African Journal of Biotechnology Vol. 12(7), pp. 658-664, 13 February, 2013
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB12.2015
ISSN 1684–5315 ©2013 Academic Journals

Full Length Research Paper

Genetic diversity of important rice cultivars of Kashmir valley using microsatellite markers

M. Habib*, K. A. Bhat, M. A. Bhat, G. A. Parray, Z. A. Dar, Gul Zaffar and A. Sabina

Molecular Biology Laboratory Division of Plant Breeding and Genetics, Sher-e-Kashmir University of Agricultural Sciences and Technology, Shalimar, 191121, India.

Accepted 24 January, 2013

In the present investigation, five microsatellite (SSR) markers namely RM-2,RM-27,RM-72, RM-107 and RM-154 were used to estimate the genetic diversity of eight *indica* rice cultivars significant for rice breeding programme in the temperate Kashmir Province of India. The SSR primers used, specific to five different chromosomes, revealed distinct polymorphism among the cultivars studied indicating the robust nature of micro satellite markers in revealing polymorphism. Twenty-three (23) alleles were scored for all the SSR primers across the genotypes with a mean value of 4.6. A total of 40 bands were scored and all were polymorphic. No primer revealed a monomorphic band in the present investigation. Number of alleles per locus varied from four (RM2) to six (RM127). Similarity coefficient (DICE) ranged from 0.2 to 0.6 with an average of 0.35. Polymorphism information content (PIC) value for the primers studied ranged from 0.22 (RM27) to 0.82 (RM154) with an average value of 0.62. No null allele was reported in the present investigation.

Key words: Genetic diversity, rice, microsatellite markers, Kashmir.

INTRODUCTION

Rice is one of the agronomically and nutritionally important cereal crops and is the principal staple food in developing countries in general and Kashmir in particular. Approximately one-third of the world's population relies on rice for the significant portion of their food. However, rice yields have plateaued during the last two and a half decades (Virmani, 1994), which may have resulted from the narrow genetic base of the released varieties (Carmona, 1990). This is an alarming situation and is a prime concern for breeders. Crosses between genetically diverse parents are, therefore, important in hybridization programme to increase variability within the genotypes developed.

Several approaches have been used to estimate genetic diversity to select appropriate parental genotypes in crossing programme. A commonly used measure of genetic similarity is coefficient of parentage (CP) or co-ancestry, which is defined as the probability that a random allele of one individual is identical by descent to a random allele of another individual at the same locus (Malecot, 1948).

Measurement of morphological and biochemical characteristics is another commonly used method to arrive at an estimate of genetic diversity in parental genotypes. However, their use is restricted as these suffer from many disadvantages. Compared to morphological analysis, molecular markers can reveal differences among many accessions/genotypes at the DNA level and thus provide a more direct, reliable and efficient tool for germplasm conservation and management. Besides, the molecular markers are stably inherited and are not affected by developmental stage of the genotypes under investigation.

Several types of molecular markers are now available for evaluating the extent of genetic variation in rice. These include restriction fragment length polymorphism (RFLP) (Botstein et al., 1980), random amplified polymorphic DNA (RAPD) (Welsh and McClelland, 1990; Williams, 1990), amplified fragment length polymorphism

*Corresponding author. E-mail: mehfuza_habib@yahoo.com. Tel: +919419904358.

Table 1. The pedigree record of the parents used.

S/N	Name of the parent/variety	Pedigree
1	SKAU-382	IR-32429-47-3-2-2 x K-448
2	SKAU-389	China-1007 x SKAU-27
3	Chenab (SKAU-23)	K-21-9-10-1 x IR-2053-521-1-1-2
4	Jhelum (SKAU-27)	K-448-1-2 (Jukkoku x IET-1444)
5	Shalimar Rice-1 (SR-1)	K-500-5-3-2-2 (China-1007 x IET-444)
6	SKAU-46	IET-1444 x Hayayuki
7	Ch-1039	Introduction
8	SKAU-5	China-1039 x IR-580-19-2-3-3

(AFLP) (Vos et al., 1995) and microsatellite or simple sequence repeat (SSR) (Tautz, 1989). Of these microsatellites, RFLP are co-dominant markers and their map positions, on the rice genome, are well known, while as RAPD and AFLP markers involve the use of random, largely dominant markers. Microsatellites are PCR-based markers that are both technically efficient and cost effective to use and are available for rice (Chen et al. 1997; Temnykh et al., 2000). Compared with RFLP's, microsatellite markers detect a significantly higher degree of polymorphism in rice (Wu and Tanksley, 1993; Yang et al., 1994), and are especially suitable for evaluating genetic diversity among closely related rice cultivars (Akagi et al., 1997).

