Online journal ISSN - 1676-5680
GMR Genetics and Molecular Research
EVOLUTION AND TECHNOLOGY

Agronomic and molecular characterization of introgression lines from the interspecific cross *Oryza sativa* (BG90-2) x *Oryza glumaepatula* (RS-16)

P.N. Rangel[1], R.P.V. Brondani[2], P.H.N. Rangel[3] and C. Brondani[2]

[1]Laboratório de Biotecnologia, Embrapa Arroz e Feijão,
Santo Antônio de Goiás, GO, Brasil
[2]Genética Molecular, Embrapa Arroz e Feijão,
Santo Antônio de Goiás, GO, Brasil
[3]Melhoramento de Arroz, Embrapa Arroz e Feijão,
Santo Antônio de Goiás, GO, Brasil

Corresponding author: P.N. Rangel
E-mail: pn.rangel@gmail.com

Genet. Mol. Res. 7 (1): 184-195 (2008)
Received October 15, 2007
Accepted December 22, 2007
Published February 19, 2008

ABSTRACT. The reduced genetic variability of modern rice varieties (*Oryza sativa*) is of concern because it reduces the possibilities of genetic gain in breeding programs. Introgression lines (ILs) containing genomic fragments from wild rice can be used to obtain new improved cultivars. The objective of the present study was to perform the agronomic and molecular characterizations of 35 BC_2F_8 ILs from the cross *O. glumaepatula* x *O. sativa*, aiming to select high-yielding ILs to be used in rice-breeding programs. All 35 ILs were field evaluated in the season 2002/2003 in three locations and the 15 best performing ones were evaluated in the season 2003/2004 in five locations. In 2003/2004, six ILs (CNAi 9934, CNAi 9931, CNAi 9930, CNAi 9935, CNAi 9936, and CNAi 9937) showed the highest yield means and were statistically superior to the controls Metica 1 and IRGA 417. Molecular characterization of the 35 ILs was performed with 92 microsatellite markers distributed on the 12 rice chromosomes and a simple regression

quantitative trait locus analysis was performed using the phenotypic data from 2002/2003. The six high-yielding ILs showed a low proportion of wild fragment introgressions. A total of 14 molecular markers were associated with quantitative trait loci in the three locations. The six high-yielding ILs were incorporated in the Embrapa breeding program, and the line CNAi 9930 is recommended for cultivation due to additional advantages of good grain cooking and milling qualities and high yield stability. The *O. glumaepatula*-derived ILs proved to be a source of new alleles for the development of high-yielding rice cultivars.

Key words: *Oryza glumaepatula*; Introgression lines; Simple sequence repeat markers; Yield

INTRODUCTION

Modern rice varieties (*Oryza sativa*) are the result of an extensive artificial selection process that led to an extreme pressure on a few target characteristics to rice cultivation, such as non-shattering of seeds, compact growth habit and loss of germination inhibition (Tanksley and McCouch, 1997). This strongly directional selection reduced the genetic variability of cultivated rice due to a great loss of allelic variability in detriment to the fixation of some alleles, which resulted in a phenomenon called "genetic erosion" (Gowda et al., 2003). In addition, rice-breeding programs tend to favor methods that maximize endogamy in cultivar development, which drastically reduces new recombination opportunities (Rangel and Neves, 1997). Besides, breeders usually use the same adapted genitors repeatedly in initial breeding crosses (Moncada et al., 2001). These events led to a concerning restriction of selection gains obtained by breeding programs over the past years. One of the objectives of modern breeding has been the recovery of lost diversity through the search of potentially favorable alleles in wild ancestors of rice (Gur and Zamir, 2004).

The advanced backcross quantitative trait locus (AB-QTL) analysis (Tanksley and Nelson, 1996) is a powerful strategy to exploit and use the potential of wild alleles in breeding programs. This methodology integrates the QTL analysis and the introgression of alleles from wild germplasm into elite material under the assumption that marker regions positively associated with traits of agronomic interest can be identified and transferred into elite cultivars (Bernacchi et al., 1997; Frary et al., 2004). The AB-QTL analysis comprises a set of activities that include the development of a backcross population derived from an interspecific cross followed by its molecular and phenotypic characterization for QTL analysis. Marker loci associated with favorable wild alleles can be used to select genotypes containing these specific genomic regions. After a few selfing generations, introgression lines (ILs) are obtained and can be field tested and used for variety development (Frary et al., 2004). Since they contain small wild fragments evenly distributed throughout the elite recurrent genome, ILs can be used for genetic and functional genomics studies, such as the dissection of gene functions and map-based cloning of QTLs underlying quantitative and qualitative traits (Li et al., 2004; Tian et al., 2006b). In addition, ILs are an important reservoir of alleles that can be used in breeding programs for the development of new cultivars with higher genetic diversity and that are more resistant to biotic and abiotic stresses.

Species related to the cultivated *Oryza sativa* have been used as an additional source of genetic variability in breeding programs, such as *Oryza glumaepatula* (Brondani et al., 2001), *Oryza*

rufipogon (Moncada et al., 2001; Nguyen et al., 2003; Septiningsih et al., 2003) and *Oryza glaberrima* (Aluko et al., 2004). These studies used the AB-QTL methodology to monitor the effects of wild introgressions on important agronomic traits such as grain yield, aluminum tolerance and grain quality. *O. glumaepatula* is a diploid AA species that is autogamous and has an annual life cycle (Vaughan et al., 2003). Populations of *O. glumaepatula* were identified in three Brazilian biomes (Amazon, Pantanal and Cerrados) and their adaptation to tropical soil and weather increase their chance as donors of genes related to traits of economic importance to rice (Brondani et al., 2005).

