

Análise de associação quanto à produtividade e seus caracteres componentes em linhagens e cultivares de arroz de terras altas

Clistiane dos Anjos Mendes⁽¹⁾, Tereza Cristina de Oliveira Borba⁽²⁾, Luíce Gomes Bueno⁽³⁾, Gustavo Alencastro Veiga Cruzeiro⁽⁴⁾, João Antônio Mendonça⁽²⁾, Gabriel Feresin Pantalhão⁽²⁾, Rosana Pereira Vianello⁽²⁾ e Claudio Brondani⁽²⁾

⁽¹⁾Universidade Federal de Goiás, Escola de Agronomia, Departamento de Genética e Melhoramento de Plantas, Rodovia Goiânia/Nova Veneza, Km 0, CEP 74690-900 Goiânia, GO, Brasil. E-mail: clisagroma@hotmail.com ⁽²⁾Embrapa Arroz e Feijão, Rodovia GO-462, Km 12, Zona Rural, CEP 75375-000 Santo Antônio de Goiás, GO, Brasil. E-mail: tereza.borba@embrapa.br, joao.mendonca@embrapa.br, gabrielferesin@hotmail.com, rosana.vianello@embrapa.br, claudio.brondani@embrapa.br ⁽³⁾Embrapa Caprinos e Ovinos, Estrada Sobral/Groairas, Km 4, CEP 62010-970 Sobral, CE, Brasil. E-mail: luice.bueno@embrapa.br ⁽⁴⁾Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto, Laboratório de Oncologia do Hospital das Clínicas, Avenida dos Bandeirantes, nº 3.900, Campus Universitário, Monte Alegre, CEP 14048-900 Ribeirão Preto, SP, Brasil. E-mail: gavcruzeiro@usp.br

Resumo – O objetivo deste trabalho foi identificar, por meio da análise de mapeamento associativo, os marcadores moleculares relacionados à produtividade do arroz de terras altas e aos seus caracteres componentes. Foram usadas 113 linhagens e cultivares de arroz de terras altas, da Coleção Nuclear de Arroz da Embrapa, com reduzido vínculo genético entre si. Os seguintes caracteres componentes da produtividade foram avaliados: número de panículas por metro, número de grãos por panícula e peso de 100 grãos. Dos 115 marcadores utilizados, 25 (21,7%) associaram-se significativamente a um ou mais caracteres. Entre os 29 SSR ("simple sequence repeats") colocalizados em QTL ("quantitative trait loci") de produtividade de arroz, 12 foram associados aos caracteres avaliados e considerados como candidatos para uso na seleção assistida por marcadores. Os marcadores NP914540, Q6ZGD1 e Q69JE3, associados ao número de grãos por panícula, ainda não foram anotados no arroz e podem constituir o ponto de partida para estudos de genômica funcional. Entre os marcadores derivados de sequências transcritas, NP914526 e NP914533 destacam-se por pertencer a rotas metabólicas relacionadas ao aumento do potencial produtivo de arroz.

Termos para indexação: *Oryza sativa*, coleção nuclear, desequilíbrio de ligação, genômica funcional, mapeamento associativo, seleção assistida por marcadores.

Association analysis for yield and its component traits in lines and cultivars of upland rice

Abstract – The objective of this work was to identify, through analysis of associative mapping, the molecular markers related to upland rice yield and its component traits. One hundred thirteen lines and cultivars of upland rice, with reduced admixture, from the Rice Core Collection of Embrapa, were used. The following yield component traits were evaluated: number of panicles per meter, number of grains per panicle, and weight of 100 grains. Out of the 115 used markers, 25 (21.7%) were significantly associated with one or more traits. Among the 29 SSR (simple sequence repeats) co-located in yield QTL (quantitative trait loci) in rice, 12 were associated with the evaluated traits and considered candidates for use in marker-assisted selection. The markers NP914540, Q6ZGD1, and Q69JE3, associated with the number of grains per panicle, are not yet annotated in rice and should constitute the starting point for functional genomics studies. Among the markers derived from transcribed sequences, NP914526 and NP914533 stand out for belonging to metabolic pathways related to increased yield potential of rice.

Index terms: *Oryza sativa*, core collection, linkage disequilibrium, functional genomics, association mapping, marker-assisted selection.

Introdução

O arroz (*Oryza sativa* L.) constitui importante fonte de carboidratos, sobretudo para populações de baixa renda (Bassinello & Naves, 2006). O seu cultivo no

Brasil se dá por meio de sistema irrigado (várzeas) e de terras altas. A maioria das variedades de arroz de várzea pertence à subespécie índica, enquanto a maioria das variedades de terras altas é da subespécie japônica (Khush, 1997). O sistema de cultivo irrigado por

inundação controlada está consolidado no Sul do país; ele compreende 40% da área destinada à orizicultura no Brasil e contribui com cerca de 60% da produção anual de arroz.

No sistema de cultivo em sequeiro, o arroz de terras altas está sujeito à redução da produtividade em consequência de veranicos ocasionais. Essa modalidade de cultivo concentra-se principalmente na região Centro-Oeste, nos estados de Mato Grosso e Goiás, em áreas onde predominam Latossolos, que apresentam baixa fertilidade natural e boas características físicas (Heinemann & Stone, 2009). Apesar da menor produtividade do arroz de terras altas, sua área de cultivo tem grande potencial de expansão no Brasil. No entanto, em razão de estresses abióticos diversos e da irregularidade na distribuição das precipitações pluviais, a produção desse sistema de cultivo apresenta grande variação no país (Pinheiro, 2006).

O aumento do patamar de produtividade do arroz de terras altas é um permanente desafio dos programas de melhoramento genético. A partir da década de 1980, quando culminou o processo de substituição de variedades tradicionais de arroz (ciclo longo, porte elevado e grão tradicional longo e espesso) por cultivares consideradas modernas (ciclo médio a curto, porte baixo e grão agulhinha longo e fino), teve início um processo de declínio de ganho genético, nos programas de melhoramento do arroz, atribuído ao estreitamento da base genética da espécie (Castro et al., 1999). Vinte anos depois, foram iniciadas no Brasil atividades de pré-melhoramento, com o objetivo de ampliar a base genética de genótipos-elite de arroz, pelo acesso à variabilidade genética útil presente nas variedades tradicionais e linhagens oriundas de outros países (Fonseca et al., 2006). Como parte desse esforço, a Embrapa elaborou sua Coleção Nuclear de Arroz, que permitiu, pela primeira vez, estabelecer (Abadie et al., 2005) e caracterizar fenotípica (Bueno et al., 2012) e molecularmente (Borba et al., 2009) uma coleção de ampla base genética e com tamanho reduzido, para a cultura.

De acordo com dados recentes, o tamanho do genoma do arroz é de aproximadamente 370 Mpb e tem 55.986 genes, com função predita, que incluem 39.045 locos de não ETs (elementos transponíveis) que codificam 49.066 modelos gênicos, e 16.941 locos de ETs que codificam 17.272 modelos gênicos (Kawahara et al., 2013). Foram reportados, até o momento,

23 genes relacionados a caracteres componentes da produtividade (Tripathi et al., 2012). O conhecimento da base molecular desses caracteres é importante para que se possa manipular a produtividade do arroz em programas de melhoramento, via piramidação de genes e obtenção de arroz geneticamente modificado.

O mapeamento associativo, também conhecido por mapeamento por desequilíbrio de ligação, é um método de mapeamento de QTLs que utiliza o histórico de desequilíbrio de ligação de uma espécie para relacionar fenótipos a genótipos (Yu et al., 2008). A análise de associação tem potencial de identificar polimorfismos dentro de genes responsáveis por variações fenotípicas, e infere de modo prático o fenótipo a partir do genótipo. Assim, é possível identificar variações funcionais específicas, ligadas a diferenças fenotípicas relacionadas a caracteres de interesse (Flint-Garcia et al., 2003). No entanto, a interpretação das associações encontradas no mapeamento associativo e a identificação de genes relacionados a determinado caráter requerem a incorporação da análise de estrutura populacional (Huang et al., 2012).

