Journal of Crop and Weed, 14(3): 94-99 (2018)

Molecular characterization and genetic diversity analysis of aromatic rice (*Oryza sativa* L.) genotypes using SSR markers

V. KUMAR, N. K. RASTOGI, R. R. SAXENA AND [1]S. N. MANDAL

Department of Genetics and Plant Breeding, College of Agriculture
Indira Gandhi Krishi Viswavidyalaya, Raipur-492012, Chhattisgarh
[1]College of Agriculture, Bidhan Chandra Krishi Viswavidyalaya, Burdwan Sadar-713101, West Bengal

Received : 12-04-18 ; Revised : 28-08-18 ; Accepted : 02-09-18

ABSTRACT

Molecular characterization of genotypes gives precise information about the extent of genetic diversity. In the present study, a total of 25 polymorphic SSR markers were used across 24 aromatic rice genotypes from India and abroad for their characterization and discrimination. The allelic richness per locus ranged from 2 to 4 alleles with an average of 2.44 alleles. A total of six rare alleles were identified at six loci among 25 microsatellite loci. The polymorphism information content (PIC) values ranged from 0.278 to 0.642 with an average of 0.465. The PIC values revealed that RM527 might be the best marker for identification and diversity estimation of aromatic rice genotypes followed by RM1, RM22565 and RM207 markers. At a 40% level of similarity, the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) cluster dendrogram showed three major clusters (I, II and III) with additional sub-clusters, except cluster III. The cluster III had only one genotype Bas 837 which showed 63% dissimilarity with rest of the genotypes. At below sub-cluster level, about half of the genotypes exhibited more than 70% genetic similarities between genotypes which itself explains the existence of narrow genetic base among aromatic rice genotypes. The low to high degree of similarity among studied genotypes indicated high level of genetic diversity, and also the possibility of using these genotypes in breeding programmes targeted at improvement of elite aromatic rice varieties.

***Keywords*:** Aromatic rice, genetic diversity, microsatellites and molecular characterization

Rice (*Oryza sativa* L.) is a staple food for about half of the global population. Aroma is considered as one of the most important quality characters of rice and the export value of a rice variety is highly depends on this character. In contrast with the diverse geographical distribution of India, there is a prevalence of highly rich aromatic rice genetic resources. Among the different groups of aromatic rice varieties 'Basmati' is usually grown in north-western India (Nene, 1998). Besides basmati rice, hundreds of aromatic short and medium grained rice are grown in specialized pockets throughout the country (Singh *et al*., 2000).These are short and medium grained having a good aroma, which retained for a longer time when grown in prevailing subtropical warm climate. These indigenous aromatic rice genotypes are graced with high genetic variability and can be vital genetic resources for biotic and abiotic stress tolerance, and improvement of agro-morphological and nutritional characteristics (Bhagwat *et al*., 2008; Chakravorty and Ghosh, 2012). Characterization and improvement of traditional small and medium grained aromatic rice, were somewhat neglected as they lacked export value. There is very little information available on genetic diversity of traditional non-basmati indigenous aromatic rice, therefore, these varieties have to be collected and evaluated for their exploitable genetic variability and conserved. Information on the diversity in plant genetic resources is essential to framing efficient and worthwhile breeding approach and conservation strategies. The extensive utilization of traditional aromatic rice varieties into breeding pipelines could assist in the development of next generation aromatic rice varieties.

Genetic characterization at the molecular level is more promising than at the phenotypic level as the latter is highly influenced by environment. So the study of polymorphism done at the level of the alignment of nucleotide bases in DNA, even seemingly identical accessions could exhibit extensive divergence. For this purpose, identification of different genotypes at the molecular level is imperative. Molecular markers are valuable tools for meticulous characterization and discrimination of genotypes and unravel their evolutionary propinquity. Among the Polymerase Chain Reaction (PCR) based markers, Simple Sequence Repeats (SSRs) are highly polymorphic, more reproducible, co-dominant and distributed throughout the genome (McCouch *et al.*, 2002). In this context, a set of aromatic rice genotypes were evaluated for genetic variation using SSR markers in order to determine genetic diversity among these genotypes and to evaluate their potential utility in the breeding programme.