The efficiency of using best performing ILs in breeding programs depends on a complete phenotypic and molecular characterization. Phenotypic characterization, which has been a routine in breeding programs, provides a great amount of information about IL field performance, helping breeders to choose those with desirable traits. Molecular characterization can be used as a complement to field evaluation, providing information about the position and estimated size of introgressed fragments on each IL. The objective of the present study was to perform agronomical and molecular characterizations of ILs derived from an *O. glumaepatula* (RS-16) x *O. sativa* (BG90-2) interspecific cross (Brondani et al., 2002). These lines were developed at Embrapa Rice and Beans (Goiânia, GO, Brazil) over the past ten years, following the AB-QTL strategy (Brondani et al., 2001, 2002; Rangel et al., 2005).

MATERIAL AND METHODS

Development of the introgression lines

ILs were developed from an interspecific cross between the inbred line BG90-2 (*Oryza sativa*), the recurrent parent, and the wild accession RS-16 (*Oryza glumaepatula*), the donor parent, collected in the Brazilian Amazon biome. Thirty-five BC_2F_8 lines were obtained using the AB-QTL methodology, as described by Rangel et al. (2005). These lines were selected from BC_2F_2 families and field evaluated on the occasion of a QTL analysis (Brondani et al., 2002). These families were then advanced on subsequent generations using the bulk methodology, where a sample of seeds from each family was mixed and sown together to originate the next generation.

Phenotypical evaluations

The 35 ILs were characterized for grain yield and grain quality traits (amylose content, gelatinization temperature and cohesiveness) in three locations (Goianira - State of Goiás; Formoso do Araguaia - State of Tocantins, and Boa Vista - State of Roraima), as described in Rangel et al. (2005). The joint analysis of variance of these experiments was used to select the 15 best performing ILs, which were evaluated in five locations (Goianira, Formoso do Araguaia, Boa Vista, Itajaí - State of Santa Catarina, and Alegrete - State of Rio Grande do Sul), following a complete randomized block design with four replications and using four high-yielding lines as controls (BG90-2, BRS Formoso, Metica 1, and IRGA 417). The 15 lines were evaluated for total yield, measured as the weight of grains from 10 random plants in each family. The analysis of variance was performed for each environment individually and for the five environments together (joint analysis) using the Genes software (Cruz, 1997). IL trait means were compared by the Scott and Knott test ($P < 0.05$), also performed by the Genes software.

Molecular marker assay

Fresh leaves of 10 representative plants of each one of the 35 ILs were collected and the DNA was extracted in bulk of plants, according to the protocol described by Ferreira and Grattapaglia (1998). Microsatellite (simple sequence repeat) markers used in the molecular characterization were selected based on their distribution throughout the 12 rice chromosomes according to the *O. glumaepatula* x *O. sativa* linkage maps (Brondani et al., 2001; Rangel et al., 2005) and the Cornell *O. sativa* ssp *indica* (IR64) x *O. sativa* ssp *japonica* (Azucena) simple sequence repeat reference map (Singh et al., 1996; Temnykh et al., 2000, 2001; Cheng et al., 2001) available at the Gramene web site (http://www.gramene.org). A total of 92 polymorphic microsatellite markers were selected, from which 69 were fluorescent-labeled with either hexachloro-6-carboxyfluorescein or 6-carboxyfluorescein, and the other 23 markers were not labeled (Supplementary Table 1). The amplification reactions were carried out in a final volume of 15 μL containing 15 ng of total genomic DNA, 0.3 μM of each primer, 0.25 mM of each dNTP, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mg/mL bovine serum albumin and 1.0 unit Taq DNA polymerase. The amplification reactions were conducted on a GeneAmp PCR System 9700 (Applied Biosystems) with a pre-cycle of 94°C for 5 min, followed by 40 cycles of 94°C for 1 min, 1 min at the annealing temperature of each primer and 72°C for 1 min. An extension step of 72°C for 7 min was used. Electrophoresis was conducted on an ABI 3100 automated DNA analyzer (Applied Biosystems) and allele sizing was performed using the software GeneMapper 2.5 (Applied Biosystems). PCR products derived from the non-labeled markers were visualized on 6% denaturing polyacrylamide gels stained with silver nitrate, as described by Bassam et al. (1991). The software CSSL Finder (http://www.mapdisto.free.fr/CSSLFinder.htm) was used to estimate the proportion of the parents' genome in each line and to construct the graphical genotypes.

The QTL analysis was conducted using the genotypic data from the 35 ILs obtained in this study and the phenotypic data obtained in field experiments conducted in Goianira, Boa Vista and Formoso do Araguaia, as described by Rangel et al. (2005). The QTL analysis was performed by the software QGene version 2.30 for MacIntosh (Nelson, 1997) using the single-marker regression method.

RESULTS

Phenotypic evaluation of introgression lines

In the season 2002/2003, the 35 ILs were evaluated in three field experiments for yield-related traits such as grain yield, tiller number and panicle number, and had their grains evaluated for milling and cooking quality traits (Rangel et al., 2005). For the average of the three locations, the most productive lines were CNAi 9930, CNAi 9931, CNAi 9934, CNAi 9935, CNAi 9936, and CNAi 9937. Among them, only CNAi 9930 showed long, thin and loose grains after cooking, characteristics of commercial value in Brazil and most parts of Latin America, and that were not present in the genitor BG90-2. The 15 best performing ILs were selected for further evaluation in five locations in the season 2003/2004.

