International Journal of Chemical Studies

P-ISSN: 2349–8528
E-ISSN: 2321–4902
IJCS 2018; 6(2): 3014-3021

Received: 23-01-2018
Accepted: 25-02-2018

Manisha Priyadarshini
Department of Agricultural Biotechnology & Molecular Biology, Dr. Rajendra Prasad Central Agricultural University, Pusa, Bihar, India

Pankaj Kumar
Department of Agricultural Biotechnology & Molecular Biology, Dr. Rajendra Prasad Central Agricultural University, Pusa, Bihar, India

VK Sharma
Department of Agricultural Biotechnology & Molecular Biology, Dr. Rajendra Prasad Central Agricultural University, Pusa, Bihar, India

Correspondence
VK Sharma
Department of Agricultural Biotechnology & Molecular Biology, Dr. Rajendra Prasad Central Agricultural University, Pusa, Bihar, India

Molecular differentiation and classification in relation to fragrance of landraces and improved varieties of aromatic rice using microsatellite markers

Manisha Priyadarshini, Pankaj Kumar and VK Sharma

Abstract
A study was conducted to examine the nature and extent of genetic differentiation and divergence in relation to fragrance among eighteen landraces and improved varieties of aromatic rice using a specific set of eighteen microsatellite markers. Altogether 123 allelic variants including 76 shared and 47 unique alleles were detected with an average of 6.8 alleles per primer, revealing genetic polymorphism due to differences in overall size of amplified products. Appearance of more than one band in some of the entries was observed in combinations with the primers RM 223, RM 252, RM 444, RM 505, RM 252, RM 7049 and RM 7356. Considering the magnitude of polymorphic information content, number of alleles, polymorphism per cent and greater efficiency to unambiguously discriminate the entries, seven primers, namely, RM 256, RM 284, Aro7, RM 223, RM 252, RM 444 and RM 7356, appeared to be highly informative. Occurrence of null allele with primer pair RM 505 was noticed in one of the entries. Microsatellites with CT, GA and AT di-nucleotide motifs detected comparatively greater number of allelic variants. Ample genetic variation amongst the entries was inferred on the basis of similarity coefficients. Basically, the entries were divided into three groups, which were further divided into clusters and sub-clusters. Highly aromatic landraces and improved varieties were accommodated separately into two different groups and third group contained a combination of both types. Principal coordinate analysis completely supported the agglomerative hierarchical classification and neighbor joining tree based grouping of entries. Molecular analysis revealed unique or entry specific alleles which could be useful as molecular fingerprints. Conclusively, molecular profiling exhibited higher level of genetic polymorphism, allowing unique genotyping and unambiguous discrimination of traditional and improved varieties of aromatic rice.

Keywords: aromatic rice, microsatellite, genetic divergence, neighbor joining, principal coordinate

Introduction
Aromatic or scented rice constitutes a small but special sub-group of highly priced rice with characteristic fragrance. Because of the strong human preference for fragrance, consumers all over the world prefer aromatic rice due to its flavor, texture, taste and palatability. The aromatic rice is preferred over non-aromatic rice during special occasions also, commanding a premium price. But, most of the aromatic rice varieties are low yielding and an improvement in their production potential is essentially required taking into consideration the association between the component characters that contribute to yield. Conventional approaches of plant selection for aroma are not easy because of the large effects of the environment and the low narrow sense heritability of aroma. More recently, genomic markers that are genetically linked to fragrance along with the advantage of being simple, inexpensive, reproducible and unaffected across different stages, seasons, locations and agronomic practices have been developed for the selection and characterization of aromatic rice. Genomic sequences based markers are serving as a powerful tool in the assessment of genetic variation and elucidation of the genetic relationships. Among genomic sequences based molecular markers, the microsatellites reveal simple sequence length polymorphism among rice genotypes (Vhora *et al.* 2013; Samal *et al.* 2014; Saheewala *et al.* 2014; Palanga *et al.* 2016; Krupa *et al.* 2017) [24, 21, 19, 16, 12] by using a pair of primers for selective amplification of targeted genomic regions to elucidate the genetic variation owing to differences in the number of repeats at homologous sites in different accessions (Yang *et al.* 1994) [27].

Chemical analysis of more than two hundred volatile compounds present in the rice grain (Pachauri *et al*. 2010) [14] has indicated that 2-acetyl-1-pyrroline is the key aroma compound present in almost all the aromatic rice varieties (Buttery *et al*. 1988; Widjaja *et al.* 1996) [6, 25]. However, existence of noticeable variation in the type and intensity of aroma in the different groups of aromatic rice varieties suggest the involvement of additional chemical compounds in varying proportions. Studies undertaken to investigate the genetics of rice aroma and to locate the genes controlling the expression of aroma (Garland *et al*. 2000; Singh *et al*. 2007; Fitzgerald *et al*. 2008; Amarawathi *et al.* 2008) [10, 23, 9, 2] have identified three quantitative trait loci (q*aro3.1,* q*aro4.1* and q*aro 8.1*) located on three rice chromosomes (3, 4 & 8). Non-functional allele of the gene within the aroma related locus (q*aro8.1*) coding for enzyme betaine aldehyde dehydrogenase contributes to aroma development (Bradbury *et al*. 2005; Bourgis *et al*. 2008) [5, 4], whereas its functional allele makes rice non-aromatic. Similarly, specific alleles of the gene within another aroma related locus (q*aro 4.1*) show association with the aromatic rice varieties (Amarawathi *et al*. 2008) [2].