The efficiency of rice hybrid breeding programme can be potentially increased if parents of the superior crosses can be predicted, before the actual field evaluation, using molecular markers (Bhat et al., 2006). The prediction of hybrid performance is of considerable importance and has attracted much interest. For this, the initial step will be estimation of genetic diversity using molecular markers.

MATERIALS AND METHODS

Experimental materials

The basic materials for the present study consisted of eight diverse genotypes of rice (*Oryza sativa* L.) namely SKAU-382, SKAU-389, Chenab (SKAU-23), Jhelum (SKAU-27), Shalimar Rice-1 (SR-1), SKAU-46, Ch-1039 and SKAU-5 selected from the germplasm collection maintained at Regional Rice and Research Station (RR and RS), Khudwani on the basis of genetic variability for various agronomic traits and maturity. The pedigree record of the parents used in the present study is given in Table 1.

Genomic DNA extraction

Plants were grown in a greenhouse under controlled conditions of temperature and humidity. Plant DNA was isolated using cetyl trimethyl ammonium bromide (CTAB) method as modified by Maroof-Saghai et al. (1984) with some modification.

Assessment of quality and quantity of DNA

Quantity of DNA was checked by agarose gel electrophoresis. In this, 0.8 g of agarose was dissolved in 100 ml of 0.5X TBE electrophoresis buffer [(Tris base (45 mM), Boric acid (45 mM) and EDTA (1 mM)]. The mixture was heated till the agarose was dissolved completely that is, when solution became transparent and clear. It was cooled down to 55 to 60°C with constant stirring. Ethidium bromide was added to a final concentration of 0.5 µg/ml of buffer. Then, the agarose solution was poured into an already prepared gel mould with combs and was left for 20 to 30 min for solidification. DNA samples for loading were prepared by adding 2 µL loading dye (6X) (0.25% w/v bromophenol blue, 50% glycerol in sterile water) to 2 µL DNA. 6 µL sterile water was added (to make the volume 10 µL) to the DNA such that the final concentration of loading dye was 1X. The DNA samples were loaded into wells with the help of micropipette. Along with the DNA samples, marker of known concentration (uncut λ DNA of 50 ng/µL concentration) was also loaded. The gel was run for about 1 to 2 h at voltage of 5 V/cm. The gel was then visualized under UV light using photo gel documentation system (Bio Rad Pvt. Ltd.) and the DNA samples were photographed. The intensity of fluorescence is directly proportional to the total mass of DNA. The intensity of fluorescence of each sample was compared with that of a standard marker and then DNA concentration of each sample was ascertained. The quality of DNA samples was judged based on whether DNA formed a single high molecular weight band (good quality) or a smear (degraded/poor quality). After the quantification, the DNA was diluted in multi-well dilution plate with sterile water such that the final concentration of DNA became 20 ng/ µL.

SSR analysis

Selection of primers

The SSR markers utilized in the present study were obtained from Genetics Biotech Asia Private Limited, New Delhi. A set of five microsatellite markers, covering five different chromosomes were selected from website www.gramene.org using the SSRTI tool. The selected microsatellite markers along with their chromosomal location are presented in the Table 2.

PCR amplification

In vitro amplification using polymerase chain reaction (PCR) was performed in a 96 well microtiter plate (Bio Rad Pvt. Ltd.) using 75 ng of genomic DNA of each genotype in a final volume of 20 µL per reaction. It contained 2 µL of 1× buffer, 1.5 mM of $MgCl_2$, 100 µM of dNTPs, 0.25 µM of each farward and backward primer, 1 unit of taq Polymerase and 75 ng of template DNA. The reaction mixture was placed in a 96 well thermal cycler. Amplification was performed using temperature profile programmed to 1 cycle of 4 min., at 94°C as an initial hot start and strand separation step. This was followed

Table 2. Sequence information and polymorphism information content (PIC) values of SSR markers.

