Crop Breeding and Applied Biotechnology 5:157-165, 2005
Brazilian Society of Plant Breeding. Printed in Brazil

Evaluation of the number and information content of fluorescent-labeled SSR markers for rice germplasm characterization

Tereza Cristina de Oliveira Borba[1]*, Rosana Vianello Brondani[1], Paulo Hideo Nakano Rangel[1], and Claudio Brondani[1]

Received 3 March 2005

Accepted 3 June 2005

ABSTRACT - *The use of fluorescent-labeled simple sequence repeat (SSR) markers allows the simultaneous analysis of diverse loci in multiplex series, providing an enhanced precision of the determination of the allele size when a large number of genotypes is used. Twenty-five fluorescent-labeled SSR markers were used to genotype 242 accessions of the Brazilian Rice Core Collection. Based on the Polymorphism Information Content (PIC) values of each marker, groups with 5, 10, 15 and 20 markers were formed. The average PIC values varied from 0.67 to 0.89 and the number of alleles sampled per group, from 70 to 377. In the case of rice, which has a narrow genetic base, the conclusion was drawn that it is of fundamental importance that the genotypes must be evaluated by the highest possible number of regularly spaced SSR markers in the genome, and that these, preferentially, should offer a high information content.*

Key words: *Oryza sativa*, genetic diversity, genetic resources.

INTRODUCTION

The use of genetic resources for rice has been suggested as the most relevant alternative to raise the productivity levels achieved by modern cultivars (Brondani et al. 2003). However, before using these resources, it is fundamental to compile in depth knowledge on the existing genetic variability in these accessions to identify genotypes that could ultimately contribute effectively to genetic gains for traits of interest in rice breeding programs. It is estimated that over 200.000 accessions are stored in rice germplasm banks all over the world (Jackson 1997), which makes the choice of top potential genotypes difficult. An alternative is the formation of core collections which aim to represent over 80% of the genetic variability of the species in a 10% sample of the total collection (Frankel and Brown 1984). Aiming at the establishment of a representative sample of the genetic variability in the Germplasm Bank of Embrapa Arroz e Feijão, Abadie et al. (2002) compiled the Brazilian Rice Core Collection (BRCC), comprising 550 accessions (5.6% of the accessions of the complete collection), divided in three groups: a) traditional varieties (308 accessions); b) germplasm improved in Brazil (94 accessions); and c)

[1]Embrapa Arroz e Feijão, Rodovia Goiânia a Nova Veneza, km 12, Zona Rural, 75375-000, Santo Antônio de Goiás, GO, Brasil. *E-mail: oliveiraborba@yahoo.com.br

germplasm introduced from other countries (148 accessions). The BRCC is being characterized by SSR (Simple Sequence Repeat) markers, the most informative in comparison with other marker types (Powell et al. 1995).

The SSR markers or microsatellites are considered an important molecular tools since they are codominant, distributed all over the genome, multiallelic, and PCR-(*Polymerase Chain Reaction*)-based (Mullis and Faloona 1987). Besides, owing to their characteristically high degree of polymorphism, microsatellites are ideal markers to study genetic diversity and germplasm characterization. The use of SSR permits the estimation of different parameters of the genetic variability among and within the conserved accessions such as: percentage of polymorphic loci; total number and mean of alleles per locus; mean of alleles heterogeneity and PIC (Polymorphism Information Content) values; identification of private alleles, that are, alleles found only in one of the analyzed genotypes; and the determination of genetic distance values (Hayward and Hamilton 1997).