MATERIALS AND METHODS

Plant materials

A total of 24 aromatic rice genotypes evaluated in the present study (Table 1) were selected from 556 genotypes on the basis of aroma and yield performance (data not shown) from experimental materials received

Email: snmandaledu@gmail.com

from Indian Institute of Rice Research, Hyderabad under Network Project of DBT entitled "Metabolic and Molecular profiling of aromatic rice germplasm of India for gaining insights about aroma". The experiment was conducted at plant molecular biology laboratory, R.H. Richharia Research Laboratory, Indira Gandhi Agricultural University, Raipur (Chhattisgarh) during 2015-16.

Genomic DNA extraction and PCR amplification

Total genomic DNA was extracted from four weeks old seedling leaves from each genotype using the modified CTAB method (Murray and Thompson, 1980). The quality of DNA was also checked by DNA quantification using a Thermo Scientific NanoDrop™ 1000 spectrophotometer (Thermo Fisher Scientific, USA). Twenty seven SSR primer pairs were selected from the list of Cornell's rice microsatellite markers displayed at the Cornell University Rice Genes web site http://www.gramene.org/microsat/ssr.html. PCR was performed in 20 µl reaction mixture containing 2 µl of diluted template DNA, 0.5 µl of each forward and reverse primer, 0.25 µl of 10 mM dNTPs, 2 µl of 10 x buffers, 0.5 µl of *Taq* polymerase and 13.5 µl Nanopure H_2O.About, 2 µl of diluted template DNA (50 çg/µl) of each genotype was dispensed in the bottom of 96 well PCR plates (AXYGEN-MAKE). About 18 µl of the cocktail was added to each tube to make final volume 20 µl. The basic PCR program to amplify DNA was as follow: an initial denaturation step for 5 minutes at 95°C (hot start and strand separation), followed by 34 cycles of 1 minute denaturation at 94°C, 1 minute annealing at the temperature depending on the marker used (55°C, 61°C and 67°C), and 2 minutes primer extension at 72°C and then a final extension at 72°C for 5 minutes. Amplified products were stored at -20°C until further use.

Gel electrophoresis and visualization of amplified products

Prior to electrophoresis, each PCR product was mixed with gel loading dye (bromophenol blue and sucrose) and electrophoresis was carried out in a mini vertical electrophoresis tank (CBS Scientific Co Inc., CA. USA) run on 5 per cent polyacrylamide gel in TBE buffer. Four microliters of the sample were loaded in each well and run at 80 V for 90 minutes. The gel, after electrophoresis was stained with ethidium bromide for 10 minutes, and then visualized under UV light using the Gel documentation unit linked to a PC.

Scoring and analysis of data

Polymorphic product from microsatellite analyses were scored qualitatively for presence (score '1') and absence (score '0') for each marker allele-genotype combination. The data entry was done into binary data matrix as discrete variables. The polymorphic information content (PIC) was calculated, according to the method of Anderson *et al.* (1993).

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

Where, P_{ij} is the frequency of *j*th allele for the *i*th marker, and is summed over *n* alleles. The calculation was based on the number of alleles per locus. Genetic similarities were estimated from the matrix of binary data using Jaccard's coefficient. The similarity coefficients were used for cluster analysis of the rice genotypes using the unweighted pair group method with arithmetic averages (UPGMA). The analysis and dendrogram construction were performed using the NTSYS-pc version 2.2 (Rohlf, 2002). For further confirmation, the genetic diversity was also analyzed using Shannon diversity index (Hutchenson, 1970). Shannon diversity index (H) was calculated by the following formula :