S/N	Primer	Chromosome Number	Primer sequence		Repeat motif	PIC value	Number of alleles/locus
			Forward	Reverse			
1	RM-2	7	ACGTGTCACCGCTTCCT	ATGTCCGGGATCTCATCG	$(GA)_{13}$	0.75	4
2	RM-27	4	GTGGGATAGCTGCGTCGCGTCG	AGGCCAGGGTGTTGGCATGCTG	$(AGG)_8$	0.22	6
3	RM-72	8	CCGGCGATAAAACAATGAG	GCATCGGTCCTAACTAAGGG	$(TAT)_5C(ATT)_{15}$	0.65	5
4	RM-107	9	AGATCGAAGCATCGCGCCCGAG	ACTGCGTCCTCTGGGTTCCCGG	$(GA)_7$	0.66	4
5	RM-154	2	ACCCTCTCCGCCTCGCCTCCTC	CTCCTCCTCCTGCGACCGCTCC	$(GA)_{21}$	0.82	4

by 35 cycles of 1 min at 94°C for denaturation, 2 min at 55 to 60°C for annealing and 2 min at 72°C for primer elongation and finally 1 cycle of 7 min at 72°C. After amplification to 20 µL of the amplified product, 3.3 µL of 6X loading dye was added so as to make the final concentration of the loading buffer in the reaction samples to 1X. The PCR products were resolved on 3.5 to 4.0% agarose gel prepared in 0.5X TBE buffer with 1 kb ladder as molecular marker (Gibco BRL). Ethidium bromide was added at concentration of 0.5 g/L. The gel was run for 3.5 h at 85 volts and visualized under UV transilluminator and photographed by using BIO-RAD gel documentation system.

Scoring of SSR allele profile

All the genotypes were scored for the presence or absence of the SSR bands, and the data were entered into a binary matrix as discrete variables namely; 1 for presence and 0 for absence of the character. The data matrix was subjected to further analysis. The Excel file containing the binary data was imported to NT Edit of NTSYS-pc 2.02e. The 0/1 matrix was used to calculate similarity using DICE coefficient and SIMQUAL subroutine in SIMILARITY routine. The resultant similarity matrix was employed to construct dendrograms using Sequential Agglomerative Hierarchical Nesting (SAHN) and Unweighted Pair Group Method with Arithmetic Means (UPGMA) options to infer genetic relationships.

Polymorphism information content (PIC) values

The PIC values described by Botstein et al. (1980) were used to refer to the relative value of each marker with respect to the amount of polymorphism exhibited. PIC values for each of the 5 primers were estimated using formula given by Nei (1973).

$$PIC = 1 - \sum_{i=1}^{n} (P_{ij})^2$$

Where P_{ij} is the frequency of jth allele in ith primer and summation extends over 'n' patterns. PIC is synonymous with the term 'gene diversity' as described by Weir (1990). The PIC takes into account not only the number of alleles that are expressed but also the relative frequencies of those alleles (Smith et al., 1997).

Cluster analysis

Cluster analysis of the germplasm, using data generated by microsatellite markers was conducted using computer software programme Numerical Taxonomic and Multivariate Analysis System (NTSYS-pc) version 2.02e (Rohlf, 1997). Microsatellite marker amplification profile for all the genotypes under study was compared to each other and DNA fragments were scored as present (scored as 1) or absent (scored as 0). Data from all the five primers were used to estimate the similarity based on the number of shared amplified bands. Similarity matrix value based on Nei and Li (1979) coefficient of similarity (D_{ij}) were calculated as:

$$D_{ij} = 2a/(2a+b+c)$$

Where, 'a' represents matched fragments, b and c are unmatched fragments. The 2a+b+c are the total number of fragments amplified in a particular set. The similarity matrix was thus generated and dendrogram was constructed using UPGMA available in NTSYS. In UPGMA similarity/dissimilarity between a genotype and an established cluster is the average similarity/dissimilarity of that genotype with all genotypes in that cluster. This means that after fusion of two most similar genotypes, clustering continues between two next closest genotypes or between any unplaced genotypes and the established cluster. An unplaced genotype can join a cluster if its average similarity to all members of the cluster is small enough in comparison with any other pair of unplaced genotype. The process is repeated until all clusters join one cluster. This is an unweighted method, because it gives equal weight to each genotype within a cluster. It is considered as only a 'moderate type of clustering' (Sneath and Sokal, 1973; Panchen, 1992).