Em arroz, o mapeamento associativo já foi utilizado para estudo da produtividade de grãos (Huang et al., 2012), e tem-se beneficiado do número massivo de marcadores SNPs ("single nucleotide polymorphisms"), derivados de sequenciamentos de baixa resolução resultantes da análise GBS ("genotyping by sequencing" (Elshire et al., 2011)) em plantas. A partir de um conjunto de dados de 3,6 milhões de SNP, obtidos do sequenciamento de 517 variedades tradicionais de arroz, Huang et al. (2010) encontraram, via mapeamento associativo, 37 marcadores SNPs associados a 12 caracteres de interesse agrônomico.

Em estudos de genética associativa, alternativamente a essa abordagem com grande densidade de marcadores, pode-se utilizar a abordagem de genes candidatos, uma vez que o mapeamento de associação fica restrito a locos relevantes, possivelmente envolvidos no controle de caracteres de interesse (Ehrenreich et al., 2009). Essa abordagem é inédita para a cultura do arroz, mas já foi previamente utilizada para a cultura do milho (Harjes et al., 2008), Pinus (González-Martínez et al., 2007) e batata (Malosetti et al., 2007).

O objetivo deste trabalho foi identificar, por meio da análise de mapeamento associativo, os marcadores moleculares relacionados à produtividade do arroz de terras altas e aos seus caracteres componentes.

Material e Métodos

A partir do conjunto de 133 linhagens e cultivares de arroz de terras altas, da Coleção Nuclear de Arroz

da Embrapa (CNAE), foram selecionados 113 acessos que não possuíam ancestralidade múltipla, dos quais 94 são linhagens e 19 são cultivares comerciais (Tabela 1).

Tabela 1. Identificação dos acessos da Coleção Nuclear de Arroz da Embrapa (CNAE), utilizados na análise de associação.

N	Código	Origem ⁽¹⁾	Nome comum	N	Código	Origem	Nome comum
1	CNA0000482	EUA	Bluebonnet	58	CNA0005972	Nigéria	Farox 301
2	CNA0000963	Ipeaco	Amarelão X Guedes	59	CNA0005975	Colômbia	IDS A 6
3	CNA0000969	Ipeaco	Honduras X Matão	60	CNA0006030	Filipinas	Padi Senemok
4	CNA0000976	Ipeaco	Saturno X Pratão Precoce	61	CNA0006034	Nigéria	ITA 150
5	CNA0000994	Ipeaco	Esav X Matão	62	CNA0006035	Nigéria	ITA 225
6	CNA0001006	EUA	DAWN	63	CNA0006174	Embrapa	LS 85-158
7	CNA0001118	EEPG	EEPG-1-269-Furnas	64	CNA0006187	Embrapa	Caiapó
8	CNA0001347	Ipeaco	Ipeaco-SL 2270	65	CNA0006217	IAC	Iapar 9
9	CNA0001350	Ipeaco	Ipeaco-SL 1970	66	CNA0006406	IAC	LS 86-68
10	CNA0002075	IAC	IAC 165	67	CNA0006413	Iapar	L 85-20
11	CNA0002123	Iapar	Japonês X Praiana	68	CNA0006422	Iapar	Iapar L 99-98
12	CNA0002524	França	Moroberekan	69	CNA0006572	França	IREM 195 (ECAD)
13	CNA0003005	Índia	Sona	70	CNA0006574	França	IRAT 112 (ECAD)
14	CNA0003281	Embrapa	IRAT 177	71	CNA0006701	Embrapa	CNAX 1762-A-41-2-1
15	CNA0003287	França	IREM 123-2-1	72	CNA0006940	EUA	Lemont
16	CNA0003288	França	IREM 293-B	73	CNA0006941	EUA	New Bonnet
17	CNA0003289	França	IREM 247	74	CNA0007119	Embrapa	CNARR 2888-B-47
18	CNA0003362	França	IRAT 142	75	CNA0007425	Embrapa	Canastra
19	CNA0003372	França	IRAT 10	76	CNA0007706	Embrapa	CNARR 2888-B-12-1-1
20	CNA0003375	França	IRAT 13	77	CNA0007799	Embrapa	IAC 1191
21	CNA0003397	França	IRAT 144	78	CNA0007937	Embrapa	Progresso
22	CNA0003403	Nigéria	TOX 490-3-108-D1-B-B	79	CNA0008070	Embrapa	CNAX 3608-6-1-2-1
23	CNA0003490	França	Mearin	80	CNA0008092	França	L 141
24	CNA0004078	Embrapa	UPR 82-1-7	81	CNA0008093	França	L 285
25	CNA0004098	Embrapa	CNA 511-12-B-5	82	CNA0008172	Embrapa	Bonanca
26	CNA0004120	Embrapa	CNA 092-BM10-BM27p-3	83	CNA0008305	Embrapa	Carisma
27	CNA0004121	Embrapa	CNA 095-BM30-BM9-28	84	CNA0008309	Embrapa	L 92-342
28	CNA0004141	Embrapa	CNA 511-12-B-2	85	CNA0008412	EUA	Bluebonnet 50
29	CNA0004168	Embrapa	L 80-68	86	CNA0008432	EUA	Lacassine
30	CNA0004172	Embrapa	CNA 511-12-B-4	87	CNA0008533	Embrapa	Maravilha
31	CNA0004193	França	IREM 656	88	CNA0008540	Embrapa	Talento
32	CNA0004206	Embrapa	Araguaia	89	CNA0008545	Colômbia	CT 11216-10-12-B-BRM-10
33	CNA0004243	Filipinas	OR 63-252	90	CNA0008711	Embrapa	CNAX 4914-36-M-M-3
34	CNA0004428	França	N.7384 [RPL X Daniela]	91	CNA0009113	Colômbia	CT10037-9-4-M-1-1P-2-M
35	CNA0004480	França	IRAT 124	92	CNA0009115	Colômbia	CT11632-3-3-M
36	CNA0004487	França	Makouta	93	CNA0009123	Colômbia	CT11891-3-3-3-M
37	CNA0004543	Nigéria	TOX 1012-12-3-1	94	CNA0009124	Colômbia	CT13364-7-1
38	CNA0004617	Nigéria	TOX 1011-4-2	95	CNA0009197	Colômbia	CT13377-8-4
39	CNA0004640	Nigéria	TOX 1785-19-18	96	CNA0009199	Colômbia	CT13381-1-3
40	CNA0004697	Colômbia	N.2583	97	CNA0009223	Colômbia	CT13569-5-7
41	CNA0004748	Embrapa	CNA 104-B-2-43-2	98	CNA0009227	Colômbia	CT13570-3-2
42	CNA0004759	Nigéria	TOX 514-16-101-1	99	CNA0009240	Colômbia	CT13572-6-2
43	CNA0004788	Nigéria	TOX 503-4-115-B-B	100	CNA0009364	Colômbia	CT13581-5-2
44	CNA0004796	Nigéria	TOX 516-28-10B-D2-B-B	101	CNA0009415	Colômbia	CT13582-11-4
45	CNA0005180	Embrapa	CNA 539-23-B-4	102	CNA0009561	Colômbia	CT13584-12-9
46	CNA0005287	França	IRAT 162	103	CNA0009591	Colômbia	CT13585-12-3
47	CNA0005334	Nigéria	TOX 1871-29	104	CNA0010451	China	YUNLU N 7
48	CNA0005342	IRAT	Rio Verde	105	CNA0010469	Filipinas	TB47H-MR-11-51-3
49	CNA0005461	Índia	CO 18	106	CNA0010476	Filipinas	B8503E-TB-19-B-3
50	CNA0005462	Índia	MTU 15	107	CNA0010503	Filipinas	YN1905-UUL-62
51	CNA0005465	Índia	W 1253	108	CNA0010527	Filipinas	BR4742-B-19-23
52	CNA0005650	Ipeaco	Ipeaco 11P	109	CNA0010533	Filipinas	1R65907-188-1-B
53	CNA0005660	Ipeaco	Ipeaco 77P	110	CNA0010618	IAC	IAC 201
54	CNA0005672	Embrapa	L 82-192	111	CNA0010671	Iapar	Iapar 62
55	CNA0005673	Embrapa	IAC 81-176	112	CNA0010679	Iapar	Iapar 63
56	CNA0005901	Embrapa	IAC 84-198	113	CNA0010682	IAC	IAC 202
57	CNA0005970	Nigéria	Farox 299	-	-	-	-

⁽¹⁾Embrapa, Empresa Brasileira de Pesquisa Agropecuária; IAC, Instituto Agronômico de Campinas; Iapar, Instituto Agronômico do Paraná; Ipeaco, Instituto Pesquisa Agropecuária do Centro Oeste.

