

Autores

Clarissa L. Falcão
Embrapa Recursos Genéticos
e Biotecnologia

Marília C.R. Pappas
Embrapa Recursos Genéticos
e Biotecnologia

Rodrigo T. Lourenço
Embrapa Recursos Genéticos
e Biotecnologia

Maurício M. De Alencar
Embrapa Pecuária Sudeste
Bolsista CNPq

Adriana R.S. Batista
Embrapa Recursos Genéticos
e Biotecnologia
Programa de Ciências Genômicas
e Biotecnologia - Universidade
Católica de Brasília

Georgios J. Pappas Jr
Embrapa Recursos Genéticos
e Biotecnologia

Dario Grattapaglia
Embrapa Recursos Genéticos
e Biotecnologia

Desenvolvimento e mapeamento de microssatélites derivados de ESTs em *Eucalyptus*

Marcadores microssatélites (SSR) são utilizados de forma crescente como uma ferramenta útil em programas de melhoramento e produção florestal, além de representar um recurso experimental fundamental para iniciativas genômicas envolvendo mapeamento genético de QTL (quantitative trait loci) e ancoragem de mapas genéticos com mapas físicos. Em eucalipto, sua aplicabilidade abrange um espectro desde o direcionamento de cruzamentos entre indivíduos contrastantes até a geração de informações associadas com caracteres de interesse ao melhoramento. No âmbito do Projeto GENOLYPTUS, apesar de existirem mais de 400 marcadores desenvolvidos e mapeados, foi verificada a necessidade do desenvolvimento de mais marcadores de forma a permitir uma ampla flexibilidade na seleção de marcadores a serem utilizados no mapeamento das descendências. Como uma das metas do GENOLYPTUS, definiu-se então obter um total de 1000 microssatélites operacionais, mapeados e transferíveis inter-especificamente no intuito de que, com este número, seja plenamente possível cobrir todo o genoma com uma densidade de marcadores suficiente para poder fazer mapeamento de QTLs com precisão, bem como a ancoragem com o mapa físico. O Projeto GENOLYPTUS gerou um banco de dados com cerca de 120.000 leituras de ESTs (expressed sequence tags) de diferentes espécies e tecidos de *Eucalyptus* com múltiplos objetivos de desenvolvimento em pesquisa genômica e investigação funcional. Este banco de dados de seqüências representa uma fonte extremamente útil e importante para o descobrimento de novos microssatélites, além daqueles desenvolvidos a partir de bibliotecas genômicas enriquecidas (BRONDANI et al., 1998; BRONDANI et al., 2002). A vantagem adicional do desenvolvimento de microssatélites a partir de ESTs é o fato de que, ao mapear estes microssatélites, automaticamente são mapeados genes no mapa genético. Alguns destes genes mapeados podem constituir potenciais candidatos para funções importantes devido à sua co-localização com QTLs para características fenotípicas de importância econômica.

OBJETIVOS

- ✓ Minerar o banco de dados de ESTs do GENOLYPTUS para seqüências simples repetitivas (SSR) utilizando seqüências consenso – região de seqüência comum a várias leituras de um mesmo transcrito – interespecíficas visando o desenvolvimento de microssatélites de elevada transferibilidade interespecífica.
- ✓ Selecionar microssatélites em genes com funções potenciais relevantes para desenvolvimento de marcadores visando o mapeamento de genes candidatos.
- ✓ Realizar uma triagem de uma bateria inicial de cerca de 200 novos locos microssatélites para a determinação da qualidade de amplificação, polimorfismo e transferibilidade em um painel de 8 árvores envolvendo 5

espécies diferentes utilizadas como parentais nas populações segregantes no GENOLYPTUS.

MATERIAL E MÉTODOS

Material genético

O material genético foi definido a partir do desenho experimental de cruzamento do GENOLYPTUS, onde são utilizados 12 parentais geneticamente não relacionados de 5 espécies diferentes de *Eucalyptus* (*E. urophylla*, *E. grandis*, *E. globulus*, *E. calmadulensis* e *E. dunnii*) e 2 híbridos (*E. dunnii* x *E. grandis* e *E. urophylla* x *E. globulus*). Para este experimento, foram selecionados 8 parentais denominados como C1 (*E. calmadulensis*), DxG (Híbrido de *E. dunnii* e *E. grandis*), G1 (*E. grandis*), G2 (*E. grandis*), GL1 (*E. globulus*), GL2 (*E. globulus*), U2 (*E. urophylla*) e U15 (*E. urophylla*).