RESULTS AND DISCUSSION

Genetic diversity analysis using SSR markers

In the present study, a set of five SSR markers, covering five different chromosomes, were used to assess the genetic diversity among eight parental lines of rice belonging to *indica* type. All the five RM primers showed polymorphism among the eight rice cultivars. A total of 40 bands were scored and all the bands were found to be polymorphic. A total of 23 alleles were detected in the eight cultivars studied. Summarized data for

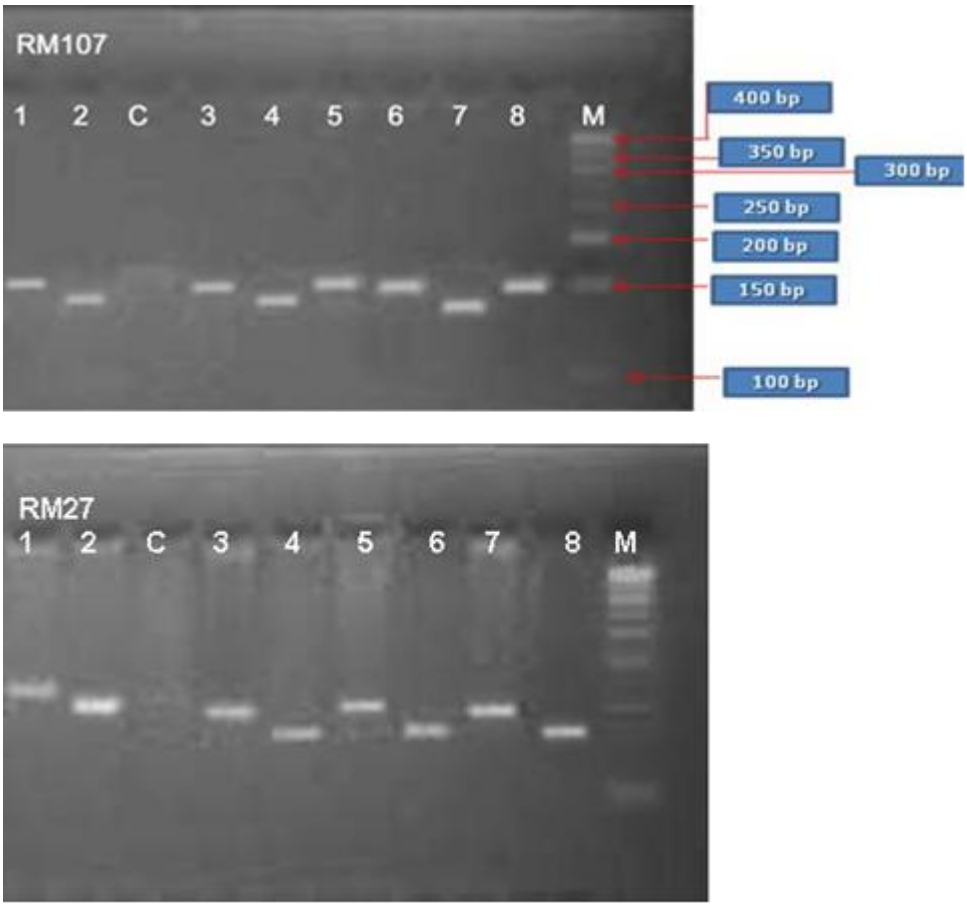

Plate 1. Ethidium bromide stained DNA amplification profile of the eight parental lines of rice (*O. sativa L.*) using SSR markers. M, ladder marker; C, control.

the number of alleles detected and the polymorphism information content (PIC) values for each of the 5 SSR primers are presented in the Table 2. The average number of SSR alleles per locus was 4.6, with a range of four alleles (RM-2) to six alleles (RM-27) (Plates 1 and 2). The genetic similarity (GS) between pairs of cultivars averaged 0.33, with a range of (0.2) to (0.6). PIC values for SSR loci across all the 8 genotypes ranged from 0.22 (RM-27) to 0.82 (RM-154) with an average of 0.62. A comparison of PIC values for different repeat sizes revealed that SSR primers with di-repeats revealed a maximum mean PIC value of 0.74, followed by tri-repeats (0.43). The present investigation revealed that the di-repeats with larger size, such as [$(GA)_{21}$–RM-154] had higher PIC value as compared to shorter di-repeats [$(GA)_{7}$–RM-107)]. The level of polymorphism displayed by the SSR markers was high in the present analysis as reflected by the mean PIC value of 0.62 and the highest PIC value of 0.81 (RM-154).