Os 113 acessos selecionados foram avaliados em experimento de campo, na safra 2007/2008, no sistema de cultivo de terras altas. O experimento foi realizado na estação experimental da Embrapa Arroz e Feijão, no Município de Santo Antônio de Goiás (GO). Utilizou-se o delineamento de blocos incompletos, com quatro linhas de 5 m, em área útil compreendida pelas duas linhas centrais, no total de 4,5 m². O espaçamento entre as linhas foi de 0,35 metros, e utilizou-se irrigação suplementar por pivô central. As cultivares BR/Irga 409, Metica 1 e Caiapó foram utilizadas como testemunhas.

Os seguintes caracteres quantitativos foram avaliados: número de panículas por metro (NPM), número de grãos por panícula (NGP), peso de 100 grãos (PCG) e produtividade (kg ha⁻¹). A análise estatística dos dados fenotípicos foi realizada com o programa SAS, versão 6, por meio do procedimento Proc GLM (SAS Institute, Cary, NC, EUA), tendo-se utilizado os dados médios da parcela para cada tratamento (genótipo). As médias ajustadas foram utilizadas para compor os dados fenotípicos da análise de associação.

Inicialmente, foram identificadas regiões do genoma do arroz previamente relacionadas à produtividade e aos componentes desta, oriundas de 27 mapas de QTL já publicados, disponibilizados no Gramene (2014). A partir da localização de cada marcador molecular relacionado a um QTL (marcador âncora) no genoma, foram identificados os transcritos situados a 200 kpb "upstream" e 200 kpb "downstream". A sequência de cada transcrito foi analisada, tendo-se selecionado os que apresentaram DNA repetitivo. A partir das sequências dos transcritos selecionados, foram desenhados pares de iniciadores com o programa Primer 3 (Untergasser et al., 2012). Esses marcadores foram analisados em conjunto com as informações de 86 marcadores SSR obtidas por Borba et al. (2010).

A extração de DNA foi realizada pelo protocolo CTAB modificado, descrito por Ferreira & Grattapaglia (1998). A reação em cadeia da polimerase (PCR) foi realizada para todos os marcadores SSR avaliados, em volume final de 15 µL, com 5 µL de DNA 3 ng µL⁻¹, 4,3 µL de cada iniciador (0,9 µmol L⁻¹), 1,5 µL de tampão 10 x, 1,3 µL de DMSO (50%), 0,2 µL de Taq DNA polimerase (5 unidades µL⁻¹), 1,3 µL de dNTP (2,5 mol L⁻¹) e 1,4 µL de água Milli-Q autoclavada. As reações foram conduzidas em termociclador PTC 100 (MJ Research, Quebec, Canadá). Os passos da amplificação foram: 94°C por 5 min, seguido por

40 ciclos de 94°C por 1 min, temperatura de anelamento específica de cada par de iniciador por 1 min e 72°C por 1 min, além de uma extensão final a 72°C por 7 min. As reações foram submetidas à eletroforese, em géis de poliacrilamida a 6% desnaturante, corados com solução de nitrato de prata, de acordo com protocolo descrito por Creste et al. (2001).

O número de alelos, os alelos privados (ou exclusivos) e a estimativa da informação genética de cada marcador, avaliada pelo PIC ("polymorphism information content"), foram obtidos com o programa PowerMarker, versão 3.23 (Liu & Muse, 2005).

A análise de associação entre marcadores e caracteres fenotípicos foi conduzida com uso do modelo linear misto (MLM), disponível no programa Tassel versão 2.1 (Bradbury et al., 2007). O modelo misto utilizado considerou, como efeito fixo, os marcadores moleculares e os dados de estruturação (matriz Q) obtidos pelo programa Structure versão 2.3.4 (Pritchard et al., 2000), enquanto a matriz de parentesco ("kinship"), obtida pelo programa SPAGeDi versão 1.2 (Hardy & Vekemans, 2002), foi considerada como efeito aleatório (Borba et al., 2010). Os alelos com frequência menor ou igual a 0,05 foram considerados como alelos nulos para a análise de associação, de acordo com estratégia estabelecida por Breseghello & Sorrells (2006).

O método de correção adotado foi o FDR ("false discovery rate"), e a probabilidade (p-valor) ajustada foi utilizada como o novo nível de significância, para cada associação identificada.

Resultados e Discussão

Todas as variáveis mensuradas (produtividade, PCG, NPM e NGP) apresentaram efeito significativo (Tabela 2). Os valores dos coeficientes de variação (CV) estiveram entre 3,98%, para o peso de 100 grãos, e 25,16%, para o número de grãos por panícula, e podem ser considerados normais (Balbinot Junior et al., 2003).

A avaliação individual dos componentes de produtividade visou diminuir a complexidade da análise, uma vez que esses caracteres têm controle genético quantitativo, e alguns apresentam baixa herdabilidade (Sürek & Beşer, 2005). A produtividade correlacionou-se positivamente ao NGP (0,42; p<0,01), semelhantemente ao observado por Cho et al. (2007), que estudaram 164 RILs (linhagens endogâmicas

recombinantes) de arroz, derivadas de cruzamento biparental. Os caracteres NPM e PCG apresentaram correlação negativa (-0,21, $p < 0,01$), o que corrobora a observação de Peng et al. (1999) de que o aumento da produtividade está negativamente relacionado ao aumento do número de panículas.

Sessenta e sete marcadores derivados de sequências transcritas com DNA repetitivo foram desenhados para a genotipagem dos 113 acessos de arroz. Desse total, 44 marcadores (63,7%) apresentaram padrão de amplificação satisfatório, com distribuição nos cromossomos 1, 2, 4, 5, 7, 8, 9, 10 e 11 (Tabela 3). Vinte e nove marcadores foram polimórficos e produziram 160 alelos, com média de 3,63 alelos por loco, a qual variou de três (marcadores Q6YUX0, Q5VRY4, Q5VRY1, NP914533, Q69JE3 e NP914540) a dez alelos (Q75HV7). O PIC médio foi 0,37, tendo variado de 0,15 (NP914524) a 0,84 (Q75HV7). Observaram-se 18 alelos privados, dos quais três provieram do acesso CNA0004121. Borba et al. (2010), com uso de 86 marcadores SSR, encontraram média de 4,76 alelos por loco e um PIC médio de 0,56, para 133 variedades de arroz de terras altas. A diferença desses resultados com os do presente trabalho é consequência do menor polimorfismo dos marcadores derivados de transcritos com DNA repetitivo, em comparação a SSR derivados de sequências do genoma estrutural. Apesar de os marcadores derivados de sequências gênicas serem frequentemente menos polimórficos do que os genômicos, eles apresentam uma série de vantagens, como o excelente índice de transferibilidade entre espécies, alto potencial de associação a genes que afetam caracteres de interesse e facilidade de acesso

via bancos públicos de sequências (Varshney et al., 2005; Castillo et al., 2008; Ramu et al., 2009). Além disso, outra grande vantagem de se utilizar marcadores derivados do genoma transcrito é a de que a variabilidade encontrada entre dois genótipos pode estar associada a diferenças no fenótipo, o que resulta em marcadores passíveis de utilização direta na seleção assistida.

Foram utilizados 115 marcadores na análise de associação: 86 SSR genômicos e 29 SSR de sequências transcritas. Alelos raros, com frequência $\leq 5\%$, foram eliminados da análise, para evitar problemas nas estimativas do desequilíbrio de ligação (Remington et al., 2001). Esses alelos tendem a aumentar o erro tipo I e a indicar a existência de subestruturação na população (Flint-Garcia et al., 2003). Dos 115 marcadores, 25 (21,74%) foram associados significativamente a um ou mais caracteres avaliados. Entre os 29 marcadores de sequências transcritas, 12 (41,38%) foram associados significativamente a um ou mais caracteres avaliados, enquanto apenas 13 (15,12%) dos 86 SSR do genoma estrutural foram associados significativamente a um ou mais caracteres (Tabela 4). O fato de estarem localizados em regiões previamente associadas a QTL indica que as sequências transcritas constituem marcadores com possibilidade real de estarem relacionados aos caracteres fenotípicos avaliados. O emprego potencial desses marcadores, em programas de melhoramento, vai desde a possibilidade de desenvolver marcadores para a seleção assistida até o desenvolvimento de plantas geneticamente modificadas que superexpressem o transcrito identificado.