Rice aroma is widely considered as a complex quantitative trait and there is need to understand the variation in its type, intensity and stability using advanced genomic approaches to assist in the formulation and implementation of appropriate and precise strategies for effective transfer of aroma genes in the genetic background of high yielding rice varieties using molecular breeding methods. Keeping into consideration that accumulated genetic information on aromatic rice and the aroma of rice offer great opportunities for the development and deployment of molecular markers for breeding applications, the present investigation was conducted to characterize and differentiate aromatic rice genotypes on the basis of fragrance related microsatellite markers and to determine the nature and extent of genetic differentiation and divergence among the aromatic landraces and improved varieties.

Materials and Methods

Experimental materials comprised eighteen landraces and improved varieties, namely, Shyam Jeera, Champaran Basmati, Tarrori Basmati, Narendra Lalmati, Malida, Baharni, Jeerabati, Marcha, Birsamati, Rajendra Kasturi, Pusa Sugandh-1, Pusa Sugandh-2, Pusa Sugandh-3, Pusa Sugandh-5, RAU 3036, Pusa 1121, Pusa1509 and RAU 3055, of aromatic rice. Seeds of each of these entries were planted in aluminium containers filled with soil. The immature leaves from two to three weeks old seedlings were collected from each genotype and used for DNA extraction in the Molecular Biology Laboratory of the Department of Agricultural Biotechnology and Molecular Biology, Dr. Rajendra Prasad Central Agricultural University, Pusa, Bihar. The extraction buffer was prepared by using 100 mM Tris-HCl (pH 8.0), 20 mM Na-EDTA salt (pH 8.0), 2 M NaCl, 3% (w/v) CTAB, 2% (w/v) PVP and 1% β-Mercaptoethanol. Total genomic DNA was isolated by using CTAB method (Doyle and Doyle, 1990) with slight modifications. The isolated DNA samples were subjected to agarose gel electrophoresis to assess their quality. Appearance of a single, sharp band of high molecular weight without smearing indicated the extraction of a good quality DNA sample with less damage during extraction process. All such samples were selected for their utilization during amplification.

Using standard protocol of polymerase chain reaction adjusted to laboratory conditions, a panel of eighteen aroma related microsatellite primers (Table 1) based amplification of targeted genomic regions was carried out in 15 µl reaction mixture with the help of a thermal cycler (Biometra), The primer vials were centrifuged before and after the addition of 1X TE buffer to the vials and sub-stocks were prepared from the stock. The reaction mixture was prepared by a combination of 2.6 µl water (Protease and Nuclease Free), 3.0 µl 5X buffer, 1.5 µl $MgCl_2$ (10 mM), 3.0 µl dNTPs (200 µM), 1.2 µl Primer F (5 µM), 1.2 µl Primer R (5 µM), 0.5 µl Taq DNA polymerase (1 unit) and 2.0 µl DNA sample.

Annealing temperature for different primer pairs was kept approximately 5^0C less than their melting temperature (Tm) and amplification was carried out with an initial denaturation for 2 min at 94 ^{0}C followed by 35 cycles of denaturation for 1 min at 94 ^{0}C, primer annealing for 1 min at 48^0-60 ^{0}C and primer extension for 2 min at 72 ^{0}C and subsequently final extension for 10 min at 72 ^{0}C. The products generated by amplification reaction were subjected to agarose (2%) gel electrophoresis at 100 V for one and half hours and then visualized and documented under gel documentation system. Size of amplified product was determined in relation to the size of markers in 50 bp DNA ladder (Fermentas) with the help of alpha view gel reader. The position of the amplified products corresponded to the location of the bands along y-axis (ranging from 0 to 1030). The Rf value for each band was determined assuming the location of well as initial position (Rf=0) and the position of migrated dye as final position (Rf=1) in a frame of reference.

Efficiency of primer pairs in generation of unique and genotype specific allele(s) was assessed by computing the polymorphism percentage as under:

$$PP = \frac{\text{Numeber of unique alleles generated}}{\text{Total number of alleles detected}} \text{ x } 100$$

The data on polymorphism information content was obtained by calculating the value according to the formula (Anderson *et al*., 1993) [1] as under:

$$PICi = 1 - \sum_{j=1}^{k} \mathrm{P}^2 ij$$

Where, k is the total number of alleles detected for a marker and Pij is the frequency of the j^{th} allele for i^{th} marker and summation extends over k alleles

Efficacy of individual primers in differentiation of the genotypes (Yadav *et al.* 2013) [26] was tested by computing the discrimination coefficient (DC) and non-discrimination coefficient (NDC) as follows:

$$DC = \frac{\text{Number of piars of genotypes discriminated}}{\text{Total number of pairs of genotypes}}$$

$$NDC = \frac{\text{Number of piars of genotypes not discriminated}}{\text{Total number of pairs of genotypes}}$$

Binary data matrix was generated by scoring for the presence and absence of bands in all the entries. The data matrix was subjected to cluster analysis. Similarity coefficients (Dice, 1945) [7] were computed for pair-wise comparisons based on the proportions of shared bands produced by the primers as follows:

$$Similarity\ coefficient = \frac{2a}{(2a + b + c)}$$

Where, a= Number of bands shared between J^{th} and K^{th} genotypes; b= Number of bands present in J^{th} genotype but absent in K^{th} genotype; and c= Number of bands absent in J^{th} genotype but present in K^{th} genotype.

Hierarchical cluster analysis was performed using the data on similarity coefficients. Tree building in the cluster analysis involved sequential agglomerative hierarchical non-overlapping (SAHN) clustering. The dendrogram based on similarity indices was obtained by un-weighted pair-group method using arithmetic mean (UPGMA). Analysis was performed with the help of NTSYS-pc software (Rohlf, 1997) [18]. The nature of differentiation and divergence at molecular level was examined by identifying the clusters. Principal coordinate analysis of the microsatellite primers dependent genetic profiles was conducted and compared with the results obtained from cluster analysis and neighbor joining tree.

