Crop Breeding and Applied Biotechnology 6:9-17, 2006
Brazilian Society of Plant Breeding. Printed in Brazil

Genetic variability analysis of elite upland rice genotypes with SSR markers

Claudio Brondani[1*], Kamylla da Silva Caldeira[1], Tereza Cristina Oliveira Borba[1], Priscila Nascimento Rangel[1], Orlando Peixoto de Morais[1], Emílio da Maia de Castro[1], Paulo Hideo Nakano Rangel[1], João Antônio Mendonça[1], and Rosana Vianello Brondani[1]

Received 15 June 2005

Accepted 20 December 2005

ABSTRACT - *Rice is one of the staple foods consumed worldwide, and rice breeding programs have become important to warrant high yield levels and grain quality in upland rice. This study aimed at insights on the genetic variability of 30 elite genotypes from a VCU trial, of the upland rice breeding program of Embrapa, using 25 SSR markers. One hundred and thirty-one alleles were obtained, an average of 5.2 alleles per locus, and mean PIC equal to 0.61. The results indicated that genetically different elite parents from the breeding program and selection in segregating families have given rise to broad-based rice genotypes. Analyzing different combinations of 10 SSR markers, we observed that the use of more informative markers is essential to explain the genetic divergence consistently with the pedigree of each rice genotype.*

Key words: Genetic resources, *Oryza sativa*, microsatellite markers, genetic distance.

INTRODUCTION

Rice is the main food source for most part of the world population. Per-capita consumption in Brazil is high, approximately 49 kg per inhabitant year^{-1} of milled polished rice (Castro et al. 1999). The annual rice production in Brazil is around 12 million tons, which is very close to the quantity consumed in the country. Population increase and the possibility of occurrence of a production drop due to climatic factors or occurrence of diseases call for the uninterrupted development of more productive cultivars, more tolerant to biotic and abiotic stress. In Brazil, rice is cultivated in two systems: upland rice, which covers 40% of the production, and irrigated rice (Yokoyama et al. 1999). Due to the limited capacity of expansion of cultivation areas of irrigated rice and the increasing costs for the establishment of the plantations, prospects are that the present production in Brazil of upland rice will at least be kept up.

VCU (Value for Cultivation and Use) trials represent a fundamental stage in the genetic breeding program of rice since they evaluate, over at least two consecutive years, the performance of best lines for traits such as yield, disease resistance, grain quality, among others (Morais et al. 2005). At the end of each cycle of VCU trials, normally one or two lines with excellent agronomic performance are released as new commercial cultivars. These trials, besides their overall goal of characterizing the potential of lines as cultivars, also represent one of the main sources of parents for the development of new populations that can be exploited in the breeding program. The value of a population as source of new lines depends, among other parameters, on the magnitude of the genetic variability that it has for the traits of interest, which in turn depends on the genetic divergence of the parents involved in the crossings that originated a population. One of the forms of evaluating the level of genetic variability available for the breeding program is the characterization of genotypes

[1] Laboratório de Biotecnologia, Embrapa Arroz e Feijão, C. P. 179, 74.001-970, Goiânia, GO, Brasil. E-mail: brondani@cnpaf.embrapa.br

by molecular markers (Brondani et al. 2003, Lu et al. 2005). SSR (Simple Sequence Repeats) markers (Weber and May 1989), also known as microsatellites, have been used widely in the genetic characterization of rice and have, as main advantage, over 2,000 of these markers for the species available (McCouch et al. 2002), are more informative in comparison to other classes of molecular markers (Powell et al. 1995), and are obtained by PCR (Polymerase Chain Reaction).

This study aimed to evaluate, by means of SSR markers, the genetic variability of the genotypes that participate in the VCU trial of the Embrapa genetic breeding program for upland rice.

MATERIAL AND METHODS

Plant material

Twenty lines and 10 commercial rice cultivars were used in this study, components of the VCU trials in 2003 and 2004, of the genetic breeding program of upland rice of the Embrapa (Table 1).