Coefficients of variation of the experiments conducted in the season 2003/2004 ranged from 7% (Itajaí) to 14% (Alegrete) and were under the expected range for this kind of experiment. According to the joint analysis of variance obtained for the five locations, lines CNAi 9934, CNAi

9931, CNAi 9930, CNAi 9935, CNAi 9936, and CNAi 9937 showed the highest yield average and were statistically superior to the controls Metica 1 and IRGA 417 (Table 1). In the experiment conducted in Itajaí, lines CNAi 9931 and CNAi 9930 showed yield results that were statistically superior to the parent BG90-2. All lines, except CNAi 9934-85, CNAi 9924-92 and CNAi 9924-3, were statistically superior to the controls Metica 1 and IRGA 417. In Alegrete and Boa Vista, lines CNAi 9934, CNAi 9931, CNAi 9930, CNAi 9935, CNAi 9936, and CNAi 9937 showed yield averages that were statistically superior to the controls IRGA 417 and BRS Formoso, which are two of the most productive and extensively planted cultivars at each location, respectively (Table 1).

Table 1. Grain yield (kg/ha) obtained in the trials conducted in the season 2003/2004 in five locations: Goianira, State of Goiás; Itajaí, State of Santa Catarina; Formoso do Araguaia, State of Tocantins; Alegrete, State of Rio Grande do Sul, and Boa Vista, State of Roraima, Brazil. The number of wild fragments and proportions of homozygous and heterogeneous fragments in the high-yielding introgression lines are shown.

Lines	Proportion of homozygous fragments	Proportion of heterogeneous fragments	Total mean*	Goianira	Itajaí	Formoso do Araguaia	Alegrete	Boa Vista
CNAi 9934	1.09	3.26	8946^{a}	7000^{a}	7772^{b}	7548^{a}	12131^{a}	10277^{a}
CNAi 9931	1.09	2.17	8882^{a}	6836^{a}	8586^{a}	6747^{a}	11722^{a}	10517^{a}
BG90-2	-	-	8598^{a}	7941^{a}	7750^{b}	6380^{a}	10357^{a}	10561^{a}
CNAi 9930	-	4.35	8574^{a}	7586^{a}	8295^{a}	6371^{a}	10973^{a}	9644^{a}
CNAi 9935	-	5.43	8554^{a}	7278^{a}	7817^{b}	6919^{a}	10247^{a}	10511^{a}
BRS Formoso	-	-	8479^{a}	7495^{a}	8747^{a}	7867^{a}	9755^{a}	8533^{b}
CNAi 9936	2.17	5.43	8438^{a}	7383^{a}	7847^{b}	7077^{a}	10003^{a}	9879^{a}
CNAi 9937	1.09	6.52	8390^{a}	6461^{b}	7921^{b}	6558^{a}	9949^{a}	11064^{a}
CNAi 9924-117	-	14.43	7742^{b}	6464^{b}	7643^{b}	5359^{b}	10389^{a}	8856^{b}
CNAi 9933	-	7.61	7577^{b}	7109^{a}	7528^{b}	5820^{b}	8308^{b}	9121^{b}
CNAi 9920-82	2.17	9.78	7394^{b}	6020^{b}	6787^{c}	5523^{b}	9774^{a}	8869^{b}
CNAi 9932	-	2.17	7344^{b}	5556^{b}	6899^{c}	5820^{b}	9397^{a}	9048^{b}
CNAi 9924-105	2.17	10.87	7317^{b}	7158^{a}	6300^{c}	5165^{b}	9489^{a}	8475^{b}
CNAi 9920-88	1.09	4.85	6961^{c}	5625^{b}	6640^{c}	5751^{b}	8349^{b}	8442^{b}
CNAi 9924-85	-	4.35	6957^{c}	7352^{a}	5726^{d}	4580^{c}	8901^{b}	8225^{b}
CNAi 9924-92	2.17	5.43	6916^{c}	6158^{b}	5349^{d}	5017^{b}	9714^{a}	8343^{b}
Metica 1	-	-	6877^{c}	6714^{a}	5244^{d}	5855^{b}	6742^{b}	9829^{a}
IRGA 417	-	-	6872^{c}	5006^{b}	5125^{d}	6998^{a}	8170^{b}	9061^{b}
CNAi 9924-3	3.26	9.78	6287^{c}	6069^{b}	5925^{d}	3973^{c}	8124^{b}	7343^{b}
Mean			7750	6752	7146	6070	9667	9290
CV%			11	11	7	12	14	8

*Total mean was obtained by the analysis of variance of the five locations together (joint analysis). CV = coefficient of variation. Superscribed letters represent statistical differences between the means according to the Scott and Knott test ($P < 0.05$).

Molecular characterization of introgression lines

The 35 ILs were genotyped with 92 microsatellite markers distributed through the 12 rice chromosomes, with an average of 7 markers on each chromosome (Figure 1). A total of 30 wild alleles were detected among the 35 ILs on all chromosomes. The highest number of wild alleles was detected on chromosome 8 (10 alleles) and no homozygous wild allele was detected on chromosomes 5, 6, and 7. The average introgression proportion of homozygous wild alleles was 1.12% and ranged from 1.09 to 3.26%. Heterogeneous fragments ranged from 21.74 (line CNAi 9920-78) to 2.17% (lines CNAi 9930 and CNAi 9932), with an average of 8.18%.