Results and Discussion

Computational analysis of the molecular size of different bands reflected different levels of molecular genetic polymorphism amongst the eighteen aromatic rice entries (Fig. 1). Several allelic variants were generated by some of the primer pairs as a result of variation in the length of microsatellites among the entries (Palanga *et al.* 2016) [16], while others generated only a few alleles. Further, the number and proportion of unique alleles varied to a considerably greater extent with the primer pairs. The differences in the size of amplified products, as recognized in the present study, reflect the variability in the length of the simple sequence repeats at a microsatellite site, which arises as a result of differences in the number of repeats existing in different entries and revealed by different primer pairs at the microsatellite locus in question. Therefore, the results of the molecular analysis clearly indicated the existence of ample genetic variability in terms of differences in the number of repeats at the primer specific microsatellite locus in the different aromatic rice entries under evaluation.

A survey of the molecular profiles generated by the evaluation of amplified products clearly indicated that altogether 123 alleles were detected with an average of 6.8 alleles per primer (Table 1). The number of alleles per primer pair ranged from three in case of E 03-92.0 to ten in the cases of RM 252, RM 256, RM 284 and RM 7356. Seven primer pairs, namely, RM 223, RM 252, RM 444, RM 505, RM 7049, RM 7356 and BAD 2a, generated more than one amplified product in some of the entries most probably due to residual heterozygosity. While classifying the allelic variants into shared alleles and unique alleles, it was observed that a total of 76 shared and 47 unique allelic variants were generated in the form of amplified products. The amplification by the eight primer pairs, namely, RM 7049, RM 8264, RM 23097, ARSSR 3, BAD 2a, BAD 2b, E 03-92.0 and SCUSSR 1 yielded only one unique allele out of six, seven, four, six, four, five, three and five allelic variants generated, respectively. The primer pairs RM 7356, RM 252, RM 284, RM 223, RM 444, Aro7, E11.44.5, RM256 and RM 505 generated considerably greater percentage of unique alleles and consequently recorded considerably higher polymorphism per cent.

Occurrence of null allele was inferred due to failure of amplification of a particular repeat locus specific to the unique flanking sequences of the microsatellite. Accordingly, the genotype was assigned a null allele for a microsatellite locus whenever an amplification product (s) was not detected in combination with a particular primer used during amplification. The null allele was shown only by the microsatellite locus associated with RM 505 in combination with one of the entries. Absence of amplified product due to failure of amplification can be attributed to the failure of primer annealing due to variation in primer binding site specific unique flanking sequences of the concerned microsatellite. Using different panels of primer pairs and different genotypes of rice, similar results have been reported earlier (Vhora *et al*. 2013) [24].

Polymorphic information content of the primers revealing their allelic diversity and frequency among the entries varied from 0.512 in the case of BAD 2a to 0.882 in the case of RM 256 with an average of 0.741 across the primers. Considerably greater value (Table 2) was obtained in the cases of the primer pairs RM 256, RM 284, Aro7, RM 8264, RM 223, RM 252, RM 444, RM 7356, ARSSR-3 and RM 515, indicating higher level of polymorphism among the entries for the marker locus in question. Among these ten primers, RM 256, RM 284, Aro7, RM 223, RM 252, RM 444 and RM 7356 generated considerably greater number of allelic variants with higher polymorphism per cent due to variation in the length of microsatellite repeats based amplified products as a consequence of amplification of the microsatellite repeats flanked by them and appeared to be highly informative primers for the purpose of molecular characterization of aromatic rice entries. Relative efficiency of these seven primer pairs in respect of their ability to unambiguously discriminate the aromatic rice entries under evaluation, as inferred on the basis of magnitude of discrimination and non-discrimination coefficients, was greater. The values obtained in respect of the number of pairs of entries not distinguished by these primer pairs were relatively smaller in magnitude. Average polymorphic information content value obtained in this study confirms that the primers utilized for molecular characterization were, in general, highly informative. Interestingly, RM 223 was also found to be the best primer to identify genetic diversity among aromatic rice varieties in earlier studies (Kibria *et al.* 2009) [11].

Simple sequence repeat loci with tetra-nucleotide repeat motifs detected greater number of alleles than the repeat loci with di-nucleotide and tri-nucleotide repeat motifs and complex repeat motifs (Table 2). Generally the primers amplifying di-nucleotide repeat motifs are found to be more polymorphic than those with tri-nucleotide, tetra-nucleotide and complex repeat motifs (Lapitan *et al.* 2007) [13]. Since, a change in the number of repeats results in the generation of allelic variants because of variation in the size of alleles exhibiting sequence length polymorphism, the total repeat count of di-nucleotide repeat loci is mostly found to be associated with large number of alleles (Lapitan *et al*. 2007; Saheewala *et al*. 2014) [19, 13]. However, the results obtained in the present study are not in complete accordance with the results of earlier reports that microsatellite loci with di-nucleotide repeat motifs, in general, tend to detect greater number of alleles than the repeat loci with tri-nucleotide and tetra-nucleotide repeat motifs. The discrepancy may be probably due to relatively less number of primers used in the present study, in addition to different panel of primers and different set of entries.