SSR markers are available in large number for rice (McCouch et al. 2002), which allows the choice of the markers with the greatest information potential and simultaneously, with a wide dispersion across the genome. The use of fluorescent-labeled SSR markers has allowed the simultaneous genotyping of diverse loci in multiplex series, which makes the process easier and quicker. Fluorescent-labeled markers increase the precision in the determination of the allele size and the comparison between the alleles of different individuals, above all in the characterization of large germplasm collections. This methodology has found wide-spread use in the characterization of rice genotypes (Coburn et al. 2002, Blair et al. 2002). The use of fluorescent-labeled SSR markers however entails costs with the labeling of the primer pairs, the use of an automatic DNA sequencing machine for the detection of the PCR product, as well as the need for the acquisition of expensive specific reagents for the electrophoresis of the fragments. SSR type markers varying from 10 (Yang et al. 1994) to 159 genotyped loci (Coburn et al. 2002) have been used in molecular characterization studies of rice genotypes.

The objectives of this study were to: a) optimize the conditions of automatized amplification and detection for the set of 25 fluorescence-tagged SSR markers and identify the markers that can be evaluated simultaneously in the multiplex type reaction, b) evaluate the viability of analyzing each accession represented by a bulk of plants and c) establish the number of SSR markers necessary to precisely estimate the content of genetic information of rice.

MATERIAL AND METHODS

Plant material

The analysis involved 242 accessions of the Brazilian Rice Core Collection (BRCC) (Abadie et al. 2002) of the two groups: germplasm improved in Brazil (94 genotypes-57 of upland rice and 37 of irrigated rice) and germplasm introduced from other countries (148 genotypes-75 of upland rice and 73 of irrigated rice), both groups formed by inbred and cultivars. The accessions were evaluated in four-plant-bulks. The total genomic DNA was extracted according to Doyle and Doyle (1987). When heterozygosis was detected in a bulk for two or more SSR loci, the bulk was opened, that is, each one of the four plants of the bulk was genotyped individually for each locus marker that detected this heterozygosis.

Optimization of the multilocus amplification conditions

In order to amplify more than one SSR locus simultaneously in the same PCR reaction, in an attempt to reduce the costs for reagents and the time spent to prepare the reaction, different DNA and primer concentrations were tested. The reaction was realized with a final volume of 15 μL: 1X reaction buffer (50mM of KCl, 10mM Tris-HCl pH 8.4, 0.1% Triton X-100 and 1.5mM of $MgCl_2$), 0.22 mM of dNTP and one unit of the enzyme Taq Polymerase. Three different primer concentrations SSR with forward and reverse sequences, 0.03, 0.07 and 0.13 μM, as well as DNA at concentrations of 6, 15 and 24ng tested. The amplification conditions were: a pre-cycle of 5 min at 94 °C, followed by 29 cycles for 1 min at 94 °C, 1 min at 56 °C or 52 °C (depending on the marker locus), 1 min at 72 °C, and a final step of 7 min at 72 °C. The automatized genotyping was conducted on an ABI 3100 (Applied Biosystems, Foster City) automatic DNA sequencer, followed by the analysis with software GeneMapper 3.5 (Applied Biosystems, Foster City). The obtained data were transferred to a spread sheet in Excel format.

SSR markers

Twenty-five SSR markers were used for the genetic analysis (Table 1). The forward sequence of 13 primers were labeled with fluorochrome 6-carboxyfluorescein (6-Fam), and 12 with fluorochrome hexachloro-6-carboxyfluorescein (Hex) (RW genes, Rio de Janeiro), the ones which amplify PCR products without overlapping of fragment size, were marked with the same fluorescence and later evaluated for their simultaneous amplification capacity in PCR by a multiplex type reaction.

Table 1. Microsatellite markers used in the evaluation of 242 accessions of the Brazilian Rice Core Collection