Mineração do banco de EST e desenho de primers

O agrupamento das cerca de 120.000 ESTs originais resultou em um conjunto de 22.298 consensos envolvendo seqüências de diferentes espécies. Com o software MREPS (KOLPAKOV et al., 2003), seqüências simples repetidas foram identificadas em 1.765 consensos de acordo com os seguintes parâmetros: 2-6 motivos de bases SSR, estrutura perfeita, i.e. sem interrupções microvariantes, e com um core mínimo de 12 bases. Pares de primers – iniciadores – flanqueando esses microssatélites foram desenhados para 1.261 marcadores utilizando o programa Primer3 com a mesma estrigência de parâmetros e gerando produtos esperados entre 80 e 350 pares de bases. A primeira bateria de 232 microssatélites foi testada envolvendo exclusivamente seqüências similares a genes interessantes para serem posicionados no mapa genético que apresentaram alto *e-value hit* – valor de significância atribuído à similaridade entre seqüências apresentado no resultado do programa BLAST. Entre esses genes estão várias classes de fatores de transcrição, proteínas kinases, enzimas envolvidas no metabolismo de polissacarídeos e lignina, assim como análogos a proteínas de resistência e choque térmico – *heat shock*.

Triagem de microssatélites

Os primers desenhados foram utilizados para a PCR dos 232 locos SSR. Destes, 210 pares foram sintetizados sem marcação fluorescente visando a triagem de sucesso de amplificação e detecção de polimorfismo. Para 22 marcadores, os primers diretos foram marcados com fluorocromos 6-FAM, HEX e NED (Applied Biosystems) visando a avaliação de polimorfismo e verificação de herança, segregação e mapeamento. Nesta bateria, a eletroforese, em sistema triplex previamente desenvolvido, foi realizada na plataforma ABI 3100 da Applied Biosystems. Os dados foram coletados automaticamente pela detecção das diferentes fluorescências e analisados com os softwares GeneScan/Genotyper (Applied Biosystems).

Na bateria de 210 microssatélites sem marcação, os produtos das reações foram submetidos a uma eletroforese em gel de agarose 3%, corado com brometo de etídeo 1mg/ml (SAMBROOK et al., 1989), por 2h a 100 volts constantes. Em cada gel, foram aplicados 100ng de um DNA padrão de fragmentos de tamanhos conhecidos (DNA Ladder 1Kb Plus - Invitrogen). Cada loco foi amplificado pelo menos 2x para confirmação do polimorfismo.

Na mineração e desenho de primers, os marcadores microssatélites foram previamente divididos em três grupos distintos, segundo os resultados do BLAST, de acordo com a sua posição em comparação com o genoma de *Arabidopsis*. As três categorias foram: SSR na região 5' UTR (untranslated region), SSR na ORF (open reading frame) e SSR de localização não determinada. Testes exatos de Fisher foram realizados visando avaliar a existência de diferenças significativas entre estas três classes no que se refere ao sucesso de amplificação, detecção de polimorfismo, bem como avaliar diferenças nestes parâmetros entre marcadores baseados em di e trinucleotídeos.

RESULTADOS

Frequência de microssatélites em EST

Uma mineração estridente foi realizada para microssatélites com base nas 22.298 seqüências consenso. Somente microssatélites perfeitos e com uma extensão mínima de 12 bases foram selecionados. Um total de 1765 microssatélites foi identificado, o que constitui um número mínimo, uma vez que microssatélites imperfeitos, interrompidos ou compostos foram excluídos. Ainda assim, para 1261 microssatélites (71,4%) foi possível desenhar primers com sucesso em uma mesma temperatura de anelamento (Figura 1). Dos 1261 microssatélites, 281 estavam localizados na região 5'UTR, 224 em ORFs e para 756 não foi possível determinar a localização precisa, tendo em vista a impossibilidade de inferir a posição do ATG inicial – início da tradução (Tabela 1). Considerando a distribuição relativamente equitativa entre SSR em 5'UTR e ORF, estes dados sugerem que provavelmente dentro da classe de SSR com localização indeterminada, cerca de metade ou pouco mais (55%) sejam oriundos de regiões 5'UTR e a outra metade de ORFs. Um outro aspecto interessante é o fato que, dos 1261 consensos nos quais microssatélites foram identificados, apenas 166 não apresentaram similaridade com nenhuma outra seqüência no GenBank, 233 apresentaram similaridade com seqüências do banco de proteínas de *Arabidopsis thaliana* e 52, com os bancos de proteínas de função não conhecida de *A. thaliana* ou *Oryza*. Ou seja, a maioria – cerca de 65% – destes marcadores microssatélites potencialmente permitirá a localização de um gene com função putativa, posto que sejam polimórficos e segreguem em uma família em mapeamento.