$$H = \sum_{j=1}^{n} - P_{ij}^{*} \ln P_{ij}$$

Where,

H = Shannon diversity index

$P_{ij} = P_{ij}$ is the frequency of j^{th} allele for the i^{th} marker

n = number of alleles

Σ = sum from allele 1 to allele n

ln = natural logarithm

RESULTS AND DISCUSSION

Number of alleles, frequent and rare alleles

A total of 61 alleles were detected for the 25 SSR markers used across 24 aromatic rice genotypes (Table 2). The allelic richness per locus varied from 2 to 4 alleles with an average of 2.44 alleles. Among the polymorphic markers, 16 markers produced two alleles each, 7 produced three alleles each, and 2 generated four alleles each. The maximum number of alleles was observed by RM1 and RM22565. The result revealed that all the primers showed distinct polymorphisms among the rice genotypes under study, showing the ability of microsatellites in detecting polymorphism. The frequency of the most common allele at each locus ranged from 41.20% (RM527) to 83.30% (RM210, RM 316 and RM455). On an average, 65.20% of the aromatic rice genotypes shared a common major allele at any given locus. A total of six rare alleles were identified at six loci among 25 microsatellite loci, with an average of 0.63 rare alleles per locus (Table 2). The numbers of rare alleles were present in aromatic rice accessions Super Basmati, Pusa Basmati 1121, Maguraphulla,

Table 1: List of aromatic rice genotypes used in present study

Sl. No.	Genotypes	Source/ Contribution	Sl. No.	Genotype	Source/Contribution
1.	Maguraphulla	OD/IIRR	13.	Loktimachi	OD/IIRR
2.	Basuabhava	OD/IIRR	14.	Shitalbhog	WB/IIRR
3.	Kapoor Kranti	RPR/IIRR	15.	Domsaih	IRAN
4.	AS GPC-19	UP/IIRR	16.	Basmati 564	J&K/IIRR
5.	Rajabhog	MP/IIRR	17.	Pusa Sugandh-5	IARI/IIRR
6.	Loung Choosi- B	UP/IIRR	18.	Vasumati	IIRR
7.	Govindbhog	WB/IIRR	19.	R-1498-778-388-1-1	IIRR
8.	Kamini Joha	AS/IIRR	20.	NDR-8497-2	UP/IIRR
9.	KDML 105	THILAND	21.	Basmati 11	PU/IIRR
10.	Bindli	NAGINA	22.	Karpura Basa	OD/IIRR
11.	Improved Sarbati	PANTNAGAR	23.	Bas 837	US PATENTED LINE
12.	Super Basmati	IARI/IIRR	24.	Pusa Basmati 1121	IARI/IIRR

Abbreviations: *AS- Assam, OD- Odisha, IARI- Indian Agricultural Research Institute, IIRR- Indian Institute of Rice Research, J&K- Jammu & Kashmir, MP- Madhya Pradesh, PU- Punjab, RPR- Raipur, UP- Uttar Pradesh, WB- West Bengal*

Table 2: SSR markers with chromosome locations, number of alleles per locus, number of rare alleles, major allele frequency, gene diversity and polymorphism information content (PIC) in promising aromatic rice accessions

S. No.	Marker	Chr. No.	No. of alleles	No. of rare alleles	Allele frequency	Gene diversity*	PIC
1.	RM 1	1	4	1	0.430	1.070	0.612
2.	RM 5	1	2	0	0.542	0.690	0.496
3.	RM 283	1	2	0	0.583	0.679	0.486
4.	RM 154	2	3	1	0.750	0.675	0.392
5.	RM 207	2	3	1	0.500	0.836	0.538
6.	RM 514	3	3	1	0.565	0.826	0.525
7.	OSR 13	3	2	0	0.565	0.685	0.491
8.	RM 162	6	2	0	0.667	0.636	0.444
9.	RM 527	6	3	0	0.412	1.060	0.642
10.	RM 11	7	3	1	0.583	0.815	0.517
11.	RM 248	7	2	0	0.583	0.679	0.486
12.	RM 455	7	2	0	0.833	0.450	0.278
13.	RM 25	8	2	0	0.667	0.636	0.444
14.	RM 152	8	2	0	0.792	0.512	0.330
15.	RM 210	8	2	0	0.833	0.450	0.278
16.	RM 408	8	2	0	0.583	0.679	0.486
17.	RM 447	8	3	1	0.625	0.792	0.496
18.	RM 22565	8	4	0	0.608	1.010	0.578
19.	RM 22710	8	3	0	0.652	0.841	0.499
20.	RM 215	9	2	0	0.542	0.690	0.496
21.	RM 242	9	2	0	0.583	0.679	0.486
22.	RM 316	9	2	0	0.833	0.450	0.278
23.	RM 19	12	2	0	0.667	0.636	0.444
24.	RM 277	12	2	0	0.652	0.646	0.454
25.	RM 519	12	2	0	0.652	0.646	0.454
	Total		**61**	**6**	**–**	**–**	**–**
	Average		**2.44**	**0.24**	**0.628**	**0.711**	**0.465**

