation values indicated that there was high level of diversity detected by SSR (0.89) than InDel (0.81) as the number of InDel markers used in study was less and target site was more specific than microsatellite markers.

Hence the overall results indicated that the application of SSR and InDel marker have proved to be useful in determining the genetic diversity among the red rice genotypes and primers used complement each other to get perfect results. SSR markers such as RM5, RM 210, RM 251, RM 261, RM 490, RM 3859 and InDel marker RID 12, RID 13, RID 19 and RID 25 were capable to distinguish colored rice genotypes. This study would contribute to predict potential genetic gains of coloured rice.

Reference

1. **Ahmadikhah, A. and G. I. Karlov. 2006.** Molecular mapping of the fertility restoration gene *Rf4* for *WA* cytoplasmic male sterility in rice. *Plant Breed*. **125**:363-367.

2. **Bate-Smith, E.C. 1973.** Tannins of herbaceous leguminosae. *Phytochemistry*.**12**: 1809–1812.

3. **Buu, B. B. and N. Lang. 1999.** Using molecular markers in study of rice genetic diversity. *Omonrice*. **7**:15-25.

4. **Cao, Q., Lu, B., Hui, X., Rong, J., Sala, F. and Grassi, F. 2006**. Genetic Diversity and Origin of Weedy Rice (*Oryza sativa* f. spontanea) Populations Found in North-eastern China Revealed by Simple Sequence Repeat (SSR) Markers. *Annals of Botany*. **98**: 1241–1252.

5. **Dhurai, S., Bhatil, S. P. and Saroj S. 2014**.Studies on genetic variability for yield and qualitycharacters in rice (*Oryza sativa* L.) under integrated fertilizer management.*The Bioscan.* **9**(**2**): 745-748.

6. **Excoffier, L., Smouse, P.E. and Quattro, J.M. 1992.** Analysis of molecular variance inferred from metric distances among DNA haplo-types: application to human mitochondrial DNA restriction data. *Genetics*.**131**: 479-491.

7. **Estorninos, L. E., Burgos, N. R., Gealy, D. R., Stewart, J. M. and Talbert, R. E. 2000.** Genetic characterization of red rice populations using molecular markers. Pages 76–84 in R. **Norman, J. and Beyrouty, C. A.** eds. B. R. Wells Rice Research Studies—1999. Series 476. Arkansas: Arkansas Agricultural Experiment Station. pp: 76-84.

8. **Gealy, D.R., Tai, T.H. and Sneller, C.H. 2002**. Identification of red rice, rice, and hybrid populations using microsatellite markers. *Weed Sci*. **50**: 333–339.

9. **Joshi, S.P., Gupta, V.S., Aggarwal, R.K., Ranjekar, P.K. and Brar, D.S. 2000.** Genetic diversity and phylogenetic relationship as revealed by Inter simple sequence repeat polymorphism in the genus. *Oryza. Theor. Appl. Genet.***100**: 1311-1320.

10. **Kato, S. and Ishikawa, J. 1921.** On the inheritance of the pigment of red rice. *Japan J. Genet*.**1**:1–7.

11. **Lee, D., Lupotto, E. and Powell, W. 2009.** G-string slippage turns white rice red. *Genome*, **52**: 490–493.

12. **O'Hanlon, P.C., Peakall, R. and Briese D. 2000.** A review of new PCR-based genetic markers and their utility to weed ecology. *Weed Research***40**: 239–254.

13. **Oki, T., Masuda, M., Kobayashi, M., Nishiba, Y., Furuta, S., Suda, I. and Sato, T. 2002.** Polymeric procyanidins as radical-scavenging components in red-hulled rice. *J. Agric. Food Chem*.**50**: 7524–7529.

14. **Peakall, R., Smouse, P.E. and Huff, D.R. 1995.** Evolutionary implications of allozyme and RAPD variation in diploid populations of buffalograss Buchloe dactyloides (Nutt. Engelm.). *Molec Ecol*. **4**: 135-147.

15. **Peakall, R. and Smouse, P.E. 2006.** GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*. **6**: 288-295.

16. **Prathepha, Preecha. 2011.** Microsatellite analysis of weedy rice (*Oryza sativa* f. spontanea) from Thailand and Lao PDR. *AJCS*. **5**(**1**):49-54.

17. **Ramaiah, K. and Rao. 1953.** Rice Breeding and Genetics. ICAR Science Monograph 19. Indian Council of Agricultural Research, New Delhi, India.

18. **Sweeney, M., Thomson,M., Pfeil, B. and McCouch, S. 2006.** Caught Red-Handed: Rc Encodes a Basic Helix-Loop-Helix Protein Conditioning Red Pericarp in Rice. *The Plant Cell*. **18**: 283–294.

19. **Sweeney, M., Thomson, M., Cho, Y., Park, Y., Williamson, S., Bustamante, C. and McCouch, S. 2007.** Global Dissemination of a Single Mutation Conferring White Pericarp in Global dissemination of a single mutation conferring white pericarp in rice. *PLoS Genet*, **3**:1418-1424.

20. **Vaughan, L. K., Ottis, B. V., Prazak-Havey, A. M., Bormans, C. A., Sneller, C., Chandler, J. M. and Park, W. D. 2001.** Is all red rice found in commercial rice really Oryza sativa? *Weed Sci*.**49**:468–476.

21. **Vhora, Z., Trivedi, R., Chakraborty, S., Ravikiran, R. and Sasidharan, N. 2013.** Molecular studies of aromatic and non aromatic rice (*Oryza sativa* L.) genotypes for quality traits using Microsatellite markers. *The Bioscan*. **8**(**2**): 359-362.

22. **Yu, Yanchun., Tang, Tian., Qian Qian., Wang, Yonghong., Yan, Meixian., Zeng, Dali., Han, Bin., Wu,Chung-I., Shi, Suhua. and Li, Jiayang. 2008.** Independent Losses of Function in a Polyphenol Oxidase in Rice: Differentiation in Grain Discoloration between Subspecies and the Role of Positive Selection under Domestication. *The Plant Cell*. **20**: 2946–2959.

23. **Zeng, Y. X., Wen, Z. H., Ma, L.Y., Ji, Z. J., Li, X. M. and Yang. C. D. 2013.** Development of 1047 insertion deletion markers for rice genetic studies and breeding. *Genet. Mol. Res*. **12**(**4**):5226-5235.

24. **Zhou, H., Xie, Z. and Ge, S. 2003.** Microsatellite analysis of genetic diversity and population genetic structure of a wild rice (*Oryzarufipogon* Griff.) in China. *Theor. Appl. Genet*. **107**:332-339.