A determinação da função dos marcadores derivados do transcriptoma foi realizada com análise de Blast para o arroz (Kawahara et al., 2013) e *Arabidopsis* (Lamesch et al., 2011), e permitiu inferir de forma direta a função putativa desses transcritos, o que é uma vantagem interessante em relação a marcadores SSR do genoma estrutural, que podem perder sua ligação física aos genes de interesse pela recombinação.

Para o caráter NGP, foram observadas 17 associações significativas (Tabela 4), localizadas nos cromossomos 1, 2, 6, 7, 9, 10 e 12. No cromossomo 1, o marcador NP914526 (LOC_Os01g09260, de acordo com a notação utilizada pelo Rice Genome Annotation Project) está relacionado à enzima citocinina desidrogenase (EC 1.5.99.12), que atua na degradação do hormônio citocinina, responsável pelo crescimento,

Tabela 2. Quadrados médios da análise de variância da produtividade e de seus caracteres componentes, em linhagens de arroz de terras altas⁽¹⁾.

Fonte de variação	GL	Produtividade (kg ha ⁻¹)	PCG (g)	NPM	NGP
Tratamento	100	4.292.406**	0,757**	670,89**	1.696,95**
Bloco	30	3.733.881**	0,394**	441,29	1.027,79**
Repetição	2	1.271.017*	0,002	842,91	121,90
Erro	-	414.879	0,010	312,56	414,50
Média	-	2.891,34	2,515	84,968	80,923
CV (%)	-	22,27	3,98	20,80	25,16

⁽¹⁾Delineamento de blocos incompletos ao acaso. * e **Significativo pelo teste F a 5 e 1% de probabilidade, respectivamente. GL, graus de liberdade; PCG, peso de 100 grãos; NPM, número de panículas por metro; e NGP, número de grãos por panícula.

diferenciação celular, dominância apical e senescência das folhas. Ashikari et al. (2005) relacionaram a redução de expressão dessa enzima ao aumento da produtividade em arroz, pelo acúmulo de citocinina no

meristema das inflorescências e aumento no número de órgãos reprodutivos. Nesse caso, a busca por alelos menos eficientes desse gene, em acessos do banco de germoplasma de arroz ou pela utilização de RNA

Tabela 3. Sequências de iniciadores de marcadores desenvolvidos a partir de sequências transcritas, provenientes de regiões adjacentes de QTL identificados previamente na literatura.

Marcador	Cromossomo	Marcador âncora	Iniciador forward	Iniciador reverse	Ta (°C)
NP914540	1	RM1	TCTTCGTCTCGCCTTTGTTT	GGAATCTCCGTTTCCTTGGT	50
Q5VRY1	1	RM1	ACCTTTCGGAACCTTCCACT	TCGATCGCCTCTTCTTCTC	50
Q5VRY4	1	RM1	GTGGTTAATCCCTTCGCTGA	CAAACCACGAGTTGTCCAAA	56
NP914524	1	RM1	TCTAGCGGCTGGAGTCTGAT	GAGTAGTCGTGCCACGTCAG	50
NP914526	1	RM1	GGACAAGTACGACCCCAAGA	AGATGCGAGGAGGAGTACA	56
NP914533	1	RM1	CTCCTCCCTCCCATTACCC	TCGAGGGGAGTACGTACCAG	50
NP917594	1	RM5	TTTTGTTCCGCCAATTAGA	TGCACAAGGTCAAAGAATGG	58
NP917595	1	RM5	ACAACAACAGCGCAGAGATG	ATTGACACCTCGTCCAAACC	50
NP917599	1	RM5	ATCTTGATGGGGTGCAGAAG	ACCCGCTGTTACAGGAGCTA	60
Q5ZE11	1	RM5	CGGAAAAAGGGTTTTTGGTT	CCCTCCTCCTCCTCCTCTTA	50
NP917589	1	RM5	CGAACTATCGCGGTCTGCTA	GATGATTGGGGATGACAGGT	52
NP917593	1	RM5	TTTTGTTCCGCCAATTAGA	TGCACAAGGTCAAAGAATGG	52
NP916369	1	RM212	TGAAACTGCCAATACAAACAGG	TGCAGTTCCATAGGACACA	56
Q5VZAV5	1	RM212	AGGAGGTGGAGGAGGAATTG	TCTGCAGGAATACGGGAAAC	56
Q5ZCG1	1	RM212	TGATACCACTGCCAAGTCCA	ATTTGAGCCACGAATTCCAC	54
Q5VR71	1	RM220	AAGTCAACGGCGTCAAACAT	TGACTGTCCGGTACTGTCTGT	56
Q5VRY8	1	RM220	CATCACTCCCCATCACTCCT	CTGCCACAGTGACACGTTCT	50
Q6K700	2	RM208	GTGCACGTGGAACCTGTGTTG	GCGAGAAGAAGCGAGAAGAA	50
Q6K701	2	RM208	GTGCACGTGGAACCTGTGTTG	GCGAGAAGAAGCGAGAAGAA	50
Q6K704	2	RM208	AAAACCAGCAAGGACGAGGT	CTCCTGCAGAAGGTGCAGTT	56
Q6YUXO	2	RM240	ATCGAGCGAACCTTCAACAC	CCAAACACGCCTCTCTTCA	50
Q6YUX6	2	RM240	GGTCTCGATGTGCAACCTTC	TCTGCTGTTGGTACTGTGTC	50
Q6YZ50	2	RM250	TTCGAGTGCCACTACTGCTG	CGATGTGGTGGTGGTAGTTG	56
Q6YZ56	2	RM250	CGTCGCATCATCTTCATCAG	CCTGCCATTTCTGTTCAATTT	50
Q6Z7K9	2	RM250	CGATCACCTTGTGCGTGATA	CCTCGATAATTTCCGGTTTG	50
Q6Z7L6	2	RM250	GCATAAAAGGCCAACCAAAA	GATGGATCTGAGGAGGGATG	52
Q6ZGD1	2	RM250	CCCATCTGATCCAAGATTGC	ACGTTGAGCTGCCAATGTAA	54
KS1	4	RM255	TGGGACAACCTGACAAGAGA	CCAAACAAATGGCCTCAGAT	52
Q7XLD8	4	RM255	TGCATATGGAGCCTTGGATT	CTCCCAAACCACAACAACC	50
Q84NC2	4	RM255	TGGGACAACCTGACAAGAGA	CCAAACAAATGGCCTCAGAT	56
QX7PL3	4	RM255	AACGAAACCGCTCAGTTCAA	GGGACCAGGAGAAAAATTCG	52
Q6AU68	5	RM164	GGCACGAGGAACTGAAGAAC	CTGACAAACCTCTCGTGCAG	58
Q6OER7	5	RM164	AGCGGCTCCTCTCTCTCTCT	CCTATGCGAAGATCACGACA	50
Q6OES1	5	RM164	GGAGAGAAGCATCCAACCAA	TCCGTGTTTGTCCGTGTTTA	50
Q75HV7	5	RM164	CAGTGCTGAATTGTCAGTGA	CGTGGGAGCAACTGGTTATT	52
Q7EZ30	7	RM11	CAACCACCACCCTTCTTC	CAACCACCACCCTTCTTC	50
Q7XIR5	7	RM11	CCTCCCTCTCTCTTGTGA	GGACGCCACACAAATATTGC	50
Q8LIJO	7	RM11	CTCACCCACCCTGTTCATCT	GGAAGGTTTCGCTGACAAGAG	50
Q6Z473	7	RM11	TCATGTTACCCCTCCCTCT	ACACACACACACACACACAC	50
Q6ZCO3	8	RM210	GGAAGAGGCGCTGACAATC	GTCACACGGACCATGTATGC	50
Q69JE3	9	RM215	GTCGTGGACGACCAACTCCT	GAAGGGTCCCAGAAGTCC	56
Q69JZ8	9	RM215	GCTTCTGCAACCTGGTAAG	AAGTTCCCTCTCGTTGTCA	52
NP920800	10	RM311	GAGAAAGTACGTGCGGTCAA	ATGAAATTGCCACCTGTGTG	58
Q53KP1	11	RM231	ACCAACTTTAGCGCCCTAT	CTGGTCGAGGACCAACAGT	50

Ta, temperatura de anelamento.

interferente, poderia aumentar o potencial produtivo da cultura.