Características de microssatélites em EST

Analisando globalmente, o motivo mais freqüente observado entre os 1261 microssatélites identificados e para os quais primers foram desenvolvidos foi de trinucleotídeo com 54,7%, seguido de dinucleotídeos com 24,2%, hexa com 9,2% e tetra com apenas 6,5% (Tabela 1). Analisando esta mesma distribuição por posicionamento do SSR, diferenças importantes são observadas com uma quantidade maior de tri e hexanucleotídeos e menor de di e tetranucleotídeos em ORF em relação a 5'UTR. Esta diferença não é uma surpresa considerando que repetições de trinucleotídeos em regiões codantes são comumente observadas, gerando repetições de um resíduo de aminoácido, sem alterar a fase de leitura na tradução. Já, em regiões 5'UTR, há maior flexibilidade para a ocorrência de mutações por imperfeições na replicação, independente do número de bases, sem que haja erro na tradução.

Teste de herança e segregação

No teste de herança e segregação com os 22 locos SSR marcados com fluorescência, 12 mostraram-se robustos, polimórficos e com completa transferibilidade dos marcadores em uma família de *E. grandis* x *E. urophylla*. Os locos que apresentaram segregação na família referência utilizada foram mapeados com sucesso.

Sucesso de amplificação e detecção de polimorfismo

A análise dos 210 locos selecionados, para verificar o aspecto operacional para genotipagem e mapeamento, revelou que 175 locos (83,3%) amplificaram com sucesso, na mesma temperatura de anelamento, em todos os oito indivíduos do painel envolvendo as 5 espécies de *Eucalyptus*. Isso demonstra que o processo de produção de microssatélites envolvendo o agrupamento de seqüências – cluster – e desenho de primers a partir dos consensos é eficiente para gerar microssatélites com alta transferibilidade interespecífica. Em apenas 35 locos (16,6%), não houve amplificação ou a amplificação foi inespecífica demandando mais esforço de otimização. Dos 175 locos que amplificaram com sucesso produtos de PCR discretos, 69 não foram polimórficos e 106 revelaram polimorfismo neste painel de 8 indivíduos (Figura 2). Resulta-se, assim, em uma eficiência de 60,6% de polimorfismo, contados a partir de locos amplificados com sucesso, e de

50,5% de novos locos microssatélites operacionais e polimórficos a partir dos 210 testados (Tabela 2). Polimorfismos ainda poderão ser detectados ao se ampliar este painel de indivíduos melhorando a eficiência de recuperação de locos polimórficos. É relevante e interessante comparar esta eficiência em relação àquela obtida com o processo tradicional de desenvolvimento de microssatélites a partir de bibliotecas enriquecidas. Uma eficiência de 63% foi obtida partindo de informação de seqüências genômicas até marcadores microssatélites operacionais e polimórficos em uma bateria de 180 pares de primers desenhados (BRONDANI et al.,1998), ou seja, um pouco superior mas não de forma substancial. Claramente, o número de seqüências geradas na abordagem de biblioteca enriquecida é muito menor, mais focado e economicamente mais barato. Por outro lado, o desenvolvimento e mapeamento de microssatélites derivados de ESTs permitem a localização de genes no mapa genético e, conseqüentemente, de genes candidatos.