Primer	Fluorescence	Multiplex Panel	Range	Chromosome	Author
RM1	FAM	01	76 – 119	01	Panaud et al. 1996
RM5	FAM	02	109 – 133	01	Panaud et al. 1996
RM11	HEX	03	123 – 147	07	Panaud et al. 1996
RM13	FAM	03	129 –164	05	Panaud et al. 1996
RM38	HEX	01	246 – 278	08	Chen et al.1997
RM204	HEX	04	106 –194	06	Chen et al.1997
RM207	FAM	05	84 – 158	02	Chen et al.1997
RM222	FAM	01	199 – 215	10	Chen et al.1997
RM223	HEX	06	139 – 163	08	Chen et al.1997
RM224	HEX	07	124 – 158	11	Chen et al.1997
RM229	HEX	05	106 – 131	11	Chen et al.1997
RM231	FAM	08	170 – 196	03	Chen et al.1997
RM234	FAM	09	126 – 156	07	Chen et al.1997
RM248	HEX	01	80 – 104	07	Chen et al.1997
RM252	FAM	04	193 – 277	04	Chen et al.1997
RM253	HEX	02	89 – 119	06	Chen et al.1997
RM257	HEX	06	104 – 192	09	Chen et al.1997
RM277	FAM	04	110 – 126	12	Temnykh et al. 2000
RM304	HEX	08	82 – 198	10	Temnykh et al. 2000
OG17	FAM	10	114 – 180	02	Brondani et al. 2001
OG44	HEX	10	152 – 182	03	Brondani et al.2001
OG61	FAM	09	96 – 152	05	Brondani et al.2001
OG106	FAM	11	200 – 250	09	Brondani et al.2001
4653	FAM	07	80 - 170	12	Rangel 2005
4879	HEX	11	108 – 150	04	Rangel 2005

Statistical analyses

The statistical analyses were conducted using the spread sheet with data information from the set of 25 SSR loci, including all accessions. Once the individual PIC values for each evaluated marker had been determined, different spread sheet were set up for the genotyping of: 1) SSR marker groups that presented high PIC estimates in the analysis of the 242 accessions; 2) SSR marker groups with low PIC estimates; 3) SSR marker groups with intermediate PIC estimates; and 4) SSR marker groups with high and low PIC estimates. The number of SSR markers present in each analysis was also varied within each group described above by using groups with 5, 10, 15, and 20 markers (Table 3). Altogether, 16 combinations of SSR marker groups were evaluated in two categories: number of markers (5 to 20) and mean PIC of the SSR of each group (high, mean, mixture of high and low, and low), to characterize the 242 rice accessions in the groups Introduced and Improved of the Brazilian Rice Core Collection (BRCC) (Table 3).

The number of alleles per locus were determined and the PIC values estimated by the program Genetic Data Analysis (GDA) (Lewis and Zaykin 2001). The PIC was determined by:

$$PIC_i = 1 - \sum_{j=1}^{n} x_{ij}^2$$

where X is the frequency of the j^{th} allele of the i^{th} marker.

The cophenetic coefficients (CC) used to compare the degree of similarity between the matrices of genetic distances were calculated by the program NTSYS (Rohlf 1989) with the option MxComp. We used the genetic distance of Rogers modified by Wright (RW) (Wright 1978).

The program Identity (Horst and Kristina 1999) was used to obtain the values of Probability of Identity (PI) and the number of pairs of identical individuals. The probability of identity was defined as the probability to find two individuals with the same genotype for the determined locus by chance, calculated as follows:

$$PI = \sum_{i=1}^{n} (pi^2)^2 + \sum_{i=1}^{n} \sum_{j=i+1}^{n} (2pipj)^2$$

where p_i and p_j are the frequencies of alleles *i* and *j*, and with $i \neq j$. The combined probability of identity was obtained by the multiplication of the P.I. values obtained for each SSR locus.

RESULTS AND DISCUSSION

Optimization of the multiplexes of fluorescent-labeled SSR markers

The formation of multiplexes was based on the expected mean fragment size of each marker. However, the overlapping of the amplification range of the markers somewhat limited their combinations in the reactions. Altogether, 9 duplex compositions were determined, one triplex and one quadruplex (Table 1).

The obtained results indicated the SSR primer concentrations of 0.5 and 0.9 ì M as the best for the use in multiplex reactions combining two (duplex) and three (triplex) loci, respectively. The ideal DNA concentration was considered 15 ng reaction^{-1} for having resulted in a greater specificity of the amplified product and a lower number of failures. When more than one SSR locus was amplified in the same reaction, no occurrence of interference of one locus on the other was observed.