Generally, the total repeat count of the di-nucleotide simple sequence repeat loci are seemed to be associated with the

number of alleles detected per locus (Sajib *et al*. 2012; Saheewala *et al*. 2014) [20, 19]. Therefore, the repeat number involved in the simple sequence repeats directly relates to the number of identified alleles and the relationship, in general, reflects that larger the repeat number involved in the simple sequence repeat locus, the larger will be the number of detected alleles. However, the total repeat count of the di-nucleotide simple sequence repeat loci, in general, was not found to be associated with the number of alleles in the present study. Therefore, the results did not indicate a direct relationship between the repeat number involved in the simple sequence repeat locus with di-nucleotide repeat motif and the number of identified alleles. Further, the relationship, in general, did not reflect that larger the repeat number involved in the simple sequence repeat loci, larger was the number of identified alleles. These findings are not in complete agreement with the earlier reports (Rabbani *et al*. 2010; Sajib *et al*. 2012; Saheewala *et al*. 2014; Palanga *et al*. 2016) [17, 20, 19, 16] and the non-correspondence may be due to the use of different panels of primers and a different set of entries in the present study.

Taking into account the nucleotide composition of the di-nucleotide repeat motifs, a comparison was made to reveal the relation between repeat motifs having a particular type of nucleotide composition and the number of alleles generated by primer pairs (Table 3). It was interesting to note that the simple sequence repeat loci with CT, GA and AT di-nucleotide repeat motifs detected greater number of alleles. Contrarily, the loci with TA and TC di-nucleotide repeat motifs appeared to detect relatively lesser number of alleles among the primers having di-nucleotide repeat motifs. The perfect di-nucleotide repeat motif GA has been reported to display high level of variation among the rice genotypes (Sajib *et al*. 2012) [20].

Microsatellites based genetic similarity between Pusa 1121 and Pusa 1509(0.612) was found to be the maximum amongst pair-wise combinations of the entries. This was followed by remarkably higher magnitude of similarity coefficient between Jeerabati and Marcha (0.540), Pusa Sugandh-1 and Pusa Sugandh-2 (0.513), Birsamati and Pusa Sugandh-2(0.462), Pusa Sugandh-5 and RAU-3036 (0.458), RAU-3036 and Pusa 1509(0.458), Jeerabati and Birsamati (0.445), Rajendra Kasturi and Pusa Sugandh-1(0.445), Marcha and Birsamati (0.433). The magnitude of similarity coefficient between Shyam Jeera and Pusa Sugandh-1, Shyam Jeera and Pusa 1121 and Jeerabati and Pusa 1509 was equal to zero. Apparently therefore, the microsatellite markers based molecular profiles did not exhibit any similarity between the entries involved in these three pair-wise combinations. Overall, the results revealed enormous diversity at the molecular level (Rabbani *et al*. 2010; Behera *et al.* 2012; Pachauri *et al.* 2013; Saheewala *et al.* 2014; Samal *et al.* 2014; Shamim *et al.* 2015; Palanga *et al.* 2016; Singh *et al.* 2016; Krupa *et al.* 2017) [10, 3, 15, 19, 21, 22, 16, 28, 12] amongst the entries evaluated in the present analysis.

Principal coordinate analysis (Fig. 2) based two dimensional plots of eighteen microsatellite primers dependent genetic profiles revealed the spatial distribution pattern of entries along the two principal axes. Although some of the entries were placed far away from the centroid of the clusters and remaining entries were placed more or less around the centroid, it is evident that entries were distinctly divided into three major genotypic groups (Table 4). The first multi-genotypic group consisted of six genetically improved aromatic entries, namely, Pusa Sugandh-3, Pusa Sugandh-5, RAU-3036, Pusa-1121, Pusa-1509 and RAU-3055. While, the second multi-genotypic group constituted of seven entries including traditional and genetically improved aromatic varieties, namely, Baharni, Jeerabati, Marcha, Birsamati, Rajendra Kasturi, Pusa Sugandh-1 and Pusa Sugandh-2. Lastly, the third multi-genotypic group, which consisted of five aromatic entries, namely, Shyam Jeera, Narendra Lalmati, Champaran Basmati, Tarrori Basmati and Malida, was dominated by traditional cultivars. Inferences derived from the results of similarity coefficients based hierarchical classification were completely corroborated by the principal coordinate analysis based two dimensional plots of eighteen microsatellite primer pairs dependent genetic profiles from eighteen aromatic rice genotypes (Fig. 2). A perusal of the neighbor joining tree (Fig. 3) and the binary data based dendrogram (Fig. 4) clearly reflects the inter-relationships amongst the entries that were basically divided into three groups. Therefore, the inferences derived from the similarity coefficients based hierarchical classification are completely supported by the principal coordinate analysis.

Molecular profiling made it possible to discern sufficient variation at the molecular level amongst the aromatic rice entries, which exhibit close similarity in respect of several agro-morphological and physico-chemical attributes. The elucidation of variability at the molecular level among these aromatic rice entries yielding ample molecular polymorphism and genetic differentiation allowed their unique genotyping. Remarkably greater discrimination potential of microsatellite markers characterized by their polymorphic nature and allelic diversity, as noticed in the present study and also reported earlier by several research workers (Rabbani *et al*. 2010; Behera *et al.* 2012; Pachauri *et al.* 2013; Saheewala *et al.* 2014; Singh *et al.* 2016; Palanga *et al*. 2016; Krupa *et al.* 2017), [10, 3, 15, 19, 28, 16, 12] indicates their usefulness and powerfulness in discrimination and unambiguous classification of aromatic rice genotypes.