**Computed based on Shannon's diversity index*

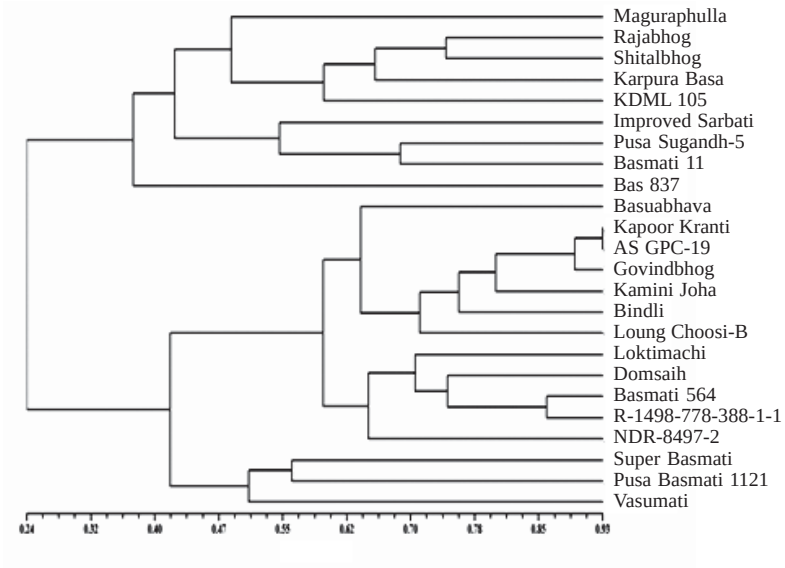

Fig. 1: Dendogram of 24 aromatic rice genotypes derived from UPGMA cluster analysis using Jaccard's coefficient based on twenty five polymorphic SSR markers

Improved Sarbati and Bas 837. Rare allele for Bas 837 was detected by RM154 (160 bp) and RM447 (146 bp). Rare allele for Maguraphulla, Super Basmati, Pusa Basmati 1121 and Improved Sarbati was detected by RM207 (65 bp), RM1 (105 bp), RM11 (120 bp), RM514 (320 bp), respectively. These genotype specific markers may be worthwhile in establishing varietal identity. Rahman *et al.* (2012) reported a total of 57 rare alleles at 24 loci, while a total of 147 rare alleles at 89 loci were reported by Lapitan *et al.* (2007) which was markedly higher than the present result. PCR amplification profile of 24 aromatic rice genotypesusing SSR markers RM19, RM527, RM207 and RM447 are presented in the fig. 2.

Polymorphism and diversity

The polymorphism information content (PIC) values ranged from 0.278 to 0.642 with an average of 0.465 (Table 2). The highest PIC value 0.642 was obtained by RM527 followed by RM1 (0.612), RM22565 (0.578), RM207 (0.538) and RM514 (0.525). The PIC values revealed that RM527 might be the best marker for identification and diversity estimation of aromatic rice genotypes followed by RM1, RM22565 and RM207 markers. The primers showed an average PIC value of 0.465 which indicated that SSR markers used in this study were not highly informative because only PIC values higher than 0.5 indicate high polymorphism. In present study, the level of polymorphism determined by the PIC value (mean= 0.465) corroborated with the reported mean PIC value in previous works (Pal *et al.*,2004; Hossain *et al.*, 2007; Sajib *et al.*, 2012). However, Nadia *et al.* (2014) reported the average PIC value of 0.84, which is markedly higher than the present average PIC value.