Analysis of genetic relationship among cultivars

Genetic distance (GD) between the cultivars was studied using Dice coefficient (Nei and Li, 1979) available in NTSYSpc 2.02e (Rohlf, 1997). The GD matrix was used as input file for cluster analysis using clustering algorithms- UPGMA available in SAHN module. Cluster analysis was used to group the varieties and to construct a dendrogram. The similarity matrix representing the DICE coefficient was used to cluster the data using the UPGMA algorithm. A total of three distinct groups resulted out of analysis of pooled SSR marker data (Figure 1). This dendrogram revealed that the genotypes that are derivatives of genetically similar types clustered more together.

Considering coefficient of similarity as 0.60, it was possible to assign the cultivars to three clusters (CI, CII and CIII). The coefficients of similarity are presented in the form of a matrix (8 × 8) (Table 3). Perusal of the similarity values revealed that high degree of similarity to the extent of 60% existed between Ch-1039 and SKAU-46, whereas very low level of similarity of 20% existed between SKAU-382 and SKAU-46; Jhelum and SKAU-389 and SKAU-23 and SKAU-389. The similarity coefficient of 0.20 to 0.40 between most of the parental pairs indicated that the materials in descent are not

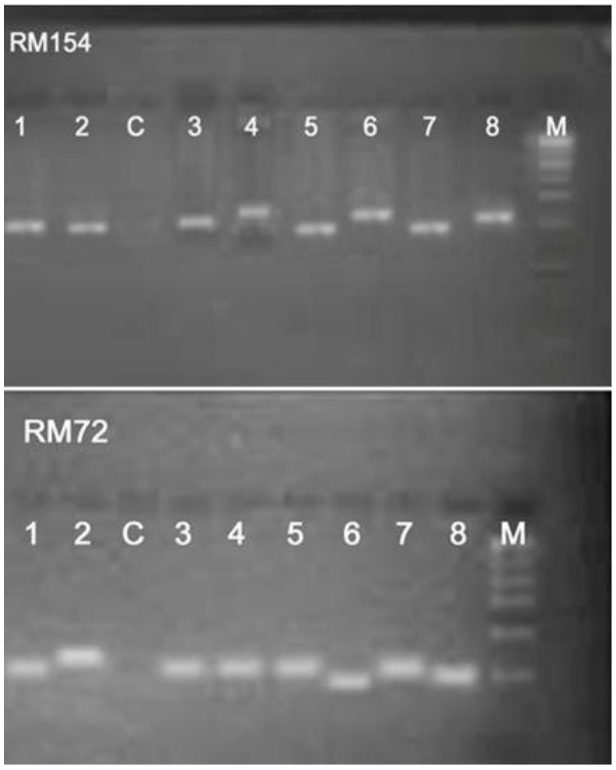

Plate 2. Ethidium bromide stained DNA amplification profile of the eight genotypes of rice (*O.sativa L.*) using SSR markers.