Thus, a high level of genetic polymorphism at the molecular level was discerned by employing a purposefully chosen set of microsatellite markers in the present study that provided a molecular database of the eighteen aromatic rice entries evaluated. The microsatellite markers based identification and differentiation of aromatic rice could be helpful to preserve the integrity of these high quality aromatic rice genotypes to benefit farmers, breeders and consumers. Additionally, the entries from different groups and these markers can be effectively and efficiently utilized in marker assisted breeding for discrimination and unambiguous identification of promising recombinants during genetic enhancement of aromatic rice.

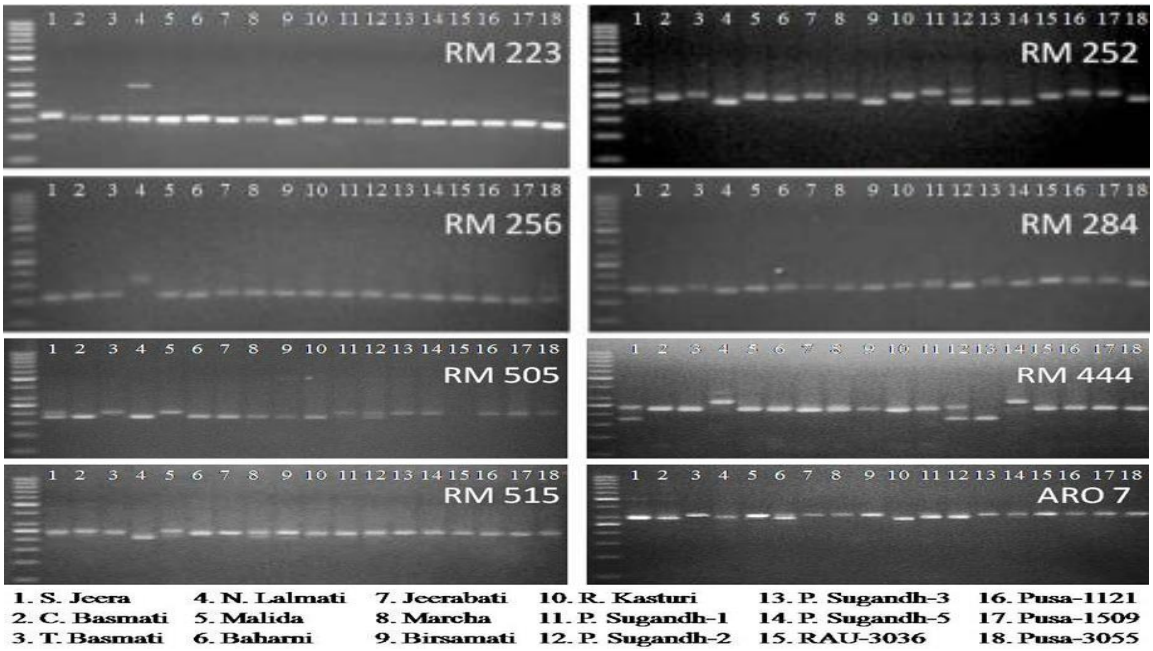

Fig 1: Microsatellite primers based amplification patterns of targeted genomic regions in eighteen entries of aromatic rice

Table 1: Analysis of primer pairs used for the amplification of targeted genomic regions in the eighteen aromatic entries of rice

Primer	Allele size difference (bp)	Allele size range (bp)	No. of alleles	No. of unique alleles	No. of shared alleles	PP	PIC	DC	NDC
RM 223	158	136 - 294	9	4	5	44.44	0.836	0.895	0.105
RM 252	054	216 - 270	10	5	5	50.00	0.808	0.882	0.118
RM 256	071	102 - 173	10	4	6	40.00	0.882	0.934	0.065
RM 284	044	139 - 183	10	5	5	50.00	0.870	0.934	0.065
RM 444	087	178 - 265	9	4	5	44.44	0.790	0.849	0.150
RM 505	034	181 - 215	5	2	3	40.00	0.657	0.745	0.255
RM 515	037	213 - 250	6	2	4	33.33	0.759	0.804	0.196
RM 7049	031	130 - 161	6	1	5	16.66	0.728	0.830	0.170
RM 7356	083	153 - 236	10	7	3	70.00	0.787	0.843	0.157
RM 8264	021	200 - 221	7	1	6	14.28	0.845	0.895	0.105
RM 23097	041	371 - 412	4	1	3	25.00	0.641	0.680	0.320
Aro7	045	288 - 333	9	4	5	44.44	0.858	0.908	0.092
ARSSR 3	016	171 - 187	6	1	5	16.66	0.783	0.830	0.170
BAD2a	014	245 - 259	4	1	3	25.00	0.512	0.660	0.340
BAD2b	022	344 - 366	5	1	4	20.00	0.728	0.765	0.236
E03-92.0	008	176 - 184	3	1	2	33.33	0.549	0.582	0.418
E11-44.5	016	188 - 204	5	2	3	40.00	0.580	0.614	0.386
SCUSSR1	011	133 - 144	5	1	4	20.00	0.728	0.771	0.229

PP: Polymorphism per cent; PIC: Polymorphism information content; DC: Discrimination coefficient; NDC: Non-discrimination coefficient

Table 2: Allelic diversity of primers in relation to repeat motifs involved in amplification

Repeat motif of primers	No. of primers	No. of alleles	No. of alleles/primer
Dinucleotide	10	74	7.4
Trinucleotide	2	12	6.0
Tetra nucleotide	2	16	8.0
Complex	1	7	7.0

Table 3: Allelic diversity of primers with di-nucleotide motifs involved in amplification

Di-nucleotide repeat motifs of primer	Number of primer	Number of allele	Number of allele/primer
GA	2	16	8.0
TA	1	06	6.0
AT	2	13	6.5
CT	4	34	8.5
TC	1	05	5.0