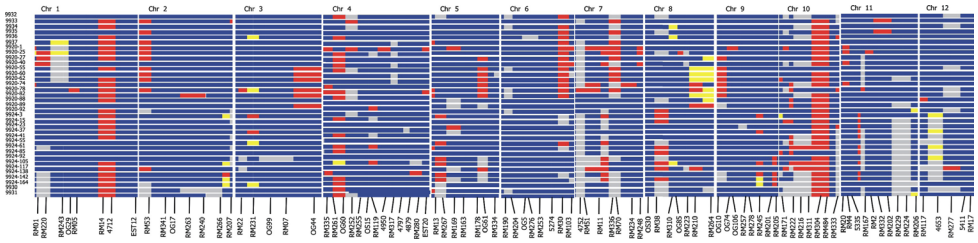

Figure 1. Graphical genotypes of the 35 introgression lines showing the 12 rice chromosomes. Blue squares represent the recurrent parent (*Oryza sativa*) proportion of the genome. Yellow squares represent the wild (*Oryza glumaepatula*) homozygous introgressions, red squares are the heterogeneous introgressions and gray squares represent missing data. Simple sequence repeat marker distributions on each chromosome are shown.

Lines CNAi 9934, CNAi 9931, CNAi 9930, CNAi 9935, CNAi 9936, and CNAi 9937 showed the best yield performance in the joint analysis of experiments and showed heterogeneous introgression proportions of 3.26, 2.17, 4.35, 5.43, 5.43, and 6.52%, respectively. Homozygous wild introgression proportions were 1.09, 1.09, 2.17, and 1.09% (Table 1).

In an attempt to identify wild alleles related to high yield, a single-regression QTL analysis was performed using the molecular data obtained in the present study and phenotypic data for the 35 ILs obtained by Rangel et al. (2005). The analysis was performed using phenotypic data for grain yield measured in three locations: Goianira, Boa Vista and Formoso do Araguaia. In Goianira, six markers were associated with grain yield ($P < 0.01$) and explained from 21.58 (RM30) to 36.69% (RM1 and 5335) of the phenotypic variation (PV) of the trait. In Boa Vista, six markers were identified and the PV explained by each one ranged from 15.34 (RM264) to 28.32% (OG10). Two markers were associated with grain yield in Formoso do Araguaia and explained 21.59 (OG10) and 17.58% (RM310) of the PV. In all QTLs, the alleles from BG90-2 were responsible for the positive effects on the trait, except the allele from the RM310 marker, identified in Formoso do Araguaia (Table 2).

Table 2. Quantitative trait loci detected for trait grain yield in 35 rice introgression lines with wild genomic fragments. Quantitative trait loci were detected under the minimum threshold of $P < 0.01$ in three locations: Goianira (State of Goiás), Boa Vista (State of Roraima) and Formoso do Araguaia (State of Tocantins), Brazil.

Location	Marker	Chromosome	Source	PV(%)	P
Goianira	RM1	1	BG90-2	36.68	0.0001
	5335	11	BG90-2	36.68	0.0001
	RM248	7	BG90-2	30.90	0.0005
	RM220	1	BG90-2	30.43	0.0019
	RM103	6	BG90-2	23.42	0.0043
	RM30	6	BG90-2	21.58	0.0085
Boa Vista	OG10	9	BG90-2	28.32	0.0017
	RM178	5	BG90-2	23.41	0.0067
	OG44	3	BG90-2	17.46	0.0125
	RM210	8	BG90-2	16.43	0.0157
	RM267	5	BG90-2	18.01	0.0173
	RM264	8	BG90-2	15.34	0.022
Formoso do Araguaia	OG10	9	BG90-2	21.59	0.0074
	RM310	8	RS-16	17.58	0.0122

PV is the phenotypic variation explained by each marker.

QTLs detected in this study were compared to the ones detected for the BC_2F_2 population (Brondani et al., 2002) that derived the 35 ILs (BC_2F_8). This analysis revealed that the RM1 marker (chromosome 1) was associated with grain yield measured in Goianira for BC_2F_2 (PV% = 39.11) and BC_2F_8 (PV% = 36.68). In both analyses, the allele that was responsible for the positive effect came from the cultivated genitor BG90-2.

DISCUSSION

The agronomic and molecular characterizations of 35 ILs derived from the interspecific cross *O. glumaepatula* x *O. sativa* revealed that CNAi 9934, CNAi 9931, CNAi 9930, CNAi 9935, CNAi 9936, and CNAi 9937 were the most productive lines and that CNAi 9931 and CNAi 9930 were statistically superior to the parent BG90-2 in the trial conducted in Itajaí. These lines showed heterogeneous introgression proportions ranging from 2.17 to 6.42% and homozygous wild introgressions ranging from 1.09 to 2.17%. In fact, most of the wild introgressions were represented by heterogeneous fragments with an average proportion of 8.18 against 1.12% of homozygous introgressions. Since the lines were advanced in bulk in each selfing generation and the DNA was extracted in bulk from 10 plants, the presence of heterogeneous loci was expected, due to the possibility of occurrence, on each locus, of a mixture of heterozygous plants and/or homozygous plants for alleles from *O. glumaepatula* and *O. sativa*.

Lines with higher grain yield showed low introgression proportions. One of the main reasons may be the lower linkage drag in these lines, reducing the possibility of the presence of wild alleles with deleterious effects in genes related to traits of agronomic importance. Tian et al. (2006a) developed ILs containing introgressions from the wild *O. rufipogon* and also observed that the high-yielding ILs had the lowest number of introgressed fragments. The backcrosses performed as part of the AB-QTL methodology played an important role in reducing the linkage drag because they allowed a progressive breakage of wild fragments in each cross and the recovery of the cultivated genetic background. The results observed for the high-yielding lines CNAi 9934, CNAi 9931, CNAi 9930, CNAi 9935, CNAi 9936, and CNAi 9937 showed that the methodology was efficient in introducing favorable wild alleles while maintaining the best features of BG90-2.