Using Shannon's diversity index, an overall genetic diversity of 0.711 was obtained from the analysis, indicating a high level of genetic variation among these genotypes (Table 2). Values for gene diversity explained the probability that two randomly chosen alleles originated from different ancestors and mirrored those of PIC. Similar genetic diversity of 0.71 was reported by Lapitan *et al.* (2007) through evaluation of twenty four rice cultivars carrying good quality traits using 164 SSR markers.

Clustering of aromatic rice genotypes

A cluster analysis using UPGMA based on Jaccard's similarity coefficients was done to study the phylogenetic relationships among the different aromatic rice genotypes. At a 40% level of similarity, the UPGMA cluster diagram showed three major clusters (I, II and III) with additional sub-clusters, except cluster III (Fig. 1). This dendogram revealed that the genotypes derived from a genetically similar type clustered together. Cluster I comprised 15 genotypes and formed the largest cluster, whereas cluster II comprised 8 genotypes. The cluster III had only one genotype Bas 837 (US patented line) which showed 63% dissimilarity with rest of the genotypes.

The first cluster was sub-divided into two clusters, IA and IB consisting of three and twelve genotypes, respectively. Of these, the sub-cluster IA included three evolved Basmati *viz.*, Pusa Basmati 1121, Super Basmati and Vasumati with 51% overall similarity within the sub-cluster. The sub-cluster IB consisted of twelve accessions