Allelic variation of the genotypes in the groups Improved and Introduced of the BRCC

The size of the alleles amplified by the 25 SSR loci varied from 68 (RM 248) to 274 base pairs (RM 38). The number of detected alleles varied from 9 (RM05 and RM257) to 30 alleles (RM 204), with an average of 17.8 alleles/ locus and a total of 446 alleles. For the 25 SSR markers used in the genotyping of the 242 bulk accessions, PIC values that varied from 0.64 (RM252) to 0.92 (OG106 and RM257) were observed with the estimate mean of 0.82 (Table 2). Eighty-nine private or exclusive alleles were

Table 2. Estimate of values of probability obtained for the 25 microsatellite markers used in the analysis of the accession groups Improved and Introduced of the BRCC

Marker	Chromosome	Number of detected alleles		PIC	
		in Bulk	open Bulk	in Bulk	open Bulk
RM1	01	20	20	0.782	0.743
RM5	01	9	9	0.732	0.731
RM11	07	12	12	0.796	0.780
RM13	05	17	17	0.737	0.720
RM38	08	18	18	0.908	0.894
RM204	02	30	30	0.884	0.871
RM207	06	22	23	0.861	0.861
RM222	10	17	17	0.839	0.845
RM223	08	16	16	0.704	0.639
RM224	11	15	15	0.877	0.874
RM229	11	11	11	0.821	0.833
RM231	03	16	16	0.813	0.804
RM234	07	16	16	0.841	0.831
RM248	07	12	12	0.841	0.820
RM252	04	22	22	0.645	0.633
RM253	06	11	11	0.773	0.755
RM257	09	25	25	0.921	0.912
RM277	12	9	9	0.706	0.663
RM304	10	21	21	0.871	0.852
OG17	02	24	24	0.779	0.787
OG44	03	17	17	0.855	0.853
OG61	05	25	25	0.886	0.884
OG106	09	26	26	0.922	0.912
4653	12	21	21	0.883	0.875
4879	04	14	14	0.821	0.837
MEAN	-	17.84	17.88	0.820	0.808

identified among the 242 genotypes and only the markers RM5 and RM277 did not detect private alleles. The number of exclusive alleles detected among the markers varied from 1 (RM224, RM248, RM253, and RM257) to 9 (RM252 and RM204).

Analysis of the individuals in the bulks of accessions

The bulks of genotyped accessions consisted of 4 individuals, since a previous evaluation had shown that above this number, alleles that occured in low frequency might not be detected by the PCR reaction (data not published); besides, the accessions evaluated in our study were lines and cultivars, so the great genetic variability between plants of a same accession would not have been expected. Nevertheless, the bulk analysis of the 242 genotypes allowed the identification of accessions with two or more alleles per SSR marker derived from the mixture of homozygotic individuals or of the presence of heterozygous individuals for each locus. Of the 25 genotyped SSR markers, 10 detected the presence of more than one allele per bulk (RM05, RM11, RM13, RM207, RM229, RM231, RM253, RM304, OG17, and OG44). Altogether, 103 accessions were identified with one or more loci in heterozygosis. Among these, 65 (63%) were composed by accessions that presented bulk in heterozygosis for only one locus. Thirteen accession bulks were heterozygous for two loci, 15 were heterozygous for three loci, five heterozygous for four loci, two heterozygous for five loci, two heterozygous for six loci, and two heterozygous for seven loci. Each one of the four individuals in the accession bulks which presented at least two bands for more than two loci was genotyped individually for each marker, resulting in the opening of 39 bulks, making up a total of 359 genotyped individuals. For 19 opened bulks, at least one individual in heterozygosis was identified.