Comparação de sucesso de amplificação e nível de polimorfismo entre microssatélites

Os dados levantados foram organizados em tabelas de contingência envolvendo as três categorias, segundo a posição do marcador SSR, visando a realização dos testes exatos de Fisher. A hipótese nula nos testes referentes ao sucesso de amplificação foi de que não haveria diferença significativa em sucesso de amplificação entre marcadores de microssatélites posicionados na região 5'UTR, ORF ou posição indeterminada no gene, tendo em vista que os primers foram desenhados com base em seqüências consenso interespecíficas, o que levaria a uma elevada robustez na amplificação, independente da espécie ou região na qual o microssatélite se encontra. No que se refere à existência de polimorfismo, a hipótese nula foi de que microssatélites em ORFs seriam menos polimórficos do que aqueles em 5'UTR, tendo em vista a pressão de seleção em ORFs e conseqüente conservação de seqüência, restringindo processo de mutação por imperfeições na replicação.

A comparação entre repetições dinucleotídicas e trinucleotídicas, respectivamente, em relação ao sucesso de amplificação e detecção de polimorfismo revelou não existir diferença significativa entre as três classes de localização ($p>0,05$). Isso sugere que estas características operacionais no desenvolvimento de microssatélites independem do tipo de motivo. Para as repetições tetra, penta e hexanucleotídicas, não foi possível aplicar o teste exato de Fisher devido ao baixo número de representantes por grupo de localização.

DISCUSSÃO E PERSPECTIVAS

Os dados experimentais obtidos com os 210 microssatélites efetivamente testados, juntamente com os dados derivados da mineração e desenvolvimento de primers para 1261 novos marcadores a partir do banco de ESTs, permitem fazer algumas inferências em relação às perspectivas deste trabalho:

- ✓ O sucesso de amplificação não varia com a localização do microssatélite no gene, o que mostra que o processo de desenho de primers é eficiente para qualquer região, possivelmente pelo fato de ser desenvolvido a partir de consensos e não de seqüências individuais.

- ✓ Entretanto, mais marcadores baseados em microssatélites localizados na região 5'UTR são polimórficos em comparação àqueles em região de ORF com uma diferença pequena (51% contra 43%), porém, significativa ($p=0,0083$).

- ✓ Um grande número de microssatélites baseados em trinucleotídeos pode ser extraído do banco de ESTs (690 neste estudo). Estes são mais facilmente genotipados do que dinucleotídeos pelo fato da diferença no tamanho de alelos ser de 3 bases. Interessante é o fato de que não há diferença na proporção de marcadores polimórficos derivados de di ou de trinucleotídeos.

- ✓ Considerando que, do teste com 210 marcadores, resultaram 106 marcadores robustos, polimórficos e transferíveis interespecificamente, é possível prever que do total de 1244 (incluindo apenas di a hexanucleotídeos) cerca de 600 marcadores microssatélites operacionais poderão ser obtidos. Somados aos pouco mais de 400 disponíveis hoje, derivados de bibliotecas enriquecidas (BRONDANI et al., 1998; BRONDANI et al., 2002) e de clones genômicos gerados a partir de fragmentação aleatória (LOURENÇO, 2004), estes marcadores completarão a meta de 1000 microssatélites estabelecida no Projeto GENOLYPTUS. Um ponto importante é o fato de que, comparando-se as seqüências de marcadores microssatélites derivados

de seqüências ESTs e genômicos, nenhum marcador foi redundante. Além disso, uma nova fonte de microssatélites, em desenvolvimento no momento, é constituída por mais de 20.000 seqüências de pontas de BAC. Marcadores derivados destas pontas serão úteis para o processo de ancoragem de mapa físico ao mapa genético.

✓ Uma vertente interessante deste trabalho será a comparação da abundância e nível de polimorfismo de microssatélites localizados em genes e aqueles localizados em seqüências não transcritas. No Projeto GENOLYPTUS, mais de 20.000 seqüências genômicas estão disponíveis para serem comparadas com os cerca de 23.000 consensos. Espera-se que microssatélites estejam preferencialmente associados com regiões codantes conforme foi observado para genomas de plantas, independente do tamanho do genoma e espécie (MORGANTE et al., 2002); e que os microssatélites em regiões transcritas, pré-datando a expansão do genoma de plantas, sejam evolutivamente mais antigos do que aqueles em regiões não codantes tendendo, assim, a serem mais polimórficos.