Molecular analysis

Twenty-five SSR markers (Akagi et al. 1996, Chen et al. 1997, Brondani et al. 2001) were used, selected in view of their high capacity of detecting polymorphism in rice and for being distributed across the 12 linkage groups of this species (Table 2). Amplification reactions were realized in a final volume of 13 mL, containing 0.3 mM of each primer, 1 U of the enzyme Taq polymerase, 0.2 mM of each dNTP, 1 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 1.3 mL of DMSO (50%), and 7.5 ng of DNA template. The PCR analyses were performed in a PT-100 (MJ Research) thermocycler with the following program: one pre-cycle at 96 ºC for 2 minutes, followed by 30 cycles at 94 ºC for 1 minute, 56 ºC for 1 minute and 72 ºC for 1 minute, and a final extension step of 7 minutes at 72 ºC. The PCR product was subjected to electrophoresis under denaturing conditions in 6% acrylamide gels (containing 7 M urea), which were silver-stained, according to the procedure described by Bassam et al. (1991).

Statistical analysis

The number of alleles per locus, private alleles and PIC (Polymorphism Information Content) were estimated using the software GDA (Lewis and Zaykin 2000). The dendrograms were constructed based on the genetic distance matrix obtained by the distance coefficient of Rogers, modified by Wright (1978), henceforth called Rogers-W, and the lines clustered by the UPGMA method, available in software NTSYS (Rohlf 1989). This program was also used for calculations of the correlation coefficient *r* between the distance matrices. Values of probability of identity and the determination of private alleles were supplied by the program Identity (Sefc et al. 1997). The bootstrap analysis, a computer technique that determines the precision of most statistical estimates (Efron et al. 1985), was used to determine the proportion (by means of a re-sampling process) in which each cluster containing the same genotypes was obtained in the dendrogram, on software Bood (Coelho 2000).

RESULTS AND DISCUSSION

One of the main applications of SSR markers in genomic analysis is the characterization of genetic resources to support genebank administration (Yang et al. 1994), as much as to estimate the genetic variability in genotypes used in genetic breeding programs (Lu et al. 2005). Specifically for this case, the degree of relatedness between cultivars and lines of the program can be precisely determined, as well as the degree of genetic purity in advanced lines and cultivars.

Genetic variability

The number of alleles per SSR marker varied from 2 to 9, with a mean of 5.2 alleles per marker, making up a total of 131 alleles. The PIC (Polymorphism Information Content) varied from 0.13 (SSR markers 4797 and 5132) to 0.84 (RM229) with a general mean of 0.61 (Table 2). The results were similar to those found by Ni et al. (2002), who obtained 6.8 alleles per locus and PIC equal to 0.62 when they evaluated 38 rice cultivars of the subspecies *indica* and *japonica* with 111 SSR markers. The subspecies *indica* and *japonica* were originated from a common ancestor, a million of years ago, and have the greatest possible genetic variability within the species *Oryza sativa* (Han and Xue 2003). The Ni et al. (2002)' results serve as parameter to indicate the high genetic variability of the genotypes in the VCU trial of upland rice.

The markers that detected genotypes in heterozygosis were OG61 (BRA01644), RM22 (IAC202), RM38 (CNA9019), RM222 (BRA01596), and RM223 (CNA9025). Lines in evaluation in the VCU experiment, that have at least 8 generations of selfing, i.e., theoretically have only a mean of 0.78% of the genome in heterozygosis, are therefore already considered fixed lines. The identification of loci in heterozygosis does not necessarily

Table 1. Lines and cultivars in the VCU trial of upland rice of the Embrapa Rice and Beans in 2003 and 2004