The QTL analysis performed for the 35 ILs showed that markers were associated with grain yield in the three sites of evaluation. The comparison of QTLs detected in the ILs (BC_2F_8) and in the BC_2F_2 families in Goianira revealed that the RM1 marker was associated, on both generations, with grain yield. The proportions of phenotypic variation explained by this marker was high in BC_2F_2 and BC_2F_8 generations (39.11 and 36.68, respectively), indicating that this marker was strongly associated with rice grain yield. The genomic region in the vicinity of the RM1 locus could be targeted to QTL fine mapping, in order to discover the gene responsible for the favorable expression of the trait. Since this marker was never identified in a QTL analysis involving intraspecific crosses, it can be implied that there are genomic regions from *O. glumaepatula* that could act positively with BG90-2 alleles to increase yield. There is also an opportunity to search for wild genomic fragments that could be involved as a trans-acting element to increase the effect of the BG90-2 allele at the RM1 locus.

Agronomic and molecular characterizations of the ILs allowed a better knowledge of their genomic composition and performance in the field. This strategy is being used routinely in the development of lines and cultivars originated from broad crosses in rice, mainly those involving interspecific crosses with *O. glumaepatula*, which is today an important source of

genetic variability for the breeding program at Embrapa for traits such as yield and drought tolerance. A considerable number of rice ILs have been developed in recent years (Wan et al., 2004; Ebitani et al., 2005; Tian et al., 2006b; Wang et al., 2007). The power of QTL detection is higher in ILs than in primary mapping populations such as F_2 or recombinant inbred lines, because, in ILs, it is possible to compare phenotypic effects between alleles on the substituted segments (Ebitani et al., 2005). Therefore, this type of population could be used to detect and confirm QTLs for important agronomic traits. Wan et al. (2004) used 66 ILs and detected six QTLs for grain quality traits that behaved as non-environment specific and that, according to the authors, could be used for marker-assisted selection. The molecular characterization of the 35 *O. glumaepatula*-derived ILs will allow a rapid identification of wild fragments facilitating the selection of ILs containing homozygous introgressions in regions of interest. Molecular markers allow the identification of introgressed fragments to the early selection of plants with smaller fragments, contributing to a faster and less expensive process. In addition, the effect of the incorporation of small fragments on the traits of interest can be measured, and the favorable alleles, either from the cultivated or the wild parent, can be detected.

Lines CNAi 9934, CNAi 9931, CNAi 9930, CNAi 9935, CNAi 9936, and CNAi 9937 showed high yield performance in five replicated experiments. These high-yielding lines have different fragments of *O. glumaepatula*, combined in different individuals, which confer genetic variability that is useful for low-input agriculture. Among them, CNAi 9930 is ready to be released to small farmers, due to grain quality traits and high production in ratooning, which increases the total yield by 30% (Rangel et al., 2005). The best performing ILs are available for rice breeders as a source of new allelic variation for the development of high-yielding cultivars. The generation of new ILs derived from *O. glumaepatula* x *O. sativa* crosses is under way, to continuously offer genetic materials of broadened genetic basis for rice breeding programs.

REFERENCES

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

Aluko G, Martinez C, Tohme J, Castano C, et al. (2004). QTL mapping of grain quality traits from the interspecific cross *Oryza sativa* x *O. glaberrima*. *Theor. Appl. Genet.* 109: 630-639.

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

Bernacchi D, Beck-Bunn T, Eshed Y, Inai S, et al. (1997). Advanced backcross QTL analysis in tomato. I. Identification of QTL for traits of agronomic importance from *Lycopersicon hirsutum*. *Theor. Appl. Genet.* 97: 381-397.

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

Brondani C, Rangel N, Brondani V and Ferreira E (2002). QTL mapping and introgression of yield-related traits from *Oryza glumaepatula* to cultivated rice (*Oryza sativa*) using microsatellite markers. *Theor. Appl. Genet.* 104: 1192-1203.

Brondani RP, Zucchi MI, Brondani C, Nakano Rangel PH, et al. (2005). Genetic structure of wild rice *Oryza glumaepatula* populations in three Brazilian biomes using microsatellite markers. *Genetica* 125: 115-123.

Cheng Z, Presting GG, Buell CR, Wing RA, et al. (2001). High-resolution pachytene chromosome mapping of bacterial artificial chromosomes anchored by genetic markers reveals the centromere location and the distribution of genetic recombination along chromosome 10 of rice. *Genetics* 157: 1749-1757.

Cruz CD (1997). Programa Genes: aplicativo computacional em genética e estatística. UFV, Viçosa.

Ebitani T, Takeuchi Y, Nonoue Y, Yamamoto T, et al. (2005). Construction and evaluation of chromosome segment substitution lines carrying overlapping chromosome segments of *indica* rice cultivar Kasalath in a genetic background of *japonica* elite cultivar Kohihikari. *Breeding Sci*. 55: 65-73.

Ferreira ME and Grattapaglia D (1998). Introdução ao uso de marcadores moleculares em análise genética. Embrapa Cenargen, Brasília.

Frary A, Fulton TM, Zamir D and Tanksley SD (2004). Advanced backcross QTL analysis of a *Lycopersicon esculentum* x *L. pennellii* cross and identification of possible orthologs in the Solanaceae. *Theor. Appl. Genet.* 108: 485-496.

Gowda M, Venu RC, Roopalakshmi K, Sreerekha MV, et al. (2003). Advances in rice breeding, genetics and genomics. *Mol. Breeding* 11: 337-352.