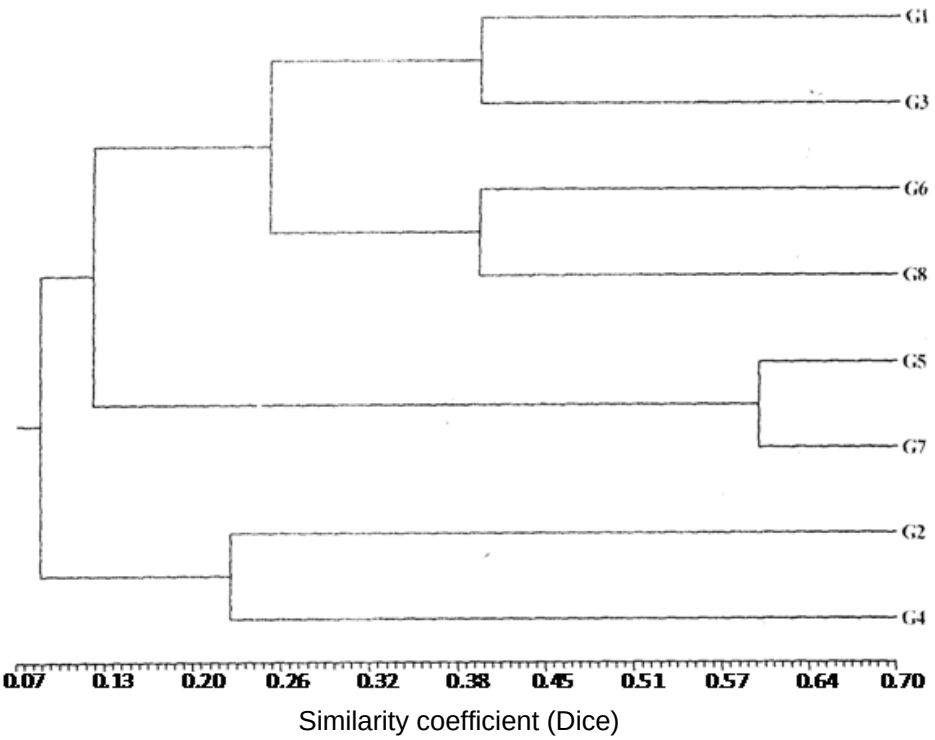

Figure 1. Dendrogram depicting the genetic relationships among the eight genotypes of rice (*Oryza sativa L.*) base on SSR data using UPGMA (Dice coefficient).

Table 3. Similarity matrix (8 × 8) for 8-parental lines used for SSR analysis using UPGMA (Dice coefficient).

Parameter	*G1	G2	G3	G4	G5	G6	G7	G8
G1	1.00							
G2	0.22	1.00						
G3	0.40	0.45	1.00					
G4	0.35	0.26	0.38	1.00				
G5	0.25	0.42	0.23	0.26	1.00			
G6	0.29	0.32	0.26	0.36	0.28	1.00		
G7	0.20	0.42	0.50	0.29	**0.60**	0.23	1.00	
G8	0.28	0.20	0.20	0.25	**0.48**	0.40	**0.47**	1.00

*G, Genotype; G1, SKAU-382; G2, Jhelum; G3, SKAU-23; G4, SR-1 (Shalimar Rice-1); G5, Ch-1039; G6, SKAU-5; G7, SKAU-46;G8, SKAU-389.

related that much and posses diverse genetic background and allelic resources. The possible reason for low similarity coefficient in our investigation can be: i) low number of SSR primers used; ii) only limited number of genotypes were studied; iii) geographical distribution of the varieties; iv) pedigree of the genotypes and common parents in some genotypes.

Introgression and subsequent recovery of pure lines are expected to produce a varietal profile that would be useful so far as future needs of rice production and productivity are concerned. The cultivars assigned to different clusters were as under: Cluster 1: SKAU-382, SKAU-23, SKAU-5 and SKAU-389; Cluster 2: China-1039 and SKAU-46 and Cluster3: Jhelum and SR-1

Polymorphism information content (PIC) value for the primers studied ranged from 0.22 (RM27) to 0.82 (RM154); with an average value of 0.62. Comparison of the PIC values for different repeat sizes revealed that SSR primers with di-repeats depicted maximum mean PIC value (0.72) followed by tri-repeats (0.43). Our results are in conformity with Ni et al. (2002) who also reported that the larger the maximum repeat number in microsatellite DNA, larger is the number of alleles detected. No null allele was reported in the present investigation. In case of null alleles, PCR amplification fails to produce a reproducible band and can arise as a result of point mutation(s) in one or both of the primer sites. Ni et al. (2002) reported low percentage of null alleles in *Oryza rufipogon* and *Oryza nivara*, which are both close relatives of *Oryza sativa*. The possible reason for not detecting null alleles in present study is that only a limited number of SSR primers were used.

Various research workers (Ravi et al., 2003; Sarla et al., 2003; Shivapriya and Hittalmani, 2006; Barooah and Sarma, 2008; Rabbani et al., 2010) have also studied the genetic diversity in rice cultivars using different molecular markers.

REFERENCES

Akagi H, Yokozeki Y, Inagaki A, Fujimura T (1997). Highly polymorphic microsatellites of rice consist of AT repeats, and a classification of closely related cultivars with these microsatellite loci. Theor. Appl. Genet. 94:61-67.