Table 4: Composition of different groups based on microsatellite specific polymorphism

Group	Entries included in each group
I	Pusa-1121, Pusa-1509, RAU-3036, Pusa Sugandh-3, Pusa Sugandh-5, RAU-3055
II	Jeerabati, Marcha, Birsamati, Pusa Sugandh-1, Pusa Sugandh-2, Rajendra Kasturi, Baharni
III	Tarrori Basmati, Malida, Champaran Basmati, Shyam Jeera, Narendra Lalmati

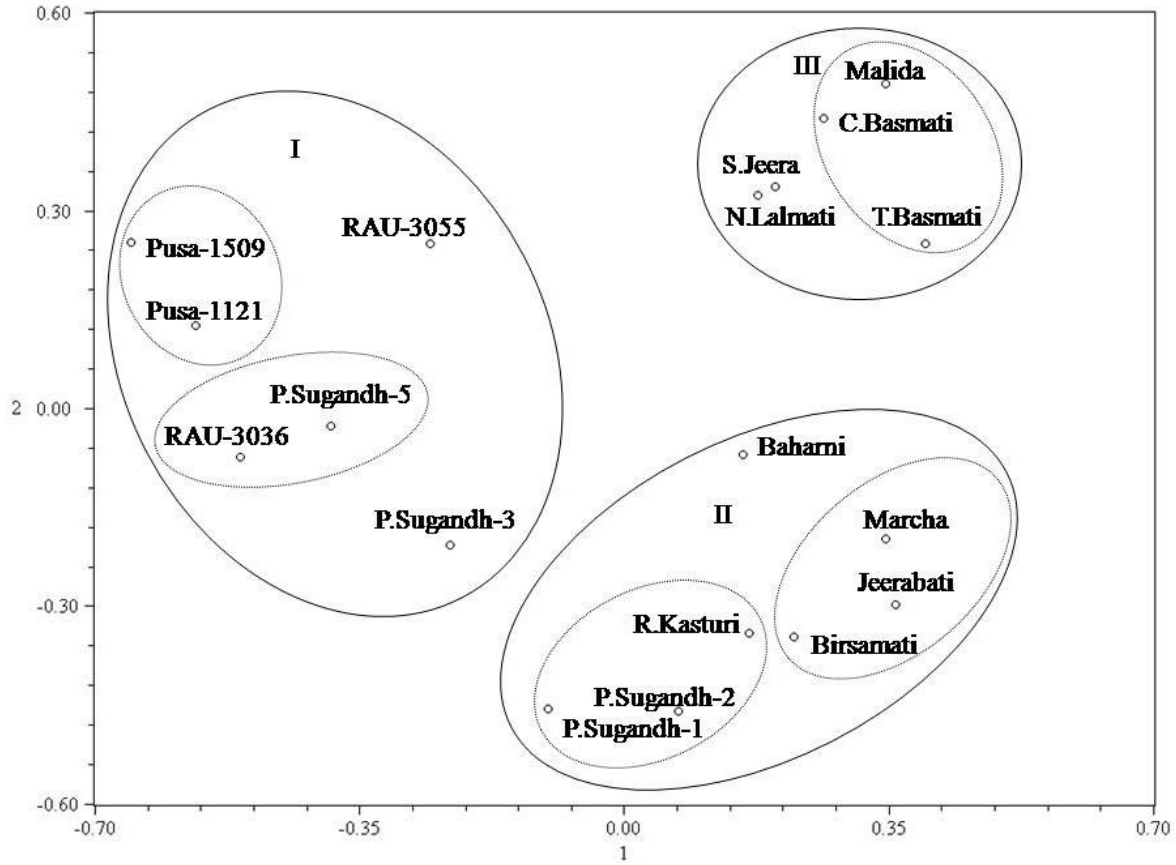

Fig 2: Principal coordinate analysis based two-dimensional ordinations of the genetic profiles of the eighteen entries of aromatic rice