Genotype	Germplasm	Original crossing
CNA10217	Line	Mearim/CT6196-33-11-2-3-B//Carajás
CNA10222	Line	Mearim/CT6196-33-11-2-3-B//Carajás
CNA10260	Line	Katy/CNA 7706
CNA9019	Line	CNA7680/CNA7726
CNA9023	Line	CNA7680/CNA7726
CNA9025	Line	CNA6710/IAC1150//150144/CNAx4036-5-1-1
CNA9026	Line	CNA6710/IAC1150//150144/CNAx4036-5-1-1
CNA9045	Line	CNA7914/CNAx3031-13-B-1-1//CNA7455/Colômbia1
BRA01504	Line	CNA8077/CIRAD 141//CNA8198/Lemont
BRA01506	Line	CNA8077/CIRAD 141//CNA8198/Lemont
BRA01545	Line	CNAx5626-2-M2-M/CT11251-7-2-M-M-BrM1-M
BRA01580	Line	Kaybonnet/Population CG2
BRA01593	Line	Kaybonnet/CNA7119
BRA01594	Line	Kaybonnet/CNA7119
BRA01596	Line	Kaybonnet/CNA7119
BRA01600	Line	Kaybonnet/CNA7119
BRA01618	Line	Kaybonnet/CNA7706
BRA01644	Line	Mearim/CT6196-33-11-2-3-B//Carajás
CNA8812	Line	CT9978-12-2-2P-4/CT10037-56-4-M-4-1P-1//CT9899-12-3-M3-3
CNA8817	Line	CT6516-23-10-1-1-3-B/CNAx1413-16-2-B-1
BRSMG Conai	Cultivar	IAC 164/IRAT 216
BRS Carisma	Cultivar	CT7244-9-1-5-3/CT6196-33-11-1-3//CT6946-2-5-3-3-2-M
BRS Caripuna	Cultivar	CT7244-9-2-1-52-1/CT6261-5-7-2P-5-1P//P5589-1-1-3P-4
BRS Talento (cultivar)	Cultivar	CT7244-9-1-5-3/CT6196-33-11-1-3//CT6946-2-5-3-3-2-M
BRS Bonança (cultivar)	Cultivar	CT7244-9-2-1-52-1/CT7232-5-3-7-2-1P//CT6196-33-11-1-3-AP
BRS Primavera (cultivar)	Cultivar	IRAT 10/LS85-158
BRS Liderança	Cultivar	Kaybonnet/CNA7119
BRS Colosso	Cultivar	Kaybonnet/CNA7119
Canastra (cultivar)	Cultivar	TOX939-107-2-101-1B/Colômbia 1XM312A//TOX1780-2-1-1P-4
IAC 202	Cultivar	Lebonnet/IAC25

mean that a certain line segregates for a trait of interest, which would make the commercial release unfeasible. Nevertheless, it is interesting that lines on the verge of becoming commercial cultivars have genetic homogeneity, and that this can be proved by molecular markers in a routine fingerprinting analysis. If the lines BRA01644, CNA9019, BRA01596 and CNA9025 are selected as new cultivars, the SSR markers that identify loci in heterozygosis will be used to select homozygous plants for the most frequent allele of each marker, in each one of these lines.

Thirty-one private or exclusive alleles were detected, six of them by marker RM22 (Table 3). The line with the highest number of private alleles was BRA01545, with 4 alleles; in other words, its genetic variability in the evaluated loci is unique, not found in the other genotypes of the VCU trial. Although the molecular markers are neutral, i.e., the molecular profile of a genotype does not vary from one cultivation environment to another, it is possible to draw conclusions on the association between alleles of a particular locus and the agronomic performance of any trait of interest in rice (Virk et al. 1996). These direct associations exist due to linkage disequilibrium, which in rice occurs principally due to the autogamous reproductive habit (Ford-Lloyd et al. 1997). Genotypes with private alleles can therefore be used firstly as parents of the breeding program, to amplify the genetic base of upland rice lines. On the other hand, the molecular profile of the genotypes of the VCU trial obtained by molecular markers can be used to orientate the addition of novel rice genotypes as

Table 2. SSR markers, their chromosome localization (Chrom.), PIC (Polymorphism Information Content), number of detected and private alleles, with size in base pairs (pb)