Gramene: A Resource for Comparative Grass Genomics. http://www.gramene.org/markers/index.html. Accessed April 10, 2005.

Gur A and Zamir D (2004). Unused natural variation can lift yield barriers in plant breeding. *PLOS. Biol.* 2: e245.

Li J, Xiao J, Grandillo S, Jiang L, et al. (2004). QTL detection for rice grain quality traits using an interspecific backcross population derived from cultivated Asian (*O. sativa* L.) and African (*O. glaberrima* S.) rice. *Genome* 47: 697-704.

Moncada P, Martínez CP, Borrero J, Chatel M, et al. (2001). Quantitative trait loci for yield and yield components in an *Oryza sativa* x *Oryza rufipogon* BC_2F_2 population evaluated in an upland environment. *Theor. Appl. Genet*. 102: 41-52.

Nelson JC (1997). QGENE: software for marker-based genomic analysis and breeding. *Mol. Breed* 3: 239-245.

Nguyen BD, Brar DS, Bui BC, Nguyen TV, et al. (2003). Identification and mapping of the QTL for aluminum tolerance introgressed from the new source, *Oryza rufipogon* Griff., into *indica* rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 106: 583-593.

Rangel PHN and Neves PC (1997). Selección recurrente aplicada al arroz de riego en Brasil. In: Selección recurrente en arroz (Guimarães EP, ed.). CIAT - Centro de Investigación Agrícola Tropical, Cali, 79-97.

Rangel PHN, Brondani C, Rangel PN, Brondani RPV, et al. (2005). Development of rice lines with gene introgression from the wild *Oryza glumaepatula* by the AB-QTL methodology. *CBAB* 5: 10-21.

Rangel PN, Brondani RPV, Coelho ASG, Rangel PHN, et al. (2007). Comparative linkage mapping of *Oryza glumaepatula* and *O. sativa* interspecific crosses based on microsatellite and expressed sequence tag markers. *Genet. Mol. Biol*. 30: 614-622.

Septiningsih EM, Prasetiyono J, Lubis E, Tai TH, et al. (2003). Identification of quantitative trait loci for yield and yield components in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. *Theor. Appl. Genet.* 107: 1419-1432.

Singh K, Ishii T, Parco A, Huang N, et al. (1996). Centromere mapping and orientation of the molecular linkage map of rice (*Oryza sativa* L.). *Proc. Natl. Acad. Sci. U. S. A.* 93: 6163-6168.

Tanksley SD and Nelson JC (1996). Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet*. 92: 191-203.

Tanksley SD and McCouch SR (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277: 1063-1066.

Temnykh S, Park WD, Ayres N, Cartinhour S, et al. (2000). Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theor. Appl. Genet*. 100: 697-712.

Temnykh S, DeClerck G, Lukashova A, Lipovich L, et al. (2001). Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. *Genome Res.* 11: 1441-1452.

Tian F, Li DJ, Fu Q, Zhu ZF, et al. (2006a). Construction of introgression lines carrying wild rice (*Oryza rufipogon* Griff.) segments in cultivated rice (*Oryza sativa* L.) background and characterization of introgressed segments associated with yield-related traits. *Theor. Appl. Genet.* 112: 570-580.

Tian F, Zhu Z, Zhang B, Tan L, et al. (2006b). Fine mapping of a quantitative trait locus for grain number per panicle from wild rice (*Oryza rufipogon* Griff.). *Theor. Appl. Genet.* 113: 619-629.

Wan XY, Wan JM, Su CC, Wang CM, et al. (2004). QTL detection for eating quality of cooked rice in a population of chromosome segment substitution lines. *Theor. Appl. Genet.* 110: 71-79.

Vaughan DA, Morishima H and Kadowaki K (2003). Diversity in the *Oryza* genus. *Curr. Opin. Plant Biol.* 6: 139-146.

Wang J, Wan X, Li H, Pfeiffer WH, et al. (2007). Application of identified QTL-marker associations in rice quality improvement through a design-breeding approach. *Theor. Appl. Genet.* 115: 87-100.

Supplementary Table 1. Microsatellite markers used for the molecular characterization of the 35 introgression lines.