Table 3. Groups formed by different markers based on their PIC

	Group	Markers
high PIC	1	RM01, RM253, RM11, OG17, RM231, RM248, RM234, RM229, 4879, RM222, RM304, OG44, RM207, RM204, RM224, 4653, OG61, RM38, RM257, OG106
	2	RM248, RM234, RM229, 4879, RM222, RM304, OG44, RM207, RM204, RM224, 4653, OG61, RM38, RM257, OG106
	3	RM304, OG44, RM207, RM204, RM224, T4653, OG61, RM38, RM257, OG106
	4	4653, OG61, RM38, RM257, OG106
Low PIC	5	RM252, RM223, RM277, RM13, RM05, RM01, RM253, RM11, RM252, RM223, RM277, RM13, RM05, RM01, RM253, RM11, OG17, RM231, RM248, RM234, RM229, 4879, RM222, RM304, OG44, RM207, RM204, RM224
	6	RM252, RM223, RM277, RM13, RM05, RM01, RM253, RM11, OG17, RM231, RM248, RM234, RM229, 4879, RM222
	7	RM252, RM223, RM277, RM13, RM05, RM01, RM253, RM11, OG17, RM231
	8	RM252, RM223, RM277, RM13, RM05
Intermediate PIC	9	RM13, RM05, RM01, RM253, RM11, OG17, RM231, RM248, RM234, RM229, 4879, RM222, RM304, OG44, RM207, RM204, RM224, T4653, OG61, RM38
	10	RM01, RM253, RM11, OG17, RM231, RM248, RM234, RM229, 4879, RM222, RM304, OG44, RM207, RM204, RM224
	11	RM11, OG17, RM231, RM248, RM234, RM229, 4879, RM222, RM304, OG44
	12	RM248, RM234, RM229, 4879, RM222
High and low PIC	13	RM252, RM223, RM277, RM13, RM05, RM01, RM253, RM11, OG17, RM231, RM304, OG44, RM207, RM204, RM224, 4653, OG61, RM38, RM257, OG106
	14	RM252, RM223, RM277, RM13, RM05, RM01, RM253, RM207, RM204, RM224, 4653, OG61, RM38, RM257, OG106
	15	RM252, RM223, RM277, RM13, RM05, 4653, OG61, RM38, RM257, OG106
	16	RM252, RM223, RM277, RM257, OG106

Evaluation of the number of SSR markers in the analysis of rice

Based on the PIC data of the 25 markers obtained in the analysis of the 359 individuals (242 accessions), 16 combinations were set involving different numbers of markers and respective information content (PIC) values, with the aim of determining the adequate number of markers for the BRCC analysis.

The number of detected alleles varied according to the group of markers used and its PIC estimates. Few alleles were identified in the SSR marker groups with low PIC estimate. The higher the PIC values in the marker groups, the greater the capacity of allele detection, except in the analysis with the group formed by five markers with intermediate PIC, where the lowest number of alleles was detected.

The combined probability of identity (PI) for the 25 SSR markers was 1.55×10^{-28} (Table 4). The same probability, in a separate estimation for the groups of evaluated markers varied from 6.5×10^{-24} for group 1 (20 SSR markers with average PIC of 0.84) to 2.5×10^{-4} for group 8 (5 SSR markers with average PIC of 0.67). In relation to the variation of the number of markers, group 2 with 15 markers with high PIC (mean PIC of 0.87), and the groups 9 (intermediate PIC), 5 (low PIC), and 13 (high and low PIC), formed by 20 markers, showed the lowest PI values.

The detection of identical genotype pairs was realized with the information obtained in the analysis with the 25 SSR markers, with the bulk accessions and the opened bulk accessions (Table 4). The number of identical genotype pairs varied from one identified pair, obtained by the analysis of the 242 bulk accessions evaluated based on the 25 markers, to 2446, a value obtained by the analysis of the set of 359 individuals realized by group 16 (Table 4). For the analysis of the opened bulks (359 individuals) the group 1 (20 markers) with PI of 6.5×10^{-24}, identified 90 identical genotype pairs, and among these, only one pair corresponded to two distinct accessions (cultivar Irat 142 and line YN906 UUL 65) and the remaining 89 pairs were identical individuals detected within the same accession. The analysis with all 25 SSR markers (PI of 1.55×10^{-28}), identified 70 pairs, of identical genotypes, just one of them was distinct accessions (the same detected by group 1) (Table 4). In the analysis of the 242 accessions, without opening the bulks, this pair of accessions was also considered genetically identical using the 25 SSR markers.