SSR	Chrom.	PIC	Detected alleles	Private allele (pb)	Genotypes with private allele
4653	12	0.81	7	106	BRA01545
4661	9	0.49	2	-	-
4712	4	0.82	9	182	IAC202
				198	Canastra
				196	BRA01545
				190	CNA9045
4797	4	0.13	2	-	-
4879	4	0.53	3	-	-
4961	11	0.21	2	-	-
5132	10	0.13	2	-	-
OG7	11	0.76	8	152	BRA01644
				150	BRA01580
				160	CNA9025
OG44	3	0.77	5	-	-
OG61	5	0.79	7	154	BRS Bonança
OG106	9	0.67	5	230	BRS Primavera
OS19	6	0.61	6	106	BRA01644
RM1	1	0.46	2	-	-
RM9	1	0.77	7	186	CNA8817
				192	CNA8812
				190	BRS Caripuna
RM22	3	0.64	9	160	IAC202
				150	CNA8812
				198	BRS Bonança
				178	BRA01545
				174	CNA9045
				168	CNA10217
RM38	8	0.59	5	280	BRA01545
RM204	6	0.64	4	-	-
RM207	2	0.81	6	-	-
RM222	10	0.23	2	-	-
RM223	8	0.78	8	146	BRS Primavera
				156	BRSMG Conai
				144	BRA01644
				68	CNA9045
RM224	11	0.81	8	170	BRA01545
				120	CNA10222
RM229	11	0.84	6	-	-
RM247	12	0.66	6	168	IAC202
				166	CNA8812
				164	BRS Carisma
RM248	7	0.59	6	94	CNA9023
RM263	2	0.67	4	-	-
Mean	-	0.61	5.24	-	-

parents in the genetic breeding program, provided that they have, in turn, a differentiated molecular standard. This strategy has been applied as criterion for the construction of core collections for genetic rice breeding (Xu et al. 2004) and to select controlled crossings aiming to maximize the probability of finding transgressive segregation (Spada et al. 2004).

Genetic distance

The mean Rogers-W distance coefficient was 0.74 (Table 3), and was used to establish the limiting value for formation of genetic similarity clusters in the dendrogram (Figure 1). Six Clusters (A to F), were formed and among these, genotypes were identified that have parents in common: BRS Carisma and BRS Talento (Cluster B); BRA01504 and BRA01506, and CNA9025 and CNA9026 (Cluster C); CNA9019 and CNA9023 (Cluster D); BRA01593, BRA01594, BRA01596, BRA01600, BRS Liderança and BRS Colosso (Cluster E); CNA10217 and CNA10222 (Cluster F). There were two exceptions: the first was that line BRA01644 should be located in Cluster F, and not in Cluster E since it has the same parents as the lines CNA10217 and CNA10222. The second exception were CNA8817 and Canastra, which do not have parents in common and were grouped with a Rogers-W genetic distance equal to 0.2. The use of 25 SSR markers resulted in a probability of identity of $2.87\ 10^{-15}$ (Table 3), which corresponds to the probability that two identical genotypes are randomly identified. This very low value strongly indicates that the genotypes CNA8817 and Canastra have one or more parents in common, although the pedigree registries do not indicate this (Table 1). The two exceptions may have occurred, at some stage of the breeding program, through an error of identification of the genotypes, by seed mixture, or due to cross fecundation, through pollen migration of pollen grain of another genotype planted in an adjacent area to these genotypes. The routine use of molecular markers for the genetic

Table 3. Results of the analysis of genetic diversity obtained for SSR marker groups

SSR Group	SSR	Alleles	Mean PIC[1]	Mean RW	Identical Genotypes	*r*	P.I.
Complete	All 25 SSRs	131	0.61	0.74	-	-	$2.87x10^{-15}$
Group I (PIC high)	4653, 4712, OG7, OG44, OG61, RM9, RM207, RM223, RM224, RM229	71	0.80	0.54	BRA01594 and BRA01593	0.85	$2.29x10^{-9}$
Group II (Intermediate PIC)	4661, 4879, OG106, OS19, RM22, RM38, RM204, RM247, RM248, RM263	50	0.61	0.47	BRA01504 and BRA01506 BRA01600 and BRA01593 BRS Colosso and BRS Liderança		
					Canastra and CNA8817	0.84	$7.04x10^{-6}$
Group III (Low PIC)	4661, 4797, 4879, 4961, 5132, OS19, RM1, RM38, RM222, RM248	32	0.40	0.37	CNA10260 and CNA10217 CNA9026 and CNA9025 BRA01618 and BRA01506 BRS Talento and BRS Carisma, BRS Bonança and CNA9045 BRS Colosso and BRS Liderança Canastra and CNA8817	0.57	$9.67x10^{-4}$
Group IV (high and low PIC)	4653, 4712, 4797, 4961, 5132, RM1, RM207, RM222, RM224, RM229	46	0.52	0.43	CNA9025 and CNA9026 BRS Colosso and BRS Liderança Canastra and CNA8817	0.68	$7.02x10^{-7}$