✓ Vários destes microssatélites serão mapeados nos mapas referência do GENOLYPTUS agregando importante informação transcricional e aumentando a perspectiva de descobrimento de genes candidatos via co-localização com QTLs.

REFERÊNCIAS BIBLIOGRÁFICAS

BRONDANI, R. P.; BRONDANI, C.; GRATTAPAGLIA, D. Towards a genus-wide reference linkage map for Eucalyptus based exclusively on highly informative microsatellite markers. **Molecular Genetics and Genomics**, Berlin, v. 267, p. 338-347, 2002.

BRONDANI, R. P. V.; BRONDANI, C.; TARCHINI, R.; GRATTAPAGLIA, D. Development, characterization and mapping of microsatellite markers in Eucalyptus grandis and E. urophylla. **Theoretical and Applied Genetics**, New York, v. 97, p. 816-827, 1998.

KOLPAKOV, R.; BANA, G.; KUCHEROV, G. mreps: Efficient and flexible detection of tandem repeats in DNA. **Nucleic Acids Research**, London, v. 31, p. 3672-3678, 2003.

LOURENÇO, R. T. Estrutura genômica de três megabases de DNA genômico "seqüenciamento por fragmentação randômica de DNA (Shotgun)" de Eucalyptus: conteúdo nucleotídico, seqüências repetitivas e genes. 2004. 138 p. Dissertação (Mestrado em Genética e Biologia Molecular) - Instituto de Biologia, Universidade Estadual de Campinas, Campinas.

MORGANTE, M.; HANAFEY, M.; POWELL, W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. **Nature Genetics**, New York, v. 30, p. 194-200, 2002.

ROZEN, S.; SKALETSKY, H. Primer3 on the WWW for general users and for biologist programmers. **Methods in Molecular and Cellular Biology**, New York, v. 132, p. 365-386, 2000.

SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. **Molecular cloning: a laboratory manual**. New York: Cold Spring Harbor Laboratory Press, 1989.