Barooah D, Sarma RN (2008). Genetic diversity analysis in indigenous deep water rice Assam using RAPD and ISSR markers. Indian J. Genet. 68(1):73-76.

Bhat MA (2006). Combining ability estimates and study of genetic diversity among maize (*Zea mays* L.) inbred lines using microsatellite markers. Ph.D. Thesis submitted to PAU Ludhiana, pp. 99-103.

Botstein D, White RL, Skolnik M, Davis RW (1980). Construction of genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:314-331.

Carmona PS (1990). Contribution of INGER to broaden the genetic base of rice in the state of Rio Grande do Sul, Brazil. INGER-Latin America Report. Cali, Colombia: Bilingual Edition, pp. 153-158.

Chen X, Temnykh S, Xu Y, Cho YG, McCouch SR (1997). Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). Theor. Appl. Genet. 95:553-567

Malecot G (1948). Les mathematiques de L'Heridite Masson et Cie; Paris.

Maroof Saghai MA, Soliman KM, Jorgensen RA, Anand RW, (1984). Rhibosomal DNA spacer length polymorphisms in barley: species diversity, chromosomal locations and population dynamics. Proc. Natl. Acad. Sci. USA 91:5466-5470.

Nei M (1973). Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA 70:3321-3323.

Nei M, Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76(10)5269-5273.

Ni J, Colowit PM, Mackill DJ (2002). Evaluation of genetic diversity in rice subspecies using microsatellite markers. Crop Sci. 42:601-607.

Panchen AL (1992). Classification, evolution and nature of biology. Cambridge University Press, Cambridge, England.

Rabbani MA, Masood MS, Shinwari ZK, Shinozaki KY (2010). Genetic analysis of basmati and non-basmati Pakistani rice (*Oryza sativa* L.) cultivars using microsatellite markers. Pak. J. Bot. 42(4):2551-2564.

Ravi M, Geethanjali S, Sameeyafarheen F, Maheswaran M (2003). Molecular marker based genetic diversity analysis in rice (*Oryza sativa* L.) using RAPD and SSR markers. Euphytica 133:243-252.

Rohlf FJ (1997). NTSYSpc: Numerical taxonomy and multivariate analysis systems, version 2.0. Exeter. Software, Setauket.

Sarla N, Bobby S, Siddiq EA (2003). ISSR and SSR markers based on AG and GA repeats delineate geographically diverse Oryza nivara accessions and reveal rare alleles. Curr. Sci. 84:683-690.

Shivapriya M, Hittalmani S (2006). Detection of Genotype specific finger prints and molecular diversity of selected Indian locals and landraces of rice (Oryza sativa L.) using DNA markers. Indian J. Genet 66(1):1-5.

Smith JSC, Chin ECL, Shu H, Smith OS, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J (1997). An evaluation of utility of SSR loci as molecular markers in maize (*Zea mays* L.): Comparison with data from RFLPs and pedigree. Theor. Appl. Genet. 95:163-173.

Sneath PHA, Sokal RR (1973). Numerical Taxonomy. W.H. Trueman

and Co., San Francisco, USA.
Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res 17:6463.
Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (2000). Mapping and genome organization of microsatellite sequences in rice (Oryza sativa L.). Theor. Appl. Genet. 100:697-712.
Virmani, S.S. 1994. Prospects of hybrid rice in the tropics and sub-tropics. In:S.S. Virmani (Ed.), Hybrid Rice Technology: New Developments and Future Prospects pp. 7-19. International Rice Research Institute, Manila, Philippines.
Vos P, Hogers R, Bleeker M, Reijans M, van Dee Lee T, Hornes M, Fritjers A, Pot J, Peleman J, Kuiper M, Zabeau M (1995). AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-4414.
Weir BS (1990). Genetic data analysis : Methods for discrete population genetic data. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts, p. 337.
Welsh J, McClelland M (1990). Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. 18:7213-7218
Williams JGK., Kubelik AR, Livak J, Rafalski A, Tingey SV (199)0. DNA primers identified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18:6531-6535.
Wu KS, Tanksley SD (1993). Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol. Gen. Genet. 241:225-235.
Yang GP, Maroof MA, Xu CG, Zhang QF, Biyashev RM (1994). Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. Mol. Gen. Genet. 245(2):187-194.