[1] Mean PIC: Mean of the PIC values of the evaluated SSR in each group; Mean RW: mean of the Rogers-W genetic distance coefficient of all pairwise genotype combinations; P.I.: probability of identity

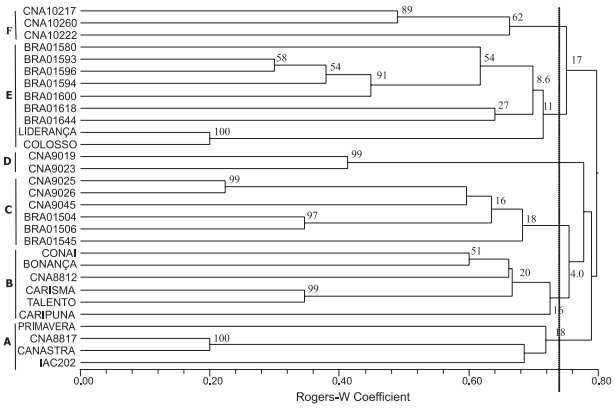

Figure 1. UPGMA based on the Rogers-W' genetic distance matrix. Values at the nodes represent the values of the bootstrap analysis. The dotted line represents the mean value of the Rogers-W genetic distance coefficient (0.74). A to F: Clusters of lowest genetic distance

characterization of all parents and the new lines and cultivars developed by the breeding program can quickly detect differences between the expected and observed pedigree of a particular genotype. Additionally, the data of molecular characterization with SSR markers can come to be used as complementary information to analyses of Distinguibility, Homogeneity and Stability (DHS) in order to index new rice cultivars (Singh et al. 2004).

Analysis of SSR markers

Data of the genotype characterization with the 25 SSR markers were used for a new analysis, involving sets of 10 markers. The objective of this analysis was to evaluate how the information content of a set of SSR markers influence the determination of the genetic distance between the genotypes of the VCU trial of upland rice, using the results obtained with the group of 25 SSR markers as a control. Four combinations were determined: Group I (SSR markers with high PIC, varying from 0.84 to 0.76, Mean PIC = 0.80), Group II (intermediate PIC, varying from 0.67 to 0.49, Mean PIC = 0.61), Group III (low PIC, varying from 0.61 to 0.13, Mean PIC = 0.4), and Group IV (mixture of markers with high and low PIC, varying from 0.84 to 0.13, Mean PIC = 0.52) (Table 3).

The highest value of the Rogers-W coefficient mean (0.54) was obtained with Group I (high PIC), where the correlation coefficient between its matrix of genetic distance and the matrix of the complete series was also highest (r= 0.85). Group I identified only one combination of identical genotypes (BRA01593 and BRA01594), and in the control dendrogram, with all 25 markers, no identical genotypes were identified (Table 3). In the dendrogram of Group I, the genotypes of the Clusters D and F of the control dendrogram were integrated in a same cluster. Three genotypes were not grouped: IAC202, CNA9045 and Caripuna (Figure 3).