O marcador NP914533 (LOC_Os01g54030) tem como produto a enzima málica dependente de NADP (EC 1.1.1.40), que participa da fotossíntese (Zhao et al., 2007). A busca por variantes alélicas desse marcador, tanto em acessos de arroz cultivado quanto de arroz silvestre de genoma AA nativo do Brasil (*O. glumaepatula*), pode possibilitar a identificação de isoformas da enzima málica com maior eficiência da fotossíntese. Esse marcador pode ser o ponto de partida para o aumento da eficiência da taxa fotossintética de arroz, uma vez que esteve associado diretamente ao caráter NGP.

O marcador NP914540, cujo transcrito é expresso no arroz (LOC_Os01g09420), ainda não foi anotado; ou seja, ainda não possui função conhecida. Mesmo com a análise de Blast, no banco de dados de *Arabidopsis*, não foi possível encontrar sequências homólogas. Como a sequência representada pelo marcador esteve associada somente ao NGP, ela é interessante

para estudar unicamente este caráter. O NP917599 (LOC_Os01g42280) também foi exclusivo para o NGP e está relacionado ao gene de repetições com pentatricopeptídeo (PPR), que tem 35 aminoácidos repetidos em tandem e está associado ao sistema restaurador da fertilidade (CMS) do arroz (Hu et al., 2012), o qual é fundamental para viabilizar a produção de híbridos de arroz.

O marcador Q5VRY1 (LOC_Os01g08860) é associado a uma proteína de choque térmico, HSP18, (Chang et al., 2007). Essa classe de proteínas normalmente está envolvida na resposta a altas temperaturas e a outros estresses (Chandel et al., 2013).

Os marcadores Q6YZ56, Q6YUX0 e Q6ZGD1 foram identificados no cromossomo 2. O marcador Q6YZ56 (LOC_Os02g53590), em arroz, está relacionado ao produto gênico defensina. Esse produto codifica proteínas ricas em cisteína, também conhecidas como polipeptídeos antimicrobianos (Tsfaye et al., 2013). O homólogo desse gene em *Arabidopsis* é o AT2G22121, que foi relacionado

Tabela 4. Marcadores significativamente associados aos caracteres avaliados e seus respectivos valores de probabilidade (p).

Marcador	Cromossomo	Repetição	Código RGAP ⁽¹⁾	Âncora	Nº Alelos	p (NGP)	p (PCG)	p (Produtividade)
NP914526	1	(TTAA)6	LOC_Os01g09260	RM1	4	1,14x10 ⁻⁴	1,29x10 ⁻⁴	-
NP914533	1	(CCG)6	LOC_Os01g09320	RM1	2	2,79x10 ⁻⁴	1,71x10 ⁻⁴	-
NP914540	1	(GCC)9	LOC_Os01g09420	RM1	2	3,46x10 ⁻⁶	-	-
NP917599	1	(AG)15	LOC_Os01g66490	RM5	6	1,25x10 ⁻⁴	-	-
OG106	9	(AG)27	-	-	8	-	7,30x10 ⁻⁵	-
Q5VRY1	1	(AGA)10	LOC_Os01g08860	RM1	2	7,98x10 ⁻⁷	3,86x10 ⁻⁵	-
Q69JE3	9	(CCG)5	LOC_Os09g36760	RM215	2	6,29x10 ⁻⁵	7,24x10 ⁻⁶	-
Q69JZ8	9	(CT)5	LOC_Os09g36470	RM215	2	-	1,12x10 ⁻⁴	-
Q6YUXO	2	(CT)13	LOC_Os02g51320	RM240	2	7,91x10 ⁻⁷	3,94x10 ⁻⁵	-
Q6YZ56	2	(CT)27	LOC_Os02g53590	RM250	4	2,34x10 ⁻⁶	3,36x10 ⁻⁶	-
Q6ZGD1	2	(CT)14	LOC_Os02g53700	RM250	4	1,99x10 ⁻⁵	-	-
Q8LIJO	7	(CT)13	LOC_Os07g32510	RM11	3	3,16x10 ⁻⁴	-	5,85x10 ⁻⁵
RM12	12	(AG)21	-	-	2	2,07x10 ⁻⁴	-	-
RM117	12	(AG)7	-	-	6	2,31x10 ⁻⁴	1,33x10 ⁻⁴	-
RM125	7	(GCT)8	-	-	2	5,35x10 ⁻⁶	1,93x10 ⁻⁶	-
RM144	11	(ATT)11	-	-	6	-	3,16x10 ⁻⁵	-
RM190	6	(CT)11	-	-	6	2,26x10 ⁻⁴	-	-
RM205	9	(CT)25	-	-	5	-	7,97x10 ⁻⁵	-
RM222	10	(CT)18	-	-	5	9,69x10 ⁻⁵	-	-
RM234	7	(CT)25	-	-	4	-	1,12x10 ⁻⁴	-
RM240	2	(CT)21	-	-	3	1,74x10 ⁻⁴	5,81x10 ⁻⁵	-
RM252	4	(CT)19	-	-	4	-	1,52x10 ⁻⁴	-
RM263	2	(CT)34	-	-	6	-	4,95x10 ⁻⁵	-
RM277	12	(AG)11	-	-	4	-	3,11x10 ⁻⁴	-
RM484	10	(AT)9	-	-	3	1,07x10 ⁻⁴	6,63x10 ⁻⁵	-

⁽¹⁾RGAP, Rice Genome Annotation Project. NGP, número de grãos por panícula; e PCG, peso de 100 grãos.

tanto à imunidade natural da planta quanto à comunicação célula a célula nas inflorescências (Takeuchi & Higashiyama, 2012). A relação desse gene com a capacidade de polinização em arroz poderá ser averiguada, para avaliar sua influência no aumento da eficiência da polinização, que tem reflexo direto no número de grãos por panícula. Giacomelli et al. (2012) também identificaram, em *Vitis vinifera*, a indução da expressão da família gênica defensina, tanto em tecidos ligados à reprodução quanto em tecidos infectados por doença. O marcador Q6YUX0 (LOC_Os02g51320) está associado a um gene da família bHLH ("basic helix-loop-helix"), a mais numerosa família de fatores de transcrição. Essa família tem sido reportada como envolvida em diversos processos biológicos das plantas, como sinalização da presença luz e de hormônios, resposta a estresse hídrico e desenvolvimento de flores e frutos (Carretero-Paulet et al., 2010). A associação desse fator de transcrição com o caráter NGP, dada à multiplicidade de caracteres que podem ser afetados por sua expressão, pode estar relacionada à ação na regulação de genes envolvidos com a floração. O marcador Q6ZGD1 (LOC_Os02g53700) está relacionado a uma proteína com domínio DENN, a qual não tem função definida em arroz. Contudo, há relatos de seu envolvimento em rotas de sinalização de certos tipos de proteínas quinases (Levivier et al., 2001), e seu ortólogo em *Arabidopsis* (AT5G35560) foi identificado como expresso em diversos tecidos, entre os quais, os órgãos florais.

No cromossomo 7, foi identificado o marcador Q8LIJ0 (LOC_Os07g32510), cujo produto gênico pertence à família de fatores de transcrição *Dof*, que desempenham funções críticas como reguladores transcricionais no crescimento e desenvolvimento de plantas (Yanagisawa, 2004). Em arroz, estima-se que existam 30 genes *Dof* (Hernando-Amado et al., 2012). O ortólogo desse gene em *Arabidopsis* é o AT1G28310, relacionado à regulação do tempo de florescimento. Em cereais, sua transcrição está relacionada à maturação das sementes (Hernando-Amado et al., 2012).