Table 4. Values obtained for mean PIC, average RW, identical genotype pairs, number of sampled alleles, CC, and PI for the groups of markers

SSR group	Mean PIC	average RW	Identical genotype pairs*	Number of sampled alleles	CC	PI
Group 1	0.841	0.843	90 (1) #	374	0.97	6.5×10^{-24}
Group 2	0.864	0.855	149 (8)	291	0.91	1.5×10^{-19}
Group 3	0.880	0.867	201 (49)	221	0.84	3.68×10^{-14}
Group 4	0.895	0.874	1224 (648)	115	0.74	5.3×10^{-8}
Group 5	0.787	0.817	70 (1) #	332	0.97	3.0×10^{-20}
Group 6	0.761	0.802	98 (3)	226	0.91	4.4×10^{-14}
Group 7	0.726	0.786	173 (66)	156	0.84	1.06×10^{-8}
Group 8	0.677	0.762	1648 (670)	73	0.68	2.5×10^{-4}
Group 9	0.822	0.836	70 (1) #	349	0.97	2.6×10^{-22}
Group 10	0.823	0.828	91 (2)	259	0.94	1.2×10^{-16}
Group 11	0.824	0.814	152 (22)	160	0.88	3.0×10^{-11}
Group 12	0.833	0.820	557 (221)	70	0.74	4.1×10^{-6}
Group 13	0.802	0.832	72 (2) §	377	0.98	3.9×10^{-22}
Group 14	0.798	0.840	137 (6)	287	0.94	5.3×10^{-17}
Group 15	0.786	0.831	233 (48)	188	0.89	1.3×10^{-11}
Group 16	0.752	0.795	2446 (1145)	98	0.71	6.2×10^{-6}
25 markers(closed bulk)	0.820	0.840	1	446	-	1.55×10^{-28}
25 markers(opened bulk)	0.808	0.831	70 (1) #	447	-	1.61×10^{-27}

* numbers in brackets refer to the number of pairs of identical accessions, excluding the pairs of identical individuals detected within each accession
# the pair of identical accessions corresponds to cultivar Irat 142 and line YN906 UUL 65
§ the pairs of identical accessions are cultivar Irat 142 and line YN906 UUL 65, and cultivar Irga 419 and line YN906 UUL 65
PIC - Polymorphism Information Content; CC - Cophenetic Correlation; RW - Rogers W Distance Coefficient

This result indicates that one of these accessions should be removed from the BRCC.

After the detection of bulked accessions in heterozygosis, it is fundamental to evaluate each plant component of these bulks individually. The detected heterozygosis in the bulks could be a result of the proper process of cultivar development known as residual heterozygosis (Allard 1961), or mixture of seeds, or more rarely, of a cross pollination whose pollen would have come from another rice genotype planted in a neighboring area. When opening the bulk, different alleles can be found in homozygosis, as well as heterozygous plants, that will segregate in the next generation of cultivation. Segregation in genes that are responsible for the expression of the more visible phenotypic traits, such as plant architecture or presence of awns, among others, can easily be detected in seed production fields of commercial cultivars. In cases where the genes responsible to phenotypical variation are linked to a molecular locus marker, plants segregating for particular locus or loci associated to this trait can be discarded. This shows the importance of detecting the segregating individuals, or homozygotic individuals for alleles with low frequency in a particular locus.