Group II (intermediate PIC) detected four combinations of identical genotypes, and of these, only Canastra and CNA8817 were misclassified as identical, considering that these were not derived from a crossing involving the same parents (Table 3). The dendrogram of Group II had a correlation coefficient of the distance matrix (r= 0.84) with the matrix of the control dendrogram. Clusters E and F of the control dendrogram were preserved in Group II, but were integrated in a cluster. The same happened with Clusters C and D, excepting the line BRA01545, component of Cluster C, which was not included in any cluster. Cluster B was divided among two other clusters (Conai, Caripuna and Bonança, and Carisma and Talento). Line CNA8812 of Cluster B of the control dendrogram, grouped with cultivar Primavera, which in the control dendrogram was part of Cluster A. The other genotypes of this cluster remained in the same group, in the dendrogram of Group II.

Groups III and IV had the lowest correlation coefficient values between their distance matrices and the distance matrix of the control group (r= 0.57 and r= 0.68, respectively). Group III (Low PIC) produced the smallest values for the number of detected alleles, mean PIC, mean RW and probability of identity, and identified the highest number of identical genotype combinations (seven), of which three involved genotypes derived from a same crossing (CNA9025 and CNA9026, BRS Liderança and BRS Colosso, BRS Carisma and BRS Talento), and in the other combinations, the genotypes were misclassified as identical (CNA10217 and CNA10260, CNA9045 and BRS Bonança, CNA8817 and Canastra, BRA01506 and BRA01618), for not being derived from a same crossing (Table 3). In Group III, genotype BRA01644 was included in the same cluster as the genotypes CNA10217 and 10222, which have the same common parents, which was not possible in the control dendrogram (Cluster F). Cluster D of the control dendrogram was also maintained in the dendrogram of Group III, but with the addition of genotypes derived from the Clusters A, B and C. The genotypes of the Clusters A, B, C and E were distributed in different clusters. Group IV (high and low PIC) detected three identical genotype combinations, where only Canastra and CNA8817 were misclassified as identical genotypes

(Table 3). In the dendrogram of Group IV, all genotypes were redistributed among nine clusters, in relation to the clusters obtained in the control dendrogram, except for Cluster A.

The different marker combinations evaluated in Groups I to IV produced somewhat contrasting results. The ideal set of markers for the molecular characterization of rice should be able to explain the genetic divergence consistent with the pedigree of each genotype. Comparing the results of these groups with the control group (25 SSRs), the tendency was observed that genotypes were misclassified as identical when the mean PIC was lower than 0.6 (Groups III and IV), due to the low mean information content of the markers. Since SSR markers sample specific genome regions, and with a variable degree of information content, resulting from different evaluation patterns (Li et al. 2002), it is important that SSR markers should be widely distributed in the genome.

CONCLUSIONS

- Genetically divergent parents and selection procedures of segregating families in the breeding program of upland rice of the Embrapa have brought forth broad-based inbred lines.
- For the molecular characterization of genotypes of upland rice with SSR markers, it is important that the mean PIC is higher than 0.60 and that the markers are well-distributed in the rice genome to obtain results that corroborate the pedigree of these genotypes.