No cromossomo 9, foi identificado o marcador Q69JE3 (LOC_Os09g36760), relacionado ao gene *DUF630*, sem função definida em arroz, mas identificado como expresso em bibliotecas de RNA de inflorescências, sementes e folhas – Rice Genome

Annotation Project (Kawahara et al., 2013). O ortólogo desse gene em *Arabidopsis* (AT3G51290) também não foi descrito funcionalmente, o que pode ser foco de pesquisa inédita para ambas as espécies.

Nenhum marcador esteve associado ao NPM. O aumento do número de marcadores utilizados na análise poderia identificar associação a esse caractere.

Entre os 18 marcadores significativamente associados ao caráter PCG, dez também estiveram associados ao NGP. Não houve correlação fenotípica entre esses caracteres, ao contrário do que foi observado por Cho et al. (2007), que relataram correlação negativa e significativa. Ranawake & Amarasinghe (2014) também relataram correlação negativa, mas não significativa quanto a esses caracteres. Estudos adicionais poderão validar o uso desses marcadores na seleção de indivíduos com desempenho favorável desses dois caracteres.

No cromossomo 1, os marcadores NP914526 e NP914533 estão ancorados pelo marcador SSR RM1, previamente relacionado ao caráter PCG (Yu et al., 1997; Cho et al., 2003), por meio do QTL *gw-1*. No cromossomo 2, o marcador Q6YUX0, também associado ao caráter NGP, foi ancorado pelo marcador RM240, previamente associado ao PCG por Zhuang et al. (2002), por meio do QTL *qTGWT-2-2* e, por Thomson et al. (2003), pelo QTL *gw2.1*. No cromossomo 4, o marcador RM252 havia sido previamente relacionado ao PCG, em uma população segregante do cruzamento interespecífico *O. sativa* x *O. glumaepatula* (Brondani et al., 2002). No cromossomo 7, o marcador RM125 foi previamente associado ao PCG, por meio do QTL *gw7.1* (Septiningsih et al., 2003). No cromossomo 9, o marcador RM205 foi previamente associado com PCG, por meio do QTL *qGW-9* (Cho et al., 2003). No cromossomo 10, o marcador RM484 foi previamente associado ao PCG, por meio do QTL *qTGW-10* (Hittalmani et al., 2002).

Entre os oito marcadores associados unicamente ao PCG, o Q69JZ8 (LOC_Os09g36470) é um retrotranspósom e não possui homologia de sequência com *Arabidopsis*. Retrotranspósom é a classe predominante de elementos transponíveis em plantas e são responsáveis pela grande diferença do tamanho dos genomas (Wicker & Keller, 2007). Existem evidências de que a ativação de retrotranspósoms é um dos fatores-chave para a adaptação a mudanças ambientais, uma vez que podem ser inseridos em sequências

regulatórias e, conseqüentemente, podem alterar a expressão de determinados genes (Todorovska, 2007).

Quanto ao caráter produtividade, verificou-se apenas uma associação significativa com o marcador Q8LIJO, que também está associado ao caráter NGP.

Os transcritos identificados no presente trabalho podem ser explorados em detalhe pela comunidade científica e, assim, contribuir para o aumento do conhecimento das bases moleculares do aumento de produtividade em *O. sativa*.

Conclusões

1. Entre os marcadores avaliados, três estão associados ao número de grãos por panícula (NP914540, Q6ZGD1 e Q69JE3) e, por ainda não haver anotação deles em arroz, podem constituir o ponto de partida para estudos de atribuição de função gênica.

2. Dois marcadores derivados de seqüências transcritas (NP914526 e NP914533) destacaram-se por pertencer a rotas metabólicas relacionadas ao aumento do potencial produtivo em arroz.

Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo apoio financeiro e bolsas de pesquisa.

Referências

- ABADIE, T.; CORDEIRO, C.M.T.; FONSECA, J.R.; ALVES, R. de B. das N.; BURLE, M.L.; BRONDANI, C.; RANGEL, P.H.N.; CASTRO, E. da M. de; SILVA, H.T. da; FREIRE, M.S.; ZIMMERMANN, F.J.P.; MAGALHÃES, J.R. Construção de uma coleção nuclear de arroz para o Brasil. **Pesquisa Agropecuária Brasileira**, v.40, p.129-136, 2005. DOI: 10.1590/S0100-204X2005000200005.
- ASHIKARI, M.; SAKAKIBARA, H.; LIN, S.Y.; YAMAMOTO, T.; TAKASHI, T.; NISHIMURA, A.; ANGELES, E.R.; QIAN, Q.; KITANO, H.; MATSUOKA, M. Cytokinin oxidase regulates rice grain production. **Science**, v.309, p.741-745, 2005. DOI: 10.1126/science.1113373.
- BALBINOT JUNIOR, A.A.; FLECK, N.G.; MENEZES, V.G.; AGOSTINETTO, D. Competitividade de cultivares de arroz irrigado com cultivar simuladora de arroz-vermelho. **Pesquisa Agropecuária Brasileira**, v.38, p.53-59, 2003. DOI: 10.1590/S0100-204X2003000100007.
- BASSINELLO, P.Z.; NAVES, M.M.V. Bioquímica e saúde humana. In: SANTOS, A.B. dos; STONE, L.F.; VIEIRA, N.R. de A. (Ed.). **A cultura do arroz no Brasil**. 2.ed. Santo Antônio de Goiás: Embrapa Arroz e Feijão, 2006. p.31-51.
- BORBA, T.C. de O.; BRONDANI, R.P.V.; BRESEGHELLO, F.; COELHO, A.S.G.; MENDONÇA, J.A.; RANGEL, P.H.N.; BRONDANI, C. Association mapping for yield and grain quality traits in rice (*Oryza sativa* L.). **Genetics and Molecular Biology**, v.33, p.515-524, 2010. DOI: 10.1590/S1415-47572010005000065.
- BORBA, T.C. de O.; BRONDANI, R.P.V.; RANGEL, P.H.N.; BRONDANI, C. Microsatellite marker-mediated analysis of the Embrapa Rice Core Collection genetic diversity. **Genetica**, v.137, p.293-304, 2009. DOI: 10.1007/s10709-009-9380-0.
- BRADBURY, P.J.; ZHANG, Z.; KROON, D.E.; CASSTEVEN, T.M.; RAMDOSS, Y.; BUCKLER, E.S. TASSEL: Software for association mapping of complex traits in diverse samples. **Bioinformatics**, v.23, p.2633-2635, 2007. DOI: 10.1093/bioinformatics/btm308.
- BRESEGHELLO, F.; SORRELLS, M.E. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. **Genetics**, v.172, p.1165-1177, 2006. DOI: 10.1534/genetics.105.044586.
- BRONDANI, C.; RANGEL, P.H.N.; BRONDANI, R.P.V.; FERREIRA, M.E. QTL mapping and introgression of yield-related traits from *Oryza glumaepatula* to cultivated rice (*Oryza sativa*) using microsatellite markers. **Theoretical and Applied Genetics**, v.104, p.1192-1203, 2002. DOI: 10.1007/s00122-002-0869-5.
- BUENO, L.G.; VIANELLO, R.P.; RANGEL, P.H.N.; UTUMI, M.M.; CORDEIRO, A.C.C.; PEREIRA, J.A.; FRANCO, D.F.; MOURA NETO, F.; MENDONÇA, J.A.; COELHO, A.S.G.; OLIVEIRA, J.P. de; BRONDANI, C. Adaptabilidade e estabilidade de acessos de uma coleção nuclear de arroz. **Pesquisa Agropecuária Brasileira**, v.47, p.216-226, 2012. DOI: 10.1590/S0100-204X2012000200010.
- CARRETERO-PAULET, L.; GALSTYAN, A.; ROIG-VILLANOVA, I.; MARÍNEZ-GARCÍA, BILBAO-CASTRO; ROBERTSON, D. Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in Arabidopsis, poplar, rice, moss, and algae. **Plant Physiology**, v.153, p.1398-1412, 2010. DOI: 10.1104/pp.110.153593.
- CASTILLO, A.; BUDAK, H.; VARSHNEY, R.; DORADO, G.; GRANER, A.; HERNANDEZ, P. Transferability and polymorphism of barley EST-SSR markers used for phylogenetic analysis in *Hordeum chilense*. **BMC Plant Biology**, v.8, e97, 2008. DOI: 10.1186/1471-2229-8-97.
- CASTRO, E.M.; BRESEGHELLO, F.; RANGEL, P.H.; MORAES, O.P. Melhoramento do arroz. In: BOREM, A. (Ed.). **Melhoramento de espécies cultivadas**. Viçosa: Ed. da UFV, 1999. p.95-130.
- CHANDEL, G.; DUBEY, M.; MEENA, R. Differential expression of heat shock proteins and heat stress transcription factor genes in rice exposed to different levels of heat stress. **Journal of Plant Biochemistry and Biotechnology**, v.22, p.277-285, 2013. DOI: 10.1007/s13562-012-0156-8.