Marker	Chromosome	Position	Forward sequence	Reverse sequence	Annealing temperature (°C)	Fluorescence	Reference
RM1	1	5.5	gcgaaaacacaatgcaaaaa	gcgttggttggacctgac	56	Fam	www.gramene.org
RM220	1	12.2	ggaaggtaactgtttccaac	gaaatgcttcccacatgtct	56	Fam	www.gramene.org
RM243	1	76.1	gatctgcagactgcagttgc	agctgcaacgatgttgtcc	56	Fam	www.gramene.org
OG29	1	101.8	gaccagttcaccatgcag	gagtgaggcagcaagaca	56	Fam	Brondani et al., 2001
RM5	1	132.6	tgcaacttctagctgctcga	gcatccgatcttgatggg	56	Fam	www.gramene.org
RM14	1	194.1	ccgaggagaggagttcgac	gtgccaatttcctcgaaaaa	56	Hex	www.gramene.org
4712	1	219.0	acaggctcgtgaatggta	ctcatcttcgccttcttg	48	-	Rangel et al., 2007
EST12	1	261.6	ccgcctcgagaacatgtgct	gctgtgccaattcaccgacg	48	-	Rangel et al., 2007
RM53	2	42.5	acgtctcgacgcatcaatgg	cacaagaacttcctcggtac	56	Fam	www.gramene.org
RM341	2	82.7	caagaaacctcaatccgagc	ctcctcccgatcccaatc	56	Hex	www.gramene.org
OG17	2	88.0	catgcatcaacaacgatc	gtgctcaagttagctgctc	56	Fam	Brondani et al., 2001
RM263	2	137.3	cccaggctagctcatgaacc	gctacgtttgagctaccacg	56	Hex	www.gramene.org
RM240	2	175.8	ccttaatgggtagtgtgcac	tgtaaccattccttccatcc	56	Hex	www.gramene.org
RM266	2	192.2	tagtttaaccaagactctc	ggttgaacccaaatctgca	56	Fam	www.gramene.org
RM207	2	202.0	ccattcgtgagaagatctga	cacctcatcctcgtaacgcc	56	Fam	www.gramene.org
RM22	3	46.6	ggtttgggagcccataatct	ctgggcttctttcactcgtc	56	Hex	www.gramene.org
RM231	3	57.0	ccagattatttcctgaggtc	cacttgcatagttctgcattg	56	Fam	www.gramene.org
OG99	3	79.3	gtggaagcacaagaacaaga	tcgtcatgcttcagcact	56	Fam	Brondani et al., 2001
RM7	3	103.2	ttcgccatgaagtctctcg	cctcccatcatttcgttgtt	56	Fam	www.gramene.org
OG44	3	208.0	acaccagctcagctcatc	tgtccaggtagtacaagctc	56	Hex	Brondani et al., 2001
RM335	4	0.2	gtacacacccacatcgagaag	gctctatgcgagtatccatgg	56	Fam	www.gramene.org
RM261	4	4.3	ctacttctccccttgtgtcg	tgtaccatcgccaaatctcc	56	Fam	www.gramene.org
OG60	4	28.7	acagctccactcctcacact	attgggtcacattgcagg	56	Hex	Brondani et al., 2001
RM252	4	45.0	gaatggcaatggcgctag	atgcggttcaagattcgatc	56	Fam	www.gramene.org
RM255	4	70.4	tgttgcgtgtggagatgtg	cgaaaccgctcagttcaac	56	Hex	www.gramene.org
OS15	4	75.0			56	-	Akagi et al., 1996
RM119	4	76.1	catccccctgctgctgctgctg	cgccggatgtgtgggactagcg	56	Hex	www.gramene.org
4950	4	106.3	cggaagaaggccatcgaggt	tgctcgtggtggtggtgttg	52	-	Rangel et al., 2007
RM317	4	118.3	catacttaccagttcaccgcc	ctggagagtgtcagctagttga	56	Hex	www.gramene.org
4797	4	130.4	ggagaaggcaatgcaacacg	gccattgccgccaagtacta	52	-	Rangel et al., 2007
4879	4	136.9	cagagatcgattggtagc	ccttgtactcagctccat	52	Hex	Rangel et al., 2007
RM280	4	152.3	acacgatccactttgcgc	tgtgtcttgagcagccagg	56	Hex	www.gramene.org
EST20	4	177.0	gtacgactattgcgccga	ttcacactccattctttaaatct	48	-	Rangel et al., 2007
RM13	5	15.4	tccaacatggcaagagagag	ggtggcattcgattccag	56	Hex	www.gramene.org
RM267	5	28.6	tgcagacatagagaaggaagtg	agcaacagcacaacttgatg	56	Hex	www.gramene.org

Continued on next page

Supplementary Table 1. Continued.