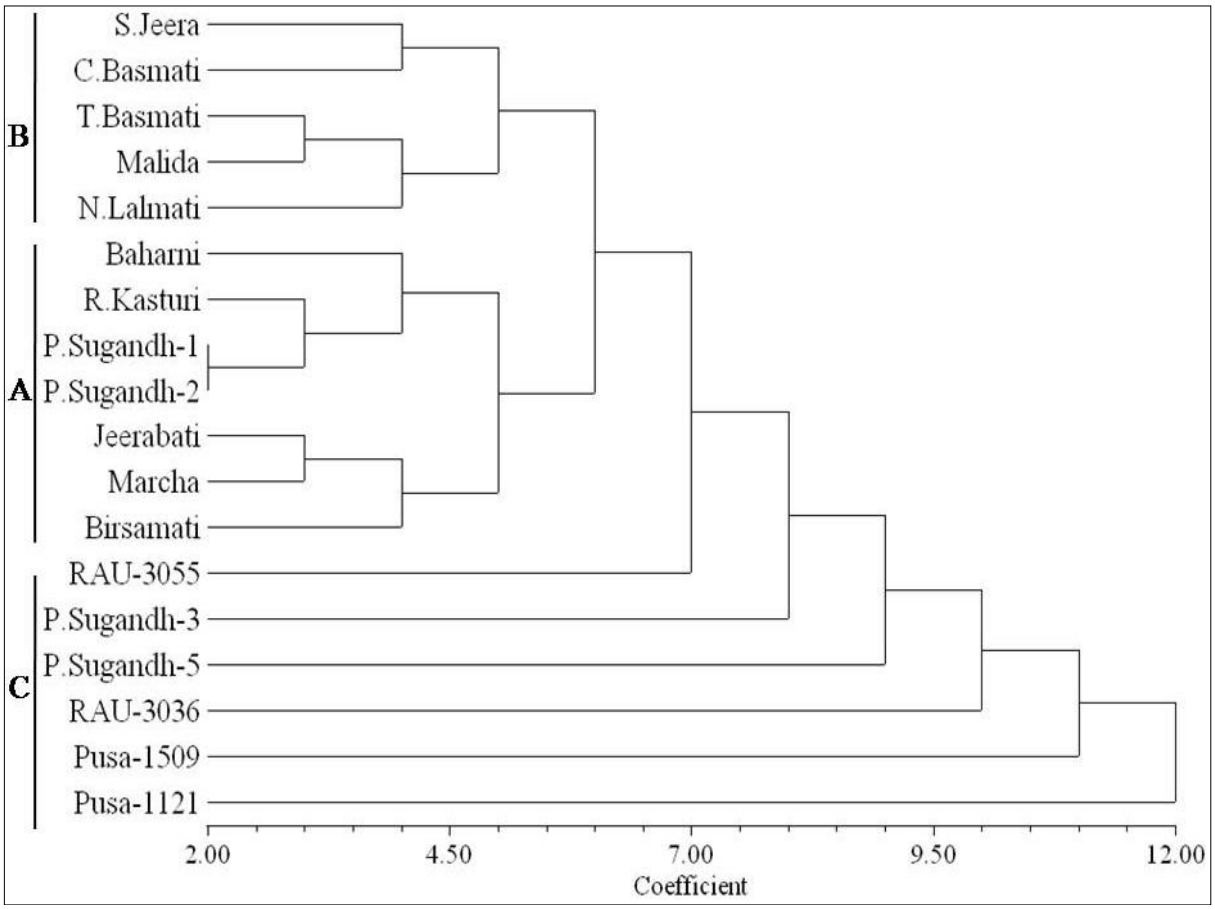

Fig 3: Similarity indices based neighbor joining tree showing inter-relationships of the eighteen entries of aromatic rice

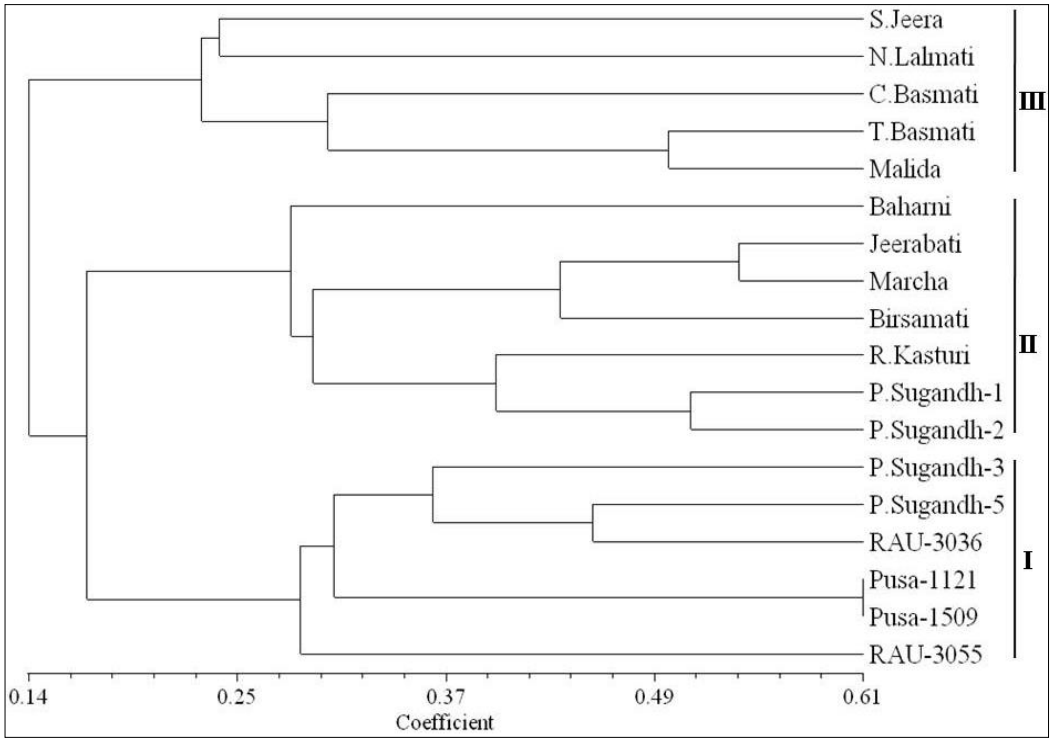

Fig 4: Dendrogram of the eighteen entries of aromatic rice using un-weighted paired group method based on mathematical averages