- CHANG, P.-F.L.; JINN, T.-L.; HUANG, W.-K.; CHEN, Y.; CHANG, H.-M.; WANG, C.-W. Induction of a cDNA clone from rice encoding a class II small heat shock protein by heat stress, mechanical injury, and salicylic acid. **Plant Science**, v.172, p.64-75, 2007. DOI: 10.1016/j.plantsci.2006.07.017.
- CHO, Y.C.; SUH, J.P.; CHOI, I.S.; HONG, H.C.; BAEK, M.K.; KANG, K.H.; KIM, Y.G.; AHN, S.N.; CHOI, H.C.; HWANG, H.G.; MOON, H.P. QTLs analysis of yield and its related traits in wild rice relative *Oryza rufipogon*. **Treatment of Crop Research**, v.4, p.19-29, 2003.
- CHO, Y.-G.; KANG, H.-J.; LEE, J.-S.; LEE, Y.-T.; LIM, S.-J.; GAUCH, H.; EUN, M.-Y.; MCCOUCH, S.R. Identification of quantitative trait loci in rice for yield, yield components, and agronomic traits across years and locations. **Crop Science**, v.47, p.2403-2417, 2007. DOI: 10.2135/cropsci2006.08.0509.
- CRESTE, B.; TULMANN NETO, A.; FIGUEIRA, A. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. **Plant Molecular Biology Reporter**, v.19, p.299-306, 2001. DOI: 10.1007/BF02772828.
- EHRENREICH, I.M.; HANZAWA, Y.; CHOU, L.; ROE, J.L.; KOVER, P.X.; PURUGGANAN, M.D. Candidate gene association mapping of Arabidopsis flowering time. **Genetics**, v.183, p.325-335, 2009. DOI: 10.1534/genetics.109.105189.
- ELSHIRE, R.J.; GLAUBITZ, J.C.; SUN, Q.; POLAND, J.A.; KAWAMOTO, K. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. **PLoS ONE**, v.6, e19379, 2011. DOI: 10.1371/journal.pone.0019379.
- FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3.ed. Brasília: Embrapa-Cernagen, 1998. 220p. (Embrapa-Cernagen. Documentos, 20).
- FLINT-GARCIA, S.A.; THORNSBERRY, J.M.; BUCKLER, E.S. Structure of linkage disequilibrium in plants. **Annual Review on Plant Biology**, v.54, p.357-374, 2003. DOI: 10.1146/annurev.arplant.54.031902.134907.
- FONSECA, J.R.; BRONDANI, C.; BRONDANI, R.P.V.; RANGEL, P.H. Recursos genéticos do arroz. In: SANTOS, A.B. dos; STONE, L.F.; VIEIRA, N.R. de A. (Ed.). **A cultura do arroz no Brasil**. 2.ed. Santo Antônio de Goiás: Embrapa Arroz e Feijão, 2006. p.257-288.
- GIACOMELLI, L.; NANNI, V.; LENZI, L.; ZHUANG, J.; DALLA SERRA, M.; BANFIELD, M.J.; TOWN, C.D.; SIVERSTEIN, K.A.; BARALDI, E.; MOSER, C. Identification and characterization of the defending-like gene family of grapevine. **Molecular Plant-Microbe Interactions**, v.25, p.1118-1131, 2012. DOI: 10.1094/MPMI-12-11-0323.
- GONZÁLEZ-MARTÍNEZ, S.C.; WHEELER, N.C.; ERSOZ, E.; NELSON, C.D.; NEALE, D.B. Association genetics in *Pinus taeda* L.I. wood property traits. **Genetics**, v.175, p.399-409, 2007. DOI: 10.1534/genetics.106.061127.
- GRAMENE. **Gramene**: a comparative resource for plants. 2014. Available at: <<http://www.gramene.org/>>. Accessed on: 9 Sept. 2014
- HARDY, O.J.; VEKEMANS, X. SPAGeDi: a versatile computer program to analyze spatial genetic structure at the individual or population levels. **Molecular Ecology Notes**, v.2, p.618-620, 2002. DOI: 10.1046/j.1471-8286.2002.00305.x.
- HARJES, C.E.; ROCHEFORD, T.R.; BAI, L.; BRUTNELL, T.P.; KANDIANIS, C.B.; SOWINSKI, S.G.; STAPLETON, A.E.; VALLABHANENI, R.; WILLIAMS, M.; WURTZEL, E.T.; YAN, J.; BUCKLER, E.S. Natural genetic variation in *lycopen epsilon cyclase* tapped for maize biofortification. **Science**, v.319, p.330-333, 2008. DOI: 10.1126/science.1150255.
- HEINEMANN, A.B.; STONE, L.F. Efeito da deficiência hídrica no desenvolvimento e rendimento de quatro cultivares de arroz de terras altas. **Pesquisa Agropecuária Tropical**, v.39, p.134-139, 2009.
- HERNANDO-AMADO, S.; GONZÁLEZ-CALLE, V.; CARBONERO, P.; BARRERO-SICILIA, C. The family of DOF transcription factors in *Brachypodium distachyon*: phylogenetic comparison with rice and barley DOFs and expression profiling. **BMC Plant Biology**, v.12, p.202-215, 2012. DOI: 10.1186/1471-2229-12-202.
- HITTALMANI, S.; SHASHIDHAR, H.E.; BAGALI, P.G.; HUANG, N.; SIDHU, J.S.; SINGH, V.P.; KHUSH, G.S. Molecular mapping of quantitative trait loci for plant growth, yield and yield related traits across three diverse locations in a doubled haploid rice population. **Euphytica**, v.125, p.207-214, 2002. DOI: 10.1023/A:1015890125247.
- HU, J.; WANG, K.; HUANG, W.; LIU, G.; GAO, Y.; WANG, J. The rice pentatricopeptide repeat protein rf5 restores fertility in Hong-Lian cytoplasmic male-sterile lines via a complex with the glycine-rich protein GRP162. **Plant Cell**, v.24, p.109-122, 2012. DOI: 10.1105/tpc.111.093211.
- HUANG, X.; WEI, X.; SANG, T.; ZHAO, Q.; FENG, Q.; ZHAO, Y.; LI, X. Genome-wide association studies of 14 agronomic traits in rice landraces. **Nature Genetics**, v.42, p.961-967, 2010. DOI: 10.1038/ng.695.
- HUANG, X.; ZHAO, Y.; WEI, X.; LI, C.; WANG, A. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. **Nature Genetics**, v.44, p.32-40, 2012. DOI: 10.1038/ng.1018.
- KAWAHARA, Y.; BASTIDE, M. de la; HAMILTON, J.P.; KANAMORI, H.; MCCOMBIE, W.R.; OUYANG, S.; SCHWARTZ, D.C.; TANAKA, T.; WU, J.; ZHOU, S.; CHILDS, K.L.; DAVIDSON, R.M.; LIN, H.; QUESADA-OCAMPO, L.; VAILLANCOURT, B.; SAKAI, H.; LEE, S.S.; KIM, J.; NUMA, H.; ITOH, T.; BUELL, C.R.; MATSUMOTO, T. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. **Rice**, v.6, p.4-14, 2013. DOI: 10.1186/1939-8433-6-4.
- KHUSH, G.S. Origin, dispersal, cultivation and variation of rice. **Plant Molecular Biology**, v.35, p.25-34, 1997. DOI: 10.1023/A:1005810616885.
- LAMESCH, P.; BERARDINI, T.Z.; LI, D.; SWARBRECK, D.; WILKS, C.; SASIDHARAN, R.; MULLER, R.; DREHER, K.; ALEXANDER, D.L.; GARCIA-HERNANDEZ, M.; KARTHIKEYAN, A.S.; LEE, C.H.; NELSON, W.D.; PLOETZ, L.; SINGH, S.; WENSEL, A.; HUALA, E. The Arabidopsis