In relation to the average RW distance observed for each group of markers, the highest value was found for group 4 (consisting of 5 SSR markers, PIC= 0.895, RW= 0.874), followed by the other groups of high PIC: 3 (10 markers, PIC= 0.880 and RW= 0.867), 2 (15 markers, PIC= 0.860 and RW= 0.855) and 1 (20 markers, PIC= 0.840 and RW= 0.843). This order followed the PIC reduction of each one of these groups, but was a tendency that did not continue, following the order of the decreasing RW values: group 14 (10 SSR markers, PIC= 0.798 and RW= 0.840); group 9 (20 SSR markers, PIC= 0.82 and RW= 0.836); and group 13 (20 SSR markers, PIC= 0.02 and RW= 0.32). The extreme range of average RW values obtained by the different groups (0.874 and 0.762) were not pronouncedly discrepant, which is a result of the characteristically high information content inherent to the SSR markers (Table 4). Although the highest mean RW value had been obtained with only 5 markers, that is, only 5 loci in the genome were used for the characterization of the accessions, it is not recommendable that the number of markers is defined based on this parameter only, since the number of evaluated genomic regions is drastically reduced.

The highest index of cophenetic correlation between the genetic distance matrices constructed with the complete series of 25 markers for the 359 individuals was obtained with the groups of 20 markers, independently of the mean PIC values of the groups (Table 4). For the analysis of genetic similarity, it was observed that a higher number of SSR markers should be used instead of few markers with a high information content per locus. The reason is that the analysis of a greater number of markers allows the evaluation of the genetic variability in diverse points of the genome.

Summing up, for the case of the rice with a relatively narrow genetic base, the use of the highest possible number of markers within the technically and economically viable conditions of each laboratory is recommended. In a second phase of this study, the 550 accessions of the BRCC are to be genotyped with 72 fluorescent-labeled SSR markers (6 markers per linkage group) to obtain a more robust estimate of the genetic variability. Based on these data, new simulations will be realized based on the analysis with 72 markers to determine the minimal series of SSR markers necessary for a precise estimation of the genetic variability of rice genotypes.

The analysis with fluorescent-labeled SSR markers allowed a fast and precise evaluation of the 242 accessions of the groups of Improved and Introduced accessions of the BRCC. The characterization of the cultivars and lines by means of bulks was fundamental to evaluate the level of genetic variability among and within these accessions. Although the lines and cultivars were subjected to successive generations of self-pollination and selection, the SSR markers detected the presence of genotypes as mixture of pure or heterozygotic plants. It is fundamental that the number of SSR markers for the molecular characterization of rice is high and that they are regularly spaced and representative of all chromosomes, independent of their degree of information content. However, if two SSR markers are localized in the same genomic region, the more informative one should certainly be chosen for the analysis.

ACKNOWLEDGEMENTS

CNPq supported this research financially and TCOB with a Master of Science scholarship.

Avaliação do número e informatividade de marcadores SSR fluorescentes para caracterização do germoplasma de arroz

RESUMO - *A utilização de marcadores SSR fluorescentes permite a análise simultânea de diversos locos em séries multiplex, propiciando um aumento na precisão da determinação do tamanho dos alelos quando utilizados grande número de genótipos. Utilizaram-se 25 marcadores SSR fluorescentes para a genotipagem de 242 acessos da coleção nuclear brasileira de arroz. Baseando-se nos valores de PIC de cada marcador, formaram-se grupos com 5, 10, 15 e 20 marcadores. Os valores de PIC médio variaram de 0,67 a 0,89 e o número de alelos amostrado por grupo variou de 70 a 377. Para o caso de arroz, que possui uma base genética estreita, concluiu-se que é fundamental que os genótipos sejam avaliados pelo maior número possível de marcadores SSR regularmente espaçados no genoma, e que estes possuam, preferencialmente, maior informatividade.*

Palavras-chave: *Oryza sativa*, diversidade genética, recursos genéticos.

REFERENCES

Abadie TE, Cordeiro CMT, Fonseca JR, Freire MS, Alves RBN, Burle ML, Brondani C, Rangel PHN, Castro EM and Silva HT (2002) Desenvolvendo uma coleção nuclear. In: **I Congresso da cadeia produtiva de arroz/VII Reunião nacional de pesquisa de arroz – RENAPA**. Embrapa, Florianópolis, p. 259-261.