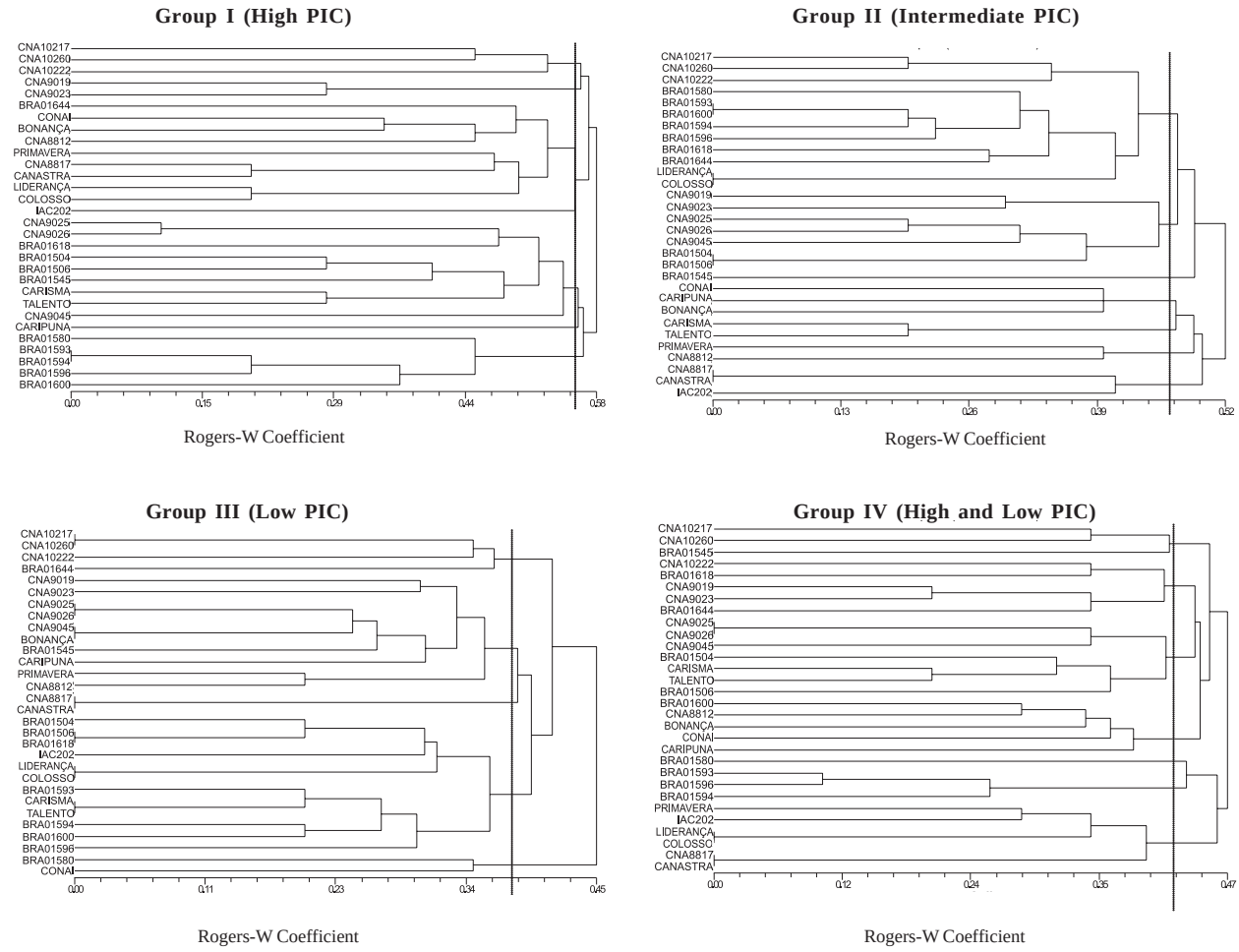

Figure 2. UPGMA obtained for Groups I to IV of SSR markers. The dotted line represents the mean value of the Rogers-W genetic distance coefficient: Group I, Mean RW= 0.54; Group II, Mean RW= 0.47; Group III, Mean RW= 0.37; Group IV, Mean RW= 0.43

Análise da variabilidade genética de genótipos-elite de arroz de terras altas com marcadores SSR

RESUMO - *O arroz é um dos principais alimentos consumidos no mundo e o melhoramento genético tem sido importante para garantir altos patamares de produtividade e qualidade de grão no sistema de cultivo de terras altas. Este trabalho objetivou inferir a variabilidade genética de 30 genótipos-elite do ensaio de avaliação do valor de cultivo e uso (VCU), do programa de melhoramento de arroz de terras altas da Embrapa, utilizando 25 marcadores SSR. Foram obtidos 131 alelos, com média de 5,2 alelos por loco e PIC médio de 0,61. Estes resultados indicam que os genitores-elite geneticamente divergentes do programa de melhoramento e a seleção de famílias segregantes têm produzido genótipos de arroz com ampla base genética. Analisando diferentes combinações de 10 marcadores SSRs, observou-se que a utilização de marcadores mais informativos é fundamental para explicar a divergência genética de modo consistente com a genealogia de cada genótipo de arroz.*

Palavras-chave: Recursos genéticos, *Oryza sativa*, marcadores microssatélites, distância genética.