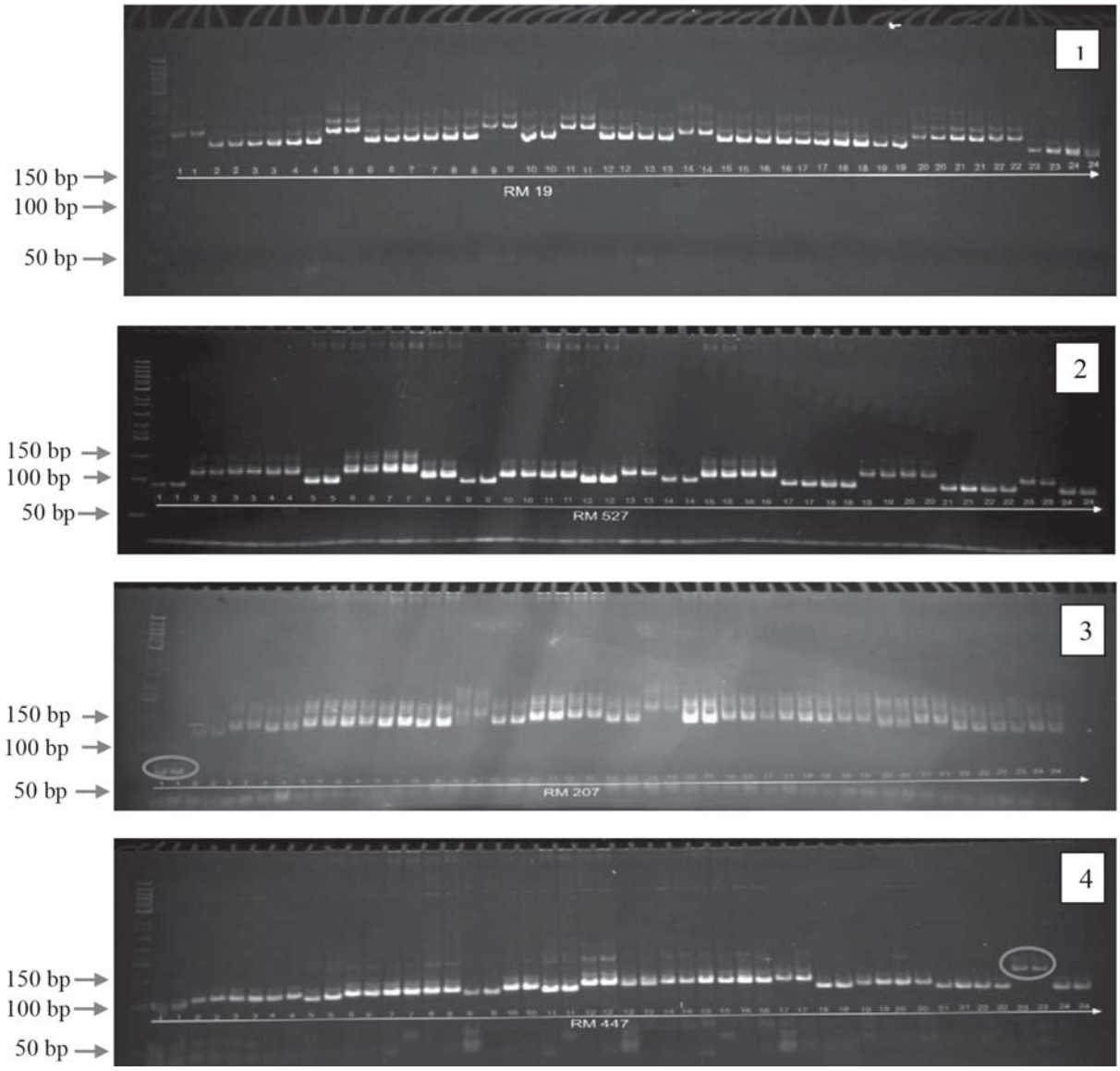

Fig. 2: Plate 1-4 showing PCR amplification profile of 24 aromatic rice genotypes using SSR makers (red circle showing rare allele)

Sequence of aromatic rice genotypes in gel images1. Maguraphulla, 2. Basuabhava, 3. Kapoor Kranti, 4. AS GPC 19, 5. Rajabhog, 6. Loung Choosi B, 7. Govindbhog, 8. Kamini Joha, 9. KDML 105, 10. Bindli, 11. Improved Sarbati, 12. Super Basmati, 13. Loktimachi, 14. Shitalbhog, 15. Domsiah, 16. Basmati 564, 17. Pusa Sugandh 5, 18. Vasumati, 19. R-1498-778-388-1-1, 20. NDR -8497-2, 21. Basmati 11, 22. Karpura Basa, 23. Bas 837, 24. Pusa Basmati 1121

which were observed to possess 60% similarity within the sub-cluster. The sub-cluster IB again sub-divided into two sub-sub-clusters, IB_1 and IB_2 at 63 % genetic similarity level. The sub-sub-cluster IB_1 included seven indigenous aromatic short grain genotypes *viz.*, Basuabhava, Kapoor Kranti, AS GPC 19, Govindbhog, Kaminijoha, Bindli and Loung Choosi B. The sub-sub-cluster IB_2 consisted of five genotypes such as Loktimachi, Domsiah, Basmati 564, R 1498-778-388-1-1 and NDR 8497-2 belonging to aromatic short and long grain type of India and abroad. The second cluster was again sub-divided into two clusters, IIA and IIB consisting of five and three genotypes, respectively. Of these, the sub-cluster IIA included Maguraphulla, Rajabhog, Shitalbhog, KDML 105 and Karpura Basa belonging to aromatic short and long grain type with 48.66% genetic similarity within the sub-cluster. Rajabhog and Shitalbhog within the sub-cluster IIA

exhibited 74 per cent genetic similarities. The sub-cluster IIB consisted of three accessions, Improved Sarbati, Pusa Sugandh 5 and Basmati 11 which were observed to possess 54.41per cent similarity within the sub-cluster, and belongs to traditional and evolved Basmati.

About half of the genotypes below sub-cluster level exhibited more than 70 per cent genetic similarities between genotypes which itself explains the existence of narrow genetic base among aromatic rice genotypes. SSR markers provided adequate power of resolution to discriminate among crop accessions and it could serve as a potential tool in the identification and characterization of genetically distant accessions from various sources. The low to high degree of similarity among studied genotypes indicated high level of genetic diversity, and also the possibility of using these genotypes in breeding programmes targeted at improvement of elite aromatic rice varieties. With the aid of microsatellite makers and clustering data, different distantly related rice genotypes may be combined by crossing, for instances, Bas 837 × Kapoor Kranti, Bas 837 × AS GPC 19, Bas 837 × Govindbhog from different clusters to get hybrid varieties with the highest heterosis. The microsatellite assay generated genotype specific alleles in some of the genotypes screened; these may be used as DNA fingerprints for cultivar identification and differentiation of aromatic rice. This would be of extensively helpful for the establishment and protection of proprietary rights and the diagnosis of cultivar purity.