References

1. Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME. Optimization parental selection for genetic linkage maps. Genome. 1993: 36:181-186.
2. Amarawathi Y, Singh R, Singh AK, Singh VP, Mohapatra T, Sharma TR *et al*. Mapping of quantitative trait loci for basmati quality traits in rice (*Oryza sativa* L). Mol. Breed. 2008; 21:49-65.
3. Behera L, Sahu RK, Datta S, Nanda A, Sahu SC, Pradhan SK *et al*. DNA fingerprinting and assessment of genetic diversity of high yielding rice varieties using microsatellite markers. Int. J of Agric. Env. Biotech. 2012; 5:89-99.
4. Bourgis F, Guyot R, Gherbi H, Tailliez E, Amabile I, Salse J, *et al*. Ghesquière A. Characterization of the major fragrance gene from an aromatic japonica rice and analysis of its diversity in Asian cultivated rice. Theor. Appl. Genet. 2008; 117:353-368.
5. Bradbury LM, Fitzgerald TL, Henry RJ, Jin Q, Waters DL. The gene for fragrance in rice. Plant Biotech. J, 2005; 3:363-370.
6. Buttery RG, Turnbaugh JC, Ling LC. Contribution of volatiles to rice aroma. J Agric Food Chem. 1988; 36:1006-1009.
7. Dice LR. Measures of the amount of ecologic association between species. Ecology. 1945; 26:297-302.
8. Doyle JJ, Doyle JL. A rapid total DNA preparation procedure for fresh plant tissue. Focus. 1990; 12:13-15.
9. Fitzgerald MA, Sackville Hamilton NR, Calingacion MN, Verhoeven HA, Butardo VM. Is there a second fragrance gene in rice? Plant Biotech J, 2008; 6:416-423.
10. Garland S, Lewin L, Blakeney A, Reinke R, Henry R. PCR based molecular markers for the fragrance gene in rice (*Oryza sativa* L). Theor. Appl. Genet. 2000; 101:364-371.
11. Kibria K, Nur F, Begum SN, Islam MM, Paul SK, Rahman KS *et al*. Molecular marker based genetic diversity analysis in aromatic rice genotypes using SSR and RAPD markers. Int. J Sustain. Crop Prod. 2009; 4:23-34.
12. Krupa KN, Shashidhar HE, Ningaraj D, Reddy M, Kumara V, Swamy HV. Molecular marker based genetic diversity analysis in rice genotypes (*Oryza sativa* L.) using SSR markers. Int. J Pure App. Biosci. 2017; 5:668-674.
13. Lapitan VC, Brar DS, Abe T, Redona ED. Assessment of genetic diversity of Philippine rice cultivars carrying good quality traits using SSR markers. Breeding Science. 2007; 57:263-270.
14. Pachauri V, Singh MK, Singh AK, Singh S, Shakeel NA, Singh VP. Singh NK. Origin and genetic diversity of aromatic rice varieties, molecular breeding and chemical and genetic basis of rice arroma. J. Plant Biochemistry & Biotechnology. 2010; 19:127-143.
15. Pachauri V, Taneja N, Vikram P, Singh NK, Singh S. Molecular and morphological characterization of Indian farmers rice varieties (*Oryza sativa* L.). Australian J Crop Sci. 2013; **7**:923-932.
16. Palanga KK, Traore K, Bimpong K, Jamshed K, Mkulama MAP. Genetic diversity studies on selected rice varieties grown in Africa based on aroma, cooking and eating quality. Afr. J Biotechnol. 2016; 15:1136-1146.
17. Rabbani MA, Masood MS, Shinwari ZK, Shinozaki KY. Genetic analysis of Basmati and Non-Basmati Pakistani rice (*Oryza sativa* L.) cultivars using microsatellite markers. Pak. J Bot. 2010; 42:2551-2564.
18. Rohlf PJ, NTSYS pc: Numerical Taxonomy and Multivariate Analysis System. ver. 2.1. Exeter Software, Setauket, New York, USA, 1997.
19. Saheewala H, Sharma VK, Shahi VK. Genetic diversity assessment in aromatic rice using microsatellite markers. Proceed. National Symposium on Crop Improvement for Inclusive Sustainable Development. November 7-9, PAU, Ludhiana, 2014, 238-241.
20. Sajib AM, Hossain MM, Mosnaz ATMZ, Hossain H, Islam MM, Ali MS, *et al*. SSR-marker based molecular characterization and genetic diversity analysis of

aromatic landraces of rice (*Oryza sativa* L.). J.Biosci. Biotech. 2012; 2:107-116.
21. Samal KC, Meti N, Bastia DN, Rout GR. Genetic diversity analysis of traditional aromatic rice using molecular markers. BioMed Research. 2014; 1:1-14.
22. Shamim MZ, Manzar H, Sharma VK, Kumar P. Microsatellite marker based characterization and divergence analysis among rice varieties. Indian J Biotechnol. 2015; 15:182-189.
23. Singh R, Singh AK, Sharma TR, Singh A, Singh NK. Fine mapping of aroma QTLs in basmati rice (Oryza sativa L) on chromosome 3, 4 and 8. J Plant Biochemistry & Biotechnology. 2007; 16:75-82.
24. Vhora Z, Trivedi R, Chakraborty S, Ravikiran R, Sasidharan N. Molecular studies on aromatic and non-aromatic rice (*Oryza sativa* L.) genotypes for quality traits using microsatellite markers. Bioscan. 2013; 8:359-362.
25. Widjaja R, Craske JD, Wootton M. Comparative studies on volatile components of non-fragrant and fragnant rices. J Sci. Food Agric. 1996; 70:151-161.
26. Yadav S, Singh A, Singh MR, Goel N, Vinod KK, Mohapatra T, *et al*. Assessment of genetic diversity in Indian rice germplasm (*Oryza sativa* L.): use of random versus trait-linked microsatellite markers. J Genet. 2013; 92:545-557.
27. Yang GP, Maroof MAS, Xu CG, Zhang Q. Biyashcv RM. Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. Mol. Gen. 1994; 245:187-194.
28. Singh N. Choudhury DD. Tiwari G. Singh AK. Kumar S. Srinivasan K. Tyagi RK. Sharma AD. Singh NK. Singh R. Genetic diversity trend in Indian rice varieties: an analysis using SSR markers. *BMC Genetics*. 2016; 17: 127-140.