1	Name	Primer	TM	Product_size	repeat	repeat	le	SSR	SSR_length	start_pos	ATG_founc	SSR_locat	blast_best_hit
291	EMBRA944R	AGCAATTCATT	59.797	280	AG	2	AGAGAG	19	303	0	0	0	expressed protein [Arabidopsis
292	EMBRA945F	GCAAGGTTCC	60.053	104	AG	2	AGAGAG	23	71	1	5'	UTR	P0482C06.20 [Oryza sativa (jap
293	EMBRA945R	GCTGCTGCTGC	60.187	104	AG	2	AGAGAG	23	71	1	5'	UTR	P0482C06.20 [Oryza sativa (jap
294	EMBRA946F	TCATCTGCTGT	59.984	289	AGC	3	AGCAGC	15	97	1	5'	UTR	P0482C06.20 [Oryza sativa (jap
295	EMBRA946R	CCTTCTCGCC	60.481	289	AGC	3	AGCAGC	15	97	1	5'	UTR	P0482C06.20 [Oryza sativa (jap
296	EMBRA947F	TATGCTGAAGG	60.481	140	GCC	3	GCCGCC	14	366	1	ORF	P0482C06.20 [Oryza sativa (jap	
297	EMBRA947R	AGTCCCCTATG	60.142	140	GCC	3	GCCGCC	14	366	1	ORF	P0482C06.20 [Oryza sativa (jap	
298	EMBRA948F	TAGTGA AAAAC	60.110	334	CCT	3	CCTCCT	17	479	1	ORF	cysteine protease [Anthurium a	
299	EMBRA948R	CAGCAACCGTA	60.011	334	CCT	3	CCTCCT	17	479	1	ORF	cysteine protease [Anthurium a	
300	EMBRA949F	CGTCCGCTCCA	62.445	284	CT	2	CTCTCT	32	4	1	5'	UTR	Putative protein with similarity t
301	EMBRA949R	ACTTGCCGTAC	60.134	284	CT	2	CTCTCT	32	4	1	5'	UTR	Putative protein with similarity t
302	EMBRA950F	CGTCCGCTCAA	60.723	176	CTCG	4	CTCGCT	18	16	1	5'	UTR	expressed protein [Arabidopsis
303	EMBRA950R	GAATCGAGCTC	59.957	176	CTCG	4	CTCGCT	18	16	1	5'	UTR	expressed protein [Arabidopsis
304	EMBRA951F	ACAAAACAAC	60.008	256	AGA	3	AGAAGA	12	528	1	ORF	hypothetical protein L73G19.50	
305	EMBRA951R	TGCATCATCCG	59.794	256	AGA	3	AGAAGA	12	528	1	ORF	hypothetical protein L73G19.50	
306	EMBRA952F	CATCAATCAG	59.799	185	GAC	3	GACGAC	13	238	0	0	DHHC-type zinc finger domain-	
307	EMBRA952R	TGGGGAAGAT	59.920	185	GAC	3	GACGAC	13	238	0	0	DHHC-type zinc finger domain-	
308	EMBRA953F	TCCGTCATCAC	60.112	218	CTG	3	CTGCTG	21	466	0	0	HB2 homeodomain protein [Poj	
309	EMBRA953R	GAATGGGAGAJ	60.195	218	CTG	3	CTGCTG	21	466	0	0	HB2 homeodomain protein [Poj	
310	EMBRA954F	TGTTCTGCTTC	60.195	172	CTGC	4	CTGCCT	17	664	0	0	HB2 homeodomain protein [Poj	
311	EMBRA954R	AAAAATACTCC	59.554	172	CTGC	4	CTGCCT	17	664	0	0	HB2 homeodomain protein [Poj	
312	EMBRA955F	AGAAGCTGCC	60.823	292	GA	2	GAGAGA	26	8	0	0	Ribulose-phosphate 3-epimeras	
313	EMBRA955R	AATCACATCCT	59.934	292	GA	2	GAGAGA	26	8	0	0	Ribulose-phosphate 3-epimeras	
314	EMBRA956F	TTGACTTCTGC	60.111	252	GAA	3	GAAGAA	18	733	1	ORF	expressed protein [Arabidopsis	
315	EMBRA956R	CTTGAACCTCG	60.255	252	GAA	3	GAAGAA	18	733	1	ORF	expressed protein [Arabidopsis	
316	EMBRA957F	ATCAGCTCGTT	60.961	110	AGAC	4	AGACAGA	19	15	1	5'	UTR	Thioredoxin M-type, chloroplast
317	EMBRA957R	CTGCAAGACTC	60.005	110	AGAC	4	AGACAGA	19	15	1	5'	UTR	Thioredoxin M-type, chloroplast
318	EMBRA958F	CGAGGGGTAG	60.005	198	CTC	3	CTCCTC	12	105	1	ORF	Thioredoxin M-type, chloroplast	
319	EMBRA958R	GGCCCTTGCA	60.196	198	CTC	3	CTCCTC	12	105	1	ORF	Thioredoxin M-type, chloroplast	
320	EMBRA959F	CAAGAATGCC	59.948	244	TGT	3	TGTTGTT	12	553	1	ORF	Thioredoxin M-type, chloroplast	
321	EMBRA959R	TTCACCCACAA	60.388	244	TGT	3	TGTTGTT	12	553	1	ORF	Thioredoxin M-type, chloroplast	
322	EMBRA960F	AGTGATCGATC	59.685	350	GGA	3	GGAGGA	12	49	1	5'	UTR	expressed protein [Arabidopsis
323	EMBRA960R	TTTTGGAGATG	60.195	350	GGA	3	GGAGGA	12	49	1	5'	UTR	expressed protein [Arabidopsis

Figura 1: Lista dos locos microssatélites indicando o nome do loco, os primers desenhados, a Tm, tamanho esperado do produto de PCR, o motivo e o tipo de microssatélite - di, tri, tetra ou mais - o comprimento do microssatélite no consenso, a presença do ATG, a localização do SSR e o melhor hit do BLAST.