ACKNOWLEDGEMENTS

We thank Department of Science and Technology, Govt. of India for providing the senior author (V Kumar) an INSPIRE Fellowship during the study and Department of Biotechnology, Govt. of India, for funding the work.

REFERENCES

Anderson, J. A., Churchill, G. A., Autrique, J. E., Tanksley, S. D. and Sorrells, M. E. 1993. Optimizing parental selection for genetic linkage maps. *Genome*, **36**: 181-86.

Bhagwat, A., Suseelan, K. N., Shinde, P. and Gopalakrishna, T. 2008. Molecular marker based diversity studies in Indian landraces of rice (*Oryza sativa* L.). *SABRAO J Breed Genet*, **40**: 9-25.

Chakravorty, A. and Ghosh, P. D. 2012. Genetic divergence in landraces of rice (*O. sativa* L.) of West Bengal. *J. Crop and Weed*,**8** :23-28.

Hossain, M. Z., Rasul, M. G., Ali, M. S., Iftekharuddaula, K. M. and Miana M. A. K. 2007. Molecular characterization and genetic diversity in fine grain and aromatic landraces of rice using microsatellite markers. *Bangladesh J. Genet. Pl. Breed.*, **20** : 1-10.

Hutchenson, K. 1970. A test for comparing diversities based on the Shannon formula. *J. Theor. Biol.*, **29**: 151-54.

Lapitan, V. C., Brar, D. S., Abe, T. and Redona, E. D. 2007. Assessment of genetic diversity of Philippine rice cultivars carrying good quality traits using SSR markers. *Breed Sci.*, **57**: 263-70.

McCouch, S. R., Teytelman, L., Xu, Y., Lobos, K. B., Clare, K., Walton, M., Fu, B., Maghirang, R., Li, Z., Xinfg, Y., Zhang, Q., Kono, I., Yano, M., Fjellstorm, R., DeClerck, G., Schneider, D., Cartinhour, S., Ware, D. and Stein, L. 2002. Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res.*, **9**: 199-207.

Murray, M. G. and Thompson, W. F. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.*, **8**: 4321-25.

Nadia, I., Mohiuddin, A. K. M., Sultana, S. and Ferdous, J. 2014. Diversity analysis of *indica* rice accessions using morphological and SSR markers. *Ann. Biol. Res.*, **5**: 20-31.

Nene, Y. L. 1998. Basmati rice a distinct variety (cultivars) of the Indian sub- continent. *Asian Agric. Hist.*, **2**: 175-88.

Pal, S., Jain, S., Saini, N., Aarti, and Jain, R. K. 2004. Identification of microsattelite markers for differentiating some Basmati and non-Basmati rice varieties. *Indian J. Biotech.*, **3**: 519-26.

Rahman, M. M., Rasaul, M. G., Hossain, M. A., Iftekharuddaula, K. M. and Hasegawa, H. 2012. Molecular characterization and genetic diversity analysis of rice using SSR markers. *J. Crop Improv.*, **26**: 244-57.

Rohlf, F. J. 2002. NTSYS-pc: Numerical taxonomy and multivariate analysis system (Ed. 2.2), Department of Ecology and evolution, State University of New York, Stony Brook.

Sajib, A. M., Hossain, M. M., Mosnaz, A. T. M. J., Islam, M. M., Ali, M. S. and Prodhan, S. M. 2012. SSR marker-based molecular characterization and genetic diversity analysis of aromatic landraces of rice (*Oryza sativa* L.). *J. BioSci. Biotech.*, **1**: 107-16.

Singh, R. K., Singh, U. S., Khush, G. S., Rohila, R., Singh, J. P., Singh, G. and Sekhar, K. S. 2000. Small and medium grained aromatic rice of India. In: aromatic rices (Eds. R.K. Singh, U.S. Singh and G.S. Khush). Oxford and IBH Publishing Co. Pvt. Ltd., New Delhi, pp.154-74.