REFERENCES

Akagi H, Yokozeki Y, Inagaki A and Fujimura T (1996) Microsatellite DNA markers for rice chromosomes. **Theoretical and Applied Genetics 93**: 1071-1077.

Bassam BJ, Caetano-Anolles G and Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrilamide gels. **Analytical Biochemistry 196**: 80-83.

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

Brondani C, Brondani RPV 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.

Castro EM, Vieira NRA, Rabelo RR and Silva S.A. (1999) **Qualidade de grãos em arroz**. Embrapa Arroz e Feijão, Santo Antônio de Goiás, 30p.

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

Coelho ASG (2000) Bood Software Version 3.04. Available upon request at coelho@icb.ufg.br.

Efron B, Halloran E and Holmes S (1996) Bootstrap confidence levels for phylogenetic trees. **Proceedings of the National Academy of Sciences 93**: 13429-13434.

Ford-Lloyd BV, Jackson MT and Newbury HJ (1997) Molecular markers and the management of genetic resources in seed genebanks: a case study of rice. In: Callow JA, Ford-Lloyd BV and Newbury HJ (eds.) **Biotechnology and plant genetic resources – conservation and use**. CAB, Wallingford, p. 103-118.

Han B and Xue Y (2003) Genome-wide intraspecific DNA-sequence variations in rice. **Current Opinion in Plant Biology 6**: 134-138.

Lewis PO and Zaykin D (2000) Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d15). Available at http://alleyn.eeb.uconn.edu/gda/2000.

Li YC, Korol AB, Fahima T, Beiles A and Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. **Molecular Ecology 11**: 2453-2465.

Lu H, Redus MA, Coburn JR, Rutger JN, McCouch SR and Tai TH (2005) Population structure and breeding patterns of 145 U.S. rice cultivars based on SSR marker analysis. **Crop Science 45**: 66-76.

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, DeClerk 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.

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

Morais OP, Rangel PHN, Fagundes PRR, Castro EM, Neves PCF, Brondani C, Prabhu AS and Magalhães AM (2005) Avanços do Melhoramento Genético do Arroz no Brasil. **Anais do III Congresso Brasileiro de Melhoramento de Plantas**. SBMP, Viçosa, p. 1-37. (CD-ROM).

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

Rohlf FJ (1989) **NTSys-Pc: Numerical Taxonomy and Multivariate Analysis System**. Exeter Publisher, New York, 130p.

Sefc KM, Steinkellner H, Wagner HW, Glössl J and Regner F (1997) Application of microsatellite markers to parentage studies in grapevine. **Vitis 36**: 179-183.

Singh RK, Sharma RK, Singh AK, Singh VP, Singh NK, Tiwari SP and Mohapatra T (2004). Suitability of mapped sequence tagged microsatellite markers for establishing distinctness, uniformity and stability in aromatic rice. **Euphytica 135**: 135-143.

Spada A, Mantegazza R, Biloni M, Caporali E and Sala F (2004) Italian rice varieties: historical data, molecular markers and pedigrees to reveal their genetic relationships. **Plant Breeding 123**: 105-111.

Virk PS, Ford-Lloyd BV, Jackson MT, Pooni HS, Clemeno TP and Newbury HJ (1996) Predicting quantitative variation within rice germplasm using molecular markers. **Heredity 76**: 296-304.

Xu Y, Beachell H and McCouch SR (2004) A marker-based approach to broadening the genetic base of rice in the USA. **Crop Science 44**: 1947-1959.

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

Yokoyama LP, Rucatti EG and Kluthcouski J (1999) Economia da produção: conjuntura, mercado e custos. In: Vieira NR, Santos AB and Sant'ana EP (eds.) **A cultura do arroz no Brasil**. Embrapa Arroz e Feijão, Santo Antônio de Goiás, p. 37-57.

Weber JL and May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the Polymerase Chain Reaction. **American Journal of Human Genetics 44**: 388-396.

Wright S (1978) Evolution and the genetics of populations (vol 4). **Variability within and among natural populations**. University of Chicago, Chicago, 254p.