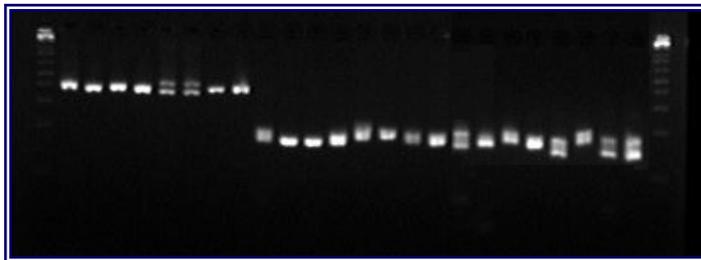


Figura 2: Perfil de 3 locos SSR polimórficos nos 8 parentais. Pistas 2 à 9, EMBRA 979; 10 à 17, EMBRA 931; 18, à 25 EMBRA 1468; pistas 1 e 26: 1 kb plus ladder.

Tabela 1: Distribuição de frequência de microssatélites em diferentes categorias de motivos nas três classes de posicionamento no gene identificados com o MREPS.

CLASSIFICAÇÃO	SSR EM 5'-UTR	%	SSR EM ORF	%	SSR EM POS. IND.	%	TOTAL	%
Dinucleotídeos	82	29,2	25	11,2	198	26,2	305	24,2
Trinucleotídeos	136	48,4	161	71,9	393	52,0	690	54,7
Tetranucleotídeos	21	7,5	6	2,7	55	7,3	82	6,5
Pentanucleotídeos	9	3,2	8	3,6	34	4,5	51	4,0
Hexanucleotídeos	30	10,7	21	9,4	65	8,6	116	9,2
Heptanucleotídeos+	3	1,1	3	1,3	11	1,5	17	1,3
TOTAL	281	100,0	224	100,0	756	100,0	1261	100

Tabela 2: Resultado da triagem de marcadores microssatélites por localização no gene quanto ao sucesso de amplificação e detecção de polimorfismos.

CLASSIFICAÇÃO	SSR EM 5'-UTR	%	SSR EM ORF	%	SSR EM POS. IND.	%	TOTAL	%
Dinucleotídeos	28	28,6	5	16,7	32	39,0	65	31,0
Trinucleotídeos	51	52,0	23	76,7	37	45,1	111	52,9
Tetranucleotídeos	8	8,2	1	3,3	8	9,8	17	8,1
Pentanucleotídeos	5	5,1	0	0,0	2	2,4	7	3,3
Hexanucleotídeos	6	6,1	1	3,3	3	3,7	10	4,8
TOTAL	98	100,0	30	100,0	82	100,0	210	100,0
Sucesso na amplif.	78	79,6	28	93,3	69	84,1	175	83,3
Polimórficos	50	51,0	13	43,3	43	52,4	106	50,5

Tabela 3: Valores p dos testes exatos de Fisher para as comparações de sucesso de amplificação (acima da diagonal) e detecção de polimorfismo (abaixo da diagonal) entre as três categorias de posicionamento do microssatélite no gene.

	5' UTR	ORF	SSR localização indeterminada
5' UTR		0,1003	0,4476
ORF	0,0083		0,3470
SSR localização não determinada	0,0986	0,1776	

<p>Circular Técnica 32</p> <p>Ministério da Agricultura, Pecuária e Abastecimento</p>	<p>Exemplares desta edição podem ser adquiridos na Embrapa Recursos Genéticos e Biotecnologia Serviço de Atendimento ao Cidadão Parque Estação Biológica, Av. W/5 Norte (Final) – Brasília, DF CEP 70770-900 – Caixa Postal 02372 PABX: (61) 448-4600 Fax: (61) 340-3666 http://www.cenargen.embrapa.br e.mail:sac@cenargen.embrapa.br</p> <p>1ª edição 1ª impressão (2004): 150 unidades</p>	<p>Comitê de Publicações</p> <p>Expediente</p>	<p>Presidente: <i>Maria Isabel de Oliveira Penteadó</i> Secretário-Executivo: <i>Maria da Graça Simões Pires Negrão</i> Membros: Arthur da Silva Mariante Maria Alice Bianchi Maria da Graça S. P. Negrão Maria de Fátima Batista Maria Isabel de O. Penteadó Maurício Machain Franco Regina Maria Dechechi Carneiro Sueli Correa Marques de Mello Vera Tavares de Campos Carneiro Supervisor editorial: <i>Maria da Graça S. P. Negrão</i> Normalização Bibliográfica: <i>Maria Alice Bianchi e Maria Iara Pereira Machado</i> Editoração eletrônica: <i>Maria da Graça Simões Pires Negrão</i></p>
---	--	--	---