Marker	Chromosome	Position	Forward sequence	Reverse sequence	Annealing temperature (°C)	Fluorescence	Reference
RM169	5	57.9	tggctggctccgtgggtagctg	tcccgttgccgttcatccctcc	56	Fam	www.gramene.org
RM163	5	82.4	atccatgtgcgcctttatgagga	cgctacctccttcacttactagt	56	-	www.gramene.org
RM178	5	118.8	tcgcgtgaaagataagcggcgc	gatcaccgttccctccgcctgc	56	Hex	www.gramene.org
OG61	5	123.6	gcatgctgatgactgaagg	gaaacgaacggatggaca	56	Fam	Brondani et al., 2001
RM334	5	141.8	gttcagtgttcagtgccacc	gactttgatctttggtggacg	56	Fam	www.gramene.org
RM190	6	7.4	ctttgtctatctcaagacac	ttgcagatgttcttcctgatg	56	Fam	www.gramene.org
RM204	6	16.4	gtgactgacttggtcataggg	gctagccatgctctcgtacc	56	Hex	www.gramene.org
OG5	6	39.3	ttcatactttcatagaaaccg	tccaatgtgtcttgtctaatag	56	-	Brondani et al., 2001
RM276	6	40.3	ctcaacgttgacacctcgtg	tcctccatcgagcagtatca	56	Hex	www.gramene.org
RM253	6	51.3	tccttcaagagtgcaaaacc	gcattgtcatgtcgaagcc	56	Hex	www.gramene.org
5274	6	101.2	ctcgacctccatggcgaact	attgcagacgctcggagagg	48	-	Rangel et al., 2007
RM30	6	125.4	ggttaggcatcgtcacgg	tcacctcaccacacgacacg	56	Hex	www.gramene.org
RM103	6	143.7	cttccaattcaggccggctggc	cgccacagctgaccatgcatgc	56	Fam	www.gramene.org
4752	7	0.1	tcctgaccatctcaacctgc	gccggagagatgatcgagta	50	-	Rangel et al., 2007
RM51	7	1.0	tctcgattcaatgtcctcgg	ctacgtcatcatcgtcttccc	56	Fam	www.gramene.org
RM11	7	33.7	tctcctcttccccgatc	atagcgggcgaggcttag	56	Fam	www.gramene.org
RM336	7	61.0	cttacagagaaacggcatcg	gctggtttgtttcaggttcg	56	Hex	www.gramene.org
RM70	7	64.6	gtggacttcatttcaactcg	gatgtataagatagtccc	56	Fam	www.gramene.org
RM234	7	104.8	acagtatccaaggccctgg	cacgtgagacaaagacggag	56	Hex	www.gramene.org
RM248	7	126.5	tccttgtgaaatctggtccc	gtagcctagcatggtgcatg	56	Hex	www.gramene.org
OS30	8	0.1			56	-	Akagi et al., 1996
RM38	8	33.1	acgagctctcgatcagccta	tcggtctccatgtcccac	56	Hex	www.gramene.org
RM310	8	57.0	ccaaaacatttaaaatatcatg	gcttgttggtcattaccattc	56	Hex	www.gramene.org
OG85	8	63.9	ctttcttgtaataggggg	agactcacgagaacagat	50	-	Brondani et al., 2001
RM223	8	77.1	gagtgagcttgggctgaaac	gaaggcaagtcttggcactg	56	Hex	www.gramene.org
RM210	8	86.9	tcacattcggtggcattg	cgaggatggttgttcacttg	56	Fam	www.gramene.org
RM264	8	128.6	gttgcgtcctactgctacttc	gatccgtgtcgatgattagc	56	Fam	www.gramene.org
OG10	9	1.1	tggtgaatacaatctaccaat	tgattttatttttgtgctaaag	52	Hex	Brondani et al., 2001
OG74	9	37.3	ttgccatcacttagccacagtc	gcgtaaatgcccggagg	56	-	Brondani et al., 2001
OG106	9	43.8	ggccgtgtcaccatcttctcta	ggggatctgacatggcatatga	56	Fam	Brondani et al., 2001
RM257	9	70.9	cagttccgagcaagagtactc	ggatcggacgtggcatatg	56	Fam	www.gramene.org
RM278	9	77.5	gtagtgagcctaacaataatc	tcaactcagcatctctgtcc	56	Hex	www.gramene.org
RM245	9	96.0	atgccgccagtgaatagc	ctgagaatccaattatctgggg	56	-	www.gramene.org
RM201	9	111.2	ctcgtttattacctacagtacc	ctacctcctttctagaccgata	56	-	www.gramene.org
RM205	9	114.7	ctggttctgtatgggagcag	ctggcccttcacgtttcagtg	56	Hex	www.gramene.org
RM171	10	0.1	aacgcgaggacacgtacttac	acgagatacgtacgcctttg	56	Fam	www.gramene.org

Continued on next page

Supplementary Table 1. Continued.

Marker	Chromosome	Position	Forward sequence	Reverse sequence	Annealing temperature (°C)	Fluorescence	Reference
RM222	10	11.3	cttaaatgggccacatgcg	caaagcttcggccaaaag	56	Fam	www.gramene.org
RM216	10	17.6	gcatggccgatggtaaag	tgtataaaaccacacggcca	56	Hex	www.gramene.org
RM311	10	25.2	tggtagtataggtactaaacat	tcctatacacatacaaacatac	56	Hex	www.gramene.org
RM304	10	73.0	tcaaaccggcacatataagac	gataggagctgaaggagatg	56	Hex	www.gramene.org
RM484	10	79.0	tctccctcctcacattgtc	tgctgccctctctctctc	56	Hex	www.gramene.org
RM333	10	91.7	gtacgactacgagtgtcaccaa	gtcttcgcgatcactcgc	56	Hex	www.gramene.org
RM20	11	0.1	atcttgtccctgcaggtcat	gaaacagaggcacatttcattg	56	-	www.gramene.org
RM4	11	1.0	ttgacgaggtcagcactgac	aggctgtatccgactcatcg	56	-	www.gramene.org
5335	11	2.0	ttacggcgctaggcaagag	gtcgagtggagcacaaggaa	52	-	Rangel et al., 2007
RM167	11	20.4	gatccagcgtgaggaacacgt	agtccgaccacaaggtgcgttgtc	56	-	www.gramene.org
RM2	11	23.9	acgtgtcacgcgcttcctc	atgtccgggatctcatcg	56	-	www.gramene.org
RM332	11	27.9	gcgaaggcgaaggtgaag	catgagtgatctcactcaccc	56	Hex	www.gramene.org
RM202	11	34.8	cagattggagatgaagtcctcc	ccagcaagcatgtcaatgta	56	-	www.gramene.org
RM229	11	51.1	cactcacacgaacgactgac	cgcaggttcttgtgaaatgt	56	Hex	www.gramene.org
RM224	11	83.4	atcgatcgatcttcacgagg	tgctataaaaggcattcggg	56	Hex	www.gramene.org
RM206	11	102.9	ccatgcgtttaactattct	cgttccatcgatccgtatgg	56	Fam	www.gramene.org
RM117	12	0.2	cgatccattccgctgctcgcg	cgcccccatgcatgagaagacg	56	Fam	www.gramene.org
4653	12	32.3	ctcggacaagcatgatct	aaccgatgcagatcagag	52	Fam	Rangel et al., 2007
RM277	12	57.2	cggtcaaatcatcacctgac	caaggcttgcaagggaag	56	Fam	www.gramene.org
5411	12	102.0	ggtattgtcggtgttcagg	gtgaaggctgtaccatcca	52	-	Rangel et al., 2007
RM17	12	105.3	tgccctgttatttttcttctctc	ggtgatcctttcccatttca	56	-	www.gramene.org

Fam (6-carboxyfluorescein) and Hex (hexachloro-6-carboxyfluorescein) are fluorescent-labeled markers.