Allard RW (1961) Relationship between diversity and consistency of performance in length environments **Crop Science 1**: 127-133.

Blair MW, Hedetale V and McCouch SR (2002) Fluorescent-labeled microsatellite panels useful for detecting allelic diversity in cultivated rice (*Oryza sativa* L.). **Theoretical Applied Genetics 105**: 449-457.

Brondani C, Brondani, RPV, Rangel PHN and Ferreira ME (2001) Development and mapping of *Oryza glumaepatula*- derived microsatellite markers in the interspecific cross *O. sativa* x *Oryza glumaepatula* **Hereditas 134**: 59-71.

Brondani C, Brondani RVP and Rangel PHN (2003) **Utilização de marcadores moleculares em programas de ampliação da base genética de espécies cultivadas**. Embrapa Arroz e Feijão, Santo Antônio de Goiás, 36p.

Chen X, Temnykh SV and McCouch SR (1997) Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L) **Theoretical Applied Genetics 95**: 553-567.

Coburn JR, Temnykh SV, Paul EM and McCouch SR (2002) Design and application of microsatellite marker panels for semiautomated genotyping of rice (*Oryza sativa* L.) **Crop Science 42**: 2092-2099.

Doyle JJ and Doyle JL (1987) Isolation of plant DNA from fresh tissue. **Focus 12**: 13-15.

Frankel OH and Brown AHD (1984) Plant genetic resources today: a critical appraisal. In: Holden JHW and Williams JT (eds.) **Crop genetic resources: conservation and evaluation**. IBPGR, Winchester, p. 249-257.

Hayard MD and Hamilton NRS (1997) Genetic diversity: population structure and conservation. In: Callow BV, Ford-Lloyd B and Newbury HJ (eds.) **Biotechnology and plant genetic resources: conservation and use**. CAB international, London, p. 52-83.

Horst WW and Kristina M (1999) **Identity**. Version 1.0.

Jackson MT (1997) Conservation of rice genetic resources: the role of the international rice genebank at IRRI. **Plant Molecular Biology 35**: 61-67.

Lewis PO and Zaykin D (2001) **Genetic data analysis: computer program for the analysis of allelic data**. Version 1.0.

McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, Li Z, Xing Y, Zhang Q, Kono I, Yano M, Fjellstrom 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 Research 9**: 199-207.

Mullis K and Fallona F (1987) Specific synthesis of DNA *in vitro* via polymerase catalyzed chain reaction. **Methods Enzymology 55**: 335-350.

Nei M (1973) Analysis of gene diversity in subdivided populations **Procedures of National Academy of Science 70**: 3321-3323.

Panoud O, Chen X and McCouch SR (1996) Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.) **Molecular and General Genetics 252**: 597-607.

Powell W, Orozco-Castillo C, Chalmers KJ, Provan J and Waugh R (1995) Polimerase chain reaction – based assays for the characterization of plant genetic resources. **Eletrophoresis 16**: 1726-1730.

Rangel P (2005) **Construção de mapa e mapeamento de QTLs utilizando marcadores SSRs, ESTs e SNPs em cruzamento interespecífico *Oryza glumaepatula* e *Oryza sativa***. Msc dissertation , Universidade Federal de Goiás, Goiânia, 111p.

Rohlf FJ (1989) **NTSYS - pc: numerical taxonomy and multivariate analysis system**. Version 2.02.

Temnykh S, Park WD, Ayres N, Cartinhour S, Huack N, Lipovich L, Cho YG, Ishii T and McCouch SR (2000) Mapping and genome organization of microsatellites sequences in rice (*Oryza sativa* L.) **Theoretical and Applied Genetics 100**: 697-712.

Wrigth S (1978) **Variability within and among natural populations 4**. Chicago Press, Chicago, 352p.

Yang GP, Maroof MAS, Xu CG, Zhang, QF and Biyashev RM (1994) Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. **Molecular and General Genetics 245**: 187-194.