- Information Resource (TAIR): improved gene annotation and new tools. **Nucleic Acids Research**, v.40, p.D1202-D1210, 2011. DOI: 10.1093/nar/gkr1090.
- LEVIVIER, E.; GOUD, B.; SOUCHET, M.; CALMELS, T.P.G.; MORNON, J.-P.; CALLEBAUT, I. uDENN, DENN, and dDENN: indissociable domains in Rab and MAP kinases signaling pathways. **Biochemical and Biophysical Research Communications**, v.287, p.688-695, 2001. DOI: 10.1006/bbrc.2001.5652.
- LIU, K.; MUSE, S.V. PowerMarker: an integrated analysis environment for genetic marker analysis. **Bioinformatics**, v.21, p.2128-2129, 2005. DOI: 10.1093/bioinformatics/bti282.
- MALOSETTI, M.; VAN DER LINDEN, C.G.; VOSMAN, B.; VAN EEUWIJK, F.A. A mixed-model approach to association mapping using pedigree information with an illustration of resistance to *Phytophthora infestans* in potato. **Genetics**, v.175, p.879-889, 2007. DOI: 10.1534/genetics.105.054932.
- PENG, S.; CASSMAN, K.G.; VIRMANI, S.S.; SHEEHY, J.; KHUSH, G.S. Yield potential trends of tropical rice since the release of IR8 and the challenge of increasing rice yield potential. **Crop Science**, v.39, p.1552-1559, 1999. DOI: 10.2135/cropsci1999.3961552x.
- PINHEIRO, B. Características morfológicas da planta relacionadas à produtividade. In: SANTOS, A.B. dos; STONE, L.F.; VIEIRA, N.R. de A. (Ed.). **A cultura do arroz no Brasil**. 2.ed. rev. Santo Antônio de Goiás: Embrapa Arroz e Feijão, 2006. 1000p.
- PRITCHARD, J.K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v.155, p.945-959, 2000.
- RAMU, P.; KASSAHUN, B.; SENTHILVEL, S.; ASHOK, K.; UMAR, C.; JAYASHREE, B.; FOLKERTSMA, R.; REDDY, L.; KURUVINASHETTI, M.; HAUSMANN, B.; HASH, C. Exploiting rice-sorghum synteny for targeted development of EST-SSRs to enrich the sorghum genetic linkage map. **Theoretical and Applied Genetics**, v.119, p.1193-1204, 2009. DOI: 10.1007/s00122-009-1120-4.
- RANAWAKE, A.L.; AMARASINGHE, U.G.S. Relationship of yield and yield related traits of some traditional rice cultivars in Sri Lanka as described by correlation analysis. **Journal of Scientific Research and Reports**, v.3, p.2395-2403, 2014. DOI: 10.9734/JSRR/2014/12050.
- REMYNGTON, D.L.; THORNSBERRY, J.M.; MATSUOKA, Y.; WILSON, L.M.; WHITT, S.R.; DOEBLEY, J.; KRESOVICH, S.; GOODMAN, M.M.; BUCKLER, E.S. Structure of linkage disequilibrium and phenotypic associations in the maize genome. **Proceedings of the National Academy of Sciences of the United States of America**, v.98, p.11479-11484, 2001. DOI: 10.1073/pnas.201394398.
- SEPTININGSIH, E.M.; PRASETIYONO, J.; LUBIS, E.; TAI, T.H.; TJUBARYAT, T.; MOELJOPAWIRO, S.; MCCOUCH, S.R. 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*. **Theoretical and applied genetics**, v.107, p.1419-1432, 2003. DOI: 10.1007/s00122-003-1376-z.
- SÜREK, H.; BEŞER, N. Selection for grain yield & yield components in early generations for temperate rice. **Philippine Journal of Crop Science**, v.28, p.3-15, 2005.
- TAKEUCHI, H.; HIGASHIYAMA, T. A species-specific cluster of defensin-like genes encodes diffusible pollen tube attractants in *Arabidopsis*. **PLoS Biology**, v.10, e1001449, 2012. DOI: 10.1371/journal.pbio.1001449.
- TESFAYE, M.; SILVERSTEIN, K.A.T.; NALLU, S.; WANG, L.; BOTANGA, C.J.; Gomez, S.K.; COSTA, L.M.; HARRISON, M.J.; SAMAC, D.A.; GLAZEBROOK, J.; KATAGIRI, F.; GUTIERREZ-MARCOS, J.F.; VANDENBOSCH, K.A. Spatio-temporal expression patterns of *Arabidopsis thaliana* and *Medicago truncatula* defensin-like genes. **PLoS ONE**, v.8, e58992, 2013. DOI: 10.1371/journal.pone.0058992.
- THOMSON, M.J.; TAI, T.H.; MCCLUNG, A.M.; LAI, X.H.; HINGA, M.E.; LOBOS, K.B.; XU, Y.; MARTINEZ, C.P.; MCCOUCH, S.R. Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. **Theoretical and applied genetics**, v.107, p.479-493, 2003. DOI: 10.1007/s00122-003-1270-8.
- TODOROVSKA, E. Retrotransposons and their role in plant-genome evolution. **Biotechnology and Biotechnology Equipment**, v.21, p.294-305, 2007. DOI: 10.1080/13102818.2007.10817464.
- TRIPATHI, A.K.; PAREEK, A.; SOPORY, S.K.; SINGLA-PAREEK, S.L. Narrowing down the targets for yield improvement in rice under normal and abiotic stress conditions via expression profiling of yield-related genes. **Rice**, v.5, p.37-49, 2012. DOI: 10.1186/1939-8433-5-37.
- VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **Trends in Biotechnology**, v.23, p.48-55, 2005. DOI: 10.1016/j.tibtech.2004.11.005.
- UNTERGASSER, A.; CUTCUTACHE, I.; KORESSAAR, T.; YE, J.; FAIRCLOTH, B.C.; REMM, M.; ROZEN, S.G. Primer3 - new capabilities and interfaces. **Nucleic Acids Research**, v.40, e115, 2012. DOI: 10.1093/nar/gks596.
- WICKER, T.; KELLER, B. Genome-wide comparative analysis of *copia* retrotransposons in Triticeae, rice, and *Arabidopsis* reveals conserved ancient evolutionary lineages and distinct dynamics of individual *copia* families. **Genome Research**, v.17, p.1072-1081, 2007. DOI: 10.1101/gr.6214107.
- YANAGISAWA, S. Dof domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants. **Plant Cell Physiology**, v.45, p.386-391, 2004. DOI: 10.1093/pcp/pch055.
- YU, J.; HOLLAND, J.B.; MCMULLEN, M.D.; BUCKLER, E.S. Genetic design and statistical power of nested association mapping in maize. **Genetics**, v.178, p.539-551, 2008. DOI: 10.1534/genetics.107.074245.
- YU, S.B.; LI, J.X.; XU, C.G.; TAN, Y.F.; GAO, Y.J.; LI, X.H.; ZHANG, Q.; MAROOF, M.A. Importance of epistasis as the

genetic basis of heterosis in an elite rice hybrid. **Proceedings of the National Academy of Sciences of the United States of America**, v.94, p.9226-9231, 1997. DOI: 10.1073/pnas.94.17.9226.

ZHAO, C.; ZHAO, B.; REN, Y.; TONG, W.; WANG, J.; ZHAO, K.; SHU, S.; XU, N.; LIU, S. Seeking transformation markers: an analysis of differential tissue proteomes on the rice germplasm generated from transformation of *Echinochloa crusgalli* genomic

DNA. **Journal of Proteome Research**, v.6, p.1354-1363, 2007. DOI: 10.1021/pr0605015.

ZHUANG, J.Y.; FAN, Y.Y.; RAO, Z.M.; WU, J.L.; XIA, Y.W.; ZHENG, K.L. Analysis on additive effects and additive-by-additive epistatic effects of QTLs for yield traits in a recombinant inbred line population of rice. **Theoretical and Applied Genetics**, v.105, p.1137-1145, 2002. DOI: 10.1007/s00122-002-0974-5.

Recebido em 8 de maio de 2014 e aprovado em 10 de setembro de 2014